Vers des modes de scrutin moins manipulables

François Durand

▶ **To cite this version:**

François Durand. Vers des modes de scrutin moins manipulables. Modélisation et simulation. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066519 . tel-01242440v2

HAL Id: tel-01242440
https://inria.hal.science/tel-01242440v2

Submitted on 11 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité
Informatique

École doctorale Informatique,
Télécommunications et Électronique (Paris)

Présentée par

François Durand

Pour obtenir le grade de

**DOCTEUR DE l'UNIVERSITÉ
PIERRE ET MARIE CURIE**

Vers des modes de scrutin moins manipulables

Soutenue le 24 septembre 2015 devant le jury composé de :

M. Markus Brill, Associé postdoctoral, Duke University :	examinateur
M. Jérôme Lang, Directeur de recherche, CNRS :	rapporteur
M. Jean-François Laslier, Directeur de recherche, CNRS :	examinateur
M. Fabien Mathieu, Ingénieur de recherche, ALBLF :	co-encadrant
M. Nicolas Maudet, Professeur, UPMC :	examinateur
M. Vincent Merlin, Directeur de recherche, CNRS :	rapporteur
M. Ludovic Noirie, Ingénieur de recherche, ALBLF :	co-encadrant
M. Sébastien Tixeuil, Professeur, UPMC :	directeur

Résumé

Nous étudions la manipulation par coalition des modes de scrutin : est-ce qu'un sous-ensemble des électeurs, en votant de façon stratégique, peut faire élire un candidat qu'ils préfèrent tous au candidat qui aurait été vainqueur si tous les électeurs avaient voté sincèrement ?

D'un point de vue théorique, nous développons un formalisme qui permet d'étudier tous les modes de scrutin, que les bulletins soient des ordres de préférences sur les candidats (systèmes *ordinaux*), des notes ou des valeurs d'approbation (systèmes *cardinaux*) ou des objets encore plus généraux. Nous montrons que pour la quasi-totalité des modes de scrutin classiques, on peut réduire strictement leur manipulabilité en ajoutant un test préliminaire visant à élire le vainqueur de Condorcet s'il en existe un. Pour les autres modes de scrutin, nous définissons la *condorcification généralisée* qui permet d'obtenir des résultats similaires. Puis nous définissons la notion de culture *décomposable*, une hypothèse dont l'indépendance probabiliste des électeurs est un cas particulier. Sous cette hypothèse, nous prouvons que, pour tout mode de scrutin, il existe un mode de scrutin qui est ordinal, qui partage certaines propriétés avec le mode de scrutin original et qui est au plus aussi manipulable. Ainsi, la recherche d'un mode de scrutin de manipulabilité minimale (dans une classe de systèmes raisonnables) peut être restreinte à ceux qui sont ordinaux et vérifient le critère de Condorcet.

Afin de permettre à tous d'examiner ces phénomènes en pratique, nous présentons SVAMP, un package Python de notre cru dédié à l'étude des modes de scrutin et de leur manipulabilité. Puis nous l'utilisons pour comparer la manipulabilité par coalition de divers modes de scrutin dans plusieurs types de cultures, c'est-à-dire des modèles probabilistes permettant de générer des populations d'électeurs munis de préférences aléatoires. Nous complétons ensuite l'analyse avec des élections issues d'expériences réelles. Enfin, nous déterminons les modes de scrutin de manipulabilité minimale pour de très faibles valeurs du nombre d'électeurs et du nombre de candidats et nous les comparons avec les modes de scrutin classiques. De manière générale, nous établissons que la méthode de Borda, le vote par notation et le vote par assentiment sont particulièrement manipulables. À l'inverse, nous montrons l'excellente résistance à la manipulation du système appelé *VTI*, également connu par son acronyme anglophone *STV* ou *IRV*, et de sa variante *Condorcet-VTI*.

Abstract

We investigate the coalitional manipulation of voting systems: is there a subset of voters who, by producing an insincere ballot, can secure an outcome that they strictly prefer to the candidate who wins if all voters provide a sincere ballot?

From a theoretical point of view, we develop a framework that allows us to study all kinds of voting systems: ballots can be linear orders of preferences over the candidates (*ordinal* systems), grades or approval values (*cardinal* systems) or even more general objects. We prove that for almost all voting systems from literature and real life, manipulability can be strictly diminished by adding a preliminary test that elects the Condorcet winner if one exists. Then we define the notion of *decomposable* culture and prove that it is met, in particular, when voters are independent. Under this assumption, we prove that for any voting system, there exists a voting system that is ordinal, has some common properties with the original voting system and is at most as manipulable. As a consequence of these theoretical results, when searching for a voting system whose manipulability is minimal (in a class of reasonable systems), investigation can be restricted to those that are ordinal and meet the Condorcet criterion.

In order to provide a tool to investigate these questions in practice, we present SVVAMP, a Python package we designed to study voting systems and their manipulability. We use it to compare the coalitional manipulability of several voting systems in a variety of cultures, i.e. probabilistic models generating populations of voters with random preferences. Then we perform the same kind of analysis on real elections. Lastly, we determine voting systems with minimal manipulability for very small values of the number of voters and the number of candidates and we compare them with classical voting systems from literature and real life. Generally speaking, we show that the Borda count, Range voting and Approval voting are especially vulnerable to manipulation. In contrast, we find an excellent resilience to manipulation for the voting system called *IRV* (also known as *STV*) and its variant *Condorcet-IRV*.

Remerciements

Merci à Sébastien d'avoir accepté de diriger cette thèse. Merci à Ludovic, qui s'est plongé avec passion dans ce domaine de recherche qui n'était pas le sien au départ et qui a apporté des contributions issus de champs aussi divers que l'optique, la mécanique quantique et les télécommunications. Merci à Fabien pour l'appui constant qu'il a apporté pendant ces années, l'appréciable complémentarité de nos approches, le sérieux et la compétence avec lequel il a guidé mon travail et surtout pour l'authenticité de ses chemises.

Merci à Jérôme et Vincent pour leur lecture attentive et la pertinence de leurs remarques. Merci à Nicolas, Jean-François et Markus pour avoir accepté de participer à ce jury de thèse.

Merci à ma famille et mes amis pour leur support, en particulier pendant la période intense de rédaction. Merci à Émilie pour son soutien permanent et pour sa relecture attentive de ce mémoire.

Merci à Anne-Laure et Arvind pour nos discussions autour du lemme de la variable aléatoire complémentaire. Merci à Laurent pour nos discussions et pour avoir suggéré le mode de scrutin VTID. Merci à Benoît d'avoir accepté de se pencher sur une nappe en papier pour examiner une question que j'aurais été bien incapable de résoudre tout seul.

Merci aux membres de Displexity, Carole, Hugues et tous les autres, qui m'ont accueilli à bras ouverts et ont prêté une oreille attentive à mes exposés. Merci aux organisateurs d'Algotel 2012, à Nicolas et Pierre pour avoir contribué à organiser nos expériences. Merci à Alix pour m'avoir communiqué la base de données de mon site marron préféré.

Merci à Pupusse et à ses parents adoptifs pour les régulières et intenses séances de réflexion que nous avons partagées sur la manipulabilité des systèmes de vote : Fabien, Rim, Alonso, Jose, Ludovic, Marco, Rémy, Loïc, Stéphane, Pierre, Akram et tous les autres qui se reconnaîtront.

Table des matières

Introduction **1**
Le vote majoritaire simple . . . 1
Paradoxes du choix social . . . 3
Paradoxe de Condorcet . . . 3
Théorème d'Arrow . . . 4
Comparaison interpersonnelle des utilités . . . 5
Théorème de Gibbard-Satterthwaite . . . 6
Théorème de Gibbard non déterministe . . . 7
Manipuler, c'est bien. La manipulabilité, c'est mal. . . . 7
Défense de la manipulation . . . 8
Manipulabilité et équilibres de Nash forts (ENF) . . . 9
Manipulabilité et *straightforwardness* . . . 10
Indicateurs de manipulabilité . . . 12
Nos objectifs principaux . . . 13
Contributions et plan . . . 14
Étude théorique de la manipulabilité . . . 14
Étude assistée par ordinateur de la manipulabilité . . . 15
Annexes . . . 16

I Étude théorique de la manipulabilité **17**

1 Formalisme **19**
1.1 Espaces électoraux . . . 19
1.1.1 Relations binaires . . . 20
1.1.2 Profils . . . 20
1.1.3 Définition d'un espace électoral . . . 21
1.1.4 Exemples d'espaces électoraux . . . 22
1.1.5 Propriétés basiques d'un espace électoral . . . 24
1.1.6 Espace électoral probabilisé . . . 26
1.2 Systèmes de vote basés sur l'état (SVBE) . . . 27
1.2.1 Définition . . . 27
1.2.2 Critères basiques pour un SVBE . . . 28
1.3 Manipulabilité . . . 32
1.4 Systèmes de vote généraux . . . 36
1.4.1 Définition . . . 36
1.4.2 Réduction d'un système de vote général à un SVBE . . . 39
1.4.3 Implantation canonique d'un SVBE . . . 40
1.5 Matrice des duels et notions Condorcet . . . 42
1.6 Zoologie des modes de scrutin . . . 45
1.6.1 Modes de scrutin cardinaux . . . 46

1.6.2 Règles positionnelles à score (RPS) . . . 47
1.6.3 Méthode de Bucklin . . . 48
1.6.4 RPS itérées à élimination simple . . . 49
1.6.5 RPS itérées à élimination par la moyenne . . . 50
1.6.6 Méthodes par éliminations diverses . . . 51
1.6.7 Méthodes Condorcet . . . 51

2 Condorcification **55**
2.1 Critère de la coalition majoritaire informée (**CMInf**) . . . 56
2.2 Liens entre manipulation et résultats des duels électoraux . . . 58
2.3 Définition de la condorcification . . . 59
2.4 Théorème faible de condorcification . . . 60
2.5 Variantes de condorcification . . . 61
2.5.1 Définition des variantes de condorcification . . . 61
2.5.2 Variantes de condorcification : résultat général . . . 62
2.5.3 Variantes de condorcification pour l'uninominal . . . 63
2.5.4 Variantes de condorcification pour U2TI . . . 66
2.5.5 Condorcification relative pour VTI . . . 68
2.5.6 Condorcification relative pour les systèmes cardinaux . . . 69
2.6 Vainqueur de Condorcet résistant . . . 70
2.7 Théorème fort de condorcification . . . 75
2.8 Borne supérieure de manipulabilité pour les systèmes Condorcet . . . 75
2.9 Condorcification et systèmes optimaux . . . 76

3 Critères majoritaires **79**
3.1 Définition des critères majoritaires . . . 81
3.2 Implications entre les critères majoritaires . . . 82
3.3 Lien avec les équilibres de Nash forts (ENF) . . . 83
3.4 Critères majoritaires vérifiés par les modes de scrutin classiques . . . 89
3.4.1 Modes de scrutin cardinaux . . . 90
3.4.2 Uninominal à un ou plusieurs tours . . . 91
3.4.3 Règles positionnelles à score (RPS) . . . 92
3.4.4 RPS itérées à élimination simple . . . 95
3.4.5 RPS itérées à élimination par la moyenne . . . 98
3.4.6 Méthode de Bucklin simple ou itérée . . . 100
3.5 Aspect informationnel des critères majoritaires . . . 102

4 Condorcification généralisée **105**
4.1 Famille de collections de coalitions . . . 106
4.1.1 Définition et propriétés basiques d'une famille . . . 107
4.1.2 Victoire et notions Condorcet généralisées . . . 108
4.1.3 Candidat $\mathcal{M}$-favori . . . 110
4.2 Critères associés à une famille . . . 111
4.2.1 Définition des critères associés à une famille . . . 111
4.2.2 Implications entre les critères associés à une famille . . . 113
4.3 Théorème de condorcification généralisée . . . 114
4.4 Théorème de condorcification comparée . . . 115
4.5 Théorème de condorcification maximale . . . 118
4.6 Exemples de condorcification généralisée . . . 121
4.6.1 Systèmes de vote usuels vérifiant CMInf . . . 121
4.6.2 Veto . . . 122
4.6.3 Vote de parité . . . 125
4.6.4 Vote d'une loi . . . 126

4.6.5 Uninominal avec un pouvoir d'imposition . . . 127
4.6.6 Élection du Secrétaire général de l'ONU . . . 128

5 Tranchage 131
5.1 Tranches d'un mode de scrutin . . . 133
5.2 Espace électoral décomposable . . . 135
5.3 Théorème de tranchage . . . 138
5.4 Condorcification et tranchage . . . 139
5.5 Théorème d'optimalité . . . 140
5.6 Décomposabilité d'un ensemble probabilisé quelconque . . . 141
5.6.1 Cas unidimensionnel : le lemme de la variable aléatoire complémentaire . . . 141
5.6.2 Cas multidimensionnel : la décomposabilité dans le cas général 143
5.7 Discussion des hypothèses et des conclusions du théorème de tranchage . . . 145
5.7.1 Diminution de la manipulabilité au sens probabiliste ou ensembliste ? . . . 145
5.7.2 Hypothèse de décomposabilité . . . 147

II Étude assistée par ordinateur de la manipulabilité 149

6 SVAMP : *Simulator of Various Voting Algorithms in Manipulating Populations* 151
6.1 Population . . . 153
6.1.1 Obtenir des informations sur une population . . . 153
6.1.2 Travailler avec des utilités . . . 154
6.1.3 Créer un objet Population . . . 155
6.2 Elections . . . 156
6.2.1 Résultat d'une élection . . . 157
6.2.2 Manipulation par coalition (MC) . . . 158
6.2.3 Variantes de la MC (MCI, MU, MT) . . . 158
6.2.4 Manipulation individuelle (MI) . . . 159
6.2.5 Indépendance des alternatives non pertinentes (IANP) . . . 160
6.2.6 Critères vérifiés par un système de vote . . . 160
6.3 Implantation de MC . . . 161
6.3.1 Minimisation des calculs . . . 161
6.3.2 Nombre de manipulateurs . . . 162
6.3.3 Anonymat et caractère résolu . . . 162
6.3.4 Sous-fonctions de MC . . . 163
6.3.5 Performances de l'algorithme générique polynomial . . . 164
6.4 Algorithmes spécifiques pour certains modes de scrutin . . . 165
6.5 Algorithmes spécifiques pour VTI et ses variantes . . . 166
6.5.1 Algorithmes de MC pour SE . . . 167
6.5.2 Algorithmes de MC pour VTI . . . 170
6.5.3 Algorithmes de MC pour CVTI . . . 172

7 Étude de cultures sphéroïdales 175
7.1 Présentation des cultures sphéroïdales . . . 175
7.2 Scénario de référence . . . 177
7.3 Culture sphérique : nombre de candidats C . . . 180
7.4 Culture sphérique : nombre d'électeurs V . . . 185
7.4.1 Nombre d'électeurs V impair . . . 185

7.4.2 Influence de la parité de V . . . 186
7.5 Culture monopolaire : concentration κ . . . 189
7.6 Culture monopolaire : position du pôle . . . 191
7.7 Culture multipolaire : nombre de pôles . . . 193
7.8 Culture multipolaire : positions relatives des pôles . . . 194
7.9 Méta-analyse dans les cultures sphéroïdales . . . 195
7.9.1 Méta-analyse de MC dans les cultures sphéroïdales . . . 196
7.9.2 Méta-analyse de MT dans les cultures sphéroïdales . . . 198

8 Étude de cultures basées sur un spectre politique 201
8.1 Puits gaussien unidimensionnel : nombre de candidats C . . . 202
8.2 Puits gaussien unidimensionnel : nombre d'électeurs V . . . 204
8.2.1 Nombre d'électeurs V impair . . . 204
8.2.2 Influence de la parité de V . . . 205
8.3 Puits gaussien unidimensionnel : décalage y_0 . . . 206
8.4 Puits gaussien multidimensionnel : nombre de dimensions n . . . 207
8.5 Comparaison avec une boîte euclidienne . . . 209
8.6 Méta-analyse en culture unidimensionnelle . . . 210
8.6.1 Méta-analyse de MC en culture unidimensionnelle . . . 210
8.6.2 Méta-analyse de MT en culture unidimensionnelle . . . 211
8.7 Méta-analyse dans l'ensemble des cultures étudiées . . . 212
8.7.1 Méta-analyse de MC dans l'ensemble des cultures étudiées . 212
8.7.2 Méta-analyse de MT dans l'ensemble des cultures étudiées . 213

9 Analyses de données expérimentales 215
9.1 Présentation des données expérimentales . . . 215
9.1.1 Expériences réalisées . . . 215
9.1.2 Site internet www.bdtheque.com . . . 216
9.1.3 Jugement de Paris . . . 217
9.1.4 Base de données PrefLib . . . 218
9.2 Méthodologie . . . 218
9.3 Résultats « sincères » et notions Condorcet . . . 221
9.4 Manipulation par coalition . . . 223
9.4.1 Taux de MC moyens . . . 223
9.4.2 MC comparée des systèmes de vote . . . 226
9.4.3 MC par candidat . . . 227
9.5 Manipulation triviale . . . 229
9.5.1 Taux de MT moyens . . . 229
9.5.2 MT comparée des systèmes de vote . . . 230
9.5.3 MT par candidat . . . 232
9.6 Synthèse des résultats . . . 233

10 Modes de scrutin optimaux 235
10.1 Graphe d'opportunité . . . 236
10.1.1 Définition . . . 236
10.1.2 Restriction aux modes de scrutin Condorcet . . . 237
10.1.3 Configurations semi-Condorcet de probabilité nulle . . . 238
10.1.4 Restriction aux modes de scrutin ordinaux . . . 240
10.1.5 Algorithme glouton . . . 240
10.1.6 Approche exacte : optimisation linéaire en nombres entiers 241
10.2 Modes de scrutin optimaux pour de faibles valeurs de V et C . . . 241
10.2.1 $V = 3$ et $C = 3$: des optima comme s'il en pleuvait . . . 242
10.2.2 $V = 5$ et $C = 3$: CVTI et c'est tout ! . . . 244

10.2.3 $V = 7$ et $C = 3$: bien choisir la règle de départage de CVTI 246
10.2.4 $V = 3$ et $C = 4$: un complexe ensemble d'optima 250
10.3 Comparaison entre l'optimum et les modes de scrutin classiques . . 253

Conclusion et perspectives **257**

Annexes **269**

A Espaces probabilisables et mesurabilité **271**

B Géométrie de l'espace des utilités **273**
B.1 Modèle de Von Neumann-Morgenstern 275
B.2 Dualité avec l'hyperplan tangent aux loteries 278
B.3 Opérateurs opposé et somme 280
B.4 Représentation riemannienne de l'espace des utilités 282
B.5 Application : mesures de probabilité sur l'espace des utilités 285
B.6 Conclusion 288
B.7 Extension : espace des utilités avec limite d'approbation 288
B.7.1 Modèles avec options virtuelles 288
B.7.2 Correspondance entre les utilités avec limite d'approbation et les formes affines 293
B.7.3 Opérateurs opposé et somme dans l'espace des utilités avec limite d'approbation 295
B.7.4 Représentation riemannienne de l'espace des utilités avec limite d'approbation 296
B.7.5 La représentation riemannienne en pratique 297

C Systèmes de vote appliqués aux réseaux de télécommunication **299**
C.1 Présentation du modèle d'établissement d'un chemin par plusieurs opérateurs 300
C.1.1 Modélisation sous forme d'élection 300
C.1.2 Sélection des chemins candidats 301
C.1.3 Modèle des coûts et des gains pour les opérateurs 302
C.1.4 Systèmes de vote étudiés 303
C.1.5 Algorithmes de manipulation 304
C.2 Manipulabilité et efficacité des systèmes de vote étudiés 305
C.2.1 Scénario de référence 305
C.2.2 Influence des paramètres 306
C.3 Conclusion 308

Glossaire des notations **311**
Symboles non alphabétiques 311
Alphabet grec 311
Alphabet latin 311
Sigles et abréviations 313

Bibliographie **315**

Publications

Communications dans un congrès

François Durand, Benoît Kloeckner, Fabien Mathieu et Ludovic Noirie : Geometry on the utility sphere. *In Proceedings of the 4th International Conference on Algorithmic Decision Theory (ADT)*, 2015.

François Durand, Fabien Mathieu et Ludovic Noirie : Élection du best paper AlgoTel 2012 : étude de la manipulabilité. *In AlgoTel 2014 – 16èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications*, 2014.

François Durand, Fabien Mathieu et Ludovic Noirie : Élection d'un chemin dans un réseau : étude de la manipulabilité. *In AlgoTel 2014 – 16èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications*, 2014.

François Durand, Fabien Mathieu et Ludovic Noirie : On the manipulability of voting systems : application to multi-operator networks. *In Proceedings of the 9th International Conference on Network and Service Management (CNSM)*, pages 292–297. IEEE, 2013.

Poster

François Durand, Fabien Mathieu et Ludovic Noirie : Reducing manipulability. Poster présenté au 5th International Workshop on Computational Social Choice (COMSOC), 2014.

Communication dans un groupe de travail

François Durand, Fabien Mathieu et Ludovic Noirie : Manipulability of voting systems. Groupe de travail Displexity, http://www.liafa.univ-paris-diderot.fr/~displexity/docpub/6mois/votes.pdf, 2012.

Rapports de recherche

François Durand, Fabien Mathieu et Ludovic Noirie : Making most voting systems meet the Condorcet criterion reduces their manipulability. https://hal.inria.fr/hal-01009134, 2014.

François Durand, Fabien Mathieu et Ludovic Noirie : Making a voting system depend only on orders of preference reduces its manipulability rate. https://hal.inria.fr/hal-01009136, 2014.

Introduction

Quand on évoque les modes de scrutin, on pense immédiatement aux élections rythmant notre vie publique. Cependant, la pratique des élections est loin d'être limitée au domaine politique : elles peuvent être utilisées dans toute situation où un certain nombre d'agents, les *électeurs*, souhaitent collectivement choisir une option parmi plusieurs possibilités exposées à leur sagacité, les *candidats*. Ainsi, les élections peuvent être utilisées pour désigner le meilleur restaurant ou le meilleur film. Elles peuvent être utilisées dans des associations ou des organisations professionnelles. Et le développement d'organisations passant les frontières physiques grâce à Internet et ayant également besoin de désigner des représentants rend d'autant plus riche la vie publique et une forme de démocratie indépendante des États. Par exemple, le projet Debian[1], la communauté Ubuntu[2], la fondation Wikipédia France[3], le site `www.boardgamegeek.com` et de nombreux autres utilisent la méthode de Schulze, un mode de scrutin développé en 1997, et sont ainsi à la pointe de l'expérimentation en matière de systèmes de vote innovants.

À l'heure où l'on constate une hausse permanente de l'abstention dans le pays d'origine de la Déclaration des Droits de l'Homme et du Citoyen, pendant que dans d'autres pays, les peuples luttent pour obtenir le droit de vote ; à l'heure où les électeurs de nos pays montrent une méfiance croissante à l'égard de la classe politique, et où les modes de scrutin classiques ne semblent pas satisfaire la volonté d'expression des électeurs ; à l'heure, enfin, où les possibilités encore balbutiantes offertes par le vote électronique permettent d'envisager, dans un futur proche, l'utilisation de modes de scrutin au dépouillement plus complexe, il nous semble plus que jamais urgent de prendre le temps de réfléchir aux procédures utilisées pour voter, c'est-à-dire aux modes de scrutin eux-mêmes.

Le vote majoritaire simple

Dans cette thèse, nous utiliserons toujours le mot *candidats* pour désigner les différentes options qui sont proposées aux électeurs, même quand il ne s'agit pas de personnes qui se présentent à une charge publique. La situation la plus simple survient lorsque les électeurs doivent répondre à une question fermée par une réponse *oui* ou *non*.

Depuis au moins la démocratie athénienne du siècle de Périclès (Ve siècle av. J.C.), on a utilisé, pour de telles décisions, le vote majoritaire simple. Il en est ainsi dans les principales assemblées athéniennes :

— l'*Ecclésia* (κκλησία), l'assemblée de tous les citoyens, réunie sur la colline de la Pnyx environ 40 fois par an ;

1. `www.debian.org/`.
2. `www.ubuntu-fr.org/`.
3. `fr.wikipedia.org/`.

— la *Boulè* (βουλή), un conseil restreint constitué de 50 membres tirés au sort qui se réunissent quotidiennement pour traiter les affaire courantes ;
— et les *dikastéria* (δικαστήρια), des jurés populaires tirés au sort, chargés aussi bien de juger des contentieux précis que de discuter la légalité des décrets, préfigurant ainsi les jurés populaires aussi bien que des institutions comme l'actuel Conseil Constitutionnel (pour la France) ou la Cour européenne des droits de l'Homme (pour l'Union Européenne).

Dans le cas d'une question fermée, le vote uninominal majoritaire s'est donc très tôt présenté comme une évidence.

Bien sûr, il existe un autre cas où il y a deux candidats seulement : quand deux véritables candidats au sens courant, c'est-à-dire deux êtres humains, sont placés en concurrence et soumis à l'avis des électeurs.

May (1952) a formalisé cette supériorité du vote majoritaire simple par une approche axiomatique : dans le cas où il y a deux candidats, c'est le seul mode de scrutin qui possède les propriétés suivantes.

1. Il est *anonyme*, c'est-à-dire qu'il traite tous les électeurs de manière égale.
2. Il est *neutre*, c'est-à-dire qu'il traite les deux candidats de manière égale.
3. Il *réagit positivement*[4], c'est-à-dire que si un électeur préfère le candidat a au candidat b, il n'y a aucun cas où il a intérêt à voter pour b.
4. De manière implicite, May suppose que le mode de scrutin est *ordinal*. Ce terme signifie que l'électeur peut établir un ordre de préférence sur les candidats : soit il préfère le candidat a à b, soit l'inverse, soit il les apprécie à égalité. Mais il ne peut exprimer une opinion plus nuancée : il lui est ainsi impossible de s'exprimer différemment selon qu'il préfère fortement a à b ou qu'il préfère légèrement a à b.

Les deux premières hypothèses, l'anonymat et la neutralité, semblent évidentes en pratique, en tout cas dans les contextes d'application où les électeurs, d'une part, et les candidats, d'autre part, sont supposés égaux en droit. Il n'en est pas toujours ainsi, et parfois de manière défendable, comme au sein d'une assemblée d'actionnaires ou d'une fédération d'États de tailles diverses. Cependant, il existe un large champ d'application où ces hypothèses sont frappées du sceau de l'évidence.

La troisième hypothèse, la réactivité positive, semble aussi découler du simple bon sens. Elle implique, en particulier, que pour tout électeur et quelle que soit ses opinions, il peut émettre un bulletin qui les défendra au mieux, quels que soient les bulletins des autres électeurs : dans la terminologie de la théorie des jeux, on dit qu'il dispose toujours d'une *stratégie dominante*. Nous reviendrons sur cette notion car elle est profondément liée à la manipulabilité, qui constituera la problématique centrale de notre étude.

La quatrième hypothèse, quasi-implicite dans la formulation de May, est l'ordinalité. Elle semble assez intuitive, également pour des considérations de stratégie : imaginons qu'un électeur puisse renforcer son vote pour a en affirmant le préférer fortement à b. S'il ne préfère que légèrement a à b, il est possible qu'il affirme quand même le préférer fortement, simplement pour donner de meilleures chances à son opinion de l'emporter. En utilisant un tel système, on risque d'aboutir à deux situations.

— Si tous les électeurs raisonnent de cette manière, on aboutit à une situation où les bulletins proclamant une préférence légère ne sont plus du tout utilisés, et on est ramené à un système ordinal où le bulletin « je préfère fortement » est simplement utilisé pour signifier « je préfère ».

4. En anglais, on dit que le mode de scrutin est *positively responsive*.

— Alternativement, si certains électeurs raisonnent ainsi mais pas tous, la situation est encore plus néfaste, puisque les premiers, qui votent stratégiquement, ont plus de pouvoir que les seconds, qui votent de façon naïve ; une telle situation compromet *de facto* le principe d'égalité entre les électeurs.

Il est aussi possible de justifier l'ordinalité par d'autres arguments très profonds. Nous y reviendrons.

Paradoxes du choix social

Nous avons vu que quand il y a deux candidats, on dispose d'un mode de scrutin, le vote majoritaire simple, qui possède des propriétés satisfaisantes, et que ce mode de scrutin est unique sous des hypothèses assez raisonnables. Mais dès qu'il y a trois candidats, les choses se gâtent.

Paradoxe de Condorcet

La question qui se pose est la suivante : puisque nous disposons d'un système possédant de bonnes propriétés pour deux candidats et qu'il est dans une certaine mesure unique, comment l'étendre de façon satisfaisante aux cas où il y a davantage de candidats ?

Pour répondre à cette question, une idée naturelle est d'exiger l'*indépendance des alternatives non pertinentes* (IANP), qui signifie intuitivement que la présence ou l'absence de candidats perdants n'influence pas le résultat du scrutin. Plus précisément, on peut formuler ce principe de deux façons équivalentes.

1. Si on supprime un candidat perdant et qu'on effectue de nouveau l'élection (avec les mêmes électeurs munis des mêmes opinions), alors le vainqueur ne doit pas changer.
2. Si on ajoute un candidat et qu'on effectue de nouveau l'élection (avec les mêmes électeurs munis des mêmes opinions), alors le nouveau vainqueur doit être soit le même que dans l'élection initiale, soit le candidat ajouté.

Là encore, ce principe semble empreint de bon sens. Si la collectivité des électeurs estime que le candidat a est la meilleure option parmi a, b et c, alors il semble évident qu'en supprimant l'alternative non pertinente c, la collectivité devrait estimer que a est la meilleure option si on doit choisir entre a et b.

Pour illustrer ce principe, on utilise souvent la plaisanterie suivante. Dans un restaurant, le serveur informe un client qu'il a le choix entre du poulet et du bœuf. Le client commande alors du poulet. Quelques minutes plus tard, le serveur revient : « Au fait, j'ai oublié de vous dire qu'il y a aussi du poisson ». Et le client répond : « Très bien, alors je prendrai du bœuf ». Ici, le client semble préférer le bœuf quand les trois options sont disponibles, mais le poulet lorsque seuls le bœuf et le poulet sont proposés, violant ainsi l'hypothèse d'IANP.

Cette plaisanterie permet à la fois de montrer l'aspect relativement naturel de l'hypothèse d'IANP et de discuter sa signification exacte. Dans ce cas particulier, on pourrait, par exemple, imaginer que le client a tout simplement changé d'avis pendant les cinq minutes d'absence du serveur, indépendamment de l'ajout de poisson au menu. Dans ce cas, le changement de décision ne paraît pas absurde. Mais l'IANP concerne un cas plus restreint et plus naturel : on demande simplement que, si les électeurs conservent les mêmes opinions, l'absence ou la présence d'un candidat non choisi n'influence pas le résultat.

On pourrait aussi imaginer que la présence de poisson au menu informe le client qu'il s'agit d'un restaurant d'une catégorie supérieure et que, dans ce cas, il préfère prendre du bœuf. Avec cette interprétation, l'opinion du client ne varie pas dans le temps, ce sont les options examinées qui changent : dans un premier temps, il pense avoir le choix entre « poulet dans un restaurant quelconque » et « bœuf dans un restaurant quelconque » ; mais dans un second temps, il pense avoir le choix entre « poulet dans un restaurant supérieur », « bœuf dans un restaurant supérieur » et « poisson dans un restaurant supérieur ». Il n'y a donc rien de contradictoire dans son changement de décision et elle ne viole qu'en apparence le principe d'IANP.

On pourrait aussi expliquer le comportement du client par diverses autres explications, impliquant, par exemple, des préférences qui ne sont pas transitives. Ainsi, nous verrons une autre interprétation possible de cette expérience de pensée dans l'exemple 1.8. Ceci dit, l'IANP semble tout de même une propriété souhaitable en général, qui garantit une certaine cohérence dans les choix effectués.

Afin d'étendre le vote majoritaire simple tout en respectant le principe d'IANP, il est nécessaire d'élire un candidat w qui, comparé à tout autre candidat c, est préféré par une majorité d'électeurs. En effet, si on enlève tous les autres candidats non pertinents, alors, par principe d'IANP, il est nécessaire que w remporte le vote majoritaire simple contre c. Quand un candidat vérifie cette condition, on dit qu'il est *vainqueur de Condorcet*.

Ramon Llull, érudit majorquin du XIIIe siècle, semble avoir été le premier à décrire des modes de scrutin qui ont la propriété d'élire le vainqueur de Condorcet, comme l'analyse notamment McLean (1990) en s'appuyant sur les écrits originaux de Llull (c. 1285, 1299). Mais, à notre connaissance, c'est Nicolas de Condorcet (1785) qui, le premier, semble avoir formulé explicitement ce principe directeur et surtout avoir remarqué qu'il n'existe pas toujours un tel candidat. En effet, considérons trois électeurs possédant les préférences suivantes.

a	b	c
b	c	a
c	a	b

Dans la notation ci-dessus, chaque colonne représente un électeur. Par exemple, le premier électeur préfère le candidat a au candidat b, qu'il préfère lui-même au candidat c. Avec les préférences ci-dessus, un majorité d'électeurs préfère a à b ; une majorité (qui n'est pas constituée des mêmes électeurs) préfère b à c et une majorité préfère c à a. Il n'y a donc pas de vainqueur de Condorcet : c'est ce qu'on appelle le *paradoxe de Condorcet*. Ce phénomène possède une telle importance en choix social qu'on l'appelle parfois tout simplement le *paradoxe du vote*.

En particulier, l'exemple ci-dessus montre qu'il est impossible d'étendre le vote majoritaire simple tout en respectant le principe d'IANP (qui n'était ni nommé ni formulé aussi explicitement à l'époque de Condorcet). Par exemple, si on décide que a est vainqueur, alors ce n'est pas cohérent avec le résultat d'une élection qui se déroulerait entre a et c, puisque le candidat c sortirait vainqueur de ce duel électoral. De la même manière, aucun candidat n'est un vainqueur cohérent avec le vote majoritaire simple et le principe d'IANP.

Théorème d'Arrow

Arrow (1950) a en quelque sorte généralisé l'observation de Condorcet. Même si son théorème original porte sur des *fonctions de bien-être social*, qui permettent

de classer complètement les options proposées aux agents, il possède une transcription immédiate pour les modes de scrutin, qui se contentent de désigner une option en fonction des préférences de la communauté[5].

En étant relativement modestes dans nos exigences, nous pouvons demander qu'un mode de scrutin possède les propriétés suivantes.

1. Il vérifie l'IANP.
2. Il est *unanime* : si tous les électeurs préfèrent le candidat a, alors celui-ci doit être élu.
3. Il n'est pas *dictatorial* : il n'existe pas un électeur qui possède le pouvoir de décider seul du résultat.
4. De manière implicite, Arrow suppose que le mode de scrutin est *ordinal*.
5. De manière explicite, Arrow suppose que le mode de scrutin est *universel*, au sens où il identifie un candidat vainqueur pour toute combinaison de bulletins autorisés. Dans ce mémoire, nous ferons toujours cette hypothèse implicitement.

Ainsi, on se pose la même question que précédemment : comment obtenir un système de vote qui vérifie l'IANP ? Mais, au lieu de faire l'hypothèse que le mode de scrutin se réduise au vote majoritaire quand il y a deux candidats, on fait les hypothèses *a priori* moins exigeantes[6] qu'il soit unanime, non dictatorial et ordinal.

Pourtant, même avec ces exigences plus faibles, le théorème d'Arrow nous apprend qu'il n'existe aucun mode de scrutin qui les satisfasse (pour un nombre de candidats supérieur ou égal à 3)[7].

En ce qui concerne le caractère unanime et non dictatorial, ces hypothèses semblent très difficiles à abandonner. Pour l'ordinalité, nous y reviendrons. Reste la propriété d'IANP, qui semble donc devoir être sacrifiée pour préserver les autres hypothèses. En pratique, tous les modes de scrutin ordinaux usuels violent effectivement la même hypothèse : l'IANP[8].

Il est important de ne pas réduire le théorème d'Arrow à une simple question procédurale, excluant l'existence d'un mode de scrutin parfaitement satisfaisant en pratique. La conséquence la plus profonde est que, dans une approche ordinale, il n'existe pas de notion canonique de « candidat préféré par la population », si on émet le souhait bien compréhensible que cette notion possède des propriétés naturelles.

Comparaison interpersonnelle des utilités

Pour résoudre les problèmes rencontrés précédemment, il semblerait donc que supprimer l'hypothèse d'ordinalité soit un bon angle d'attaque. Une façon simple de modéliser l'intensité des préférences d'un électeur consiste à recourir à des utilités cardinales : l'intérêt de chaque électeur pour chaque candidat est représenté

5. Nous n'utiliserons pas le terme *fonction de choix social* car il est habituellement réservé à des systèmes qui sont ordinaux. Comme nous le développerons dans la section 1.4, nous utilisons l'expression *mode de scrutin* dans une acception plus large, qui comprend des systèmes non ordinaux.

6. Les hypothèses citées sont *a priori* moins exigeantes car on a supposé l'IANP. Ce n'est pas le cas sinon.

7. On pourra notamment consulter Geanakoplos (2005) pour d'élégantes preuves du théorème d'Arrow, qui ont par ailleurs inspiré plusieurs variantes de la preuve du théorème de Gibbard-Satterthwaite, sur lequel nous reviendrons dans un instant. Par ailleurs, Mossel (2012) et Keller (2012) fournissent des versions « quantitatives » du théorème d'Arrow qui discutent avec quelle probabilité on observe une violation des propriétés exigées.

8. Pour une discussion des hypothèses du théorème d'Arrow, voir Gibbard (2014).

par un nombre réel. Le terme *cardinal* signifie ici, non seulement, que la comparaison entre deux utilités (laquelle est la plus élevée) révèle quel candidat est préféré par l'électeur mais, également, que la valeur numérique des utilités traduit une intensité de préférence.

Plusieurs variantes de ce modèle existent, mais la plus courante est celle de Von Neumann et Morgenstern (1944). Dans ce modèle, les utilités d'un électeur sont définies à une constante additive et à une constante multiplicative strictement positive près. Intuitivement, si on fait l'analogie entre la mesure de la position d'un candidat sur l'axe abstrait des préférences d'un électeur (c'est-à-dire l'axe des utilités) et la mesure de la position d'un objet concret sur une droite, l'électeur peut choisir arbitrairement où il place l'origine du repère (d'où la constante additive) et quelle unité de longueur il utilise (d'où la constante multiplicative).

Dans le cas d'une mesure de longueur, un agent peut prêter sa règle graduée à un autre, ce qui permet au moins d'utiliser la même unité. Mais dans le cas des préférences, l'instrument de mesure reste dans l'univers mental de chaque agent, et il est impossible de savoir si on utilise la même unité de longueur qu'un autre. La question n'a même pas de sens, puisqu'il est impossible de placer les axes de préférences de deux agents dans un même univers mental afin d'y comparer les distances mesurées.

De manière plus générale, réaliser une comparaison interpersonnelle des utilités, c'est-à-dire des intensités de préférence des agents, est impossible sans effectuer une hypothèse supplémentaire nécessairement arbitraire qui, *in fine*, revient à favoriser tel ou tel type d'agent. Nous ne développerons pas davantage ces question complexes ; pour un bon aperçu, le lecteur est invité à consulter Hammond (1991) [9].

On en retiendra, en tout cas, que supprimer l'hypothèse d'ordinalité est loin d'être aussi anodin qu'il n'y paraît. Par ailleurs, nous verrons d'autres raisons de favoriser les modes de scrutin ordinaux dans la suite de ce manuscrit [10].

Théorème de Gibbard-Satterthwaite

De notre point de vue, la principale conclusion du théorème d'Arrow et des problèmes fondamentaux posés par une comparaison interpersonnelle des utilités est qu'il n'existe pas de notion canonique et incontestable de « candidat préféré par la population ». On peut donc se poser la question d'un autre point de vue : en pratique, comment se comporte le mode de scrutin ? En particulier, est-ce qu'il donne le même pouvoir à tous les électeurs en fonction de leur niveau d'information ? D'un point de vue de théorie des jeux, est-ce qu'il est facile d'atteindre des situations d'équilibre ?

De manière plus précise, on dit qu'un mode de scrutin est *manipulable* dans une certaine configuration de préférence des électeurs si, et seulement si, un sous-ensemble des électeurs, en votant de façon non sincère, peut faire élire un candidat qu'ils préfèrent au résultat du vote sincère (en supposant que les autres électeurs votent sincèrement de toute façon). Il paraît intuitivement évident que la manipulabilité est une propriété néfaste et nous discuterons bientôt pourquoi de façon plus détaillée.

Malheureusement, Gibbard (1973) a montré que dès qu'il y a 3 candidats éligibles ou davantage, pour tout mode de scrutin non dictatorial, il existe au

9. Chichilnisky (1985) prouve également, par d'autres arguments, qu'il est essentiellement impossible d'agréger des préférences cardinales par une méthode possédant des propriétés raisonnables.

10. On pourra notamment consulter les résultats théoriques du chapitre 5, les simulations des chapitres 7 et 8 et les expériences du chapitre 9.

moins une configuration où le mode de scrutin est non seulement manipulable, mais par une coalition constituée d'un seul électeur !

Nous nous conformerons à la tradition qui consiste à nommer ce résultat le théorème de Gibbard-Satterthwaite. Cependant, le résultat de Gibbard est à la fois antérieur et plus fort que celui de Satterthwaite (1975) : en effet, il s'applique à tout type de mode de scrutin (ou *game form*, voir la section 1.4), alors que celui de Satterthwaite s'applique seulement aux modes de scrutin ordinaux.

Théorème de Gibbard non déterministe

Si l'on accepte sans réserve le principe d'un recours au hasard, alors il existe des systèmes satisfaisants, contrairement au cas déterministe où nous avons vu que le chemin est pavé de théorèmes d'impossibilité.

Premièrement, on peut recourir au hasard de façon inconditionnelle, sans prendre en compte les préférences des électeurs. Là encore, on trouve des exemples dans la démocratie athénienne antique, où de nombreuses charges étaient attribuées par tirage au sort parmi un ensemble de candidats éligibles.

Deuxièmement, on peut combiner les préférences des électeurs et le hasard. Dans ce cas, il est naturel d'exiger que le mode de scrutin soit non manipulable, unanime et anonyme. Gibbard (1977, 1978) a montré, d'abord pour les modes de scrutin ordinaux puis en général, que le seul mode de scrutin qui vérifie ces hypothèse est la *dictature aléatoire* : chaque électeur vote pour un candidat, puis on tire un bulletin au sort de façon équiprobable et le candidat indiqué sur ce bulletin est élu [11].

Pour certaines applications, ce système peut être intéressant. En particulier, s'il s'agit d'une décision collective qui entraîne des conséquences mesurées et/ou qui sera renouvelée fréquemment, par exemple le choix d'un restaurant par un groupe d'amis, alors il permet un équilibre des pouvoirs tout en éliminant la question du vote stratégique.

En revanche, dans d'autres contextes, un tel usage du hasard restera toujours débattu, hormis en tant que règle complémentaire pour départager des candidats en cas d'égalité parfaite dans les bulletins des électeurs. Dans ce mémoire, nous nous concentrerons exclusivement sur des modes de scrutin *déterministes*, c'est-à-dire que le candidat vainqueur ne dépend que des bulletins des électeurs, sans qu'on fasse appel à un élément aléatoire.

Manipuler, c'est bien. La manipulabilité, c'est mal.

Nous avons dit rapidement que la manipulabilité était une propriété néfaste d'un mode de scrutin. Nous allons maintenant discuter cette question fondamentale.

Ce que nous appelons *manipulation* est la pratique par certains électeurs du vote stratégique. Il nous semble important de la distinguer de la *manipulabilité*, qui est le fait que la manipulation puisse fonctionner, c'est-à-dire mener à un résultat différent du vote sincère. Autrement dit, la manipulabilité d'un mode de scrutin dans une certaine configuration des préférences des électeurs est le fait que le vote sincère ne soit pas un *équilibre de Nash fort* (ENF) : une coalition d'électeurs peut dévier du vote sincère et obtenir un résultat qu'ils jugent préférable au vainqueur résultant d'un vote sincère.

11. Plus généralement, Gibbard (1978) montre, en autorisant les modes de scrutin non déterministes, quels sont tous les systèmes non manipulables. Le cas anonyme et unanime que nous citons pour simplifier n'est qu'un corollaire de ce résultat général.

Défense de la manipulation

La manipulation, c'est-à-dire la pratique du vote stratégique, est parfois vue d'un mauvais œil. Voici les principaux arguments de ce point de vue, que nous allons discuter.

1. Les manipulateurs sont des tricheurs.
2. La manipulation conduit à un résultat « incorrect » pour l'élection.
3. La manipulation s'exerce au détriment de la communauté.

L'argument 1 est défendable si on attache une dimension morale au vote sincère, mais elle s'oppose à tous les points de vue modernes sur le *mechanism design* et la *Nash-implémentabilité*[12] : en général, on considère de nos jours que les agents sont stratégiques et que la problématique est de trouver une règle du jeu qui conduise à un résultat satisfaisant pour le bien-être social, en acceptant — et la plupart du temps, en exploitant — ce comportement stratégique. En économie, un tel point de vue remonte essentiellement à la « main invisible » d'Adam Smith, et cette idée peut être naturellement transposée en théorie du vote : de notre point de vue, l'électeur stratège n'est donc pas un tricheur mais un agent qui contribue à rechercher et peut-être à obtenir un équilibre de Nash.

En ce qui concerne l'argument 2, on peut lui opposer le théorème d'Arrow (dans un cadre ordinal) et les problèmes fondamentaux de comparaison interpersonnelle des utilités (dans un cadre cardinal) : il n'y a pas de notion canonique de « candidat préféré par la population ». Parler de résultat « incorrect » n'a donc pas vraiment de sens.

L'argument 3 présente des similarités avec le précédent. Mais alors que l'argument 2 suppose l'existence d'une sorte de vérité supérieure (le résultat « correct », étant donné les préférences sincères de la population), l'argument 3 est plus pragmatique et revient finalement à la question suivante : quels électeurs préfèrent le vainqueur de l'élection sincère, et lesquels préfèrent le vainqueur résultant d'une manipulation ?

Dans le cas général, il est clairement possible que certains électeurs soient moins satisfaits par le vainqueur alternatif que par le vainqueur sincère. Mais cela ne peut pas être le cas pour tous les électeurs : en effet, au moins les manipulateurs préfèrent le vainqueur alternatif par définition. En revanche, nous allons voir que l'inverse peut se produire : il est possible que la totalité des électeurs préfèrent le vainqueur alternatif au vainqueur sincère (c'est-à-dire que le premier *Pareto-domine* le second).

En effet, considérons l'exemple suivant. On dispose de 26 candidats, désignés par les lettres de l'alphabet. On utilise le mode de scrutin Veto, muni de la règle de départage alphabétique.

1. Chaque électeur émet un veto[13], c'est-à-dire vote contre un unique candidat.
2. Le candidat recevant le moins de vetos est élu.
3. En cas d'égalité, le premier candidat dans l'ordre alphabétique, parmi les ex aequo, est déclaré vainqueur.

Supposons qu'il y a 25 électeurs et qu'ils préfèrent unanimement les candidats dans l'ordre inverse alphabétique (ils préfèrent z et détestent a). S'ils votent sincèrement, il émettent tous un veto contre a et, par règle de départage, c'est b qui est élu : il s'agit du deuxième pire candidat pour tous les électeurs !

12. Voir par exemple Feldman (1980).

13. Dans ce mémoire, nous appelons *veto*, avec une minuscule, un vote négatif contre un candidat. Le mot *Veto*, avec une majuscule, désigne le mode de scrutin décrit ci-dessus, connu dans la littérature anglophone sous la désignation d'*Antiplurality* ou simplement *Veto*.

Si un électeur est conscient de cet état de fait, il peut décider de voter stratégiquement contre b et c'est alors c qui est élu, améliorant le sort de toute la communauté. Cette manipulation n'est donc pas néfaste. On peut même aller plus loin : si les électeurs se coordonnent habilement pour voter contre tous les candidats sauf z, alors celui-ci est élu, ce qui satisfait au mieux tous les électeurs.

Pour résumer, il est donc possible qu'une manipulation s'exerce au bénéfice de toute la communauté. À l'inverse, même s'il est possible qu'elle s'exerce au détriment d'une partie de la communauté, elle ne peut pas nuire à toute la communauté. L'argument 3 est donc assez circonstanciel et discutable en général[14].

Manipulabilité et équilibres de Nash forts (ENF)

Continuons avec un exemple plus pratique, puisqu'il appartient désormais à l'Histoire. L'élection présidentielle française de 2002 s'est déroulée en utilisant l'uninominal à deux tours. Au premier tour, Jacques Chirac (droite) a reçu 19,9% des votes, Jean-Marie Le Pen (extrême-droite) 16,9%, Lionel Jospin (gauche) 16,2% et 13 candidats divers ont partagé le reste. Au second tour, Chirac fut vainqueur par 82,2% des voix contre Le Pen.

Cependant, selon certains sondages de l'époque, Jospin aurait remporté le second tour contre n'importe quel adversaire, c'est-à-dire qu'il était vainqueur de Condorcet. Il est impossible de vérifier si c'était effectivement le cas sur l'ensemble de l'électorat français ; mais cela permet de savoir que la réalité était suffisamment proche de ce cas de figure pour qu'on puisse considérer comme un exemple réaliste une situation où les votes observés auraient été sincères et où Jospin aurait été vainqueur de Condorcet.

Sous cette hypothèse, il existait une possibilité de manipulation : si tous les électeurs préférant Jospin à Chirac avaient voté pour Jospin au premier tour, alors Jospin aurait été élu directement au premier tour. Autrement dit, le vote sincère n'était pas un équilibre de Nash fort (ENF). Mais les électeurs n'ont pas accompli cette manipulation, et c'est une observation essentielle sur laquelle nous reviendrons.

Cette situation pose principalement deux types de problèmes :

1. le sentiment *a posteriori* des électeurs ;
2. la légitimité du résultat.

D'une part, après l'élection, certains électeurs sincères peuvent ressentir des regrets sur le choix de leur bulletin et aussi un sentiment d'injustice : puisque des bulletins non sincères auraient mieux défendu leur opinion, ils peuvent estimer que leurs bulletins sincères n'ont pas eu l'impact qu'ils méritaient. Ils peuvent en retirer une méfiance à l'égard du mode de scrutin utilisé lui-même : le fait qu'un bulletin sincère ne défende pas au mieux l'opinion qu'il exprime apparaît comme un *bug* du système[15]. Cette expérience peut même développer ou renforcer leur méfiance de la pratique élective en général.

D'autre part, on peut craindre pour le résultat d'une telle élection. Sur cet aspect, cet exemple peut recevoir deux interprétations, mais elles mènent toutes deux à condamner la manipulabilité. Si on estime que le résultat du vote sincère

14. Pour une étude de l'impact négatif potentiel d'une manipulation sur le bien-être social, on pourra consulter Brânzei *et al.* (2013).

15. Qu'on me pardonne une anecdote personnelle : quand j'ai commencé à m'intéresser à la théorie du vote, à la suite de ce fameux 21 avril 2002, j'avais retiré de cette élection l'intime conviction que l'uninominal à deux tours était un mode de scrutin catastrophique en termes de manipulabilité. Nous verrons cependant dans les chapitres 7, 8 et 9 que, même si ce n'est pas le mode de scrutin le moins manipulable en général, c'est pourtant loin d'être le pire, même pour une quinzaine de candidats.

représente par définition au mieux les opinions des électeurs, alors Chirac était le vainqueur légitime ; mais dans ce cas, la manipulabilité de la situation aurait risqué de conduire à l'élection de Jospin si les électeurs concernés avaient été plus habiles. À l'inverse, si on estime qu'un résultat manipulé peut être meilleur en termes de bien-être collectif (comme le défendraient en l'occurrence les partisans du critère de Condorcet), alors la manipulation elle-même n'est pas indésirable, mais la *manipulabilité* de la situation l'est tout de même : en effet, elle rend ce « meilleur » résultat plus difficile à identifier et à produire. Par exemple, si toute la population vote sincèrement, il n'est pas atteint.

Ainsi, il semble qu'en l'occurrence, il y avait une différence entre le résultat constaté de l'élection et celui d'un éventuel ENF. Il est amusant de constater que, dans d'autres contextes de théorie des jeux, il peut arriver d'étudier les ENF parce qu'on pense que ce sont les configurations vers lesquelles les agents vont naturellement converger. En théorie du vote, il nous semble que la situation est légèrement différente : ce sont des situations vers lesquelles il serait souhaitable de converger pour les raisons exposées ci-dessus ; mais en pratique, cet objectif peut être difficile à atteindre car le vote sincère ne mène pas nécessairement à un ENF. Or ce problème correspond précisément à la définition de la manipulabilité.

Comme tout exemple réel, l'élection présidentielle française de 2002 doit être examinée avec plus de précautions qu'un exemple artificiel, qu'il serait possible d'ajuster à loisir pour illustrer un argument de façon paradigmatique. En particulier, il faut ici bien distinguer le fait que la situation ne soit pas un ENF du fait que le vainqueur ne soit pas le vainqueur de Condorcet (bien que les deux aspects soient liés). Le premier peut entraîner divers types de problèmes, que nous avons décrits ; le second peut également poser un problème de légitimité du vainqueur, puisqu'une majorité d'électeurs préfère le vainqueur de Condorcet au candidat élu. Heureusement, nous verrons (notamment au chapitre 2) que lutter contre ces deux types de problèmes n'est pas incompatible, bien au contraire.

Par ailleurs, la situation que nous avons décrite a eu aussi une dimension symbolique importante, liée au déroulement pratique de l'uninominal à deux tours et à son analogie de principe avec certaines compétitions sportives. Ainsi, il est apparu choquant à certains électeurs qu'un candidat d'extrême-droite parvienne en « finale » de la compétition et apparaisse d'une certaine façon comme deuxième dans le « classement » final. Nous ne développerons cependant pas ce point outre mesure : tout en gardant en tête que la dimension symbolique et la perception publique des événements est toujours importante, particulièrement pour une élection politique, nous considérerons dans ce mémoire que le résultat principal d'un scrutin est tout de même le choix du candidat vainqueur.

Manipulabilité et *straightforwardness*

Par définition, la manipulabilité d'un mode de scrutin signifie que la situation de vote sincère n'est pas forcément un ENF. Ceci pose plusieurs problèmes.

1. Avant l'élection, l'électeur est confronté à un dilemme : voter sincèrement ou stratégiquement ? S'il accorde une vertu morale au vote sincère, il peut s'agir d'un cas de conscience. Sinon, il reste tout de même un problème pratique : comment choisir un vote stratégique adapté au mieux à la situation ?
2. Ces aspects stratégiques induisent *de facto* une asymétrie de pouvoir entre les électeurs informés et stratèges et ceux qui ne sont ni l'un ni l'autre. Comme le dit Dodgson dans une phrase rendue célèbre par Black (1958), le

vote devient alors « davantage un jeu d'habileté qu'un test réel des souhaits des électeurs » [16].

3. Comme on l'a vu dans l'exemple de vote en Veto, si la situation est manipulable, alors les électeurs peuvent avoir besoin d'information, de pouvoir de calcul et de coordination pour aboutir à un ENF (même dans une situation, *a priori* favorable, où ils sont tous d'accord). À l'inverse, si la situation est non manipulable, il suffit de voter sincèrement pour atteindre un tel équilibre ; aucun échange d'information, aucun calcul et aucune coordination n'est alors nécessaire.
4. Ce besoin d'information, individuel ou collectif, octroie un pouvoir discutable aux sources d'information, comme les médias et les instituts de sondage.

Nous allons montrer que ces problèmes sont profondément liés à une question intimement connectée à la manipulabilité, la *straightforwardness*. On dit qu'un mode de scrutin est *straightforward* (Gibbard, 1973) si tout électeur, quelles que soient ses opinions, dispose d'une stratégie dominante. Afin de bien comprendre la nuance entre cette notion et la non-manipulabilité, considérons le cas simple où il n'y a que deux candidats et examinons les trois modes de scrutin suivants. Les deux derniers sont plutôt des curiosités théoriques, mais ils vont nous permettre d'illustrer notre point de vue.

Vote majoritaire simple Chaque électeur est censé annoncer son candidat préféré, a ou b, et le candidat recevant le plus de voix est élu.

Vote majoritaire inversé Chaque électeur est censé annoncer son candidat préféré, a ou b, mais son vote est comptabilisé en faveur de l'autre candidat.

Vote de parité Chaque électeur place une boule blanche ou une boule noire dans l'urne. S'il y a un nombre impair de boules noires, alors a est élu. Dans le cas contraire, b est élu.

Le vote majoritaire simple est non manipulable. En particulier, comme nous l'avons déjà remarqué, chaque électeur dispose toujours d'une stratégie dominante (qui est simplement le vote sincère) : autrement dit, ce mode de scrutin est *straightforward*.

Le vote majoritaire inversé, lui, est généralement manipulable : en effet, si un électeur préfère a, il a intérêt à délaisser un bulletin avec l'étiquette a pour lui préférer un bulletin avec l'étiquette b, puisque celui-ci sera comptabilisé en faveur de a. En revanche, ce mode de scrutin est *straightforward* : un électeur qui préfère a a toujours intérêt à utiliser un bulletin b, et réciproquement. Or ce mode de scrutin, s'il peut, à la limite, causer une certaine confusion chez l'électeur, ne pose aucun des autres problèmes mentionnés : il est facile de trouver des ENF sans aucun calcul, sans échange d'information et sans coordination particulière. Ces problèmes sont donc bien liés aux défauts de *straightforwardness* et pas à la manipulabilité. On voit d'ailleurs sur cet exemple que la notion de vote sincère est en partie conventionnelle : on aurait pu, au contraire, interpréter le bulletin avec l'étiquette a comme un veto contre a. Avec cette convention de langage, on aurait conclu que ce mode de scrutin n'est jamais manipulable.

Examinons à présent le vote de parité. Si un électeur connaît les bulletins des autres électeurs, alors il peut toujours décider du candidat vainqueur. La plupart du temps, ce mode de scrutin est donc manipulable. Et il n'est particulièrement pas *straightforward* : si un électeur préfère a, il n'a essentiellement aucune indication pour savoir si une boule blanche ou noire défendra au mieux son opinion.

16. En version anglaise originale : « more of a game of skill than a real test of the wishes of the electors ».

Dans ce mode de scrutin, tous les problèmes évoqués précédemment sont exacerbés : par exemple, même si tous les électeurs préfèrent le même candidat, il leur faut une coordination parfaite pour l'élire. Par ailleurs, l'équilibre des pouvoirs entre les électeurs est particulièrement détruit : un électeur disposant d'une information parfaite possède un pouvoir absolu, alors qu'un électeur ignorant ne possède essentiellement aucun pouvoir de décision.

Résumons : nous avons examiné un mode de scrutin non manipulable (donc *straightforward*), un mode de scrutin *straightforward* mais manipulable, et un mode de scrutin non *straightforward* (donc manipulable). Et nous avons constaté que les problèmes évoqués sont en fait davantage liés au défaut de *straightforwardness* qu'à la manipulabilité, puisque ces problèmes sont absents du second mode de scrutin que nous avons examiné.

Cependant, il y a un lien profond entre la manipulabilité et la *straightforwardness*. En effet, quitte à définir conventionnellement le vote sincère comme l'utilisation de la stratégie dominante (comme nous l'avons fait en transformant le vote majoritaire inversé en Veto), alors un mode de scrutin *straightforward* est non-manipulable. Réciproquement, si un mode de scrutin n'est pas *straightforward*, alors il n'existe aucune façon de définir le vote sincère qui le rend non-manipulable.

En résumé, par un choix adéquat de la *fonction de sincérité*, c'est-à-dire de la façon conventionnelle dont on associe un vote réputé sincère à une opinion, la question de la *straightforwardness*, qui est au fond le point fondamental, peut donc se ramener à la question de la manipulabilité, plus commode à appréhender en pratique.

Indicateurs de manipulabilité

Le théorème de Gibbard (1973) nous enseigne qu'aucun mode de scrutin non trivial[17], qu'il soit ordinal ou cardinal, ne peut être *straightforward*. Autrement dit, quelle que soit la fonction de sincérité utilisée, un système de vote non trivial est nécessairement manipulable.

Tout ce que l'on peut espérer, c'est donc de limiter l'ampleur du problème, en étudiant dans quelle mesure les modes de scrutin classiques sont manipulables et en identifiant des procédés permettant de concevoir des modes de scrutin moins manipulables.

À cet effet, notre indicateur de référence sera le taux de manipulabilité d'un mode de scrutin : dépendant de la *culture*, c'est-à-dire la distribution probabiliste des préférences de la population, ce taux est défini comme la probabilité que le mode de scrutin soit manipulable dans une configuration tirée au hasard. Autrement dit, c'est la probabilité qu'un vote sans échange d'information conduise à un ENF.

Dans la littérature, on rencontre principalement deux autres façons d'estimer la manipulabilité d'un mode de scrutin :

17. Ici, nous utilisons le raccourci de langage *non trivial* pour signifier : non réduit à deux candidats éligibles et non dictatorial.

1. Le nombre de manipulateurs nécessaires, ainsi que d'autres types d'indicateurs similaires qui quantifient la manipulabilité quand celle-ci est possible[18].
2. La complexité algorithmique de la manipulation[19].

Dans les deux cas, on estime généralement que plus la manipulation est difficile, plus il s'agit d'une propriété louable du mode de scrutin. Tout en comprenant et en respectant ce point de vue, il nous paraît intéressant et pertinent de défendre précisément l'opinion inverse.

En effet, nous avons défendu l'idée que parvenir facilement à un ENF est une bonne propriété pour un mode de scrutin. Pour cela, le meilleur cas est celui où le mode de scrutin n'est pas manipulable, puisque cela signifie qu'on est capable d'atteindre un équilibre sans échange d'information et sans calcul particulier. Mais dans les autres cas, pour avoir les meilleures chances d'atteindre un équilibre, il vaut mieux que voter stratégiquement soit peu coûteux en nombre d'électeurs, en information, en complexité de calcul et en communication.

Repensons à l'exemple de Veto cité précédemment. Dans la situation que nous avions décrite, pour atteindre un équilibre, il faut que 24 électeurs sur 25 votent stratégiquement. Certains pourraient estimer que c'est une bonne chose, car cette manipulation est très difficile à effectuer. Au contraire, nous pensons que c'est une propriété néfaste, puisqu'il faut beaucoup d'efforts pour atteindre le seul résultat raisonnable !

Considérons un autre exemple, le vote par assentiment. Chaque électeur vote pour autant de candidats qu'il le souhaite, et le candidat recevant le plus de votes est déclaré vainqueur. En pratique, nous verrons, dans cette thèse, que ce mode de scrutin est souvent manipulable. Cependant, il possède au moins l'avantage que la question stratégique y est relativement simple : comme l'expose Laslier (2009), il suffit d'avoir accès à des sondages donnant les deux candidats favoris pour pouvoir utiliser une stratégie efficace, la *Leader rule*. En outre, l'utilisation de ce comportement stratégique conduit à élire le vainqueur de Condorcet s'il existe. Ainsi, même dans les cas où le vote sincère ne conduit pas à un équilibre, il est possible d'en trouver un avec un coût relativement faible[20].

En résumé, il nous semble qu'une bonne propriété est que la manipulation soit facile. Et la manipulation idéale, puisqu'elle ne demande aucune information, c'est tout simplement le vote sincère.

Nos objectifs principaux

Toutes ces considérations étant exposées, il est temps de présenter les buts principaux qui ont orienté notre recherche.

18. On pourra consulter Chamberlin *et al.* (1984); Saari (1990); Lepelley et Valognes (1999); Slinko (2004); Favardin et Lepelley (2006); Pritchard et Slinko (2006); Pritchard et Wilson (2007); Xia et Conitzer (2008); Lepelley *et al.* (2008); Reyhani *et al.* (2009); Reyhani (2013); Green-Armytage (2014). Pour divers indicateurs de manipulabilité liés à la manipulation par un seul électeur, on pourra notamment lire Aleskerov et Kurbanov (1999).
19. Cet axe de recherche a été initié par Bartholdi *et al.* (1989a); Bartholdi et Orlin (1991). Depuis, divers résultats de complexité de manipulation ont été prouvés pour des systèmes de vote classiques, sur lesquels nous reviendrons. Dans le même esprit, Conitzer et Sandholm (2003), Elkind et Lipmaa (2005a), Elkind et Lipmaa (2005b) proposent des méthodes pour transformer un mode de scrutin de façon à augmenter la complexité de manipulation. Cependant, Conitzer et Sandholm (2006); Procaccia et Rosenschein (2006, 2007); Faliszewski et Procaccia (2010) montrent qu'il est essentiellement impossible d'avoir un système de vote qui possède des propriétés raisonnables et qu'il est algorithmiquement difficile de manipuler en moyenne.
20. On pourra également consulter Myerson (1999).

Premièrement, nous souhaitons quantifier la manipulabilité. Est-elle fréquente en pratique ? Parmi les modes de scrutin, lesquels sont les moins manipulables ? Les différences sont-elles importantes ?

Deuxièmement, en étudiant les taux de manipulabilité, nous avons rapidement constaté le problème suivant : on ne connaît pas le taux minimal qu'il est possible d'atteindre dans une classe donnée de modes de scrutin[21]. Par conséquent, on peut comparer les systèmes de vote entre eux, mais on ne peut pas dire si les taux de manipulabilité constatés sont éloignés d'un optimum. Dans l'idéal, nous souhaiterions identifier un système de vote de manipulabilité minimale, ou au moins d'estimer le taux de manipulabilité correspondant : même si le mode de scrutin obtenu était trop complexe pour être utilisé en pratique, il nous fournirait un mètre étalon pour jauger les taux de manipulabilité des autres modes de scrutin.

Contributions et plan

Étude théorique de la manipulabilité

Le chapitre 1 développe le formalisme des *espaces électoraux*, qui permet d'appréhender tous les types de modes de scrutin, y compris les systèmes cardinaux et même d'autres types de systèmes. Nous profitons de ce chapitre pour rappeler la définition formelle des notions Condorcet et présenter les modes de scrutin étudiés dans cette thèse. Nous présentons notamment le système *VTI*, aussi connu par ses acronymes anglophones *STV* ou *IRV*, qui joue un rôle important dans les chapitres suivants. Nous introduisons aussi le système *VTID*, une variante de VTI qui nous a été suggérée par Laurent Viennot. Enfin, le système *BI* est, à notre connaissance, une contribution originale.

Dans le chapitre 2, nous montrons que, pour tous les modes de scrutin classiques sauf Veto, on peut rendre le système moins manipulable en ajoutant un test préliminaire visant à élire le vainqueur de Condorcet s'il existe[22]. Nous appelons ce système la *condorcification* du mode de scrutin initial. Nous discutons précisément avec quelles notions Condorcet ce résultat est valide et nous montrons que la diminution de manipulabilité obtenue est stricte pour tous les modes de scrutin classiques sauf Veto, en utilisant une notion que nous introduisons et caractérisons, le *vainqueur de Condorcet résistant*.

Les théorèmes de condorcification supposent que le mode de scrutin étudié vérifie une propriété que nous appelons le *critère de la coalition majoritaire informée* (**CMInf**), qui signifie qu'une majorité stricte d'électeurs a toujours le pouvoir de choisir le résultat de l'élection, s'ils connaissent les bulletins des autres électeurs. En pratique, tous les modes de scrutin classiques vérifient cette hypothèse, sauf Veto. Dans le chapitre 3, nous définissons plusieurs autres critères majoritaires, dont certains sont des contributions originales, et nous étudions leur lien avec la difficulté de trouver des ENF. Pour l'ensemble des modes de scrutin classiques, nous étudions quels sont les critères vérifiés par chacun.

21. Chaque fois que nous mentionnons la notion de manipulabilité minimale, c'est dans une certaine classe de modes de scrutin « raisonnables », que nous définirons un peu plus tard. En effet, si l'on considérait tous les modes de scrutin, la question serait triviale, puisque la dictature n'est pas manipulable du tout.

22. Un autre mode de scrutin un peu moins classique n'est pas concerné par ce résultat, tout comme Veto : il s'agit de la méthode de Kim-Roush, qui est inspirée de Veto et dont nous rappellerons la définition. Dans la suite, nous verrons que la plupart des théorèmes qui ne s'appliquent pas à Veto ne s'appliquent pas non plus à la méthode de Kim-Roush.

Même si une large catégorie de systèmes vérifie l'hypothèse **CMInf**, ce n'est pas toujours le cas. Dans le chapitre 4, nous étudions comment diminuer la manipulabilité en général, grâce à un processus que nous appelons *condorcification généralisée*, en utilisant une approche inspirée de la théorie des *jeux simples*. Pour les systèmes usuels vérifiant **CMInf**, nous utilisons ce formalisme pour montrer que la condorcification au sens usuel, c'est-à-dire basée sur le vainqueur de Condorcet, est, en un certain sens, optimale.

Dans le chapitre 5, nous étudions l'influence du caractère ordinal d'un mode de scrutin sur sa manipulabilité. Nous montrons que, pour tout mode de scrutin non ordinal, il existe un mode de scrutin ordinal qui possède certaines propriétés en commun avec le mode de scrutin original et qui est au plus aussi manipulable, à condition que la culture vérifie une condition que nous introduisons, la *décomposabilité*. En particulier, nous montrons que ce théorème est applicable quand les électeurs sont indépendants au sens probabiliste. En combinant ce résultat avec les théorèmes de condorcification, nous concluons que la recherche d'un mode de scrutin de manipulabilité minimale (au sein de la classe des systèmes vérifiant **CMInf**) peut être restreinte aux modes de scrutin qui sont ordinaux et qui vérifient le critère de Condorcet.

Étude assistée par ordinateur de la manipulabilité

À partir du chapitre 6, on met provisoirement de côté la recherche d'un mode de scrutin de manipulabilité minimale et on tente de quantifier la manipulabilité des modes de scrutin classiques. Pour cela, on présente SVVAMP, un package Python de notre cru dédié à l'étude des modes de scrutin et de leur manipulation. Son architecture logicielle est modulaire, permettant l'implantation rapide de nouveaux modes de scrutin. Utilisant les critères définis au chapitre 3, SVVAMP possède des algorithmes génériques de manipulation et des algorithmes spécifiques pour certains modes de scrutin, soit issus de la littérature, soit conçus spécialement pour ce package logiciel. SVVAMP sera utilisé dans tous les chapitres suivants.

Dans le chapitre 7, nous étudions le taux de manipulabilité des différents modes de scrutin dans des cultures *sphéroïdales*, qui généralisent le modèle habituellement connu sous le nom de *culture impartiale* (voir par exemple Nitzan, 1985). Notamment, nous utilisons, pour la première fois, le modèle de Von Mises–Fisher pour modéliser des préférences et nous expliquons les raisons de ce choix. Nous étudions l'effet de la variation des différents paramètres et nous introduisons des diagrammes de *méta-analyse*, qui permettent de comparer la manipulabilité des modes de scrutin étudiés. Nous montrons notamment que le mode de scrutin *CVTI*, obtenu par condorcification de VTI, est généralement le moins manipulable.

Dans le chapitre 8, nous étudions un autre modèle, basé sur un spectre politique abstrait et généralisant la notion de culture *single-peaked* introduite par Black (1958). Ce cadre de travail différent nous permet de nuancer notre conclusion sur le système CVTI : en effet, sa suprématie est alors plus discutable que dans les cultures sphéroïdales. En particulier, d'autres modes de scrutin présentent des performances intéressantes, comme CSD, BI, l'uninominal à deux tours et d'autres.

Dans le chapitre 9, nous analysons des expériences réelles, issues de contextes très différents, y compris des contextes de révélation de préférences qui ne sont pas des élections. Ces expériences nous permettent de confirmer des résultats antérieurs selon lesquels il existe fréquemment un vainqueur de Condorcet (par exemple Tideman, 2006) et d'établir que le système CVTI se distingue par sa faible manipulabilité.

Le chapitre 10 reprend la recherche d'un mode de scrutin de manipulabilité minimale (dans la classe des systèmes vérifiant **CMInf**). On définit le *graphe d'opportunité* d'un espace électoral et on montre que la question peut se ramener à un problème d'optimisation linéaire en nombres entiers, qu'on peut étudier pour de très modestes valeurs des paramètres en combinant une approche théorique et l'utilisation d'un logiciel dédié (CPLEX). Enfin, on utilise SVAMP pour comparer les modes de scrutin classiques avec l'optimum.

Annexes

Dans l'annexe A, nous étudions les questions de mesurabilité, c'est-à-dire un point technique lié à l'utilisation d'espaces probabilisés, principalement utile pour le chapitre 5 sur le tranchage.

Dans l'annexe B, nous exposons un travail réalisé en collaboration avec Benoît Kloeckner sur la géométrie de l'espace des utilités. Nous y motivons notamment l'utilisation des cultures sphéroïdales du chapitre 7.

Dans l'annexe C, nous illustrons comment l'étude des modes de scrutin et de leur manipulabilité peut apporter des réponses à des questions de télécommunication, en utilisant un modèle initialement conçu par Ludovic Noirie. Nous y montrons notamment que le système VTI permet de concilier la non-manipulabilité et l'efficacité économique.

Dans le glossaire (page 311), le lecteur pourra trouver un résumé des principales notations et des sigles utilisés dans cette thèse.

Nous vous souhaitons une bonne lecture !

Première partie

Étude théorique de la manipulabilité

Chapitre 1

Formalisme

Dans la littérature de choix social arrovien, les préférences d'un électeur sont généralement représentées par un ordre strict de préférence sur les candidats (Arrow, 1950). Un mode de scrutin (ou *fonction de choix social*) est une fonction qui, aux préférences de tous les électeurs, associe un candidat vainqueur. Cependant, avec ce formalisme, il est impossible d'appréhender correctement des modes de scrutin comme le *vote par assentiment* ou le *vote par notation*.

Nous développons le formalisme des *espaces électoraux* (section 1.1), une classe de modèles capables d'utiliser des objets mathématiques arbitraires pour représenter les préférences des électeurs, tout en permettant de prouver des théorèmes généraux sur ces modèles. Dans ce cadre, nous définissons en section 1.2 les *systèmes de vote basés sur l'état* (SVBE), qui constitueront l'objet principal de notre étude, et nous traduisons la définition habituelle de la *manipulabilité* dans ce modèle (section 1.3). Puis nous définissons les *systèmes de vote généraux*, qui permettent de représenter tous les modes de scrutin imaginables en pratique (section 1.4). Nous exposons comment on peut naviguer entre systèmes de vote généraux et SVBE et nous montrons pourquoi l'étude peut être restreinte à ces derniers.

Nous profitons ensuite de ce premier chapitre généraliste pour rappeler les notions de matrice des duels et de vainqueur de Condorcet (section 1.5). Nous prenons garde à bien distinguer les différentes variantes de cette notion, liées au fait que nous autorisons des préférences qui sont des relations binaires quelconques, par exemple des ordres faibles. En particulier, nous généralisons la notion habituelle de *vainqueur de Condorcet faible* en définissant la notion de candidat *Condorcet-admissible*, dont les chapitres suivants révèleront qu'il s'agit d'une notion centrale pour la manipulabilité de presque tous les modes de scrutin usuels.

Enfin, nous présentons les modes de scrutin que nous étudierons dans cette thèse (section 1.6). Parmi eux, le système *BI* est, à notre connaissance, une contribution originale. Il en est de même des systèmes *Condorcet-doyen* et *Condorcet-dictature*, qui servent avant tout de matière première commode pour des exemples théoriques. Nous introduisons également le système *VTID* ; ce mode de scrutin nous a été suggéré par Laurent Viennot, que nous remercions.

1.1 Espaces électoraux

Dans cette section, nous introduisons le formalisme des *espaces électoraux*, qui nous permettra à la fois de représenter les préférences et les bulletins des électeurs.

1.1.1 Relations binaires

Avant toute chose, quelques rappels s'imposent sur les relations binaires. Soit E un ensemble et $\mathrm{P}_0 \in \mathcal{P}(E^2)$ une relation binaire sur E. On dit que P_0 est :

- *réflexive* ssi $\forall c \in E, c\ \mathrm{P}_0\ c$;
- *irréflexive* ssi $\forall c \in E, \mathrm{non}(c\ \mathrm{P}_0\ c)$ (i.e. ssi non P_0 est réflexive) ;
- *faiblement complète*, ou simplement *complète*[1], ssi $\forall (c,d) \in E^2, c \neq d \Rightarrow c\ \mathrm{P}_0\ d$ ou $d\ \mathrm{P}_0\ c$;
- *antisymétrique* ssi $\forall (c,d) \in E^2, c\ \mathrm{P}_0\ d$ et $d\ \mathrm{P}_0\ c \Rightarrow c = d$ (i.e. ssi non P_0 est complète) ;
- *transitive* ssi $\forall (c,d,e) \in E^3, c\ \mathrm{P}_0\ d$ et $d\ \mathrm{P}_0\ e \Rightarrow c\ \mathrm{P}_0\ e$;
- *négativement transitive* ssi non P_0 est transitive.

On dit que P_0 est un *ordre strict faible* ssi elle est négativement transitive, irréflexive et antisymétrique ; un *ordre strict total* ssi elle est transitive, irréflexive et faiblement totale.

1.1.2 Profils

Nous pouvons, à présent, représenter les préférences des électeurs sur les candidats.

Soit V et C deux entiers strictement positifs. Notons $\mathcal{V} = [\![1, V]\!]$ l'ensemble des indices des *électeurs* et $\mathcal{C} = [\![1, C]\!]$ l'ensemble des indices des *candidats*, où la notation avec crochet ajouré désigne l'ensemble des entiers inclus dans l'intervalle fermé. Les candidats peuvent être des électeurs eux-mêmes, sans que cela n'altère nos résultats[2].

On note $\mathcal{R}_\mathcal{C}$ l'ensemble des relations binaires sur $\mathcal{C}$: un élément de $\mathcal{R}_\mathcal{C}$ représente la relation binaire de préférence d'un électeur sur les candidats. On note $\mathcal{F}_\mathcal{C}$ l'ensemble des ordres stricts faibles et $\mathcal{L}_\mathcal{C}$ l'ensemble des ordres stricts totaux sur $\mathcal{C}$.

Notons $\mathcal{R} = (\mathcal{R}_\mathcal{C})^V$. Comme il est d'usage en choix social, nous appelons un élément de $\mathcal{R}$ un *profil* ; pour chaque électeur, il donne sa relation binaire de préférence sur les candidats.

Quand les relations de préférence des électeurs sont des ordres stricts (totaux ou faibles), on représente un profil comme dans l'exemple suivant.

44	32	24
a	b	c
b, c	c	a
	a	b

En l'occurrence, les 44 premiers électeurs préfèrent a et sont indifférents entre b et c, les 32 suivants préfèrent b à c et c à a, etc. Dans les cas où chaque colonne correspond à un seul électeur, on omettra la première ligne, qui indique dans le cas général le nombre d'électeurs de chaque colonne. On pourra également indiquer le nombre d'électeurs de chaque colonne en pourcentage du nombre total d'électeurs V.

1. On dit que P_0 est *fortement complète* quand elle est également réflexive. Selon les ouvrages, le qualificatif *complète*, utilisé seul, peut désigner l'une ou l'autre notion. Comme nous utiliserons généralement le mot *complète* pour des relations irréflexives, il signifiera nécessairement *faiblement complète* donc il n'y aura pas de confusion possible.

2. En revanche, nous considérons que l'ensemble des candidats est fixé *a priori*. Une extension intéressante du modèle, que nous n'aborderons pas dans ce mémoire, consiste à autoriser les candidats potentiels à entrer en lice ou à se retirer pour des raisons stratégiques (Dutta *et al.*, 2001, 2002; Lang *et al.*, 2013; Brill et Conitzer, 2015).

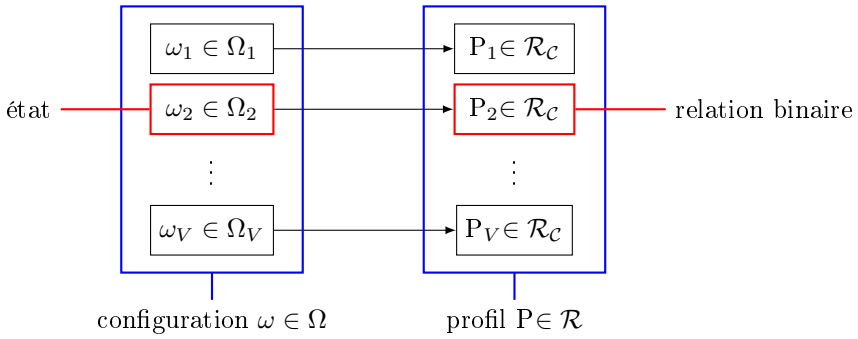


FIGURE 1.1 – Configuration et profil des électeurs

En abrégé, on pourra également écrire que l'ordre faible de préférence des 44 premiers électeurs est $(a \succ b \sim c)$, où le symbole $\succ$ représente une préférence stricte et le symbole $\sim$ une indifférence.

1.1.3 Définition d'un espace électoral

Avant de définir la notion d'*espace électoral* en général, nous présentons un modèle qui permet d'étudier tous les modes de scrutin courants et que nous utiliserons fréquemment pour illustrer nos résultats.

Définition 1.1 (espace électoral de référence)
Dans le modèle que nous appelons *espace électoral de référence*, chaque électeur v est capable d'établir mentalement :
— un ordre strict faible de préférence $p_v \in \mathcal{F}_\mathcal{C}$ sur les candidats,
— un vecteur $u_v \in [0,1]^C$ de *notes* sur les candidats,
— et un vecteur $a_v \in \{0,1\}^C$ de *valeurs d'approbation* sur les candidats.
Le triplet $\omega_v = (p_v, u_v, a_v)$ sera appelé son *état* sincère et nous noterons P_v la fonction qui extrait le premier élément de ce triplet : $\mathrm{P}_v(\omega_v) = p_v$.

De manière générale, nous noterons Ω_v l'ensemble des états possibles pour l'électeur v. Dans le premier modèle ci-dessus, nous avons $\Omega_v = \mathcal{F}_\mathcal{C} \times [0,1]^C \times \{0,1\}^C$. Nous suggérons au lecteur de voir cet ensemble comme l'analogue d'un univers en théorie des probabilités : pour la plupart des problèmes, il n'est pas nécessaire de spécifier exactement son contenu. Le point essentiel est la possibilité de définir des fonctions qui extraient telle ou telle information à propos de l'électeur v, comme le ferait une variable aléatoire en théorie des probabilités : par exemple, la fonction P_v permet d'accéder à sa relation binaire de préférence.

Par analogie avec les notations habituelles pour les variables aléatoires, nous écrirons fréquemment P_v comme raccourci pour $\mathrm{P}_v(\omega_v)$: par conséquent, les expressions $c\ \mathrm{P}_v(\omega_v)\ d$ et $c\ \mathrm{P}_v\ d$ sont des synonymes qui signifient que l'électeur v (dans l'état ω_v) préfère le candidat c au candidat d. Pour une certaine relation binaire de préférence sur les candidats p_v, la notation $\mathrm{P}_v(\omega_v) = p_v$ ou, en abrégé, $\mathrm{P}_v = p_v$, signifie que la variable P_v a la valeur p_v (dans l'état ω_v).

Nous appelons *configuration* un V-uplet $\omega = (\omega_1, \ldots, \omega_V)$ donnant l'état de chaque électeur, comme l'illustre la figure 1.1. Nous notons $\Omega = \prod_{v \in \mathcal{V}} \Omega_v$ le produit cartésien qui contient toutes les configurations possibles et $\mathrm{P} = (\mathrm{P}_1, \ldots, \mathrm{P}_V)$ la fonction à V variables qui, à une configuration ω, associe le profil qui lui correspond : $(\mathrm{P}_1(\omega_1), \ldots, \mathrm{P}_V(\omega_V))$.

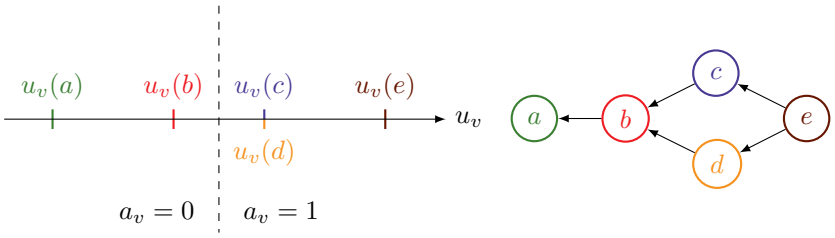


FIGURE 1.2 – État d'un électeur dans l'espace électoral de référence

Nous avons maintenant tous les éléments nécessaires pour définir un *espace électoral*, c'est-à-dire un modèle mathématique représentant les préférences des électeurs.

Définition 1.2 (espace électoral)
Un *espace électoral* est donné par :
— deux entiers positifs V et C,
— pour chaque électeur $v \in \mathcal{V}$, un ensemble non vide Ω_v de ses états possibles,
— pour chaque électeur $v \in \mathcal{V}$, une fonction $\mathrm{P}_v : \Omega_v \to \mathcal{R}_\mathcal{C}$, dont le résultat est sa relation binaire de préférence.

Un tel espace électoral est noté $(V, C, \Omega, \mathrm{P})$, ou simplement Ω quand il n'y a pas d'ambiguïté.

Revenons à l'exemple de l'espace électoral de référence, où $\omega_v = (p_v, u_v, a_v)$. Pour plus de réalisme, le planificateur social [3] a la possibilité d'adopter l'hypothèse que les notes et les valeurs d'approbation de chaque électeur sont cohérents avec sa relation binaire de préférence : si $c \, p_v \, d$, alors $u_v(c) > u_v(d)$ et $a_v(c) \geq a_v(d)$. La figure 1.2 illustre l'état d'un électeur vérifiant ces hypothèses. La partie gauche de la figure représente les positions des candidats sur son axe des utilités u_v, ainsi que la limite entre les candidats qu'il désapprouve ($a_v = 0$) et ceux qu'il approuve ($a_v = 1$). La partie droite représente le graphe de sa relation binaire de préférence. En l'occurrence, comme celle-ci est un ordre strict faible par hypothèse, la transitivité est sous-entendue dans la figure.

Il est facile d'inclure cette hypothèse dans le modèle, en définissant Ω_v comme l'ensemble des triplets (p_v, u_v, a_v) qui vérifient les conditions demandées. D'autres hypothèses peuvent être ajoutées de la même façon, en choisissant un ensemble Ω_v adapté pour chaque électeur.

1.1.4 Exemples d'espaces électoraux

Un autre modèle est très important pour la théorie : dans le choix social traditionnel arrovien, il est courant de ne représenter les opinions de chaque électeur que par un ordre strict total sur les candidats. Cette pratique correspond à l'espace électoral suivant.

3. Dans les sciences économiques et sociales, le *planificateur social* est un décideur qui essaie d'obtenir le meilleur résultat possible pour tous les acteurs concernés. Dans le cas qui nous intéresse, il s'agit d'un personne abstraite qui peut représenter à la fois les chercheurs en choix social et toutes les personnes qui peuvent avoir une influence pour choisir le système de vote utilisé dans une organisation humaine donnée. Par hypothèse, le planificateur social essaie d'établir un système de vote ayant d'aussi bonnes propriétés que possibles. Pour ce faire, il est nécessaire au préalable de cerner le domaine d'application privilégié en définissant l'espace électoral utilisé.

Définition 1.3 (espace électoral des ordres stricts totaux)
Pour tout $v \in \mathcal{V}$, notons $\Omega_v = \mathcal{L}_\mathcal{C}$ et P_v la fonction identité[4]. Ce modèle est appelé *espace électoral des ordres strict totaux pour V et C*.

Ce modèle est parfois élargi en autorisant les ordres stricts faibles. On peut aussi considérer des relations binaires quelconques.

Définition 1.4 (espace électoral des ordres stricts faibles)
Pour tout $v \in \mathcal{V}$, notons $\Omega_v = \mathcal{F}_\mathcal{C}$ et P_v la fonction identité. Ce modèle est appelé *espace électoral des ordres strict faibles pour V et C*.

Définition 1.5 (espace électoral des relations binaires)
Pour tout $v \in \mathcal{V}$, notons $\Omega_v = \mathcal{R}_\mathcal{C}$ et P_v la fonction identité. Ce modèle est appelé *espace électoral des relations binaires pour V et C*.

Le choix de l'espace électoral ne garantit rien sur l'expressivité d'un éventuel système de vote qu'on souhaiterait étudier par la suite. Par exemple, si on choisit l'espace électoral des ordres stricts faibles, cela ne signifie pas que le système de vote étudié permettra à un électeur de transmettre un ordre strict faible dans son bulletin. Cela signifie juste que le planificateur social admet qu'un électeur puisse avoir comme préférences un ordre faible sur les candidats et considère impossible qu'il ait des préférences d'une autre forme (approche *descriptive*). Ou, plus raisonnablement, cela signifie que les bonnes propriétés qu'il saurait prouver pour le système de vote étudié sont garanties si les électeurs ont des ordres stricts faibles de préférence, mais *a priori* pas dans les autres cas (approche *normative*).

Continuons à explorer les possibilités offertes par les espaces électoraux avec quelques exemples.

Exemple 1.6 (utilité avec marge d'incertitude)
Supposons que chaque électeur v soit capable d'établir un vecteur de notes $u_v \in [0,1]^C$ et un réel $\varepsilon \geq 0$, interprété comme une incertitude, tel qu'il préfère un candidat c à un candidat d ssi $u_v(c) > u_v(d) + \varepsilon$. Son espace d'états est alors $\Omega_v = [0,1]^C \times \mathbb{R}$ et sa fonction P_v est définie par l'inégalité précédente[5].

Exemple 1.7 (intervalles d'utilité)
Il est facile de généraliser l'exemple précédent. À chaque candidat c, l'électeur v associe un intervalle non vide $[\underline{u_v}(c), \overline{u_v}(c)]$. On l'interprète de la façon suivante : v situe son utilité pour le candidat c dans cet intervalle, mais pas plus précisément (par incapacité d'être plus précis, par manque d'intérêt ou parce que cela représente un coût cognitif trop élevé). Il préfère le candidat c à d ssi $\underline{u_v}(c) > \overline{u_v}(d)$. La figure 1.3 représente un exemple d'état pour un tel électeur. Dans la partie gauche, on a représenté son intervalle d'utilité pour chaque candidat ; dans la partie droite, le graphe de sa relation binaire de préférence. Celle-ci est irréflexive, antisymétrique et transitive ; c'est pourquoi on n'a pas représenté explicitement les arêtes de d ou e vers a, qui se déduisent des autres arêtes par transitivité. En revanche, la relation n'est pas négativement transitive : en effet, l'électeur ne préfère ni e à d ni d à c, et pourtant il préfère e à c. Ce n'est donc pas un ordre strict faible.

Exemple 1.8 (préférences multicritères)
À présent, chaque électeur note mentalement chaque candidat selon trois critères dans l'intervalle $[0,1]$: son espace d'états Ω_v est l'ensemble des matrices de

4. C'est un abus de langage que nous commettrons sans scrupule : pour être tout à fait rigoureux, P_v est l'injection canonique de $\mathcal{L}_\mathcal{C}$ dans $\mathcal{R}_\mathcal{C}$.

5. On pourra rapprocher cette idée des travaux de Ehlers *et al.* (2004), dans lesquels on considère une manipulabilité avec un seuil, c'est-à-dire où les électeurs souhaitent au moins améliorer leur utilité par une certaine amplitude.

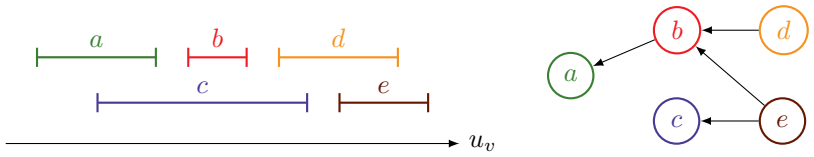


FIGURE 1.3 – Un état d'électeur dans l'espace électoral des intervalles d'utilité

		Candidat		
		Poulet (a)	Bœuf (b)	Poisson (c)
Critère	Bon marché	1	0,5	0
	Sain	0,5	0	1
	Moelleux	0	1	0,5

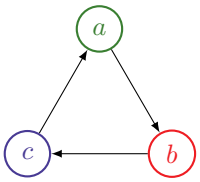


FIGURE 1.4 – État d'un électeur avec des préférences multicritères

taille $3 \times C$ à valeurs dans $[0, 1]$. On suppose qu'il préfère un certain candidat à un autre ssi le premier est strictement mieux noté que le second selon au moins deux critères : ceci définit la fonction P_v. Considérons un électeur dans l'état ω_v représenté par la figure 1.4. Il préfère a à b, b à c et c à a : sa relation binaire de préférence $\mathrm{P}_v(\omega_v)$ n'est donc pas transitive. D'une certaine façon, un tel électeur réalise un paradoxe de Condorcet à lui tout seul : la multiplicité des critères qu'il considère, et le fait qu'il décide ses préférences selon la majorité des critères, imite le comportement de trois électeurs distincts utilisant le vote majoritaire simple pour chaque couple de candidats.

Si on demande à un tel électeur de désigner un préféré parmi les trois candidats, alors il viole nécessairement à lui tout seul la propriété d'indépendance des alternatives non pertinentes (IANP) que nous évoquions en introduction : par exemple, s'il désigne le bœuf (b) comme son candidat préféré parmi les trois, cela semble contradictoire avec le fait qu'il choisisse le poulet quand il s'agit de choisir entre poulet et bœuf.

Exemple 1.9 (générateur aléatoire décentralisé)
Chaque électeur possède une relation binaire de préférence p_v sur les candidats et dispose d'une pièce de monnaie. Après avoir tiré à pile ou face, son état (sincère) est (p_v, x_v), où $x_v \in \{\text{Pile}, \text{Face}\}$. On suppose qu'il est le seul à connaître le résultat de son tirage aléatoire : par conséquent, s'il doit communiquer son état à un mode de scrutin, il peut mentir sur le résultat du tirage aussi bien que sur son ordre de préférence.

1.1.5 Propriétés basiques d'un espace électoral

Bien que nous interprétions P_v comme une préférence *stricte* dans la plupart de nos exemples [6], l'antisymétrie n'est pas exigée par la définition 1.2 d'un espace électoral. Nous discuterons en section 1.3 (page 34) quelle interprétation on peut

6. Jusqu'à présent, le seul cas où l'antisymétrie n'était pas vérifiée était l'espace électoral des relations binaires (définition 1.5).

donner à la relation P_v quand elle n'est pas antisymétrique. Si le lecteur est perturbé par cette possibilité, il peut lire tout ce qui suit en ajoutant une hypothèse d'antisymétrie. Cependant, en toute généralité, nous noterons :

— $c\ \mathrm{I}_v\ d$ ssi non $c\ \mathrm{P}_v\ d$ et non $d\ \mathrm{P}_v\ c$ (indifférence),
— $c\ \mathrm{PP}_v\ d$ ssi $c\ \mathrm{P}_v\ d$ et non $d\ \mathrm{P}_v\ c$ (préférence antisymétrique),
— $c\ \mathrm{PM}_v\ d$ ssi $c\ \mathrm{P}_v\ d$ et $d\ \mathrm{P}_v\ c$ (préférence mutuelle).

Si la relation P_v est antisymétrique, ce qui est une hypothèse courante, il n'y a que trois possibilités mutuellement exclusives : $c\ \mathrm{P}_v\ d$ (équivalent à $c\ \mathrm{PP}_v\ d$ dans ce cas), $d\ \mathrm{P}_v\ c$ et $c\ \mathrm{I}_v\ d$.

Pour certains résultats, nous supposerons que les électeurs ont une certaine liberté d'opinion (sans préjuger des possibilités d'exprimer cette opinion dans leurs bulletins).

Définition 1.10 (richesse d'un espace électoral)
Nous disons que :

1. Ω *comprend tous les ordres stricts totaux* ssi tout électeur peut avoir tout ordre strict total comme relation binaire de préférence ; c'est-à-dire, $\forall (v, p_v) \in \mathcal{V} \times \mathcal{L}_{\mathcal{C}}, \exists \omega_v \in \Omega_v$ t.q. $\mathrm{P}_v(\omega_v) = p_v$;
2. Ω *autorise tout candidat comme préféré* ssi tout électeur peut préférer strictement n'importe quel candidat à tous les autres ; c'est-à-dire, $\forall (v, c) \in \mathcal{V} \times \mathcal{C}, \exists \omega_v \in \Omega_v$ t.q. $\forall d \in \mathcal{C} \setminus \{c\}, c\ \mathrm{PP}_v\ d$.

L'implication 1 $\Rightarrow$ 2 est triviale.

Par exemple, dans l'espace électoral de référence, où $\omega_v = (p_v, u_v, a_v)$, la relation binaire de préférence p_v d'un électeur v peut être n'importe quel ordre strict faible ; en particulier, tout ordre strict total est possible. Par conséquent, cet espace électoral vérifie les propriétés 1 et 2.

Nous insistons de nouveau sur le fait que ces propriétés n'ont rien à voir avec l'expressivité d'un éventuel système de vote qu'on souhaiterait étudier. Par exemple, si on suppose que l'espace électoral comprend tous les ordres stricts totaux, cela ne signifie pas que le système de vote permet à l'électeur d'exprimer son ordre strict total de préférence s'il en a un ; cela signifie juste qu'une telle opinion est considérée comme possible *a priori*.

Alors que limiter un espace électoral aux ordres stricts totaux est discutable, l'hypothèse plus faible qu'il comprenne les ordres stricts totaux semble en général assez naturelle. Cependant, on peut facilement concevoir des modèles raisonnables qui ne vérifient pas cette propriété. Par exemple, s'il y a un très grand nombre de candidats, on peut arguer qu'il est cognitivement impossible pour un électeur d'établir un ordre strict total de préférence sur les candidats. Nous adressons tout de même une critique à ce type de modèle : même si l'idée est intéressante d'un point de vue *descriptif*, il est légitime d'ajouter la possibilité d'établir un ordre total pour des raisons *normatives*. En effet, même si on estime qu'un électeur ne peut atteindre cet état de connaissance complète sur ses préférences, on n'a aucune raison (et aucun moyen pratique) de le lui interdire *a priori*.

Avant de voir un autre exemple d'espace électoral qui ne comprend pas tous les ordres stricts totaux, rappelons la notion classique de *single-peakedness* (Black, 1958). L'intuition est la suivante : supposons qu'il existe une manière « naturelle » de placer tous les candidats sur un axe abstrait, par exemple un axe gauche-droite en politique. Ceci donne un ordre de référence sur les candidats que nous noterons $\mathrm{P}_{\mathrm{ref}}$.

Imaginons que cet axe possède la propriété suivante : chaque électeur possède une valeur maximale d'utilité pour un certain candidat (son pic) et son utilité

est décroissante lorsqu'on s'éloigne de ce candidat, aussi bien vers la droite que vers la gauche. Alors son ordre de préférence P_v commence par le candidat qui représente son pic d'utilité, puis il place tous les candidats plus à gauche dans l'ordre d'éloignement du pic, et de même pour les candidats de droite. Il est tout à fait possible qu'il préfère tous les candidats à gauche du pic puis tous ceux à droite, ou l'inverse, ou que ces candidats soient entrelacés dans l'ordre P_v.

La propriété caractéristique d'un tel ordre de préférence P_v est qu'entre un certain candidat d et deux autres candidats c et e qui sont respectivement à sa gauche et à sa droite (au sens de $\mathrm{P}_{\mathrm{ref}}$), il est impossible d'apprécier d moins que c et d à la fois : en effet, ceci impliquerait que le pic de préférence de l'électeur est à la fois à gauche et à droite de d, alors que ce pic est unique par hypothèse. Cette caractérisation présente l'avantage de ne pas recourir à la notion d'utilité et c'est pourquoi elle est couramment utilisée comme définition de la *single-peakedness*.

Définition 1.11 (single-peakedness)
Pour P_v et $\mathrm{P}_{\mathrm{ref}}$ deux ordres stricts totaux sur les candidats, on dit que P_v est *single-peaked* par rapport à $\mathrm{P}_{\mathrm{ref}}$ ssi :

$$c\ \mathrm{P}_{\mathrm{ref}}\ d \text{ et } d\ \mathrm{P}_{\mathrm{ref}}\ e \Rightarrow \mathrm{non}(c\ \mathrm{P}_v\ d \text{ et } e\ \mathrm{P}_v\ d).$$

Pour P un profil et $\mathrm{P}_{\mathrm{ref}}$ un ordre strict total, on dit que P est *single-peaked* par rapport à $\mathrm{P}_{\mathrm{ref}}$ ssi pour tout électeur v, P_v est un ordre strict total qui est *single-peaked* par rapport à $\mathrm{P}_{\mathrm{ref}}$.

On dit qu'un profil P est *single-peaked* ssi il existe un ordre strict total $\mathrm{P}_{\mathrm{ref}}$ tel que P est *single-peaked* par rapport à $\mathrm{P}_{\mathrm{ref}}$.

On dit qu'un espace électoral Ω est *single-peaked* ssi tout profil P de cet espace électoral est *single-peaked*.

Exemple 1.12 (température d'une pièce)
Les occupants d'une pièce s'apprêtent à choisir la température du thermostat parmi les options candidates $\{17°, 18°, 19°, 20°\}$. Le planificateur social fait l'hypothèse que chaque électeur peut *a priori* avoir n'importe quel ordre de préférence à condition qu'il soit *single-peaked* par rapport à l'ordre naturel sur les températures. Cet espace électoral autorise tout candidat comme préféré mais ne comprend pas tous les ordres stricts totaux : par exemple, l'ordre de préférence $20° \succ 17° \succ 19° \succ 18°$ est exclu par hypothèse. En effet, dans ce cadre, si la température idéale d'un électeur est 20°, alors il ne peut pas estimer que 17° est préférable à 19°.

1.1.6 Espace électoral probabilisé

À présent, nous allons munir Ω d'une mesure de probabilité, ou *culture*. Afin de manipuler les notions probabilistes de façon rigoureuse, nous devons considérer des tribus, des ensembles mesurables et des événements probabilistes. Cependant, la mesurabilité n'est pas un problème crucial en pratique : par exemple, sans l'axiome du choix, tout sous-ensemble de $\mathbb{R}^C$ est Lebesgue-mesurable. C'est pourquoi ces questions techniques ne seront discutées que dans l'annexe A.

Définition 1.13 (espace électoral probabilisé)
Un *espace électoral probabilisé*, ou *EEP*, est défini par la donnée d'un espace électoral $(V, C, \Omega, \mathrm{P})$ et d'une mesure de probabilité π sur Ω, appelée *culture*.

Un tel EEP est noté $(V, C, \Omega, \mathrm{P}, \pi)$, ou simplement (Ω, π).

Nous notons μ la loi de la variable aléatoire P selon la culture π.

Par exemple, considérons l'espace électoral de référence. Indépendamment pour chaque électeur v :

— tirons un vecteur de notes u_v uniformément dans $[0,1]^C$;
— définissons p_v comme l'ordre strict faible naturellement induit par u_v, au sens où $c\, p_v\, d \Leftrightarrow u_v(c) > u_v(d)$;
— pour chaque candidat c, définissons la valeur d'approbation $a_v(c)$ comme l'arrondi de $u_v(c)$ à l'entier le plus proche, 0 or 1.

Alors nous avons défini un exemple de culture π, c'est-à-dire une mesure de probabilité sur l'espace électoral Ω. Implicitement, ceci définit une loi μ pour le profil P.

Un exemple classique de culture est la *culture impartiale*, dont nous rappelons la définition.

Définition 1.14 (culture impartiale)
Sur l'espace électoral des ordres stricts totaux, on appelle *culture impartiale* la loi selon laquelle l'ordre de préférence de chaque électeur est tiré indépendamment et uniformément dans $\mathcal{L}_C$.

Terminons avec un exemple plus complexe. Supposons que nous avons 3 électeurs et 3 candidats. Nous souhaiterions étudier un espace électoral où chaque électeur est caractérisé par un ordre strict total de préférence et où tous les profils *single-peaked*, et seulement ceux-là, sont possibles, quel que soit l'ordre $\mathrm{P}_{\mathrm{ref}}$ de référence. Nous allons voir que c'est impossible, pourquoi il est souhaitable que ce soit impossible, et comment on peut transformer le problème.

Pour tout ordre strict total de préférence p_0, le profil où tous les électeurs ont cet ordre de préférence est manifestement *single-peaked* (pour p_0 par exemple), donc autorisé. Par conséquent, tout ordre p_0 appartient à l'ensemble des états possibles Ω_v de tout électeur v. Donc le profil suivant est autorisé.

a	b	c
b	c	a
c	a	b

Mais ce profil n'est pas *single-peaked*[7] : en effet, le candidat second dans l'ordre $\mathrm{P}_{\mathrm{ref}}$ ne peut pas être c (resp. a, b) à cause de l'électeur 1 (resp. 2, 3).

Ceci est dû au fait que nous avons défini Ω comme un produit cartésien : si une opinion particulière ω_v est possible *a priori* pour l'électeur v, nous considérons qu'elle reste possible, autorisée, quelles que soient les opinions des autres électeurs. *It's not a bug, it's a feature* : il s'agit d'une propriété souhaitable du modèle.

En revanche, la probabilité d'une certaine opinion ω_v peut varier en fonction de l'état des autres électeurs et même devenir nulle. Pour étudier toutes les configurations *single-peaked*, on peut donc utiliser un EEP afin de réduire à zéro la probabilité des autres configurations : par exemple, on peut considérer l'espace électoral des ordres stricts totaux muni d'une culture π dont le support est égal à l'ensemble des configurations *single-peaked*.

1.2 Systèmes de vote basés sur l'état (SVBE)

1.2.1 Définition

Dans cette section, nous modélisons les modes de scrutin en définissant les *systèmes de vote basés sur l'état* (SVBE), qui constitueront ensuite l'objet principal

7. Le lecteur habitué à ces notions aura également remarqué que le profil ne peut pas être *single-peaked* car il n'y a pas de vainqueur de Condorcet ; nous y reviendrons.

de notre étude. À première vue, ce modèle ne semble pas permettre de représenter tous les modes de scrutin imaginables. Mais nous verrons en section 1.4 que, dans l'optique de limiter la manipulabilité, on peut restreindre l'étude aux SVBE[8], ce qui permet d'alléger le formalisme.

Définition 1.15 (système de vote basé sur l'état)
Un *système de vote basé sur l'état* sur l'espace électoral Ω, ou *SVBE*, est une fonction $f : \Omega \to \mathcal{C}$.

Dans la suite, sauf mention explicite du contraire, f désignera toujours un SVBE sur un espace électoral Ω.

Par exemple, considérons une des variantes possibles du système de vote appelé *vote par notation*, dans l'espace électoral de référence, où $\omega_v = (p_v, u_v, a_v)$.

— Chaque électeur v communique un état appartenant à Ω_v ;
— On dit qu'il vote *sincèrement* ssi il communique son véritable état ω_v ;
— La fonction f ne prend en compte que les vecteurs de notes u_v communiqués par les électeurs et renvoie le candidat avec la plus grande note totale.

Pour terminer la description de ce SVBE, il est nécessaire de donner une *règle de départage*, c'est-à-dire une procédure qui résout les cas d'égalité. Dire que nous avons défini un SVBE est donc un léger abus de langage : il s'agit plutôt d'une classe de SVBE, dont chaque membre donne le même résultat quand il n'y a pas d'égalité.

Dans la littérature, on autorise parfois qu'un mode de scrutin renvoie, comme résultat, un sous-ensemble des candidats, et on précise alors qu'il est *résolu* si son résultat est toujours un candidat unique. Dans cette thèse, l'hypothèse que le mode de scrutin soit résolu fait partie intégrante de la définition d'un SVBE. Nous rejoignons, en cela, le point de vue de Gibbard (1973) : si l'objectif d'une élection est de choisir une unique option parmi plusieurs, alors le système n'est totalement défini que si on inclut dans sa description la règle de départage éventuelle. Celle-ci peut être importante pour les propriétés étudiées : en particulier, elle peut avoir une influence sur les manipulations possibles[9].

Par ailleurs, pour implanter en pratique un système de vote comme le vote par notation présenté ci-dessus, il est suffisant que les bulletins ne contiennent que l'information effectivement utilisée par la fonction f, c'est-à-dire les notes dans cet exemple. Nous reviendrons sur ce point dans la section 1.4. Mais ce formalisme basé sur l'état facilite les analyses générales, indépendamment de l'implantation pratique du mode de scrutin. D'une part, il évite le besoin d'une fastidieuse *fonction de sincérité* qui associe à un état d'opinion le bulletin sincère correspondant (comme nous l'avons évoqué en introduction et comme nous le verrons de façon plus précise dans la section 1.4). D'autre part, il permet de définir facilement des transformations de systèmes de vote, comme la condorcification que nous verrons au chapitre 2 et le tranchage au chapitre 5.

1.2.2 Critères basiques pour un SVBE

Tout d'abord, on dit qu'un système est *unanime* s'il vérifie la propriété suivante : pour tout candidat c, si tous les électeurs préfèrent strictement c aux autres candidats et votent sincèrement, alors c est élu.

8. Cette notion de *système de vote basé sur l'état* est une généralisation de ce qui est appelé *procédure de vote élémentaire* par Moulin (1978, chapitre II, définition 2). L'auteur remarque déjà que considérer une telle procédure ne peut que diminuer les possibilités stratégiques des électeurs.

9. Nous reviendrons sur la question complexe de définir la manipulabilité pour des modes de scrutin non résolus ou non déterministes dans la section 6.3.3.

Nous allons aussi définir l'*anonymat* et de la *neutralité*. Comme le modèle des espaces électoraux est plus abstrait et plus général que la simple utilisation d'ordres stricts totaux de préférence, ces définitions y exigent un peu plus de précautions que la définition informelle que nous avons donnée dans l'introduction de ce mémoire.

On dit que l'espace électoral Ω lui-même est *anonyme* ssi tous les électeurs ont le même espace d'états : $\forall (v, v') \in \mathcal{V}^2, \Omega_v = \Omega_{v'}$ et $\mathrm{P}_v = \mathrm{P}_{v'}$. On notera qu'il s'agit ici d'une égalité de P_v et $\mathrm{P}_{v'}$ en tant que fonctions, et non de l'égalité entre deux relations de préférence.

On dit qu'un SVBE f sur l'espace électoral Ω est *anonyme* ssi :

1. Ω est anonyme ;
2. Pour tout état $\omega \in \Omega$, pour toute permutation $\sigma \in \mathfrak{S}_{\mathcal{V}}$ des électeurs, en notant $\omega_\sigma = (\omega_{\sigma(1)}, \ldots, \omega_{\sigma(V)})$, on a $f(\omega) = f(\omega_\sigma)$.

Autrement dit, le vainqueur ne dépend que des bulletins reçus, et non de l'identité des électeurs qui les ont émis.

Pour définir la neutralité, notons Φ une *action* du groupe des permutations des candidats $\mathfrak{S}_{\mathcal{C}}$ sur l'espace d'état Ω_v de chaque électeur v. Ceci signifie qu'à tout état ω_v et à toute permutation σ des candidats, on associe un état $\Phi_\sigma(\omega_v)$, que nous notons simplement $\sigma(\omega_v)$; et qu'on exige que cette transformation soit compatible avec la structure de groupe de $\mathfrak{S}_C$: ceci signifie par exemple qu'en notant σ' une autre permutation et $\circ$ l'opérateur de composition, on exige que $\sigma(\sigma'(\omega_v)) = (\sigma \circ \sigma')(\omega_v)$.

On exige, en outre, que Φ soit compatible avec P_v, au sens où $\mathrm{P}_v(\sigma(\omega_v)) = \sigma(\mathrm{P}_v(\omega_v))$. Dans le membre de droite de cette égalité, la relation $\sigma(\mathrm{P}_v(\omega_v))$ est définie naturellement par : $c\ \mathrm{P}_v(\omega_v)\ d \Leftrightarrow \sigma(c)\ \sigma(\mathrm{P}_v(\omega_v))\ \sigma(d)$, c'est-à-dire qu'on considère la même relation de préférence en permutant simplement les noms des candidats.

Quand une telle action de groupe est définie sur l'espace d'état de chaque électeur, elle induit une action de groupe sur l'espace électoral Ω tout entier : pour toute configuration $\omega \in \Omega$, il suffit de poser $\sigma(\omega) = (\sigma(\omega_1), \ldots, \sigma(\omega_V))$.

On dit que f est *neutre* (par rapport à l'action de groupe Φ) ssi $\forall \sigma \in \mathfrak{S}_{\mathcal{C}}, \forall \omega \in \Omega, f(\sigma(\omega)) = \sigma(f(\omega))$. Autrement dit, le candidat vainqueur ne dépend pas des étiquettes qui servent à nommer les candidats : si on change les étiquettes de tous les candidats, il existe une façon de ré-étiqueter les bulletins (Φ) qui permet d'obtenir le même système de vote.

En général, l'action de groupe Φ qu'on considère est intuitivement évidente : par exemple, dans l'espace électoral de référence, on appliquera σ à l'ordre de préférence p_v, on permutera les coordonnées du vecteur de notes u_v et on fera de même pour le vecteur de valeurs d'approbation a_v.

Cependant, en toute généralité, il peut être nécessaire de préciser l'action de groupe Φ utilisée. Reprenons l'exemple 1.9 du générateur aléatoire décentralisé, où chaque électeur possède le résultat d'un lancer de pièce. On considère un cas où il y a $C = 2$ candidats ; par conséquent, pour définir l'action de groupe Φ, il suffit de donner son effet pour la seule permutation non triviale $\sigma = (1 \leftrightarrow 2)$, qui consiste à échanger les deux candidats. Son effet sur les ordres de préférence est automatique, puisqu'on a exigé que Ψ soit compatible avec les fonctions P_v. Il ne reste donc qu'à définir l'effet de σ sur Pile et Face. Par ailleurs, on suppose qu'il y a 3 électeurs (pour éviter les questions d'égalité).

Examinons, tout d'abord, le SVBE suivant : s'il y a davantage de Pile, c'est le candidat 1 qui gagne ; sinon, c'est le candidat 2 qui est vainqueur. Ce système

est d'ailleurs particulièrement manipulable : si un électeur préfère le candidat 1 et que sa pièce est tombée sur Face, il a toujours intérêt à mentir et à annoncer Pile ; mais c'est la neutralité de ce système qui nous intéresse pour le moment. Si on définit l'action de $\sigma = (1 \leftrightarrow 2)$ comme inversant Pile et Face (en plus de son effet sur les relations binaires de préférence), alors ce mode de scrutin est neutre par rapport à Φ.

Considérons, à présent, un autre SVBE. Chaque électeur communique son ordre de préférence et l'état de sa pièce de monnaie. S'il s'agit de Pile, on compte 1 point en faveur de son candidat préféré ; si c'est Face, on compte 3 points en faveur de son candidat préféré. Puis le candidat ayant le plus de points est élu. Ce système est similaire à un vote par notation où Pile signifie « préférer légèrement » et Face, « préférer fortement ». D'ailleurs, comme nous l'avons déjà remarqué, un électeur qui possède Pile dans son état a toujours intérêt à mentir et à annoncer Face. En ce qui concerne la neutralité, si on définit l'action de $\sigma = (1 \leftrightarrow 2)$ comme laissant Pile et Face inchangés (en plus de son effet sur les relations binaires de préférence), alors ce SVBE est neutre par rapport à Φ.

Ainsi, dans le cas général, il est possible d'avoir divers modes de scrutin qui sont neutres en considérant des actions de groupes différentes et non rendues évidentes *a priori* par la seule définition de l'espace électoral considéré. Ceci dit, dans la plupart des cas d'études pratiques, l'action de groupe Φ utilisée sera intuitivement évidente, comme nous l'avons vu dans l'espace électoral de référence.

Par la suite (notamment dans le chapitre 10), il nous arrivera parfois de nous concentrer sur l'espace électoral des ordres stricts totaux. Une question naturelle se pose alors : à quelle condition existe-t-il un mode de scrutin anonyme et neutre ?

Proposition 1.16

On se place dans l'espace électoral des ordres stricts totaux.

1. *Les conditions suivantes sont équivalentes.*
   (a) *Il existe un SVBE anonyme et neutre.*
   (b) *Il est impossible d'écrire C comme une somme de diviseurs de V supérieurs à 1. Autrement dit, il n'existe pas de séquence d'entiers naturels $(k_1, \dots, k_n)$ telle que :*

$$\left[\forall i \in [\![1, n]\!], k_i > 1 \text{ et } k_i \text{ divise } V\right] \text{ et } \sum_{i=1}^{n} k_i = C.$$

   (c) *Il est impossible d'écrire C comme une somme de facteurs premiers de V. Autrement dit, il n'existe pas de séquence d'entiers naturels $(k_1, \dots, k_n)$ telle que :*

$$\left[\forall i \in [\![1, n]\!], k_i \text{ est premier et } k_i \text{ divise } V\right] \text{ et } \sum_{i=1}^{n} k_i = C.$$

2. *Pour qu'il existe un SVBE anonyme et neutre, il est nécessaire mais pas suffisant que V et C soient premiers entre eux.*

Démonstration. Nous allons d'abord supposer 1 et en déduire 2, ce qui donnera une première intuition pour la preuve de 1 que nous donnerons ensuite.

2. Montrons d'abord que cette condition simplifiée n'est pas suffisante : considérons $V = 6$ et $C = 7$, qui sont bien premiers entre eux. Il est possible d'écrire C comme une somme de facteurs de V supérieurs à 1 : $C = 3 + 2 + 2$, où 3 divise V

et 2 divise V. Donc la condition 1b n'est pas vérifiée, ce qui implique qu'il n'existe pas de SVBE anonyme et neutre.

Pour s'en convaincre intuitivement, examinons le profil suivant, qui donnera une piste pour la preuve de l'équivalence énoncée dans le point 1.

1	2	3	1	2	3
2	3	1	2	3	1
3	1	2	3	1	2
4	5	4	5	4	5
5	4	5	4	5	4
6	7	6	7	6	7
7	6	7	6	7	6

Ce profil est obtenu en partant du classement du premier électeur et en appliquant des permutations circulaires aux trois premiers candidats, aux deux suivants et aux deux derniers. On remarquera notamment que les permutations des deux dernières paires sont toujours simultanées (si 4 est devant 5, alors 6 est devant 7), ce qui permet d'avoir 3×2 électeurs seulement et non $3 \times 2 \times 2 = 12$ (ce qui serait le cas si on considérait les images du premier électeur par toutes les permutations circulaires mentionnées, de façon indépendante les unes des autres). On pourra d'ailleurs noter qu'on a, en l'occurrence, $V = \text{ppcm}(3, 2, 2)$.

Il est, alors, facile de voir que, quel que soit le candidat vainqueur dans ce profil, le SVBE ne peut être anonyme et neutre : par exemple, si 1 est élu, alors 2 et 3 devraient l'être également par anonymat et neutralité.

Prouvons, à présent, qu'il est nécessaire que V et C soient premiers entre eux. Si ce n'est pas le cas, soit k un facteur premier commun de C et V. Notons $n = \frac{C}{k}$ et $k_i = k$ pour $i \in [\![1, n]\!]$. Alors $\sum_{i=1}^{n} k_i = C$, où tous les k_i sont premiers et divisent V. Donc la condition 1c n'est pas vérifiée, ce qui implique qu'il n'existe pas de SVBE anonyme et neutre.

1a $\Rightarrow$ 1b. Si la condition 1b n'est pas vérifiée, généralisons l'exemple précédent. Considérons une séquence $(k_1, \ldots, k_n)$ telle que $\sum_{i=1}^{n} k_i = C$, tous les k_i sont supérieurs à 1 et divisent V. Cette dernière propriété est équivalente à : $\text{ppcm}_i(k_i)$ divise V. Considérons un ordre de préférence quelconque (par exemple $1 \succ 2 \succ \ldots \succ C$) puis toutes ses images en appliquant simultanément une permutation circulaire sur les k_1 premiers candidats, sur les k_2 candidats suivants, etc. L'orbite, c'est-à-dire l'ensemble des classements images obtenus, est de taille $\text{ppcm}_i(k_i)$. En copiant ces images suffisamment de fois, on obtient un profil à V électeurs. Si l'un des k_1 premiers candidats est vainqueur, alors par anonymat et neutralité, chacun des k_1 premiers candidats devrait être vainqueur également, ce qui contredit l'unicité du résultat car $k_1 > 1$. Il en est de même pour les k_2 candidats suivants, etc. Donc il n'existe pas de SVBE anonyme et neutre : la condition 1a n'est pas vérifiée.

1b $\Rightarrow$ 1a. Supposons la condition 1b vérifiée. Soit P un profil. On dit que deux candidats sont équivalents dans P s'il existe une permutation des candidats qui envoie l'un sur l'autre et qui laisse P stable, à permutation près des électeurs. Alors notons $k_1, \ldots, k_n$ les cardinaux des classes d'équivalence : on a $\sum_{i=1}^{n} k_i = C$. Considérons une certaine classe d'équivalence de taille k_i : comme tout candidat de la classe prend le meilleur rang (au sein de la classe) chez le même nombre d'électeurs, k_i divise V. Comme la condition 1b est vérifiée, il existe donc un k_1 égal à 1. On peut alors choisir le candidat correspondant comme vainqueur sans violer l'anonymat et la neutralité (ce qui fixe le vainqueur dans tout autre

profil obtenu à partir de P par permutation des candidats et/ou des électeurs). En effectuant le même raisonnement pour tout profil P dont le vainqueur n'est pas encore choisi, on obtient un SVBE anonyme et neutre.

1b ⇒ 1c. C'est immédiat car tout entier naturel premier est supérieur à 1.

1c ⇒ 1b. Si la condition 1b n'est par vérifiée, alors il est possible d'écrire C comme une somme $C = \sum_{i=1}^{n} k_i$, où tous les k_i divisent V et sont différents de 1. Pour chaque k_i, notons k'_i un facteur premier arbitraire de k_i, qui est donc également un facteur premier de V, et posons $n_i = \frac{k_i}{k'_i}$. On a alors $C = \sum_{i=1}^{n} \sum_{j=1}^{n_i} k'_i$, qui est une somme de facteurs premiers de V. Donc la condition 1c n'est pas vérifiée. □

Cette proposition est à mettre en parallèle avec un résultat classique (Moulin, 1978), qui concerne les modes de scrutin non seulement anonymes et neutres, mais également *efficients*. Ce qualificatif signifie qu'on a la propriété suivante : si tous les électeurs préfèrent strictement un certain candidat c à un certain d (c'est-à-dire si d est *Pareto-dominé* par c), alors d ne peut pas être élu. Le résultat précédemment connu est qu'il existe un SVBE neutre, anonyme et efficient (sur l'espace électoral des ordres stricts totaux) ssi $C!$ et V sont premiers entre eux, c'est-à-dire ssi tout entier inférieur ou égal à C est premier avec V.

Afin de vérifier la cohérence entre ce résultat et la proposition 1.16, on peut supposer que cette condition est vérifiée et montrer qu'elle implique bien la condition 1b. Si on écrit C comme une somme d'entiers naturels supérieurs à 1, alors chacun d'eux est inférieur ou égal à C, donc il est premier avec V ; comme il est supérieur à 1, ceci signifie qu'il ne divise pas V. Donc la condition 1b est vérifiée.

À notre connaissance, si on supprime l'hypothèse d'efficience, il n'existe pas de résultat précédent dans la littérature donnant une condition nécessaire et suffisante sur le couple (V, C) pour qu'il existe un SVBE anonyme et neutre (sur l'espace électoral des ordres stricts totaux), comme le fait la proposition 1.16.

1.3 Manipulabilité

À présent, nous allons traduire la définition habituelle de la manipulabilité dans le cadre des espaces électoraux.

Pour deux candidats w et c, on note :

$$\mathrm{Manip}_{\omega}(\mathrm{w} \to c) = \{v \in \mathcal{V} \text{ t.q. } c \, \mathrm{P}_v(\omega_v) \, \mathrm{w}\}.$$

C'est l'ensemble des électeurs qui préfèrent c à w ; si w est vainqueur du vote sincère, ces électeurs sont donc intéressés par une manipulation en faveur de c. À l'inverse, on note :

$$\mathrm{Sinc}_{\omega}(\mathrm{w} \to c) = \{v \in \mathcal{V} \text{ t.q. } \mathrm{non}(c \, \mathrm{P}_v(\omega_v) \, \mathrm{w})\}.$$

Il s'agit des électeurs qui ne sont pas intéressés par une manipulation visant à faire gagner c au lieu de w. Pour ces deux notions, on sous-entendra généralement la dépendance en ω et on notera simplement $\mathrm{Manip}(\mathrm{w} \to c)$ et $\mathrm{Sinc}(\mathrm{w} \to c)$.

Définition 1.17 (manipulabilité)
Pour $(\omega, \psi) \in \Omega^2$, un sous-ensemble d'électeurs $M \in \mathcal{P}(\mathcal{V})$ et un candidat $c \in \mathcal{C}$, on dit que f est *manipulable* dans la configuration ω par la coalition M

vers la configuration ψ en faveur du candidat c ssi :

$$\begin{cases} c \neq f(\omega), \\ f(\psi) = c, \\ M \subseteq \mathrm{Manip}(f(\omega) \to c), \\ \forall v \in \mathcal{V} \setminus M, \psi_v = \omega_v. \end{cases}$$

Dans la suite de cette thèse, nous dirons indifféremment *manipulable par coalition* (MC), *coalition-manipulable* ou, simplement, *manipulable* sans davantage de précision : en effet, c'est la notion la plus générale de manipulabilité, au sens où elle est impliquée par toutes les autres formes de manipulabilité que nous mentionnerons plus tard, notamment à partir du chapitre 6. Nous utiliserons le sigle MC indifféremment comme adjectif (*manipulable* par coalition) ou comme substantif (*manipulation* par coalition).

Quand on dit que f est manipulable en ω sans préciser ψ (resp. M, c), cela signifie qu'il existe ψ (resp. M, c) qui vérifie la définition précédente.

Ainsi, pour $(\omega, \psi) \in \Omega^2$ et un candidat $c \in \mathcal{C}$, on dit que f est *manipulable* dans la configuration ω vers la configuration ψ en faveur du candidat c ssi il existe une coalition M telle que les conditions précédentes soient vérifiées, ce qui se traduit par les relations suivantes.

$$\begin{cases} c \neq f(\omega), \\ f(\psi) = c, \\ \forall v \in \mathrm{Sinc}(f(\omega) \to c), \psi_v = \omega_v. \end{cases}$$

De même, pour $(\omega, \psi) \in \Omega^2$, on dit que f est *manipulable* dans la configuration ω vers la configuration ψ ssi il existe un candidat c telle que les relations précédentes soient vérifiées, ce qui s'écrit :

$$\begin{cases} f(\psi) \neq f(\omega), \\ \forall v \in \mathrm{Sinc}(f(\omega) \to f(\psi)), \psi_v = \omega_v. \end{cases}$$

Enfin, pour $\omega \in \Omega$, on dit que f est *manipulable* dans la configuration ω ssi il existe une configuration $\psi \in \Omega$ telle que f est manipulable en ω vers ψ. Par commodité de langage, il nous arrivera alternativement de dire que c'est la configuration ω qui est manipulable (dans le cadre du système de vote f).

On note MC_f l'ensemble des configurations ω où f est manipulable et on utilise la même notation pour la fonction indicatrice de cet ensemble :

$$\mathrm{MC}_f : \left| \begin{array}{ccc} \Omega & \to & \{0,1\} \\ \omega & \to & \left| \begin{array}{l} 1 \text{ si } f \text{ est manipulable en } \omega, \\ 0 \text{ sinon.} \end{array} \right. \end{array} \right.$$

Pour une culture π, nous appelons *taux de MC de f pour π* (à condition que MC_f soit mesurable, cf. appendice A) :

$$\begin{aligned} \tau^{\pi}_{\mathrm{MC}}(f) &= \pi(f \text{ est manipulable en } \omega) \\ &= \int_{\omega \in \Omega} \mathrm{MC}_f(\omega) \pi(\mathrm{d}\omega). \end{aligned}$$

Quand il n'y a pas d'ambiguïté sur la culture π utilisée, nous notons simplement ce taux $\tau_{\mathrm{MC}}(f)$.

Pour deux SVBE f et g, on dit que f est *au plus aussi manipulable que g au sens ensembliste*[10] ssi $\mathrm{MC}_f \subseteq \mathrm{MC}_g$; et on dit que f est *au plus aussi manipulable que g au sens probabiliste* ssi $\tau^{\pi}_{\mathrm{MC}}(f) \leq \tau^{\pi}_{\mathrm{MC}}(g)$. La première propriété est très forte puisqu'elle implique la seconde pour toute culture π : si f est moins manipulable que g au sens ensembliste, alors, dans toute culture, f est moins manipulable que g au sens probabiliste[11].

Cette définition de la manipulabilité est la première où nous exploitons réellement les relations binaires de préférence P_v. Profitons de cette occasion pour revenir sur le fait que nous ne les avons pas supposées antisymétriques.

Considérons deux candidats a et b. Supposons que pour chaque électeur v, son état ω_v est un couple (t_v, x_v) où :

— $t_v \in \{a, b, \varnothing\}$ est égal soit à son candidat favori (son « top », qui peut être a ou b), soit à $\varnothing$ s'il n'est partisan d'aucun des deux ;

— x_v est une variable booléenne qui représente si l'électeur est corruptible ou non, typiquement dans le cas où un candidat lui proposerait un pot-de-vin.

Nous nous posons la question suivante : étant donné le système de vote utilisé, le résultat du vote est-il résistant aux effets combinés de la corruption par des pots-de-vin et du vote stratégique au sens habituel ?

Si a (resp. b) est le vainqueur du vote sincère, les électeurs susceptibles de tenter de changer le résultat sont ceux qui sont corruptibles, quel que soit leur candidat favori, et ceux dont le candidat favori est vraiment b (resp. a). Remarquons alors que ce problème peut être traité avec notre formalisme sur la manipulation en utilisant des préférences qui ne sont pas antisymétriques. En effet, pour tout électeur v, définissons la relation binaire P_v comme dans le tableau suivant.

		t_v (candidat favori)		
		a	$\varnothing$	b
x_v (corruptible)	Faux	$a\ \mathrm{PP}_v\ b$	$a\ \mathrm{I}_v\ b$	$b\ \mathrm{PP}_v\ a$
	Vrai	$a\ \mathrm{PM}_v\ b$		

Dans l'espace électoral ainsi défini, la *manipulabilité* (au sens formel de la définition 1.17) coïncide avec la vulnérabilité aux effets combinés de la corruption et du vote stratégique au sens habituel.

Par ailleurs, l'absence d'hypothèse d'antisymétrie permet d'envisager un électeur *jamais content*, pour qui on a $c\ \mathrm{P}_v\ d$ pour tout couple de candidats (c, d). Un tel électeur est prêt à participer à n'importe quelle tentative de manipulation, introduisant ainsi une instabilité intéressante dans le système. On retiendra de cet exemple plaisant que, quand on veut étudier la manipulation, la question pertinente pour un électeur donné peut être formulée ainsi : s'il sait que le vainqueur par défaut est un certain w, est-ce qu'il est prêt à agir pour qu'un certain candidat c soit vainqueur à la place ? Tandis qu'exiger que la relation obtenue ait des propriétés intéressantes (antisymétrie, transitivité) peut avoir un intérêt pour attribuer une certaine cohérence aux actions d'un électeur, de telles hypothèses de rationalité ne sont pas nécessaires pour étudier la manipulation.

Nous avons deux autres raisons de ne pas supposer l'antisymétrie en général. D'une part, la plupart de nos résultats ne nécessitent pas cette hypothèse ; il serait donc dommage de prouver des théorèmes dans un cas particulier alors qu'ils sont vrais dans le cas général. D'autre part, supprimer cette hypothèse permet d'avoir

10. Cette notion est définie par Lepelley et Mbih (1994).

11. Si Ω est muni d'une tribu telle que tout singleton est mesurable (ce qui est généralement le cas), ce n'est pas seulement une implication mais une équivalence.

des définitions et des énoncés qui se prêtent plus facilement à d'autres types de généralisation, notamment celles du chapitre 4.

Pour les lecteurs qui pourraient être perturbés par un tel modèle, tout ce qui suit peut être lu en ajoutant l'hypothèse d'antisymétrie et l'interprétation habituelle de la manipulabilité.

Le théorème de Gibbard (1973), que nous avons déjà mentionné en introduction, implique que dans un espace électoral qui comprend tous les ordres stricts totaux, pour tout SVBE f non trivial (c'est-à-dire qui n'est pas dictatorial et dont l'image contient au moins 3 candidats), on a $\mathrm{MC}_f \neq \varnothing$. Le théorème de Satterthwaite (1975) énonce essentiellement le même résultat dans l'espace électoral des ordres stricts faibles[12]. En outre, nous verrons que le théorème de Gibbard concerne également les *systèmes de vote généraux*, que nous allons définir dans la section 1.4.

Si l'espace électoral ne comprend pas les ordres stricts totaux, il est possible d'avoir un SVBE non trivial et non manipulable. Par exemple, dans l'espace des températures de l'exemple 1.12, et plus généralement dans un espace *single-peaked* (définition 1.11), il est classique et facile à prouver que le mode de scrutin suivant n'est pas manipulable : chaque électeur indique sa température préférée, puis la température médiane est élue[13]. Ainsi, quand l'ordre de référence $\mathrm{P}_{\mathrm{ref}}$ est fixé *a priori* et connu au moment où on fixe la règle de vote, il existe au moins un système non manipulable.

Cependant, pour de nombreuses applications, une telle hypothèse est déraisonnable, à la fois pour des raisons descriptives et normatives. Considérons l'exemple d'une élection politique. D'un point de vue descriptif, il n'y a pas forcément de façon évidente et canonique de placer les candidats sur un axe (typiquement gauche-droite, au sens politique du terme) ; et il est loin d'être garanti que chaque électeur ait des préférences *single-peaked* par rapport à cet ordre de référence. D'un point de vue normatif, utiliser un mode de scrutin qui présuppose un ordre de référence fixé viole nécessairement la neutralité : en effet, on interdit à un électeur de préférer un candidat de gauche, puis un candidat de droite, puis un centriste, alors que l'ordre inverse est autorisé.

Pour de telles applications, même si on considère en première approximation que les préférences sont *single-peaked* avec probabilité 1, l'ordre $\mathrm{P}_{\mathrm{ref}}$ n'est pas fixé *a priori*, et il est impossible d'en tirer avantage dans le choix du système de vote, comme on pouvait le faire dans l'exemple des températures. Sous ces hypothèses, on constate que la conclusion du théorème de Gibbard-Satterthwaite reste vraie :

12. Pour des preuves alternatives du théorèmes de Gibbard-Satterthwaite, on pourra notamment consulter Gärdenfors (1977); Barberá (1983); Benoît (2000); Sen (2001). Plusieurs approches, notamment celle de Reny (2001), permettent de comprendre le lien profond entre ce résultat et le théorème d'Arrow. Muller et Satterthwaite (1977) montrent qu'il est aussi intimement lié à l'impossibilité qu'un mode de scrutin non trivial vérifie la *strong positive association*, une condition de monotonie par rapport aux préférences qui généralise la propriété de *réponse positive* que nous avons vue pour 2 candidats. Weber (2009) souligne, comme l'avaient fait Gibbard et Satterthwaite, qu'il n'est pas nécessaire de supposer que le mode de scrutin est surjectif (une formulation du théorème souvent rencontrée), mais simplement qu'au moins 3 candidats sont dans son image, c'est-à-dire sont éligibles. Duggan et Schwartz (2000) généralisent le résultat à un système de vote non résolu, c'est-à-dire sans règle de départage explicite ; Brandt et Brill (2011) discutent ce qu'il en est lorsqu'on considère des notions de non-manipulabilité moins exigeantes. Slinko et White (2008) montrent que le théorème reste vrai même en se limitant à des manipulations dites *sûres* (*safe*). Barberá et Peleg (1990) proposent une variante du théorème pour un nombre infini de candidats et une hypothèse de continuité sur les préférences des électeurs. On pourra consulter Taylor (2005) pour une vue d'ensemble de divers résultats d'impossibilité du même type.

13. Moulin (1980) montre une extension de ce résultat, qui est elle-même généralisée par Border et Jordan (1983) et Barberà (2001) dans un cadre multi-dimensionnel.

tout système non trivial est manipulable (Penn *et al.*, 2011). Contrairement à ce qu'on peut lire parfois, il est donc faux de dire qu'en culture single-peaked, tout système Condorcet est non-manipulable : ceci n'est vrai que si l'ordre de référence est fixé, à la fois dans les préférences sincères et dans les bulletins autorisés [14].

On peut donc se demander s'il existe une condition nécessaire et suffisante simple sur l'espace électoral pour que les conclusions du théorème de Gibbard-Satterthwaite s'appliquent, c'est-à-dire qu'il n'existe pas de système de vote non trivial et non manipulable. Aswal *et al.* (2003) répondent en partie à cette question en fournissant une condition suffisante (plutôt technique) mais, à notre connaissance, le problème reste ouvert. Dans ce mémoire, nous considérerons souvent des espaces électoraux qui comprennent tous les ordres stricts totaux, où le théorème de Gibbard-Satterthwaite s'applique ; ceci dit, la plupart de nos résultats théoriques concerneront un espace électoral quelconque.

1.4 Systèmes de vote généraux

Jusqu'à présent, nous avons défini des systèmes de vote basés sur l'état, ou SVBE. Cependant, les modes de scrutin réels sont parfois plus complexes : notamment, ils peuvent se dérouler en plusieurs tours et la forme des bulletins ne coïncide pas forcément avec l'objet mathématique utilisé pour modéliser les préférences des électeurs. Nous présentons ici un formalisme qui permet d'étudier n'importe quel type de mode de scrutin et dont les systèmes de vote basés sur l'état (SVBE) ne sont que des cas particuliers. Puis nous prouvons que, dans l'optique de limiter la manipulabilité, nous pouvons restreindre l'étude aux SVBE.

1.4.1 Définition

Pour commencer, considérons une des variantes possibles pour le mode de scrutin appelé *vote par notation*. Chaque électeur v peut choisir dans un ensemble de stratégies $\mathcal{S}_v = [0,1]^C$: il doit attribuer une note à chaque candidat. Une fois ces notes communiquées, on utilise une règle de dépouillement f qui renvoie le candidat avec la meilleure note moyenne (et utilise, en cas d'égalité, une règle déterministe choisie à l'avance). Généralisons ce cas particulier pour définir un *mode de scrutin*.

Définition 1.18 (mode de scrutin)
Un *mode de scrutin* (pour V et C) est défini par la donnée de :
— pour chaque électeur $v \in \mathcal{V}$, un ensemble non vide $\mathcal{S}_v$ dont les éléments sont appelés *stratégies* [15] ;
— et une fonction $f : \mathcal{S}_1 \times \ldots \times \mathcal{S}_V \to \mathcal{C}$ que nous appelons *règle de dépouillement*.

En anglais, un mode de scrutin tel que nous le définissons est appelé *game form* dans l'article fondateur de Gibbard (1973).

Quand un électeur v est dans l'état $\omega_v \in \Omega_v$, on peut se demander quelle stratégie $S_v \in \mathcal{S}_v$ doit être considérée comme sa stratégie sincère. Comme le

14. Avec un cadre légèrement différent, Blin et Satterthwaite (1976) discutent précisément les hypothèses impliquées et montrent notamment qu'il est crucial que non seulement les bulletins ne puissent exprimer que des préférences single-peaked par rapport à l'ordre de référence, mais aussi que les électeurs ne puissent avoir sincèrement que de telles préférences.

15. Ici, les stratégies correspondent à ce qu'on appelle « stratégies pures » en théorie des jeux. Il peut s'agir de bulletins, d'arbres de décisions dans un processus en plusieurs étapes ou, plus généralement de n'importe quel type d'objet.

remarque Gibbard (1973), il n'y a pas de façon générale, canonique, de définir le vote sincère en se basant seulement sur le mode de scrutin[16]. Voyons pourquoi avec deux exemples.

Exemple 1.19 (vote de parité)
Nous avons déjà évoqué ce système en introduction. On considère $C = 2$ candidats. Chaque électeur met une boule blanche ou noire dans l'urne. Si le nombre de boules noires est impaire, alors le candidat 1 est élu. Dans le cas contraire, c'est le candidat 2. Étant donné le candidat préféré par l'électeur v, on ne peut pas dire que celui-ci ait un bulletin sincère naturel !

On pourrait objecter que cet exemple repose sur une façon de voter contre-intuitive (bien qu'il s'agisse d'un mode de scrutin tout à fait conforme à la définition 1.18). Mais l'exemple suivant repose sur un mode de scrutin très classique.

Exemple 1.20 (uninominal sans ordre de préférence)
Dans l'espace électoral des ordres stricts faibles, considérons le mode de scrutin *uninominal* (à un tour) : chaque électeur vote pour un candidat, et le candidat qui reçoit le plus de voix est élu. Si un électeur ne préfère pas un candidat à tous les autres, il n'y a pas de façon évidente, canonique, de définir son vote sincère.

Ces exemples montrent que la notion de stratégie sincère est plus évanescente qu'elle n'y paraît de prime abord. Par conséquent, elle doit être définie de façon extrinsèque au mode de scrutin par des fonctions $s_v : \Omega_v \to \mathcal{S}_v$ qui, à chaque état d'opinion ω_v, associent un bulletin $S_v = s_v(\omega_v)$.

Définition 1.21 (système de vote)
Un *système de vote* F (sur l'espace électoral Ω) est défini par la donnée de :
— un mode de scrutin $((\mathcal{S}_v)_{v\in\mathcal{V}}, f)$;
— et pour chaque électeur $v \in \mathcal{V}$, une fonction $s_v : \omega_v \to \mathcal{S}_v$ qui est appelée *fonction de sincérité*.

Selon une première interprétation, la fonction s_v représente un consensus social sur l'esprit selon lequel on doit utiliser le mode de scrutin, ou un mode d'emploi du système communiqué par le planificateur social.

D'un point de vue légèrement différent, considérons un électeur qui choisit toujours son bulletin de façon déterministe, en se basant uniquement sur son état d'opinion, sans aucune information sur les opinions et les stratégies des autres électeurs. Cela peut provenir du fait qu'il n'a pas accès à ce type d'information avant l'élection, ou qu'il refuse de dépendre de ses sources d'information, comme les instituts de sondage. Alors, par définition, il utilise précisément une fonction $s_v : \Omega_v \to \mathcal{S}_v$ qui peut être vue, dans ce cas, comme une heuristique pour choisir son bulletin en fonction de sa propre opinion, mais sans information externe.

Dans l'exemple du vote par notation, il est possible de choisir, par exemple, la fonction de sincérité définie par $s_v(p_v, u_v, a_v) = u_v$ (dans l'espace électoral de référence). Si on suit la première interprétation, le planificateur social transmet le message aux électeurs que leur vecteur sincère u_v est considéré comme un bulletin approprié. Selon la seconde interprétation, s_v est une façon heuristique de choisir son bulletin en absence d'information sur les autres électeurs. Quelle que soit l'interprétation choisie, d'autres choix sont possibles : par exemple, nous pourrions décider que le bulletin sincère d'un électeur consiste à appliquer une transformation linéaire positive à son vecteur de notes de telle manière que son minimum soit 0 et que son maximum soit 1. Dans ce cas, nous étudierions un autre *système de vote* basé sur le même *mode de scrutin*.

16. Il en présente un exemple éclairant, la *fête sans alcool*.

La distinction formelle entre *mode de scrutin* et *système de vote* est donc importante quand on envisage d'utiliser un certain mode de scrutin en faisant varier la fonction de sincérité utilisée. Dans les autres cas, cette distinction n'est pas nécessaire et, par commodité de langage, nous emploierons les termes *mode de scrutin* et *système de vote* indifféremment.

Dans le monde réel, on se contente souvent de définir le mode de scrutin utilisé, sans définir explicitement une fonction de sincérité. Dans l'exemple de la France, on trouve sur les sites gouvernementaux les indications suivantes [17].

> Le principe du scrutin majoritaire est simple. Le ou les candidats qui obtiennent la majorité des suffrages exprimés sont élus. [...] Le scrutin peut être uninominal s'il y a un siège à pourvoir par circonscription. Les électeurs votent alors pour un seul candidat.

À notre connaissance, aucun texte officiel ne précise qu'il est recommandé de voter pour son candidat préféré : le choix de la fonction de sincérité éventuelle ou du comportement stratégique à adopter est laissé à l'entière appréciation des électeurs. Ainsi, les textes définissent un mode de scrutin, mais pas un système de vote.

Définissons à présent la manipulabilité d'un système de vote en général.

Pour $\omega \in \Omega$ et $S = (S_1, \dots, S_V) \in \mathcal{S}_1 \times \dots \times \mathcal{S}_V$, on dit que F est *manipulable en ω vers S* ssi :

$$\begin{cases} f(S_1, \dots, S_V) \neq f\big(s_1(\omega_1), \dots, s_V(\omega_V)\big), \\ \forall v \in \mathrm{Sinc}\Big(f\big(s_1(\omega_1), \dots, s_V(\omega_V)\big) \to f(S_1, \dots, S_V)\Big), S_v = s_v(\omega_v). \end{cases}$$

On dit que F est *manipulable en ω* ssi il existe $S \in \mathcal{S}_1 \times \dots \times \mathcal{S}_V$ tel que F est manipulable en ω vers S. L'ensemble des configurations manipulables de F est noté MC_F.

Si F n'est pas manipulable en ω, cela signifie qu'avec l'heuristique s_v (qui peut être fournie par le planificateur social ou choisie individuellement par chaque électeur), les électeurs sont capables de trouver un équilibre de Nash fort sans échange préliminaire d'information. Par conséquent, même si tous les bulletins sont révélés après l'élection, ils n'auront aucun regret sur le bulletin qu'ils ont choisi.

Dans son article fondateur, Gibbard (1973) présente un premier théorème fondamental qui peut être reformulé de la façon suivante. On dit qu'un mode de scrutin est *straightforward* ssi pour tout ordre strict total sur les candidats, chaque électeur possède une stratégie dominante quand il possède cet ordre de préférence. Le résultat est le suivant : il n'existe pas de mode de scrutin non trivial (c'est-à-dire qui n'est pas dictatorial et dont l'image contient au moins 3 candidats) qui soit *straightforward*. Autrement dit, il est impossible que tous les électeurs mettent au point une fonction de sincérité qui leur permettra de toujours défendre au mieux leurs opinions (même sans aucune information sur les bulletins des autres électeurs).

De ce premier théorème, Gibbard déduit un second résultat qu'on peut reformuler ainsi : dans un espace électoral qui comprend tous les ordres stricts totaux, pour tout système de vote F non trivial (c'est-à-dire qui n'est pas dictatorial et dont l'image contient au moins 3 candidats), on a $\mathrm{MC}_F \neq \varnothing$. Quand on restreint ce second théorème aux SVBE dans l'espace électoral des ordres stricts faibles, on obtient le théorème de Gibbard-Satterthwaite dans sa version la plus souvent citée.

17. http://www.interieur.gouv.fr/Elections/Les-elections-en-France/Les-modalites-d-elections/Les-differents-modes-de-scrutins.

1.4.2 Réduction d'un système de vote général à un SVBE

À présent, nous allons prouver que, dans l'optique de diminuer la manipulabilité, nous pouvons restreindre notre étude aux SVBE. La proposition qui suit est essentiellement un résultat de Moulin (1978, chapitre II). L'adapter au formalisme des espaces électoraux permet de le généraliser à des électeurs dont les préférences sont représentées par des objets mathématiques quelconques, et pas obligatoirement des relations d'ordre.

Commençons par un exemple. Dans l'espace électoral des ordres stricts totaux, examinons l'uninominal à deux tours (U2T) avec sa fonction de sincérité la plus intuitive : à chaque tour, chaque électeur vote pour son candidat préféré parmi ceux qui sont encore en lice. Si un électeur vote pour le candidat c au premier tour et pour le candidat d au second tour alors que c est encore en lice, alors il n'est manifestement pas sincère.

Pour empêcher ce type de comportement, on peut modifier le système de vote : les électeurs communiquent leur ordre de préférence ; puis, l'uninominal à deux tours est simulé avec les stratégies sincères correspondantes. Cela n'empêche pas les électeurs de mentir sur leur ordre de préférence, mais en tout cas, il ne peuvent plus voter pour c au premier tour et voter pour d contre c au second. Nous appelons le système obtenu l'*uninominal à deux tours instantanés (U2TI)*. C'est cette idée simple que nous allons généraliser.

Proposition 1.22 (version basée sur l'état)
Étant donné un système de vote F, considérons le système de vote $F' = ((\mathcal{S}'_v)_{v\in\mathcal{V}}, f', (s'_v)_{v\in\mathcal{V}})$ défini de la façon suivante.

1. *Chaque électeur communique un état : $\forall v \in \mathcal{V}, \mathcal{S}'_v = \Omega_v$.*
2. *La sincérité consiste à donner son véritable état : $\forall v \in \mathcal{V}, s'_v = \mathrm{Id}$.*
3. *Pour obtenir le résultat, on utilise l'ancienne règle de dépouillement f, en considérant que chaque électeur utilise la stratégie sincère correspondant à l'état qu'il a communiqué : $f'(\omega) = f(s_1(\omega_1), \ldots, s_V(\omega_V))$.*

Alors, le système de vote F', appelé la version basée sur l'état *de F, est au plus aussi manipulable que F : $\mathrm{MC}_{F'} \subseteq \mathrm{MC}_F$.*

Démonstration. Dans F', le vote sincère conduit au même résultat que dans F, mais les manipulateurs ont accès à au plus les mêmes bulletins. □

Un système de vote F est appelé un *système de vote basé sur l'état* (SVBE) ssi il est égal à sa version basée sur l'état, c'est-à-dire ssi, pour tout électeur $v \in \mathcal{V}$, on a $\mathcal{S}_v = \Omega_v$ et $s_v = \mathrm{Id}$. Par raccourci de notation, un SVBE peut donc être désigné uniquement par sa fonction de dépouillement f, comme nous l'avons déjà fait jusqu'à présent.

Considérer la version basée sur l'état d'un système de vote présente divers avantages.

— Cela peut empêcher les électeurs de recourir à des stratégies qui ne sont manifestement pas sincères, comme dans l'exemple de l'uninominal à deux tours.
— Cela simplifie le formalisme en identifiant les états d'opinion $\omega_v \in \Omega_v$ et les stratégies autorisées $S_v \in \mathcal{S}_v$, ce qui permet d'omettre les fonctions de sincérité s_v.
— Dans un SVBE, le bulletin inclut la relation binaire de préférence, même si ce n'était pas le cas dans le mode de scrutin original. Pour des systèmes de vote comme le vote par assentiment, cette étape est nécessaire avant de définir la *condorcification*, comme nous le ferons dans le chapitre suivant.

1	4	8	4
w	w	c	d
c	d	w	w
d	c	d	c

TABLE 1.1 – Exemple : U2T est manipulable mais pas U2TI.

Le bulletin utilisé dans ce cadre doit être vu comme une abstraction théorique, plus que comme un bulletin réel. Pour une implantation pratique du système de vote considéré, on pourra simplifier le bulletin utilisé, comme nous le verrons en section 1.4.3.

Ainsi, la proposition 1.22 nous apprend que considérer la version basée sur l'état ne peut pas augmenter la manipulabilité. Dans la table 1.1, on montre que ce procédé peut même la diminuer strictement. On considère les systèmes U2T et U2TI. En vote sincère, le second tour oppose w à c et c'est w qui est élu.

En U2T ou en U2TI, il est impossible de manipuler pour d : en effet, les manipulateurs concernés ne peuvent éviter que d soit éliminé au terme du premier tour. Examinons donc la manipulabilité en faveur de c. Les manipulateurs doivent éviter un second tour contre w, sinon celui-ci est élu. Ils ont donc besoin d'un second tour entre c et d. Pour cela, ils doivent donner au moins une voix à d. En U2TI, cela signifie qu'au second tour, d a au moins 8 voix venant des électeurs sincères (ceux qui ne préfèrent pas c à w), ainsi qu'au moins 1 voix des manipulateurs, donc d est élu et la manipulation échoue.

En revanche, en U2T, 6 manipulateurs votent pour c et 2 votent pour d au premier tour, ce qui fait 6 voix au total pour c et d, contre 5 pour w. Lors du second tour entre c et d, tous les manipulateurs peuvent voter pour c, y compris ceux qui avaient voté pour d au premier tour. Alors c est élu par 9 voix contre 8. Ainsi, U2T est manipulable en faveur de c, tandis qu'U2TI n'est pas manipulable dans la configuration étudiée.

1.4.3 Implantation canonique d'un SVBE

Dans la définition 1.15 d'un SVBE, nous avons vu qu'un électeur est censé communiquer toute l'information à propos de son état. Pour l'exemple du vote par notation dans l'espace électoral de référence, cela signifie qu'il doit communiquer son ordre faible et ses valeurs d'approbation, ce qui est manifestement inutile puisque ces informations n'ont aucune influence sur le résultat du vote. Nous allons simplement formaliser le fait que, pour implanter un tel système de vote en pratique, on ne demandera pas ces informations inutiles.

Pour la définition suivante, nous aurons besoin d'une notation pour décrire les états d'un sous-ensemble d'électeurs. Soit un sous-ensemble d'électeurs $M \in \mathcal{P}(\mathcal{V})$. On note $\Omega_M = \prod_{v \in M} \Omega_v$ l'ensemble des états possibles pour ces électeurs. La notation ω_M désignera un élément de cet ensemble. Si une configuration $\omega \in \Omega$ a été préalablement définie, alors ω_M désigne la restriction de la configuration ω aux électeurs de M.

Pour $M \in \mathcal{P}(\mathcal{V})$, $\omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}$ et $\psi_M \in \Omega_M$, la notation $(\omega_{\mathcal{V}\setminus M}, \psi_M)$ désigne la configuration obtenue en rassemblant les électeurs de $\omega_{\mathcal{V}\setminus M}$ et de ψ_M. Par raccourci de notation, on écrira $f(\omega_{\mathcal{V}\setminus M}, \psi_M)$ sans ajouter une seconde paire de parenthèses.

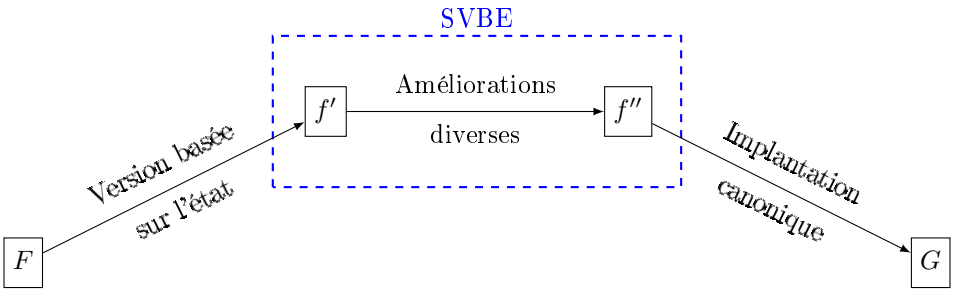


FIGURE 1.5 – Transformations de systèmes de vote

Proposition 1.23 (implantation canonique)
Soit f un SVBE et v un électeur. On dit que deux états $(\omega_v, \psi_v) \in (\Omega_v)^2$ sont indistinguables *pour f, et on note $\omega_v \equiv_f \psi_v$, ssi pour tout état des autres électeurs $\phi_{\mathcal{V}\setminus\{v\}} \in \Omega_{\mathcal{V}\setminus\{v\}}$, on a $f(\phi_{\mathcal{V}\setminus\{v\}}, \omega_v) = f(\phi_{\mathcal{V}\setminus\{v\}}, \psi_v)$. En toute rigueur, la relation $\equiv_f$ devrait aussi être indexée par l'électeur v, mais en pratique il n'y aura pas d'ambiguïté.*

On appelle implantation canonique *de f le système de vote $G = ((\mathcal{S}_v)_{v\in\mathcal{V}}, g, (s_v)_{v\in\mathcal{V}})$ défini de la façon suivante.*

1. *Pour tout électeur $v \in \mathcal{V}$, son ensemble de stratégies $\mathcal{S}_v$ est le quotient de Ω_v par la relation d'équivalence $\equiv_f$.*
2. *Pour tout électeur $v \in \mathcal{V}$, sa fonction de sincérité s_v est la projection canonique de Ω_v vers le quotient $\mathcal{S}_v = (\Omega_v / \equiv_f)$.*
3. *La fonction de dépouillement $g : \mathcal{S}_1 \times \ldots \times \mathcal{S}_V \to \mathcal{C}$ est le passage de f au quotient.*

Un SVBE f et son implantation canonique G ont la même manipulabilité : $\mathrm{MC}_G = \mathrm{MC}_f$.

Démonstration. Il suffit de remarquer que la définition de la manipulabilité est inchangée par passage au quotient. □

Examinons un autre exemple. Dans l'espace électoral des ordres stricts totaux (avec $V \geq 3$), considérons l'uninominal, muni d'une règle de départage qui n'utilise pas d'autre information d'état que le candidat préféré des électeurs. Voici deux exemples de telles règles.

1. Parmi les candidats ex aequo, le vainqueur est celui qui est premier dans l'ordre des indices.
2. Si le premier électeur dans l'ordre des indices vote pour un des candidats ex-aequo, celui-ci est élu. Sinon, on examine le bulletin du deuxième électeur, puis du troisième, etc. Autrement dit, l'électeur d'indice minimal parmi ceux qui votent pour un des candidats ex-aequo départage le vote en faveur de son favori.

Dans le SVBE représentant ce système de vote, chaque électeur communique son ordre de préférence. Dans l'implantation canonique, il communique uniquement son candidat préféré.

Quand nous étudierons un système de vote F, nous utiliserons presque toujours (et en général implicitement) le mécanisme représenté dans la figure 1.5 : tout d'abord, nous considérons sa version basée sur l'état, notée ici f', ce qui ne peut

augmenter la manipulabilité. Puis nous lui appliquons diverses transformations pour diminuer la manipulabilité et nous obtenons un nouveau SVBE f''. Si on souhaite mettre en pratique ce système de vote, on utilisera son implantation canonique G, ce qui ne change pas la manipulabilité.

1.5 Matrice des duels et notions Condorcet

À présent, nous rappelons les notions classiques de *matrice des duels*, de *vainqueur de Condorcet* et de *vainqueur de Condorcet faible*. Comme nous autorisons des relations binaires de préférence qui peuvent être quelconques, nous accordons une attention particulière à différentes variantes de ces notions, même si certaines d'entre elles peuvent coïncider dans certains espaces électoraux, par exemple quand on n'autorise que des ordres stricts totaux. En particulier, nous introduisons la notion de *candidat Condorcet-admissible*, qui sera une notion centrale pour étudier la manipulabilité d'une large classe de systèmes de vote.

Quand $\mathcal{A}(v)$ est une assertion qui dépend de l'électeur v, nous notons $|\mathcal{A}(v)| = \text{card}\{v \in \mathcal{V} \text{ t.q. } \mathcal{A}(v)\}$. Par exemple, pour un couple de candidats (c, d), la notation $|c\ \mathrm{P}_v\ d|$ désigne le nombre d'électeurs qui préfèrent c à d.

Définition 1.24 (matrice des duels)
Pour toute paire de candidats (c, d), on note $D_{cd}(\omega) = |c\ \mathrm{P}_v\ d|$, c'est-à-dire le nombre d'électeurs qui préfèrent c à d. La matrice $D(\omega)$ est appelée *matrice des duels* de ω.

La relation de *victoire relative* $\mathrm{P}_{\mathrm{rel}}(\omega)$ est définie par :

$$c\ \mathrm{P}_{\mathrm{rel}}\ d \Leftrightarrow D_{cd} > D_{dc}.$$

Il est facile de vérifier que cette relation est irréflexive, antisymétrique mais pas nécessairement complète. S'il y a un nombre impair d'électeurs et si leurs relations de préférence P_v sont des ordres strict totaux, alors on ne peut pas avoir $D_{cd} = D_{dc}$ donc l'un des candidats est nécessairement vainqueur : par conséquent, $\mathrm{P}_{\mathrm{rel}}$ est complète. De la même façon que pour les préférences individuelles, on définit $\mathrm{I}_{\mathrm{rel}}$ (absence de victoire relative). Les relations $\mathrm{PP}_{\mathrm{rel}}$ et $\mathrm{PM}_{\mathrm{rel}}$ n'ont guère d'intérêt : comme $\mathrm{P}_{\mathrm{rel}}$ est antisymétrique, $\mathrm{PP}_{\mathrm{rel}}$ est identique à $\mathrm{P}_{\mathrm{rel}}$; et $\mathrm{PM}_{\mathrm{rel}}$ est la relation vide.

La relation de *victoire absolue* $\mathrm{P}_{\mathrm{abs}}(\omega)$ est définie par :

$$c\ \mathrm{P}_{\mathrm{abs}}\ d \Leftrightarrow D_{cd} > \frac{V}{2}.$$

Quand cette relation est vérifiée, nous dirons indifféremment que c a une *victoire absolue* contre d dans ω, ou que d a une *défaite absolue* contre c dans ω. Quand nous parlerons de victoire (resp. défaite) sans davantage de précision, il s'agira toujours de victoire (resp. défaite) absolue.

À partir de $\mathrm{P}_{\mathrm{abs}}(\omega)$, on définit les relations $\mathrm{I}_{\mathrm{abs}}$ (absence de victoire), $\mathrm{PP}_{\mathrm{abs}}$ (victoire stricte) et $\mathrm{PM}_{\mathrm{abs}}$ (victoires mutuelles) comme nous l'avons fait pour les préférences individuelles.

Attention, quand nous disons que c a une victoire absolue contre d, ce qui est noté $c\ \mathrm{P}_{\mathrm{abs}}\ d$, l'inégalité :

$$D_{cd} > \frac{V}{2}$$

est déjà stricte par définition. Quand nous disons que c a une victoire absolue **stricte** contre d, ce qui est noté $c\ \mathrm{PP}_{\mathrm{abs}}\ d$, cela ne concerne donc pas le signe de

cette inégalité mais ajoute la condition que d n'a pas de victoire absolue contre c, ce qui signifie que :

$$D_{dc} \leq \frac{V}{2}.$$

Heureusement, ces subtilités se dissipent dès que les électeurs ont des préférences antisymétriques, comme nous allons le voir. Nous verrons cependant, dans le cadre de la condorcification généralisée (chapitre 4), que formaliser les choses ainsi facilitera leur généralisation.

Nous utiliserons plusieurs fois le résultat suivant (qui est trivial).

Proposition 1.25 (nombre de points total d'un duel)

$$D_{cd} + D_{dc} = V + |c \,\mathrm{PM}_v\, d| - |c \,\mathrm{I}_v\, d|.$$

Il en découle immédiatement que, si les relations P_v sont antisymétriques (ce qui est une hypothèse courante), alors la relation $\mathrm{P}_{\mathrm{abs}}$ est antisymétrique : deux candidats c et d ne peuvent pas avoir de victoires mutuelles. En d'autres mots, avec l'hypothèse classique d'antisymétrie, toute victoire (absolue) est stricte.

Par ailleurs, il découle également de cette proposition que, si les relations P_v sont complètes et si le nombre d'électeurs V est impair, alors la relation $\mathrm{P}_{\mathrm{abs}}$ est complète : entre deux candidats distincts c et d, il ne peut y avoir une absence de victoire.

À présent, nous allons définir diverses variantes de la notion de vainqueur de Condorcet.

Définition 1.26 (vainqueur de Condorcet absolu)

Soit $\omega \in \Omega$ et $c \in \mathcal{C}$. On dit que c est *vainqueur de Condorcet absolu*, ou simplement *vainqueur de Condorcet*, dans ω, ssi c a une victoire absolue stricte contre tout autre candidat. C'est-à-dire que pour tout autre candidat d, on a $c \,\mathrm{PP}_{\mathrm{abs}}\, d$, ce qui s'écrit également :

$$\begin{cases} |c \,\mathrm{P}_v\, d| > \dfrac{V}{2}, & (1.1) \\ |d \,\mathrm{P}_v\, c| \leq \dfrac{V}{2}. & (1.2) \end{cases}$$

Si les relations P_v sont antisymétriques (ce qui est une hypothèse courante), alors nous avons remarqué que toute victoire absolue est stricte. Dans ce cas, c est vainqueur de Condorcet ssi pour tout autre candidat d :

$$|c \,\mathrm{P}_v\, d| > \frac{V}{2}.$$

Définition 1.27 (vainqueur de Condorcet relatif)

Soit $\omega \in \Omega$ et $c \in \mathcal{C}$. On dit que c est *vainqueur de Condorcet relatif* dans ω ssi c a une victoire relative contre tout autre candidat. C'est-à-dire que pour tout autre candidat d, on a $c \,\mathrm{P}_{\mathrm{rel}}\, d$, ce qui s'écrit également :

$$|c \,\mathrm{P}_v\, d| > |d \,\mathrm{P}_v\, c|.$$

Définition 1.28 (vainqueur de Condorcet faible)

Soit $\omega \in \Omega$ et $c \in \mathcal{C}$. On dit que c est *vainqueur de Condorcet faible* dans ω ssi c n'a aucune défaite relative. C'est-à-dire que, pour tout autre candidat d, on a $\mathrm{non}(d \,\mathrm{P}_{\mathrm{rel}}\, c)$, ce qui s'écrit également :

$$|c \,\mathrm{P}_v\, d| \geq |d \,\mathrm{P}_v\, c|.$$

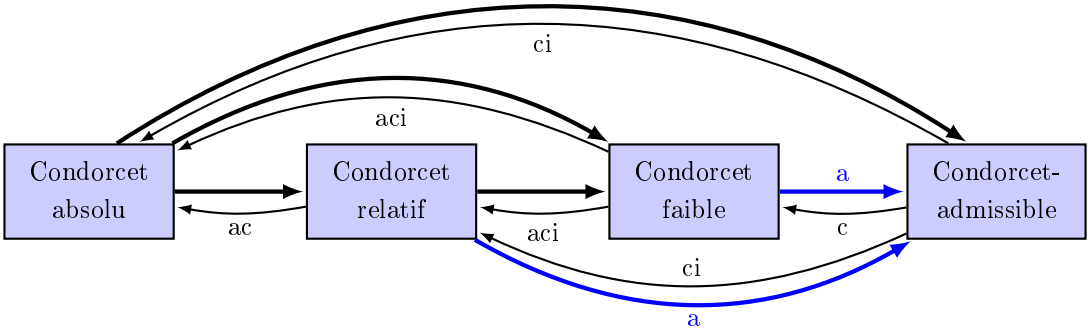


FIGURE 1.6 – Implication entre les notions Condorcet (a : préférences antisymétriques, c : préférences complètes, i : nombre impair d'électeurs).

Définition 1.29 (candidat Condorcet-admissible)
Soit $\omega \in \Omega$ et $c \in \mathcal{C}$. On dit que c est *Condorcet-admissible* dans ω ssi c n'a aucune défaite absolue. C'est-à-dire que, pour tout autre candidat d, on a non($d \, \mathrm{P}_{\mathrm{abs}} \, c$), ce qui s'écrit également :

$$|d \, \mathrm{P}_v \, c| \leq \frac{V}{2}.$$

La proposition suivante indique les liens entre ces différentes notions Condorcet. Toutes les implications mentionnées dans la proposition, ainsi que celles qui en découlent immédiatement, sont représentées à titre de référence dans la figure 1.6. On retiendra notamment que, dans le cas où les préférences sont antisymétriques, la notion de vainqueur de Condorcet absolu est plus forte que celle de vainqueur de Condorcet relatif, qui est plus forte que celle de vainqueur de Condorcet faible, qui est plus forte que celle de candidat Condorcet-admissible.

Proposition 1.30 (implications entre les notions Condorcet)
Si c est vainqueur de Condorcet (absolu), alors c est vainqueur de Condorcet relatif. Si les relations de préférence P_v sont antisymétriques et complètes, alors la réciproque est vraie.

Si c est vainqueur de Condorcet relatif, alors c est vainqueur de Condorcet faible. Si les relations de préférence P_v sont antisymétriques et complètes et si le nombre d'électeurs est impair, alors la réciproque est vraie.

Si c est vainqueur de Condorcet faible et si les relations de préférence P_v sont antisymétriques, alors c est Condorcet-admissible. Si c est Condorcet-admissible et si les relations de préférence P_v sont complètes, alors c est vainqueur de Condorcet faible.

Si c est vainqueur de Condorcet (absolu), alors c est Condorcet-admissible. Si c est Condorcet-admissible, si les relations de préférence P_v sont complètes et si le nombre d'électeurs est impair, alors c est vainqueur de Condorcet (absolu).

Démonstration. Chacune de ces propriétés est triviale ou découle immédiatement de la proposition 1.25. □

Si les relations de préférence P_v sont antisymétriques et complètes (en particulier si ce sont des ordres stricts totaux), alors ces quatre notions se ramènent

à deux seulement : d'une part, vainqueurs de Condorcet absolu et relatif sont équivalents ; d'autre part, vainqueur de Condorcet faible et candidat Condorcet-admissible le sont aussi. Si, en outre, il y a un nombre impair d'électeurs, alors les quatre notions sont équivalentes.

Dans la littérature, l'expression *vainqueur de Condorcet* (sans précision) désigne plutôt, habituellement, le vainqueur de Condorcet relatif. Dans cette thèse, nous utiliserons, cependant, cette expression pour désigner le vainqueur de Condorcet absolu, puisque nous utiliserons cette notion beaucoup plus souvent.

Clairement, le vainqueur de Condorcet absolu ou relatif, s'il existe, est unique. On a également la propriété suivante, tout aussi évidente mais qui nous sera très utile.

Proposition 1.31
Si un candidat est vainqueur de Condorcet (absolu), alors aucun autre candidat n'est Condorcet-admissible.

À présent, on peut définir formellement le critère de Condorcet.

Définition 1.32 (critère de Condorcet)
On dit qu'un SVBE f vérifie le *critère de Condorcet* ssi, pour toute configuration $\omega \in \Omega$ et pour tout candidat $c \in \mathcal{C}$, si c est vainqueur de Condorcet dans ω, alors $f(\omega) = c$. On note indifféremment $\mathbf{Cond}(\Omega)$ ou simplement **Cond** l'ensemble des SVBE sur Ω qui vérifient ce critère, aussi bien que le critère lui-même.

Par commodité de langage, on utilisera aussi le mot *Condorcet* comme un adjectif : ainsi, on dira que f est un mode de scrutin Condorcet. Dans le même esprit, en parlant d'une configuration ω, on dira qu'elle est :
- *Condorcet* s'il existe un vainqueur de Condorcet dans ω ;
- *semi-Condorcet* s'il existe au moins un candidat Condorcet-admissible mais pas de vainqueur de Condorcet dans ω ;
- *non admissible* s'il n'existe aucun candidat Condorcet-admissible dans ω.

De même, nous dirons que la configuration ω est *admissible* s'il existe un candidat Condorcet-admissible (c'est-à-dire si elle est Condorcet ou semi-Condorcet) et qu'elle est *non-Condorcet* s'il n'existe pas de vainqueur de Condorcet (c'est-à-dire si elle est semi-Condorcet ou non admissible).

1.6 Zoologie des modes de scrutin

Pour terminer ce premier chapitre généraliste, nous allons présenter les systèmes de vote fréquemment utilisés dans cette thèse. Afin de ne pas alourdir outre mesure le formalisme, nous ne donnons pas toujours une définition mathématique complète de chacun de ces systèmes. Pour plus de détails, le lecteur pourra se référer par exemple à Tideman (2006). À chaque mode de scrutin, nous attribuons également un nom court (sigle ou abréviation), qui nous servira notamment dans les tableaux et les graphiques de la seconde partie de ce mémoire.

Dans les définitions suivantes, quand un système de vote définit un score, le candidat possédant le score maximal est élu. En cas d'égalité, une règle de départage déterministe permet de choisir un candidat parmi les ex-aequo. Comme nous l'avons remarqué précédemment à propos du vote par notation, ce que nous définissons à chaque fois n'est pas rigoureusement un système de vote, mais une classe de systèmes de vote qui renvoient le même résultat sauf dans un certain ensemble de cas (en général raisonnablement limité), pour lesquels il est nécessaire de préciser la règle de départage.

1.6.1 Modes de scrutin cardinaux

Nous disons qu'un système de vote est *ordinal* ssi la donnée de la relation binaire de préférence d'un électeur est suffisante pour déterminer sa stratégie sincère, c'est-à-dire qu'on a l'implication $[\mathrm{P}_v(\omega_v) = \mathrm{P}_v(\omega'_v)] \Rightarrow [s_v(\omega_v) = s_v(\omega'_v)]$. Dans le cas contraire, nous disons qu'il est *non ordinal*. Par commodité de langage et de façon légèrement informelle, nous disons qu'un *mode de scrutin* est non ordinal s'il n'y a pas de fonction de sincérité relativement naturelle pour laquelle le *système de vote* obtenu est ordinal. Parmi les systèmes non ordinaux, on distingue notamment les systèmes *cardinaux*, où le bulletin de chaque électeur consiste à attribuer une note (ou un objet mathématique similaire) à chaque candidat. On a notamment les trois systèmes suivants.

***Vote par notation* (VN)** Chaque électeur attribue une note à chaque candidat, parmi un ensemble de notes autorisées (qui sont des nombres réels). Le score d'un candidat est sa note moyenne.

***Jugement majoritaire* (JM)** Chaque électeur attribue une appréciation à chaque candidat, parmi un ensemble d'appréciations autorisées. Les appréciations sont des objets de nature mathématique quelconque, munis d'un ordre total, par exemple des nombres réels. En pratique, dans cette thèse, nous représenterons ces appréciations par des nombres réels que nous appellerons des notes, comme en vote par notation, même si seules les positions relatives sont importantes et pas les nombres eux-mêmes [18]. Le score d'un candidat est sa note médiane. Si le nombre d'électeurs est pair, on considère la médiane défavorable : ainsi, si les deux électeurs médians pour un certain candidat lui attribuent les notes 0,4 et 0,5, alors son score est 0,4. Pour la règle utilisée au cas où plusieurs candidats possèdent la même note médiane, voir Balinski et Laraki (2010).

Pour les deux modes de scrutin précédents, on supposera toujours que l'ensemble des notes ou des appréciations autorisées contient au moins deux éléments distincts.

***Vote par assentiment* (VA)** Quand les notes autorisées sont uniquement 0 et 1, les deux systèmes précédents deviennent équivalents [19]. Le mode de scrutin obtenu est appelé le *vote par assentiment*. Pour plus de détails, on pourra consulter, par exemple, Brams et Fishburn (1978).

Dans ces modes de scrutin cardinaux, on représentera une configuration sous la forme suivante (à condition que les relations binaires de préférence soient des ordres stricts faibles cohérents avec les notes).

2	3	4
$a : 1$	$a : 1$	$c : 1$
$b : 0$	$b, c : 0$	$b : 1$
$c : 0$		$a : 0$

Dans cet exemple, les deux premiers électeurs (colonne de gauche) préfèrent le candidat a, puis b, puis c. Ils attribuent la note (ou la valeur d'assentiment) 1

18. On considère donc l'ensemble $\mathbb{R}$ comme muni de sa structure topologique de l'ordre, mais pas comme un espace métrique.

19. Pour le prouver, il est nécessaire de préciser la règle de départage des égalités du jugement majoritaire. Pour plus de précisions sur ce point, on pourra consulter Balinski et Laraki (2010).

à a et 0 aux candidats b et c. Les trois électeurs suivants (colonne du milieu) préfèrent le candidat a, auxquelles ils attribuent la note 1 ; ils sont indifférents entre les candidats b et c, auxquels ils attribuent la note 0. Etc.

À présent, nous allons nous intéresser aux modes de scrutin ordinaux. En général, on supposera que chaque électeur fournit un ordre strict total de préférence. Si l'espace électoral autorise des préférences qui ne sont pas des ordres stricts totaux, une manière simple d'adapter ces règles consiste à décider que, si un électeur ne fournit pas un ordre total, alors son bulletin n'est pas comptabilisé. Dans certains cas, cependant, nous donnerons d'autres généralisations possibles.

1.6.2 Règles positionnelles à score (RPS)

Définition 1.33 (règle positionnelle à score)
Soit $\mathbf{x} = (x_1, \ldots, x_C) \in \mathbb{R}^C$ une liste de nombre réels appelés *poids*. Nous supposerons toujours qu'un tel vecteur est décroissant et qu'en outre, il est non constant, c'est-à-dire que $x_C < x_1$.

Pour tout électeur v et pour tout candidat c, notons $r(v, c)$ le rang de c dans l'ordre de préférence de v (1 pour le meilleur candidat, C pour le pire). Le score du candidat c est défini par :

$$\text{score}(c) = \sum_{v \in \mathcal{V}} x_{r(v,c)}.$$

On appelle *règle positionnelle à score (RPS)* de poids $\mathbf{x}$ le mode de scrutin qui élit le candidat possédant le meilleur score au sens ci-dessus.

Les RPS ont fait l'objet de nombreuses caractérisations, dont on pourra trouver une vue d'ensemble présentée par Merlin (2003). Les trois cas particuliers suivants de RPS sont particulièrement célèbres.

Uninominal* (Uni.) ou *Plurality RPS de poids $\mathbf{x} = (1, 0, \ldots, 0)$.

Veto* ou *Antiplurality RPS de poids $\mathbf{x} = (0, \ldots, 0, -1)$.

***Méthode de Borda* (Bor.)** RPS de poids $\mathbf{x} = (C-1, C-2, \ldots, 0)$. Cette procédure a fait l'objet d'innombrables études. On pourra notamment consulter Fishburn et Gehrlein (1976); Saari (2012).

Ces trois systèmes particuliers possèdent des généralisations assez naturelles au cas où un électeur ne fournit pas un ordre strict total.

Pour l'uninominal, si v préfère strictement un candidat c à tout autre candidat d, au sens où $c \, \text{PP}_v \, d$, alors il donne un point à c ; dans le cas contraire, il ne donne un point à personne. Ainsi, s'il préfère strictement c mais possède un cycle de préférence parmi les autres candidats, sa préférence pour c est tout de même prise en compte. Pour Veto, il est facile de proposer une adaptation similaire, *mutatis mutandis*.

Quand les électeurs possèdent des ordres faibles de préférence, une autre possibilité pour l'uninominal (resp. Veto) consiste à diviser la voix d'un électeur équitablement entre les candidats qu'il place en tête (resp. en bas) de son bulletin.

Pour la méthode de Borda, sauf mention explicite du contraire, nous utiliserons toujours la généralisation suivante. Chaque électeur v apporte au score du candidat c :

	Moyenne	Médiane
Rang	Méthode de Borda	Méthode de Bucklin
Note	Vote par notation	Jugement majoritaire

TABLE 1.2 – Comparaison de principe entre les méthodes de Borda et Bucklin, le vote par notation et le jugement majoritaire

— 1 point pour chaque candidat d tel que $c \,\mathrm{PP}_v\, d$;
— et 0,5 point pour chaque candidat d tel que $c \,\mathrm{I}_v\, d$ ou $c \,\mathrm{PM}_v\, d$.

Cette généralisation est assez naturelle dans la mesure où le nombre total de points donnés par un électeur est constant, que ses préférences soient un ordre strict total ou pas. En particulier, certaines propriétés impliquant le score moyen d'un candidat resteront vraies avec cette généralisation.

1.6.3 Méthode de Bucklin

En *méthode de Bucklin* (Buck.), le dépouillement se déroule par tours. Au tour t, le score provisoire d'un candidat c est le nombre d'électeurs qui placent c entre le rang 1 et le rang t sur leur bulletin. Si au moins un candidat a un score strictement supérieur à $\frac{V}{2}$, alors le candidat possédant le plus haut score est élu. Sinon, on procède au tour de dépouillement $t+1$.

Ce système élit le candidat qui possède le meilleur rang médian, à condition de considérer la médiane défavorable quand le nombre d'électeurs est pair. Par exemple, si les deux rangs médians pour un certain candidat c sont les rangs 2 et 3, alors le rang médian au sens de Bucklin est 3.

Sur le principe, la méthode de Bucklin est à la méthode de Borda ce que le jugement majoritaire est au vote par notation, comme l'illustre la table 1.2 : en effet, la méthode de Bucklin désigne le candidat avec le meilleur rang médian, alors que la méthode de Borda élit celui qui a le meilleur rang moyen. Le jugement majoritaire désigne le candidat avec la meilleure note médiane, alors que le vote par notation élit celui avec la meilleure note moyenne.

La règle énoncée ci-dessus est plus précise que choisir simplement le candidat possédant le meilleur rang médian, puisqu'elle permet, en général, de départager deux candidats ayant le même rang médian. Pour le voir, il suffit d'observer que la définition suivante est équivalente à celle que nous avons donnée.

Pour tout candidat c, notons m_c son rang médian au sens de Bucklin (c'est-à-dire défavorable quand le nombre d'électeurs est pair) et x_c le nombre d'électeurs qui lui attribuent un rang égal ou meilleur (c'est-à-dire moins élevé). Posons :

$$\mathrm{score}(c) = (m_c, x_c).$$

L'ensemble des rangs médians m_c est muni de la relation d'ordre décroissante (un rang moins élevé est meilleur) et l'ensemble des valeurs x_c est muni de la relation d'ordre naturelle (un nombre plus élevé est meilleur). Alors il est facile de montrer que le candidat possédant le meilleur score au sens lexicographique (rang moins élevé et, en cas d'égalité, valeur x_c plus élevée) est bien le vainqueur par la méthode de Bucklin, au sens de la première définition que nous avons donnée.

Cette variante de la définition permet d'étendre naturellement la méthode de Bucklin au cas où les relations binaires de préférence ne sont pas des ordres stricts totaux. En effet, une possibilité élégante consiste à généraliser conventionnellement le rang $r(v,c)$ attribué par l'électeur v au candidat c par $r(v,c) = C - B(v,c)$,

où $B(v, c)$ est le score de Borda tel que nous l'avons généralisé. Ainsi, si l'électeur place deux candidats à égalité devant les autres, on considère qu'il les place au rang 1,5. S'il place trois candidats dans un cycle de préférence non transitif a PP_v b PP_v c PP_v a au-dessus de tous les autres candidats, on considère qu'il les place au rang 2. On peut, ensuite, calculer les scores tels que nous les avons définis ci-dessus.

1.6.4 RPS itérées à élimination simple

Dans cette section et les deux suivantes, nous allons nous intéresser aux méthodes par élimination. De façon informelle, ces systèmes fonctionnent de la façon suivante. On attribue un score à chaque candidat selon une certaine règle. Puis on élimine un certain nombre de candidats qui ont les plus mauvais scores. On itère ce processus jusqu'à ce qu'il ne reste plus qu'un candidat, qui est alors déclaré vainqueur.

En général, nous considérerons le cas où les tours sont virtuels : les électeurs donnent toute l'information utile dans leur bulletin dès le début et les tours sont émulés automatiquement lors du dépouillement (comme pour le système U2TI, évoqué en section 1.4.2). Mais les tours peuvent aussi être réels (comme pour U2T), c'est-à-dire que les électeurs retournent aux urnes à chaque fois. Dans ce cas, le système ne peut généralement pas être décrit par un SVBE, mais par le formalisme d'un système de vote général.

Parmi les modes de scrutin par élimination, on distingue notamment les *RPS itérées* ou *RPSI* : on utilise une règle positionnelle à score et on élimine un certain nombre de candidats à chaque tour. Deux façons de faire, notamment, sont classiques[20] : l'élimination simple et celle basée sur la moyenne.

Définition 1.34 (RPSI à élimination simple)
Soit $\mathbf{x}$ une fonction qui, à chaque entier k compris entre 2 et C, associe un vecteur $\mathbf{x}^k = (x_1^k, \ldots, x_k^k) \in \mathbb{R}^k$.

Le dépouillement se déroule en tours numérotés de 1 à $C-1$. Au tour t, le score de chaque candidat non éliminé est calculé en utilisant la RPS de vecteur $\mathbf{x}^{C+1-t}$ et le candidat avec le score le plus faible est éliminé. Le candidat non éliminé à la fin du dernier tour est déclaré vainqueur.

Par commodité, le score d'un candidat c au tour t est noté $\mathrm{score}_k(c)$, où k est le nombre de candidats encore en lice. Cette convention d'écriture permet que $\mathrm{score}_k(c)$ soit calculé en utilisant le vecteur de poids $\mathbf{x}^k$ qui possède k composantes.

Ce système est appelé la *RPS itérée à élimination simple (RPSI-ES)* de vecteurs de poids $\mathbf{x}$.

Dans cette famille de modes de scrutin, les trois systèmes suivants sont particulièrement étudiés.

Vote à tours instantanés (VTI) La RPSI-ES de l'uninominal est connue sous les noms d'*Instant-Runoff Voting*, *Single Transferable Voting*, *Hare method*, *Alternative vote* ou, en français, *Vote à tours instantanés*[21].

20. Pour les RPSI, nous reprenons en partie une classification proposée par Favardin et Lepelley (2006) pour le cas $C = 3$.

21. La dénomination *Single Transferable Vote* peut mener à une confusion avec sa version à plusieurs vainqueurs, utilisée par exemple en Australie, alors que, à notre connaissance, la terminologie *Instant-Runoff Voting* n'est utilisée que pour la version à un seul vainqueur. C'est pourquoi nous préférons utiliser la seconde, ainsi que sa traduction française : *Vote à tours instantanés*.

***Méthode de Coombs* (Coo.)** RPSI-ES de Veto.

***Méthode de Baldwin* (Bald.)** RPSI-ES de la méthode de Borda. L'intérêt de ce système réside dans le fait qu'un vainqueur de Condorcet a toujours un score de Borda strictement supérieur au score moyen (car il s'agit de la somme de sa ligne dans la matrice des duels). Par conséquent, ce système vérifie le critère de Condorcet.

1.6.5 RPS itérées à élimination par la moyenne

Le principe de ces systèmes est similaire aux RPSI-ES mais, à chaque tour de dépouillement, on élimine en même temps tous les candidats qui ont un score inférieur au score moyen.

Définition 1.35 (RPSI à élimination par la moyenne)
Soit $\mathbf{x}$ une fonction qui, à chaque entier k compris entre 2 et C, associe un vecteur $\mathbf{x}^k = (x_1^k, \ldots, x_k^k) \in \mathbb{R}^k$.

Le dépouillement se déroule en tours numérotés. Le nombre total de tours n'est pas fixé à l'avance. En notant k le nombre de candidats encore en lice au début d'un tour, le score pour ce tour de chaque candidat non éliminé est calculé en utilisant la RPS de vecteur $\mathbf{x}^k$. Chaque candidat possédant un score strictement inférieur au score moyen est éliminé. Le candidat non éliminé à la fin du dernier tour est déclaré vainqueur.

De même que pour les RPSI-ES, le score d'un candidat c à un certain tour de dépouillement est noté $\text{score}_k(c)$, où k est le nombre de candidats encore en lice. Ainsi, $\text{score}_k(c)$ est calculé en utilisant le vecteur de poids $\mathbf{x}^k$ qui possède k composantes.

Ce système est appelé la *RPS itérée à élimination par la moyenne (RPSI-EM)* de vecteurs de poids $\mathbf{x}$.

Deux exemples, notamment, ont fait l'objet d'une attention particulière.

***Méthode de Nanson* (Nan.)** RPSI-EM de la méthode de Borda. Pour les mêmes raisons que la méthode de Baldwin, ce système vérifie le critère de Condorcet. Historiquement, il s'agit de la motivation initiale pour définir la méthode de Nanson, qui inspira ensuite la classe des RPSI-EM par généralisation.

***Méthode de Kim-Roush* (KR)** RPSI-EM de Veto. On pourra consulter notamment Kim et Roush (1996) et Lepelley et Valognes (1999).

Dans la même lignée, il est naturel de considérer le système suivant.

***VTI basé sur la moyenne* (VTIM)** RPSI-EM de l'uninominal.

Dans nos simulations informatiques (chapitre 6 et suivants), nous n'inclurons pas KR et VTIM, bien que ces systèmes soient disponibles dans notre package SVAMP. En effet, ils ont été implantés lors de la phase finale de rédaction, alors que les simulations avaient déjà été réalisées et les chapitres correspondants rédigés. Pour des raisons de temps de calcul et afin d'éviter l'ajout de coquilles de dernière minute dans ces chapitres, nous avons préféré nous en tenir aux simulations déjà effectuées. Nous donnerons tout de même quelques indications sur les premiers résultats de simulation obtenus avec ces deux systèmes dans la section 7.2 et en conclusion de ce mémoire.

1.6.6 Méthodes par éliminations diverses

Les modes de scrutin suivants n'entrent pas dans les familles des RPSI-ES ou des RPSI-EM mais ils utilisent tout de même des éliminations successives.

***Scrutin exhaustif* (SE)** Il s'agit de la variante de VTI avec des tours réels. À chaque tour, chaque électeur vote pour un candidat et le candidat qui reçoit le moins de voix est éliminé. Si les électeurs votent sincèrement, ce mode de scrutin est équivalent à VTI. Quand on utilise ce mode de scrutin en pratique, on ajoute généralement la clause suivante : dès qu'un candidat possède une majorité absolue, il est déclaré vainqueur. Si les électeurs votent sincèrement, cette clause ne change pas le résultat et permet d'économiser des tours inutiles.

***Uninominal à deux tours* (U2T)** Nous avons déjà mentionné plusieurs fois ce système. Quand il y a trois candidats, il est équivalent au scrutin exhaustif ; et sa version instantanée, que nous avons déjà nommée U2TI (section 1.4.2), est alors équivalente à VTI. En particulier, l'exemple de la table 1.1 en section 1.4.2, qui concernait U2T et U2TI avec 3 candidats, démontre également que VTI est en général strictement moins manipulable que SE.

***Méthode de Bucklin itérée* (BI)** On élimine le candidat le moins bon au sens de Bucklin, c'est-à-dire ayant le pire rang médian et, en cas d'égalité, celui qui a le moins d'électeurs lui attribuant son rang médian ou un meilleur rang. Puis on itère, en réévaluant les rangs à chaque tour d'élimination à cause des candidats éliminés.

1.6.7 Méthodes Condorcet

Nous avons déjà évoqué les méthodes de Baldwin et Nanson, qui faisaient respectivement partie des RPSI-ES et des RPSI-EM. Voici, à présent, diverses autres méthodes qui vérifient le critère de Condorcet.

***Condorcet-doyen* (CDoy.)** S'il y a un vainqueur de Condorcet, il est élu. Sinon, un candidat fixé à l'avance, appelé *doyen*, est déclaré vainqueur.

***Condorcet-dictature* (CDict.)** S'il y a un vainqueur de Condorcet, il est élu. Sinon, on déclare vainqueur le candidat qui est en tête du bulletin d'un électeur fixé à l'avance.

Méthode de Black (Condorcet-Borda ou CBor.) S'il y a un vainqueur de Condorcet, il est élu. Sinon, le candidat possédant le plus haut score de Borda est élu. Ce mode de scrutin a été proposé par Black (1958). On pourra également consulter Blin et Satterthwaite (1976).

***VTI à Duels* (VTID)** Le principe de ce système est inspiré de VTI et il nous a été suggéré par Laurent Viennot, que nous remercions. Lors de chaque tour de dépouillement, les deux candidats qui sont en tête dans le moins de bulletins sont sélectionnés pour un duel. Celui des deux qui perd le duel (au sens de la matrice des duels) est éliminé. Puis on procède au tour de dépouillement suivant.

Méthode de Copeland (Cop.) Le score d'un candidat c est égal à son nombre de victoires dans la matrice des duels. Le principal inconvénient de cette méthode est qu'elle mène très souvent à des égalités : par exemple, quel que soit le nombre d'électeurs, s'il y a 3 ou 4 candidats, alors dès qu'il n'y a pas de vainqueur de Condorcet, il y a une égalité à départager. Ainsi, la méthode de Copeland couvre une importante variété de modes de scrutin, suivant la règle de départage utilisée (par exemple au meilleur score de Borda, puis par ordre alphabétique sur les candidats).

***Maximin* (Max.)** Le score d'un candidat c est égal au coefficient non-diagonal minimal de sa ligne dans la matrice des duels ; autrement dit, son score dans son plus mauvais duel. Si les bulletins ne sont pas des ordres stricts totaux, plusieurs variantes sont naturelles, selon qu'on considère la matrice des duels en pourcentages exprimés ou en nombre d'électeurs, le score brut de chaque duel ou le différentiel à l'adversaire. Nous ne nous attarderons pas davantage sur ces variantes.

***Méthode de Kemeny* (Kem.)** La *distance tau de Kendall* entre deux ordres de préférence de $\mathcal{L}_\mathcal{C}$ consiste à compter 1 point pour chaque couple de candidats (c, d) tel que le premier ordre place c devant d et que le second ordre fait l'inverse.

Pour chaque ordre $p_0 \in \mathcal{L}_\mathcal{C}$ sur les candidats, son score de Kemeny est :

$$\text{score}(p_0) = -\sum_{v \in \mathcal{V}} \delta(\mathrm{P}_v, p_0),$$

où δ désigne la distance tau de Kendall. L'ordre possédant le meilleur score est choisi, et le premier candidat de cet ordre est déclaré vainqueur de l'élection. On consultera notamment Kemeny (1959); Young et Levenglick (1978); Saari et Merlin (2000).

On peut aussi exprimer le problème de la façon suivante : il s'agit de trouver une permutation p_0 telle qu'en l'appliquant simultanément aux lignes et aux colonnes de la matrice des duels, la somme des coefficients au-dessous de la diagonale soit minimale. Cette reformulation donne, d'ailleurs, une généralisation naturelle de la méthode quand les préférences ne sont pas des ordres stricts totaux.

***Méthode de Dodgson* (Dodg.)** On appelle *action élémentaire* le fait d'échanger deux candidats consécutifs dans l'ordre de préférence d'un électeur. Le score d'un candidat c est, en négatif, le nombre minimal d'actions élémentaires qu'il faut appliquer au profil de la population pour que c devienne vainqueur de Condorcet.

Autrement dit, on cherche le profil le plus proche de celui de la population (au sens des distances tau de Kendall) parmi ceux qui possèdent un vainqueur de Condorcet et dont les préférences sont des ordres stricts totaux.

Charles Lutwidge Dodgson, qui a conçu ce système de vote, est plus connu sous son nom d'écrivain, Lewis Carroll, avec lequel il a signé *Alice au pays des merveilles*.

Les méthodes de Kemeny et de Dodgson possèdent un inconvénient important : Bartholdi *et al.* (1989b) ont montré qu'il est $\mathcal{NP}$-difficile de déterminer le vainqueur d'une élection dans ces modes de scrutin [22].

22. On pourra également consulter Caragiannis *et al.* (2009) et Betzler *et al.* (2010) pour la complexité de déterminer le vainqueur dans ces systèmes de vote.

***Méthode de Condorcet avec somme des défaites* (CSD)** À présent, on appelle *action élémentaire* le fait d'inverser les préférences d'un électeur à propos d'une paire de candidats (mais sans exiger que sa relation de préférence reste transitive). Le score d'un candidat c est, en négatif, le nombre minimal d'actions élémentaires qu'il faut appliquer au profil de la population pour que c devienne vainqueur de Condorcet.

Autrement dit, on cherche le profil le plus proche de celui de la population (au sens des distances tau de Kendall) parmi ceux qui possèdent un vainqueur de Condorcet, mais sans imposer que les préférences soient des ordres stricts totaux.

En pratique, le score d'un candidat c est :

$$\text{score}(c) = -\sum_{c \text{ ne vainc pas } d} \left(\left\lfloor \frac{V}{2} \right\rfloor + 1 - D_{cd} \right).$$

Pour chaque duel qui n'est pas une victoire, le candidat c perd le nombre de points qui lui manquent pour vaincre. C'est pourquoi, par raccourci de langage, nous appelons ce système la *méthode de Condorcet avec somme des défaites* (*CSD*).

Dans son approche, cette méthode est similaire à la méthode de Dodgson, mais elle possède l'avantage que l'on peut déterminer le vainqueur en temps polynomial[23].

***Méthode par paires ordonnées* (PO)** On construit un graphe orienté dont les sommets sont les candidats. Pour cela, on examine tous les duels entre candidats (dans la matrice des duels) par amplitude décroissante ; par exemple, on commence par le duel (c, d) pour lequel $D_{cd} - D_{dc}$ est maximal. Pour chaque duel, on ajoute une arête au graphe dans le sens de la victoire, sauf si ajouter cette arête crée un cycle dans le graphe.

À la fin du processus, on obtient un graphe orienté transitif dont la relation d'adjacence est incluse dans la relation de victoire. Le candidat qui est le nœud maximal de ce graphe par ordre topologique est déclaré vainqueur (Tideman, 1987). Lorsque plusieurs duels présentent la même amplitude, on se donne généralement une règle de départage qui permet de savoir lequel on examine en premier. C'est ce que nous ferons dans ce mémoire, en particulier pour les simulations de la seconde partie, car la détermination du vainqueur est alors clairement faisable en temps polynomial. En revanche, si on veut calculer l'ensemble des vainqueurs possibles pour toutes les règles de départage, alors le problème est $\mathcal{NP}$-complet (Brill et Fischer, 2012).

***Méthode de Schulze* (Sch.)** On considère le graphe capacitaire défini par la matrice des duels : pour chaque couple de candidats (c, d), il y a une arête dont le poids[24] est D_{cd}. Il est conseillé de voir une arête comme un tuyau unidirectionnel de c vers d et le poids comme la largeur de ce tuyau. On définit la largeur d'un chemin comme le poids de l'arête minimale de ce chemin. On note $\text{score}(c, d)$ la largeur du plus large chemin du candidat c vers le candidat d. Le candidat c est

23. Le système CSD correspond au système noté V par Caragiannis *et al.* (2014) (cet article utilise une autre notation pour le nombre d'électeurs). Les auteurs font remarquer que CSD donne une approximation du score de Dodgson à un facteur multiplicatif près égal au nombre de candidats. Pour un nombre d'électeurs impair, CSD est équivalent à la *règle de Dodgson simplifiée* de Tideman (2006). Merci à Jérôme Lang pour nous avoir indiqué ces antécédents.

24. Il est aussi possible de prendre les poids $D_{cd} - D_{dc}$. Comme nous considérons des ordres stricts totaux ici, on a $D_{cd} - D_{dc} = 2D_{cd} - V$, ce qui est équivalent à D_{cd} à multiplication par la constante positive 2 et soustraction de la constante V près. Pour cette raison, les deux définitions aboutissent au même système de vote.

dit *meilleur* que d ssi $\text{score}(c, d) \geq \text{score}(d, c)$. Le candidat c est un *vainqueur potentiel* si aucun candidat d n'est meilleur que c.

Schulze (2011) prouve que l'ensemble des vainqueurs potentiels est toujours non vide. Il est facile de voir que, s'il existe un vainqueur de Condorcet, alors c'est le seul vainqueur potentiel.

Dans la méthode de Schulze telle qu'elle est promue par son inventeur, le vainqueur est tiré au hasard parmi les vainqueurs potentiels. Dans ce mémoire, en revanche, nous considérerons que la règle de départage utilisée est déterministe.

Chapitre 2

Condorcification

Certains auteurs, comme Chamberlin *et al.* (1984), Smith (1999), Favardin *et al.* (2002), Lepelley et Valognes (2003), Favardin et Lepelley (2006) ou Tideman (2006) ont émis l'intuition que les modes de scrutin qui vérifient le critère de Condorcet ont une tendance générale à être moins manipulables que les autres. Dans ce chapitre, nous examinons cet aspect et nous montrons qu'à condition d'être précisée, cette intuition est justifiée. Nous reprenons et développons ici les travaux présentés par Durand *et al.* (2012, 2014d,e).

Tout d'abord (section 2.1), nous définissons le *critère de la coalition majoritaire informée* (**CMInf**), qui sera utilisé comme hypothèse pour les principaux théorèmes de ce chapitre. Comme ce critère est vérifié par la plupart des modes de scrutin classiques, à l'exception notable de Veto, ceci donnera une certaine généralité à nos résultats.

Dans la section 2.2, nous présentons une série de lemmes simples que nous utiliserons souvent par la suite, notamment pour prouver les théorèmes de ce chapitre. Ils permettent d'établir des liens entre la manipulabilité et les résultats des duels électoraux, avant et après la manipulation. Ces résultats sont bien connus dans la littérature dédiée à la manipulation ; nous les adaptons simplement au cas général où les relations binaires de préférence sont quelconques.

Ensuite, nous définissons la *condorcification* d'un mode de scrutin (section 2.3) : dans le nouveau système, on ajoute simplement un test préliminaire sur l'existence d'un vainqueur de Condorcet et, dans ce cas, il est déclaré vainqueur ; sinon, on utilise le mode de scrutin original. Nous systématisons ainsi le procédé utilisé pour définir la méthode de Black (1958) à partir de celle de Borda.

Ce système hybride pourrait être vu comme une construction artificielle et, en tant que telle, il n'est pas évident, *a priori*, qu'il possède de bonnes propriétés. Mais, de façon assez surprenante, nous montrons en section 2.4 que, si un système de vote vérifie le critère courant **CMInf**, alors sa condorcification est au plus aussi manipulable que le système original : c'est l'objet du *théorème faible de condorcification* (théorème 2.9).

Dans le cas général, ce théorème suppose que la condorcification d'un système de vote est réalisée en utilisant la notion de vainqueur de Condorcet absolu. Nous examinons ce qu'il en est avec des notions moins exigeantes, comme le vainqueur de Condorcet relatif (section 2.5). Nous montrons que, même si ces variantes de condorcification ne permettent pas de diminuer la manipulabilité dans le cas général, celle basée sur le vainqueur de Condorcet relatif permet de diminuer la

manipulabilité pour l'uninominal et U2TI, mais pas pour VTI ou le vote par assentiment par exemple.

Notre but, ensuite, est d'examiner si la condorcification permet de réduire *strictement* la manipulabilité. Pour cela, nous définissons la notion de *vainqueur de Condorcet résistant*, un candidat qui, pour toute paire d'autres candidats, dispose d'une majorité d'électeurs qui le préfèrent simultanément aux deux membres de la paire (section 2.6). Nous montrons que cette définition est équivalente à la propriété suivante : dans tout mode de scrutin respectant le critère de Condorcet, ce candidat est élu et la configuration est non manipulable.

À l'aide de la notion de vainqueur de Condorcet résistant, nous montrons en section 2.7 que pour tous les modes de scrutin classiques sauf Veto, leur condorcification est *strictement* moins manipulable que le mode de scrutin original : c'est l'objet du *théorème fort de condorcification* (théorème 2.20).

En section 2.8, nous exploitons la notion de vainqueur de Condorcet résistant pour donner une borne supérieure de manipulabilité pour les systèmes Condorcet, et nous montrons que, dans l'espace électoral des ordres stricts totaux avec $C \geq 6$, cette borne supérieure est atteinte.

Enfin, en section 2.9, nous soulignons une conséquence importante des théorèmes de condorcification : la recherche d'un mode de scrutin de manipulabilité minimale (au sein de **CMInf**) peut être restreinte aux systèmes Condorcet.

Indépendamment de nos recherches, Green-Armytage *et al.* (2014) ont également introduit, dans un rapport de recherche disponible en ligne, le critère de la coalition majoritaire informée sous le nom de *conditional majority determination* (CMD) et énoncé une version du théorème faible de condorcification. Cependant, d'une part, nous nous permettons de revendiquer l'antériorité (Durand *et al.*, 2012). D'autre part, leur preuve n'est pas correcte si les relations de préférence ne sont pas des ordres stricts, hypothèse qui est pourtant absente de leur formulation. En outre, quand les préférences ne sont pas des ordres stricts, il n'est pas évident de savoir quelle notion de vainqueur de Condorcet ils utilisent ; l'interprétation la plus naturelle consiste à comprendre qu'il s'agit du vainqueur de Condorcet relatif [1], mais on verra en section 2.5 que le résultat est alors erroné. À la décharge des auteurs, l'objet central de leur rapport n'est pas la condorcification mais l'étude par simulation de la manipulabilité de divers systèmes de vote pour $C = 3$ candidats. Nous avons eu de cordiaux et enrichissants échanges avec les auteurs de ce rapport en juin 2014 ; mais, à notre connaissance, la version de leur rapport qui est en ligne à l'heure où nous écrivons ce mémoire n'a pas encore inclus ces corrections et la référence à notre travail. Enfin, nous allons plus loin que ce résultat dans la section 2.5 et les suivantes, notamment en examinant les variantes de condorcification et en prouvant la version forte du théorème.

2.1 Critère de la coalition majoritaire informée (CMInf)

Nous avons déjà présenté le critère de Condorcet (définition 1.32). Nous allons maintenant présenter un critère plus faible, le *critère de la coalition majoritaire informée* (**CMInf**).

1. Leur définition est la suivante : « A 'Condorcet winner' is a candidate who, according to the ballot data, would win a two-person race against any other candidate ». C'est-à-dire : « Un *vainqueur de Condorcet* est un candidat qui, selon les données des bulletins, remporterait un duel électoral contre tout autre candidat ».

Définition 2.1 (critère de la coalition majoritaire informée)
Soit f un SVBE.

On dit que f vérifie le *critère de la coalition majoritaire informée* (**CMInf**) ssi toute coalition majoritaire et informée de ce que font les autres électeurs peut décider du résultat [2]. Formellement, $\forall M \in \mathcal{P}(\mathcal{V})$, si $\text{card}(M) > \frac{V}{2}$, alors : $\forall c \in \mathcal{C}$, $\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}$, $\exists \omega_M \in \Omega_M$ t.q. $f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c$.

Comme pour le critère de Condorcet, la notation **CMInf** désigne indifféremment le critère lui-même ou l'ensemble des SVBE (sur Ω) qui le vérifient.

Proposition 2.2
Si l'espace électoral autorise tout candidat comme préféré (ce qui est une hypothèse courante, cf. définition 1.10), alors **Cond** $\subseteq$ **CMInf**.

Démonstration. Si un mode de scrutin vérifie le critère de Condorcet, il suffit à une coalition majoritaire de prétendre préférer un candidat c à tous les autres candidats pour que c apparaisse comme un vainqueur de Condorcet et soit élu. □

Il est facile d'étendre les définitions de **CMInf** et de **Cond** à un système de vote général (section 1.4). Dans ce cas, on verrait que **CMInf** est une propriété du *mode de scrutin*, exprimant le pouvoir accordé par la règle de dépouillement à une majorité stricte d'électeurs, alors que **Cond** est une propriété du *système de vote* puisqu'elle relie les préférences et le vainqueur du vote sincère. Par conséquent, on peut dire qu'un mode de scrutin vérifie **CMInf** sans qu'il soit crucial de préciser l'espace électoral utilisé ; mais, pour définir **Cond**, il est nécessaire d'expliciter l'espace électoral et les fonctions de sincérité.

Proposition 2.3
Les systèmes de vote suivants vérifient **CMInf** *: BI, le jugement majoritaire, l'uninominal, U2TI, VTI, VTIM, le vote par assentiment, le vote par notation et les méthodes de Borda, Bucklin et Coombs, ainsi que tous les systèmes Condorcet.*

Cette propriété n'est pas difficile à démontrer mais nous attendrons le chapitre 3 afin d'inclure ce résultat dans des propositions plus détaillées sur chacun de ces modes de scrutin.

Nous verrons aussi dans le chapitre 3 que Veto est l'un des rares systèmes de vote souvent étudiés ne vérifiant pas **CMInf** [3] ; cependant, il n'est guère utilisé en pratique.

Parmi les systèmes de vote courants, il est intéressant de constater que même ceux dont les arguments habituels ne reposent pas sur la notion de majorité (VA, VN ou JM par exemple) vérifient **CMInf**, qui est clairement une propriété majoritaire.

En pratique, ces observations donnent une large gamme d'applications aux résultats de ce chapitre. D'un point de vue théorique, on peut se demander s'il y a une raison profonde pour laquelle la plupart des systèmes de vote classiques vérifient **CMInf** [4]. Il nous semble que cette question est une problématique intéressante pour de futurs travaux, que ce soit avec une approche de sciences dures ou de sciences humaines. Nous y reviendrons également dans le chapitre 3, où nous étudierons de manière plus générale des critères liés à la notion de majorité.

2. Cette notion est basée sur la même idée que ce que Peleg (1984) appelle le *troisième jeu simple* associé à un système de vote.

3. De manière plus générale, nous donnerons dans la proposition 3.13 une condition nécessaire et une condition suffisante pour qu'une RPS vérifie **CMInf**.

4. Au moins, c'est le cas pour les systèmes de vote utilisés en pratique qui traitent les électeurs et les candidats de façon symétrique, aux cas d'égalité près. Nous évoquerons certains modes de scrutin délibérément asymétriques dans le chapitre 4.

2.2 Liens entre manipulation et résultats des duels électoraux

À présent, nous présentons une série de lemmes très simples qui permettent d'établir des liens entre la manipulabilité et le résultat des duels électoraux, avant et après la manipulation. Les démonstrations découlent immédiatement des définitions et sont omises. Tous ces résultats font partie d'une gymnastique habituelle dans la littérature sur la manipulation ; nous les adaptons simplement au cas général où les relations binaires de préférence ne sont pas forcément des ordres stricts totaux.

Lemme 2.4
Soit $(\omega, \psi) \in \Omega^2$. On note $\mathrm{w} = f(\omega)$ et $c = f(\psi)$ et on suppose que f est manipulable en ω vers ψ.

Le résultat du duel de c contre w ne peut pas être meilleur dans ψ que dans ω, c'est-à-dire qu'on a :

$$D_{c\mathrm{w}}(\psi) \leq D_{c\mathrm{w}}(\omega).$$

Par conséquent, si w est Condorcet-admissible en ω, alors c ne peut pas avoir une victoire contre w dans ψ ; en particulier, c n'est pas vainqueur de Condorcet en ψ.

On remarquera que pour le lemme 2.4, aucune hypothèse n'est faite sur le système de vote f : en particulier, il n'est pas nécessaire qu'il vérifie **CMInf** ou **Cond**. Par ailleurs, ce lemme entraîne immédiatement le suivant.

Lemme 2.5
Si le système de vote vérifie le critère de Condorcet, alors une configuration Condorcet ne peut pas être manipulable vers une autre configuration Condorcet.

En revanche, une configuration admissible ω peut être manipulable vers une configuration Condorcet ψ, mais seulement si un candidat non-admissible est élu en ω.

Les lemmes 2.4 et 2.5 généralisent un résultat classique de Moulin (1978, chapitre I, théorème 1). Contrairement à ce qu'on peut lire parfois, ce résultat n'empêche absolument pas une configuration Condorcet d'être manipulable.

Dans le même esprit, on dispose du lemme suivant, qui s'applique lorsque les préférences sont antisymétriques (ce qui est une hypothèse courante) et qui se concentre sur le résultat $D_{\mathrm{w}c}$ du duel de w contre c (au lieu de $D_{c\mathrm{w}}$ dans le lemme 2.4).

Ce résultat ne sera pas nécessaire pour prouver les théorèmes de condorcification (théorèmes 2.9 et 2.20), mais il aurait été dommage d'énoncer le lemme 2.4 sans en profiter pour mentionner le lemme 2.6, qui est extrêmement similaire et que nous utiliserons également à plusieurs occasions.

Lemme 2.6
Soit $(\omega, \psi) \in \Omega^2$. On note $\mathrm{w} = f(\omega)$ et $c = f(\psi)$ et on suppose que f est manipulable en ω vers ψ. On suppose, en outre, que les relations binaires de préférence sont antisymétriques dans ω.

Le résultat du duel de w contre c ne peut pas être moins bon dans ψ que dans ω, c'est-à-dire qu'on a :

$$D_{\mathrm{w}c}(\psi) \geq D_{\mathrm{w}c}(\omega).$$

Par conséquent, si w est vainqueur de Condorcet en ω, alors c conserve une défaite contre w en ψ ; en particulier, c n'est pas Condorcet-admissible en ψ.

Enfin, voici l'autre lemme qui sera utilisé pour prouver les théorèmes de condorcification. Sa preuve est tout aussi immédiate.

Lemme 2.7
Soit $\omega \in \Omega$. On suppose que f vérifie **CMInf**.
Si $f(\omega)$ n'est pas Condorcet-admissible en ω, alors f est manipulable en ω.

Une conséquence immédiate de ce lemme est que, si une configuration ω est non admissible, alors tout système de vote f vérifiant **CMInf** est manipulable en ω.

2.3 Définition de la condorcification

L'idée de la *condorcification* remonte à Black (1958), qui l'a appliquée à la méthode de Borda pour définir la méthode qui porte désormais son nom. Cependant, il semble que, depuis, ce principe n'ait été ni exploité pour d'autres modes de scrutin ni étudié en général avant les travaux de Green-Armytage *et al.* (2014) et ceux de Durand *et al.* (2012). Il nous semble que cela s'explique par le fait que ce procédé, produisant des modes de scrutin hybrides qui semblent artificiels et manquent peut-être d'élégance mathématique, ne possédait pas de bonne propriété connue jusqu'à présent. Nous espérons remédier à ce fait dans le présent chapitre.

Définition 2.8 (condorcification)
On appelle *condorcification absolue de f*, ou simplement *condorcification de f*, le SVBE suivant :

$$f^* : \left| \begin{array}{ccc} \Omega & \to & \mathcal{C} \\ \omega & \to & \left| \begin{array}{l} \text{si } \omega \text{ a un vainqueur de Condorcet (absolu) } c, \text{alors } c, \\ \text{sinon, } f(\omega). \end{array} \right. \end{array} \right.$$

Par définition, f^* vérifie le critère de Condorcet. Par conséquent, si on fait l'hypothèse courante que l'espace électoral autorise tout candidat comme préféré, alors f^* vérifie **CMInf** (proposition 2.2).

Afin de comprendre comment cette définition s'applique dans le cas d'un mode de scrutin cardinal, examinons la condorcification du vote par notation, dans l'espace électoral de référence où $\omega_v = (p_v, u_v, a_v)$.

— Chaque électeur communique un état.
— S'il y a un vainqueur de Condorcet (calculé avec les relations p_v communiquées), alors il est élu.
— Sinon, le candidat avec la note moyenne maximale (calculée avec les vecteurs de notes u_v communiqués) est élu.

Pour implanter ce système en pratique, il est maintenant nécessaire d'inclure les notes (comme pour le vote par notation original, voir les sections 1.2.1 et 1.4.3) mais aussi les ordres de préférence dans les bulletins. Cependant, si le planificateur social a adopté l'hypothèse que $c \; p_v \; d \Leftrightarrow u_v(c) > u_v(d)$ (que nous avions mentionnée dans la section 1.1.3), alors il est suffisant de demander les notes, puisque les ordres de préférence s'en déduisent immédiatement.

Pour chaque mode de scrutin, nous préfixerons son nom par « Condorcet » et son abréviation par la lettre C pour désigner sa condorcification. Par exemple, la condorcification de la méthode de Bucklin sera appelée Condorcet-Bucklin et abrégée en CBuck. Nous étudierons également de manière approfondie la condorcification de VTI, notée CVTI.

En fait, nous avons déjà utilisé implicitement ce procédé pour définir le système Condorcet-doyen à partir d'un mode de scrutin au résultat constant (qui élit toujours le même candidat appelé *doyen*) et pour définir le système Condorcet-dictature à partir d'un mode de scrutin dictatorial.

2.4 Théorème faible de condorcification

Le théorème suivant montre un avantage important de la condorcification : pour un système vérifiant **CMInf**, sa condorcification est au plus aussi manipulable que le système original.

Théorème 2.9 (théorème faible de condorcification)
Soit f un SVBE. On suppose que f vérifie **CMInf**.
Alors sa condorcification f^* est au plus aussi manipulable que f :

$$\mathrm{MC}_{f^*} \subseteq \mathrm{MC}_f.$$

Pour donner l'intuition de la preuve, rappelons que les configurations ω sont de trois types.

Condorcet Il y a un vainqueur de Condorcet.

Semi-Condorcet Il y a au moins un candidat Condorcet-admissible mais pas de vainqueur de Condorcet.

Non admissible Aucun candidat n'est Condorcet-admissible.

Si le système de vote vérifie **CMInf**, nous allons prouver que, lorsqu'une configuration est Condorcet, alors on ne peut pas empirer la manipulabilité du système en choisissant celui-ci comme vainqueur de l'élection. En effet, d'une part, nous verrons que cela ne peut pas empirer la manipulabilité de la configuration considérée elle-même : si le vainqueur était différent dans le système original, alors la configuration était manipulable en faveur du vainqueur de Condorcet (ce sera une conséquence du lemme 2.7), donc on ne peut pas empirer la situation. D'autre part, cela ne peut pas rendre une autre configuration Condorcet manipulable (lemme 2.5, qui découle du lemme 2.4) ; ça ne peut amener une configuration semi-Condorcet à être manipulable que si elle l'était déjà au départ (lemme 2.4) ; et les configurations non admissibles sont condamnées à être manipulables de toute façon (lemme 2.7).

Démonstration. Supposons que f^* soit manipulable en ω vers ψ, mais que f n'est pas manipulable en ω.

Puisque f n'est pas manipulable en ω, le lemme 2.7 assure que $f(\omega)$ est Condorcet-admissible en ω. S'il est vainqueur de Condorcet en ω, alors $f^*(\omega) = f(\omega)$; sinon, il n'y a pas de vainqueur de Condorcet en ω (proposition 1.31) donc, par définition de f^*, on a aussi $f^*(\omega) = f(\omega)$.

À présent, notons $\mathrm{w} = f^*(\omega) = f(\omega)$ et $c = f^*(\psi)$. Comme w est Condorcet-admissible en ω, le lemme 2.4 (appliqué à f^*) assure que c n'est pas vainqueur de Condorcet en ψ. Donc, par définition de f^*, on a $f^*(\psi) = f(\psi)$.

Par conséquent, on a $f(\omega) = f^*(\omega)$ et $f(\psi) = f^*(\psi)$, donc f est manipulable en ω : c'est absurde ! □

La démonstration de Green-Armytage *et al.* (2014) est articulée en 5 points numérotés. À la fin du point 2, les auteurs établissent qu'il n'existe aucun candidat B tel qu'une majorité absolue stricte d'électeurs préfère B à un certain candidat A ; c'est-à-dire que le candidat A est Condorcet-admissible, dans notre

terminologie. Au point 3, ils en déduisent qu'aucun candidat B distinct de A ne peut être vainqueur de Condorcet (apparemment relatif).

Cependant, en général, cette implication n'est pas correcte : en effet, un candidat B peut être vainqueur de Condorcet relatif en étant préféré à A par 45 % des électeurs pendant que A est préféré à B par 40 % des électeurs. Pour que leur démonstration soit correcte, il faut donc, soit la limiter à des ordres stricts totaux (ce qui semble implicitement être leur cas), soit utiliser la notion de vainqueur de Condorcet absolu.

En fait, nous montrerons dans la section 2.5 que, si on effectue la condorcification en utilisant la notion de vainqueur de Condorcet relatif, alors non seulement leur démonstration devient incorrecte, mais le théorème de condorcification lui-même devient faux dans le cas général : pour certains modes de scrutin, il est possible de rendre certaines configurations manipulables alors qu'elles ne l'étaient pas.

Par ailleurs, une objection classique aux systèmes de vote Condorcet est qu'ils souffrent tous du *no-show paradox* (Moulin, 1988; Pérez, 2001) : il existe au moins une configuration où un électeur peut manipuler par abstention[5]. Cependant, le théorème faible de condorcification 2.9 nuance cette critique : lorsque l'on compare un système f et sa condorcification f^*, si on est dans une configuration où les manipulateurs peuvent exploiter l'abstention stratégique dans f^*, cela signifie que le système de vote original f est manipulable de toute façon (pas forcément par abstention). Pour le voir, il suffit de considérer un espace électoral où l'abstention est un état sincère possible et par conséquent un bulletin disponible dans un SVBE[6].

2.5 Variantes de condorcification

À présent, nous montrons que le théorème faible de condorcification 2.9 est bien ajusté, dans le sens suivant : il ne se généralise pas quand on utilise une version étendue de la condorcification impliquant des vainqueurs de Condorcet relatifs, faibles ou des candidats Condorcet-admissibles.

2.5.1 Définition des variantes de condorcification

Depuis le début de ce chapitre, nous nous sommes concentrés sur la notion de vainqueur de Condorcet absolu, qui est basée sur des victoires absolues : dans le duel d'un tel candidat c contre tout adversaire d, on a $|c \mathrel{\mathrm{P}_v} d| > \frac{V}{2}$ et $|d \mathrel{\mathrm{P}_v} c| \leq \frac{V}{2}$, ces deux conditions se résumant à la première quand les préférences sont antisymétriques (ce qui est une hypothèse courante).

5. De nombreux autres systèmes de vote souffrent de ce paradoxe, par exemple les RPSI-ES. Mais dans celles-ci, une situation manipulable par abstention est toujours également manipulable en utilisant des bulletins stratégiques qui ne sont pas des abstentions. En revanche, dans la méthode de Kim-Roush par exemple, il existe des profils qui ne sont manipulables que par abstention. Pour plus de détails, voir Lepelley et Merlin (2001).

6. Par exemple, on peut considérer un espace électoral où l'état d'un électeur est une paire constituée d'une relation de préférence sur les candidats et d'une variable booléenne qui représente s'il s'abstient ou pas. Une autre possibilité consiste à considérer un espace électoral où l'ensemble des états possibles d'un électeurs est l'ensemble des ordres stricts totaux, auxquels on ajoute la relation binaire vide (indifférence parfaite), qui représente l'abstention sincère mais qui peut aussi être utilisée comme comportement stratégique. Enfin, il est possible de supposer que l'abstention sincère n'existe pas mais que l'abstention stratégique est possible ; pour cela, il faut utiliser un système de vote général (et non un SVBE) où l'abstention est une des stratégies autorisées.

À la place, nous aurions pu considérer les variantes plus faibles de cette notion, dont on peut rappeler informellement les définitions.

Vainqueur de Condorcet relatif $|c\,\mathrm{P}_v\,d| > |d\,\mathrm{P}_v\,c|$.

Vainqueur de Condorcet faible $|c\,\mathrm{P}_v\,d| \geq |d\,\mathrm{P}_v\,c|$.

Candidat Condorcet-admissible $|d\,\mathrm{P}_v\,c| \leq \frac{V}{2}$.

Pour tout SVBE f, de la même façon qu'on a défini sa condorcification f^*, on définit les systèmes de vote suivants.

f^{rel} : s'il y a un vainqueur de Condorcet relatif, il est élu, sinon on applique f. Ce système sera appelé la *condorcification relative* de f.

$f^{!\mathrm{faible}}$: s'il y a un unique vainqueur de Condorcet faible, il est élu, sinon on applique f.

f^{faible} : s'il y a au moins un vainqueur de Condorcet faible, on en choisit un arbitrairement, sinon on applique f.

$f^{!\mathrm{adm}}$: s'il y a un unique candidat Condorcet-admissible, il est élu, sinon on applique f.

f^{adm} : s'il y a au moins un candidat Condorcet-admissible, on en choisit un arbitrairement, sinon on applique f.

Dans le cas de f^{faible} ou f^{adm}, ceci ne définit pas un SVBE à proprement parler mais plutôt une classe de SVBE, puisque la définition laisse des choix arbitraires à faire. Ce raccourci de notation ne posera pas de problème en pratique car les résultats que nous prouverons seront vrais pour tout SVBE répondant à cette définition, quels que soient les choix opérés.

Nous allons examiner si l'une ou l'autre de ces notions permet d'obtenir un résultat similaire au théorème faible de condorcification 2.9. Cependant, nous allons voir que ce n'est pas le cas.

À tout moment de cette section, on se place dans un espace électoral où l'ensemble des relations binaires de préférence possibles pour chaque électeur est égal à l'ensemble des ordres stricts faibles (pour éviter, notamment, que les notions de vainqueur de Condorcet absolu ou relatif ne soient confondues, ce qui ne nous permettrait pas de voir quelle est la notion qui rend le théorème valide). En particulier, on se place dans un cas favorable où les relations binaires de préférence sont antisymétriques, ce qui permet de simplifier les raisonnements (alors qu'on se souviendra que cette hypothèse n'est pas nécessaire pour le théorème faible de condorcification 2.9).

2.5.2 Variantes de condorcification : résultat général

Commençons par l'exemple d'un mode de scrutin qui est, certes, un peu artificiel mais qui va nous permettre de prouver de façon assez concise qu'aucune des notions ci-dessus ne permet de généraliser le théorème faible de condorcification 2.9.

Proposition 2.10

Le SVBE f considéré est Condorcet-doyen.

Il existe au moins une valeur de (V, C) telle que pour tout système g choisi parmi les variantes f^{rel}, $f^{!faible}$, f^{faible}, $f^{!adm}$ ou f^{adm}, on n'a pas $\mathrm{MC}_g \subseteq \mathrm{MC}_f$.

Démonstration. Considérons $V = 5$ électeurs, $C = 3$ candidats notés a, b, c. Le doyen est a. Considérons la configuration ω suivante et sa matrice des duels $D(\omega)$.

a	a, c	b	b	c
b	b	a, c	c	a
c			a	b

$D(\omega)$	a	b	c
a	–	3	1
b	2	–	3
c	2	2	–

Il n'y a pas de vainqueur de Condorcet absolu : en effet, il faudrait qu'un candidat ait au moins 3 points dans chaque duel de sa ligne de $D(\omega)$, ce qui n'est pas le cas. Donc le doyen est élu : $f(\omega) = a$. Il est facile de voir que f n'est pas manipulable en ω : si c'était le cas, d'après le lemme 2.5, ce serait vers une configuration sans vainqueur de Condorcet. Mais, dans toutes les configurations de ce type, c'est a qui est élu, donc la manipulation échoue.

Les candidats b et c ne sont même pas Condorcet-admissibles (ils ont chacun une défaite absolue dans leur colonne). Par conséquent, quelle que soit la variante g étudiée, on a $g(\omega) = a$.

Considérons maintenant la configuration suivante ψ, où les deux derniers électeurs tentent de faire gagner c au lieu de a.

a	a, c	b	**c**	c
b	b	a, c	**b**	**b**
c			a	**a**

$D(\psi)$	a	b	c
a	–	2	1
b	3	–	2
c	2	3	–

À présent, c est un vainqueur de Condorcet relatif, donc également un vainqueur de Condorcet faible et un candidat Condorcet-admissible. Les candidats a et b ne sont même pas Condorcet-admissibles (ils ont chacun une défaite dans leur colonne). Par conséquent, quelle que soit la variante g utilisée, on a $g(\psi) = c$.

En conclusion, g est manipulable en ω vers ψ en faveur de c, alors que f n'est pas manipulable en ω. □

Cette proposition montre pourquoi la condorcification que nous avons utilisée dans le théorème faible de condorcification 2.9 repose sur la notion de vainqueur de Condorcet absolu. En effet, si on conserve la même hypothèse $f \in$ **CMInf** (qui est trivialement vérifiée par Condorcet-doyen), le théorème n'est pas généralisable aux vainqueurs de Condorcet relatifs, faibles ou aux candidats Condorcet-admissibles.

Une autre possibilité, pour généraliser le théorème en utilisant la notion de vainqueur de Condorcet relatif, de vainqueur de Condorcet faible ou de candidat Condorcet-admissible, consisterait à remplacer l'hypothèse $f \in$ **CMInf** par une autre hypothèse. Ceci dit, pour que le théorème obtenu ait un domaine de validité aussi intéressant que le théorème 2.9, il faudrait qu'il soit au moins applicable à des systèmes de vote usuels, comme l'uninominal, U2TI, VTI ou le vote par assentiment.

Nous allons donc examiner ces systèmes de vote particuliers. Nous verrons qu'en général, on ne peut avoir une telle généralisation du théorème 2.9 ; mais que pour certains de ces systèmes, la condorcification relative f^{rel} peut être moins manipulable que le système original f, et même que sa condorcification f^*.

2.5.3 Variantes de condorcification pour l'uninominal

Tout d'abord, précisons comment nous généralisons l'uninominal, U2TI et VTI si les électeurs peuvent avoir des ordres stricts faibles de préférence. Nous allons considérer les deux règles que nous avons présentées dans la section 1.6.2 : si un électeur a plusieurs candidats préférés à égalité (parmi les candidats non éliminés),

on peut décider que sa voix est divisée également entre eux, ou n'est pas comptée du tout. Nos résultats seront valables dans les deux cas.

Proposition 2.11
Le SVBE f considéré est l'uninominal.

1. *On a :* $\mathrm{MC}_{f^{\mathrm{rel}}} \subseteq \mathrm{MC}_{f^*}$.
2. *Il existe (V, C) t.q. l'inclusion précédente est stricte.*
3. *Il existe (V, C) t.q. pour tout système g choisi parmi les variantes $f^{!faible}$, f^{faible}, $f^{!adm}$ ou f^{adm}, on n'a pas* $\mathrm{MC}_g \subseteq \mathrm{MC}_f$.

Afin de démontrer le point 1, nous allons d'abord prouver le lemme suivant, que nous appliquerons ensuite à $h = f^*$.

Lemme 2.12
Soit h un SVBE. On suppose que, pour toute configuration ω et tout candidat $d \neq h(\omega)$, si $h(\omega)$ a une défaite relative contre d, alors h est manipulable en ω.
Alors $\mathrm{MC}_{h^{\mathrm{rel}}} \subseteq \mathrm{MC}_h$.

Démonstration. Supposons qu'il existe une configuration ω où h n'est pas manipulable mais où h^{rel} est manipulable vers une certaine configuration ψ.

Notons $c = h(\omega)$. Comme h n'est pas manipulable en ω, l'hypothèse du lemme implique que c n'a pas de défaite relative, donc aucun autre candidat ne peut être vainqueur de Condorcet relatif. Par conséquent, $h^{\mathrm{rel}}(\omega) = c$.

Notons $d = h^{\mathrm{rel}}(\psi)$. Comme le résultat relatif du duel de d contre c ne peut pas avoir été amélioré par la manipulation (lemmes 2.4 et 2.6), d n'a toujours pas de victoire relative contre c, donc il n'est pas vainqueur de Condorcet relatif. Par définition de h^{rel}, on en déduit que $h(\psi) = d$.

Ainsi, $h(\omega) = h^{\mathrm{rel}}(\omega) = c$ et $h(\psi) = h^{\mathrm{rel}}(\psi) = d$ donc h est manipulable en ω vers ψ : c'est absurde ! □

Nous pouvons maintenant prouver la proposition 2.11.

Démonstration. 1. On envisage d'appliquer le lemme à $h = f^*$ et $h^{\mathrm{rel}} = (f^*)^{\mathrm{rel}} = f^{\mathrm{rel}}$. Il suffit de montrer que $h = f^*$ vérifie l'hypothèse du lemme pour en conclure que $\mathrm{MC}_{f^{\mathrm{rel}}} \subseteq \mathrm{MC}_{f^*}$.

Soit ω une configuration, $c = f^*(\omega)$ et d un autre candidat. Supposons que c a une défaite relative contre d et montrons que f^* est manipulable. Comme c a une défaite relative, il n'est pas vainqueur de Condorcet absolu. Et puisqu'il est vainqueur par f^*, aucun autre candidat non plus n'est vainqueur de Condorcet absolu dans ω.

Définissons ψ ainsi : tous les électeurs qui préféraient d à c affirment maintenant préférer d à tous les autres candidats, sans modifier leurs autres préférences.

Alors, à présent, d est vainqueur en uninominal. En effet, on peut raisonner en excluant les électeurs qui placent c et d en tête à égalité : selon la généralisation d'uninominal choisie, ils répartissent leurs voix également entre c et d ou ne leur en donnent aucune. À ces électeurs près, on a $\mathrm{score}(d) = |d \ \mathrm{P}_v \ c| > |c \ \mathrm{P}_v \ d| \geq \mathrm{score}(c)$. Or le score de c n'a pas diminué par rapport au vote sincère, donc $f(\omega) = d$.

De plus, les seuls duels dont le résultat s'est amélioré sont ceux de d, donc aucun autre candidat ne peut être vainqueur de Condorcet absolu. Donc $f^*(\psi) = d$. Par conséquent, f^* est manipulable en ω vers ψ en faveur de d.

2. Nous allons exhiber une configuration où f^{rel} n'est pas manipulable mais où f^* l'est. Considérons la configuration ω suivante.

24	19	19	19	19
a	d_1	d_2	d_3	d_4
c	c	c	a, c	a, c
Autres	Autres	Autres	Autres	Autres

En uninominal, on a $f(\omega) = a$. Par ailleurs, c a une victoire relative contre a (38 voix contre 24) et une victoire absolue contre chaque d_i (81 voix contre 19), donc il est vainqueur de Condorcet relatif mais pas absolu. Par conséquent, $f^*(\omega) = a$ et $f^{\text{rel}}(\omega) = c$.

Montrons que f^{rel} n'est pas manipulable en ω. Il est impossible que le nouveau vainqueur soit vainqueur de Condorcet relatif (car il ne peut améliorer son duel contre c), donc il faut qu'il soit vainqueur en uninominal et qu'il n'y ait plus de vainqueur de Condorcet relatif. Seul a peut devenir (en fait, rester) vainqueur en uninominal. Mais, en cas de manipulation pour a, on ne peut éviter que c reste vainqueur de Condorcet relatif, donc la manipulation échoue.

Reste à montrer que f^* est manipulable en ω. Considérons la configuration suivante ψ, qui est une tentative de manipulation en faveur de c.

24	19	19	19	19
a	**c**	**c**	d_3	d_4
c			a, c	a, c
Autres	Autres	Autres	Autres	Autres

En uninominal, on a $f(\omega) = c$. Par ailleurs, c est toujours vainqueur de Condorcet relatif, donc aucun autre candidat n'est vainqueur de Condorcet absolu et on a $f^*(\omega) = c$. Ainsi, f^* est manipulable en ω vers ψ en faveur de c.

3. Montrons, tout d'abord, que pour $g = f^{\text{ladm}}$ ou $g = f^{\text{adm}}$, on peut exhiber une configuration où g est manipulable, mais où f ne l'est pas. Considérons la configuration ω suivante.

4	3	2
a	b	d
d	a, c	c
c	d	a, b
b		

$D(\omega)$	a	b	c	d
a	–	4	4	7
b	3	–	3	3
c	2	6	–	3
d	2	6	6	–

En uninominal, le candidat a est élu et il est facile de voir que ce n'est pas manipulable. Par ailleurs, comme a est vainqueur de Condorcet relatif, il reste élu quelle que soit la variante g utilisée.

Considérons maintenant la situation ψ suivante, où les 2 derniers électeurs ont changé leur bulletin dans le but de faire gagner c.

4	3	2
a	b	**c**
d	a, c	**d**
c	d	**b**
b		**a**

$D(\omega)$	a	b	c	d
a	–	4	4	7
b	5	–	3	3
c	2	6	–	5
d	2	6	4	–

À présent, c est le seul candidat Condorcet-admissible (chaque autre candidat a une défaite absolue dans sa colonne), donc a $g(\psi) = c$, aussi bien pour $g = f^{!\mathrm{adm}}$ que pour $g = f^{\mathrm{adm}}$.

Par conséquent, f^{adm} et $f^{!\mathrm{adm}}$ sont manipulables en ω vers ψ en faveur de c, alors que f n'est pas manipulable en ω.

Afin de montrer le résultat pour $f^{!\mathrm{faible}}$ et f^{faible}, comme nous allons utiliser des égalités, nous devons préciser la règle de départage[7]. On peut décider, par exemple, qu'en cas d'égalité dans f, les candidats sont départagés par ordre alphabétique ; ou par l'ordre donné par le bulletin d'un électeur choisi à l'avance, quitte à utiliser une règle complémentaire si son bulletin est un ordre faible.

Redéfinissons les configurations ω et ψ précédentes avec 3 électeurs dans chaque groupe. Si on a choisi de départager par un électeur privilégié, on considère le cas où il fait partie du premier groupe, c'est-à-dire les partisans de a. Si on a choisi de départager les candidats par ordre alphabétique, aucune précaution n'est nécessaire, puisque a est favorisé de toute façon. Par le même raisonnement que précédemment, on a $f(\omega) = a$ et ce n'est pas manipulable. Mais $g(\omega) = a$ et $g(\psi) = c$ (car c est le seul vainqueur de Condorcet faible dans ψ) donc g est manipulable en ω. □

Pour l'uninominal, nous avons donc vu qu'on ne peut pas généraliser le théorème faible de condorcification 2.9 aux vainqueurs de Condorcet faibles ou aux candidats Condorcet-admissible. En revanche, cela fonctionne pour le vainqueur de Condorcet relatif, et la condorcification relative est même moins manipulable que la condorcification absolue.

2.5.4 Variantes de condorcification pour U2TI

Comme la condorcification relative fonctionne pour l'uninominal, on peut se demander si l'uninominal à deux tours instantanés (U2TI) est à même de nous fournir un contre-exemple pour une éventuelle généralisation du théorème 2.9 reposant sur la notion de vainqueur de Condorcet relatif. Mais nous allons voir que ce n'est pas le cas : en effet, la condorcification relative fonctionne également pour U2TI.

Proposition 2.13
Le SVBE f considéré est U2TI.

1. *On a :* $\mathrm{MC}_{f^{rel}} \subseteq \mathrm{MC}_{f^*}$.
2. *Il existe (V, C) t.q. l'inclusion précédente est stricte.*
3. *Il existe (V, C) t.q. pour tout système g choisi parmi les variantes $f^{!faible}$, f^{faible}, $f^{!adm}$ ou f^{adm}, on n'a pas* $\mathrm{MC}_g \subseteq \mathrm{MC}_f$.

7. C'est un mal nécessaire ici. En effet, en toute généralité, on veut exhiber une configuration ω où f n'est pas manipulable mais où g est manipulable vers une certaine configuration ψ. Notons $a = f(\omega)$.

Si un candidat $b \neq a$ est vainqueur de Condorcet faible, considérons la configuration ϕ où ses partisans utilisent la stratégie triviale (b en tête) : alors b reçoit au moins autant de voix que a. Mais comme f est supposée non manipulable en ω, ceci nécessite de préciser la règle de départage (favorable à a en ϕ).

À l'inverse, supposons qu'aucun candidat différent de a ne soit vainqueur de Condorcet faible. Alors $g(\omega) = a$. Pour que le contre-exemple fonctionne, il faut donc qu'en notant $c = g(\psi)$, on ait $c \neq f(\psi)$. Donc c est vainqueur de Condorcet faible dans ψ. Comme la manipulation ne peut avoir amélioré son duel contre a, il avait déjà une non-défaite relative contre a dans ω. Considérons à présent la situation χ où les électeurs préférant c à a utilisent la stratégie triviale (c en tête). Alors c a au moins autant de voix que a dans f. Mais comme f est supposée non manipulable en ω, ceci nécessite de préciser la règle de départage (favorable à a en χ).

Démonstration. 1. Supposons qu'il existe une configuration ω où f^* est non manipulable mais où f^{rel} est manipulable vers une certaine configuration ψ. Notons $a = f^*(\omega)$.

Cas 1 On suppose qu'on a $f(\omega) \neq a$.

Par définition de f^*, ceci signifie que a est un vainqueur de Condorcet absolu, donc un vainqueur de Condorcet relatif (proposition 1.30), ce qui entraîne $f^{\text{rel}}(\omega) = f^*(\omega) = a$. Notons $d = f^{\text{rel}}(\psi)$. Il est nécessaire que $d \neq f^*(\psi)$, sinon f^* serait aussi manipulable en ω. Par conséquent, d est vainqueur de Condorcet relatif (mais pas absolu) dans ψ. Or a était vainqueur de Condorcet relatif dans ω et la manipulation ne peut avoir amélioré le résultat du duel de d contre a : c'est absurde ! On pourra d'ailleurs remarquer que ce cas 1 n'utilise aucune propriété particulière d'U2TI : il s'applique de la même façon à tout système de vote f.

Cas 2 On suppose qu'on a $f(\omega) = a$.

Pour montrer qu'on a aussi $f^{\text{rel}}(\omega) = a$, il nous suffit de montrer qu'aucun autre candidat ne peut être vainqueur de Condorcet relatif. Notons b l'adversaire de a au second tour (dans f). On sait qu'il n'est pas vainqueur de Condorcet relatif, sinon il gagnerait le second tour contre a. Supposons qu'un candidat $c \notin \{a, b\}$ soit vainqueur de Condorcet relatif en ω. Considérons alors la situation ϕ où chaque électeur préférant c à a tente de manipuler pour c en utilisant la *stratégie de compromis* (mettre c en première position, sans changer le reste de son bulletin).

- Si on utilise f, le candidat c est sélectionné pour le second tour. En effet, on peut raisonner en excluant les électeurs qui placent a et c en tête à égalité : selon la généralisation d'U2TI choisie, ils répartissent leurs voix également entre a et c ou ne leur en donnent aucune. À ces électeurs près, on a $\text{score}(c) = |c \text{ P}_v \ a| > |a \text{ P}_v \ c| \geq \text{score}(a)$. Or le score de a n'a pas diminué par rapport au vote sincère, donc c est sélectionné pour le second tour.
- Seuls les duels de c ont été améliorés par rapport à la configuration ω, donc c reste vainqueur de Condorcet relatif. Par conséquent, il gagne le second tour donc on a $f(\phi) = c$. De plus, aucun autre candidat ne peut être vainqueur de Condorcet absolu, donc $f^*(\phi) = c$.
- Donc f^* est manipulable en ω vers ϕ en faveur de c, ce qui est exclu.

Ainsi, on a : $a = f(\omega) = f^*(\omega) = f^{\text{rel}}(\omega)$.

Notons $d = f^{\text{rel}}(\psi)$. Il est nécessaire que $d \neq f^*(\psi)$, sinon f^* serait aussi manipulable en ω (cf. cas 1). Par conséquent, d est vainqueur de Condorcet relatif (mais pas absolu) en ψ. On en déduit que d possédait une victoire relative contre a dans ω : en effet, la manipulation ne peut avoir amélioré le résultat de ce duel.

Considérons, à présent, la configuration χ où les électeurs qui préfèrent d à a utilisent la stratégie de compromis (mettre d en tête, sans changer le reste de leur bulletin). Par le même raisonnement que ci-dessus, on a $\text{score}(d) = |d \text{ P}_v \ a| > |a \text{ P}_v \ d| \geq \text{score}(a)$, donc d est sélectionné pour le second tour. Or, dans χ, le candidat d apparaît comme un vainqueur de Condorcet relatif : en effet, tous ses duels sont au moins aussi bons que dans ψ. Donc il gagne le second tour et on a $f(\chi) = d$. Comme d est vainqueur de Condorcet relatif, aucun autre candidat ne peut être vainqueur de Condorcet absolu et on a $f^*(\chi) = d$. Donc f^* est manipulable en ω vers χ en faveur de d : c'est absurde !

2. Considérons la configuration ω suivante.

23	20	19	19	19
a	d_1	d_2	d_3	d_4
c	c	c	a, c	a, c
	a	a		
Autres	Autres	Autres	Autres	Autres

Au second tour, a affronte d_1 et gagne (80 voix contre 20) : on a $f(\omega) = a$.

Le candidat c a une victoire relative contre a (39 voix contre 23) et une victoire absolue contre tout candidat d_i (par 80 ou 81 voix). Donc c est vainqueur de Condorcet relatif, mais pas absolu. On a donc $f^*(\omega) = a$ et $f^{\text{rel}}(\omega) = c$.

Montrons que f^{rel} n'est pas manipulable en ω.

— En faveur de a : le candidat c reste vainqueur de Condorcet relatif, donc la manipulation échoue.

— En faveur d'un d_i : il est facile de voir qu'aucun d'entre eux ne peut gagner en U2TI. En effet, même en manipulant, seul d_1 peut être sélectionné pour le second tour, mais il perd alors contre a. Par ailleurs, aucun d_i ne peut devenir vainqueur de Condorcet relatif, car il aura toujours une défaite relative (et même absolue) contre c.

À présent, considérons la configuration ψ suivante, qui est une tentative pour manipuler f^* en faveur de c.

23	20	19	19	19
a	$\mathbf{c}$	$\mathbf{c}$	d_3	d_4
c	$\mathbf{d_1}$	$\mathbf{d_2}$	a, c	a, c
	a	a		
Autres	Autres	Autres	Autres	Autres

En uninominal à deux tours, il est facile de voir qu'on a $f(\psi) = c$. De plus, comme les duels de c n'ont été qu'améliorés, il reste vainqueur de Condorcet relatif et aucun autre candidat n'est vainqueur de Condorcet absolu. Donc $f^*(\psi) = c$.

Ainsi, f^* est manipulable en ω vers ψ en faveur de c, alors que f^{rel} n'est pas manipulable en ω.

3. On peut utiliser le même exemple que pour le point 3 de la proposition 2.11. Pour $f^{!\text{faible}}$ ou f^{faible}, là encore, on considère 3 électeurs par groupe et il est nécessaire de préciser la règle de départage, par exemple alphabétique sur les candidats. □

Ainsi, si f est l'uninominal à deux tours instantané (U2TI), on a le même résultat que pour l'uninominal : le système de vote f^{rel} est moins manipulable que f, et même que la condorcification f^*.

2.5.5 Condorcification relative pour VTI

À ce stade, on pourrait presque se demander si l'échec de la condorcification relative n'est pas limitée à des modes de scrutin exotiques comme Condorcet-doyen, puisqu'elle fonctionne pour l'uninominal et U2TI. Mais VTI va nous arrêter dans notre élan : dans ce cas, la condorcification relative n'est pas moins manipulable que le mode de scrutin original.

Proposition 2.14
Le SVBE f considéré est VTI.
Il existe (V, C) t.q. on n'a pas $\mathrm{MC}_{f^{rel}} \subseteq \mathrm{MC}_f$.

Démonstration. Considérons la configuration ω suivante.

12	11	25	12	12	28
a	a	b	b	c	d
b	d	a, c	c	a	a, c
c	c	d	a	d	b
d	b		d	b	

$D(\omega)$	a	b	c	d
a	–	63	23	72
b	37	–	49	49
c	24	51	–	61
d	28	51	39	–

En VTI, les candidats c, d puis b sont successivement éliminés donc $f(\omega) = a$. Montrons tout d'abord que f n'est pas manipulable en ω.

— En faveur de b : même si b parvenait au dernier tour, il perdrait contre tout autre candidat (par 63 ou 51 voix).
— En faveur de c : les quatrième et cinquième groupes sont intéressés ($12+12$ électeurs). Pour que c ne soit pas éliminé au premier tour, il faut que 23 ou 24 manipulateurs mettent c en tête de bulletin et que a soit éliminé (il n'est pas possible d'éliminer b ou d). Au deuxième tour, comme b et d ont chacun plus du tiers des voix (37 et 39 respectivement), c est éliminé.
— En faveur de d : seul le sixième groupe (28 électeurs) est intéressé. Au premier tour, il ne peuvent à la fois sauver c et d : en effet, comme a reçoit 23 voix, il faudrait que les manipulateurs et les électeurs qui votent sincèrement pour c et d totalisent au moins $2 \times 23 = 46$ voix, mais ils n'en ont que $12 + 28 = 40$. Comme d doit rester en lice, c'est donc c qui doit être éliminé au premier tour. Au second tour, a et b ont chacun plus du tiers des voix (35 et 37 respectivement), donc d est éliminé.

Comme c est vainqueur de Condorcet relatif, on a $f^{\mathrm{rel}}(\omega) = c$. Considérons la configuration ψ suivante, qui est une tentative de manipuler en faveur de a.

12	11	25	12	12	28
a	a	b	b	c	d
b	d	a, c	c	a	a, c
d	**b**	d	a	d	b
c	**c**		d	b	

$D(\omega)$	a	b	c	d
a	–	63	23	72
b	37	–	60	49
c	24	40	–	49
d	28	51	51	–

Du point de vue de VTI, le dépouillement se déroule de la même façon et on a $f(\psi) = a$. Et comme les autres candidats ne sont même pas Condorcet-admissibles (ils ont tous une défaite dans leur colonne), on a aussi $f^{\mathrm{rel}}(\psi) = a$. Donc f^{rel} est manipulable en ω vers ψ en faveur de a. □

2.5.6 Condorcification relative pour les systèmes cardinaux

Nous allons, à présent, voir que le résultat que nous venons de montrer pour VTI est également valable pour le vote par assentiment, le vote par notation et le jugement majoritaire : pour ces systèmes, la condorcification relative ne permet pas de diminuer la manipulabilité. Dans les trois cas, les notes extrêmes autorisées seront 0 et 1 par convention.

Dans une configuration où toutes les notes sont 0 et 1 (même si d'autres notes sont autorisées), on remarquera que ces trois modes de scrutin renvoient le même vainqueur, c'est-à-dire le candidat qui a la note 1 dans le plus grand nombre de

bulletins (voir section 1.6.1). Ceci va nous permettre de traiter ces trois systèmes en une seule démonstration.

Proposition 2.15
Le SVBE f considéré est le vote par assentiment, le vote par notation ou le jugement majoritaire.
Il existe (V, C) t.q. on n'a pas $\mathrm{MC}_{f^{rel}} \subseteq \mathrm{MC}_f$.

Démonstration. Considérons la configuration ω suivante.

2	3	4
$a:1$	$c:1$	$b:1$
$c:0$	$a:1$	$a,c:0$
$b:0$	$b:0$	

$D(\omega)$	a	b	c
a	–	5	2
b	4	–	4
c	3	5	–

On a $\mathrm{score}(a) = 5$, $\mathrm{score}(b) = 4$ et $\mathrm{score}(c) = 3$, donc $f(\omega) = a$. En cas de tentative de manipulation pour c, les électeurs intéressés (second groupe) ne peuvent faire mieux que descendre la note de a à 0, mais dans ce cas c'est b qui gagne. Quant à une manipulation pour b, les électeurs intéressés (troisième groupe) ne peuvent faire mieux que leur bulletin actuel. Donc f n'est pas manipulable en ω.

Le candidat c est vainqueur de Condorcet relatif, donc $f^{\mathrm{rel}}(\omega) = c$. Considérons maintenant la configuration ψ suivante, où le premier groupe d'électeurs tente de faire gagner a.

2	3	4
$a:1$	$c:1$	$b:1$
$\mathbf{b}:0$	$a:1$	$a,c:0$
$\mathbf{c}:0$	$b:0$	

$D(\omega)$	a	b	c
a	–	5	2
b	4	–	6
c	3	3	–

On a $f(\psi) = a$ et les autres candidats ne sont même pas Condorcet-admissibles (ils ont tous une défaite dans leur colonne), donc $f^{\mathrm{rel}}(\omega) = a$. Donc f^{rel} est manipulable en ω vers ψ en faveur de a, alors que f n'est pas manipulable en ω. □

2.6 Vainqueur de Condorcet résistant

Le théorème faible de condorcification 2.9 énonce que la condorcification f^* est au plus aussi manipulable que le système original f (si celui-ci vérifie **CMInf**). Dans le théorème fort de condorcification 2.20, nous montrerons qu'en fait, cette diminution de manipulabilité est stricte pour de nombreux systèmes de vote. Afin de nous donner les outils pour le démontrer, nous allons à présent définir la notion originale de *vainqueur de Condorcet résistant* et nous allons prouver qu'un tel candidat est caractérisé par une forme d'immunité à la manipulation.

Voici l'intuition amenant à la notion de *vainqueur de Condorcet résistant*. Si f vérifie le critère de Condorcet et s'il y a un vainqueur de Condorcet c, alors les manipulateurs en faveur de d ont besoin d'empêcher c d'apparaître comme un vainqueur de Condorcet. Par conséquent, ils doivent empêcher une victoire stricte de c contre un certain candidat $e \neq c$. Cependant, ce plan est voué à l'échec si les électeurs sincères (ceux qui ne préfèrent pas d à c) assurent déjà : (1) une victoire de c contre e ; et (2) une non-victoire de e contre c. Cette observation conduit à la définition suivante.

Définition 2.16 (vainqueur de Condorcet résistant)
Soit $\omega \in \Omega$ et $c \in \mathcal{C}$.
On dit que c est *vainqueur de Condorcet résistant* en ω ssi $\forall (d, e) \in (\mathcal{C} \setminus \{c\})^2$:

$$\begin{cases} |\text{non}(d \, \mathrm{P}_v \, c) \text{ et } c \, \mathrm{P}_v \, e| > \dfrac{V}{2}, & (2.1) \\ |\text{non}(d \, \mathrm{P}_v \, c) \text{ et non}(e \, \mathrm{P}_v \, c)| \geq \dfrac{V}{2}, & (2.2) \end{cases}$$

Nous dirons qu'une configuration ω est *résistante* ssi elle possède un vainqueur de Condorcet résistant.

Si toutes les relations binaires P_v sont antisymétriques (ce qui est une hypothèse courante), alors la condition (2.2) devient redondante. Dans ce cadre, c est vainqueur de Condorcet résistant ssi $\forall (d, e) \in (\mathcal{C} \setminus \{c\})^2$:

$$|\text{non}(d \, \mathrm{P}_v \, c) \text{ et } c \, \mathrm{P}_v \, e| > \frac{V}{2}.$$

Si, en outre, toutes les relations binaires P_v sont complètes, alors la définition est encore plus simple puisque la relation ci-dessus devient symétrique par rapport à d et e :

$$|c \, \mathrm{P}_v \, d \text{ et } c \, \mathrm{P}_v \, e| > \frac{V}{2}.$$

Autrement dit, pour toute paire d'autres candidats (d, e), il y a une majorité stricte d'électeurs qui préfèrent simultanément c à d et c à e.

Clairement, un vainqueur de Condorcet résistant est également vainqueur de Condorcet. Par conséquent, s'il existe un vainqueur de Condorcet résistant, il est unique.

Ainsi, nous avons choisi de définir le *vainqueur de Condorcet résistant* en terme de préférences des électeurs dans la configuration ω. Le théorème suivant reprend formellement la propriété d'immunité à la manipulation que nous avons mentionnée pour introduire la notion et montre que cette propriété est caractéristique (c'est-à-dire qu'elle n'est vérifiée que par un vainqueur de Condorcet résistant), sous l'hypothèse courante où l'espace électoral comprend tous les ordres stricts totaux.

Théorème 2.17 (caractérisation du vainqueur de Condorcet résistant)
Soit $\omega \in \Omega$ et $c \in \mathcal{C}$. On considère les conditions suivantes.

1. *Le candidat c est vainqueur de Condorcet résistant en ω.*
2. *Pour tout SVBE f respectant le critère de Condorcet, c est élu par le vote sincère et f n'est pas manipulable en ω.*

On a l'implication 1 $\Rightarrow$ 2.
Si l'espace électoral comprend tous les ordres stricts totaux (ce qui est une hypothèse courante, cf. définition 1.10), alors la réciproque 2 $\Rightarrow$ 1 est vraie.

Démonstration. 1 $\Rightarrow$ 2 : puisque c est vainqueur de Condorcet résistant, même après une tentative de manipulation en faveur d'un autre candidat d, les électeurs sincères assurent que c a encore une victoire stricte contre tout candidat $e \neq c$; par conséquent, c apparaît toujours comme un vainqueur de Condorcet et il est élu.

À présent, nous allons montrer la réciproque 2 $\Rightarrow$ 1. Supposons que c n'est pas vainqueur de Condorcet résistant. Si c n'est pas vainqueur de Condorcet, c'est

immédiat : on a nécessairement $C \geq 2$ (car il existe un candidat $d \neq c$ tel que c n'ait pas une victoire stricte contre d) et il existe donc un système Condorcet où c n'est pas vainqueur du vote sincère. Nous pouvons donc nous concentrer sur le cas où c est vainqueur de Condorcet (mais pas résistant). Nous allons alors prouver qu'il existe un SVBE f respectant le critère de Condorcet et manipulable en ω.

Puisque c n'est pas vainqueur de Condorcet résistant, au moins l'une des conditions (2.1) ou (2.2) de la définition 2.16 n'est pas vérifiée. Nous distinguons trois cas : la condition (2.1) n'est pas vérifiée pour un certain $e = d$; la condition (2.2) n'est pas vérifiée; ou la condition (2.1) n'est pas vérifiée avec $e \neq d$.

Dans chacun des trois cas, le principe de la preuve est le même : on exhibe une configuration ψ qui n'a pas de vainqueur de Condorcet et qui diffère de ω seulement par des électeurs qui préfèrent d à c. Par conséquent, il est possible de choisir un SVBE f qui vérifie le critère de Condorcet et tel que $f(\psi) = d$. On en déduit que f est manipulable en ω vers ψ en faveur du candidat d.

Cas 1 Si la condition (2.1) n'est pas vérifiée pour un certain $e = d$, cela signifie que $|\text{non}(d \ \mathrm{P}_v \ c)$ et $c \ \mathrm{P}_v \ d| \leq \frac{V}{2}$. Soit p_0 un ordre strict total de la forme : $(d \succ c \succ$ autres candidats). Puisque l'espace électoral comprend tous les ordres stricts totaux, pour chaque manipulateur $v \in \text{Manip}(c \to d)$, on peut choisir un bulletin ψ_v tel que $\mathrm{P}_v(\psi_v) = p_0$. Pour chaque électeur sincère $v \in \text{Sinc}(c \to d)$, on pose $\psi_v = \omega_v$. Dans la nouvelle configuration ψ, le candidat c ne peut apparaître comme un vainqueur de Condorcet car $|c \ \mathrm{P}_v(\psi_v) \ d| = |\text{non}(d \ \mathrm{P}_v(\omega_v) \ c)$ et $c \ \mathrm{P}_v(\omega_v) \ d| \leq \frac{V}{2}$. Le candidat d ne peut apparaître comme un vainqueur de Condorcet (lemme 2.4) et aucun autre candidat non plus, car le nombre d'électeurs qui affirment leur préférer c n'a pas diminué.

Cas 2 Si la condition (2.2) n'est pas vérifiée pour un certain d et un certain e, cela signifie que $|\text{non}(d \ \mathrm{P}_v \ c)$ et $\text{non}(e \ \mathrm{P}_v \ c)| < \frac{V}{2}$. Remarquons que $e \neq d$, car sinon c ne serait pas vainqueur de Condorcet. Quitte à échanger les rôles entre d et e, on peut supposer que e n'a pas de victoire stricte contre d. Soit p_0 un ordre strict total de la forme : $(d \succ e \succ c \succ$ autres candidats). Pour chaque manipulateur $v \in \text{Manip}(c \to d)$, on peut choisir un bulletin ψ_v tel que $\mathrm{P}_v(\psi_v) = p_0$. Pour chaque électeur sincère $v \in \text{Sinc}(c \to d)$, on pose $\psi_v = \omega_v$. Dans la nouvelle configuration ψ, le candidat c ne peut apparaître comme vainqueur de Condorcet car il a une défaite contre e : en effet, $|\text{non}(e \ \mathrm{P}_v(\psi_v) \ c)| = |\text{non}(d \ \mathrm{P}_v(\omega_v) \ c)$ et $\text{non}(e \ \mathrm{P}_v(\omega_v) \ c)| < \frac{V}{2}$. Le candidat d ne peut apparaître comme vainqueur de Condorcet (lemme 2.4), le candidat e non plus parce qu'il n'a toujours pas de victoire stricte contre d et aucun autre candidat non plus, car le nombre d'électeurs qui affirment leur préférer c n'a pas diminué.

Cas 3 Reste le cas où la condition (2.1) n'est pas vérifiée, avec $e \neq d$. En notant $B_{de} = |\text{non}(d \ \mathrm{P}_v \ c)$ et $c \ \mathrm{P}_v \ e|$, cela signifie que $B_{de} \leq \frac{V}{2}$. En utilisant le cas précédent, on peut supposer, cependant, que la condition (2.2) est vérifiée.

Nous allons voir que dans la configuration finale ψ, on peut assurer qu'il n'y ait ni une victoire de c contre e, ni une victoire de e contre c.

Soit p_0 un ordre strict total de la forme : $(d \succ e \succ c \succ$ autres candidats).

Soit p_0' un ordre strict total de la forme : $(d \succ c \succ e \succ$ autres candidats).

Comme c est un vainqueur de Condorcet, on a $|c \ \mathrm{P}_v \ e| > \frac{V}{2}$ donc :

$$|d \ \mathrm{P}_v \ c \text{ et } c \ \mathrm{P}_v \ e| > \frac{V}{2} - B_{de} \geq 0.$$

Par conséquent, on peut choisir $\lfloor \frac{V}{2} \rfloor - B_{de}$ électeurs parmi les manipulateurs (les électeurs qui préfèrent d à c) ; pour chacun d'entre eux, noté v, choisissons ψ_v tel que $\mathrm{P}_v(\psi_v) = p'_0$. Pour chaque autre manipulateur v, choisissons ψ_v tel que $\mathrm{P}_v(\psi_v) = p_0$. Enfin, pour chaque électeur sincère $v \in \mathrm{Sinc}(c \to d)$, notons $\psi_v = \omega_v$.

On a alors :

$$D_{ce}(\psi) = B_{de} + \left(\left\lfloor \frac{V}{2} \right\rfloor - B_{de} \right) = \left\lfloor \frac{V}{2} \right\rfloor,$$

donc c n'a pas de victoire contre e.

Par ailleurs, la condition (2.1) n'est pas vérifiée pour cette paire (d, e) mais la condition (2.2) est vérifiée. Donc on a :

$$\begin{cases} |\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{PP}_v\ e| + |\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{PM}_v\ e| \leq \left\lfloor \frac{V}{2} \right\rfloor, \\ |\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{PP}_v\ e| + |\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{I}_v\ e| \quad \geq \left\lceil \frac{V}{2} \right\rceil, \end{cases}$$

donc, par soustraction :

$$|\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{PM}_v\ e| - |\mathrm{non}(d\ \mathrm{P}_v\ c) \text{ et } c\ \mathrm{I}_v\ e| \leq \left\lfloor \frac{V}{2} \right\rfloor - \left\lceil \frac{V}{2} \right\rceil.$$

En utilisant la proposition 1.25, on en déduit :

$$\begin{aligned} D_{ec}(\psi) &= V + |c\ \mathrm{PM}_v(\psi_v)\ e| - |c\ \mathrm{I}_v(\psi_v)\ e| - D_{ce}(\psi), \\ &\leq V + \left\lfloor \frac{V}{2} \right\rfloor - \left\lceil \frac{V}{2} \right\rceil - \left\lfloor \frac{V}{2} \right\rfloor = \left\lfloor \frac{V}{2} \right\rfloor, \end{aligned}$$

donc e n'a pas de victoire contre c.

En résumé, ni c ni e ne peut être vainqueur de Condorcet. Pour les mêmes raisons que dans les cas précédents, d et les autres candidats non plus. □

Dans la condition 2, il est nécessaire d'exiger que tous les systèmes de vote Condorcet f aient le même résultat c ou, de façon équivalente, que c est vainqueur de Condorcet. Sinon, la réciproque 2 ⇒ 1 n'est pas vraie. En effet, considérons une configuration où chaque électeur est indifférent entre tous les candidats : alors tout système de vote est non manipulable, mais il n'y a pas de vainqueur de Condorcet résistant.

Par ailleurs, si on suppose que les préférences sont complètes et antisymétriques (en particulier si ce sont des ordres stricts totaux), alors la démonstration est simplifiée : les conditions (2.1) et (2.2) de la définition 2.16 sont équivalentes, donc la démonstration se limite au cas 2.

Définition 2.18 (critère de Condorcet-résistant [8])

On dit que f vérifie le *critère de Condorcet-résistant* (**rCond**) ssi, pour toute configuration $\omega \in \Omega$ et tout candidat $c \in \mathcal{C}$, si c est vainqueur de Condorcet résistant en ω, alors $f(\omega) = c$.

Il est clair que vérifier le critère de Condorcet implique de vérifier le critère de Condorcet-résistant.

8. Nous écrivons cette expression avec un trait d'union pour insister sur le fait que l'adjectif *résistant* s'applique à la notion de vainqueur de Condorcet et non au substantif *critère*.

Proposition 2.19

On considère un espace électoral qui comprend tous les ordres stricts totaux. Pour chaque système de vote dans la liste suivante, il existe des valeurs de V et C pour lesquelles le système ne vérifie pas le critère de Condorcet-résistant : BI, le jugement majoritaire, l'uninominal, U2TI, Veto, VTI, VTIM, le vote par assentiment, le vote par notation et les méthodes de Borda, Bucklin, Coombs et Kim-Roush.

Démonstration. Nous allons donner un seul contre-exemple qui permet de couvrir tous ces systèmes de vote sauf BI.

17	13	14	14	14	14	14
a	c	d_1	d_2	d_4	d_7	d_{11}
	a	c	d_3	d_5	d_8	d_{12}
		a	c	d_6	d_9	d_{13}
			a	c	d_{10}	d_{14}
				a	c	d_{15}
					a	c
Autres	Autres	Autres	Autres	Autres	Autres	Autres
c	d_1	d_2	d_4	d_7	d_{11}	a

Le candidat c est préféré à toute paire de candidats (a, d_i) par 69 électeurs (sur 100) et à toute paire de candidats (d_i, d_j) par au moins 55 électeurs, donc c est vainqueur de Condorcet résistant.

En jugement majoritaire, en vote par assentiment ou en vote par notation, il suffit de considérer le cas où le vote sincère de chaque électeur consiste à attribuer la note maximale à son candidat favori et la note minimale aux autres candidats. Alors c'est a qui est élu (comme en uninominal).

Dans une RPS de vecteur de poids $\mathbf{x}$, on a $\text{score}(a) - \text{score}(c) = 3(x_1 - x_C) + 1(x_1 - x_2) > 0$, donc c ne peut pas être élu. En particulier, c'est le cas pour l'uninominal, Veto et la méthode de Borda.

En U2TI, en VTI ou en VTIM, c est éliminé au premier tour.

En méthode de Bucklin, on a $\text{score}(a) = (4\,;\,58)$ et $\text{score}(c) = (4\,;\,55)$ donc c ne peut être élu.

En méthode de Coombs ou de Kim-Roush, c est éliminé au premier tour.

Dans le chapitre 3, nous verrons des résultats plus détaillés pour chacun de ces systèmes de vote en fonction du nombre de candidats. Nous prouverons également qu'en général, BI ne vérifie pas **rCond**, que la seule RPSI-ES qui vérifie **rCond** est la méthode de Baldwin, et que la seule RPSI-EM qui vérifie **rCond** est la méthode de Nanson. □

À notre connaissance, il n'y a pas de système de vote classique (dans la littérature ou les applications pratiques) qui vérifie le critère de Condorcet-résistant mais pas le critère de Condorcet. Il est facile de définir des contre-exemples artificiels, comme le système *Résistant-Doyen* (RDoy.) : on élit le vainqueur de Condorcet résistant quand il existe, et un candidat fixé à l'avance dans le cas contraire. Mais cette observation tend à montrer qu'un système de vote qui n'a pas été conçu pour élire tous les vainqueurs de Condorcet n'a généralement aucune raison « naturelle » d'élire ceux qui sont résistants.

2.7 Théorème fort de condorcification

Grâce à la notion de vainqueur de Condorcet résistant, nous disposons de l'outil nécessaire pour montrer que, pour de nombreux systèmes de vote, la condorcification f^* est *strictement* moins manipulable que le système de vote original f.

Théorème 2.20 (théorème fort de condorcification)
Soit Ω un espace électoral et f un SVBE. On suppose que f vérifie **CMInf** *mais pas* **rCond**.
Alors sa condorcification f^* est strictement moins manipulable que f :

$$\mathrm{MC}_{f^*} \subsetneq \mathrm{MC}_f.$$

Démonstration. Le théorème de condorcification 2.9 garantit l'inclusion.

Puisque f ne vérifie pas **rCond**, il existe une configuration $\omega \in \Omega$ et un candidat $c \in \mathcal{C}$, vainqueur de Condorcet résistant dans ω, tel que $f(\omega) \neq c$. Par la proposition 1.31, on sait que $f(\omega)$ n'est pas Condorcet-admissible en ω, donc le lemme 2.7 assure que f est manipulable en ω. Mais le théorème 2.17 assure que f^* n'est pas manipulable en ω. Par conséquent, l'inclusion est stricte. □

Corollaire 2.21
On considère un espace électoral qui comprend tous les ordres stricts totaux. Pour chaque système de vote f dans la liste suivante, il existe des valeurs de V et C telles que la condorcification f^* est strictement moins manipulable que f : BI, le jugement majoritaire, l'uninominal, U2TI, VTI, VTIM, le vote par assentiment, le vote par notation et les méthodes de Borda, Bucklin et Coombs.

Démonstration. Comme ces systèmes vérifient **CMInf** (proposition 2.3) mais pas **rCond** (proposition 2.19), on peut leur appliquer le théorème fort de condorcification 2.20. □

2.8 Borne supérieure de manipulabilité pour les systèmes Condorcet

Puisque tout SVBE Condorcet est non-manipulable dans les configurations résistantes, ceci nous fournit une borne supérieure de manipulabilité. Plus précisément, notons τ^{π}_{CR} la probabilité qu'une configuration soit résistante dans la culture π. Alors, pour tout SVBE f vérifiant le critère de Condorcet, on a :

$$\tau^{\pi}_{\mathrm{MC}}(f) \leq 1 - \tau^{\pi}_{\mathrm{CR}}.$$

On peut donc se demander si cette borne supérieure est ajustée, c'est-à-dire s'il existe un système de vote Condorcet qui est manipulable dans toutes les configurations non résistantes. Nous allons voir que, dans une importante classe d'espaces électoraux, un tel système existe.

Considérons l'espace électoral des ordres stricts totaux pour $C \geq 6$ et V quelconque. Le système de vote f est défini de la façon suivante. Au préalable, on choisit une bijection arbitraire entre les permutations de $C-3$ éléments et les entiers de l'intervalle $[\![1, (C-3)![\![$. Puis les électeurs fournissent leurs bulletin et on calcule le résultat de la façon suivante.

— S'il y a un vainqueur de Condorcet, alors il est élu.

— Sinon, pour chaque électeur v, on note x_v l'entier associé à son ordre de préférence sur les $C-3$ derniers candidats de son bulletin. On déclare vainqueur le candidat $\sum x_v$, en comptant modulo C pour le ramener dans l'intervalle $[\![1, C]\!]$.

Considérons une configuration non résistante ω et montrons qu'elle est manipulable. Soit $\mathrm{w} = f(\omega)$ le vainqueur sincère.

Cas 1 Si w est vainqueur de Condorcet mais pas vainqueur de Condorcet résistant, alors il existe deux candidats distincts d et e tels que les électeurs préférant simultanément w à d et e ne forment pas une majorité stricte (en effet, w ne peut pas violer la définition de vainqueur de Condorcet résistant pour $d = e$, sinon il ne serait pas vainqueur de Condorcet). Quitte à échanger les rôles de d et e, on peut supposer que d a une victoire contre e. Nous allons construire une manipulation en faveur de d.

Tous les manipulateurs (les électeurs préférant d à c) mettent d en tête de bulletin, puis e, puis c. Ceci assure qu'aucun candidat n'est vainqueur de Condorcet (comme dans le cas 2 de la preuve du théorème 2.17, on utilise le fait que e n'a pas de victoire contre d). Tous les manipulateurs, sauf un, mettent les autres candidats en bas de leur bulletin dans un ordre arbitraire. Pour le dernier manipulateur v, en choisissant de façon adéquate l'ordre de ses trois derniers candidats, il peut rendre sa valeur x_v égale à n'importe quel entier de l'intervalle $[\![1, (C-3)!]\!]$ et puisque $(C-3)! \geq C$, il peut choisir n'importe quel vainqueur, en particulier d.

Cas 2 Si w est Condorcet-admissible mais pas vainqueur de Condorcet, le processus est similaire. Il existe un candidat c tel que $|c \ \mathrm{P}_v \ \mathrm{w}| = \frac{V}{2}$. On construit une manipulation en faveur de c. Les manipulateurs placent c en tête, ce qui assure qu'aucun autre candidat n'est vainqueur de Condorcet (et c lui-même non plus puisqu'il n'a toujours pas de victoire contre w). Puis tous les manipulateurs choisissent de façon arbitraire l'ordre dans lequel ils placent les autres candidats, sauf le dernier manipulateur v dont le choix de x_v permet de choisir n'importe quel vainqueur, en particulier c.

Cas 3 Enfin, si w n'est pas Condorcet-admissible, alors on sait que ω est manipulable (lemme 2.7).

Par conséquent, le système de vote f est aussi manipulable que possible pour un système de Condorcet : en effet, il est non-manipulable seulement dans les configurations résistantes ! Ainsi, dans toute culture π, il atteint la borne supérieure de manipulabilité que nous avons donnée : $\tau^{\pi}_{\mathrm{MC}}(f) = 1 - \tau_{\mathrm{CR}}$.

Nous verrons cependant qu'exhiber un système de vote Condorcet qui est manipulable dans toutes les configurations non-résistantes n'est pas possible pour toutes les valeurs de (V, C) : dans l'espace électoral des ordres stricts totaux avec $V = 3$ électeurs et $C = 3$ candidats, la section 10.2.1 révèlera qu'un tel système n'existe pas.

2.9 Condorcification et systèmes optimaux

Jusqu'à présent, nous considérions généralement un système de vote f donné et nous comparions l'ensemble des configurations manipulables pour f et pour sa condorcification f^*. Nous avons montré que, sous certaines hypothèses, f^* était au plus aussi manipulable que f (théorème faible de condorcification 2.9), voire strictement moins manipulable (théorème fort de condorcification 2.20). Ces

résultats pourraient suggérer d'utiliser des systèmes comme la condorcification de l'uninominal, de VTI, etc.

Si on ne se concentre pas sur un système de vote particulier, les théorèmes de condorcification ont également des conséquences profondes qui concernent tout planificateur social qui souhaiterait trouver un système de vote acceptable dont la manipulabilité soit minimale.

Corollaire 2.22

Considérons la fonction :

$$\mathrm{MC} : \left| \begin{array}{ccc} \textbf{CMInf} & \to & \mathcal{P}(\Omega) \\ f & \to & \mathrm{MC}_f, \end{array} \right.$$

qui renvoie, pour chaque SVBE f vérifiant **CMInf***, l'ensemble de ses configurations manipulables.*

Soit $A \in \mathcal{P}(\Omega)$ une valeur minimale de MC *(s'il en existe), c'est-à-dire un sous-ensemble de Ω tel qu'au moins un système $f \in$* **CMInf** *vérifie $\mathrm{MC}_f = A$, mais aucun système $f \in$* **CMInf** *ne vérifie $\mathrm{MC}_f \subsetneq A$. Alors :*

— *Tout système $f \in$* **CMInf** *vérifiant $\mathrm{MC}_f = A$ vérifie* **rCond** *;*

— *Il existe $f \in$* **Cond** *t.q. $\mathrm{MC}_f = A$.*

Pour bien comprendre la portée de ce corollaire, il faut remarquer que la fonction MC peut avoir plusieurs minima qui sont non comparables, puisque la relation d'inclusion sur $\mathcal{P}(\Omega)$ n'est pas un ordre total. Il peut ainsi exister deux systèmes f et g tels qu'aucun système n'est moins manipulable que f ou g, mais dont les ensembles de configurations manipulables sont non inclus l'un dans l'autre.

Le corollaire 2.22 peut être résumé de la façon suivante : si on cherche un système de vote vérifiant **CMInf** et dont la manipulabilité est minimale (au sens ensembliste), alors l'étude *doit* être restreinte à **rCond** et *peut* être restreinte à **Cond**.

De même, les théorèmes de condorcification entraînent le corollaire suivant.

Corollaire 2.23

Pour une culture donnée π, considérons la fonction :

$$\tau^{\pi}_{\mathrm{MC}} : \left| \begin{array}{ccc} \textbf{CMInf} & \to & [0,1] \\ f & \to & \tau^{\pi}_{\mathrm{MC}}(f). \end{array} \right.$$

Soit τ_0 la borne inférieure de τ^{π}_{MC}. Si elle est atteinte, alors il existe $f \in$ **Cond** *t.q. $\tau^{\pi}_{\mathrm{MC}}(f) = \tau_0$.*

Contrairement à ce qui se passe dans le corollaire 2.22, il n'est *a priori* pas nécessaire qu'un système de manipulabilité minimale (au sens probabiliste) vérifie **rCond** : en effet, si certaines configurations résistantes sont de mesure nulle, alors elles peuvent être manipulables sans altérer le taux de manipulabilité. Cependant, pour tout système de taux de manipulabilité minimal (au sein de **CMInf**), on peut le condorcifier pour obtenir un optimum qui vérifie **Cond** et *a fortiori* **rCond**.

Le corollaire 2.23 peut être résumé de la façon suivante : si on cherche un système de vote vérifiant **CMInf** et dont le taux de manipulabilité soit minimal, alors l'étude peut être restreinte à **Cond**.

Dans les deux corollaires 2.22 et 2.23, on peut se demander si un optimum tel que celui mentionné existe. Si on considère un espace électoral fini (en particulier un espace électoral ordinal, comme celui des ordres stricts totaux ou celui des

ordres stricts faibles), alors il existe un nombre fini de SVBE possibles, *a fortiori* si on exige qu'ils vérifient **CMInf**, donc il existe un optimum.

À l'inverse, si on considère un espace électoral infini (donc nécessairement non limité à l'aspect ordinal, comme l'espace électoral de référence), alors l'existence d'un optimum n'est pas garantie *a priori*. Dans le chapitre 5 consacré à l'étude des systèmes non ordinaux, nous nous intéresserons à cette question pour le taux de manipulabilité : nous donnerons une condition suffisante pour qu'il existe un SVBE qui (au sein de la classe **CMInf**) minimise le taux de manipulabilité dans une culture donnée.

Chapitre 3

Critères majoritaires

Dans les chapitres précédents, nous avons rappelé le critère de Condorcet (**Cond**) puis nous avons introduit le critère de la coalition majoritaire informée (**CMInf**) et le critère de Condorcet-résistant (**rCond**). Nous allons maintenant voir comment ces propriétés s'inscrivent dans une famille plus large de critères majoritaires et développer des liens avec des concepts de théorie des jeux comme l'ensemble des équilibres de Nash forts (ENF) et la capacité à les atteindre.

Dans la section 3.1, nous définissons d'autres critères liés à la notion de majorité. En théorie du vote, il est classique de considérer le critère majoritaire, dont nous rappelons la définition et que nous appelons *critère du favori majoritaire* (**FavMaj**) afin de le distinguer des autres critères majoritaires. Nous introduisons également le *critère de la coalition majoritaire ignorante* et le *critère du bulletin majoritaire*. La motivation initiale de ces définitions est simplement pratique : il s'agit de se doter de critères faciles à tester qui permettent de prouver qu'un système de vote vérifie **CMInf**.

En effet, nous montrons dans la section 3.2 que tous les autres critères évoqués impliquent **CMInf**. Nous montrons, en outre, qu'ils forment une chaîne d'implications [1], du critère le plus fort (**Cond**) au critère le plus faible (**CMInf**).

En introduction, nous avons déjà évoqué les équilibres de Nash forts (ENF) et discuté l'intérêt d'une telle notion [2]. En particulier, la non-manipulabilité d'une configuration signifie exactement que le vote sincère est un ENF pour les préférences correspondantes. Brill et Conitzer (2015) ont montré que, pour un système de vote vérifiant **CMInf**, s'il existe un vainqueur de Condorcet pour les préférences sincères des électeurs, alors il est le seul à pouvoir être vainqueur d'un équilibre de Nash fort. Ce lien déjà connu entre un critère majoritaire et une notion d'équilibre nous amène en section 3.3 à nous intéresser à plusieurs critères d'équilibre pour un système de vote : le fait que l'existence d'un ENF soit garanti par l'existence d'un vainqueur de Condorcet (**XENFC**) ou par celle d'un candidat Condorcet-admissible (**XENFA**) et la restriction des ENF aux vainqueurs de Condorcet (**RENFC**) ou aux candidats Condorcet-admissible (**RENFA**). Si les premiers sont des critères d'existence, les seconds peuvent être vus comme une version faible de critères d'unicité. Nous révélons les relations d'implication entre ces critères d'équilibre et les critères majoritaires. En particulier, non seulement

1. Pour être tout à fait rigoureux, nous verrons que la chaîne d'implications est complète si l'espace électoral autorise tout candidat comme préféré, ce qui est une hypothèse courante (cf. définition 1.10).

2. À propos des ENF et de diverses variantes de ce concept, on pourra consulter Bernheim *et al.* (1987).

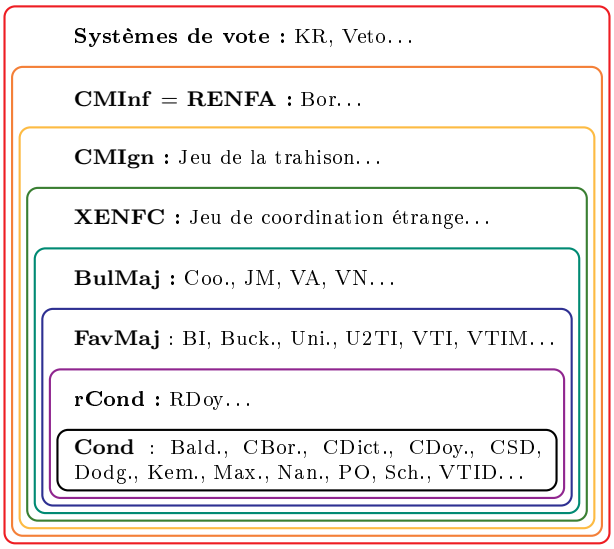


FIGURE 3.1 – Diagramme d'inclusion des critères majoritaires. On suppose que les préférences sont antisymétriques et que l'espace électoral autorise tout candidat comme préféré.

la simple formulation des critères permet de généraliser le résultat de Brill et Conitzer en montrant que **CMInf** implique **RENFA** mais nous montrons que ces deux critères sont en fait équivalents.

Dans la proposition 2.3, nous avions énoncé sans démonstration que presque tous les systèmes classiques vérifient **CMInf**. Dans la section 3.4, nous prouvons et précisons ce résultat en étudiant quels critères sont vérifiés par les modes de scrutin classiques et nous détaillons cette étude en fonction du nombre de candidats C. Ceci nous permettra d'établir progressivement le diagramme d'inclusion de la figure 3.1, qui présente l'inconvénient d'interrompre un suspense haletant mais l'avantage de fournir une carte d'orientation au lecteur pendant la lecture de ce chapitre. Ce diagramme se lit de la façon suivante. Par exemple, l'ensemble des systèmes de vote qui vérifient **FavMaj** est inclus (en général, strictement) dans l'ensemble de ceux qui vérifient **BulMaj** ; la méthode de Coombs appartient au second, mais généralement pas au premier (sauf dans des espaces électoraux particuliers, par exemple un espace des ordres stricts totaux avec 2 candidats).

Cette section 3.4 consacrée aux critères vérifiés par différents systèmes opère la synthèse de résultats classiques et de contributions originales. À notre connaissance, les résultats sur **CMInf**, **CMIgn**, **BulMaj** et **rCond** sont tous originaux puisque nous avons introduit ces critères ; ceci dit, les résultats sur **CMIgn** et **BulMaj** découlent essentiellement de la définition, puisque ces critères sont justement conçus pour être faciles à tester. Les résultats sur **CMInf** et **rCond** demandent en général plus d'effort. Les résultats sur **FavMaj** et **Cond** sont des résultats classiques, avec les nuances suivantes. Premièrement, nous n'avons pas trouvé d'étude exhaustive de **FavMaj** pour les RPSI-EM dans la littérature ; ceci dit, il serait surprenant que les résultats que nous présentons n'aient pas été énoncés, puisqu'il s'agit d'un critère et de modes de scrutin très classiques. Deuxièmement, les résultats sur la méthode Bucklin itérée sont tous originaux, puisque ce système de vote est une contribution de ce mémoire.

Enfin, dans la section 3.5, nous proposons une réflexion sur les critères étudiés en terme de quantité d'information nécessaire pour coordonner des stratégies de manipulation et pour obtenir des ENF. Ceci permet de discuter en quoi ces différents critères peuvent être considérés comme souhaitables pour un système de vote.

3.1 Définition des critères majoritaires

Dans les chapitres précédents, nous avons déjà rappelé le critère de Condorcet (définition 1.32) et introduit le critère de la coalition majoritaire informée (définition 2.1) et le critère de Condorcet-résistant (définition 2.18). Nous définissons, à présent, trois autres critères, dont on verra par la suite qu'ils sont des outils commodes pour prouver qu'un système de vote donné vérifie **CMInf** (et qu'il est, par conséquent, concerné par le théorème faible de condorcification 2.9).

Avant de passer aux critères proprement dits, définissons la notion de *favori majoritaire*.

Définition 3.1 (favori majoritaire)
Pour une configuration $\omega \in \Omega$ et un candidat $c \in \mathcal{C}$, on dit que c est un *favori majoritaire* dans ω ssi une majorité stricte d'électeurs préfère strictement c à tout autre candidat : $|\forall d \in \mathcal{C} \setminus \{c\}, c \ \mathrm{PP}_v \ d| > \frac{V}{2}$. Quand les préférences sont des ordres stricts (faibles ou totaux), cela signifie, simplement, que plus de la moitié des électeurs placent c en première position dans leur ordre de préférence, sans égalité avec d'autres candidats.

Si c est favori majoritaire, alors il est immédiat que c est un vainqueur de Condorcet résistant.

Définition 3.2 (critères majoritaires)
On dit que f vérifie le *critère du favori majoritaire* (**FavMaj**) ssi, pour toute configuration $\omega \in \Omega$ et pour tout candidat $c \in \mathcal{C}$, si c est favori majoritaire dans ω, alors $f(\omega) = c$.

On dit que f vérifie le *critère du bulletin majoritaire* (**BulMaj**) ssi, pour tout candidat c, il existe une assignation de bulletins aux électeurs vérifiant la propriété suivante : si elle est respectée par une majorité stricte d'électeurs, alors c est élu. Formellement, cette condition s'écrit : $\forall c \in \mathcal{C}, \exists \psi^c \in \Omega$ t.q. $\forall \omega \in \Omega, \left[|\omega_v = \psi^c_v| > \frac{V}{2} \Rightarrow f(\omega) = c\right]$.

On dit que f vérifie le *critère de la coalition majoritaire ignorante* (**CMIgn**) ssi toute coalition majoritaire peut décider du résultat, quoi que fassent les autres électeurs. Formellement, $\forall M \in \mathcal{P}(\mathcal{V})$, si $\mathrm{card}(M) > \frac{V}{2}$ alors : $\forall c \in \mathcal{C}, \exists \omega_M \in \Omega_M$ t.q. $\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}, f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c$.

Comme pour **CMInf**, **Cond** ou **rCond**, chaque notation **FavMaj**, **BulMaj** ou **CMIgn** désigne indifféremment le critère lui-même ou l'ensemble des SVBE (sur Ω) qui le vérifient.

Il est assez facile d'étendre les critères majoritaires aux systèmes de vote généraux définis en section 1.4. Dans ce cas, il apparaît que **CMInf**, **CMIgn** et **BulMaj** sont des propriétés du *mode de scrutin* : ces critères décrivent le pouvoir que la règle de dépouillement alloue à une majorité stricte d'électeurs. En revanche, **FavMaj**, **rCond** et **Cond** sont des propriétés du *système de vote* : elles relient les préférences des électeurs et le résultat sincère du vote. Pour ces derniers critères, il est donc nécessaire d'expliciter l'espace électoral et les fonctions de sincérité.

Dans un espace électoral anonyme (section 1.2.2), on peut également définir le *critère du bulletin majoritaire unisson* (**BulMajUni**) par la propriété suivante : pour tout candidat c, il existe un bulletin ψ_0^c (appartenant à un Ω_v quelconque, puisqu'ils sont tous identiques) tel que, s'il est utilisé par une majorité stricte d'électeurs, c est élu. Nous avons choisi de ne pas inclure cette notion dans le diagramme d'inclusion de la figure 3.1 car elle n'est pas définie dans tous les espaces électoraux.

La définition plus générale de **BulMaj** évite d'avoir besoin de supposer que l'espace électoral est anonyme. Cette formulation permet aussi d'appliquer ce critère à des modes de scrutin non anonymes (même si l'espace électoral l'est) : on peut, en effet, donner une consigne différente à chaque électeur ; tout ce que le critère exige, c'est qu'en cas de respect de la consigne par une majorité stricte d'électeurs, le candidat c désiré soit élu. Dans le cas d'un système de vote général (section 1.4), cette formulation présente, en outre, l'avantage qu'elle est stable par *isomorphisme de mode de scrutin*, au sens suivant : si on change les étiquettes des bulletins en adaptant la règle de dépouillement en conséquence, cela n'a pas d'impact sur le fait que ce critère soit vérifié ou non.

Le critère **CMIgn** est très semblable à **CMInf** (définition 2.1) : la seule différence réside dans l'échange des quantificateurs « $\exists \omega_M \in \Omega_M$ » et « $\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}$ ». En pratique, si les quantificateurs sont dans cet ordre, qui est celui de **CMIgn**, la manipulation est plus difficile car les manipulateurs votent d'abord et les autres électeurs peuvent riposter ; dans l'ordre inverse qui est celui de **CMInf**, les bulletins des électeurs sincères sont connus d'abord et les manipulateurs peuvent choisir leur vote en fonction de ceux-ci. Le critère **CMIgn** est donc plus exigeant que **CMInf**.

3.2 Implications entre les critères majoritaires

La proposition suivante établit une hiérarchie entre les six critères majoritaires que nous avons rencontrés. Nous les présentons sous forme d'inclusion afin d'établir le lien avec le diagramme 3.1 mais il est équivalent de les formuler en termes d'implications. Par exemple, l'inclusion **Cond** $\subseteq$ **rCond** est une autre façon d'énoncer pour tout SVBE f, on a : $f \in$ **Cond** $\Rightarrow f \in$ **rCond**.

Proposition 3.3
On a les inclusions suivantes.

1. **Cond** $\subseteq$ **rCond** $\subseteq$ **FavMaj**.
2. **BulMaj** $\subseteq$ **CMIgn** $\subseteq$ **CMInf**.

Si l'espace électoral autorise tout candidat comme préféré (ce qui est une hypothèse courante, cf. définition 1.10), alors on a également **FavMaj** $\subseteq$ **BulMaj**. *Sous cette hypothèse, on a donc :*

$$\textbf{Cond} \subseteq \textbf{rCond} \subseteq \textbf{FavMaj} \subseteq \textbf{BulMaj} \subseteq \textbf{CMIgn} \subseteq \textbf{CMInf}.$$

Toutes les inclusions sont strictes en général, c'est-à-dire que les implications réciproques sont toutes fausses.

Démonstration. **Cond** $\subseteq$ **rCond** $\subseteq$ **FavMaj** : cette chaîne d'inclusions découle immédiatement du fait qu'un favori majoritaire est un vainqueur de Condorcet résistant et qu'un vainqueur de Condorcet résistant est un vainqueur de Condorcet.

BulMaj $\subseteq$ **CMIgn** : pour tout candidat c, **BulMaj** assure l'existence d'une assignation de bulletins ψ^c qui permet d'élire c si elle est utilisée par une majorité stricte d'électeurs. Si une coalition strictement majoritaire souhaite faire élire c, il leur suffit d'utiliser ces bulletins pour que c soit élu, quelle que soit la réaction des autres électeurs. Donc le système utilisé vérifie **CMIgn**.

CMIgn $\subseteq$ **CMInf** : cette inclusion découle immédiatement de la remarque que nous avons faite sur l'ordre des quantificateurs dans les définitions de ces critères.

FavMaj $\subseteq$ **BulMaj** : considérons une assignation de bulletins où chaque électeur affirme strictement préférer un certain candidat c à tout autre candidat, ce qui est possible puisque l'espace électoral autorise tout candidat comme favori. Si le système considéré respecte **FavMaj**, il suffit qu'une majorité stricte d'électeurs respecte cette assignation pour que c soit élu. Donc le système considéré vérifie **BulMaj**.

Le fait que les réciproques soient fausses découlera de l'étude des modes de scrutin classiques que nous ferons en section 3.4. On peut déjà avoir un aperçu des contre-exemples de chaque inclusion réciproque en consultant le diagramme d'inclusion de la figure 3.1. □

Dans un espace électoral anonyme, il est clair qu'on a également **BulMajUni** $\subseteq$ **BulMaj**. Si, en outre, l'espace électoral autorise tout candidat comme favori, alors on a aussi **FavMaj** $\subseteq$ **BulMajUni**.

3.3 Lien avec les équilibres de Nash forts (ENF)

Comme le remarque Gibbard (1973), la donnée d'un mode de scrutin et d'une configuration de préférence ω pour les électeurs définit un jeu, au sens habituel de la théorie des jeux : chaque joueur-électeur dispose d'un ensemble de stratégies, d'objectifs bien déterminés et il existe une règle qui permet de décider le résultat en fonction des stratégies des joueurs. Il est donc immédiat d'adapter la notion classique d'*équilibre de Nash fort* dans ce cadre.

On pourra d'ailleurs remarquer qu'en utilisant le formalisme des systèmes de vote généraux (section 1.4), tout jeu peut être exprimé sous forme d'un mode de scrutin et d'un état de préférence des électeurs. C'est ce qui amène Gibbard à appeler un mode de scrutin (au sens de la définition 1.18) un *game form* : il s'agit d'un proto-jeu qui définit une procédure à suivre et auquel il ne manque que la donnée des préférences des joueurs sur les issues possibles du jeu.

Définition 3.4 (équilibre de Nash fort)
Soit f un SVBE et $(\omega, \psi) \in \Omega^2$.

On dit que ψ est un *équilibre de Nash fort* (ENF) pour les préférences ω dans le système f ssi ψ est un équilibre de Nash fort pour le jeu défini par (f, ω). C'est-à-dire qu'il n'existe pas de configuration ϕ telle que :

$$\begin{cases} f(\phi) \neq f(\psi), \\ \forall v \in \mathrm{Sinc}_\omega(f(\psi) \to f(\phi)), \phi_v = \psi_v. \end{cases}$$

Dans le cas d'un système de vote général, en utilisant les notations de la section 1.4, on adapterait la définition en disant qu'un vecteur $S = (S_1, \ldots, S_V)$ de stratégies des électeurs (c'est-à-dire de bulletins) est un ENF pour une configuration de préférence ω. Comme nous l'avons déjà remarqué, le fait qu'une configuration ω soit non-manipulable signifie simplement que le vote sincère $s(\omega)$ est un ENF pour ω.

Pour un SVBE, le formalisme est simplifié puisque les notions de bulletin autorisé et d'état de préférence possible sont confondues. Le fait qu'une configuration ω soit non-manipulable signifie alors que la configuration ω (vue comme la donnée des bulletins) est un ENF pour ω (vue comme les préférences des électeurs).

Comme nous l'évoquions en introduisant ce chapitre, Brill et Conitzer (2015) ont montré que, dans un système qui vérifie **CMInf**, si une configuration est Condorcet, alors seul le vainqueur de Condorcet peut être le vainqueur d'un ENF. Ce résultat suggère un lien profond entre certains critères majoritaires, non seulement avec la manipulabilité, mais avec la notion d'ENF en général. À la suite de la lecture de cet article et d'enrichissants échanges autour d'une boisson probablement sans alcool avec Markus Brill, que nous remercions ici, nous avons examiné quelques question liées à cette problématique. Nous avons donc défini les quatre critères d'équilibre suivants et nous avons étudié leur lien avec les critères majoritaires.

Définition 3.5 (critères d'équilibre)
On dit que f vérifie le critère de *restriction des ENF possibles aux candidats Condorcet-admissibles* (**RENFA**) ssi pour tout $(\omega, \psi) \in \Omega^2$: si ψ est un ENF pour les préférences ω, alors $f(\psi)$ est Condorcet-admissible en ω.

On dit que f vérifie le critère de *restriction des ENF possibles aux vainqueurs de Condorcet* (**RENFC**) ssi pour tout $(\omega, \psi) \in \Omega^2$: si ψ est un ENF pour les préférences ω, alors $f(\psi)$ est vainqueur de Condorcet en ω.

On dit que f vérifie le critère d'*existence d'ENF pour un vainqueur de Condorcet* (**XENFC**) ssi $\forall(\omega, c) \in \Omega \times \mathcal{C}$: si c est vainqueur de Condorcet dans ω, alors il existe ψ qui est un ENF pour ω et tel que $f(\psi) = c$.

On dit que f vérifie le critère d'*existence d'ENF pour un candidat Condorcet-admissible* (**XENFA**) ssi $\forall(\omega, c) \in \Omega \times \mathcal{C}$: si c est Condorcet-admissible dans ω, alors il existe ψ qui un ENF pour ω et tel que $f(\psi) = c$.

Comme nous le remarquions en introduction de ce chapitre, le critère **RENFC** peut être vu comme un critère d'unicité, non pas de l'ENF, mais du candidat vainqueur dans un ENF. Dans un espace électoral où il existe au moins une configuration semi-admissible (c'est-à-dire s'il est possible d'avoir un candidat Condorcet-admissible qui n'est pas vainqueur de Condorcet, comme dans l'espace électoral des ordres stricts totaux avec un nombre pair d'électeurs), le critère **RENFA** en est une version affaiblie, dont nous verrons qu'elle est vérifiée par la plupart des systèmes de vote usuels. S'il n'y a pas de configuration semi-Condorcet (par exemple dans l'espace électoral des ordres stricts totaux avec un nombre impair d'électeurs), les deux notions sont équivalentes. De même, dans ce cas, les critères d'existence **XENFC** et **XENFA** sont équivalents.

Le critère **RENFA** implique en particulier le critère de Brill et Conitzer : quand il est vérifié, si ω est une configuration Condorcet, alors le vainqueur de Condorcet est le seul candidat Condorcet-admissible (proposition 1.31) donc il est le seul vainqueur possible d'un ENF.

À présent, nous étudions les liens entre ces critères d'équilibre et les critères majoritaires. Nous nous concentrons d'abord sur **RENFA** et **XENFC**, dont nous verrons qu'ils s'intègrent naturellement dans la chaîne d'inclusion de la proposition 3.3, ce qui justifie leur position dans le diagramme d'inclusion de la figure 3.1.

Proposition 3.6
On suppose que l'espace électoral autorise tout candidat comme préféré.
Alors **CMInf** = **RENFA**.

Dans le même esprit, Sertel et Sanver (2004) montrent que dans l'espace électoral des ordres stricts faibles, si un système de vote vérifie **FavMaj**, alors le vainqueur d'un ENF est nécessairement un vainqueur de Condorcet faible (ce qui est équivalent à un candidat Condorcet-admissible dans ce contexte) : autrement dit, **FavMaj** $\subseteq$ **RENFA**. La proposition 3.6 est donc une généralisation de ce résultat (au sens où elle l'implique). Les auteurs, quant à eux, le généralisent dans une autre direction en considérant une famille de variantes des critères **FavMaj** et **RENFA**.

Démonstration. Si un système de vote f vérifie **CMInf**, alors il vérifie **RENFA** : en effet, si le candidat vainqueur d'une configuration ψ n'est pas Condorcet-admissible dans ω, alors une majorité stricte d'électeurs préfèrent un autre candidat (au sens de ω) et sont capables de le faire gagner en vertu de **CMInf**. Donc ψ ne peut pas être un ENF pour ω.

À présent, considérons un système de vote qui ne vérifie pas **CMInf** et montrons qu'il ne vérifie pas **RENFA**. Par définition, il existe un candidat c, une coalition strictement majoritaire M, une configuration $\psi_{\mathcal{V}\setminus M}$ des autres électeurs tels que quels que soient les bulletins ψ_M de la coalition, $f(\psi_M, \psi_{\mathcal{V}\setminus M}) \neq c$.

Considérons l'état $\psi_{\mathcal{V}\setminus M}$ mentionné ci-dessus (tel que la minorité empêche l'élection de c) et un état arbitraire ψ_M de la coalition majoritaire M. Notons $a = f(\psi_M, \psi_{\mathcal{V}\setminus M})$, qui est par hypothèse distinct de c.

Considérons à présent le profil de préférence ω suivant.

Électeurs de M (majorité)	Électeurs de $\mathcal{V} \setminus M$ (minorité)
c	a
a	c
Autres	Autres

Clairement, c est vainqueur de Condorcet, donc c'est le seul candidat Condorcet-admissible.

Si les bulletins sont ψ, alors le candidat a est vainqueur par hypothèse. Les seules déviations possibles proviennent des électeurs de M, et leur seul souhait est de faire gagner c, mais c'est impossible. Par conséquent, ψ est un ENF pour ω dont le vainqueur n'est pas Condorcet-admissible en ω. Ainsi, le système de vote considéré ne vérifie pas **RENFA**. □

Proposition 3.7

On suppose que les relations binaires de préférence sont antisymétriques.
On a : **BulMaj** $\subseteq$ **XENFC**. *La réciproque est fausse en général.*

Démonstration. Soit f un SVBE qui vérifie **BulMaj** et ω une configuration. S'il existe un vainqueur de Condorcet c en ω, alors considérons la situation ψ^c dont l'existence est garantie par **BulMaj** : si une majorité d'électeurs respecte cette assignation de bulletins, alors c est élu. Il est facile de voir que c'est un ENF pour ω : en effet, si un sous-ensemble d'électeurs souhaite faire élire un certain candidat d au lieu de c, alors comme c est vainqueur de Condorcet et comme les préférences sont antisymétriques, ces électeurs sont strictement minoritaires. Donc, par définition de ψ^c, le candidat c reste élu. Ainsi, f vérifie **XENFC**.

Pour montrer que la réciproque est fausse, considérons l'espace électoral des ordres stricts totaux pour $V = 3$ électeurs et $C = 3$ candidats. Nous allons définir un système de vote que nous appelons le *jeu de coordination étrange* et montrer qu'il vérifie **XENFC** mais pas **BulMaj**.

Voici la règle utilisée.

1. S'il existe au moins un doubleton d'électeurs dont les deux membres mettent le même candidat en tête de bulletin mais leurs deux derniers candidats dans des ordres différents, alors leur candidat préféré est élu.
2. Dans tous les autres cas, le doyen (c'est-à-dire un candidat fixé à l'avance) est élu.

Ainsi, si deux électeurs souhaitent faire élire un candidat c qui n'est pas le doyen, il doivent se coordonner habilement pour placer c en tête et surtout pour s'assurer que leurs bulletins ne sont pas identiques.

Montrons que ce système vérifie **XENFC**. Considérons une configuration ω Condorcet. Comme $V = 3$ et $C = 3$, un vainqueur de Condorcet est forcément le favori d'au moins un électeur (sinon un autre candidat est au moins deux fois en tête de bulletin donc il est favori majoritaire et *a fortiori* vainqueur de Condorcet).

Si le vainqueur de Condorcet est favori d'au moins deux électeurs, alors il leur suffit de désynchroniser le bas de leurs bulletins pour qu'il soit élu, et la situation est alors un ENF : ils sont parfaitement satisfaits, et si ce n'est pas le cas du troisième électeur, celui-ci n'y peut rien.

Si le vainqueur de Condorcet est favori d'un électeur exactement, alors il est facile de montrer qu'à permutation près des électeurs et des candidats, on est dans un profil de préférence ω du type suivant.

c	a	b
Autres	c	c
	b	a

Le candidat c est vainqueur de Condorcet. Quitte à échanger les rôles de a et b et ceux des deux derniers électeurs, on peut supposer que le doyen n'est pas b : c'est donc c ou a.

Considérons alors le profil ψ suivant.

c	a	**c**
a	c	**b**
b	b	a

Le candidat c est vainqueur et c'est un ENF pour ω : en effet, le second électeur ne peut pas changer le résultat et le troisième électeur ne peut pas faire gagner b (il ne peut bénéficier ni de la règle 1 car b ne peut devenir favori majoritaire, ni de la règle 2 car b n'est pas doyen). Ainsi, le jeu de coordination étrange vérifie **XENFC**.

À présent, prouvons qu'il ne vérifie pas **BulMaj**. Considérons un candidat c distinct du doyen et essayons de concevoir une assignation des bulletins ψ^c qui assure une victoire de c dès qu'elle est respectée par une majorité stricte d'électeurs. Pour chaque paire d'électeurs, il faut exactement un bulletin $c \succ a \succ b$ et un bulletin $c \succ b \succ a$, ce qui est impossible : en effet, on ne s'autorise que deux bulletins différents pour trois électeurs, donc il existe deux électeurs qui ont le même bulletin. Si ces deux électeurs utilisent ce bulletin identique, il suffit à l'autre électeur de ne pas placer c en tête pour que le doyen soit élu, ce qui contredit la définition de ψ^c. Ainsi, le jeu de coordination étrange ne vérifie pas **BulMaj**. □

Proposition 3.8

On suppose que l'espace électoral autorise tout candidat comme préféré.
Alors **XENFC** $\subseteq$ **CMIgn**. *La réciproque est fausse en général.*

Démonstration. Considérons un système de vote qui ne vérifie pas **CMIgn**. Par définition, il existe un candidat c et une coalition strictement majoritaire M tels que pour tout allocation ψ_M de bulletins pour la coalition, il existe une réponse $\psi_{\mathcal{V}\setminus M}$ des autres électeurs telle que $f(\psi_M, \psi_{\mathcal{V}\setminus M}) \neq c$.

Considérons un profil de préférence ω du type suivant.

Électeurs de M (majorité)	Électeurs de $\mathcal{V} \setminus M$ (minorité)
c	Autres
Autres	c

Clairement, c est vainqueur de Condorcet.

Dans un état ψ qui est un ENF pour ω, il est impossible que c soit vainqueur : en effet, les membres de $\mathcal{V} \setminus M$ peuvent empêcher une victoire de c et ils sont intéressés quel que soit le nouveau résultat. Ainsi, le système de vote ne vérifie pas **XENFC**, ce qui prouve l'inclusion **XENFC** $\subseteq$ **CMIgn** par contraposition.

Afin de prouver que la réciproque est fausse, nous allons définir un système de vote que nous appelons le *jeu de la trahison* et montrer qu'il vérifie **CMIgn** mais pas **XENFC**. Nous allons utiliser un système de vote non ordinal afin de mieux révéler l'intuition sous-jacente, puis nous montrerons qu'avec un peu d'effort, on peut adapter ce contre-exemple à système ordinal.

Voici la règle utilisée : chaque électeur v produit un bulletin du type $(M_v, c_v) \in \mathcal{P}(\mathcal{V}) \times \mathcal{C}$: il annonce une coalition, c'est-à-dire un ensemble d'électeurs, et l'indice d'un candidat. Afin de se fixer les idées, on peut interpréter M_v comme l'ensemble des électeurs que v considère comme ses « amis ». On pourrait exiger que la coalition M_v contienne obligatoirement v lui-même, mais cette hypothèse supplémentaire n'a aucune incidence sur notre démonstration. À présent, voici comment on détermine le candidat vainqueur.

1. S'il existe une coalition strictement majoritaire M et un candidat c tels que chaque électeur de M produit le bulletin (M, c), alors c est élu (il n'y a pas d'ambiguïté car une telle coalition est nécessairement unique).
2. Dans tous les autres cas, le résultat est $\sum_{v\in\mathcal{V}} c_v$, ramené modulo C à l'intervalle $[\![1, C]\!]$.

Clairement, ce mode de scrutin vérifie **CMIgn** : en effet, si une coalition strictement majoritaire M souhaite faire élire un candidat c, il leur suffit de se coordonner pour annoncer le bulletin (M, c), ce qui garantit l'élection de c, quels que soient les bulletins des autres électeurs.

Afin de montrer que le jeu de la trahison ne vérifie pas **XENFC**, considérons le profil de préférence suivant, où le candidat c est vainqueur de Condorcet.

c	a	b
a	c	c
b	b	a

Peut-il y avoir un ENF où c est élu ? Si c est élu en vertu de la règle 1, alors le deuxième électeur, le troisième électeur ou les deux participent nécessairement à la coalition M puisque celle-ci est supposée majoritaire. N'importe lequel d'entre eux peut alors trahir la coalition M en nommant la coalition constituée de lui-même et faire gagner a ou b en choisissant le modulo adéquat.

L'autre possibilité est que c soit élu en vertu de la règle 2. Alors le deuxième électeur peut nommer la coalition constituée de lui-même et faire gagner a en choisissant le modulo adéquat.

Par conséquent, il n'y a pas d'ENF où c est élu[3], ce qui prouve que le jeu de la trahison ne vérifie pas **XENFC**.

Si l'on souhaite avoir un exemple purement ordinal, il suffit de considérer des candidats supplémentaires $d_1, \dots, d_4$ et de coder la coalition M par l'ordre donné sur les quatre derniers candidats d'un bulletin (4 candidats fournissent une expressivité suffisante car $4! = 24 > 2^3 = 8$), selon un procédé similaire à celui utilisé dans la section 2.8 pour construire un système Condorcet très manipulable. On considère alors une variante du profil de préférence ci-dessus où les candidats $d_1, \dots, d_4$ sont placés en dernier chez tous les électeurs et on montre de la même façon qu'il n'existe pas d'ENF où c est élu. □

Jusqu'à présent, nous nous sommes concentrés sur les critères **RENFA** et **XENFC** et nous avons démontré leur positionnement dans le diagramme d'inclusion de la figure 3.1. Les critères **RENFC** et **XENFA** en sont manifestement des renforcements respectifs, au sens où on a **RENFC** $\subseteq$ **RENFA** et **XENFA** $\subseteq$ **XENFC**. Si l'espace électoral ne contient aucune configuration semi-Condorcet (par exemple l'espace électoral des ordres stricts totaux avec un nombre impair d'électeurs), alors il est immédiat que ces inclusions sont des égalités : **RENFC** = **RENFA** et **XENFA** = **XENFC**.

Pour tous les critères à l'intérieur de la zone **RENFA** (resp. **XENFC**), en particulier **Cond**, il existe donc des SVBE qui vérifient également **RENFC** (resp. **XENFA**) : en effet, il suffit de considérer un espace électoral sans configuration semi-Condorcet. Dans ce cas, tout système vérifiant un critère plus fort que **RENFA** (resp. **XENFC**) vérifie également **RENFC** (resp. **XENFA**) puisque les notions sont équivalentes.

D'un autre côté, pour tous les critères à l'intérieur de la zone **RENFA** (resp. **XENFC**), en particulier **Cond**, il existe également des SVBE qui ne vérifient pas **RENFC** (resp. **XENFA**). En effet, considérons l'espace électoral des ordres stricts totaux avec un nombre pair d'électeurs et le système de vote Condorcet-doyen, où le doyen est noté a.

1. Si la moitié des électeurs préfèrent un candidat c distinct du doyen aux autres candidats et si l'autre moitié préfère tous les autres candidats à c, alors c est Condorcet-admissible mais il n'y a pas d'ENF où il est élu (il y a toujours une déviation possible en faveur de a). Donc ce système ne vérifie pas **XENFA**.
2. Si la moitié des électeurs préfèrent le doyen a aux autres candidat et si l'autre moitié place a en dernier dans leur ordre de préférence, alors le vote sincère est un ENF où a est élu, bien qu'il soit seulement Condorcet-admissible mais pas vainqueur de Condorcet. Donc ce système ne vérifie pas **RENFC**.

Par conséquent, **XENFA** est inclus dans **XENFC** mais il n'entretient, en général, pas de relation d'inclusion simple avec les critères à l'intérieur de la zone **XENFC** : par exemple, il peut posséder une intersection non vide avec **Cond** mais aussi avec son complémentaire. Il est est de même pour **RENFC**, qui est inclus dans **RENFA** mais qui n'a en général pas de relation d'inclusion simple avec les autres critères plus forts que **RENFA**.

Par ailleurs, on pourra remarquer que, dans un espace électoral où il existe au moins une configuration semi-Condorcet, les critères **XENFA** et **RENFC** sont

3. En fait, dans cette configuration de préférence, on peut même montrer qu'il n'y a pas d'ENF du tout.

incompatibles : en effet, dans une situation semi-Condorcet, **XENFA** exige qu'il existe un ENF où un candidat Condorcet-admissible est élu, mais comme celui-ci n'est pas vainqueur de Condorcet, ceci contredit **RENFC**.

Pour toutes ces raisons, il semble que les critères **XENFA** et **RENFC** sont moins « naturels » que leurs versions faibles **XENFC** et **RENFA** et nous avons choisi de ne pas les inclure dans le diagramme de la figure 3.1 afin de ne pas l'alourdir inutilement.

Pour résumer cette section, nous avons étudié « l'unicité » et l'existence d'ENF en lien avec les notions Condorcet.

Pour « l'unicité » de l'équilibre, nous avons défini le critère **RENFC**, qui exige que seul un vainqueur de Condorcet puisse bénéficier d'un ENF. Mais finalement, il semble que **RENFA** soit une propriété plus naturelle que sa version faible **RENFC**, au sens où elle est plus souvent vérifiée par les systèmes usuels, ces notions étant bien sûr équivalentes quand il n'y a pas de configuration semi-Condorcet. S'il était assez clair dès la définition que **CMInf** impliquait **RENFA**, nous avons montré qu'en fait, les deux critères sont équivalents.

Nous souhaitions aussi savoir ce qui garantit l'existence d'un ENF quand les préférences sont Condorcet, c'est-à-dire le critère **XENFC**. Nous avons vu que ce critère est plus naturel que sa version forte **XENFA**. Nous avons déterminé que **BulMaj** est suffisant mais pas nécessaire (comme le montre le jeu de coordination étrange) et que **CMIgn** est nécessaire mais pas suffisant (comme le montre le jeu de la trahison).

3.4 Critères majoritaires vérifiés par les modes de scrutin classiques

Nous comprenons mieux, à présent, la position des critères majoritaires et des critères d'équilibre dans le diagramme d'inclusion de la figure 3.1. Avant de proposer une réflexion sur la motivation et les conséquences de ces critères, nous allons étudier quels critères sont vérifiés par les systèmes de vote usuels. Dans toute cette section, nous nous placerons à nombre de candidats C fixé mais à nombre d'électeurs V variable : ainsi, on dira qu'un système de vote[4] vérifie un certain critère pour une valeur de C donnée s'il le vérifie pour tout nombre d'électeurs ; et qu'il ne le vérifie pas s'il existe une configuration qui contredit le critère mentionné, quel que soit le nombre d'électeurs mentionné.

La motivation de ce choix est double. D'une part, quand on organise une élection en pratique, on connaît presque toujours le nombre de candidats (au moins au moment où les bulletins doivent être donnés), alors qu'il est fréquent qu'on ne connaisse pas à l'avance le nombre d'électeurs. D'autre part, fixer le nombre d'électeurs peut donner lieu à des questions difficiles de quantification : ainsi, il peut arriver qu'un certain système respecte un certain critère pour une valeur de (V, C) pour des raisons liées à des problèmes de sac à dos, alors qu'il ne vérifie pas le critère en général pour cette valeur de C (en particulier pour un nombre d'électeurs suffisamment grand). Autrement dit, il est pragmatique et techniquement plus accessible de laisser varier le nombre d'électeurs. En cela,

4. Afin de respecter une terminologie mathématiquement rigoureuse, nous devrions dire un *méta-système de vote*, puisqu'en faisant varier le nombre d'électeurs, nous considérons des systèmes qui sont définis sur plusieurs espaces électoraux différents. Nous éviterons, cependant, cette coquetterie de langage et nous continuerons à parler de *systèmes de vote*, même dans ce cas.

nous rejoignons, par exemple, l'approche de Smith (1973), qui a étudié certaines de ces questions pour les RPS avec un nombre d'électeurs variable.

3.4.1 Modes de scrutin cardinaux

Proposition 3.9

1. *Le vote par notation, le jugement majoritaire et le vote par assentiment vérifient* **BulMaj**.
2. *Soit $C \geq 3$. Supposons que, pour tout électeur v et pour tout ordre strict total p_v sur les candidats, l'électeur v dispose au moins :*
   — *d'un état ω_v tel que $\mathrm{P}_v(\omega_v) = p_v$ où il attribue la note maximale à son candidat préféré et la note minimale aux autres ;*
   — *et d'un état ω'_v tel que $\mathrm{P}_v(\omega'_v) = p_v$ où il attribue la note maximale à ses deux candidats préférés et la note minimale aux autres.*
   Alors ces modes de scrutin ne vérifient pas **FavMaj**.

Démonstration. 1. Soit c un candidat. Considérons un bulletin qui consiste à attribuer la note maximale à c et la note minimale aux autres candidats. Si une majorité stricte d'électeurs utilisent ce bulletin, alors c est élu. Ainsi, ces modes de scrutin vérifient **BulMaj**. Si on définit ces systèmes de vote raisonnablement dans un espace électoral anonyme, alors nous venons de démontrer qu'ils vérifient même **BulMajUni**.

2. Considérons la configuration suivante où les candidats sont notés $c, d_1, \ldots, d_{C-1}$ et où les notes minimale et maximale sont respectivement fixées à 0 et 1 par convention.

2	1
$c : 1$	$d_1 : 1$
$d_1 : 1$	$d_2 : 0$
$d_2 : 0$	$\vdots$
$\vdots$	$d_{C-1} : 0$
$d_{C-1} : 0$	$c : 0$

Alors le candidat c est favori majoritaire mais c'est d_1 qui est élu. □

La propriété 2 est également vraie pour $C = 2$ mais nous avons préféré ne pas mentionner ce cas dans la proposition car l'hypothèse qui y est faite devient, alors, assez discutable. En effet, s'il n'y a que deux candidats et si les relations binaires de préférence sont des ordres stricts totaux, alors il y a une fonction de sincérité naturelle puisque chaque électeur dispose d'une stratégie dominante : attribuer la note maximale à son candidat préféré et la note minimale à l'autre. En utilisant cette fonction de sincérité, ces systèmes de vote cardinaux deviennent équivalents au vote majoritaire simple (évoqué en introduction) et ils vérifient donc trivialement **FavMaj** (et même **Cond**). En revanche, dès que $C \geq 3$, l'hypothèse de pouvoir sincèrement voter pour son candidat préféré ou pour ses deux candidats préférés est assez naturelle en pratique et conforme à l'esprit du vote par assentiment, du vote par notation et du jugement majoritaire.

Dans tout le reste de cette section 3.4 consacrée aux systèmes de vote usuels, on se place dans l'espace électoral des ordres stricts totaux.

Pour mémoire, on rappelle que Condorcet-dictature, Condorcet-doyen, CSD, Maximin, PO, VTID et les méthodes de Baldwin, Black (Condorcet-Borda), Copeland, Dodgson, Kemeny, Nanson et Schulze vérifient **Cond**, ce qui détermine immédiatement leur place dans le diagramme d'inclusion de la figure 3.1.

3.4.2 Uninominal à un ou plusieurs tours

Même si nous reverrons l'uninominal comme un cas particulier de la proposition 3.13 sur les RPS, VTI comme un cas particulier de la proposition 3.16 sur les RPSI-ES et VTIM comme un cas particulier de la proposition 3.21 sur les RPSI-EM, nous allons d'abord étudier ces systèmes de vote à part et profiter de l'occasion pour examiner l'uninominal à deux tours. D'une part, certains de ces systèmes de vote sont très utilisés en pratique (tous sauf VTIM) et ils présentent, tous, des similitudes de principe. D'autre part, cela nous permettra de montrer dans un contexte simple des résultats utiles pour la suite.

Proposition 3.10

1. *L'uninominal, l'uninominal à deux tours (U2T ou U2TI), VTI et VTIM vérifient* **FavMaj**.
2. *Ils vérifient* **rCond** *ssi* $C \leq 4$.
3. *Ils ne vérifient pas* **Cond** *(sauf dans le cas trivial* $C \leq 2$*).*

Pour prouver cette proposition, nous allons exploiter les lemmes suivants.

Lemme 3.11

On se place dans l'espace électoral des ordres stricts totaux et on suppose qu'un certain candidat c est vainqueur de Condorcet résistant.

Si $C \leq 4$, alors le score de c au sens de l'uninominal est strictement le plus élevé.

Si $C \geq 5$, alors le score de c au sens de l'uninominal peut être strictement le plus faible.

Démonstration. Pour $C \leq 3$, un vainqueur de Condorcet résistant est, par définition, un favori majoritaire, donc l'affirmation est évidente. Supposons donc $C = 4$. Nommons les candidats c, d_1, d_2, d_3 et notons γ (resp. δ_1, δ_2, δ_3) le nombre d'électeurs qui ont c (resp. d_1, d_2, d_3) en tête dans leur ordre de préférence. On a donc le profil suivant.

γ	δ_1	δ_2	δ_3
c	d_1	d_2	d_3
Autres	Autres	Autres	Autres

Comme c est vainqueur de Condorcet résistant, une majorité stricte d'électeurs le préfère simultanément à d_1 et d_2. Or on sait qu'au moins $\delta_1 + \delta_2$ électeurs ne vérifient pas cette condition. Par conséquent, on a $\gamma + \delta_3 > \delta_1 + \delta_2$. Des relations similaires sont obtenues par permutation des δ_i.

Supposons qu'un candidat, par exemple d_1, possède un score supérieur ou égal à celui de c au sens de l'uninominal, c'est-à-dire qu'on a $\delta_1 \geq \gamma$. En combinant cette relation avec les précédentes, on obtient successivement $\delta_1 + \delta_2 \geq \gamma + \delta_2 > \delta_1 + \delta_3 \geq \gamma + \delta_3 > \delta_1 + \delta_2$: c'est absurde !

Pour $C \geq 5$, considérons un profil du type suivant.

1	2	2	2	2
c	d_1	d_2	d_3	d_4
	c	c	c	c
Autres	Autres	Autres	Autres	Autres

Le candidat c est préféré à tout autre couple d'adversaires par au moins 5 électeurs sur 9, donc il est vainqueur de Condorcet résistant. Mais son score en uninominal est strictement le plus faible. □

Lemme 3.12

On se place dans l'espace électoral des ordres stricts totaux et on suppose qu'un certain candidat c est vainqueur de Condorcet.

Si $C \geq 3$, le score de c au sens de l'uninominal peut être strictement le plus faible.

Démonstration. Cette propriété classique est un cas particulier du lemme de Smith (1973), qui montre le même résultat non seulement pour l'uninominal mais pour toutes les RPS. Nous y reviendrons dans le lemme 3.17. Dans l'immédiat, il suffit, par exemple, de considérer le profil suivant.

1	2	2
c	d_1	d_2
	c	c
Autres	Autres	Autres

□

Nous avons, maintenant, tous les éléments pour prouver la proposition 3.10.

Démonstration. 1. Il est clair que l'uninominal, l'uninominal à deux tours et VTI vérifient **FavMaj**.

2. Le lemme 3.11 prouve qu'avec 4 candidats ou moins, un vainqueur de Condorcet résistant a toujours strictement le meilleur score en uninominal. En uninominal à deux tours, cela implique qu'il est sélectionné pour le second tour ; comme il est vainqueur de Condorcet, il en sort vainqueur. En VTI ou en VTIM, cela implique qu'il a le meilleur score à chaque tour ; en particulier, il remporte l'élection.

Le lemme 3.11 énonce également qu'avec 5 candidats ou plus, un vainqueur de Condorcet résistant peut avoir strictement le pire score en uninominal. Donc il peut perdre en uninominal et être éliminé au premier tour en uninominal à deux tours, en VTI ou en VTIM.

3. Le lemme 3.12 prouve qu'avec 3 candidats ou plus, un vainqueur de Condorcet peut avoir strictement le pire score en uninominal, donc perdre également en uninominal à deux tours, en VTI ou en VTIM. □

3.4.3 Règles positionnelles à score (RPS)

Nous connaissons déjà les critères vérifiés par l'uninominal. Généralisons maintenant notre investigation aux RPS, qui comprennent notamment la méthode de Borda et Veto.

Pour étudier les RPS et plus tard les RPS itérées, il sera souvent commode d'énoncer les résultats en utilisant les moyennes arithmétiques de tous les poids ou de certains d'entre eux. Nous utiliserons la convention suivante : étant donné des réels $x_1, \ldots, x_n$, on note $\mathrm{moy}(x_1, \ldots, x_n)$ leur moyenne arithmétique.

Proposition 3.13

Soit f une RPS de vecteur de poids $\mathbf{x}$.

1. *Pour que f vérifie* **CMInf**, *il est nécessaire que pour tout $k \in [\![2, C]\!]$, on ait :*

$$\mathrm{moy}(x_1, \ldots, x_k) + \mathrm{moy}(x_{C-k+1}, \ldots, x_C) \leq x_1 + x_C.$$

2. *Pour que f vérifie* **CMInf**, *il est suffisant que pour tout $k \in [\![2, C]\!]$, on ait :*
$$x_k + x_{C-k+2} \leq x_1 + x_C$$
et qu'au moins $j-1$ telles inégalités soient strictes, où j est le plus grand entier tel que $x_j = x_1$.
3. *f vérifie* **CMIgn** *ssi c'est l'uninominal. Dans ce cas, elle vérifie aussi* **FavMaj**.
4. *f vérifie* **rCond** *ssi c'est l'uninominal avec $C \leq 4$.*
5. *f ne vérifie pas* **Cond** *(sauf dans le cas trivial $C \leq 2$).*

Tout d'abord, remarquons que le cas particulier $k = 1$ n'est pas mentionné dans la propriété 1. C'est, en effet, inutile car la condition énoncée devient $x_1 + x_C \leq x_1 + x_C$, ce qui est toujours vérifié.

Pour la propriété 2, on considère la somme de x_k et x_{C-k+2}, c'est-à-dire qu'on additionne x_2 et x_C, ou x_3 et x_{C-1}, etc. Le cas particulier $k = 1$ n'aurait donc tout simplement pas de sens, puisque le poids x_{C+1} n'existe pas.

Démonstration. 1. Nous traiterons le cas particulier $k = C$ en dernier. Supposons qu'il existe $k \in [\![2, C-1]\!]$ tel que :
$$\mathrm{moy}(x_1, \ldots, x_k) + \mathrm{moy}(x_{C-k+1}, \ldots, x_C) > x_1 + x_C.$$
Soit α un entier naturel. Considérons α électeurs sincères qui placent toujours certains candidats $d_1, \ldots, d_k$ en tête (pas forcément dans le même ordre) et le candidat c en dernier. Si $\alpha+1$ électeurs supplémentaires souhaitent faire gagner c, il est nécessaire que c ait un meilleur score que les candidats $d_1, \ldots, d_k$. Pour réaliser cette condition nécessaire, les manipulateurs ne peuvent mieux faire que mettre c en tête et $d_1, \ldots, d_k$ en bas de bulletin (pas forcément dans le même ordre). Mais alors, on a :
$$\begin{aligned}
&\mathrm{moy}(\mathrm{score}(d_1), \ldots, \mathrm{score}(d_k)) - \mathrm{score}(c) \\
&= \alpha\,\mathrm{moy}(x_1, \ldots, x_k) + (\alpha+1)\,\mathrm{moy}(x_{C-k+1}, \ldots, x_C) - (\alpha+1)x_1 - \alpha x_C \\
&= (\alpha+1)\left[\frac{\alpha}{\alpha+1}\,\mathrm{moy}(x_1, \ldots, x_k) + \mathrm{moy}(x_{C-k+1}, \ldots, x_C) - x_1 - \frac{\alpha}{\alpha+1}x_C\right].
\end{aligned}$$
Pour α suffisamment grand, cette quantité est strictement positive, donc au moins un candidat d_i possède un meilleur score que c et celui-ci ne peut pas être élu. Ainsi, f ne vérifie pas **CMInf**.

Reste le cas particulier $k = C$. Supposons que f vérifie **CMInf**. Nous savons, à présent, qu'elle vérifie alors la condition énoncée pour $k = C-1$, qu'on peut reformuler ainsi :
$$x_1 + 2x_2 + \ldots + 2x_{C-1} + x_C \leq (C-1)(x_1 + x_C).$$
En ajoutant l'inégalité triviale $x_1 + x_C \leq x_1 + x_C$, on obtient $2\sum_i x_i \leq C(x_1 + x_C)$, ce qui prouve que la condition énoncée est également vérifiée pour $k = C$.

On remarque au passage que, pour prouver qu'un système de vote vérifie l'ensemble des conditions nécessaires mentionnées, il n'est pas nécessaire de tester le cas particulier $k = C$ qui découle directement de la condition pour $k = C-1$. En revanche, pour montrer qu'un système ne vérifie pas **CMInf**, il peut être commode de tester d'abord la condition pour $k = C$ si celle-ci s'avère non vérifiée.

2. Montrons qu'avec les conditions énoncées, f vérifie **CMInf**. S'il y a α électeurs sincères et β manipulateurs en faveur d'un certain candidat c, avec $\beta > \alpha$,

distinguons deux sous-groupes parmi les β manipulateurs : les α premiers d'entre eux sont chargés de donner à c un score supérieur ou égal à ceux des autres candidats et les $\beta - \alpha$ manipulateurs restants vont permettre de rendre cette inégalité stricte.

Les α premiers manipulateurs se mettent en bijection avec les électeurs sincères et chacun utilise, comme bulletin, l'ordre inverse de son alter ego, tout en faisant remonter c en première position. Les inégalités supposées par hypothèse assurent alors que les candidats adverses ont au plus le score de c et qu'au moins $j-1$ d'entre eux ont un score strictement inférieur.

Quand on ajoute un manipulateur supplémentaire, il peut donner x_1 points à c et à ces $j-1$ candidats et strictement moins aux autres, donc le candidat c possède strictement le meilleur score. Si $\beta - \alpha > 1$, les manipulateurs restants n'ont qu'à mettre c en tête pour maintenir l'inégalité stricte.

3. Montrons que la seule RPS qui vérifie **CMIgn** est l'uninominal. Quitte à translater les poids par une constante et à les multiplier par une constante strictement positive, on peut supposer $x_1 = 1$ et $x_C = 0$.

Considérons $\frac{V}{2} + \varepsilon$ électeurs qui souhaitent faire gagner c et $\frac{V}{2} - \varepsilon$ électeurs qui votent ensuite et dont le seul but est de les en empêcher, avec $\varepsilon > 0$. Pour faire élire c, ses partisans ont intérêt à le mettre en tête, et les autres ont intérêt à le mettre en dernier. On a, ainsi, $\text{score}(c) = \frac{V}{2} + \varepsilon$.

Chaque partisan de c distribue au moins x_2 points aux candidats adversaires de c, donc au moins $\frac{x_2}{C-1}$ points en moyenne à chacun d'entre eux. Parmi ces candidats, soit d celui qui reçoit le plus de points des partisans de c. Pour empêcher la victoire de c, les électeurs opposants ont intérêt à mettre d en tête de bulletin. On a alors :

$$\text{score}(d) \geq \left(\frac{V}{2} + \varepsilon\right) \frac{x_2}{C-1} + \left(\frac{V}{2} - \varepsilon\right).$$

Pour que f vérifie **CMIgn**, il faut que même pour ε négligeable devant V, on ait $\text{score}(d) < \text{score}(c)$. Par passage à la limite $\varepsilon \to 0$, on obtient après simplification $x_2 \leq 0$, donc $x_2 = \ldots = x_C = 0$ (car les poids sont toujours supposés décroissants, cf. définition 1.33). Ainsi, f est l'uninominal.

Réciproquement, l'uninominal vérifie **CMIgn** et même **FavMaj** (proposition 3.10).

4. Pour que f vérifie **rCond**, il est nécessaire qu'elle vérifie **FavMaj** donc, d'après le point précédent, qu'il s'agisse de l'uninominal. Or on sait (proposition 3.10) que l'uninominal vérifie **rCond** ssi $C \leq 4$.

5. Enfin, c'est une propriété classique que f ne vérifie pas **Cond** (sauf si $C \leq 2$). Une preuve possible est la suivante. Pour que f vérifie **Cond**, elle doit vérifier **FavMaj**, donc f est nécessairement l'uninominal. Or on sait (proposition 3.10) que l'uninominal ne vérifie par **Cond** (sauf dans le cas trivial $C \leq 2$). □

Le fait que la seule RPS vérifiant **FavMaj** soit l'uninominal est déjà prouvé par Lepelley et Merlin (1998). Le point 3 montre qu'en fait, c'est la seule RPS qui vérifie le critère **CMIgn** (*a priori* moins exigeant).

Afin de compléter la proposition 3.13 sur les RPS, l'idéal serait de trouver une condition nécessaire et suffisante pour qu'une RPS vérifie **CMInf**. D'après nos recherches en cours sur le sujet, nous conjecturons qu'il est possible de trouver une condition suffisante dont la forme se rapproche de celle de la condition nécessaire que nous avons donnée. Cette question est laissée ouverte pour des travaux futurs.

De la proposition 3.13, on déduit immédiatement les deux corollaires suivants pour la méthode de Borda et pour Veto.

Corollaire 3.14

1. *La méthode de Borda vérifie* **CMInf**.
2. *Elle ne vérifie pas* **CMIgn** *(sauf dans le cas trivial $C \leq 2$).*

Corollaire 3.15

Veto ne vérifie pas **CMInf** *(sauf dans le cas trivial $C \leq 2$).*

Dans les deux cas, il est possible de prouver le résultat directement, sans recourir à la proposition 3.13. Pour Veto, on peut ainsi donner une preuve directe particulièrement concise. Soit $C \geq 3$ et $V = 7$. Supposons que 3 électeurs sincères votent contre le candidat c. Alors c a plus de vetos que la moyenne (qui est $\frac{V}{C}$), donc il ne peut pas être élu, quels que soient les bulletins des 4 autres électeurs (malgré le fait que ceux-ci forment une majorité stricte).

3.4.4 RPS itérées à élimination simple

Nous avons déjà étudié le cas de VTI dans la proposition 3.10. Nous allons, à présent, généraliser ce résultat aux RPSI-ES, qui comprennent notamment les méthodes de Baldwin et Coombs.

Proposition 3.16

Soit f une RPSI-ES, de vecteurs de poids $(\mathbf{x}^k)_{k \leq C}$.

1. *f vérifie* **BulMajUni**.
2. *f vérifie* **FavMaj** *ssi pour tout $k \in [\![3, C]\!]$:*

$$\text{moy}(x_1^k, x_k^k) \geq \text{moy}(x_1^k, \ldots, x_k^k).$$

   Si cette condition est vérifiée et si $C \leq 4$, alors f vérifie aussi **rCond**.

3. *f vérifie* **Cond** *ssi c'est la méthode de Baldwin.*

La condition énoncée pour qu'une RPSI-ES vérifie **FavMaj** est prouvée par Lepelley et Merlin (1998). La condition pour vérifier **Cond** est due à Smith (1973). Nous allons tout de même en mentionner une esquisse de démonstration afin de donner une vue d'ensemble de la proposition 3.16.

Pour prouver ces résultats et, plus tard, pour étudier les RPSI-EM, nous utiliserons plusieurs fois le lemme suivant.

Lemme 3.17 (Smith)

Soit f une RPS (pour C fixé et V variable).

Si f est la méthode de Borda, alors un vainqueur de Condorcet a nécessairement un score strictement supérieur au score moyen.

Sinon, un vainqueur de Condorcet peut avoir un score qui est strictement le plus faible.

La première affirmation ci-dessus est indiquée pour mémoire : elle découle du fait que le score de Borda d'un candidat est la somme des scores de sa ligne dans la matrice des duels. La seconde est due à Smith (1973).

Le lemme suivant montre un résultat du même type pour le favori majoritaire.

Lemme 3.18

Soit f une RPS de vecteur de poids $\mathbf{x}$ (avec $C \geq 2$).

Considérons la condition suivante :

$$\text{moy}(x_1, x_C) \geq \text{moy}(x_1, \ldots, x_C).$$

Si elle est vérifiée, alors un favori majoritaire a nécessairement un score strictement supérieur au score moyen.

Dans le cas contraire, un favori majoritaire peut avoir un score qui est strictement le plus faible.

Démonstration. Le score minimal qu'un favori majoritaire c peut atteindre est $\frac{V}{2}(x_1 + x_C) + \varepsilon(x_1 - x_C)$, où $\varepsilon > 0$ peut être rendu négligeable devant V. Et le score moyen est $V \operatorname{moy}(x_1, \ldots, x_C)$. Pour que le score de c soit toujours strictement supérieur au score moyen, il faut et il suffit que $\frac{1}{2}(x_1 + x_C) \geq \operatorname{moy}(x_1, \ldots, x_C)$.

Si la condition n'est pas vérifiée, remarquons qu'il est possible que le score d'un favori majoritaire c soit non seulement inférieur ou égal, mais même strictement inférieur au score moyen (en prenant ε suffisamment petit). Considérons un tel profil. Quitte à en réaliser $C - 1$ copies par permutation circulaire des autres candidats et à joindre les profils obtenus, les adversaires ont tous le même score, alors que c a un score inférieur à la moyenne : il a donc strictement le moins bon score. Or, c est toujours favori majoritaire dans le profil obtenu. □

Enfin, pour $C \leq 4$, le lemme suivant montre que la même condition permet d'avoir un résultat similaire pour un vainqueur de Condorcet résistant.

Lemme 3.19

Soit f une RPS de vecteur de poids $\mathbf{x}$.

Considérons la condition suivante :

$$\operatorname{moy}(x_1, x_C) \geq \operatorname{moy}(x_1, \ldots, x_C).$$

Si elle est vérifiée et si $C \leq 4$, alors un vainqueur de Condorcet résistant a nécessairement un score strictement supérieur au score moyen.

Si elle n'est pas vérifiée (quel que soit C), alors un vainqueur de Condorcet résistant peut avoir un score qui est strictement le plus faible[5].

Démonstration. La deuxième affirmation est la plus facile : si la condition n'est pas vérifiée, alors d'après le lemme 3.18, un favori majoritaire (qui est en particulier vainqueur de Condorcet résistant) peut avoir un score qui est strictement le plus faible. Quant à la première affirmation, si $C \leq 3$, elle découle également du lemme 3.18 car les notions de favori majoritaire et de vainqueur de Condorcet résistant sont alors équivalentes.

Il nous reste à prouver la première affirmation pour $C = 4$. Quitte à soustraire x_4 de chaque composante du vecteur de poids, on peut supposer $x_4 = 0$. Traduisons la condition vérifiée par hypothèse :

$$\frac{x_1 + x_2 + x_3 + 0}{4} \leq \frac{x_1 + 0}{2},$$

c'est-à-dire $x_2 + x_3 \leq x_1$.

Pour v un électeur et $\{d, e\}$ un doubleton de candidats distincts, on dit que $(v, \{d, e\})$ rapporte une *suprématie* à c ssi v place simultanément c devant d et e. Pour être vainqueur de Condorcet résistant, il est nécessaire d'avoir strictement plus de $\frac{3}{2}V$ suprématies, puisqu'il faut une majorité absolue contre chaque doubleton d'adversaires.

Si un électeur place c en tête, cela rapporte trois suprématies à c : une sur le doubleton d'adversaires $\{d_1, d_2\}$, une sur $\{d_1, d_3\}$ et une sur $\{d_2, d_3\}$. Si un électeur place c en deuxième position, cela lui rapporte une seule suprématie,

5. Pour que la condition ne soit pas vérifiée, on pourra remarquer qu'il est nécessaire d'avoir $C \geq 3$.

contre les deux candidats moins bien placés que lui. Si un électeur place c en troisième ou quatrième position, cela ne lui rapporte aucune suprématie. Ainsi, le « prix unitaire » d'une suprématie (en points de score marqués) est $\frac{1}{3}x_1$ quand c est en première position et x_2 quand c est en seconde position.

Cas 1 Supposons $\frac{1}{3}x_1 \leq x_2$. Alors, pour être vainqueur de Condorcet résistant en marquant le moins de points possible, l'optimum est d'être en première position pour une majorité stricte d'électeurs (jamais en deuxième position) ; autrement dit, d'être vainqueur majoritaire. Donc, d'après le lemme 3.18, le score de c est strictement supérieur au score moyen.

Cas 2 Supposons $x_2 < \frac{1}{3}x_1$. À présent, pour être vainqueur de Condorcet résistant en marquant le moins de points possibles, l'optimum est d'être en deuxième position aussi souvent que possible. Cependant, être toujours deuxième ne suffit pas, car cela ne donne que V suprématies (et non les $\frac{3}{2}V$ suprématies nécessaires). Pour avoir assez de suprématies, il est donc nécessaire d'être en première position chez $\frac{V}{4} + \varepsilon$ électeurs, pour un certain $\varepsilon > 0$. Avec cette notation, il faut être en deuxième position chez au moins $\frac{3V}{4} - 3\varepsilon$ électeurs (il reste donc, d'ailleurs, 2ε électeurs qui peuvent faire n'importe quoi dans leur coin). On a alors, en notant moy(score) le score moyen de tous les candidats :

$$\begin{aligned}\text{score}\,(c) - \text{moy(score)} &\geq \frac{V}{4}x_1 + \frac{3V}{4}x_2 + \varepsilon\,(x_1 - 3x_2) - V\,\text{moy}(x_1, \ldots, x_4) \\ &\geq \frac{V}{4}x_2 + \frac{V}{4}(x_2 - x_3) + \varepsilon\,(x_1 - 3x_2)\,.\end{aligned}$$

Or les deux premiers termes sont positifs par définition d'une RPS et le dernier est strictement positif par hypothèse. Donc c a un score strictement supérieur au score moyen. □

La première affirmation du lemme 3.19 n'est pas généralisable au cas $C \geq 5$: dans ce cas, un vainqueur de Condorcet résistant peut même avoir un score strictement inférieur au score moyen. En effet, considérons l'uninominal : $\mathbf{x} = (1, 0, \ldots, 0)$. La condition $\text{moy}(x_1, x_C) \geq \text{moy}(x_1, \ldots, x_C)$ est alors clairement vérifiée. Cependant, nous avons déjà montré dans le lemme 3.11 qu'un vainqueur de Condorcet résistant peut avoir un score en uninominal qui est strictement le plus faible.

Nous disposons à présent de tous les éléments pour prouver la proposition 3.16 sur les RPSI-ES.

Démonstration. 1. Montrons que f vérifie **BulMajUni**. Si une coalition majoritaire souhaite faire gagner le candidat c, il leur suffit de placer c en tête de bulletin et tous les autres candidats dans un même ordre arbitraire. Au tour de dépouillement où il reste k candidats en lice, notons d_k le candidat placé dernier par ces partisans de c. Même si tous autres les électeurs placent d_k en premier et c en dernier sur leurs bulletins, on a $\text{score}_k(c) > \frac{V}{2}(x_1^k + x_k^k) > \text{score}_k(d_k)$, donc c ne peut pas être éliminé. Il est donc vainqueur de l'élection.

2. Du lemme 3.18, il découle immédiatement que f vérifie **FavMaj** ssi pour tout $k \in [\![3, C]\!]$, $\text{moy}(x_1^k, x_k^k) \geq \text{moy}(x_1^k, \ldots, x_k^k)$. Si cette condition est vérifiée et si $C \leq 4$, le lemme 3.19 nous apprend qu'un vainqueur de Condorcet résistant ne peut pas être éliminé, donc f vérifie **rCond**.

3. Le fait que la seule RPSI-ES qui vérifie **Cond** soit la méthode de Baldwin est dû à Smith (1973) : c'est une conséquence directe du lemme 3.17. □

Afin de compléter la proposition 3.16 sur les RPSI-ES, il reste donc à trouver un critère général pour **rCond**, mais il nous semble compromis d'obtenir une relation simple valide pour tout nombre de candidats. On entrevoit le problème dans la démonstration du lemme 3.19 : plus il y a de candidats, plus il y a de façons possibles de gagner des suprématies. En fonction des valeurs des poids, il peut être plus économique (en score) d'être souvent placé en première position, en deuxième, etc. De plus, s'il est plus économique d'être souvent en k-ième position pour un certain $k > 1$, il est possible qu'on soit obligé de compléter avec des électeurs qui placent le candidat dans des positions plus hautes du bulletin, comme nous l'avons fait dans le second cas de la preuve. Il y a, alors, une sous-distinction de cas à faire en fonction du prix unitaire de la suprématie dans ces différentes positions. On peut donc craindre qu'il faille distinguer tous les sous-ensembles non vides des $C-2$ premières positions comme réservoirs potentiels de suprématies, ce qui peut conduire, dans le pire des cas, à $2^{C-2}-1$ inégalités.

On sait, en tout cas, que pour $C \geq 5$, il est possible pour une RPSI-ES de vérifier **FavMaj** sans vérifier **rCond**, puisque c'est le cas de VTI (proposition 3.10).

De la proposition 3.16, il découle immédiatement le corollaire suivant pour la méthode de Coombs.

Corollaire 3.20

1. *La méthode de Coombs vérifie* **BulMaj**.
2. *Elle ne vérifie pas* **FavMaj** *(sauf dans le cas trivial $C \leq 2$).*

De même que pour le corollaire 3.15, on peut prouver le dernier résultat de façon plus concise qu'en recourant à la proposition 3.16. Même s'il est classique que la méthode de Coombs ne vérifie pas **FavMaj**, voici un contre-exemple pour mémoire.

2	2	3
a	a	b
b	c	c
Autres	Autres	Autres
c	b	a

Le candidat a est favori majoritaire mais il est éliminé dès le premier tour.

3.4.5 RPS itérées à élimination par la moyenne

Nous avons déjà étudié le cas de VTIM dans la proposition 3.10 et nous allons maintenant étudier les RPSI-EM en général, qui comprennent notamment les méthodes de Nanson et Kim-Roush. Pour cela, nous allons pouvoir exploiter certains lemmes démontrés pour traiter le cas des RPSI-ES.

Proposition 3.21

Soit f_C une RPSI-EM, de vecteurs de poids $(\mathbf{x}^k)_{k\leq C}$. Pour tout $C' < C$, on note $f_{C'}$ la RPSI-EM de vecteurs de poids $(\mathbf{x}^k)_{k\leq C'}$.

1. *Les conditions suivantes sont équivalentes.*
   (a) *Pour tout $C' \leq C$, $f_{C'}$ vérifie* **CMInf**.
   (b) *Pour tout $k \in [\![3, C]\!]$, on a :*
   $$\mathrm{moy}(x_1^k, x_k^k) \geq \mathrm{moy}(x_1^k, \ldots, x_k^k).$$
   Quand elles sont vérifiées, chaque $f_{C'}$ vérifie également **FavMaj** *et, pour $C' \leq 4$,* **rCond**.
2. *f_C vérifie* **Cond** *ssi c'est la méthode de Nanson.*

On remarquera que, si on affaiblit la condition 1a en exigeant seulement que f_C vérifie **CMInf**, alors la condition 1b n'est plus nécessaire. En effet, soit $C = 4$, $\mathbf{x}^4 = (1, 0, 0, 0)$ (uninominal) et $\mathbf{x}^3 = (1, 1, 0)$ (Veto). La condition 1b n'est donc pas vérifiée pour $k = 3$.

Nous allons montrer que f vérifie cependant **FavMaj**, donc **CMInf**. Supposons qu'une majorité stricte d'électeurs place c en tête de bulletin. Comme il y a 4 candidats au début du dépouillement, un candidat doit avoir au moins $\frac{V}{4}$ voix pour passer au tour suivant : comme c a une majorité stricte de voix, au plus un adversaire de c peut atteindre ce seuil, donc au moins deux adversaires sont éliminés dès le premier tour. Par conséquent, il n'y a pas de tour pour $k = 3$ et c est assuré de la victoire. Ainsi, f_C vérifie **FavMaj** et les critères plus faibles, en particulier **CMInf**, malgré le fait que la condition 1b de la proposition précédente n'est pas vérifiée.

Démonstration. non 1b $\Rightarrow$ non 1a : soit k tel que $\text{moy}(x_1^k, x_k^k) < \text{moy}(x_1^k, \dots, x_k^k)$. Considérons le système de vote $f_{C'}$, avec $C' = k$. Supposons que $\frac{V}{2} - \varepsilon$ électeurs sincères placent le candidat c en dernière position, avec $\varepsilon > 0$. Pour éviter une élimination immédiate de c, les $\frac{V}{2} + \varepsilon$ manipulateurs ne peuvent faire mieux que le mettre en tête de leur bulletin. On a alors, en notant $\text{moy}(\text{score}_k)$ la moyenne des scores de tous les candidats :

$$\begin{aligned}\text{moy}(\text{score}_k) - \text{score}_k(c) &= V\,\text{moy}(x_1^k, \dots, x_k^k) - \left[\left(\frac{V}{2} - \varepsilon\right) x_k^k + \left(\frac{V}{2} + \varepsilon\right) x_1^k\right] \\ &= V\big[\text{moy}(x_1^k, \dots, x_k^k) - \text{moy}(x_1^k, x_k^k)\big] - \varepsilon\big[x_1^k - x_k^k\big],\end{aligned}$$

Si ε est suffisamment petit devant V, alors la quantité ci-dessus est strictement positive donc c est éliminé. Ainsi, $f_{C'}$ ne vérifie pas **CMInf**.

1b $\Rightarrow$ 1a : grâce au lemme 3.18, on sait que la condition 1b implique qu'un favori majoritaire a toujours un score strictement supérieur à la moyenne, donc il ne peut pas être éliminé. Ainsi, chaque $f_{C'}$ vérifie **FavMaj** donc les critères plus faibles, en particulier **CMInf**.

Supposons, à présent, que ces conditions sont vérifiées. Pour $C' \leq 4$, le lemme 3.19 assure qu'un vainqueur de Condorcet résistant a toujours un score strictement supérieur à la moyenne, donc il ne peut pas être éliminé. Ainsi, $f_{C'}$ vérifie **rCond**.

2. Le fait que la seule RPSI-EM qui vérifie **Cond** soit la méthode de Nanson est une conséquence directe du lemme 3.17 dû à Smith (1973). □

Il reste donc à trouver un critère général pour **rCond** mais, comme pour les RPSI-ES, il nous semble compromis d'obtenir une relation simple valide pour tout nombre de candidats C. Là encore, on sait en tout cas qu'il est possible pour une RPSI-EM de vérifier **FavMaj** sans vérifier **rCond**, puisque c'est le cas de VTIM pour $C \geq 5$ (proposition 3.10).

Pour terminer notre étude des RPSI-EM, examinons le cas particulier de la méthode de Kim-Roush.

Proposition 3.22

La méthode de Kim-Roush ne vérifie pas **CMInf** *(sauf dans le cas trivial $C \leq 2$).*

Il ne s'agit pas à proprement parler d'un corollaire de la proposition 3.21 (qui permet seulement de conclure que pour $C \geq 3$ donné, il existe $C' \leq C$ tel que la méthode de Kim-Roush pour C' candidats ne vérifie pas **CMInf**). Nous allons donc prouver ce résultat directement.

Démonstration. On peut utiliser le même contre-exemple que pour Veto (corollaire 3.15). Pour $C \geq 3$ et $V = 7$ électeurs, supposons que 3 électeurs votent contre un certain candidat c. Alors, quoi que fasse la majorité constituée des 4 autres électeurs, c est éliminé dès le premier tour car il reçoit plus de vetos que la moyenne, qui est $\frac{V}{C} \leq \frac{7}{3}$. □

3.4.6 Méthode de Bucklin simple ou itérée

Proposition 3.23

1. *La méthode de Bucklin vérifie* **FavMaj**.
2. *Elle vérifie* **rCond** *ssi* $C \leq 3$.
3. *Elle ne vérifie pas* **Cond** *(sauf dans le cas trivial* $C \leq 2$*).*

Démonstration. 1, 2. Il est facile de prouver les cas où les critères sont vérifiés. Pour **FavMaj**, il suffit de remarquer que lorsqu'il y a un favori majoritaire, son rang médian est 1 alors que celui des autres candidats est strictement plus élevé. Pour **rCond**, il suffit de remarquer que, pour $C \leq 3$, un vainqueur de Condorcet résistant est nécessairement un vainqueur majoritaire (dans l'espace électoral des ordres stricts totaux). À présent, montrons les cas où les critères ne sont pas vérifiés.

2. Considérons le profil suivant, où $C \geq 4$.

49	11	6	6	14	14
c	a	d_1	d_2	a	a
a	c	c	c	d_1	d_2
d_1	d_1	a	a	c	c
d_2	d_2	d_2	d_1	d_2	d_1
Autres	Autres	Autres	Autres	Autres	Autres

Le candidat c est préféré à tout couple (a, d_i) par 55 électeurs et au couple (d_1, d_2) par 60 électeurs sur 100, donc il est vainqueur de Condorcet résistant. Mais, au sens de Bucklin, on a score$(c) = (2\ ;\ 72)$ et score$(a) = (2\ ;\ 88)$, donc a est élu. Ainsi, pour $C \geq 4$, la méthode de Bucklin ne vérifie pas **rCond**.

3. Il est classique que la méthode de Bucklin ne respecte pas **Cond** (sauf dans le cas trivial $C \leq 2$) mais nous allons fournir un contre-exemple pour mémoire.

40	15	15	30
c	a	b	a
a	c	c	b
b	b	a	c
Autres	Autres	Autres	Autres

Le candidat c est vainqueur de Condorcet, mais on a score$(c) = (2\ ;\ 70)$ et score$(a) = (2\ ;\ 85)$. Donc a est élu. □

Proposition 3.24

1. *La méthode de Bucklin itérée vérifie* **FavMaj**.
2. *Elle vérifie* **rCond** *ssi* $C \leq 4$.
3. *Elle ne vérifie pas* **Cond** *(sauf dans le cas trivial* $C \leq 2$*).*

Pour prouver cette proposition, nous allons recourir au lemme suivant.

Lemme 3.25

On se place dans l'espace électoral des ordres stricts totaux avec $C \leq 4$ et on suppose qu'un certain candidat c est vainqueur de Condorcet résistant.

Alors il existe un candidat d dont le rang médian (au sens de Bucklin) est strictement supérieur à celui de c. En particulier, le score de d au sens de Bucklin est strictement moins bon que celui de c.

Démonstration. Si $C \leq 3$, cela découle du fait qu'un vainqueur de Condorcet résistant est nécessairement un vainqueur majoritaire (dans l'espace électoral des ordres stricts totaux) : son rang médian est donc 1 et c'est le seul candidat possédant cette propriété.

Si $C \leq 4$, notons les candidats c, d_1, d_2 et d_3. Comme c est vainqueur de Condorcet résistant, il est préféré par une majorité stricte d'électeurs à d_1 et d_2 simultanément : cette majorité d'électeurs place donc c au rang 1 ou 2. Par conséquent, le rang médian de c au sens de Bucklin est au pire 2 (c'est-à-dire 1 ou 2).

Supposons qu'aucun candidat n'ait un rang médian strictement supérieur. Pour cela, il est nécessaire que chacun des 4 candidats (y compris c) occupe strictement plus de $\frac{V}{2}$ positions dans les rangs 1 ou 2 des V électeurs : il y a donc strictement plus de $2V$ pigeons pour $2V$ tiroirs à chaussettes, ce qui est absurde. □

Nous pouvons maintenant prouver la proposition 3.24.

Démonstration. 1, 2. Il est clair que la méthode de Bucklin itérée vérifie **FavMaj**. Le lemme 3.25 prouve que pour $C \leq 4$, elle vérifie **rCond**. Montrons, à présent, les cas où les critères ne sont pas vérifiés.

2. Pour $C = 5$, notons les candidats $\{c, d_1, \ldots, d_4\}$. Posons $\alpha = 18$, $\beta = 4$ et $\gamma = 15$ et considérons le profil suivant. Pour la première colonne, par exemple, notre notation signifie que, pour toute permutation σ des entiers de 1 à 4, il y a α électeurs qui préfèrent c puis $d_{\sigma(1)}$, $d_{\sigma(2)}$, $d_{\sigma(3)}$ et enfin $d_{\sigma(4)}$. Pour l'ensemble de la première colonne, il y a donc $4! \times \alpha = 24\alpha$ électeurs qui placent c en première position. Au total, il y a $24(\alpha + \beta + \gamma) = 888$ électeurs.

24α	24β	24γ
c	•	•
•	•	•
•	c	•
•	•	•
•	•	c

Le candidat c est préféré à toute paire (d_i, d_j) par $24\alpha + 4\beta = 448$ électeurs donc il est vainqueur de Condorcet résistant. Au sens de Bucklin, on a $\text{score}(c) = (3\ ;\ 24\alpha + 24\beta) = (3\ ;\ 528)$ et pour tout autre candidat d_i, on a $\text{score}(d_i) = (3;\ 12\alpha + 12\beta + 18\gamma) = (3;\ 534)$. Donc c est éliminé. Ainsi, la méthode de Bucklin itérée ne vérifie pas **rCond**.

3. Pour $C = 3$, montrons que la méthode de Bucklin itérée ne vérifie pas **Cond**.

24	24	4	4	22	22
c	c	d_1	d_2	d_1	d_2
d_1	d_2	c	c	d_2	d_1
d_2	d_1	d_2	d_1	c	c

Le candidat c est préféré à tout autre candidat d_i par 52 électeurs, donc c est vainqueur de Condorcet. Mais, au sens de Bucklin, on a score$(c) = (2\ ;\ 56)$ et score$(d_i) = (2\ ;\ 72)$, donc c est éliminé.

Pour adapter ces contre-exemples à davantage de candidats, il suffit d'ajouter des candidats additionnels dans un ordre commun à la fin des préférences de chaque électeur : ces candidats fantoches seront éliminés dans les premiers tours de dépouillement et on sera ramené aux contre-exemples exposés ci-dessus. □

3.5 Aspect informationnel des critères majoritaires

Pour conclure ce chapitre, nous proposons une interprétation qualitative des critères que nous avons étudiés. Il s'agit ici d'une discussion informelle, dont le but est simplement de proposer une expérience de pensée permettant d'appréhender les problèmes d'information et de communication qui peuvent survenir pour voter stratégiquement dans un système de vote donné et potentiellement trouver un ENF ou, en tout cas, parvenir au même candidat élu que si on avait trouvé un ENF.

Pour simplifier, considérons un espace électoral où il n'existe pas de configuration semi-Condorcet, par exemple l'espace électoral des ordres stricts totaux avec un nombre impair d'électeurs : ceci permet d'identifier les critères **RENFA** et **RENFC** d'une part, et **XENFA** et **XENFC** d'autre part.

Nous nous limitons aux systèmes de vote respectant **CMInf**, puisque ceux-ci semblent favorisés par la pratique dans les cas d'application où l'on souhaite une certaine égalité entre les électeurs d'une part, et entre les candidats d'autre part (non pas que nous négligions l'intérêt d'autres types de systèmes, que nous étudierons dans le prochain chapitre). Avec cette hypothèse, nous avons vu que le critère **RENFA**, en l'occurrence équivalent à **RENFC**, est également vérifié. Donc les seules configurations ω susceptibles de posséder ENF sont les configurations Condorcet, et le seul vainqueur possible d'un ENF est le vainqueur de Condorcet.

Enfin, nous supposons que l'espace électoral est fini, ce qui assure l'existence d'un système de vote de manipulabilité minimale (au sens ensembliste) au sein de toute classe de systèmes de vote, et en particulier de la classe **CMInf** [6].

L'objectif de notre expérience de pensée est d'illustrer le fait suivant : plus le système vérifie de critères dans le diagramme d'inclusion de la figure 3.1, plus on progresse sur les aspects suivants :

— il existe plus souvent un ENF ;
— il est plus facile d'atteindre le même résultat qu'un ENF (c'est-à-dire d'élire le vainqueur de Condorcet s'il existe) ;
— il est plus facile d'atteindre un ENF.

Quand nous employons l'expression « facile », c'est du point de vue de la quantité d'information à échanger entre les agents : plus elle est élevée, moins la tâche est facile.

Voici notre expérience de pensée, dont les conclusions sont résumées dans la table 3.1. Imaginons qu'un coordinateur externe souhaite aider les électeurs à trouver le seul résultat possible d'un éventuel ENF (c'est-à-dire le vainqueur

6. Nous verrons dans le chapitre 5 une autre condition suffisante importante qui garantit l'existence d'un optimum (au sens probabiliste cette fois) dans certaines cultures particulières, même dans des espaces électoraux infinis, typiquement ceux que l'on utilise pour certains systèmes de vote cardinaux.

	CMInf	CMIgn	XENFC	BulMaj	FavMaj	rCond	Cond
Propriétés des ENF							
Maximise l'ensemble des ω possédant un ENF			✓	✓	✓	✓	✓
Il existe peut-être un système de manipulabilité minimale						✓	✓
Il existe un système de manipulabilité minimale							✓
Tout vainqueur d'ENF coïncide avec le vainqueur sincère							✓
Problématiques d'information et de communication							
Candidat cible c	✗	✗	✗	✗	✗	✗	
Assignation des bulletins	✗	✗	✗	✗			
Membres de la coalition	✗	✗	✗				
Bulletins des autres électeurs	✗						

TABLE 3.1 – Aspect informationnel des critères majoritaires et des critères d'équilibre. Chaque colonne illustre les propriétés de l'ensemble des systèmes de vote qui vérifient un critère mais ne vérifient pas les critères plus exigeants mentionnés.

de Condorcet avec nos hypothèses). On suppose qu'il existe une majorité stricte d'électeurs, que nous appelons ici *la* coalition, qui souhaitent collaborer avec le coordinateur pour réaliser cet objectif. En utilisant le critère **CMInf**, on peut imaginer le protocole suivant, où on exploite quelques dons surnaturels du coordinateur.

1. Le coordinateur a le pouvoir magique de connaître les préférences des électeurs, ce qui lui permet d'identifier quel candidat c est le vainqueur de Condorcet. C'est donc le candidat cible qu'il s'agit de faire gagner.
2. Les membres de la coalition lui envoient un message pour l'informer de leur identité et du fait qu'ils vont suivre ses instructions.
3. Le coordinateur a le pouvoir magique de connaître les bulletins des autres électeurs.
4. En exploitant **CMInf**, le coordinateur détermine des bulletins pour les membres de la coalition qui permettent de faire gagner c et il leur diffuse cette information.

Pour effectuer ce protocole, il y a quatre aspects d'information à gérer : d'une part, le coordinateur doit connaître le candidat cible c, l'identité des membres de la coalition et les bulletins des autres électeurs ; d'autre part, il doit diffuser aux membres de la coalition l'assignation de leurs bulletins.

Bien sûr, nous ne prétendons pas qu'un tel déroulement soit possible en pratique. Le but de cette expérience de pensée est simplement d'illustrer à quelles problématiques d'information et de communication peut se confronter une population d'électeurs stratèges à la recherche d'un équilibre de Nash fort ou, au moins, cherchant à élire le même candidat que s'ils avaient atteint un équilibre.

Voyons maintenant en quoi la situation s'améliore si on exige des systèmes de vote vérifiant des critères plus forts que **CMInf**. Les réflexions ci-dessous sont récapitulées dans la table 3.1.

Si le système de vote vérifie **CMIgn**, alors le coordinateur n'a plus besoin de connaître les bulletins des électeurs non membres de la coalition.

Si le système de vote vérifie **XENFC**, les échanges d'information sont *a priori* les mêmes (partie inférieure de la table). Mais le système maximise l'ensemble des configurations de préférence où un ENF existe, puisqu'on garantit que toute situation Condorcet possède un ENF et que le critère **CMInf**, équivalent à **RENFC** sous nos hypothèses, impose que ce soit les seules. Ainsi, **XENFC** ne modifie pas le protocole utilisé, mais maximise l'ensemble des configurations de préférence où il permet d'atteindre un résultat qui est le même que celui d'un ENF.

Si le système de vote vérifie **BulMaj**, alors quitte à diffuser une assignation de bulletins à tous les électeurs et non seulement aux membres de la coalition, le coordinateur n'a pas besoin de savoir à l'avance quels sont les électeurs qui vont obéir à ses consignes. En pratique, tous les modes de scrutin courants qui vérifient **BulMaj** vérifient également **BulMajUni** (par exemple les RPSI-ES, comme la méthode de Coombs), donc le coordinateur peut se contenter de *broadcaster* un seul bulletin et ne pas recourir à des messages individualisés pour chaque électeur. C'est le cas, par exemple, du vote par assentiment, pour lequel Laslier (2009) montre, par une approche différente, qu'on peut atteindre un équilibre de façon relativement peu coûteuse.

Si le système de vote vérifie **FavMaj**, alors il est inutile de diffuser une assignation de bulletins aux électeurs, on peut se contenter de leur communiquer le nom du candidat c à la place : il suffit aux électeurs obéissants de placer c en tête de leur bulletin.

Si on se place dans la classe **rCond**, les problématiques d'information sont *a priori* les mêmes (partie inférieure de la table). Mais il n'est pas exclu qu'un de ces modes de scrutin possède une manipulabilité minimale (au sens ensembliste), alors que c'est impossible pour un mode de scrutin ne vérifiant pas **rCond**, comme l'énonce le corollaire 2.22 des théorèmes de condorcification. Un tel système de manipulabilité minimale maximise l'ensemble des situations non-manipulables, c'est-à-dire l'ensemble des situations où on peut trouver un ENF sans aucun échange d'information, simplement par le vote sincère.

Si le système de vote vérifie **Cond**, alors quand il y a un ENF, son résultat coïncide avec celui du vote sincère. Le coordinateur peut alors n'échanger aucune information avec les électeurs : s'ils votent sincèrement, ils sont toujours capables de trouver le même résultat qu'un éventuel ENF sans échange d'information.

En outre, dans la classe **Cond**, l'existence d'un système de manipulabilité minimale est non seulement possible mais garantie par le corollaire 2.22. Si on utilise un tel système optimal, alors les électeurs ont la capacité de trouver un ENF aussi souvent qu'il est possible de le faire sans échange d'information.

Chapitre 4

Condorcification généralisée

Jusqu'à présent, nous avons accordé une attention importante aux systèmes de vote qui vérifient **CMInf**, qui est un critère lié à la notion de majorité. Comme nous le discutions en introduction de cette thèse, le principe majoritaire est issu du vote majoritaire simple et, comme le formalise le théorème de May (1952), il est directement issu des principes d'anonymat et de neutralité.

Cependant, il existe des exemples d'applications pratiques où un système de vote viole l'anonymat, la neutralité ou les deux, pour des raisons qui peuvent être défendues comme légitimes. Par exemple, une réunion de copropriétaires ou l'assemblée d'actionnaires d'une entreprise n'est généralement pas symétrique entre les électeurs. Lors d'une révision de la Constitution française, la nouvelle version doit obtenir deux tiers des voix, ce qui crée une asymétrie entre les candidats : en effet, l'autre candidat est l'ancienne version de la Constitution, qui a seulement besoin de dépasser un tiers des voix.

Dans le cas général, des critères comme **CMInf** ne sont pas nécessairement vérifiés par le système de vote utilisé et on ne peut donc pas appliquer le théorème de faible de condorcification 2.9 : on n'a donc aucune garantie de ne pas augmenter la manipulabilité en utilisant la condorcification habituelle, basée sur la notion de majorité. En revanche, on peut espérer qu'un théorème similaire soit valide avec une notion inspirée de la condorcification. Pour cela, nous utilisons une approche issue de la théorie des *jeux simples* et nous définissons la *condorcification généralisée*, puis nous explorons ses liens avec la manipulabilité.

Dans la section 4.1, nous définissons la notion de *famille*, qui est centrale pour ce chapitre. À chaque candidat c, une famille $\mathcal{M}$ associe un ensemble $\mathcal{M}_c$ de coalitions. Ce type d'objet permettra, ensuite, de décrire, pour chaque candidat c, un ensemble de coalitions qui ont le pouvoir de faire gagner c, dans un certain sens que nous choisirons (plus spécifiquement, de façon ignorante ou informée). Nous utilisons cette notion pour généraliser les notions de vainqueur de Condorcet, de candidat Condorcet-admissible et de favori majoritaire.

Dans la section 4.2, nous généralisons certains critères majoritaires en nous limitant à ceux qui seront utiles pour les théorèmes de ce chapitre ; en particulier, nous généralisons **CMInf** en un critère noté $\mathcal{M}$**CInf**. De plus, nous définissons deux critères entièrement nouveaux, $\mathcal{M}$**CInfA** et $\mathcal{M}$**CIgnA**, qui exigent qu'une coalition de $\mathcal{M}_c$ (respectivement informée ou ignorante) puisse non seulement faire gagner le candidat c, mais puisse de plus faire en sorte qu'il soit *admissible* au sens de la famille $\mathcal{M}$ considérée. Ces critères techniques nous permettront, ensuite, de composer les condorcifications généralisées utilisant des familles différentes. Par

ailleurs, comme nous l'avions fait pour les critères majoritaires, nous établissons les relations d'implication entre les critères étudiés.

Dans la section 4.3, nous définissons de façon naturelle la $\mathcal{M}$-condorcification d'un système de vote f, que nous notons $f^{\mathcal{M}}$. Il est alors facile d'adapter la démonstration du théorème de condorcification faible 2.9 pour montrer que, si un système de vote vérifie $\mathcal{M}$**CInf**, alors $f^{\mathcal{M}}$ est au plus aussi manipulable que f : c'est le *théorème de condorcification généralisée* 4.18.

Afin de ne pas alourdir ce chapitre, nous n'allons pas plus loin dans le parallèle avec le chapitre 2 : en particulier, nous ne généralisons par la notion de vainqueur de Condorcet résistant et le théorème de fort condorcification 2.20. Nous préférons plutôt nous concentrer sur les possibilités nouvelles offertes par la condorcification généralisée.

En particulier, celle-ci laisse le choix de la famille $\mathcal{M}$ employée pour définir $f^{\mathcal{M}}$. Dans la section 4.4, nous comparons les condorcifications réalisées au moyen de deux familles $\mathcal{M}$ et $\mathcal{M}'$ et nous introduisons le *théorème de condorcification comparée* 4.21 qui permet, sous certaines hypothèses, de montrer que $f^{\mathcal{M}'}$ est au plus aussi manipulable que $f^{\mathcal{M}}$.

Dans la section 4.5, nous nous posons la question naturelle suivante : parmi les familles $\mathcal{M}$ qui permettent de bénéficier du théorème de condorcification généralisée 4.18, y en a-t-il une qui diminue le plus la manipulabilité ? Pour y répondre, nous définissons la *famille maximale* de f : pour chaque candidat c, on considère toutes les coalitions qui sont capables de faire gagner c quand elles sont informées des bulletins des autres électeurs. Nous appelons *condorcification maximale* de f le système $f^{\mathcal{M}'}$, où $\mathcal{M}'$ est la famille maximale. Puis nous prouvons le *théorème de condorcification maximale* 4.25 qui énonce, sous certaines hypothèses, que parmi les condorcifications vérifiant les hypothèses du théorème 4.18, la condorcification maximale est la moins manipulable.

Dans la section 4.6, nous examinons divers exemples d'application de la condorcification généralisée. Pour les systèmes de vote classiques vérifiant **CMInf**, nous montrons que la condorcification maximale est la condorcification au sens habituel, c'est-à-dire majoritaire (section 4.6.1). Puis nous étudions la condorcification maximale de Veto. Enfin, nous étudions les condorcifications généralisées de divers modes de scrutin qui violent l'anonymat, la neutralité ou les deux.

4.1 Famille de collections de coalitions

À titre de rappel, la théorie des *jeux simples* est un formalisme utilisé pour étudier des jeux coopératifs dans un cadre abstrait et relativement simplifié. On dispose d'un ensemble de joueurs, qui, pour nous, sont des électeurs. Un *jeu simple* est défini par un ensemble de coalitions $\mathcal{M}$, c'est-à-dire de sous-ensembles d'électeurs. Si une coalition M appartient à $\mathcal{M}$, on dit que M est *gagnante* ; sinon, on dit qu'elle est *perdante*.

L'esprit de ce modèle est que, si les membres d'une coalition gagnante M parviennent à se mettre d'accord (selon une certaine modalité), alors ils sont vainqueurs du jeu, généralement au détriment des autres joueurs. Suivant l'ensemble de coalitions $\mathcal{M}$ qui définit le jeu simple étudié, on peut se poser diverses questions, comme les pouvoirs respectifs des différents joueurs et la façon dont on peut répartir les gains entre les vainqueurs (ce qui amène à définir des notions comme la *valeur de Shapley*).

Dans le cas des systèmes de vote, les coalitions qui permettent de faire gagner un candidat donné c dépendent *a priori* de c : dans l'exemple du changement de

constitution, il suffit d'un tiers strict des électeurs pour conserver l'ancienne, mais il faut deux tiers des électeurs pour adopter la nouvelle version. Ces observations conduisent à définir la notion de *famille*, qui est au centre de ce chapitre.

4.1.1 Définition et propriétés basiques d'une famille

Définition 4.1 (famille de collections de coalitions)
On appelle *famille de collections de coalitions*, ou *famille*, une fonction :

$$\mathcal{M} : \left| \begin{array}{ccc} \mathcal{C} & \to & \mathcal{P}(\mathcal{P}(\mathcal{V})) \\ c & \to & \mathcal{M}_c. \end{array} \right.$$

Intuitivement, nous conseillons de voir une famille comme un objet donnant l'information suivante : pour chaque candidat c, on décrit quelles sont les coalitions d'électeurs qui peuvent faire gagner c, dans différents sens que nous préciserons par la suite. Ainsi, pour une coalition d'électeurs $M \in \mathcal{P}(\mathcal{V})$, nous dirons que M est une *$\mathcal{M}$-coalition gagnante pour c* ssi $M \in \mathcal{M}_c$. Pour l'instant, il ne s'agit que d'une convention de langage dont on verra l'aspect commode par la suite.

Définition 4.2 (propriétés basiques d'une famille)
On dit que $\mathcal{M}$ est *monotone* ssi $\forall c \in \mathcal{C}, \forall (M, M') \in \mathcal{P}(\mathcal{V})^2$:

$$M \in \mathcal{M}_c \text{ et } M \subseteq M' \Rightarrow M' \in \mathcal{M}_c.$$

On dit que $\mathcal{M}$ est *exclusive* ssi $\forall (c, d) \in \mathcal{C}^2, \forall (M, M') \in \mathcal{P}(\mathcal{V})^2$, si $c \neq d$ alors :

$$M \in \mathcal{M}_c \text{ et } M' \in \mathcal{M}_d \Rightarrow M \cap M' \neq \varnothing.$$

On dit que $\mathcal{M}$ est *neutre* ssi $\forall (c, d) \in \mathcal{C}^2, \mathcal{M}_c = \mathcal{M}_d$.
On dit que $\mathcal{M}$ est *anonyme* ssi $\forall c \in \mathcal{C}, \forall \sigma \in \mathfrak{S}_{\mathcal{V}}, \forall M \in \mathcal{P}(\mathcal{V}) : M \in \mathcal{M}_c \Leftrightarrow \sigma(M) \in \mathcal{M}_c$, où $\sigma(M)$ désigne la coalition obtenue en considérant l'image des membres de la coalition M par une permutation σ de tous les électeurs.

La monotonie signifie que si M est une coalition gagnante pour c, alors toute coalition M' qui contient M est une coalition gagnante pour c. Avec l'interprétation que nous avons donnée, il s'agit donc d'une hypothèse assez naturelle.

L'exclusivité signifie que, si deux coalitions M et M' sont disjointes, alors elles ne peuvent pas être respectivement gagnantes pour des candidats c et d distincts. Cette notion est similaire à celle de jeu simple *propre*, qui impose qu'une coalition et son complémentaire ne puissent être toutes deux gagnantes. Si on suppose que les électeurs de M et ceux de M' disposent des mêmes pouvoirs, alors il est naturel de considérer que les deux coalitions ne peuvent pas faire gagner c et d simultanément, ce qui est le cas pour une manipulation ignorante. En revanche, s'il s'agit de manipulation informée, l'exclusivité n'est pas évidente : en effet, une coalition informée pourrait manipuler pour un certain candidat a ; puis une coalition informée disjointe, connaissant ces bulletins, pourrait changer leurs propres bulletins pour faire gagner un autre candidat b ; mais alors la première coalition, si elle connaît ces nouveaux bulletins, pourrait changer de bulletins pour faire gagner de nouveau a, etc. Nous verrons un tel exemple avec le vote de parité.

Enfin, la signification de la neutralité et de l'anonymat est évidente. En ce qui concerne l'anonymat, il est clairement équivalent de dire que l'appartenance d'une coalition M à $\mathcal{M}_c$ ne dépend que du cardinal de M.

Afin de disposer d'un exemple commode pour la suite, nous allons définir un particulier cas simple : les *familles à seuil*.

Définition 4.3 (famille à seuil)
Pour $\alpha \in [0, V]$, on appelle *famille de seuil* α la famille neutre constituée de toutes les coalitions comprenant strictement plus de α électeurs. Formellement :

$$\mathcal{M} : \left| \begin{array}{lcl} \mathcal{C} & \to & \mathcal{P}(\mathcal{P}(\mathcal{V})) \\ c & \to & \{M \in \mathcal{P}(\mathcal{V}) \text{ t.q. } \mathrm{card}(M) > \alpha\}. \end{array} \right.$$

Pour $x \in [0, 100]$, on parlera aussi de la *famille des x %* pour désigner la famille de seuil $V \cdot x$ %.
On appelle *famille majoritaire* la famille des 50 %, c'est-à-dire celle constituée de toutes les coalitions strictement majoritaires.

Il est facile de voir que les familles à seuil sont les seules familles neutres, anonymes et monotones. En outre, la famille majoritaire est exclusive ; parmi les familles à seuil, c'est la famille maximale (au sens de l'inclusion, c'est-à-dire possédant le seuil minimal) qui soit exclusive. Par exemple, la famille des 90 % (qui est exclusive également) ne contient pas une coalition de 70 % des électeurs, alors que la famille majoritaire la contient.

4.1.2 Victoire et notions Condorcet généralisées

Nous allons, à présent, reprendre rapidement les mêmes étapes que dans les chapitres précédents, où nous nous concentrions sur la famille majoritaire. Le fait d'avoir autorisé des préférences non antisymétriques a conduit, dans ce cas particulier, à adopter d'ores et déjà des définitions générales qui présentent l'avantage de se généraliser assez facilement à une famille quelconque.
Nous avons vu dans le chapitre 2, et notamment dans la section 2.5, que la relation de victoire absolue $\mathrm{P}_{\mathrm{abs}}$ était la notion conduisant naturellement aux théorèmes de condorcification, contrairement à la relation $\mathrm{P}_{\mathrm{rel}}$. Nous allons donc nous concentrer sur la généralisation de la relation $\mathrm{P}_{\mathrm{abs}}$.

Définition 4.4 ($\mathcal{M}$-victoire, $\mathcal{M}$-défaite)
Pour $\omega \in \Omega$, on note $\mathrm{P}_{\mathcal{M}}(\omega)$, ou plus simplement $\mathrm{P}_{\mathcal{M}}$, la relation binaire sur les candidats définie par $\forall (c, d) \in \mathcal{C}^2$:

$$c \,\mathrm{P}_{\mathcal{M}}\, d \Leftrightarrow \{v \in \mathcal{V} \text{ t.q. } c \,\mathrm{P}_v\, d\} \in \mathcal{M}_c.$$

Autrement dit, cette relation signifie que l'ensemble des électeurs qui préfèrent c à d forment une coalition gagnante pour c (au sens de la famille $\mathcal{M}$).
Quand cette relation est vérifiée, on dit que c a une *$\mathcal{M}$-victoire* contre d en ω, ou que d a une *$\mathcal{M}$-défaite* contre c en ω.
On dit que c a une *$\mathcal{M}$-victoire stricte* (resp. *$\mathcal{M}$-défaite stricte*) contre d en ω ssi il a une $\mathcal{M}$-victoire et pas de $\mathcal{M}$-défaite (resp. une $\mathcal{M}$-défaite et pas de $\mathcal{M}$-victoire) contre d.

Dans le cas de la famille majoritaire, nous avions vu (en conséquence de la proposition 1.25) que si les préférences sont antisymétriques, alors la relation de victoire l'est également. Ce n'est pas forcément le cas en général. En effet, considérons la famille des 40 % et le profil suivant :

45	10	45
c	c, d	d
d		c

Alors c possède une $\mathcal{M}$-victoire contre d mais elle n'est pas stricte, puisque d possède également une $\mathcal{M}$-victoire contre c. Il est cependant facile de donner une condition suffisante pour que la relation de victoire soit antisymétrique. C'est l'objet de la proposition suivante.

Proposition 4.5
Soit $\omega \in \Omega$. On suppose que :
— *La famille $\mathcal{M}$ est exclusive ;*
— *Les relations $\mathrm{P}_v(\omega_v)$ sont antisymétriques.*
Alors la relation $\mathrm{P}_{\mathcal{M}}(\omega)$ est antisymétrique. Autrement dit, entre deux candidats distincts, il ne peut pas y avoir des $\mathcal{M}$-victoires mutuelles.

Démonstration. Si deux candidats distincts c et d ont des $\mathcal{M}$-victoires mutuelles, alors, par définition :

$$\begin{cases} \{v \in \mathcal{V} \text{ t.q. } c \,\mathrm{P}_v\, d\} \in \mathcal{M}_c, \\ \{v \in \mathcal{V} \text{ t.q. } d \,\mathrm{P}_v\, c\} \in \mathcal{M}_d. \end{cases}$$

Donc, par exclusivité, $\{v \in \mathcal{V} \text{ t.q. } c \,\mathrm{P}_v\, d\} \cap \{v \in \mathcal{V} \text{ t.q. } d \,\mathrm{P}_v\, c\} \neq \varnothing$, ce qui contredit l'antisymétrie des relations P_v. □

Pour la famille majoritaire, nous avons déjà remarqué que la relation de victoire n'est pas forcément complète. D'une part, avec un nombre d'électeurs pairs, il est possible qu'exactement la moitié des électeurs préfèrent c à d et que l'autre moitié préfère d à c (même si les préférences vérifient des hypothèses assez fortes, comme d'être des ordres stricts totaux). D'autre part, si les préférences sont des ordres stricts faibles, par exemple, il est possible que 45 % des électeurs préfèrent c à d et que le même nombre d'électeurs aient l'opinion inverse : alors aucun des deux candidats n'a de victoire contre l'autre. La proposition suivante donne une condition suffisante pour que la relation de victoire soit complète.

Proposition 4.6
Soit $\omega \in \Omega$. On suppose que :
— *La famille $\mathcal{M}$ est telle que pour tout couple (c, d) de candidats distincts et pour tout couple $(M, M') \in \mathcal{P}(\mathcal{V})^2$, si $M \cup M' = \mathcal{V}$, alors $M \in \mathcal{M}_c$ ou $M' \in \mathcal{M}_d$;*
— *Les relations P_v sont complètes.*
Alors la relation $\mathrm{P}_{\mathcal{M}}$ est complète. Autrement dit, entre deux candidats distincts c et d, il ne peut pas y avoir une absence de $\mathcal{M}$-victoire.

Dans cette proposition, l'hypothèse faite sur la famille $\mathcal{M}$ est dans le même esprit que la notion de jeu simple *fort*, qui impose que parmi une coalition et sa coalition complémentaire, au moins l'une soit toujours gagnante.

Démonstration. Soit c et d deux candidats distincts. Notons $M = \{v \in \mathcal{V} \text{ t.q. } c \,\mathrm{P}_v\, d\}$ et $M' = \{v \in \mathcal{V} \text{ t.q. } d \,\mathrm{P}_v\, c\}$. Comme les relations P_v sont complètes, on a $M \cup M' = \mathcal{V}$. Par conséquent, en utilisant l'hypothèse sur $\mathcal{M}$, on a $M \in \mathcal{M}_c$ ou $M' \in \mathcal{M}_d$. □

Il découle de la proposition 4.6 que, pour toute famille contenant la famille majoritaire, si les préférences sont complètes, alors la relation de victoire $\mathrm{P}_{\mathcal{M}}$ est complète. Mais elle n'est pas forcément antisymétrique. Par exemple, pour la famille des 30 %, si les préférences sont complètes, alors dans tout duel, au moins un candidat possède une victoire contre l'autre, mais il est possible que l'inverse soit simultanément vrai.

Puisque nous avons généralisé la notion de victoire, il est immédiat de généraliser celles de vainqueur de Condorcet et de candidat Condorcet-admissible.

Définition 4.7 ($\mathcal{M}$-Condorcet, $\mathcal{M}$-admissible)
Soit $\omega \in \Omega$ et $c \in \mathcal{C}$.
On dit que c est *$\mathcal{M}$-Condorcet* en ω ssi c a une $\mathcal{M}$-victoire stricte contre tout autre candidat d, c'est-à-dire :

$$\forall d \in \mathcal{C} \setminus \{c\}, \begin{cases} \{v \in \mathcal{V} \text{ t.q. } c \, \mathrm{P}_v \, d\} \in \mathcal{M}_c, & (4.1) \\ \{v \in \mathcal{V} \text{ t.q. } d \, \mathrm{P}_v \, c\} \notin \mathcal{M}_d. & (4.2) \end{cases}$$

On dit que c est *$\mathcal{M}$-admissible* en ω ssi c n'a pas de $\mathcal{M}$-défaite, c'est-à-dire :

$$\forall d \in \mathcal{C} \setminus \{c\}, \{v \in \mathcal{V} \text{ t.q. } d \, \mathrm{P}_v \, c\} \notin \mathcal{M}_d.$$

Si la famille $\mathcal{M}$ est exclusive et avec l'hypothèse commune que les relations P_v sont antisymétriques, alors toute victoire est stricte (proposition 4.5) donc la condition (4.2) peut être omise : elle est impliquée par la condition (4.1). Dans le cas particulier de la famille majoritaire, nous avions déjà remarqué cette simplification pour la définition 1.26 du vainqueur de Condorcet.

La proposition suivante est une généralisation de la proposition 1.31, dont la preuve découle immédiatement des définitions.

Proposition 4.8
Soit $\omega \in \Omega$. Si un candidat est $\mathcal{M}$-Condorcet en ω, alors :
— *il est $\mathcal{M}$-admissible ;*
— *aucun autre candidat n'est $\mathcal{M}$-admissible.*
En particulier, s'il existe un candidat $\mathcal{M}$-Condorcet, alors il est unique.

On pourra d'ailleurs remarquer que la condition (4.2) de la définition 4.7 a précisément pour motivations principales d'assurer dans le cas général, d'une part, qu'un candidat $\mathcal{M}$-Condorcet est également $\mathcal{M}$-admissible, et d'autre part que le candidat $\mathcal{M}$-Condorcet éventuel est unique[1].

4.1.3 Candidat $\mathcal{M}$-favori

À présent, nous généralisons la notion de favori majoritaire. Celle-ci ne sera pas directement impliquée dans le théorème de condorcification généralisée 4.18 mais elle fournira un outil commode pour faire le lien entre les notions Condorcet généralisées et la manipulation. Nous allons voir que cette notion demande quelques précautions supplémentaires par rapport au cas particulier du favori majoritaire, en raison de la plus grande généralité de nos hypothèses sur la famille utilisée.

Définition 4.9 (candidat $\mathcal{M}$-favori)
Pour $\omega \in \Omega$ et $c \in \mathcal{C}$, on dit que c est *$\mathcal{M}$-favori* en ω ssi :

$$\{v \in \mathcal{V} \text{ t.q. } \forall d \in \mathcal{C} \setminus \{c\}, c \, \mathrm{PP}_v \, d\} \in \mathcal{M}_c.$$

Si on considère la famille des 40 %, alors il est clair que le $\mathcal{M}$-favori n'est pas forcément unique : en effet, 45 % des électeurs peuvent avoir un certain candidat c comme préféré, et autant d'électeurs peuvent avoir un autre candidat d comme

1. Au passage, on remarquera que dans le cas d'une famille neutre, son *nombre de Nakamura* (Nakamura, 1979) permet d'avoir une condition nécessaire et suffisante sur le nombre de candidats C pour que tout profil de préférences transitives possède au moins un candidat $\mathcal{M}$-admissible.

préféré. La proposition suivante donne une condition suffisante pour que le $\mathcal{M}$-favori soit unique et montre que, sous certaines hypothèses assez naturelles, cette condition est nécessaire.

Proposition 4.10
On considère les conditions suivantes.

1. *$\mathcal{M}$ est exclusive.*
2. *Dans toute configuration possédant un candidat $\mathcal{M}$-favori, celui-ci est unique.*

On a 1 $\Rightarrow$ 2.
Si on suppose que l'espace électoral autorise tout candidat comme préféré et que $\mathcal{M}$ est monotone, alors 2 $\Rightarrow$ 1.

Démonstration. L'implication 1 $\Rightarrow$ 2 découlant immédiatement des définitions, nous allons nous concentrer sur la réciproque non 1 $\Rightarrow$ non 2.

Supposons que $\mathcal{M}$ n'est pas exclusive. Alors il existe deux candidats distincts c et d, deux coalitions disjointes M et M', telles que $M \in \mathcal{M}_c$ et $M' \in \mathcal{M}_d$. Comme $\mathcal{M}$ est monotone, on a $\mathcal{V} \setminus M \in \mathcal{M}_d$. Comme l'espace électoral autorise tout candidat comme préféré, il existe une configuration où les membres de M affirment que c est leur préféré et ceux de $\mathcal{V} \setminus M$ affirment que d est leur préféré. Alors, les candidats c et d sont tous deux $\mathcal{M}$-favoris. □

Comme nous l'avons déjà dit dans les chapitres précédents, il est clair qu'un favori majoritaire est nécessairement un vainqueur de Condorcet. La proposition suivante, dont la preuve découle immédiatement des définitions, permet de généraliser cette constatation si la famille considérée est monotone et exclusive, tout comme l'est la famille majoritaire.

Proposition 4.11
On suppose que $\mathcal{M}$ est monotone et exclusive.
Pour $\omega \in \Omega$ et $c \in \mathcal{C}$, si c est $\mathcal{M}$-favori en ω, alors c est $\mathcal{M}$-Condorcet en ω.

4.2 Critères associés à une famille

4.2.1 Définition des critères associés à une famille

Nous allons généraliser les critères **Cond**, **FavMaj**, **CMIgn** et **CMInf** que nous avons définis pour traiter le cas particulier de la famille majoritaire. En outre, nous allons définir deux nouveaux critères, $\mathcal{M}$**CIgnA** et $\mathcal{M}$**CInfA**, dont nous montrerons, plus tard, l'intérêt pour combiner et comparer les condorcifications généralisées réalisées en utilisant des familles $\mathcal{M}$ et $\mathcal{M}'$ différentes pour un même système de vote. Afin de ne pas alourdir cette section inutilement, nous n'allons pas évoquer les généralisations des autres critères majoritaires.

Définition 4.12 (critères associés à une famille $\mathcal{M}$)
On dit que f vérifie le *critère de $\mathcal{M}$-Condorcet* ($\mathcal{M}$**Cond**) ssi, pour tout $\omega \in \Omega$ et $c \in \mathcal{C}$, si c est $\mathcal{M}$-Condorcet en ω, alors $f(\omega) = c$.

On dit que f vérifie le *critère du $\mathcal{M}$-favori* ($\mathcal{M}$**Fav**) ssi, pour tout $\omega \in \Omega$ et $c \in \mathcal{C}$, si c est $\mathcal{M}$-favori en ω, alors $f(\omega) = c$.

On dit que f vérifie le *critère de la $\mathcal{M}$-coalition ignorante* ($\mathcal{M}$**CIgn**) ssi $\forall c \in \mathcal{C}, \forall M \in \mathcal{M}_c, \exists \omega_M \in \Omega_M$ t.q. :

$$\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}, f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c.$$

On dit que f vérifie le *critère de la $\mathcal{M}$-coalition ignorante admissible* ($\mathcal{M}$**CIgnA**) ssi $\forall c \in \mathcal{C}, \forall M \in \mathcal{M}_c, \exists \omega_M \in \Omega_M$ t.q. :

$$\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}, \begin{cases} f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c, \\ c \text{ est } \mathcal{M}\text{-admissible en } (\omega_M, \omega_{\mathcal{V}\setminus M}). \end{cases}$$

On dit que f vérifie le *critère de la $\mathcal{M}$-coalition informée* ($\mathcal{M}$**CInf**) ssi $\forall c \in \mathcal{C}, \forall M \in \mathcal{M}_c, \forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}$:

$$\exists \omega_M \in \Omega_M \text{ t.q. } f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c$$

On dit que f vérifie le *critère de la $\mathcal{M}$-coalition informée admissible* ($\mathcal{M}$**CInfA**) ssi $\forall c \in \mathcal{C}, \forall M \in \mathcal{M}_c, \forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}$:

$$\exists \omega_M \in \Omega_M \text{ t.q. } \begin{cases} f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c, \\ c \text{ est } \mathcal{M}\text{-admissible en } (\omega_M, \omega_{\mathcal{V}\setminus M}). \end{cases}$$

Les critères $\mathcal{M}$**Cond**, $\mathcal{M}$**Fav**, $\mathcal{M}$**CIgn** et $\mathcal{M}$**CInf** sont des généralisations assez naturelles des critères correspondants pour la famille majoritaire. Le critère $\mathcal{M}$**CInfA** (resp. $\mathcal{M}$**CIgnA**) est un renforcement de $\mathcal{M}$**CInf** (resp. $\mathcal{M}$**CIgn**) qui exige que la coalition informée (resp. ignorante), en plus de réussir à faire gagner un certain candidat c, soit capable d'assurer que c apparaît comme $\mathcal{M}$-admissible. Pour les systèmes de vote courants, c'est une condition naturelle : en reprenant les notations de la définition, si les membres de M sont capables de faire gagner c, ils peuvent généralement le faire en produisant un bulletin où c est strictement préféré aux autres candidats ; dans ce cas, c est $\mathcal{M}$-favori dans la configuration obtenue par manipulation et, sous l'hypothèse courante où $\mathcal{M}$ est monotone est exclusive, cela implique que c est $\mathcal{M}$-Condorcet (proposition 4.11) et *a fortiori* $\mathcal{M}$-admissible.

Comme nous l'avions déjà remarqué de façon informelle, l'hypothèse d'exclusivité est naturelle quand il s'agit de manipulation ignorante. La proposition suivante formalise cette observation.

Proposition 4.13
*Si un système de vote vérifie $\mathcal{M}$**CIgn**, alors la famille $\mathcal{M}$ est exclusive.*

Démonstration. Si deux coalitions disjointes ont les pouvoir respectifs de faire gagner des candidats c et d distincts de manière ignorante, alors ils peuvent faire gagner simultanément c et d, ce qui est contradictoire avec l'unicité du résultat de f. □

En revanche, l'hypothèse d'exclusivité ne va pas de soi quand on parle de manipulation informée : quand un système de vote vérifie $\mathcal{M}$**CInf**, il est possible que la famille $\mathcal{M}$ ne soit pas exclusive.

En effet, considérons le vote de parité : s'il y a un nombre impair (resp. pair) de boules noires, alors le candidat a (resp. b) est élu. Considérons la famille neutre $\mathcal{M}$ constituée de toutes les coalitions non vides : alors le vote de parité vérifie $\mathcal{M}$**CInf**. En notant v et v' deux électeurs distincts (on suppose $V \geq 2$), les coalitions $\{v\}$ et $\{v'\}$ sont distinctes et sont toutes les deux gagnantes, donc la famille n'est pas exclusive.

Dans ce mode de scrutin, si l'électeur v vote en dernier en connaissant les autres bulletins, il peut décider du résultat ; et de même pour l'électeur v'. La non-exclusivité n'est donc pas contradictoire avec l'unicité du résultat.

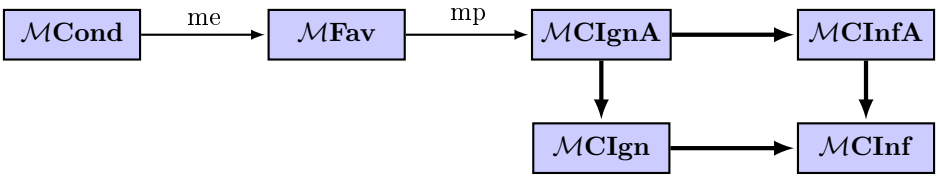


FIGURE 4.1 – Implication entre les critères liés à une famille $\mathcal{M}$ (m : famille monotone, e : famille exclusive, p : l'espace électoral autorise tout candidat comme préféré).

4.2.2 Implications entre les critères associés à une famille

La proposition suivante indique les relations d'implications entre les différents critères, c'est-à-dire les inclusions entre les ensembles de systèmes de vote correspondants. Les résultats énoncés sont résumés par le graphe d'implication de la figure 4.1.

Proposition 4.14

*Si $\mathcal{M}$ est monotone et exclusive, alors $\mathcal{M}$**Cond** $\subseteq$ $\mathcal{M}$**Fav**.*

*Si $\mathcal{M}$ est monotone et si l'espace électoral autorise tout candidat comme préféré, alors $\mathcal{M}$**Fav** $\subseteq$ $\mathcal{M}$**CIgnA**.*

On a les inclusions :

— $\mathcal{M}$**CIgnA** $\subseteq$ $\mathcal{M}$**CIgn**,

— $\mathcal{M}$**CIgnA** $\subseteq$ $\mathcal{M}$**CInfA**

— $\mathcal{M}$**CIgn** $\subseteq$ $\mathcal{M}$**CInf**,

— $\mathcal{M}$**CInfA** $\subseteq$ $\mathcal{M}$**CInf**.

Cependant, en général, on n'a :

— *ni* $\mathcal{M}$**CIgn** $\subseteq$ $\mathcal{M}$**CInfA**,

— *ni* $\mathcal{M}$**CInfA** $\subseteq$ $\mathcal{M}$**CIgn**.

Démonstration. $\mathcal{M}$**Cond** $\subseteq$ $\mathcal{M}$**Fav** : supposons qu'un système de vote vérifie $\mathcal{M}$**Cond**. Si un certain candidat c est $\mathcal{M}$-favori en ω, alors comme $\mathcal{M}$ est supposée monotone et exclusive, c est $\mathcal{M}$-Condorcet en ω (proposition 4.11), donc il est élu.

$\mathcal{M}$**Fav** $\subseteq$ $\mathcal{M}$**CIgnA** : supposons qu'un système de vote vérifie $\mathcal{M}$**Fav**. Soit $c \in \mathcal{C}$ et $M \in \mathcal{M}_c$. Il est suffisant que les membres de M affirment préférer strictement c à tous les autres candidats, ce qui est possible parce que l'espace électoral autorise tout candidat comme préféré. Quoi que fassent les autres électeurs, l'ensemble des électeurs affirmant que c est leur préféré contient M, donc il appartient à $\mathcal{M}_c$ (par monotonie), donc c est $\mathcal{M}$-favori et élu. Ceci prouve au passage que le système vérifie $\mathcal{M}$**CIgn**, donc (proposition 4.13) que la famille $\mathcal{M}$ est exclusive.

Poursuivons alors notre manipulation pour c : celui-ci est $\mathcal{M}$-favori, on sait, à présent, que la famille $\mathcal{M}$ est exclusive et elle est aussi monotone par hypothèse, donc c est $\mathcal{M}$-Condorcet (proposition 4.11) et, *a fortiori*, $\mathcal{M}$-admissible. Donc le système vérifie $\mathcal{M}$**CIgnA**.

Les autres implications découlent immédiatement des définitions.

Pour prouver qu'en général, on n'a ni $\mathcal{M}$**CIgn** $\subseteq$ $\mathcal{M}$**CInfA** ni $\mathcal{M}$**CInfA** $\subseteq$ $\mathcal{M}$**CIgn**, considérons comme contre-exemple la famille majoritaire.

Dans le corollaire 3.14 de la proposition 3.13 sur les RPS, nous avons déjà montré que la méthode de Borda vérifie $\mathcal{M}$**CInf**. Or dans la démonstration, les bulletins utilisés par les manipulateurs placent toujours le candidat c en tête de

bulletin, donc la méthode de Borda vérifie $\mathcal{M}$**CInfA**. Cependant, le corollaire 3.14 énonce aussi qu'en général, elle ne vérifie pas $\mathcal{M}$**CIgn**.

Enfin, dans l'espace électoral des ordres stricts totaux, avec $C = 2$ et V impair, considérons le système de vote suivant : le vote majoritaire inversé. Chaque électeur communique un ordre de préférence, mais le candidat a gagne ssi b en tête sur une majorité de bulletins (et vice versa). Si une coalition majoritaire souhaite faire gagner un candidat, il lui suffit de le placer toujours en bas de bulletin ; mais, par définition de ce mode de scrutin curieux, un candidat vainqueur n'apparaît jamais comme Condorcet-admissible. □

4.3 Théorème de condorcification généralisée

À présent, nous allons généraliser le théorème faible de condorcification 2.9 dans le formalisme général des familles de collections de coalitions. La démarche est essentiellement la même, c'est pourquoi nous avancerons assez rapidement.

Le lemme suivant généralise le lemme 2.4. Comme pour celui-ci, on notera qu'aucune hypothèse n'est faite sur les critères éventuels vérifiés par f.

Lemme 4.15
Soit $(\omega, \psi) \in \Omega^2$. On note $\mathrm{w} = f(\omega)$ et $c = f(\psi)$ et on suppose que f est manipulable en ω vers ψ.
On suppose que $\mathcal{M}$ est monotone.
Si w *est $\mathcal{M}$-admissible en ω, alors c ne peut pas avoir une $\mathcal{M}$-victoire contre* w *en ψ ; en particulier, c n'est pas $\mathcal{M}$-Condorcet en ψ.*

Démonstration. Comme w est $\mathcal{M}$-admissible en ω, on a : $\{v \in \mathcal{V}, c\, \mathrm{P}_v(\omega_v)\ \mathrm{w}\} \notin \mathcal{M}_c$. Mais, par définition de la manipulabilité, $\{v \in \mathcal{V}, c\ \mathrm{P}_v(\psi_v)\ \mathrm{w}\} \subseteq \{v \in \mathcal{V}, c\, \mathrm{P}_v(\omega_v)\ \mathrm{w}\}$. Donc, comme la famille $\mathcal{M}$ est monotone, $\{v \in \mathcal{V}, c\, \mathrm{P}_v(\psi_v)\ \mathrm{w}\} \notin \mathcal{M}_c$. □

Au passage, le lemme 4.15 entraîne une généralisation du lemme 2.5 : si la famille $\mathcal{M}$ est monotone et si le système de vote vérifie le critère de $\mathcal{M}$-Condorcet, alors une situation possédant un candidat $\mathcal{M}$-Condorcet ne peut pas être manipulable vers une autre situation possédant un candidat $\mathcal{M}$-Condorcet.

Le lemme suivant généralise le lemme 2.7.

Lemme 4.16
*Soit $\omega \in \Omega$. On suppose que f vérifie $\mathcal{M}$**CInf**.*
Si $f(\omega)$ n'est pas $\mathcal{M}$-admissible en ω, alors f est manipulable en ω.

Démonstration. Notons $\mathrm{w} = f(\omega)$. Comme w n'est pas $\mathcal{M}$-admissible, il existe un autre candidat c qui possède une victoire contre w :

$$\exists c \in \mathcal{C} \setminus \{\mathrm{w}\} \text{ t.q. } \{v \in \mathcal{V} \text{ t.q. } c\, \mathrm{P}_v\, \mathrm{w}\} \in \mathcal{M}_c.$$

Notons $M = \mathrm{Manip}(\mathrm{w} \to c)$ la coalition pour c contre w. Exploitant le critère $\mathcal{M}$**CInf**, on a :

$$\exists \psi_M \in \Omega_M \text{ t.q. } f(\omega_{\mathcal{V}\setminus M}, \psi_M) = c.$$

Donc f est manipulable en ω (en faveur du candidat c). □

De même que pour le lemme 2.7, on en déduit que, si une configuration ω n'est pas $\mathcal{M}$-admissible, alors tout système de vote vérifiant $\mathcal{M}$**CInf** est manipulable en ω.

Nous pouvons, à présent, généraliser la définition 2.8 de la condorcification.

Définition 4.17 ($\mathcal{M}$-condorcification)
On appelle *$\mathcal{M}$-condorcification de f* le SVBE :

$$f^{\mathcal{M}} : \left| \begin{array}{ccc} \Omega & \to & \mathcal{C} \\ \omega & \to & \left| \begin{array}{l} \text{si } \omega \text{ a un candidat } \mathcal{M}\text{-Condorcet } c, \text{ alors } c, \\ \text{sinon, } f(\omega). \end{array} \right. \end{array} \right.$$

Par définition, la $\mathcal{M}$-condorcification de f vérifie le critère de $\mathcal{M}$-Condorcet. En particulier, si la famille $\mathcal{M}$ est monotone et exclusive et si l'espace électoral autorise tout candidat comme préféré, alors elle vérifie $\mathcal{M}$**CInf** (proposition 4.14).

Dans le contexte de ce chapitre, nous appellerons *condorcification majoritaire* la condorcification f^* usuelle (définition 2.8), c'est-à-dire basée sur la famille majoritaire.

Nous avons, à présent, tous les outils pour généraliser le théorème faible de condorcification 2.9.

Théorème 4.18 (condorcification généralisée)
Soit f un SVBE et $\mathcal{M}$ une famille. On suppose que :
— *$\mathcal{M}$ est monotone ;*
— *f vérifie $\mathcal{M}$**CInf**.*
Alors $f^{\mathcal{M}}$ est au plus aussi manipulable que f :

$$\mathrm{MC}_{f^{\mathcal{M}}} \subseteq \mathrm{MC}_f.$$

Démonstration. Supposons que $f^{\mathcal{M}}$ est manipulable en ω vers ψ, mais que f n'est pas manipulable en ω.

Comme f n'est pas manipulable en ω, le lemme 4.16 assure que $f(\omega)$ est $\mathcal{M}$-admissible en ω. S'il est $\mathcal{M}$-Condorcet en ω, alors $f^{\mathcal{M}}(\omega) = f(\omega)$; sinon, il n'y a pas de candidat $\mathcal{M}$-Condorcet en ω (proposition 4.8) donc par définition de $f^{\mathcal{M}}$, on a aussi $f^{\mathcal{M}}(\omega) = f(\omega)$.

Notons $\mathrm{w} = f^{\mathcal{M}}(\omega) = f(\omega)$ et $c = f^{\mathcal{M}}(\psi)$. Comme w est $\mathcal{M}$-admissible en ω, le lemme 4.15 (appliqué à $f^{\mathcal{M}}$) assure que w n'est pas $\mathcal{M}$-Condorcet en ψ. Donc, par définition de $f^{\mathcal{M}}$, on a $f^{\mathcal{M}}(\psi) = f(\psi)$.

Par conséquent, on a $f(\omega) = f^{\mathcal{M}}(\omega)$ et $f(\psi) = f^{\mathcal{M}}(\psi)$ donc f est manipulable en ω : c'est absurde ! □

4.4 Théorème de condorcification comparée

Tout comme nous venons de généraliser le théorème faible de condorcification 2.9, on pourrait aussi généraliser le théorème fort de condorcification 2.20 en définissant une notion de candidat $\mathcal{M}$-Condorcet résistant qui généraliserait la notion de vainqueur de Condorcet résistant (définition 2.16). Nous ne développerons cependant pas ce point afin de ne pas alourdir inutilement l'exposé et nous nous concentrerons sur un point spécifique à la condorcification généralisée : le choix de la famille $\mathcal{M}$ utilisée pour condorcifier.

En effet, il est possible en général qu'un système de vote vérifie $\mathcal{M}$**CInf** pour plusieurs choix possibles de la famille $\mathcal{M}$. Dans ce cas, on peut envisager d'utiliser le théorème de condorcification généralisée 4.18 en utilisant l'une ou l'autre famille. Il se pose alors des questions naturelles : dans quels cas peut-on « condorcifier une condorcification » tout en continuant de diminuer (au sens large) la manipulabilité ? Est-ce qu'une condorcification permet de diminuer la

manipulabilité plus (ou au moins autant) qu'une autre ? Enfin, existe-t-il une famille qui permette de diminuer la manipulabilité plus (ou au moins autant) que toutes les autres ?

Sauf mention particulière, $\mathcal{M}$ et $\mathcal{M}'$ désigneront deux familles *a priori* quelconques.

Définition 4.19 (famille plus condorcifiante)
On dit que $\mathcal{M}'$ est *plus condorcifiante* que $\mathcal{M}$ (au sens large) ssi un candidat $\mathcal{M}$-Condorcet est toujours également $\mathcal{M}'$-Condorcet, c'est-à-dire :

$$\forall \omega \in \Omega, \forall c \in \mathcal{C} : c \text{ est } \mathcal{M}\text{-Condorcet en } \omega \Rightarrow c \text{ est } \mathcal{M}'\text{-Condorcet en } \omega.$$

Cette notion donne un cas particulier pour la composition de deux condorcifications : si $\mathcal{M}'$ est plus condorcifiante que $\mathcal{M}$, alors on a $(f^{\mathcal{M}})^{\mathcal{M}'} = f^{\mathcal{M}'}$. En effet, pour les configurations où $f^{\mathcal{M}}$ et f ont des vainqueurs distincts, il y a un candidat $\mathcal{M}$-Condorcet, qui est, par conséquent, $\mathcal{M}'$-Condorcet ; donc quand on condorcifie avec la famille $\mathcal{M}'$, le vainqueur est ce candidat, quel que soit le système initial.

Le lemme suivant va permettre que la première condorcification, réalisée avec la famille $\mathcal{M}$, vérifie les hypothèses qui lui permettront de bénéficier de la deuxième condorcification, réalisée avec la famille $\mathcal{M}'$. Ce lemme est la motivation principale de la définition du critère $\mathcal{M}$**CInfA**.

Lemme 4.20 (de la coalition informée admissible)
On suppose que :
— $\mathcal{M}'$ *est plus condorcifiante que* $\mathcal{M}$ *;*
— f *vérifie* $\mathcal{M}'$**CInfA**.
Alors $f^{\mathcal{M}}$ *vérifie* $\mathcal{M}'$**CInfA**.

Interprétons ce lemme à l'énoncé quelque peu obscur : on suppose que f vérifie une bonne propriété ($\mathcal{M}'$**CInfA**) pour $\mathcal{M}'$, c'est-à-dire la deuxième famille que nous allons utiliser, et on en conclut que $f^{\mathcal{M}}$ vérifie cette même propriété ($\mathcal{M}'$**CInfA**), toujours pour cette deuxième famille $\mathcal{M}'$. Intuitivement, la motivation est la suivante : dans la suite, on prendra pour $\mathcal{M}'$ une famille fixée, très condorcifiante, et il suffira de tester que f vérifie $\mathcal{M}'$**CInfA** pour cette famille fixée, donc une bonne fois pour toutes. Ensuite, la première famille $\mathcal{M}$ pourra être n'importe quelle famille moins condorcifiante que $\mathcal{M}'$: ce lemme nous permettra de montrer que, dans tous les cas, $f^{\mathcal{M}}$ vérifie la propriété qui nous permettra d'enchaîner avec la $\mathcal{M}'$-condorcification tout en continuant à diminuer la manipulabilité ; et ce, sans que nous ayons besoin de tester des conditions dépendant de $\mathcal{M}$ pour chaque famille $\mathcal{M}$ concernée. En réalité, il nous suffirait d'avoir que $f^{\mathcal{M}}$ vérifie $\mathcal{M}'$**CInf** dans la conclusion du lemme pour les applications suivantes ; le fait que $f^{\mathcal{M}}$ vérifie aussi $\mathcal{M}'$**CInfA** est offert par la maison.

Démonstration. Soit $\omega \in \Omega, c \in \mathcal{C}, M \in \mathcal{M}'_c$. Comme f vérifie $\mathcal{M}'$**CInfA**, il existe un bulletin $\psi_M \in \Omega_M$ de la coalition tel que :

$$\begin{cases} f(\omega_{\mathcal{V}\setminus M}, \psi_M) = c, \\ c \text{ est } \mathcal{M}'\text{-admissible en } (\omega_{\mathcal{V}\setminus M}, \psi_M). \end{cases}$$

Notons $d = f^{\mathcal{M}}(\omega_{\mathcal{V}\setminus M}, \psi_M)$. Si d est $\mathcal{M}$-Condorcet en $(\omega_{\mathcal{V}\setminus M}, \psi_M)$, alors comme $\mathcal{M}'$ est plus condorcifiante que $\mathcal{M}$, le candidat d est également $\mathcal{M}'$-Condorcet, donc il est le seul candidat $\mathcal{M}'$-admissible (proposition 4.8), donc

$d = c$. À l'inverse, si d n'est pas $\mathcal{M}$-Condorcet en $(\omega_{\mathcal{V}\setminus M}, \psi_M)$, alors par définition de $f^{\mathcal{M}}$, on a également $d = c$. Par conséquent, M peut faire gagner c de façon informée dans le système $f^{\mathcal{M}}$, tout en garantissant que c soit $\mathcal{M}'$-admissible. Donc $f^{\mathcal{M}}$ vérifie $\mathcal{M}'$**CInfA**. □

Les exemples que nous verrons dans la section 4.6 montrerons que les hypothèses du lemme 4.20 sont en fait assez courantes en pratique. Sous ces hypothèses, nous pouvons, à présent, comparer les condorcifications obtenues en utilisant les familles $\mathcal{M}$ et $\mathcal{M}'$. C'est l'objet du *théorème de condorcification comparée*.

Théorème 4.21 (condorcification comparée)
Soit f un SVBE, $\mathcal{M}$ et $\mathcal{M}'$ deux familles.
On suppose que :
— *$\mathcal{M}'$ est monotone ;*
— *$\mathcal{M}'$ est plus condorcifiante que $\mathcal{M}$;*
— *f vérifie $\mathcal{M}'$**CInfA**.*
Alors $f^{\mathcal{M}'}$ est au plus aussi manipulable que $f^{\mathcal{M}}$:

$$\mathrm{MC}_{f^{\mathcal{M}'}} \subseteq \mathrm{MC}_{f^{\mathcal{M}}}.$$

Démonstration. Par le lemme 4.20, on sait que $f^{\mathcal{M}}$ vérifie $\mathcal{M}'$**CInf**. Il suffit alors d'appliquer le théorème de condorcification généralisée 4.18 à $f^{\mathcal{M}}$ et $\mathcal{M}'$ et de se souvenir que $(f^{\mathcal{M}})^{\mathcal{M}'} = f^{\mathcal{M}'}$. □

Ce n'est pas évident à première vue, mais ce théorème contient implicitement l'affirmation que $f^{\mathcal{M}'}$ est moins manipulable que f (ce que nous savons par ailleurs, par le théorème de condorcification généralisée 4.18). En effet, considérons le cas particulier de la famille $\mathcal{M}$ telle que pour tout candidat c, $\mathcal{M}_c = \varnothing$. Alors un candidat est toujours $\mathcal{M}$-admissible et jamais $\mathcal{M}$-Condorcet. En particulier, toute famille est plus condorcifiante que $\mathcal{M}$ et on a $f^{\mathcal{M}} = f$. Donc la conclusion du théorème devient : $\mathrm{MC}_{f^{\mathcal{M}'}} \subseteq \mathrm{MC}_f$.

Avec seulement les hypothèses de ce théorème, on n'a pas la garantie que $f^{\mathcal{M}}$ soit moins manipulable que f : en effet, on n'a pas supposé que $\mathcal{M}$ vérifie les hypothèses du théorème de condorcification généralisée 4.18. Ceci dit, on sait l'essentiel : $f^{\mathcal{M}'}$ est au plus aussi manipulable que f et que tout système $f^{\mathcal{M}}$, où $\mathcal{M}$ est moins condorcifiante que $\mathcal{M}'$.

À la lecture de ce théorème, une question se pose immédiatement : est-ce que la conclusion est toujours vraie si on suppose seulement que f vérifie $\mathcal{M}'$**CInf** au lieu de $\mathcal{M}'$**CInfA** ? Nous allons voir que ce n'est pas le cas.

Considérons l'espace électoral des ordres avec $V = 3$ électeurs et $C = 3$ candidats nommés a, b et c. On utilise le système de vote f suivant.

1. Si tous les électeurs placent le même candidat en tête de bulletin, alors c est vainqueur.
2. Si au moins un électeur place a en tête et au moins un électeur place c en tête, alors a est vainqueur.
3. Dans tous les autres cas, b est vainqueur.

On note $\mathcal{M}$ la famille *unanime*, c'est-à-dire la famille des coalitions contenant tous les électeurs, et $\mathcal{M}'$ la famille majoritaire. Comme $\mathcal{M}'$ est une famille à seuil, elle est monotone, et il est facile de vérifier qu'elle est plus condorcifiante que $\mathcal{M}$: en effet, un candidat qui est $\mathcal{M}$-Condorcet est préféré par tous les électeurs donc il est vainqueur de Condorcet, c'est-à-dire $\mathcal{M}'$-Condorcet.

Montrons que f vérifie $\mathcal{M}'$**CInf**.

— Pour faire gagner c, il suffit que 2 manipulateurs votent comme le troisième électeur (que celui-ci soit sincère ou pas).
— Pour faire gagner a, il suffit qu'un manipulateur place a en tête et qu'un autre manipulateur place c en tête.
— Pour faire gagner b, si l'électeur sincère place b en tête, il suffit que les manipulateurs placent a et b en tête ; et si l'électeur sincère place a ou c en tête, il suffit que les manipulateurs placent tous les deux b en tête.

On remarquera que les manipulations proposées ne sont pas admissibles. Par ailleurs, la manipulation pour c crée forcément un favori unanime (préféré par tous les électeurs) et on ne peut pas toujours choisir que ce soit c : en particulier, on ne peut pas toujours faire en sorte que c soit Condorcet-admissible donc f ne vérifie pas $\mathcal{M}'$**CInfA**.

Considérons alors le profil ω suivant, que nous connaissons bien : il s'agit d'un exemple minimal de paradoxe de Condorcet.

a	b	c
b	c	a
c	a	b

Dans f, c'est le candidat a qui est vainqueur en vertu de la règle 2. Comme il n'y a ni $\mathcal{M}$-Condorcet (candidat unanimement préféré) ni $\mathcal{M}'$-Condorcet (vainqueur de Condorcet classique), le candidat a est également vainqueur dans $f^{\mathcal{M}}$ et dans $f^{\mathcal{M}'}$.

Dans $f^{\mathcal{M}'}$, c'est-à-dire la condorcification de f au sens habituel majoritaire, la configuration ω est clairement manipulable puisqu'elle est non admissible (lemme 2.7). On pourra, d'ailleurs, remarquer que la manipulation en faveur de c ne s'effectue pas de la même façon dans f et $f^{\mathcal{M}'}$. Dans f, les manipulateurs doivent mettre a en tête pour bénéficier de la règle 1. Dans $f^{\mathcal{M}'}$, ils doivent mettre c en tête pour que c soit vainqueur de Condorcet.

En revanche, dans $f^{\mathcal{M}}$ (qui utilise la famille unanime), nous allons montrer que la configuration n'est pas manipulable.

— Pour manipuler en faveur de c, les deux derniers électeurs ont besoin que c soit en tête de tous les bulletins, ce qu'ils ne peuvent garantir.
— Pour manipuler en faveur de b, seul le deuxième électeur est intéressé. Pour qu'il réussisse, il faut soit que b soit en tête de tous les bulletins dans la configuration finale (pour bénéficier de la $\mathcal{M}$-condorcification), soit éviter qu'il y ait deux bulletins avec respectivement a et c en tête (pour bénéficier de la règle originale f en évitant le cas 2) ; mais dans les deux cas, c'est impossible.

Résumons : toute les hypothèses du théorème sont vérifiées, hormis le fait qu'au lieu de vérifier $\mathcal{M}'$**CInfA**, le système f vérifie seulement $\mathcal{M}'$**CInf**. Et nous avons exhibé une configuration où $f^{\mathcal{M}'}$ est manipulable mais où $f^{\mathcal{M}}$ n'est pas manipulable. Ainsi, la conclusion du théorème n'est plus valide si l'on suppose seulement que f vérifie $\mathcal{M}'$**CInf**. Ce contre-exemple motive donc l'emploi de l'hypothèse $\mathcal{M}'$**CInfA**.

4.5 Théorème de condorcification maximale

À présent, on cherche à savoir s'il existe une « meilleure » famille pour appliquer la condorcification, c'est-à-dire pour diminuer la manipulabilité en utilisant le théorème de condorcification généralisée 4.18. L'idée naturelle est d'utiliser la famille qui, pour tout candidat c, considère les coalitions qui ont le pouvoir de faire gagner c quand elles sont informées du bulletin des autres électeurs : c'est ce

que nous appellerons la *famille maximale* d'un système de vote. Dans cette section, nous examinons sous quelles conditions cette famille permet d'obtenir une condorcification généralisée qui, en un certain sens, est optimale.

Définition 4.22 (famille maximale de coalitions informées gagnantes)
On appelle *famille maximale de coalitions informées gagnantes* de f, ou *famille maximale* de f, la famille la plus grande $\mathcal{M}'$ (au sens de l'inclusion) telle que f vérifie $\mathcal{M}'$-**CInf**, c'est-à-dire telle que $\forall c \in \mathcal{C}, \forall M \in \mathcal{P}(\mathcal{V})$:

$$M \in \mathcal{M}'_c \Leftrightarrow \left[\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}, \exists \omega_M \in \Omega_M \text{ t.q. } f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c\right].$$

Avec cette notation, le système de vote $f^{\mathcal{M}'}$ est appelé la *condorcification maximale* de f.

Par définition, la famille maximale est la plus grande famille avec laquelle on puisse envisager d'appliquer le théorème de condorcification généralisée 4.18. On peut donc se demander si c'est également celle qui permet de diminuer le plus la manipulabilité par condorcification généralisée.

Une autre convention possible serait d'appeler famille maximale la plus grande famille $\mathcal{M}'$ telle que f vérifie $\mathcal{M}'$**CInfA**, qui est, par définition, incluse dans la famille maximale au sens ci-dessus. Comme nous le verrons, le théorème de condorcification maximale 4.25 portera sur les modes de scrutin raisonnables, où les deux notions sont confondues. Or, il nous semble qu'en pratique, il est plus facile d'identifier la famille maximale $\mathcal{M}'$ (au sens ci-dessus) puis de vérifier que f vérifie aussi $\mathcal{M}'$**CInfA**, plutôt que d'identifier la plus grande famille telle que f vérifie $\mathcal{M}'$**CIgnA** puis de vérifier qu'elle est aussi maximale pour la notion $\mathcal{M}'$**CInf**.

La proposition suivante découle de la définition : en effet, si une coalition informée est toujours capable de faire gagner c, alors toute coalition informée plus grande (au sens de l'inclusion) est également capable de faire gagner c.

Proposition 4.23
La famille maximale de f est monotone.

Par conséquent, on peut appliquer le théorème de condorcification généralisée 4.18 à la famille maximale $\mathcal{M}'$, ce qui prouve que $f^{\mathcal{M}'}$ est moins manipulable que f. Nous allons montrer que, sous certaines hypothèses, la condorcification maximale $f^{\mathcal{M}'}$ est la moins manipulable des $f^{\mathcal{M}}$, où $\mathcal{M}$ est une famille telle que f vérifie $\mathcal{M}$**CInf**.

Par la proposition 4.23, on sait donc que la famille maximale est monotone. Mais elle n'est pas forcément exclusive. En effet, pour le vote de parité (avec $V \geq 2$), il est facile de montrer que la famille maximale est celle qui comprend toutes les coalitions non vides ; mais nous avons déjà remarqué que celle-ci n'est pas exclusive.

Afin d'appliquer le théorème de condorcification comparée 4.21, nous allons montrer dans la proposition suivante que, sous certaines hypothèses, toute autre famille $\mathcal{M}$ susceptible de bénéficier du théorème de condorcification généralisée 4.18 est moins condorcifiante que la famille maximale $\mathcal{M}'$.

Proposition 4.24
Soit $\mathcal{M}'$ la famille maximale de f et $\mathcal{M}$ une famille.
On suppose que :
— *les relations* P_v *sont toujours antisymétriques ;*

— $\mathcal{M}'$ *est exclusive.*

Si f *vérifie* $\mathcal{M}$**CInf**, *alors* $\mathcal{M}'$ *est plus condorcifiante que* $\mathcal{M}$.

Démonstration. Soit $\phi \in \Omega$ et $c \in \mathcal{C}$. Supposons que c est $\mathcal{M}$-Condorcet en ϕ et montrons qu'il est également $\mathcal{M}'$-Condorcet, c'est-à-dire au sens de la famille maximale de f.

Pour tout $d \in \mathcal{C} \setminus \{c\}$, le candidat c a une $\mathcal{M}$-victoire contre d :

$$M = \{v \in \mathcal{V} \text{ t.q. } c \,\mathrm{P}_v\, d\} \in \mathcal{M}_c.$$

Comme $M \in \mathcal{M}_c$ et f vérifie $\mathcal{M}$**CInf**, on a :

$$\forall \omega_{\mathcal{V}\setminus M} \in \Omega_{\mathcal{V}\setminus M}, \exists \omega_M \in \Omega_M \text{ t.q. } f(\omega_M, \omega_{\mathcal{V}\setminus M}) = c.$$

Par définition de la famille maximale $\mathcal{M}'$ de f, ceci implique que $M \in \mathcal{M}'_c$. Donc c a une $\mathcal{M}'$-victoire contre d.

Puisque $\mathcal{M}'$ est exclusive et que les préférences sont antisymétriques, la proposition 4.5 assure que cette $\mathcal{M}'$-victoire est stricte. □

En pratique, on considère généralement des préférences antisymétriques, donc l'hypothèse de la proposition 4.24 qui est la plus susceptible de ne pas être vérifiée est l'exclusivité de la famille maximale $\mathcal{M}'$. Montrons que, dans ce cas, la conclusion de la proposition 4.24 n'est pas vérifiée.

Pour cela, considérons le vote de parité (boules noires et blanches). Il est facile de voir que la famille $\mathcal{M}'$ maximale est neutre et contient toutes les coalitions non vides. Cette famille n'est pas exclusive (en supposant $V \geq 2$) car deux singletons d'électeurs distincts sont des coalitions gagnantes pour les candidats a et b respectivement. Pour que le candidat a soit $\mathcal{M}'$-Condorcet, il faut et il suffit qu'il ait une victoire, c'est-à-dire qu'au moins un électeur le préfère à b, et qu'il n'ait pas de défaite, c'est-à-dire qu'aucun électeur ne préfère b à a ; autrement dit (en supposant que les préférences sont des ordres stricts), il faut et il suffit que a soit le candidat préféré de tous les électeurs. Si on considère, à présent, la famille majoritaire $\mathcal{M}$, pour que a soit $\mathcal{M}$-Condorcet, il faut et il suffit qu'il soit préféré par une majorité stricte des électeurs, ce qui est une condition moins exigeante. Donc il est faux de dire que $\mathcal{M}'$ est plus condorcifiante que $\mathcal{M}$. En l'occurrence, on a même l'inverse : c'est $\mathcal{M}$ qui est plus condorcifiante que $\mathcal{M}'$.

Il ne nous reste plus qu'à regrouper les propriétés connues pour la famille maximale dans le théorème suivant.

Théorème 4.25 (Condorcification maximale)

Soit f *un SVBE et* $\mathcal{M}'$ *sa famille maximale.*

On suppose que f *vérifie* $\mathcal{M}'$**CInfA**.

1. *La condorcification maximale* $f^{\mathcal{M}'}$ *est au plus aussi manipulable que* f :

$$\mathrm{MC}_{f^{\mathcal{M}'}} \subseteq \mathrm{MC}_f.$$

2. *Pour toute famille* $\mathcal{M}$ *qui est moins condorcifiante que* $\mathcal{M}'$ *(au sens large), la condorcification maximale* $f^{\mathcal{M}'}$ *est au plus aussi manipulable que* $f^{\mathcal{M}}$:

$$\mathrm{MC}_{f^{\mathcal{M}'}} \subseteq \mathrm{MC}_{f^{\mathcal{M}}}.$$

3. *On suppose en outre que les préférences* P_v *sont toujours antisymétriques et que la famille maximale* $\mathcal{M}'$ *est exclusive. Alors, pour toute famille* $\mathcal{M}$ *telle que* f *vérifie* $\mathcal{M}$**CInf**, *la condorcification maximale* $f^{\mathcal{M}'}$ *est au plus aussi manipulable que* $f^{\mathcal{M}}$:

$$\mathrm{MC}_{f^{\mathcal{M}'}} \subseteq \mathrm{MC}_{f^{\mathcal{M}}}.$$

Démonstration. 1. Par la proposition 4.23, on sait que la famille maximale $\mathcal{M}'$ est monotone. Et comme f vérifie $\mathcal{M}'$**CInfA**, il vérifie *a fortiori* $\mathcal{M}'$**CInf** (proposition 4.14). On peut donc utiliser le théorème de condorcification généralisée 4.18.

2. Comme la famille maximale $\mathcal{M}'$ est monotone, on peut utiliser le théorème 4.21 de condorcification comparée.

3. Par la proposition 4.24, la famille $\mathcal{M}'$ est moins condorcifiante que $\mathcal{M}$ donc on peut utiliser le point 2 du présent théorème. □

4.6 Exemples de condorcification généralisée

À présent, nous allons appliquer la condorcification généralisée à divers systèmes de vote afin de montrer en pratique comment se comporte cette notion.

4.6.1 Systèmes de vote usuels vérifiant CMInf

Dans le cas des systèmes usuels vérifiant **CMInf**, on sait déjà qu'ils peuvent bénéficier de la condorcification majoritaire (théorème 2.9). L'objectif de cette section est d'examiner s'il s'agit de leur condorcification maximale.

Tout d'abord, remarquons que la plupart des modes de scrutin classiques vérifient **CMIgn** et même le critère de la coalition majoritaire ignorante admissible, c'est-à-dire le critère $\mathcal{M}$**CIgnA** pour la famille majoritaire. En effet, dans ces systèmes de vote raisonnables, si les manipulateurs (en particulier ceux d'une coalition majoritaire ignorante) peuvent faire gagner un certain candidat c, alors ils peuvent le faire en mettant c en tête de leur bulletin.

Pour traiter le cas de ces systèmes de vote, la proposition suivante sera utile. Son point essentiel est la conclusion 1, qui établit une implication entre $\mathcal{M}'$**CIgnA** et la maximalité de $\mathcal{M}'$ dans le cas de la famille majoritaire. Les conclusions 2, 3 et 4 reprennent les conclusions du théorème de condorcification maximale dans ce cas particulier.

Proposition 4.26
Soit f un SVBE. On suppose que :
— *les relations P_v sont toujours antisymétriques ;*
— *le nombre d'électeurs V est impair ;*
— *f vérifie le critère de la coalition majoritaire ignorante admissible.*

1. *Alors la famille maximale $\mathcal{M}'$ de f est la famille majoritaire.*
2. *La condorcification majoritaire $f^{\mathcal{M}'} = f^*$ est au plus aussi manipulable que f :*
$$\mathrm{MC}_{f^*} \subseteq \mathrm{MC}_f.$$
3. *Pour toute famille $\mathcal{M}$ qui est moins condorcifiante que la famille majoritaire $\mathcal{M}'$ (au sens large), la condorcification majoritaire f^* est au plus aussi manipulable que $f^{\mathcal{M}}$:*
$$\mathrm{MC}_{f^*} \subseteq \mathrm{MC}_{f^{\mathcal{M}}}.$$
4. *Pour toute famille $\mathcal{M}$ telle que f vérifie $\mathcal{M}$**CInf**, la condorcification majoritaire f^* est au plus aussi manipulable que $f^{\mathcal{M}}$:*
$$\mathrm{MC}_{f^*} \subseteq \mathrm{MC}_{f^{\mathcal{M}}}.$$

Démonstration. Comme f vérifie le critère de la coalition majoritaire ignorante, il est facile de voir que sa famille maximale est celle des coalitions majoritaires : avec plus de la majorité, on peut toujours imposer sa volonté ; avec moins, les autres électeurs le peuvent toujours (car V est impair donc ils ont eux-mêmes une majorité stricte dans ce cas). Ceci prouve le point 1.

Comme f vérifie $\mathcal{M}'$**CIgnA** par hypothèse, il vérifie aussi $\mathcal{M}'$**CInfA** (proposition 4.14), donc on peut utiliser le théorème de condorcification maximale 4.25. Ceci permet de conclure immédiatement pour les points 2 et 3.

Puisque les préférences sont toujours antisymétriques par hypothèse et que la famille majoritaire $\mathcal{M}'$ est exclusive (définition 4.3), le théorème de condorcification maximale 4.25 permet de prouver également le point 4. □

Corollaire 4.27

On suppose que les relations P_v *sont toujours antisymétriques et que* V *est impair.*

Alors les conclusions de la proposition 4.26 sont vraies pour BI, CSD, Maximin, le jugement majoritaire, PO, l'uninominal, U2TI, VTI, VTID, VTIM, le vote par assentiment, le vote par notation, et les méthodes de Baldwin, Borda, Black, Bucklin, Coombs, Copeland, Dodgson, Kemeny, Nanson et Schulze.

Démonstration. Tous les systèmes mentionnés sauf la méthode de Borda vérifient le critère de la coalition majoritaire ignorante admissible. Pour ceux-ci, la proposition 4.26 permet donc de conclure.

Si le système f considéré est la méthode de Borda, on ne peut pas appliquer la proposition 4.26 car ce système de vote ne vérifie pas **CMIgn** en général (corollaire 3.14 de la proposition 3.13). Cependant, nous allons voir que le résultat reste vrai comme corollaire du théorème de condorcification maximale 4.25.

En effet, f vérifie le critère de la coalition majoritaire informée admissible. Montrons que la famille majoritaire est maximale : si une coalition n'est pas majoritaire, alors les autres électeurs sont strictement majoritaires (car V est impair) ; s'ils placent un certain candidat d en premier et c en dernier sur leurs bulletins, alors c ne peut pas être élu, quoi que fassent les manipulateurs de la coalition minoritaire. On peut donc appliquer le théorème de condorcification maximale 4.25. □

Pour les systèmes qui vérifient le critère de Condorcet, la condorcification ne modifie évidemment pas le système de vote. La proposition 4.26 et le corollaire 4.27 signifient que, dès leur état original, on ne peut pas espérer diminuer la manipulabilité de ces systèmes en leur appliquant le théorème de condorcification généralisée 4.18. Autrement dit, ils sont leur propre condorcification maximale.

4.6.2 Veto

Le système Veto, ne vérifiant pas **CMInf**, ne peut bénéficier du théorème faible de condorcification (majoritaire) 2.9. À présent, nous pouvons déterminer sa famille maximale et lui appliquer le théorème de condorcification maximale 4.25.

Proposition 4.28

Le système de vote f *considéré est Veto. On suppose que les préférences sont toujours antisymétriques.*

On suppose $\mathrm{mod}(V, C) = C - 1$.

1. *Alors la famille maximale* $\mathcal{M}'$ *de Veto est la famille à seuil dont les coalitions sont de cardinal supérieur ou égal à* $(C-1)\left\lceil \frac{V}{C} \right\rceil$.

2. *Veto vérifie* $\mathcal{M}'$**CIgnA**.
3. *La famille maximale* $\mathcal{M}'$ *est exclusive.*

Donc toutes les conclusions du théorème de condorcification maximale 4.25 sont vérifiées.

De même que nous avons supposé le nombre d'électeurs impair dans la proposition 4.26 et le corollaire 4.27 pour éviter les questions d'égalité dans les modes de scrutin liés à la notion de majorité, la proposition 4.28 utilise une hypothèse de modulo pour simplifier l'étude : $\mathrm{mod}(V, C) = C - 1$. On pourra remarquer que pour $C = 2$, Veto devient équivalent au vote majoritaire simple et qu'alors cette équation signifie simplement que V est impair.

Dans la limite où V est grand devant C, on pourra considérer en première approximation que la famille maximale de Veto est celle des $\left(1 - \frac{1}{C}\right) V$, aux questions d'arrondi et de départage près. Autrement dit, pour choisir le vainqueur en Veto, il faut une très grande coalition, d'autant plus grande que C est élevé.

Démonstration. Notons $\alpha = (C-1)\left\lceil \frac{V}{C} \right\rceil$ et $\mathcal{M}'$ la famille à seuil dont les coalitions sont de cardinal supérieur ou égal à α. Nous allons montrer que $\mathcal{M}'$ vérifie les propriétés énoncées et en particulier que c'est la famille maximale de f. Pour tout le début de cette démonstration, nous n'allons pas recourir à l'hypothèse $\mathrm{mod}(V, C) = C - 1$ mais seulement à l'hypothèse plus faible que C ne divise pas V.

3. Cette hypothèse affaiblie est suffisante pour avoir $\alpha > (C-1)\frac{V}{C} \geq \frac{V}{2}$. Donc la famille $\mathcal{M}'$ est exclusive.

2. Si une coalition de cardinal $(C-1)\left\lceil \frac{V}{C} \right\rceil$ souhaite faire gagner un certain candidat c, il suffit que $\left\lceil \frac{V}{C} \right\rceil$ membres de la coalition votent contre chaque autre candidat. Alors comme C ne divise pas V, chacun de ces candidats adversaires reçoit un nombre de vetos strictement supérieur au nombre de vetos moyen, donc seul c peut être élu. Ainsi, Veto vérifie $\mathcal{M}'$**CIgn**.

Les membres de la coalition peuvent réaliser cette manipulation tout en mettant c en tête de leur bulletin. Comme la famille $\mathcal{M}'$ est exclusive, ceci assure alors que c apparaît comme un candidat $\mathcal{M}'$-admissible (aucun candidat d ne peut être placé α fois devant c donc il peut pas avoir une $\mathcal{M}'$-victoire contre c). Donc Veto vérifie $\mathcal{M}'$**CIgnA**.

1. Comme Veto vérifie $\mathcal{M}'$**CIgnA**, nous savons déjà qu'il vérifie $\mathcal{M}'$**CInfA** (proposition 4.14). Pour montrer que $\mathcal{M}'$ est maximale, il ne reste qu'à montrer qu'une coalition n'appartenant pas à $\mathcal{M}'$ n'a pas la capacité de décider du vainqueur, même en étant informée des autres votes. C'est pour ce point que nous allons exploiter l'hypothèse $\mathrm{mod}(V, C) = C - 1$. En remarquant qu'on a dans le cas général :

$$\left\lceil \frac{V}{C} \right\rceil = \frac{V + (C - \mathrm{mod}(V, C))}{C},$$

on obtient alors $\alpha = \frac{(C-1)(V+1)}{C}$.

Considérons une coalition n'appartenant pas à $\mathcal{M}'$, c'est-à-dire de taille strictement inférieure à α. Alors le cardinal de son complémentaire est au moins :

$$V + 1 - \alpha = \frac{V+1}{C} > \frac{V}{C}.$$

Donc, si tous les électeurs du complémentaire votent contre le candidat c, alors celui-ci a un nombre de vetos supérieur à la moyenne et il ne peut pas être élu. Donc la coalition considérée n'est pas toujours capable de faire gagner c, même de manière informée, ce qui prouve que $\mathcal{M}'$ est la famille maximale de Veto. □

Nous allons voir que la condorcification maximale de Veto peut être strictement moins manipulable que Veto. Pour cela, considérons $V = 7$ électeurs et $C = 4$ candidats. On a $\mod(V, C) = C - 1$, donc les hypothèses de la proposition 4.28 sont vérifiées et la famille maximale $\mathcal{M}'$ est celle des coalitions comprenant 6 électeurs ou plus. Examinons la configuration ω suivante.

3	3	1
c	c	a
a	a	b
b	d	d
d	b	c

$D(\omega)$	a	b	c	d
a	–	7	1	7
b	0	–	1	4
c	6	6	–	6
d	0	3	1	–

En Veto, le candidat a est élu et la configuration est manipulable en faveur de c : il suffit que les 6 manipulateurs émettent 2 vetos contre chacun des candidats a, b et d. Dans la condorcification maximale de Veto, le candidat c est élu et il est facile de voir que ce n'est pas manipulable : en effet, c est non seulement $\mathcal{M}'$-Condorcet mais il est même $\mathcal{M}'$-favori ; par conséquent, les électeurs qui placent c en tête, et qui ne sont donc jamais manipulateurs, suffisent à assurer qu'il reste $\mathcal{M}'$-Condorcet.

À titre de curiosité, on peut aussi montrer que la condorcification usuelle (c'est-à-dire majoritaire) de Veto ne fonctionne pas, au sens où elle peut être manipulable dans certaines configurations alors que Veto ne l'est pas. Pour cela, considérons la configuration ω suivante, avec de nouveau $V = 7$ électeurs et $C = 4$ candidats.

1	2	1	1	2
c	c	d	a	a
a	a	c	b	b
d	b	a	c	d
b	d	b	d	c

$D(\omega)$	a	b	c	d
a	–	7	3	6
b	0	–	3	5
c	4	4	–	4
d	1	2	3	–

Le candidat a est le seul à ne recevoir aucun veto, donc il est élu. Montrons que Veto n'est pas manipulable en ω. Aucun électeur ne préfère b à a. Les candidats c et d reçoivent au moins 2 vetos chacun, ce qui est davantage que le nombre de vetos moyen $\frac{V}{C} = \frac{7}{4}$, donc il est impossible de manipuler en leur faveur. Par conséquent, Veto n'est pas manipulable en ω.

À présent, examinons la condorcification majoritaire de Veto. Le candidat c est vainqueur de Condorcet en ω, donc il est élu. Examinons la configuration suivante ψ, qui est une tentative de manipulation pour a.

1	2	1	1	2
c	c	d	a	a
a	a	c	b	b
d	b	a	**d**	d
b	d	b	**c**	c

$D(\psi)$	a	b	c	d
a	–	7	3	6
b	0	–	3	5
c	4	4	–	3
d	1	2	4	–

Il n'y a alors plus de vainqueur de Condorcet et c'est a qui est élu. Ainsi, la condorcification majoritaire de Veto est manipulable en ω vers ψ en faveur de a, alors que Veto n'est pas manipulable en ω.

Dans le cas de Veto, l'intérêt de la condorcification maximale n'est pas avant tout son utilisation en pratique : en effet, la probabilité d'avoir un candidat $\mathcal{M}'$-Condorcet semble relativement faible dans les cultures raisonnables, puisqu'il faut

qu'il ait environ $\left(1 - \frac{1}{C}\right) V$ partisans contre n'importe quel adversaire. En revanche, elle permet de mieux comprendre le phénomène général et de donner une mesure du pouvoir des électeurs dans ce mode de scrutin.

4.6.3 Vote de parité

À présent, nous allons examiner un cas où on ne peut pas utiliser le théorème de condorcification maximale 4.25. Pour cela, reprenons l'exemple du vote de parité : s'il y a un nombre impair (resp. pair) de boules noires, alors le candidat a (resp. b) est élu. En parallèle, chaque électeur fournit également son ordre de préférences, mais celui-ci n'est pas pris en compte.

Nous avons déjà remarqué que la famille maximale de ce système est celle de toutes les coalitions non vides et que cette famille n'est pas exclusive. Il n'est donc pas possible d'appliquer le théorème de condorcification maximale 4.25.

Cependant, on peut remarquer que pour toute famille monotone $\mathcal{M}$ telle que f vérifie $\mathcal{M}$**CInf**, f vérifie aussi $\mathcal{M}$**CInfA**. Par théorème de condorcification comparée 4.21, si l'une de ces familles $\mathcal{M}'$ est plus condorcifiante qu'une autre de ces familles $\mathcal{M}'$, alors $f^{\mathcal{M}'}$ est au plus aussi manipulable que $f^{\mathcal{M}}$.

La relation « être plus condorcifiante » est clairement une relation d'ordre (partiel), en particulier sur l'ensemble des familles monotones $\mathcal{M}$ telle que f vérifie $\mathcal{M}$**CInf**. Les éléments maximaux de cet ordre permettent des condorcifications optimales (en termes de manipulabilité), mais qu'on ne sait *a priori* pas comparer entre elles du point de vue de la manipulabilité.

Examinons deux exemples.

En supposant V impair, considérons la famille majoritaire. Alors la condorcification généralisée est le vote majoritaire simple, qu'on sait non manipulable.

À présent, soit $\mathcal{M}_a$ l'ensemble des coalitions de plus d'un tiers des électeurs et $\mathcal{M}_b$ celle des coalitions d'au moins deux tiers des électeurs. Dans la condorcification généralisée, b gagne s'il possède au moins deux tiers de voix ; sinon c'est a. Dans ce cas également, la condorcification généralisée n'est pas manipulable.

Ces deux condorcifications sont deux façons différentes de diminuer la manipulabilité, obtenues avec deux familles qui ne sont pas comparables, c'est-à-dire telles qu'aucune des deux n'est plus condorcifiante que l'autre.

Par ailleurs, on peut examiner les condorcifications du vote de parité en utilisant les familles à seuil (c'est-à-dire neutres, anonymes et monotones).

Tout d'abord, remarquons que pour la condorcification généralisée, utiliser par exemple la famille des 30 % ou celle des 70 % revient au même (en supposant que V n'est pas divisible par 10 pour éviter les nuances entre inégalités strictes et larges). En effet, dans la famille des 30 %, un certain candidat c possède une victoire absolue stricte contre un certain d ssi plus de 30 % préfère c à d (victoire) et moins de 30 % préfère d à c (pas de défaite) ; comme les préférences sont des ordres stricts totaux, ceci revient à dire que plus de 70 % des électeurs préfèrent c à d, c'est-à-dire que c possède une victoire absolue stricte contre d au sens de la famille des 70 %.

Toujours sous l'hypothèse que les relations binaires sont des ordres stricts totaux, on voit donc que la famille des 50 % est, parmi les familles à seuil, celle qui accepte le plus de victoires (puisque même dans la famille des 30 %, il faut des victoires avec au moins 50 % d'électeurs). C'est donc la plus condorcifiante des familles monotones, neutres et anonymes.

Par conséquent, on peut retrouver le vote majoritaire simple comme la condorcification optimale du vote de parité parmi celles obtenues avec des familles monotones, anonymes et neutres.

Pour l'illustrer, considérons la configuration ω suivante, où tous les électeurs utilisent une boule blanche.

60	39
a	b
b	a

Comme il y a un nombre pair de boules noires (égal à 0), c'est le candidat b qui est élu. Si on utilise la famille $\mathcal{M}$ des 70 %, alors il n'y a pas de candidat $\mathcal{M}$-Condorcet : aucun candidat n'a une victoire contre l'autre. Si on utilise la famille des 30 %, alors il n'y a pas de candidat $\mathcal{M}$-Condorcet non plus : les candidats a et b ont une victoire mutuelle l'un contre l'autre (donc ils ne sont même pas $\mathcal{M}$-admissibles). Dans ces deux systèmes, il suffit qu'un candidat préférant a remplace sa boule blanche par une boule noire, sans changer son ordre de préférence, pour manipuler en faveur de a.

En revanche, si on utilise la famille majoritaire, alors la condorcification du vote de parité est simplement le vote majoritaire simple : a est élu et ce système n'est jamais manipulable (le vote sincère, qui consiste à voter pour son candidat préféré, est une stratégie dominante).

4.6.4 Vote d'une loi

À présent, nous allons nous intéresser à des modes de scrutin qui violent l'anonymat, la neutralité, ou les deux. Dans toute la suite de ce chapitre, on note $\mathcal{M}'$ la famille maximale du système f étudié. On suppose pour simplifier que les relations binaires de préférence sont des ordres stricts totaux.

Nous utiliserons souvent implicitement la proposition 4.13 : si un système vérifie $\mathcal{M}$**CIgn**, alors $\mathcal{M}$ est exclusive. Comme les préférences sont antisymétriques, on rappelle qu'alors la définition 4.7 de candidat $\mathcal{M}$-Condorcet se réduit à la relation (4.1) : il suffit de vérifier que le candidat concerné possède une victoire contre tout autre candidat et il est inutile de vérifier qu'aucun autre candidat n'a de victoire contre lui.

Examinons un premier système, qui permet de voter une loi. On pourra penser, par exemple, à une révision de la Constitution. On soumet au vote les trois options suivantes : deux versions a et b du projet de loi, et $\varnothing$ qui représente le statu quo, c'est-à-dire le fait qu'aucune des deux versions ne soit entérinée. Le système de vote considéré est le suivant :

1. On vote entre a et $\varnothing$. Si a reçoit au moins deux tiers des voix, alors a est élu et le processus est terminé.
2. Sinon, on vote entre b et $\varnothing$. Si b reçoit au moins deux tiers des voix, alors b est élu ; sinon, $\varnothing$ est élu.

Il s'agit d'un *système de vote général* en plusieurs tours (section 1.4). Le *mode de scrutin* utilisé pose un problème *a priori* pour définir la fonction de sincérité [2] : si un électeur a l'ordre de préférence $(b \succ a \succ \varnothing)$, il n'est pas clair de savoir quel doit être son bulletin au premier tour puisque en réalité, le choix n'est pas entre a et $\varnothing$ mais plutôt entre a et le fait d'effectuer une élection opposant b et $\varnothing$. Le *système de vote* que nous allons étudier utilise la fonction de sincérité sous-entendue par notre formulation initiale : pour le premier tour, le vote sincère d'un électeur est de voter pour a ssi il préfère a à $\varnothing$.

2. Cet exemple est très semblable à celui de la *fête sans alcool* de Gibbard (1973), qui est précisément conçu pour illustrer la problématique de la définition du vote sincère.

La première chose que nous allons faire est de considérer à la place la version basée sur l'état de ce système, qui est au plus aussi manipulable (proposition 1.4.2) : chaque électeur communique son ordre de préférence, puis on émule le système original.

Examinons à présent sa famille maximale.
Pour toute coalition M, on a :

$$M \in \mathcal{M}'_{\varnothing} \Leftrightarrow \text{card}(M) > \frac{V}{3}.$$

Pour toute coalition M et tout candidat $c \neq \varnothing$, on a :

$$M \in \mathcal{M}'_{c} \Leftrightarrow \text{card}(M) \geq \frac{2V}{3}.$$

Il est facile de vérifier que le système de vote vérifie $\mathcal{M}'$**CIgnA**. Donc la famille maximale $\mathcal{M}'$ est exclusive.

Pour comparer la manipulabilité de f et $f^{\mathcal{M}'}$, considérons le profil ω suivant.

70	30
b	a
a	b
$\varnothing$	$\varnothing$

Dans le système original (ou sa version basée sur l'état), c'est le candidat a qui gagne dès le premier vote, en tout cas si les électeurs sont sincères. Mais la configuration est manipulable en faveur de b : en effet, ses partisans peuvent échanger a et $\varnothing$ dans leur ordre de préférence ; alors le projet de loi a est rejeté puis b est accepté.
Dans la condorcification maximale, b est $\mathcal{M}'$-Condorcet donc il est élu et c'est non manipulable : en effet, le candidat b est $\mathcal{M}'$-favori donc les électeurs qui ont b pour préféré assurent à eux seuls que b soit $\mathcal{M}'$-Condorcet.
Dans cet exemple, la condorcification maximale permet donc d'avoir un système de vote strictement moins manipulable que l'original.

4.6.5 Uninominal avec un pouvoir d'imposition

Renaud et ses amis envisagent d'aller au bal (b), à l'église (e), de rendre visite à Germaine (g) ou la Pépette (p) ou de rester ensemble à parler ($\varnothing$). On vote en uninominal à l'exception suivante près. Renaud, qui possède la voiture, a le droit d'imposer le candidat $\varnothing$: s'il vote pour $\varnothing$, alors ce candidat est automatiquement élu, personne ne va nulle part et on reste ensemble à parler.
Pour toute coalition M, on a :

$$M \in \mathcal{M}'_{\varnothing} \Leftrightarrow \text{Renaud} \in M \text{ ou } \text{card}(M) > \frac{V}{2}.$$

Pour toute coalition M et tout candidat $c \neq \varnothing$, on a :

$$M \in \mathcal{M}'_{c} \Leftrightarrow \text{Renaud} \in M \text{ et } \text{card}(M) > \frac{V}{2}.$$

Il est facile de vérifier que le système de vote vérifie $\mathcal{M}'$**CIgnA**. Donc la famille maximale $\mathcal{M}'$ est exclusive.
Examinons le profil ω suivant avec $V = 9$ électeurs.

1 (Renaud)	1	1	1	1	1	1	1	1
g	$\varnothing$	b	b	b	p	p	e	e
$\varnothing$	g	$\varnothing$	$\varnothing$	e	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
e	e	g	e	$\varnothing$	g	g	g	g
p	p	p	g	g	e	e	b	b
b	b	e	p	p	b	b	p	p

$D(\omega)$	$\varnothing$	b	e	g	p
$\varnothing$	–	6*	6*	8	7*
b	3	–	3	3	5
e	3	6*	–	4	6*
g	1*	6*	5*	–	7*
p	2	4*	3	2	–

Dans la matrice des duels ci-dessus, on a conventionnellement signalé par une astérisque chaque ensemble d'électeurs qui contient Renaud.

Dans le système de vote original, on a $f(\omega) = b$, mais Renaud peut manipuler en faveur de $\varnothing$ en utilisant son pouvoir d'imposition.

Si on utilise la condorcification majoritaire f^*, alors le vainqueur sincère est $\varnothing$ car il est vainqueur de Condorcet. Mais le profil reste manipulable : en effet, considérons la configuration ψ suivante, qui est une tentative de manipulation pour b.

1 (Renaud)	1	1	1	1	1	1	1	1
g	$\varnothing$	b	b	b	p	p	e	e
$\varnothing$	g	**g**	**e**	e	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
e	e	**e**	**Ø**	$\varnothing$	g	g	g	g
p	p	**Ø**	g	g	e	e	b	b
b	b	**p**	p	p	b	b	p	p

$D(\psi)$	$\varnothing$	b	e	g	p
$\varnothing$	–	6*	4*	6	7*
b	3	–	3	3	5
e	5	6*	–	4	7*
g	3*	6*	5*	–	7*
p	2	4*	2	2	–

Il n'y a plus de vainqueur de Condorcet et c'est b qui gagne, donc la condorcification majoritaire f^* est manipulable de ω vers ψ en faveur de b.

Enfin, si on utilise la condorcification maximale, alors le vainqueur sincère en ω est également $\varnothing$ car il est $\mathcal{M}'$-Condorcet. Si on souhaite manipuler en faveur de g, alors seul Renaud est intéressé, mais il ne peut empêcher que $\varnothing$ reste $\mathcal{M}'$-Condorcet car il aura toujours une majorité stricte contre tout autre candidat ; donc la manipulation échoue. Si on souhaite manipuler en faveur d'un autre candidat que g, alors Renaud ne participe pas à la manipulation donc sa voix, à elle seule, assure que $\varnothing$ conserve une $\mathcal{M}'$-victoire contre e, p et b. Le seul espoir des manipulateurs consiste donc à empêcher $\varnothing$ d'avoir une $\mathcal{M}'$-victoire contre g. Si on considère une manipulation pour b (resp. e, p), alors on peut lire sur la matrice des duels qu'il y a 3 manipulateurs (resp. 3, 2), donc le nombre de voix de $\varnothing$ contre g, qui est 8 en vote sincère, est, après manipulation, supérieur ou égal à $8 - 3 = 5$ (resp. 5, 6), donc $\varnothing$ conserve une victoire contre g et la manipulation échoue. Ainsi, $f^{\mathcal{M}'}$ n'est pas manipulable en ω, alors que f et f^* le sont.

4.6.6 Élection du Secrétaire général de l'ONU

Nous allons à présent étudier un exemple un peu plus complexe mais issu de la politique internationale. Nous nous contenterons de donner la famille maximale, sans étudier en détails le système de vote obtenu par condorcification maximale.

L'élection du Secrétaire général de l'ONU peut être formalisée de la façon suivante. Il s'agit d'un vote par notation où les notes autorisées sont -1, 0 et 1 ; mais chaque pays membre du Conseil de Sécurité, un sous-ensemble de $\mathcal{V}$ noté CS, dispose d'un droit de rejet sur toute personne candidate. En fait, il y a donc autant de candidats que de personnes physiques briguant le poste, plus un candidat spécial $\varnothing$ qui signifie : « aucune personne physique n'est élue » (ce qui a diverses conséquences en pratique).

Ainsi, chaque membre de CS possède un droit de rejet sur tout candidat, sauf sur le candidat spécial $\varnothing$. C'est un pouvoir plus fort que le pouvoir d'imposition de Renaud vu en section 4.6.5 : en effet, un membre du Conseil de Sécurité peut non seulement imposer le candidat $\varnothing$ comme vainqueur (tout comme Renaud), mais il peut aussi, de manière plus nuancée, choisir de mettre son veto sur certains autres candidats mais pas tous. Par exemple, il peut mettre son veto sur tous les candidats sauf une certaine personne physique c et le candidat spécial $\varnothing$, ce qui restreint le choix des autres électeurs à ces deux candidats : Renaud ne dispose pas d'un tel pouvoir.

Pour toute coalition M, on a :

$$M \in \mathcal{M}'_{\varnothing} \Leftrightarrow \exists v \in \text{CS t.q. } v \in M.$$

Pour toute coalition M et tout candidat $c \neq \varnothing$, on a :

$$M \in \mathcal{M}'_{c} \Leftrightarrow \text{CS} \subseteq M \text{ et } \text{card}(M) > \frac{V}{2}.$$

Il est facile de vérifier que le système de vote vérifie $\mathcal{M}'\mathbf{CIgnA}$, donc la famille maximale $\mathcal{M}'$ est exclusive.

On peut donc appliquer le théorème de condorcification maximale 4.25. Afin de ne pas alourdir davantage, nous ne donnerons, cette fois, pas d'exemple d'application. On remarquera que, pour qu'une candidat physique soit $\mathcal{M}'$-Condorcet, il faut qu'il soit non seulement vainqueur de Condorcet au sens usuel majoritaire mais également le candidat préféré de chaque membre du Conseil de Sécurité. Pour les candidats physiques, être $\mathcal{M}'$-Condorcet est donc une condition très exigeante. En revanche, pour que le candidat spécial $\varnothing$ soit $\mathcal{M}'$-Condorcet, il suffit que pour tout candidat physique c, il existe un membre du Conseil de Sécurité qui préfère $\varnothing$ à c, ce qui est une condition particulièrement peu exigeante. L'analyse de la famille maximale montre que, dans un tel système, l'issue spéciale $\varnothing$ est beaucoup plus facile à obtenir que n'importe quel candidat physique, aussi bien dans le système initial que dans sa condorcification maximale.

Si les parties en présence ont par ailleurs un intérêt à éviter l'issue spéciale $\varnothing$, on peut imaginer, et c'est ce qui se passe en pratique, qu'il existe des processus externes au système de vote utilisé qui permettent de pré-sélectionner les candidats de façon à ne pas aboutir à un blocage lors du vote lui-même. En l'occurrence, les membres du Conseil de Sécurité mènent des négociations préalables de façon à identifier un candidat suffisamment consensuel pour éviter un blocage et le soumettent ensuite au vote de l'Assemblée Générale (dont ils font partie). À notre connaissance, lors de chaque élection passée du Secrétaire Général de l'ONU (à l'heure où nous écrivons ces lignes), le Conseil de Sécurité a pré-sélectionné un candidat, et un seul, pour l'élection finale impliquant l'ensemble des membres de l'Assemblée Générale des Nations Unies.

Chapitre 5

Tranchage

Dans les chapitres 2 et 3, nous avons examiné les liens des notions majoritaires (en particulier le vainqueur de Condorcet absolu) avec la manipulabilité et, plus généralement, avec la capacité à trouver des équilibres de Nash forts. Dans le chapitre 4, nous avons étendu les résultats de condorcification afin de diminuer (au sens large) la manipulabilité pour des systèmes où la notion de majorité ne joue pas de rôle particulier.

Dans ce chapitre, nous étudions le lien entre l'ordinalité d'un système de vote et la manipulabilité. Pour rappel, nous disons qu'un système de vote est *ordinal* si le résultat ne dépend que du *profil*, c'est-à-dire des relations binaires de préférence des électeurs, même si ce ne sont pas des ordres. Parmi les systèmes non-ordinaux, on utilise informellement le qualificatif *cardinal* pour des systèmes qui utilisent une note numérique, comme le vote par assentiment, le vote par notation et le jugement majoritaire. D'autres systèmes ne sont pas ordinaux, comme le vote de parité qui utilise des boules noires et boules n'ayant pas de signification ordinale ou cardinale particulière.

Concernant les systèmes cardinaux, on peut émettre l'intuition qu'ils ont une tendance à être plus manipulables que les systèmes ordinaux : en effet, non seulement un électeur peut mentir sur les positions relatives des candidats dans ses préférences (aspect ordinal), mais il dispose aussi d'un degré de liberté supplémentaire qui consiste à exagérer ses notes, vers le haut ou vers le bas (aspect cardinal).

Dans ce chapitre, nous allons voir qu'il est effectivement possible de donner un sens précis à cette intuition et de la généraliser à tous les types de systèmes qui ne sont pas ordinaux. Pour l'essentiel, ces travaux sont présentés par Durand *et al.* (2014c,e).

Dans la section 5.1, nous définissons les *tranches* d'un système de vote f, chacune d'elles étant un système de vote ordinal. Intuitivement, chaque tranche est de la forme suivante : considérant le profil P des préférences ordinales des électeurs, on le complète d'une manière arbitraire en un état de préférence fictif ω' et on applique le système original f à l'état obtenu. Par exemple, la méthode de Borda est une tranche du vote par notation, parmi l'infinité de ses tranches possibles : elle consiste à « plaquer » arbitrairement les notes $C-1, C-2, \ldots, 0$ sur l'ordre de préférence de chaque électeur.

Une illustration intuitive de la métaphore de la *tranche* est donnée dans la figure 5.1. En général, si on dispose d'une fonction réelle à deux variables $(x, y) \to f(x, y)$, on peut la représenter dans un espace à 3 dimensions (où elle définit une surface). Fixer y revient à considérer une tranche à 2 dimensions de

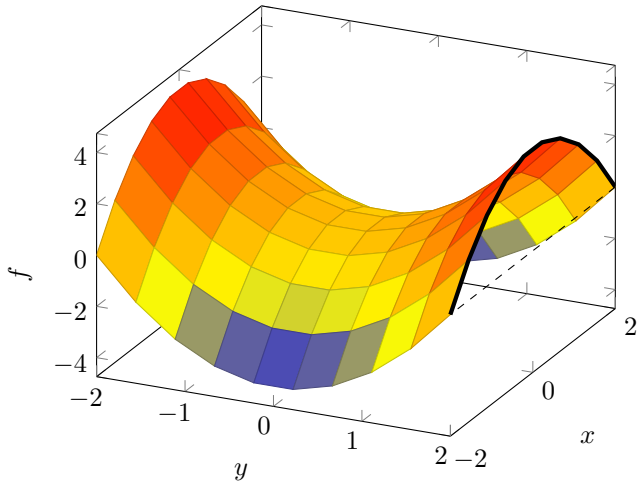


FIGURE 5.1 – En trait noir épais : tranche $y = 2$ d'une fonction réelle f à deux variables.

cet espace (dans laquelle la fonction $x \to f(x, y)$ définit une courbe). Dans cette analogie, x correspond au profil P des préférences ordinales et y est l'information complémentaire qui permet de reconstituer une configuration complète (par exemple la liste des notes). De façon informelle, une tranche du système de vote peut être vue comme une fonction $\mathrm{P} \to f(\mathrm{P}, y)$, où l'information complémentaire y est fixée arbitrairement.

Dans la section 5.2, nous introduisons la notion d'espace électoral probabilisé (EEP) *décomposable*, qui servira ensuite d'hypothèse au théorème de tranchage 5.9. Intuitivement, l'état d'un électeur est caractérisé par deux informations : ses préférences ordinales et une « information complémentaire » sur son état, par exemple une liste de notes (sans leur correspondance avec les candidats). Si on dispose des deux informations, alors on connaît son état complet. L'idée initiale de la décomposabilité consiste à demander que les deux informations soient indépendantes au sens probabiliste. Mais nous verrons que quitte à définir astucieusement l'espace des informations complémentaires, on peut étendre cette notion à une plus large variété de cas. En particulier, nous montrons que si les électeurs sont indépendants les uns des autres, alors l'espace est toujours décomposable : ceci n'est pas évident, puisque cette propriété est vérifiée même si, pour chaque électeur, ses préférences ordinales et les autres informations sur son état ne semblent pas du tout indépendantes à première vue. Nous montrons également que l'indépendance des électeurs n'est pas nécessaire et nous exhibons une classe plus large de modèles qui sont décomposables.

Dans la section 5.3, nous prouvons le résultat central de ce chapitre, le *théorème de tranchage* 5.9 : si l'espace électoral est décomposable, alors il existe une tranche de f dont le taux de manipulabilité est inférieur ou égal à celui de f. La tranche concernée est, par définition, un système ordinal.

Le fait que ce système soit une tranche de f permettra d'assurer qu'elle conserve certaines propriétés de f. En particulier, nous verrons en section 5.4 si f vérifie le critère de Condorcet, alors toute tranche de f le vérifie également. Par conséquent, on peut combiner le théorème de tranchage 5.9 avec les théorèmes de condorcification. En partant d'un système de vote f, on peut considérer

sa condorcification puis trancher celle-ci pour obtenir un système au plus aussi manipulable (au sens probabiliste).

En particulier, ceci nous amène en section 5.5 au *théorème d'optimalité* 5.15, qu'on peut voir comme la conclusion de ces cinq premiers chapitres théoriques : dans la classe **CMInf**, si l'espace électoral est décomposable, alors il existe un système de vote ordinal et vérifiant le critère de Condorcet dont le taux de manipulabilité soit minimal.

Les sections 5.6 et 5.7 peuvent être considérées comme des annexes techniques de ce chapitre. Dans la section 5.6, nous élargissons la notion de décomposabilité à un espace probabilisé qui n'est pas nécessairement un espace électoral et nous prouvons certains résultats techniques qui n'étaient que mentionnés dans la section 5.2.

Dans la section 5.7, nous discutons les hypothèses du théorème de tranchage 5.9 et nous montrons en quel sens le théorème est bien ajusté. En particulier, nous montrons qu'on ne peut espérer diminuer la manipulabilité qu'au sens probabiliste et non au sens ensembliste. Nous montrons aussi qu'il n'est pas possible de supprimer l'hypothèse de décomposabilité purement et simplement, c'est-à-dire sans la remplacer par une autre hypothèse éventuellement plus faible. En revanche, nous ignorons s'il est possible de la remplacer par une condition plus faible. Trouver une condition nécessaire et suffisante pour que tout système admette une tranche moins manipulable (au sens probabiliste) est laissé comme question ouverte.

5.1 Tranches d'un mode de scrutin

Dans l'introduction de ce chapitre, nous avons donné une première intuition du tranchage. De façon plus détaillée, chaque électeur v communique sa relation binaire de préférence $p_v = \mathrm{P}_v(\omega_v)$. Puis on utilise une méthode définie à l'avance, notée y_v, pour reconstituer un état fictif ω'_v qui est cohérent avec p_v. Enfin, on applique f à la configuration fictive $(\omega'_1, \ldots, \omega'_V)$. L'ensemble de ce processus définit un système de vote f_y qui est *ordinal* : il ne dépend que des relations binaires de préférence.

Par exemple, dans l'espace électoral de référence où $\omega_v = (p_v, u_v, a_v)$, considérons $y_v(p_v) = (p_v, u'_v, a'_v)$, où u'_v est le vecteur de scores de Borda[1] associé à p_v et a'_v est un vecteur d'assentiment avec 1 pour chaque candidat. Dans ce cas particulier, les fonctions y_v sont les mêmes pour chaque électeur v, mais ce n'est pas obligatoire en général.

Examinons la *tranche* du vote par notation par $y = (y_1, \ldots, y_V)$. Une fois que les électeurs ont communiqué des relations binaires de préférence $p = (p_1, \ldots, p_V)$, on utilise y afin de reconstituer des états fictifs ω'_v : en particulier, chaque u'_v est, à présent, un vecteur de scores de Borda. Enfin, on applique le vote par notation à cette configuration fictive : le candidat avec le plus grand score total est déclaré élu. Pour résumer, la *tranche* du vote par notation par y est la méthode de Borda. Une infinité d'autres tranches peuvent être définies, en fonction du choix de y.

1. Score de Borda du candidat c pour l'électeur v (rappel) : c reçoit un point pour chaque candidat d tel que v préfère c à d, et un demi-point pour chaque candidat que v juge incomparable ou mutuellement préférable à c. Dans cet exemple, on divise ensuite ce score par $C - 1$ pour obtenir une valeur dans l'intervalle $[0, 1]$.

Plus généralement, toute RPS peut être obtenue comme une tranche du vote par notation[2]. En revanche, l'inverse n'est pas vrai : dans une tranche du vote par notation, il est possible de compléter l'état d'un électeur avec des scores d'uninominal (1 pour le meilleur candidat, 0 pour les autres) et l'état d'un autre électeur avec des scores de Borda. Dans ce cas, la tranche obtenue n'est pas une RPS. On remarquera au passage que le tranchage ne préserve pas l'anonymat éventuel du mode de scrutin initial, mais nous verrons qu'il préserve tout de même un certaine balance des pouvoirs dans la section 5.4.

Nous pouvons maintenant donner les définitions formelles de tout cela.

Notations 5.1 (espace $\mathcal{Y}$)

Pour chaque électeur v, on note :

$$\mathcal{Y}_v = \{y_v : \mathrm{P}_v(\Omega_v) \to \Omega_v \text{ t.q. } \mathrm{P}_v \circ y_v = \mathrm{Id}\}.$$

Il s'agit de l'ensemble des fonctions y_v qui, à chaque p_v possible, associent un état fictif $\omega'_v = y_v(p_v)$ qui est cohérent avec p_v, c'est-à-dire tel que $\mathrm{P}_v(\omega'_v) = p_v$. Mathématiquement, c'est l'ensemble des inverses à droite de P_v (corestreint à son image), c'est-à-dire les fonctions qui choisissent un élément ω'_v dans $(\mathrm{P}_v)^{-1}(p_v)$.

Intuitivement, $\mathcal{Y}_v$ est l'ensemble des « informations complémentaires » possibles pour l'électeur v.

On note $\mathcal{Y} = \prod_{v=1}^{V} \mathcal{Y}_v$. Une fonction $y = (y_1, \ldots, y_V) \in \mathcal{Y}$ définit une information complémentaire pour chaque électeur, c'est-à-dire une méthode de tranchage[3] : à chaque profil possible $p = (p_1, \ldots, p_V) \in \prod_{v\in\mathcal{V}} \mathrm{P}_v(\Omega_v)$, elle associe une configuration $\omega' = y(p) = (y_1(p_1), \ldots, y_V(p_V))$ qui est cohérente avec p.

Définition 5.2 (tranche)

Pour un SVBE f et $y \in \mathcal{Y}$, on appelle *tranche de f par y* le système de vote f_y défini par :

$$f_y : \left| \begin{array}{ccc} \Omega & \to & \mathcal{C} \\ \omega & \to & f(y(\mathrm{P}(\omega))). \end{array} \right.$$

À présent, nous présentons un lemme qui donne une idée centrale pour le théorème de tranchage 5.9 : quand on est dans une configuration $y(p)$, les systèmes de vote f et f_y renvoient le même résultat ; mais pour les manipulateurs, leurs possibilités d'expression dans f_y sont incluses dans celles qu'ils ont dans f, donc ils ont moins de pouvoir.

Par exemple, considérons une configuration ω très particulière où pour chaque électeur v, son vecteur de notes sincère u_v est égal au vecteur de scores de Borda associés à la relation binaire p_v. Manifestement, si les électeurs votent sincèrement, alors le vote par notation et la méthode de Borda renvoient le même résultat. Nous remarquons simplement que si la méthode de Borda est manipulable en ω, alors le vote par notation l'est aussi : les manipulateurs peuvent utiliser les mêmes stratégies que celles qu'ils utiliseraient en méthode de Borda.

Lemme 5.3

Pour tout profil $p \in \mathrm{P}(\Omega)$ et toute méthode de tranchage $y \in \mathcal{Y}$, si f_y est manipulable en $y(p)$, alors f est manipulable en $y(p)$. Autrement dit, en utilisant l'indicatrice de manipulabilité, on a toujours :

$$\mathrm{MC}_{f_y}(y(p)) \leq \mathrm{MC}_f(y(p)).$$

2. En toute rigueur, cette affirmation est vraie si on a fait l'hypothèse que l'espace électoral autorise qu'un électeur attribue la même note à deux candidats tout en préférant l'un à l'autre. Sinon, il faut se limiter aux vecteurs de poids $\mathbf{x}$ strictement décroissants.

3. Par raccourci d'écriture, on associe un V-uplet de fonctions à la fonction multivariée correspondante.

Démonstration. Supposons que la tranche f_y est manipulable dans la configuration $\omega = y(p)$. Par définition, il existe des bulletins $\psi \in \Omega$ tels que $f_y(\psi) \neq f_y(\omega)$ et :

$$\forall v \in \mathrm{Sinc}(f_y(\omega) \to f_y(\psi)), \psi_v = \omega_v.$$

La tranche et le système de vote original ont le même résultat sincère en ω : en effet, en développant la définition de $f_y(\omega)$ et en utilisant la relation $\mathrm{P} \circ y = \mathrm{Id}$, on a $f_y(\omega) = f(y(\mathrm{P}(y(p)))) = f(y(p)) = f(\omega)$.

Notons $\phi = y(\mathrm{P}(\psi))$: il s'agit des bulletins qui sont effectivement pris en compte par f_y (dans notre exemple, la conversion des bulletins stratégiques ψ en format Borda). Par simple réécriture, on a :

$$\forall v \in \mathrm{Sinc}(f(\omega) \to f(\phi)), \psi_v = \omega_v.$$

Si $\psi_v = \omega_v$, alors $\phi_v = y_v(\mathrm{P}_v(\psi_v)) = y_v(\mathrm{P}_v(\omega_v)) = y_v(\mathrm{P}_v(y_v(p_v))) = y_v(p_v) = \omega_v$. Donc :

$$\forall v \in \mathrm{Sinc}(f(\omega) \to f(\phi)), \phi_v = \omega_v.$$

Or on a vu que $f(\phi) = f_y(\psi) \neq f_y(\omega) = f(\omega)$. Donc f est manipulable en ω vers ϕ. □

5.2 Espace électoral décomposable

En choisissant un couple (p, y) approprié, toute configuration ω peut être exprimée par $\omega = y(p)$. Par exemple, si le véritable état ω_v de chaque électeur v correspond à des scores de Borda, alors il peut être représenté par sa relation binaire de préférence sincère p_v et la méthode de tranchage y_v donnée en exemple précédemment. Par conséquent, dans cette configuration $\omega = y(p)$, le lemme 5.3 permet de comparer la manipulabilité de f et celle de f_y.

L'idée de la décomposabilité est la suivante : en tirant indépendamment p et y avec des distributions bien choisies, on souhaiterait reconstituer une configuration ω avec la bonne mesure de probabilité π. Si c'est possible, alors nous verrons (en prouvant le théorème de tranchage 5.9) que le taux de manipulabilité de f peut être comparé à une certaine moyenne de ceux de toutes les tranches possibles f_y.

Nous allons, d'abord, donner la définition formelle de la décomposabilité, puis une interprétation et un exemple. Rappelons que μ désigne la loi de P (en culture π).

Définition 5.4 (décomposabilité)
On dit que l'EEP (Ω, π) est P-*décomposable*, ou juste *décomposable*, ssi il existe une loi de probabilité ν sur $\mathcal{Y}$ telle que pour tout événement A sur Ω :

$$\pi(\omega \in A) = (\mu \times \nu)\big(y(p) \in A\big).$$

Dans ce qui suit, quand (Ω, π) est décomposable, on notera toujours ν une mesure arbitraire sur $\mathcal{Y}$, parmi celles qui vérifient cette propriété.

Cette définition exige que π soit la mesure image de $\mu \times \nu$ par l'opérateur qui, à p et y, associe $y(p)$. Autrement dit, en tirant indépendamment p et y (avec les mesures μ et ν respectivement), puis en considérant $\omega = y(p)$, on tire ω avec la mesure de probabilité correcte π. La décomposabilité est liée à la notion d'information complémentaire sur les états des électeurs : cette idée est davantage développée dans la section 5.6.

Par commodité de langage, quand l'EEP (Ω, π) est décomposable, nous dirons parfois que l'espace électoral Ω ou que la culture π est décomposable.

Exemple 5.5

Considérons $V = 2$ électeurs et $C = 2$ candidats nommés a et b. Supposons que l'état d'un électeur est un couple constitué d'un ordre strict total de préférence sur les candidats et d'une information complémentaire, « ananas » ou « banane ». Dans un cas d'étude réel, ces deux valeurs pourraient avoir une signification non alimentaire, comme « préfère passionnément » et « préfère un peu », mais cela n'a pas d'importance pour notre exemple.

Soit π la culture qui tire équiprobablement l'une des deux configurations suivantes :

1. Chaque électeur est dans l'état $\mathcal{A} = (a \succ b, \text{ananas})$;
2. Chaque électeur est dans l'état $\mathcal{B} = (b \succ a, \text{banane})$.

Pour montrer que cet EEP est décomposable, considérons la mesure ν qui tire, de façon certaine, deux fonctions identiques y_1 et y_2 telles que pour tout électeur v, $y_v(a \succ b) = \mathcal{A}$ et $y_v(b \succ a) = \mathcal{B}$.

En tirant le profil p avec la loi μ, on a de façon équiprobable $p = (a \succ b, a \succ b)$ ou $p = (b \succ a, b \succ a)$. Puis, en tirant y avec la loi (déterministe) ν, on a de façon équiprobable $y(p) = (\mathcal{A}, \mathcal{A})$ ou $y(p) = (\mathcal{B}, \mathcal{B})$, ce qui est exactement la culture π.

En bref, cet EEP (Ω, π) peut être émulé en tirant le profil $p = (p_1, p_2)$ avec la loi μ (qui est directement définie par la culture π), en tirant $y = (y_1, y_2)$ avec la loi ν (que nous avons exhibée), puis en considérant $\omega = y(p)$. Donc il est décomposable.

En général, il n'est pas trivial de déterminer si un espace électoral donné est décomposable ou pas. Pour cette raison, nous allons donner quelques conditions suffisantes ou nécessaires.

Proposition 5.6

Si les électeurs $(\omega_1, \dots, \omega_V)$ sont indépendants, alors (Ω, π) est décomposable.

Nous n'allons pas prouver la proposition 5.6 pour l'instant. C'est une conséquence de la proposition 5.18, que nous verrons en section 5.6 dans un cadre plus général.

Toutefois, l'indépendance n'est pas une condition nécessaire. En effet, dans l'exemple 5.5, les électeurs ne sont pas indépendants : ou bien ils sont tous les deux dans l'état $\mathcal{A}$, ou bien dans l'état $\mathcal{B}$. Cependant, l'EEP est décomposable.

Une autre condition suffisante est vérifiée par une classe importante de modèles. Par exemple, considérons un espace électoral où l'état de chaque électeur v est constitué d'un ordre strict total de préférence p_v et d'un entier $k_v \in [\![0, C]\!]$. Cet entier peut avoir la signification suivante : l'électeur v « approuve » les k_v premiers candidats de son ordre de préférence (quel que soit le sens exact de ce terme).

La culture π considérée est la suivante. On tire $(p_1, \dots, p_V)$ selon une certaine loi de probabilité μ sur $(\mathcal{L}_\mathcal{C})^V$. Indépendamment, on tire $(k_1, \dots, k_V)$ selon une certaine loi de probabilité ξ sur $[\![0, C]\!]^V$.

Remarquons que, aussi bien pour μ que pour ξ, les électeurs peuvent ne pas être indépendants entre eux. Mais les tirages par μ et ξ sont indépendants par hypothèse. La proposition suivante prouve qu'un tel EEP est décomposable.

Proposition 5.7

Pour tout $v \in \mathcal{V}$, soit $\mathcal{P}_v$ une partie non vide de $\mathcal{R}_\mathcal{C}$, soit $\mathcal{K}_v$ un ensemble mesurable non vide et soit $\Omega_v = \mathcal{P}_v \times \mathcal{K}_v$. Soit P_v la fonction définie par $\mathrm{P}_v(p_v, k_v) = p_v$ et K_v la fonction définie par $K_v(p_v, k_v) = k_v$.

Soit π une culture sur $\Omega = \prod_{v \in \mathcal{V}} \Omega_v$.

Si les deux variables aléatoires $\mathrm{P} = (\mathrm{P}_1, \dots, \mathrm{P}_V)$ *et* $K = (K_1, \dots, K_V)$ *sont indépendantes, alors* (Ω, π) *est décomposable.*

Démonstration. À $k_v \in \mathcal{K}_v$, on associe la fonction $\mathrm{concat}_v(k_v) \in \mathcal{Y}_v$ qui consiste à concaténer p_v et k_v de façon à reconstituer un état ω_v :

$$\mathrm{concat}_v(k_v) : \left| \begin{array}{ccl} \mathrm{P}_v(\Omega_v) & \to & \Omega_v \\ p_v & \to & \big(\mathrm{concat}_v(k_v)\big)(p_v) = (p_v, k_v). \end{array} \right.$$

À $k \in \mathcal{K}$, on associe $\mathrm{concat}(k) = (\mathrm{concat}_1(k_1), \dots, \mathrm{concat}_V(k_V)) \in \mathcal{Y}$: à chaque profil p, cette fonction se contente de juxtaposer le vecteur k de façon à renvoyer une configuration.

Alors, en notant ξ la loi de K, la mesure image ν de ξ par concat est clairement convenable pour prouver la décomposabilité. □

La condition de la proposition 5.7 est suffisante pour assurer la décomposabilité mais elle n'est pas nécessaire : quand chaque ensemble Ω_v est défini comme un produit cartésien $\mathcal{P}_v \times \mathcal{K}_v$ (où $\mathcal{P}_v \subseteq \mathcal{R}_\mathcal{C}$), il est possible que les variables aléatoires P et K ne soient pas indépendantes mais que l'espace électoral soit décomposable quand même. En effet, dans l'exemple 5.5, si $\mathrm{P} = (a \succ b, a \succ b)$, alors on sait de façon certaine que $K = (\text{ananas}, \text{ananas})$, alors que si $\mathrm{P} = (b \succ a, b \succ a)$, alors $K = (\text{banane}, \text{banane})$; par conséquent, P et K ne sont pas indépendants. Cependant, comme nous l'avons vu, cet EEP est décomposable.

Comme l'exemple didactique et fruitier 5.5, l'exemple suivant montre que, même sans vérifier les conditions suffisantes des propositions 5.6 et 5.7, l'EEP peut être décomposable. Mais cette fois, on utilise un modèle plus élaboré et plus susceptible d'être utilisé en pratique.

On tire $(p_1, \dots, p_V)$ selon une loi μ sur $(\mathcal{F}_\mathcal{C})^V$. Pour chaque électeur v, soit k_v le nombre de classes d'indifférence dans son ordre strict faible de préférence p_v ; par exemple, si l'électeur n'a pas d'égalité, le nombre de ses classes d'indifférence est simplement égal au nombre de candidats. On tire k_v notes selon une certaine loi $\xi_v(k_v)$ sur $[0,1]^{k_v}$. Cette loi $\xi_v(k_v)$ est choisie de telle sorte que les k_v notes sont presque sûrement toutes distinctes.

Étant donné un ordre strict faible p_v et un vecteur de notes $(u_1, \dots, u_{k_v})$, on construit un état ω_v en attribuant les notes aux candidats dans l'ordre de p_v.

Par exemple, si $p_v = (1 \sim 2 \succ 3 \succ 4)$, alors il y a $k_v = 3$ classes d'équivalence : $\{1,2\}$, $\{3\}$ et $\{4\}$. Donc on utilise la loi $\xi_v(3)$ pour tirer 3 notes, par exemple (0,1 ; 0,8 ; 0,2). Finalement, on a $\omega_v = (1 \sim 2 \succ 3 \succ 4, (0{,}8\ ;\ 0{,}8\ ;\ 0{,}2\ ;\ 0{,}1))$.

Pour prouver que cet EEP est décomposable, construisons une mesure ν convenable. Simultanément pour chaque entier $k_v \in [\![1, C]\!]$, tirons k_v notes selon $\xi_v(k_v)$. Ceci définit parfaitement y_v, une transformation qui associe à chaque p_v un état ω_v. Par définition, en tirant p avec la loi μ et y de cette façon, $\omega = y(p)$ est tirée avec la mesure de probabilité voulue.

Pour terminer, voici une condition nécessaire pour que l'EEP soit décomposable.

Proposition 5.8

Si (Ω, π) *est décomposable, alors pour tout sous-ensemble des électeurs* $\mathcal{V}'$*, pour tout événement* A *sur* $\prod_{v \in \mathcal{V}'} \Omega_v$ *(c'est-à-dire concernant seulement les électeurs de* $\mathcal{V}'$*), pour tout profil* $p = (p_1, \dots, p_V) \in \mathcal{R}$ *de probabilité non nulle :*

$$\pi\big(\omega_{\mathcal{V}'} \in A \mid P_{\mathcal{V}'} = p_{\mathcal{V}'}\big) = \pi\big(\omega_{\mathcal{V}'} \in A \mid P = p\big). \tag{5.1}$$

C'est une conséquence de la proposition 5.19, que nous verrons en section 5.6 dans un cadre plus général. L'interprétation intuitive est que, si on connaît les préférences ordinales p_v pour un sous-ensemble d'électeurs $\mathcal{V}'$ et si on souhaite reconstituer leurs états ω_v de façon probabiliste, alors connaître les relations p_v des autres électeurs ne fournit pas d'information. On notera qu'il n'est pas exclu, en revanche, que connaître les états complets ω_v des autres électeurs (et pas seulement leurs relations p_v) donne une information utile pour reconstituer les états des électeurs de $\mathcal{V}'$ à partir de leurs préférences ordinales.

L'avantage de cette condition est qu'en pratique, elle est plus commode à tester que la décomposabilité. Dans la section 5.6, nous montrerons que malheureusement, elle n'est pas suffisante pour assurer la décomposabilité.

5.3 Théorème de tranchage

Nous disposons, à présent, de tous les éléments pour démontrer le *théorème de tranchage*.

Théorème 5.9 (tranchage)
Soit (Ω, π) un EEP et f un SVBE dont le taux de manipulabilité est bien défini (c'est-à-dire que MC_f est mesurable, voir annexe A).

Si (Ω, π) est décomposable, alors il existe une méthode de tranchage $y \in \mathcal{Y}$ telle que la tranche de f par y a un taux de manipulabilité inférieur ou égal à celui de f :

$$\tau_{\mathrm{MC}}(f_y) \leq \tau_{\mathrm{MC}}(f).$$

Démonstration. Premièrement, montrons que le taux de manipulabilité de toute tranche f_y peut être exprimée en considérant un espace électoral ordinal (où chaque électeur est décrit uniquement par sa relation binaire de préférence), muni de la loi de probabilité μ. En effet, par définition :

$$\tau_{\mathrm{MC}}(f_y) = \int_{\omega \in \Omega} \mathrm{MC}_{f_y}(\omega) \pi(\mathrm{d}\omega).$$

Comme $\mathrm{MC}_{f_y}(\omega)$ dépend seulement du profil $\mathrm{P}(\omega)$ et comme ω et $y(\mathrm{P}(\omega))$ ont le même profil, on a :

$$\tau_{\mathrm{MC}}(f_y) = \int_{\omega \in \Omega} \mathrm{MC}_{f_y}(y(\mathrm{P}(\omega))) \pi(\mathrm{d}\omega).$$

D'où, par changement de variable :

$$\tau_{\mathrm{MC}}(f_y) = \int_{p \in \mathrm{P}(\Omega)} \mathrm{MC}_{f_y}(y(p)) \mu(\mathrm{d}p).$$

Deuxièmement, étudions la manipulabilité de f. En utilisant la décomposabilité, on a par changement de variable :

$$\tau_{\mathrm{MC}}(f) = \int_{(p,y) \in \mathrm{P}(\Omega) \times \mathcal{Y}} \mathrm{MC}_f(y(p)) (\mu \times \nu)(\mathrm{d}p, \mathrm{d}y).$$

Le théorème de Fubini-Tonelli donne :

$$\tau_{\mathrm{MC}}(f) = \int_{y \in \mathcal{Y}} \left(\int_{p \in \mathrm{P}(\Omega)} \mathrm{MC}_f(y(p)) \mu(\mathrm{d}p) \right) \nu(\mathrm{d}y).$$

Le lemme 5.3 assure que $\mathrm{MC}_f(y(p)) \geq \mathrm{MC}_{f_y}(y(p))$, ce qui conduit à :

$$\begin{aligned}\tau_{\mathrm{MC}}(f) &\geq \int_{y\in\mathcal{Y}} \left(\int_{p\in \mathrm{P}(\Omega)} \mathrm{MC}_{f_y}(y(p))\mu(\mathrm{d}p) \right) \nu(\mathrm{d}y) \\ &\geq \int_{y\in\mathcal{Y}} \tau_{\mathrm{MC}}(f_y)\nu(\mathrm{d}y).\end{aligned}$$

Ainsi, le taux de manipulabilité de f est supérieur ou égal à la moyenne (avec la pondération ν) des taux de manipulabilité des tranches f_y. Par conséquent, il existe au moins une tranche f_y telle que $\tau_{\mathrm{MC}}(f_y) \leq \tau_{\mathrm{MC}}(f)$. □

Dans la section 5.7, nous discuterons si l'hypothèse de décomposabilité est nécessaire pour que ce théorème soit valide et si on peut espérer une conclusion plus forte, à savoir la diminution de manipulabilité au sens ensembliste et non seulement probabiliste.

Dans l'immédiat, nous allons plutôt nous concentrer sur les conséquences du théorème de tranchage. En le combinant avec la proposition 5.6, on obtient le corollaire suivant.

Corollaire 5.10
Si les électeurs $(\omega_1, \ldots, \omega_V)$ sont indépendants, alors il existe $y \in \mathcal{Y}$ telle que $\tau_{\mathrm{MC}}(f_y) \leq \tau_{\mathrm{MC}}(f)$.

À ce stade, nous n'avons pas une idée précise de ce à quoi pourrait bien ressembler le système de vote f_y dont l'existence est garantie par le théorème 5.9 ou son corollaire 5.10 : nous savons seulement que c'est une tranche de f. Mais il pourrait avoir des propriétés indésirables : par exemple, il n'est pas explicitement exclu qu'il soit dictatorial, ce qui serait une façon particulièrement triviale et inintéressante de diminuer le taux de manipulabilité. Pour éviter ce type d'écueil, nous allons voir, à présent, que ces résultats de tranchage sont particulièrement pertinents quand on les combine avec les théorèmes de condorcification.

5.4 Condorcification et tranchage

Il est évident, mais essentiel, que si f vérifie le critère de Condorcet, alors il en est de même pour toute tranche f_y. En effet, s'il y a un vainqueur de Condorcet en ω, alors il l'est aussi dans $\omega' = y(\mathrm{P}(\omega))$, puisque le profil est le même dans les deux configurations ω et ω'. Ceci conduit à un autre corollaire du théorème 5.9.

Corollaire 5.11
Si (Ω, π) est décomposable et si f vérifie le critère de Condorcet, alors il existe un SVBE ordinal f', vérifiant le critère de Condorcet, tel que :

$$\tau_{\mathrm{MC}}(f') \leq \tau_{\mathrm{MC}}(f).$$

En effet, il suffit de considérer une tranche $f' = f_y$ dont le taux de manipulabilité est inférieur ou égal à celui de f. Son existence est garantie par le théorème de tranchage 5.9.

En combinant le théorème faible de condorcification 2.9 et le corollaire 5.11 du théorème de tranchage 5.9, on obtient le théorème suivant.

Théorème 5.12 (condorcification et tranchage)
Soit (Ω, π) un EEP et f un SVBE (tel que MC_f et MC_{f^*} soient mesurables, voir annexe A).
On suppose que :
— *(Ω, π) est décomposable ;*
— *f vérifie* **CMInf**.
Alors il existe un SVBE ordinal f', vérifiant le critère de Condorcet, tel que :

$$\tau_{\mathrm{MC}}(f') \leq \tau_{\mathrm{MC}}(f).$$

Démonstration. Par le théorème de Condorcification 2.9, on sait que $\tau_{\mathrm{MC}}(f^*) \leq \tau_{\mathrm{MC}}(f)$. En appliquant le corollaire 5.11 à f^*, on obtient un système de vote f' convenable. □

Par exemple, considérons de nouveau le vote par notation. Nous venons de prouver que dans toute culture décomposable, il existe un système de vote f' qui n'utilise pas des notes mais seulement des relations binaires de préférence, qui vérifie le critère de Condorcet et qui est au plus aussi manipulable que le vote par notation (au sens probabiliste).

La puissance du formalisme des espaces électoraux réside dans le fait que ce résultat ne se limite pas, par exemple, au vote par notation : il s'applique aux systèmes de vote dont les bulletins sont des notes, des valeurs d'approbation, des notes multicritères, des ananas, des bananes ou n'importe quel autre type d'objet (voir la section 1.1.4 pour divers exemples d'espaces électoraux). De plus, il n'y a absolument aucune hypothèse sur les relations binaires de préférence : elles peuvent violer la transitivité et même l'antisymétrie.

5.5 Théorème d'optimalité

La preuve du théorème de tranchage 5.9 n'est pas constructive : par conséquent, dans le théorème de condorcification et tranchage 5.12, nous ne savons malheureusement pas exhiber un système de vote f' qui soit ordinal, Condorcet et dont le taux de manipulabilité soit inférieur à celui du système original f.

Ceci dit, nous pensons que les résultats précédents ont des implications plus profondes. En effet, si l'espace électoral est décomposable et si on cherche un système de vote dans **CMInf** qui soit aussi peu manipulable que possible, alors ces résultats signifient qu'on peut restreindre l'investigation aux systèmes de vote ordinaux Condorcet.

De plus, puisqu'il existe un nombre fini de tels systèmes de vote (pour une valeur donnée de V et C), cette observation garantit l'existence d'un système de vote dont le taux de manipulabilité est minimal dans **CMInf**. Nous allons, à présent, formaliser ces constatations.

Définition 5.13 (image ordinale d'un EEP)
Pour un EEP $(V, C, \Omega, \mathrm{P}, \pi)$, son *image ordinale* est l'EEP $(V, C, \Omega', \mathrm{P}', \mu)$, où :
— Pour chaque électeur v, son espace d'états Ω'_v est constitué des préférences ordinales autorisées dans Ω_v : on a $\Omega'_v = \mathrm{P}_v(\Omega_v) \subseteq \mathcal{R}_{\mathcal{C}}$.
— Pour chaque électeur v, la fonction P'_v est la fonction identité.
— La loi μ est la loi de la variable P dans l'EEP original $(V, C, \Omega, \mathrm{P}, \pi)$.

Proposition 5.14
On considère un espace électoral (V, C, Ω', P') ordinal *: pour tout électeur v, on a $\Omega'_v \subseteq \mathcal{R}_{\mathcal{C}}$ et P'_v est la fonction identité. Soit μ une culture sur Ω'.*

Alors il existe un SVBE Condorcet g qui est le moins manipulable au sens probabiliste :

$$\tau^{\mu}_{\mathrm{MC}}(g) = \min_{\mathbf{Cond}(\Omega')}(\tau^{\mu}_{\mathrm{MC}}).$$

*On dit que g est τ^{μ}_{MC}-*optimal parmi les systèmes de vote Condorcet de Ω'.

Démonstration. Il existe un nombre fini de fonctions $g : \Omega \to \mathcal{C}$, *a fortiori* si nous exigeons qu'elles vérifient le critère de Condorcet. Par conséquent, au moins l'une d'entre elles minimise $\tau^{\mu}_{\mathrm{MC}}(g)$. □

Théorème 5.15 (optimalité)
Soit (Ω, π) un EEP. Soit g un système τ^{μ}_{MC}-optimal parmi les systèmes de vote Condorcet de l'image ordinale de (Ω, π).
Si (Ω, π) est décomposable, alors :

$$\tau^{\pi}_{\mathrm{MC}}(g \circ \mathrm{P}) = \min_{\mathbf{CMInf}(\Omega)}(\tau^{\pi}_{\mathrm{MC}}).$$

Autrement dit, g est optimal, non seulement parmi les systèmes de vote ordinaux Condorcet, mais parmi la classe plus large **CMInf** de tous les systèmes de vote qui vérifient le critère de la coalition majoritaire informée et qui peuvent ne pas être ordinaux. Comme nous l'avons vu, ceci inclut une part importante des systèmes de vote usuels de la littérature et de la vie réelle.

Démonstration. Soit $f \in \mathbf{CMInf}$. Par le théorème de condorcification et tranchage 5.12, il existe un SVBE ordinal f' (défini formellement sur Ω, mais ne dépendant que du profil P) qui vérifie le critère de Condorcet et tel que $\tau^{\pi}_{\mathrm{MC}}(f') \leq \tau^{\pi}_{\mathrm{MC}}(f)$. Et puisque g est τ^{μ}_{MC}-optimal, $\tau^{\pi}_{\mathrm{MC}}(g \circ \mathrm{P}) \leq \tau^{\pi}_{\mathrm{MC}}(f') \leq \tau^{\pi}_{\mathrm{MC}}(f)$. □

Par conséquent, dans l'optique de trouver un système de vote qui minimise le taux de manipulabilité dans **CMInf**, on peut restreindre l'investigation aux systèmes de vote ordinaux Condorcet, qui existent en nombre fini. Nous exploiterons intensivement cette propriété dans le chapitre 10, dédié à la recherche de systèmes de vote qui sont optimaux en ce sens.

5.6 Décomposabilité d'un ensemble probabilisé quelconque

Dans la section 5.2, nous avons défini un espace électoral décomposable et nous avons énoncé quelques propositions sans démonstration, afin de donner une vue d'ensemble de ces propriétés et de parvenir rapidement au théorème de tranchage 5.9 sans trop nous attarder sur des aspects techniques.

Dans cette section, nous généralisons la notion de décomposabilité à des ensembles probabilisés qui ne sont pas nécessairement des espaces électoraux et nous prouvons quelques résultats liés à cette notion dans ce cadre général.

5.6.1 Cas unidimensionnel : le lemme de la variable aléatoire complémentaire

Avant de définir un espace décomposable en toute généralité, nous allons examiner le cas particulier où il y a une seule dimension, ce qui correspond à un seul électeur pour l'application à un espace électoral. Pour cela, nous allons prouver ce que nous appelons le lemme de la *variable aléatoire complémentaire*, qui montre que dans ce cas, il existe toujours une mesure ν vérifiant une condition similaire à

celle de la définition 5.4 de la décomposabilité. Ceci nous permettra, par la suite, de traiter le cas où il y a plusieurs électeurs mais où ceux-ci sont indépendants.

Considérons un espace probabilisé (Ω, π) et une variable aléatoire X à valeurs dans un ensemble mesurable fini $\mathcal{X}$ muni de la tribu discrète (voir annexe A). On note μ la loi de X. Afin d'établir le parallèle avec les espaces électoraux, le lecteur peut se représenter $\omega \in \Omega$ comme l'état de l'unique électeur et X comme sa relation binaire de préférence.

Quand l'expérience aléatoire est réalisée, l'état du système obtenu est décrit par l'objet mathématique ω. La valeur $x = X(\omega)$ est une information partielle sur cet état : si on connaît seulement x, il manque généralement de l'information pour connaître parfaitement ω. Imaginons qu'il existe un espace $\mathcal{Y}$ qui permette d'exprimer cette information complémentaire : cela signifie que le couple (x, y) représente ω sans ambiguïté.

Imaginons, de plus, que les variables aléatoires x et y soient indépendantes : généralement, c'est une propriété puissante, puisqu'elle permet de traiter les deux variables séparément. En particulier, c'est une telle propriété d'indépendance, généralisée au cas de plusieurs électeurs, qui nous a permis de démontrer le théorème de tranchage 5.9.

La construction que nous allons considérer est une généralisation de cette notion d'information complémentaire. En effet, nous avons une liberté très importante : nous pouvons choisir l'ensemble $\mathcal{Y}$. Pour traiter la question en toute généralité, nous choisirons toujours l'ensemble des fonctions $y : X(\Omega) \to \Omega$ qui sont cohérentes avec X, au sens où $X \circ y = \mathrm{Id}$. En effet, c'est le cadre général pour que la donnée d'un x et d'un y définisse parfaitement un état ω qui soit cohérent avec x.

Le lemme suivant montre alors qu'il est toujours possible de choisir $y \in \mathcal{Y}$ qui vérifie les propriétés souhaitées. En pratique, on va anticiper chaque valeur de x possible et lui attribuer une des valeurs de ω cohérentes avec ce x, comme nous l'avons fait dans l'exemple fruitier 5.5. Comme les différentes valeurs de x ne sont jamais atteintes en même temps (l'électeur a toujours une valeur de x bien déterminée), nous allons voir qu'on peut les traiter de manière indépendante.

Lemme 5.16 (variable aléatoire complémentaire)
Soit (Ω, π) un ensemble probabilisé, X une variable aléatoire à valeurs dans un ensemble fini $\mathcal{X}$ muni de la tribu discrète, μ la loi de X et $\mathcal{Y} = \{y : X(\Omega) \to \Omega \text{ t.q. } X \circ y = Id\}$.

Alors il existe une mesure ν sur $\mathcal{Y}$ telle que pour tout événement A sur Ω :

$$\pi(\omega \in A) = (\mu \times \nu)\big(y(x) \in A\big).$$

Pour donner l'idée de la preuve, imaginons par exemple que nous ayons trois états possibles et deux valeurs de X correspondantes :

Probabilité	ω	X
π^1	ω^1	x^a
π^2	ω^2	x^b
π^3	ω^3	

Pour reconstruire un ω à partir d'un X, le plus simple est que y soit une fonction de $\mathcal{X}$ dans Ω. Il nous suffit de trouver une bonne façon de tirer aléatoirement les valeurs de cette fonction en x^a et en x^b. Quand X vaut x^a, c'est facile : on sait alors que $\omega = \omega^1$. On peut donc décider que pour tout y, on aura $y(x^a) = \omega^1$.

Quand X vaut x^b, on a $\omega = \omega^2$ avec une probabilité $\frac{\pi^2}{\pi^2+\pi^3}$ et $\omega = \omega^3$ le reste du temps. On va donc choisir $y(x^b) = \omega^2$ avec une probabilité $\frac{\pi^2}{\pi^2+\pi^3}$ et $y(x^b) = \omega^3$ le reste du temps. En résumé, on peut écrire informellement :

$$y : \left| \begin{array}{l} x^a \to \omega^1 \\ x^b \to \left| \begin{array}{l} \omega^2 \text{ avec probabilité } \frac{\pi^2}{\pi^2+\pi^3} \\ \omega^3 \text{ avec probabilité } \frac{\pi^3}{\pi^2+\pi^3}. \end{array} \right. \end{array} \right.$$

Le tirage (ici certain) de $y(x^a)$ et celui (ici incertain) de $y(x^b)$ permet de tirer une fonction y qui correspond aux conditions exigées.

S'il y avait une autre valeur image x^c, on tirerait la valeur de $y(x^c)$, indépendamment des valeurs $y(x^a)$ et $y(x^b)$, parmi les antécédents possibles de x^c, avec la loi de probabilité conditionnelle au fait que $X = x^c$.

Généralisons et formalisons cela.

Démonstration. Pour tout $x \in X(\Omega)$:

— Si $\pi(X = x) > 0$, on note π_x la mesure de probabilité conditionnelle sachant $X = x$ (restreinte à $X^{-1}(x)$) ;

— Si $\pi(X = x) = 0$, on choisit un $\omega_x \in X^{-1}(x)$ arbitrairement [4] et on note π_x la mesure de probabilité qui renvoie ω_x de façon sûre.

En identifiant toute fonction $y \in \mathcal{Y}$ à la liste de ses valeurs pour chaque argument possible x, on définit ν comme la mesure produit de toutes les π_x.

Alors, pour tout événement A sur Ω :

$$\begin{aligned} (\mu \times \nu)\big(y(x) \in A\big) &= \sum_{x \in X(\Omega)} \mu(\{x\}) \cdot \nu(y(x) \in A) \\ &= \sum_{\pi(X=x)>0} \pi(X = x) \cdot \pi(\omega \in A \mid X = x) \\ &= \pi(\omega \in A). \end{aligned}$$

Nous remercions Anne-Laure Basdevant et Arvind Singh pour de fructueuses discussions à propos de ce lemme. □

5.6.2 Cas multidimensionnel : la décomposabilité dans le cas général

À présent, nous traitons la décomposabilité dans le cas général multidimensionnel, qui correspond au cas avec plusieurs électeurs dans l'application aux espaces électoraux.

Soit $V \in \mathbb{N} \setminus \{0\}$. Pour chaque $v \in [\![1, V]\!]$, soit Ω_v un ensemble mesurable et $X_v : \Omega_v \to \mathcal{X}_v$ une fonction mesurable, où $\mathcal{X}_v$ est un ensemble fini muni de la tribu discrète.

Soit π une mesure de probabilité sur l'univers $\Omega = \prod_{v=1}^{V} \Omega_v$. On note $X = (X_1, \dots, X_V)$ et μ la loi de X.

Pour chaque v, on note $\mathcal{Y}_v = \{y_v : X_v(\Omega_v) \to \Omega_v \text{ t.q. } X_v \circ y_v = \mathrm{Id}\}$. On note $\mathcal{Y} = \prod_{v=1}^{V} \mathcal{Y}_v$. Pour $(x, y) \in X(\Omega) \times \mathcal{Y}$, on note $y(x) = (y_1(x_1), \dots, y_V(x_V))$.

Définition 5.17 (décomposabilité dans le cas général)

On dit que (Ω, π) est *X-décomposable* ssi il existe une mesure ν sur $\mathcal{Y}$ telle que pour tout événement A sur Ω :

$$\pi(\omega \in A) = (\mu \times \nu)\big(y(x) \in A\big).$$

4. Il n'y a pas de problème avec l'axiome du choix car il s'agit de faire un nombre fini de choix.

La difficulté provient de notre exigence que les variables aléatoires complémentaires y_v soient *individuelles* : y ne peut pas être une fonction quelconque de $X(\Omega)$ vers Ω, mais doit être un V-uplet de fonctions, où chaque y_v est définie de $X_v(\Omega_v)$ vers Ω_v. En effet, dans la preuve du lemme 5.3, qui est la pierre angulaire du théorème de tranchage 5.9, nous avons besoin de variables aléatoires individuelles (pour traiter les électeurs sincères). Si nous demandions une variable aléatoire *collective* y qui, à partir de x, permet de reconstituer l'état ω avec la loi de probabilité voulue, ce serait toujours possible, par une application directe du lemme 5.16 de la variable aléatoire complémentaire.

Nous pouvons à présent généraliser et démontrer la proposition 5.6 : celle-ci énonçait que si les électeurs sont indépendants, alors l'espace électoral est décomposable.

Proposition 5.18
Si les variables aléatoires $(\omega_1, \dots, \omega_V)$ sont indépendantes, alors (Ω, π) est X-décomposable.

Démonstration. Il suffit d'appliquer le lemme 5.16 pour chaque $v \in [\![1, V]\!]$, ce qui définit une mesure ν_v sur chaque ensemble $\mathcal{Y}_v$. Ensuite, on définit ν comme la mesure produit des ν_v. □

La proposition suivante généralise et prouve la proposition 5.8, qui donne une condition nécessaire de décomposabilité.

Proposition 5.19
Si (Ω, π) est X-décomposable alors, pour chaque sous-ensemble $\mathcal{V}'$ de $[\![1, V]\!]$, pour chaque événement A sur $\prod_{v \in \mathcal{V}'} \Omega_v$, pour chaque $x = (x_1, \dots, x_V)$ de probabilité non nulle :

$$\pi\big(\omega_{\mathcal{V}'} \in A \mid X_{\mathcal{V}'} = x_{\mathcal{V}'}\big) = \pi\big(\omega_{\mathcal{V}'} \in A \mid X = x\big). \tag{5.2}$$

Démonstration. D'une part :

$$\begin{aligned}\pi\big(\omega_{\mathcal{V}'} \in A \mid X = x\big) &= (\mu \times \nu)\big(y_{\mathcal{V}'}(x_{\mathcal{V}'}) \in A \mid X = x\big) \\ &= \nu\big(y_{\mathcal{V}'}(x_{\mathcal{V}'}) \in A\big).\end{aligned}$$

D'autre part, on obtient de manière similaire :

$$\pi\big(\omega_{\mathcal{V}'} \in A \mid X_{\mathcal{V}'} = x_{\mathcal{V}'}\big) = \nu\big(y_{\mathcal{V}'}(x_{\mathcal{V}'}) \in A\big).$$

□

Cependant, la condition (5.2) n'assure pas que (Ω, π) soit X-décomposable. Comme contre-exemple, prenons $V = 2$. L'état ω_v (avec $v = 1$ ou $v = 2$) peut prendre 4 valeurs, notées de ω_v^1 à ω_v^4. La variable X_v peut prendre 2 valeurs, x_v^a et x_v^b. La table suivante définit la correspondance entre les états ω_v et les variables x_v, ainsi que la mesure π.

π	$\omega_1^1 \to x_1^a$	$\omega_1^2 \to x_1^a$	$\omega_1^3 \to x_1^b$	$\omega_1^4 \to x_1^b$
$\omega_2^1 \to x_2^a$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	0
$\omega_2^2 \to x_2^a$	$\frac{1}{16}$	$\frac{1}{16}$	0	$\frac{1}{8}$
$\omega_2^3 \to x_2^b$	$\frac{1}{8}$	0	0	$\frac{1}{8}$
$\omega_2^4 \to x_2^b$	0	$\frac{1}{8}$	$\frac{1}{8}$	0

Cette table se lit de la façon suivante : dans l'en-tête de la première colonne de données, on lit que si $\omega_1 = \omega_1^1$, alors $x_1 = x_1^a$; de même, dans l'en-tête de la première ligne de données, on lit que si $\omega_2 = \omega_2^1$, alors $x_2 = x_2^a$. À l'intersection de cette colonne et de cette ligne, on voit que la probabilité de l'état (ω_1^1, ω_2^1) est $\frac{1}{16}$.

Il est fastidieux mais facile de vérifier que la condition (5.2) est vérifiée. En effet, par exemple :

$$\pi(\omega_1^1 \mid x_1^a \wedge x_2^a) = \frac{1}{2} = \pi(\omega_1^1 \mid x_1^a).$$

À présent, supposons que (Ω, π) est X-décomposable. Soit ν une mesure convenable pour cette décomposition. Pour $(\alpha, \beta, \gamma, \delta) \in \{1,2\} \times \{3,4\} \times \{1,2\} \times \{3,4\}$, adoptons le raccourci de notation suivant :

$$\nu(\alpha, \beta, \gamma, \delta) = \nu\Big(y_1(x_1^a) = \omega_1^\alpha \ \wedge \ y_1(x_1^b) = \omega_1^\beta \ \wedge \ y_2(x_2^a) = \omega_2^\gamma \ \wedge \ y_2(x_2^b) = \omega_2^\delta\Big).$$

Par exemple, $\nu(\mathbf{1}, \mathbf{4}, \mathbf{2}, \mathbf{4})$ représente la probabilité que, lorsqu'on tire y avec la mesure de probabilité ν, on ait $y_1(x_1^a) = \omega_1^{\mathbf{1}}$ et $y_1(x_1^b) = \omega_1^{\mathbf{4}}$ et $y_2(x_2^a) = \omega_2^{\mathbf{2}}$ et $y_2(x_2^b) = \omega_2^{\mathbf{4}}$: cette valeur de y correspond au choix des colonnes **1** et **4** et des lignes **2** et **4** dans la table.

On note les faits suivants.

— $0 = \pi(\omega_1^3 \wedge \omega_2^3 \mid x_1^b \wedge x_2^b) = \sum_{(\alpha,\gamma)} \nu(\alpha, 3, \gamma, 3)$. Comme tous les termes de cette somme sont positifs ou nuls, ils sont tous nuls et on a en particulier $\nu(1,3,1,3) = 0$.

— $0 = \pi(\omega_1^1 \wedge \omega_2^4 \mid x_1^a \wedge x_2^b) = \sum_{(\beta,\gamma)} \nu(1, \beta, \gamma, 4)$ donc $\nu(1,3,1,4) = 0$ et $\nu(1,4,1,4) = 0$.

— $0 = \pi(\omega_1^4 \wedge \omega_2^1 \mid x_1^b \wedge x_2^a) = \sum_{(\alpha,\delta)} \nu(\alpha, 4, 1, \delta)$ donc $\nu(1,4,1,3) = 0$.

Par conséquent, $\frac{1}{4} = \pi(\omega_1^1 \wedge \omega_2^1 \mid x_1^a \wedge x_2^a) = \sum_{(\beta,\delta)} \nu(1, \beta, 1, \delta) = \nu(1,3,1,3) + \nu(1,3,1,4) + \nu(1,4,1,3) + \nu(1,4,1,4) = 0$: cette contradiction prouve que (Ω, π) n'est pas décomposable. Ainsi, la condition (5.2) est nécessaire mais pas suffisante pour que (Ω, π) soit X-décomposable.

Dans le contre-exemple ci-dessus, on peut montrer qu'il existe une *mesure signée* ν qui vérifie la relation définissant la décomposabilité. Malheureusement, nous avons besoin d'une mesure au sens habituel, c'est-à-dire une mesure positive : en effet, la preuve du théorème de tranchage 5.9 utilise la propriété de croissance de l'intégrale, qui est elle-même basée sur sa propriété de positivité.

5.7 Discussion des hypothèses et des conclusions du théorème de tranchage

Dans cette section, nous allons examiner si le théorème de tranchage 5.9 est bien ajusté ou s'il est possible de le généraliser, soit en renforçant les conclusions, soit en affaiblissant les hypothèses.

5.7.1 Diminution de la manipulabilité au sens probabiliste ou ensembliste ?

Dans le théorème de tranchage 5.9, on conclut à l'existence d'une tranche f_y qui est moins manipulable que le système original f au sens probabiliste, c'est-à-dire telle que $\tau^\pi_{\mathrm{MC}}(f_y) \leq \tau^\pi_{\mathrm{MC}}(f)$ pour une culture π donnée.

La conclusion serait plus forte si, comme pour les théorèmes de condorcification 2.9 et 2.20, le système de vote obtenu était moins manipulable que f au sens

ensembliste, c'est-à-dire si on avait $\mathrm{MC}_{f_y} \subseteq \mathrm{MC}_f$. Avec cette conclusion plus forte, le théorème 5.9 prendrait la forme du « proto-théorème » 5.20 ci-dessous.

Proto-théorème 5.20 (tranchage, avec diminution de la manipulabilité au sens ensembliste)

Soit (Ω, π) un EEP et f un SVBE dont le taux de manipulabilité est bien défini (c'est-à-dire que MC_f est mesurable, voir annexe A).

Si (Ω, π) est décomposable, alors il existe une méthode de tranchage $y \in \mathcal{Y}$ telle que la tranche de f par y est au plus aussi manipulable que f :

$$\mathrm{MC}_{f_y} \subseteq \mathrm{MC}_f.$$

Dans ce proto-théorème, il est vraisemblable que les hypothèses de mesurabilité et de décomposabilité ne servent à rien, puisque la conclusion n'a rien de probabiliste, mais nous allons les conserver au cas (*a priori* peu crédible) où elles seraient d'une quelconque utilité. Si ce proto-théorème était vrai, alors on aurait immédiatement le corollaire suivant, qui est obtenu de la même façon que le théorème de condorcification et tranchage 5.12.

Proto-théorème 5.21 (condorcification et tranchage, avec diminution de la manipulabilité au sens ensembliste)

Soit (Ω, π) un EEP et f un SVBE (tel que MC_f et MC_{f^*} soient mesurables, voir annexe A).

On suppose que :

— *(Ω, π) est décomposable ;*

— *f vérifie* **CMInf**.

Alors il existe un SVBE ordinal f', vérifiant le critère de Condorcet, tel que :

$$\mathrm{MC}_{f'} \subseteq \mathrm{MC}_f.$$

Nous allons montrer que le proto-théorème 5.21 est faux et, par conséquent, que le proto-théorème 5.20 l'est également.

Considérons l'espace électoral suivant. On considère $V = 3$ électeurs et $C = 3$ candidats a, b, c. L'état de chaque électeur est donné par un ordre strict total de préférence et un bit dont la valeur est 0 ou 1. Cet espace électoral est similaire au générateur aléatoire décentralisé de l'exemple 1.9. Afin d'éviter les problèmes de mesurabilité, on munit l'espace électoral de la tribu discrète (cf. annexe A), ce qui assure que toute fonction est mesurable. Définissons la culture π : pour chaque électeur indépendamment, on tire son état de façon uniforme. D'après la proposition 5.6, ceci assure que l'EEP (Ω, π) est décomposable.

Voici le SVBE f que nous considérons.

1. S'il y a un vainqueur de Condorcet, il est élu.
2. Si la configuration est non Condorcet et si au moins deux bits valent 1, alors c est élu.
3. Si la configuration est non Condorcet et si au moins deux bits valent 0, alors b est élu.

Comme f vérifie le critère de Condorcet, il vérifie aussi **CMInf**. Donc les hypothèses du proto-théorème 5.21 sont bien toutes réunies. Nous allons voir que cependant, la conclusion n'est pas vérifiée.

Considérons les trois configurations suivantes ω, ϕ et ψ.

Configuration	Électeur			Vainqueur de Condorcet
	1	2	3	
ω	a c , 0 b	b c , 1 a	c a , 1 b	c
ϕ	a b , 1 c	b a , 0 c	c a , 1 b	a
ψ	a b , 0 c	b c , 0 a	c b , 1 a	b

f n'est pas manipulable en ω. En effet, d'une part, pour manipuler en faveur de b, l'unique électeur intéressé ne peut pas empêcher que a ou c soit vainqueur de Condorcet. D'autre part, pour manipuler en faveur de a, l'unique électeur intéressé ne peut rendre a vainqueur de Condorcet car il aura toujours une victoire contre c, et il n'existe pas d'autre façon de faire gagner a.

De même, on peut montrer que f n'est manipulable ni en ϕ ni en ψ.

Considérons un système f' ordinal vérifiant le critère de Condorcet, et la famille suivante de configurations χ (les bits des électeurs n'ont pas d'importance pour f').

Configuration	Électeur			Vainqueur de Condorcet
	1	2	3	
χ	a b c	b c a	c a b	Aucun

Si $f'(\chi) = a$ (resp. b, c), alors f' est manipulable en ω (resp. ϕ, ψ) vers χ. Par conséquent, f' est nécessairement manipulable en au moins une des trois configurations ω, ϕ, ψ, et il est donc impossible d'avoir $\mathrm{MC}_{f'} \subseteq \mathrm{MC}_f$. Par conséquent, le proto-théorème 5.21 est faux, ce qui entraîne que le proto-théorème 5.20 l'est également.

Ainsi, dans le théorème de tranchage 5.9 et le théorème de condorcification et tranchage 5.12, on ne peut pas renforcer la conclusion en exigeant que le système de vote obtenu soit au plus aussi manipulable que le système original f au sens ensembliste. On doit se contenter d'obtenir un système qui est au plus aussi manipulable que f au sens probabiliste.

5.7.2 Hypothèse de décomposabilité

Dans la section précédente, nous avons vu qu'on ne pouvait pas renforcer la conclusion des théorèmes 5.9 et 5.12 en exigeant une diminution de manipulabilité au sens ensembliste. À présent, examinons s'il est possible de généraliser ces théorèmes d'une autre façon : en affaiblissant les hypothèses, et en particulier en supprimant l'hypothèse centrale qu'est la décomposabilité de l'EEP (Ω, π). Sans cette hypothèse, le théorème de tranchage 5.9 prend alors la forme du proto-théorème 5.22 ci-dessous.

Proto-théorème 5.22 (tranchage, sans la décomposabilité)
Soit (Ω, π) un EEP et f un SVBE dont le taux de manipulabilité est bien défini (c'est-à-dire que MC_f est mesurable, voir annexe A).

Alors il existe une méthode de tranchage $y \in \mathcal{Y}$ telle que la tranche de f par y a un taux de manipulabilité inférieur ou égal à celui de f :

$$\tau_{\mathrm{MC}}(f_y) \leq \tau_{\mathrm{MC}}(f).$$

Si ce proto-théorème était vrai, alors on pourrait immédiatement généraliser le théorème de condorcification et tranchage 5.12 sous la forme du proto-théorème suivant.

Proto-théorème 5.23 (condorcification et tranchage, sans la décomposabilité)

⌋ Soit (Ω, π) un EEP et f un SVBE (tel que MC_f et MC_{f^*} soient mesurables, voir annexe A).

On suppose que f vérifie **CMInf**.

Alors il existe un SVBE ordinal f', vérifiant le critère de Condorcet, tel que :

$$\tau_{\mathrm{MC}}(f') \leq \tau_{\mathrm{MC}}(f).$$

Pour prouver que ce résultat est faux, reprenons le contre-exemple de la section 5.7.1. Mais définissons, à présent, la culture π par $\pi(\omega) = \pi(\phi) = \pi(\psi) = \frac{1}{3}$. Alors les hypothèses du proto-théorème 5.23 sont toutes vérifiées.

Le système de vote initial f a un taux de manipulabilité nul : bien qu'il soit manipulable dans certaines configurations, il est presque sûrement non manipulable car toute la probabilité est concentrée sur les configurations ω, ϕ et ψ, qui sont non manipulables. Cependant, nous avons vu que dans tout système f' ordinal Condorcet, au moins l'une de ces configurations est manipulable, donc f' a un taux de manipulabilité d'au moins $\frac{1}{3}$. Ceci prouve que la conclusion du proto-théorème 5.23 n'est pas vraie.

Par conséquent, le proto-théorème 5.23 est faux, ce qui prouve que le proto-théorème 5.22 l'est également.

Ainsi, le théorème de tranchage 5.9 et sa séquelle, le théorème de condorcification et tranchage 5.12, ne sont plus valides si on enlève l'hypothèse de décomposabilité purement et simplement, c'est-à-dire sans la remplacer par une autre hypothèse éventuellement plus faible.

Il reste cependant une question naturelle : est-ce que les théorèmes 5.9 et 5.12 restent vrais avec une hypothèse plus faible que la décomposabilité, par exemple la condition nécessaire (5.1) présentée dans la proposition 5.8 ? Nous laissons cette question ouverte pour de futurs travaux.

Deuxième partie

Étude assistée par ordinateur de la manipulabilité

Chapitre 6

SVAMP : *Simulator of Various Voting Algorithms in Manipulating Populations*

Comme nous l'avons déjà évoqué, l'histoire de la théorie du vote a été marquée par la découverte de plusieurs paradoxes, notamment le paradoxe de Condorcet (1785) ou « paradoxe du vote » sur les comparaisons entre paires de candidats ; le paradoxe d'Arrow (1950), dont un point clé est l'indépendance des alternatives non pertinentes (IANP) ; et le théorème d'impossibilité de Gibbard (1973) et Satterthwaite (1975) sur la manipulabilité.

Puisqu'aucun système de vote non trivial ne peut éviter ces paradoxes totalement, la probabilité d'apparition de ces cas pathologiques dans divers types de cultures a fait l'objet d'intenses investigations (Campbell et Tullock, 1965; Chamberlin *et al.*, 1984; Nitzan, 1985; Lepelley et Mbih, 1987; Kelly, 1993; Aleskerov et Kurbanov, 1999; Smith, 1999; Lepelley et Valognes, 2003; Favardin et Lepelley, 2006; Pritchard et Wilson, 2007; Aleskerov *et al.*, 2008; Reyhani *et al.*, 2009; Jennings, 2011; Reyhani, 2013; Green-Armytage, 2014; Wang *et al.*, 2014). Cependant, il reste d'importantes questions ouvertes dans le domaine, en particulier à propos des performances relatives des différents systèmes de vote selon divers critères et sous divers jeux d'hypothèses sur les préférences des électeurs.

Récemment, grâce au développement de la communauté du choix social computationnel, d'intéressants résultats ont été publiés sur les questions algorithmiques liées aux systèmes de vote et à leur manipulation. D'une part, on s'est intéressé aux problèmes de complexité, suivant la ligne de recherche initiée par Bartholdi *et al.* (1989b) puis Bartholdi et Orlin (1991), et on a montré que déterminer la manipulabilité était $\mathcal{NP}$-difficile pour certains modes de scrutin. D'autre part, certains auteurs, comme Xia *et al.* (2009); Walsh (2010b); Zuckerman *et al.* (2009, 2011); Gaspers *et al.* (2013), ont mis au point des algorithmes explicites pour traiter des questions de manipulabilité.

Cependant, à notre connaissance, il n'existait pas de logiciel disponible publiquement qui se repose sur ces techniques existantes, en particulier pour l'étude de la manipulabilité. Cette observation nous a conduit à développer SVAMP (*Simulator of Various Voting Algorithms in Manipulating Populations*), un package Python dédié à l'étude des systèmes de vote et de leur manipulabilité. Il est facile à utiliser, propose d'établir les préférences des électeurs par un large choix de modèles pour les préférences des électeurs ou par l'importation de données externes et

il implante une variété de systèmes de vote. Enfin, il est simple d'étendre SVAMP par l'adjonction de nouveaux modèles de population et de nouveaux systèmes de vote.

SVAMP est un logiciel libre, sous la *GNU General Public License version 3*. Sa documentation inclut la procédure d'installation, des tutoriels, un guide de référence et des instructions pour les contributeurs. Elle est disponible à l'adresse suivante :

https://svvamp.readthedocs.org.

Après un exemple basique destiné à donner une vue d'ensemble de l'utilisation de SVAMP, nous présentons les outils pour la classe Population, représentant un ensemble d'électeurs munis de préférences sur un ensemble de candidats (section 6.1). Puis nous présentons les outils pour la classe Election, qui est une classe abstraite dont tous les systèmes de vote héritent (section 6.2). Nous illustrons nos choix d'implantation en nous concentrant sur l'exemple de la manipulation par coalition (MC), dont l'étude est un des buts principaux du logiciel (section 6.3). Nous indiquons les algorithmes spécifiques utilisés pour certains modes de scrutin (section 6.4). Enfin, nous nous concentrons sur l'implantation de VTI et de ses variantes (section 6.5) : en effet, nous verrons dans les chapitres suivants que ce mode de scrutin est particulièrement résistant à la manipulation. C'est pourquoi il a fait l'objet d'une attention particulière.

Donnons une vue d'ensemble rapide de l'utilisation de SVAMP à travers un exemple basique.

Après avoir chargé SVAMP, définissons une population aléatoire de $V = 5$ électeurs munis de préférences sur $C = 4$ candidats en utilisant le modèle *Spheroid* (ellipsoïdal), qui est une extension de l'habituelle *culture impartiale* (définition 1.14) et que nous présenterons davantage en section 6.1.3.

```
>>> import svvamp
>>> pop = svvamp.PopulationSpheroid(V=5, C=4)
```

Ceci crée un objet `pop` de classe Population. Ensuite, nous pouvons trier les électeurs par ordre lexicographique sur leurs ordres de préférence et afficher ces ordres.

```
>>> pop.ensure_voters_sorted_by_rk()
>>> print(pop.preferences_rk)
[[0 2 1 3]
 [1 3 2 0]
 [1 3 2 0]
 [2 0 1 3]
 [3 2 0 1]]
```

Chaque ligne de la matrice attribut `preferences_rk` représente l'ordre strict total de préférence d'un électeur. Par exemple, le premier électeur préfère le candidat 0, puis 2, puis 1, puis 3. Suivant les conventions habituelles en Python, la numérotation des candidats et des électeurs démarre à 0.

Créons une élection avec cette population d'électeurs, en utilisant le système de vote *Plurality* (uninominal), et déterminons le vainqueur sincère. Celui-ci, noté w dans SVAMP et dans l'ensemble de ce chapitre, est le candidat 1 dans notre exemple.

```
>>> election = svvamp.Plurality(pop)
>>> print(election.w)
1
```

SWAMP peut nous dire que l'élection est MC et nous donner des détails sur ce type de manipulation.

```
>>> is_CM, log_CM, candidates_CM = election.CM_with_candidates
    ()
>>> print(is_CM)
True
>>> print(candidates_CM)
[ 1.  0.  1.  0.]
```

Les candidats 0 ou 2 peuvent bénéficier d'une MC, mais pas le candidat 3. Le candidat 1, c'est-à-dire w, est exclu par convention de la définition de MC.

6.1 Population

Dans cette section, nous présentons comment étudier les propriétés d'une population en poursuivant notre premier exemple (section 6.1.1). Ensuite, nous montrons que SWAMP n'est pas limité aux ordres stricts totaux de préférence mais peut également exploiter des utilités et des ordres faibles de préférence (section 6.1.2). Enfin, nous présentons un panel d'outils pour définir des populations (section 6.1.3).

6.1.1 Obtenir des informations sur une population

Un objet de classe Population possède plusieurs fonctions et attributs qui permettent d'étudier ses propriétés.

Les scores de Borda attribués par les électeurs aux candidats sont directement implantés en tant qu'attribut dans la classe Population, parce qu'ils sont utilisés par plusieurs systèmes de vote.

```
>>> print(pop.preferences_borda_rk)
[[3 1 2 0]
 [0 3 1 2]
 [0 3 1 2]
 [2 1 3 0]
 [1 0 2 3]]
```

Le score de Borda total et le score par *Plurality*[1] de chaque candidat sont disponibles de la même façon, en utilisant les attributs `borda_score_c_rk` et `plurality_scores_rk` respectivement.

D'autres attributs fournissent la *matrice des duels*, la *matrice des victoires*[2] et l'identité du vainqueur de Condorcet.

```
>>> print(pop.matrix_duels_rk)
[[0 3 1 2]
 [2 0 2 4]
 [4 3 0 2]
 [3 1 3 0]]
```

1. À titre de rappel, le *score par Plurality* (uninominal) d'un candidat c est le nombre d'électeurs qui placent c en premier dans leur ordre de préférence.
2. Pour toute paire de candidat distincts (c, d), `matrix_victories_rk[c, d]` vaut :
   — 1 ssi `matrix_duels_rk[c, d]` $> V/2$,
   — 0,5 ssi `matrix_duels_rk[c, d]` $= V/2$,
   — 0 ssi `matrix_duels_rk[c, d]` $< V/2$.

Par convention, les coefficients diagonaux valent 0.

```
>>> print(pop.matrix_victories_rk)
[[ 0.  1.  0.  0.]
 [ 0.  0.  0.  1.]
 [ 1.  1.  0.  0.]
 [ 1.  0.  1.  0.]]
>>> print(pop.condorcet_winner_rk)
nan
```

Dans cet exemple, la dernière valeur renvoyée est `nan` (*not a number*), ce qui est une convention signifiant qu'il n'y a pas de vainqueur de Condorcet.

La matrice des duels, la matrice des victoires et le vainqueur de Condorcet ont plusieurs variantes implantées dans SVAMP, qui dépendent de choix de départage des égalités. En particulier, les notions de candidat Condorcet-admissible, vainqueur de Condorcet faible et vainqueur de Condorcet relatif sont implantées. Le lecteur intéressé se reportera à la documentation de SVAMP pour davantage de détails.

6.1.2 Travailler avec des utilités

La plupart des études sur la vraisemblance d'une manipulation, citées dans l'introduction de ce chapitre, ne concernent que des systèmes de vote ordinaux. Cependant, d'intéressants systèmes de vote sont cardinaux. Un cas particulier en est le vote par assentiment, qui est un important sujet de recherche (Brams et Fishburn (1978); Fishburn et Brams (1981); Brams (1994, 2003); Laslier et Van der Straeten (2008); Laslier (2009)). Il est aussi utilisé dans des élections de la vie réelle : par exemple, il a été choisi par les théoriciens du choix social pour leurs véritables élections au sein de la *Society of Social Choice and Welfare*.

Un des objectifs de SVAMP est de permettre la comparaison de systèmes de vote ordinaux et cardinaux sur une même population d'électeurs. Pour cela, les électeurs n'ont pas seulement des ordres de préférence, mais aussi des utilités sur les candidats. Par conséquent, nous pouvons définir une population en fournissant de telles utilités.

```
>>> pop1 = svvamp.Population(preferences_ut=
                             [[1,  1,  0],
                              [2,  0, -1]])
```

Les utilités étendent l'ensemble des préférences possibles, par rapport aux ordres stricts totaux de préférence. Elles fournissent une mesure d'intensité et elle permettent à un électeur d'être indifférent entre plusieurs candidats, ce qui donne une façon pratique de manipuler des ordres faibles de préférence. Dans `pop1`, si le candidat 0 est le vainqueur sincère d'une élection, alors le premier électeur n'est pas intéressé par une manipulation en faveur du candidat 1, et vice versa.

Cependant, pour certains systèmes de vote, les électeurs sont obligés de fournir un ordre strict de préférence : dans ce cas, chaque électeur doit casser les égalités dans ses propres préférences. Pour cette raison, si une population a été définie uniquement par ses utilités, le classement de chaque électeur est obtenu en départageant ses égalités au hasard, une fois pour toutes, la première fois où l'attribut `preferences_rk` est appelé.

```
>>> print(pop1.preferences_rk)
[[1 0 2]
 [0 1 2]]
```

Le but de ce choix de spécification est que la comparaison des systèmes de vote basés sur des ordres stricts se fasse uniquement sur la base de leur comportement

spécifique, sans introduire un biais dû au fait que les électeurs sincères donnent des bulletins différents aux divers systèmes de vote.

À l'inverse, si une population est définie uniquement par des classements, SVVAMP peut remplir automatiquement la matrice des utilités. Bien sûr, choisir des utilités cohérentes avec les classement est nécessairement arbitraire ; par convention, SVVAMP utilise, dans ce cas, des utilités égales aux scores de Borda.

```
>>> pop2 = svvamp.Population(preferences_rk=
                             [[0, 1, 2, 3],
                              [2, 0, 1, 3]])
>>> print(pop2.preferences_ut)
[[3 2 1 0]
 [2 1 3 0]]
```

Il est, enfin, possible de fournir des utilités et des classements simultanément, ces derniers indiquant comment chaque électeur départage ses propres égalités dans son bulletin sincère, quand un système de vote basé sur des classements est utilisé.

```
>>> pop3 = svvamp.Population(
            preferences_ut=[[1, 1,  0],
                            [2, 0, -1]],
            preferences_rk=[[0, 1, 2],
                            [0, 1, 2]])
```

Les classements doivent être cohérents avec les utilités, au sens où si un électeur v place un candidat c devant un candidat d dans son classement, alors son utilité pour c doit être supérieure ou égale à son utilité pour d. Si ce n'est pas le cas, SVVAMP lève une erreur.

Certains attributs mentionnés dans la section 6.1.1 diffèrent selon que l'on considère des utilités ou les classements des électeurs. Par exemple, l'attribut `preferences_borda_ut` est basé sur les utilités [3] et traite les indifférences comme telles. En revanche, l'attribut `preferences_borda_rk` est basé sur les classements.

```
>>> print(pop3.preferences_borda_ut)
[[ 1.5  1.5  0. ]
 [ 2.   1.   0. ]]
>>> print(pop3.preferences_borda_rk)
[[2 1 0]
 [2 1 0]]
```

Des variantes similaires avec les suffixes `_rk` (basées sur les ordres stricts totaux) et `_ut` (basées sur les utilités et permettant des ordres faibles) sont implantées pour la matrice des duels, la matrice des victoires et les notions Condorcet.

6.1.3 Créer un objet Population

Jusqu'à présent, mis à part notre exemple introductif, nous avons défini des populations manuellement, en utilisant la fonction d'initialisation de la classe Population. Ceci fournit une façon facile de définir des exemples simples. Afin de faire des études à plus large échelle, SVVAMP permet également d'importer une population depuis un fichier externe ou d'utiliser une panel de modèles probabilistes pour générer des populations aléatoires.

3. Pour chaque électeur v et candidat c, `preferences_borda_ut[v, c]` est la somme de :
   — 1 point pour chaque candidat d tel que v a une utilité strictement plus grande pour c que pour d ;
   — Et 0,5 point pour chaque candidat d tel que v a la même utilité pour c et pour d.

Importer une population depuis un fichier externe est immédiat : SWAMP peut lire de simples fichiers au format CSV contenant les utilités de la population ou des fichiers dans le format de la base de données PrefLib (Mattei et Walsh, 2013), sur laquelle nous reviendrons dans le chapitre 9.

Pour générer des populations artificielles, SWAMP implante une variété de cultures, c'est-à-dire de générateurs aléatoires utilisant un modèle probabiliste.

Spheroid (ellipsoïdal) et *Cubic Uniform* (cubique uniforme) : ces deux modèles étendent la culture impartiale (définition 1.14) aux utilités. Pour les classements, ces deux modèles sont équivalents à la culture impartiale. Nous étudierons le modèle ellipsoïdal dans le chapitre 7.

Ladder (« échelle », au sens propre) : comme en culture impartiale, les électeurs sont indépendants et le modèle est anonyme et neutre (au sens où il traite les électeurs de façon symétrique, et de même pour les candidats), mais un électeur peut être indifférent entre plusieurs candidats. Ce modèle peut donc être vu comme une extension de la culture impartiale utilisant des ordres faibles. En pratique, l'utilité de chaque électeur pour chaque candidat est tirée de façon indépendante et uniforme dans un ensemble fini de valeurs (les « barreaux » de l'échelle). Lors de la phase de développement, cette culture est souvent utilisée pour vérifier que SWAMP a le comportement désiré quand les électeurs sont indifférents entre plusieurs candidats. Afin de ne pas alourdir les chapitres suivants, nous n'en ferons pas une étude détaillée dans ce mémoire.

Gaussian Well (puits gaussien) et *Euclidean box* (boîte euclidienne) : ces deux modèles attribuent des positions aléatoires aux électeurs et aux candidats dans un espace euclidien appelé le *spectre politique*. L'utilité d'un électeur pour un candidat est une fonction décroissante de la distance entre eux. Si la dimension du spectre politique est égale à 1, alors les préférences sont nécessairement *single-peaked* (voir définition 1.11 et Black, 1958). Ces modèles seront l'objet du chapitre 8.

Von Mises–Fisher : ce modèle est similaire au modèle de Mallows (1957), mais il est adapté à des populations définies par leurs utilités. Il y a un point spécial de l'espace des utilités, le *pôle*, où la densité de probabilité est maximale, et un paramètre de concentration κ qui permet à SWAMP de contrôler la forme de la distribution (Watson et Williams, 1956). Il est possible de créer une population constituée de plusieurs sous-groupes utilisant différents pôles avec différences concentrations. Un exemple d'une telle population est donné dans la figure 6.1, elle-même générée par SWAMP. Chaque point rouge représente le vecteur d'utilité d'un électeur. Les distributions de Von-Mises Fisher sont générées en utilisant l'algorithme d'Ulrich (1984) modifié par Wood (1994). Nous étudierons ce modèle dans le chapitre 7.

6.2 Elections

Dans SWAMP, Election est une classe abstraite. L'utilisateur final manipule toujours une de ses sous-classes implantant les élections utilisant un système de vote particulier (uninominal, assentiment, etc.), qui hérite des attributs et des méthodes de la classe Election.

Dans cette section, nous verrons comment SWAMP permet d'étudier le résultat d'une élection, différentes notions de manipulation et l'indépendance des alternatives non pertinentes (IANP ou, en anglais : *Independence of Irrelevant Alternatives*, *IIA*).

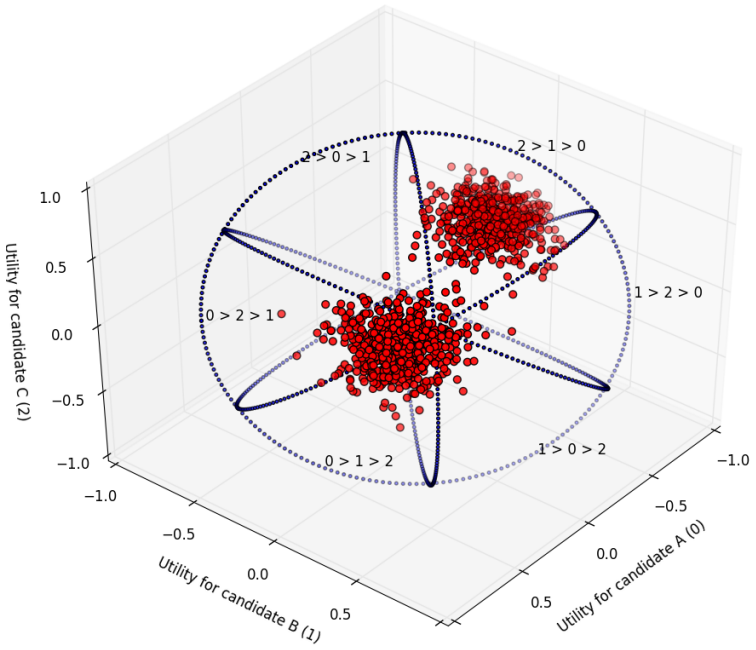


FIGURE 6.1 – Une population utilisant le modèle de Von Mises–Fisher. $V = 1000$ électeurs, $C = 3$ candidats, deux pôles aléatoires de même concentration $\kappa = 50$.

6.2.1 Résultat d'une élection

Divers attributs permettent d'étudier le résultat d'une élection quand les électeurs sont sincères.

```
>>> pop = svvamp.PopulationSpheroid(V=3, C=5)
>>> print(pop.preferences_rk)
[[0 2 3 4 1]
 [0 2 4 3 1]
 [2 1 3 4 0]]
>>> election = svvamp.Plurality(pop)
>>> print(election.ballots)
[0 0 2]
>>> print(election.scores)
[2 0 1 0 0]
>>> print(election.w)
0
```

Nous obtenons le bulletin sincère de chaque électeur, le score de chaque candidat et le vainqueur de l'élection. Les questions d'égalité seront discutées dans la section 6.3.3. SVVAMP fournit également les attributs `candidates_by_scores_best_to_worst` et `scores_best_to_worst` pour trier les candidats en fonction de leur résultat dans l'élection.

Les types des attributs `ballots`, `scores` et `scores_best_to_worst` dépendent du système de vote. Par exemple, pour le vote par assentiment (*Approval voting*), `ballots` est une matrice $V \times C$ de booléens et `ballots[v, c]` vaut **True** ssi l'électeur v approuve le candidat c.

On pourra se reporter à la documentation pour une liste exhaustive des attributs et plus de détails sur chaque système de vote.

6.2.2 Manipulation par coalition (MC)

Afin d'illustrer la manipulation par coalition, considérons l'exemple suivant. Rappelons qu'*IRV* est le sigle anglophone pour le vote à tours instantanés (VTI).

```
>>> pop = svvamp.PopulationSpheroid(V=1000,
                                    C=10)
>>> election = svvamp.IRV(pop)
```

Examinons la manipulabilité par coalition de cette élection.

```
>>> is_CM, log_CM = election.CM()
>>> print(is_CM)
nan
```

Pour chaque fonction, SVAMP utilise par défaut son algorithme le plus précis parmi ceux qui fonctionnent en temps polynomial. Dans le cas d'IRV (VTI), décider MC est $\mathcal{NP}$-complet (Bartholdi et Orlin, 1991), donc cet algorithme polynomial n'est pas exact. Pour cette raison, `is_CM` peut être un booléen (est-ce que l'élection est manipulable ou pas), ou la valeur conventionnelle `nan` signifiant que l'algorithme n'a pas été capable de décider.

On peut choisir un algorithme exact en modifiant l'attribut `CM_option`.

```
>>> election.CM_option = 'exact'
>>> is_CM, log_CM = election.CM()
>>> print(is_CM)
True
```

Pour les systèmes de vote où il y a un algorithme exact fonctionnant en temps polynomial, `'exact'` est la seule option disponible.

Comme nous l'avons vu en introduction de ce chapitre, SVAMP peut préciser en faveur de quels candidats l'élection est manipulable.

```
>>> is_CM, log_CM, candidates_CM = election.CM_with_candidates
    ()
>>> print(candidates_CM)
[ 0.  0.  0.  1.  0.  0.  0.  0.  0.  0.]
```

6.2.3 Variantes de la MC (MCI, MU, MT)

SVAMP implante également trois variantes de MC.

On dit qu'une élection est *manipulable par coalition ignorante* (MCI, ou en anglais : *Ignorant-Coalition Manipulable*, *ICM*) en faveur d'un candidat $c \neq w$ ssi les électeurs qui préfèrent c au vainqueur sincère w peuvent utiliser des stratégies telles que, quelles que soient les stratégies choisies par les autres électeurs, le candidat c soit déclaré vainqueur. C'est donc une notion de manipulation inspirée du critère **CMIgn**.

On dit qu'une élection est *manipulable-unisson* (MU, ou en anglais : *Unison-Manipulable*, *UM*) en faveur du candidat c ssi il existe une stratégie telle que, quand tous les électeurs qui préfèrent strictement c au vainqueur sincère w utilisent cette même stratégie (pendant que les autres électeurs continuent de voter sincèrement), le candidat c est déclaré vainqueur. La terminologie « unisson » est empruntée à Walsh (2010b). On peut faire le parallèle avec le critère **BulMajUni** pour l'aspect unisson, mais il est important de souligner les différences. D'une part, on impose que tous les électeurs intéressés participent à la manipulation et utilisent le même bulletin (avant tout pour permettre un calcul plus facile) ; d'autre part, il s'agit ici de manipulation informée (c'est-à-dire à bulletins des

autres électeurs constants), alors que le critère **BulMajUni** concerne une manipulation ignorante (qui fonctionne quels que soient les bulletins des électeurs sincères).

Enfin, on dit qu'une élection est *manipulable trivialement* (MT, ou en anglais : *Trivially Manipulable*, *TM*) en faveur du candidat $c \neq w$ ssi, quand tous les électeurs qui préfèrent strictement c à w utilisent leur *stratégie triviale* (pendant que les autres électeurs continuent de voter sincèrement), le candidat c est vainqueur.

Ce que nous appelons *stratégie triviale* pour l'électeur v en faveur du candidat c contre le candidat w dépend du type de bulletin. Pour les modes de scrutin ordinaux, c'est le bulletin où c est classé premier (« compromis »), w est classé dernier (« enterrement »), et où les autres candidats restent dans l'ordre sincère de préférence de v. Pour les systèmes de vote basés sur les notes, c'est le bulletin où v attribue la note maximale autorisée à c (par défaut 1) et la note minimale aux autres candidats (par défaut -1).

La stratégie triviale paraît naturelle quand l'électeur v souhaite faire gagner c et sait que w est un adversaire fort, sans aucun autre indice sur les chances des autres candidats. Elle requiert peu de coordination entre les manipulateurs. Pour cette raison, quand une MC est possible, la MT est un bon indicateur pour savoir si la manipulation est facile ou pas, en termes d'échange d'information entre les manipulateurs. L'autre avantage évident de cette notion est qu'elle peut être calculée en temps polynomial (à condition que calculer le vainqueur d'une élection soit polynomial, ce qui est généralement le cas, sauf pour des systèmes de vote comme la méthode de Kemeny).

Ces trois variantes de MC — MCI, MU et MT — utilisent la même syntaxe que MC. Par exemple, MU peut être décidée avec les fonctions `UM` et `UM_with_candidates` et son algorithme peut être choisi avec `UM_option`. Comme pour MC, leur option par défaut est d'utiliser l'algorithme polynomial le plus précis disponible.

6.2.4 Manipulation individuelle (MI)

On dit qu'une élection est *manipulable individuellement* (MI, ou en anglais : *Individually Manipulable*, *IM*) ssi un électeur v, en produisant un bulletin non sincère, peut faire gagner un candidat c qu'il préfère strictement au vainqueur sincère w (pendant que les autres électeurs continuent de voter sincèrement).

La fonction qui fournit le plus d'information sur MI est `IM_full`.

```
>>> pop = svvamp.PopulationSpheroid(V=4, C=3)
>>> election = svvamp.IRV(pop)
>>> election.w
2
>>> election.IM_option = 'exact'
>>> is_IM, log_IM, candidates_IM, voters_IM, v_IM_for_c =
    election.IM_full()
>>> print(is_IM)
True
>>> print(candidates_IM)
[ 1.  0.  0.]
>>> print(voters_IM)
[ 0.  0.  1.  1.]
>>> print(v_IM_for_c)
[[ 0.  0.  0.]
 [ 0.  0.  0.]
 [ 1.  0.  0.]
 [ 1.  0.  0.]]
```

Dans cet exemple, l'élection est MI. Plus précisément, seul le premier candidat peut en bénéficier. Cette manipulation peut être accomplie par l'un quelconque des deux derniers électeurs, alors que les deux premiers ne peuvent pas ou ne sont pas intéressés.

Si seul un sous-ensemble de ces informations est requis, des fonctions dédiées permettent de les obtenir. Elles n'effectuent que les calculs nécessaires, ce qui les rend plus rapides que `IM_full`.

6.2.5 Indépendance des alternatives non pertinentes (IANP)

Rappelons que l'élection est *indépendante des alternatives non pertinentes* (IANP, ou en anglais : *Independent of Irrelevant Alternatives*, *IIA*) ssi, quand on effectue l'élection avec les mêmes électeurs et tout sous-ensemble de candidats contenant le vainqueur sincère w, celui-ci reste vainqueur.

Ici, nous étudions l'IANP pour une instance spécifique d'élection utilisant une population donnée, pas l'IANP en tant que propriété d'un système de vote en général.

```
>>> not_IIA, log_IIA, example_subset_IIA, example_winner_IIA =
    election.not_IIA_full()
>>> print(not_IIA)
True
>>> print(example_subset_IIA)
[ True  True False]
>>> print(example_winner_IIA)
0
```

`not_IIA` vaut **True**, ce qui signifie que l'élection viole la propriété d'IANP (par cohérence avec les méthodes de manipulation comme MC, **True** correspond au comportement « indésirable »). Quand c'est le cas, `example_subset_IIA` fournit un sous-ensemble de candidats qui viole l'IANP : ici, le sous-ensemble contenant les candidats 0 et 1. `example_winner_IIA` fournit le vainqueur correspondant, en l'occurrence le candidat 0.

Si l'élection ne viole pas IANP, les variables de contre-exemple renvoient la valeur conventionnelle `nan`.

6.2.6 Critères vérifiés par un système de vote

Les systèmes de vote possèdent des attributs spéciaux qui représentent diverses propriétés. L'utilisateur final peut accéder à ces attributs à travers n'importe quel objet Election, mais ils concernent le système de vote lui-même, pas une instance particulière d'élection. Ils sont surtout utiles pour les développeurs de nouveaux systèmes de vote, qui peuvent renseigner ces attributs afin d'aider SVAMP à accélérer ses calculs (cf. section 6.3.4).

Par exemple, l'attribut suivant signifie que VTI (IRV) est basé sur des classements stricts, pas sur des utilités.

```
>>> election = svvamp.IRV(pop)
>>> print(election.is_based_on_rk)
True
```

L'attribut suivant signifie que le système de vote est basé sur des utilités et que, pour un manipulateur unique ou pour une coalition de manipulateurs en faveur du candidat c, il est optimal d'agir comme s'ils avaient une utilité de 1 pour c et -1 pour les autres candidats. Par exemple, c'est le cas pour le vote par

assentiment et le vote par notation dans leur configuration par défaut [4]. Bien sûr, cet attribut vaut **False** pour IRV.

```
>>> print(election.is_based_on_ut_minus1_1)
False
```

Les quatre critères suivants sont intimement liés et sont utilisés pour le calcul de MC et de ses variantes. Il s'agit des critères majoritaires principaux étudiés au chapitre 3.

```
>>> print(election.meets_Condorcet_c_rk)
False
>>> print(election.meets_majority_favorite_c_rk)
True
>>> print(election.meets_IgnMC_c)
True
>>> print(election.meets_InfMC_c)
True
```

Dans SVAMP, plusieurs variantes de ces critères existent, en fonction de la distinction habituelle entre classement et utilités, et en fonction de questions de départage d'égalité entre les candidats. Voir la documentation pour plus de détails.

SVAMP gère les implication entre ces critères : par exemple, un développeur peut se contenter d'informer SVAMP qu'un certain système de vote vérifie une variante donnée de **FavMaj**. SVAMP sait alors quelles autres variantes de **FavMaj** et quelles déclinaisons de **CMIgn** et **CMInf** cela implique.

6.3 Implantation de MC

Dans cette section, nous nous concentrons sur l'implantation de MC, qui illustre au mieux les techniques générales utilisées dans SVAMP.

6.3.1 Minimisation des calculs

Lors du calcul de MC et de ses variantes, si l'élection n'est pas manipulable, il est généralement nécessaire de boucler sur tous les candidats pour le prouver. Mais si l'élection est manipulable, il est suffisant de prouver la manipulation en faveur d'un seul candidat. Pour cette raison, il est intéressant de deviner pour quels candidats une manipulation est la plus vraisemblable et de tester ces candidats en premier. Dans les méthodes génériques de la classe Election, on utilise une heuristique simple : les candidats sont examinés dans l'ordre décroissant du nombre de manipulateurs qui leur correspond. Ce comportement peut facilement être redéfini quand on implante un système de vote en particulier.

De manière générale, SVAMP essaie d'être aussi paresseux que possible. Cela signifie par exemple : 1. ne pas faire de calculs inutiles et 2. ne pas faire le même calcul deux fois.

4. Dans la configuration par défaut pour le vote par notation et le jugement majoritaire, on considère que les vecteurs d'utilité sont définis à une transformation affine strictement positive près, comme dans le modèle classique de Von Neumann et Morgenstern (annexe B). Donc prétendre qu'on a une utilité de 1 pour c et -1 pour les autres revient à dire qu'on a strictement la meilleure utilité pour c et une même utilité pour les autres.

Dans la configuration par défaut du vote par assentiment, on considère que les vecteurs d'utilité sont définis à une constante multiplicative positive près, comme dans le modèle avec limite d'approbation de la section B.7, et que le bulletin sincère d'un électeur ne dépend que du signe des utilités. Donc prétendre qu'on a une utilité de 1 pour c et -1 pour les autres revient à dire qu'on donne son assentiment à c, et uniquement à c.

Ceci peut être illustré par l'exemple suivant.

```
>>> is_CM, log_CM = election.CM()
>>> print(is_CM)
True
>>> is_CM, log_CM, candidates_CM = election.CM_with_candidates
    ()
>>> print(candidates_CM)
[0. 0. 1. 0. 1.]
```

1. Lorsque `CM` est appelé, voici une exécution possible : SWAMP examine le candidat 0 en premier, parce qu'il a le plus de manipulateurs, mais la manipulation est prouvée impossible. Puis SWAMP examine le candidat 2 et trouve une manipulation. SWAMP arrête le calcul et décide que `is_CM` est **True**.
2. Lorsque `CM_with_candidates` est appelé, SWAMP se rappelle le résultat du calcul pour les candidats 0 et 2 et il effectue le calcul uniquement pour les autres candidats.

6.3.2 Nombre de manipulateurs

Quand on calcule MU ou MT pour un candidat donné c, on utilise un ensemble fixe de manipulateurs : les électeurs qui préfèrent effectivement c au vainqueur sincère w. Cependant, MC et MCI fonctionnent différemment. En fait, pour étudier MC ou MCI en faveur d'un candidat c, SWAMP traite la question suivante. Étant donné la sous-population des électeurs sincères, c'est-à-dire ceux qui ne préfèrent pas strictement c à w, quel est le nombre de manipulateurs minimal qu'il faut ajouter pour effectuer la manipulation[5] ? Notons cet entier x_c.

Pendant le calcul, SWAMP maintient à jour une borne inférieure et une borne supérieure de x_c. Dès que la borne inférieure devient plus grande que le nombre de manipulateurs, la manipulation est prouvée impossible. À l'inverse, dès que la borne supérieure devient inférieure ou égale au nombre de manipulateurs, la manipulation pour c est prouvée.

Ce mécanisme permet à SWAMP de combiner différents algorithmes et d'accélérer les calculs, même quand l'attribut `CM_option` a la valeur `'exact'`. Par exemple, en utilisant VTI en mode exact, SWAMP commence par utiliser un algorithme polynomial approché pour diminuer l'intervalle entre les bornes (nous y reviendrons en section 6.5). Si cela suffit à décider MC, le calcul s'arrête. Sinon, l'algorithme exact non-polynomial est utilisé et il exploite les bornes déjà connues pour accélérer le calcul.

6.3.3 Anonymat et caractère résolu

Le mécanisme d'ajout de manipulateurs que nous venons de décrire explique un important choix de SWAMP : les modes de scrutin doivent être *anonymes*, au sens où ils traitent tous les électeurs de la même façon. Ce choix permet à SWAMP de travailler sur le nombre de manipulateurs, pas sur leur identité. Si les électeurs avaient des poids différents, beaucoup de problèmes faciles deviendraient difficiles : par exemple, même pour un nombre fixe de 3 candidats, décider MC avec des électeurs pondérés pour Borda, Maximin, VTI ou Veto est $\mathcal{NP}$-complet (Conitzer

5. Cette question est appelée le problème d'optimisation non-pondéré constructif par coalition (*Constructive Coalitional Unweighted Optimization* problem) par Zuckerman *et al.* (2009). Il permet de décider MC pour tous les systèmes de vote actuellement implantés dans SWAMP.

et Sandholm, 2002; Conitzer *et al.*, 2003, 2007), alors qu'avec des électeurs non pondérés, le problème est dans $\mathcal{P}$ (par rapport au nombre d'électeurs V).

Nous supposons également que les modes de scrutin sont déterministes et résolus : ils élisent un unique candidat. Sinon, définir la manipulation est une question en soi, parce qu'il est nécessaire de préciser comment les préférences sur les candidats s'étendent à des préférences sur les ensembles de candidats (ou les distributions de probabilité sur les candidats), ce qui peut être fait d'une infinité de façons. On pourra en trouver des exemples chez Gärdenfors (1976); Gibbard (1978); Duggan et Schwartz (2000). Utiliser des modes de scrutin résolus permet à SVAMP de décider MC et ses variantes en bouclant sur les candidats susceptibles de bénéficier d'une manipulation (c.a.d. tous sauf le vainqueur sincère), en considérant pour chacun l'ensemble des électeurs intéressés par le candidat en question. En revanche, si le mode de scrutin renvoie un sous-ensemble de candidats (resp. une distribution de probabilité sur les candidats), alors il est déraisonnable (resp. impossible) de boucler sur tous les résultats possibles.

Puisque les modes de scrutin sont anonymes et résolus, ils sont nécessairement non neutres en général (voir proposition 1.16) : en effet, ils doivent renvoyer un unique vainqueur, y compris dans des configurations totalement symétriques comme l'exemple minimal habituel de paradoxe de Condorcet (avec $V = 3$ et $C = 3$). Dans SVAMP, la règle de départage habituelle est que le candidat avec le plus petit index est favorisé. Par exemple, s'il y a une égalité entre les candidats 0 et 1, alors le candidat 0 est déclaré vainqueur. Nous avons choisi cette règle pour sa simplicité : avec des règles plus compliquées, certains problèmes de manipulation habituellement dans $\mathcal{P}$ deviennent $\mathcal{NP}$-complets (Obraztsova *et al.*, 2011; Obraztsova et Elkind, 2012; Aziz *et al.*, 2013). Cependant, il n'y a pas de limitation architecturale dans SVAMP qui empêche d'implanter des systèmes de vote utilisant d'autres règles de départage.

6.3.4 Sous-fonctions de MC

SVAMP est écrit de façon modulaire, en utilisant une approche orientée objet qui est naturelle dans le langage Python. Par exemple, la fonction `CM` (ou sa version détaillée `CM_with_candidates`) définie dans la classe Election appelle, en fait, un certain nombre de sous-fonctions spécifiques. Chacune de ces sous-fonctions peut être redéfinie dans la sous-classe qui implante un mode de scrutin en particulier, tout en conservant les autres sous-fonctions par défaut.

À titre indicatif, on peut diviser les sous-fonctions en trois catégories : tests préliminaires généraux, tests préliminaires dédiés à un candidat, noyau de calcul.

Les tests préliminaires généraux ont pour but de décider MC immédiatement si possible.

- Si le système de vote vérifie **FavMaj** (section 6.2.6) et si w est un favori majoritaire[6], alors MC vaut **False**.
- Si le système de vote vérifie **Cond** et si w est vainqueur de Condorcet résistant, alors MC vaut **False**.
- Si le système de vote vérifie **CMInf** et si w n'est pas Condorcet-admissible, alors MC vaut **True**.
- Si MT, MU ou MCI vaut **True**, alors MC vaut **True**.

Si nécessaire, SVAMP effectue des tests préliminaires qui sont dédiés à la manipulation en faveur d'un candidat donné c.

6. Pour chaque test préliminaire, il y a diverses variantes qui dépendent de questions de départage d'égalité. Par souci de concision et de clarté, nous ne décrivons pas ces variantes ici. Pour plus de détails, on pourra consulter le code de la classe Election.

— Si le système de vote vérifie **CMInf**, alors il est suffisant d'avoir plus de manipulateurs que d'électeurs sincères.
— Si le système de vote vérifie **FavMaj**, alors il est nécessaire d'avoir assez de manipulateurs pour empêcher le vainqueur sincère w d'apparaître comme un favori majoritaire.
— Si le système de vote vérifie **Cond**, alors il est nécessaire d'avoir assez de manipulateurs pour empêcher w d'apparaître comme un vainqueur de Condorcet (on peut faire le parallèle entre ce test et la notion de vainqueur de Condorcet résistant).
— Si MT ou MU est possible pour c, alors MC est possible pour c.
— Le nombre de manipulateurs suffisant pour MCI est également suffisant pour MC.

À ce stade, si l'attribut `CM_option` a la valeur `'lazy'`, l'algorithme s'arrête, que la solution ait été trouvée ou pas. Dans ce dernier cas, on renvoie la valeur conventionnelle `nan`.

Si l'attribut `CM_option` a la valeur `'exact'`, alors la classe Election possède un algorithme par défaut, qui peut être redéfini dans chaque système de vote pour plus d'efficacité.

Si l'attribut `is_based_on_rk` vaut **True**, l'algorithme exact par défaut de la classe générique Election utilise la force brute : les manipulateurs essaient tous les bulletins possibles. Pour un candidat c, en notant n_c le nombre de manipulateurs pour c contre w, cet algorithme par défaut a une complexité en temps de l'ordre de $(C!)^{n_c}$ (multiplié par un polynôme et par le temps nécessaire pour calculer le vainqueur d'une élection), ce qui le rend utilisable seulement pour de petites configurations [7].

Si l'attribut `is_based_on_ut_minus1_1` vaut **True**, l'algorithme exact par défaut est évident : les manipulateurs ont simplement besoin de tester le bulletin sincère correspondant à une utilité de 1 pour c et une utilité de -1 pour tous les autres candidats. Ceci fonctionne en temps polynomial (multiplié par le temps nécessaire pour calculer le vainqueur d'une élection).

Dans les autres cas (c'est-à-dire pour un système qui ne repose pas sur des classements et où la stratégie optimale de manipulation pour un certain candidat c ne consiste pas à agir comme si on avait une utilité de 1 pour c et -1 pour les autres candidats), cette méthode lève une erreur NotImplemented.

Dans la sous-classe qui implante un système de vote en particulier, la sous-fonction implantant l'algorithme exact est fréquemment redéfini pour bénéficier d'un algorithme plus efficace que la force brute.

6.3.5 Performances de l'algorithme générique polynomial

En utilisant l'option `'lazy'`, l'algorithme générique pour MC fonctionne en temps polynomial mais peut ne pas décider. La figure 6.2 donne son efficacité mesurée sur des systèmes de vote pour lesquels aucun algorithme spécifique n'est implanté dans SWAMP [8]. Pour ces systèmes de vote, seule la règle de calcul du vainqueur est fournie à SWAMP, ainsi que les propriétés de base du système de vote (`is_based_on_rk`, `meets_Condorcet_c`, etc.). Par conséquent, l'algorithme

7. La complexité est en fait un peu inférieure car on teste les bulletins possibles à permutation près des manipulateurs.

8. Parmi les systèmes représentés dans la figure 6.2, on sait que le problème de la manipulation est $\mathcal{NP}$-difficile pour les méthodes de Baldwin et Nanson (Narodytska *et al.*, 2011; Davies *et al.*, 2014), PO (Xia *et al.*, 2009; Hemaspaandra *et al.*, 2013; Parkes et Xia, 2012) et la méthode de Kemeny. Pour VTID, CSD et BI, la question est ouverte.

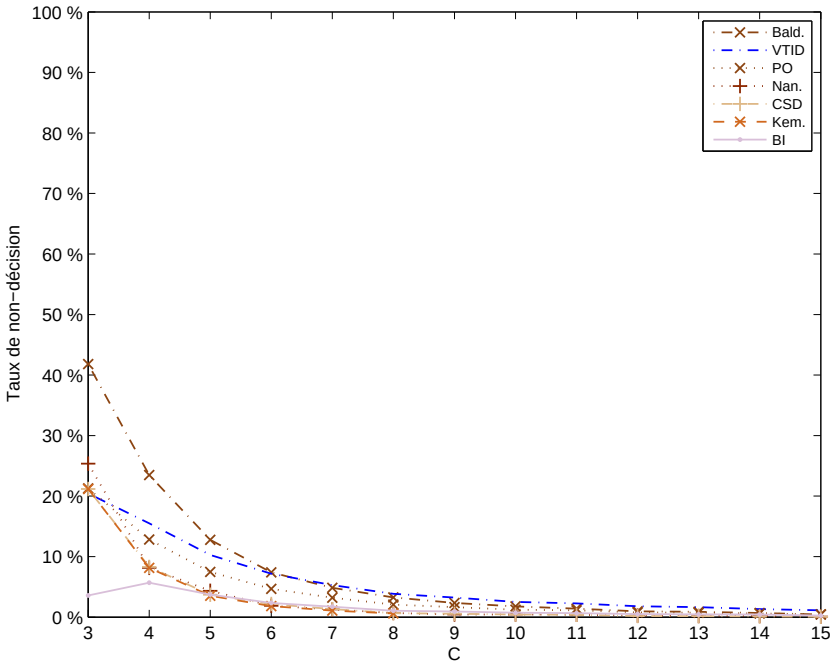


FIGURE 6.2 – Performances de l'algorithme polynomial par défaut pour MC. Modèle *Spheroid*, $V = 33$.

générique `'lazy'` repose avant tout sur les tests préliminaires exposés en 6.3.4, dont le plus important est MT.

Pour obtenir la figure 6.2, pour chaque valeur de C, 10 000 populations aléatoires ont été générées par le modèle *Spheroid* avec $V = 33$. La valeur sur l'axe des ordonnées est le ratio de populations pour lesquelles la fonction `CM` a renvoyé la valeur `nan`, signifiant qu'elle n'a été capable de prouver ni MC ni sa négation.

On voit que l'algorithme paresseux générique est assez efficace pour décider MC. Empiriquement, en comparant avec l'algorithme exact, nous avons remarqué que l'algorithme paresseux est efficace pour prouver MC, mais l'est moins pour prouver la négation de MC. Afin de ne pas faire de publicité mensongère, il nous faut mentionner que dans les simulations conduisant à cette figure, les fréquences de MC étaient relativement élevées (comme nous le verrons en reprenant cet exemple dans la section 7.3), ce qui explique les bonnes performances de l'algorithme générique approché.

6.4 Algorithmes spécifiques pour certains modes de scrutin

Pour certains modes de scrutin, des algorithmes spécifiques sont implantés. Dans cette section, nous donnons une vue d'ensemble de ces algorithmes.

Pour le vote par notation, le jugement majoritaire, le vote par assentiment, la méthode de Bucklin, l'uninominal, l'uninominal à deux tours et veto, MC est calculée exactement en temps polynomial. Pour la méthode de Bucklin, nous utilisons un algorithme de Xia *et al.* (2009). Pour les autres, l'algorithme de manipulation est évident.

Pour la méthode de Schulze, nous utilisons un algorithme de Gaspers *et al.* (2013), qui fonctionne en temps polynomial. L'article original prouve que cet algorithme est exact pour le *problème à plusieurs vainqueurs*, c'est-à-dire si les manipulateurs ont pour but qu'un certain candidat c fasse partie des *vainqueurs potentiels* (voir la définition de ce terme dans la définition de la méthode de Schulze, en section 1.6.7). Il en est de même pour le *problème à un seul vainqueur*, c'est-à-dire si les manipulateurs ont pour but qu'un certain candidat c soit l'unique vainqueur potentiel (c'est-à-dire vainqueur pour toute règle de départage). Dans SWAMP, il a une marge d'erreur d'un manipulateur à cause de la règle de départage. Par conséquent, il peut arriver (relativement rarement) que le simulateur renvoie `nan` quand on lui demande de décider la manipulabilité.

Pour la méthode de Borda, décider MC est $\mathcal{NP}$-complet, même quand il y a seulement 2 manipulateurs (Betzler *et al.*, 2011; Davies *et al.*, 2011). Nous utilisons un algorithme d'approximation de Zuckerman *et al.* (2009), qui possède une marge d'erreur d'un manipulateur.

Pour Maximin, décider MC est $\mathcal{NP}$-complet, même quand il y a seulement 2 manipulateurs (Xia *et al.*, 2009). Nous utilisons un algorithme d'approximation de Zuckerman *et al.* (2011), qui a un facteur d'erreur multiplicatif de $\frac{5}{3}$ sur le nombre de manipulateurs nécessaire.

Pour la méthode de Coombs, il y a deux options possibles pour MC : `'fast'` et `'exact'`. En mode exact, notre algorithme est similaire à celui utilisé par Walsh (2010b) pour IRV. Sa complexité en temps est d'ordre $C!$. En mode rapide, nous utilisons une heuristique originale que nous ne décrirons pas en détails mais dont le principe est similaire à celui utilisé pour VTI, que nous décrirons dans la section 6.5. La question est cependant plus simple que pour VTI car on peut remarquer que MC est équivalente à MU pour la méthode de Coombs : en effet, si les manipulateurs ont la capacité d'éliminer successivement les candidats $d_1, \ldots, d_{C-1}$, alors ils peuvent également le faire en utilisant l'ordre de préférence $(c \succ d_{C-1} \succ \ldots \succ d_1)$.

6.5 Algorithmes spécifiques pour VTI et ses variantes

VTI fut l'un des premiers systèmes de vote pour lesquels décider MI (et *a fortiori* MC) fut prouvé $\mathcal{NP}$-complet (Bartholdi et Orlin, 1991). De plus, il semble particulièrement résistant à la manipulation (Chamberlin *et al.*, 1984; Bartholdi et Orlin, 1991; Lepelley et Mbih, 1994; Lepelley et Valognes, 2003; Green-Armytage, 2011, 2014). À notre connaissance, l'algorithme de manipulation le plus rapide pour la manipulation individuelle est en $O(2^C VC^k))$ (Walsh, 2010a) et le plus rapide pour la manipulation par coalition est en $O(C!(V + Ck))$ (Coleman et Teague, 2007).

Pour ces raisons, nous avons accordé une importance particulière aux méthodes pour VTI ainsi que pour deux de ses variantes, SE et CVTI, dont nous verrons dans les chapitres suivants qu'elles sont également assez résistantes à la manipulation. Pour CVTI, on sait, par les théorèmes de condorcification 2.9 et 2.20 qu'il est strictement moins manipulable que VTI et il serait intéressant de mesurer quantitativement l'écart de manipulabilité. Pour SE, nous savons que c'est un système strictement plus manipulable que VTI (section 1.4.2), mais il présente l'avantage d'être souvent plus simple que VTI et de permettre des algorithmes plus rapides.

Par exemple, quand on veut prouver que VTI n'est pas manipulable, il est suffisant de montrer que SE ne l'est pas ; et nous verrons que c'est généralement plus rapide.

6.5.1 Algorithmes de MC pour SE

Algorithme 1 MC exact en faveur de c pour SE

$LiceIni = \mathcal{C}$
$LicesPossiblesIniTour = \{LiceIni\}$
pour $t = 1 \rightarrow C - 1$ **faire**
  $LicesPossiblesFinTour = \varnothing$
  pour tout $LiceIni \in LicesPossiblesIniTour$ **faire**
    pour tout $d \in LiceIni \setminus \{c\}$ **faire**
      $test$ = sous l'hypothèse que le tour d'élimination t se déroule entre les candidats de $LiceIni$, les manipulateurs peuvent-ils choisir leurs bulletins de façon à éliminer d ?
      si $test$ **alors**
        $LiceFin = LiceIni \setminus \{d\}$
        $LicesPossiblesFinTour = LicesPossiblesFinTour \cup \{LiceFin\}$
      fin si
    fin pour
  fin pour
  $LicesPossiblesIniTour = LicesPossiblesFinTour$
fin pour
renvoyer $(LicesPossiblesFinTour \neq \{\})$

En mode exact, SWAMP utilise l'algorithme 1, qui est une adaptation pour SE d'un algorithme initialement conçu pour VTI par Coleman et Teague (2007). À chaque tour, on détermine l'ensemble des situations atteignables par les manipulateurs à la fin du tour, c'est-à-dire les candidats restant en lice (en ignorant les situations où c est éliminé). Il est clair que cet algorithme termine et permet de décider la manipulabilité : à la fin du tour $C-1$, l'ensemble $LicesPossiblesFinTour$ est non vide si et seulement si une manipulation est possible en faveur de c. Dans ce cas, l'ensemble $LicesPossiblesFinTour$ contient nécessairement uniquement le singleton $\{c\}$.

La principale différence avec l'algorithme de Coleman et Teague (2007) est la suivante : dans VTI, le bulletin utilisé par un manipulateur au tour t en faveur d'un candidat d est bloqué jusqu'à ce que le candidat d soit éliminé. Pour cette raison, éliminer un certain candidat d puis e ou l'inverse n'est pas forcément équivalent pour les manipulateurs : même s'il reste les mêmes candidats au début du tour suivant, les manipulateurs bloqués ne sont pas forcément les mêmes. Pour cette raison, on peut être amené à explorer tous les ordres d'élimination possibles, ce qui cause une complexité en $C!$.

En revanche, en SE, la situation des manipulateurs au début d'un tour est entièrement caractérisée par les candidats restant en lice. Ainsi, l'algorithme 1 explore, au pire, tous les sous-ensembles de $\mathcal{C}$ et possède donc une complexité en 2^C « seulement ».

En mode rapide (obtenu avec `CM_option = fast`), SWAMP utilise l'algorithme 2, qui est une contribution originale. À chaque début de tour, les manipulateurs examinent tous les candidats (différents de c) qu'ils sont capables d'éliminer. Parmi eux, ils choisissent le meilleur candidat d_{meilleur}, au sens d'une

Algorithme 2 MC rapide en faveur de c pour SE

1: $Lice = \mathcal{C}$
2: **pour** $t = 1 \to C - 1$ **faire**
3: $Eliminables = \{d \in Lice \backslash \{c\}$ t.q., sous l'hypothèse que le tour l'élimination t se déroule entre les candidats de $Lice$, les manipulateurs peuvent choisir leurs bulletins de façon à éliminer $d\}$.
4: **si** $Eliminables = \varnothing$ **alors**
5: **renvoyer Maybe**
6: **sinon si** $t = C - 1$ **alors**
7: **renvoyer True**
8: **fin si**
9: $d_{\text{meilleur}} = \arg\max\{\text{Situation}(d), d \in Eliminables\}$
10: $Lice = Lice \setminus \{d_{\text{meilleur}}\}$
11: **fin pour**

fonction bien choisie notée Situation, qui est chargée d'évaluer la situation après l'élimination chaque adversaire d possible. Il s'agit donc d'un algorithme glouton.

Un choix important pour cet algorithme est donc celui de la fonction Situation. Dans SWAMP, on utilise la fonction suivante :

$$\text{Situation}(d) = \text{score}(c \mid \not{d}) - \max_{e \neq c}(\text{score}(e \mid \not{d})),$$

où $\text{score}(e \mid \not{d})$ désigne le score attribué à un candidat e, par les électeurs sincères uniquement, dans le cas où d est éliminé, c'est-à-dire dans un tour se déroulant entre les membres de $Lice \setminus \{d\}$.

Il est clair que cet algorithme termine en temps polynomial et que s'il renvoie **True**, alors le profil est manipulable en faveur du candidat c. En revanche, s'il renvoie **Maybe**, il est impossible de conclure.

La figure 6.3 montre les performances de SWAMP en mode rapide. On considère une culture impartiale, $V = 33$ électeurs et un nombre variable de candidats. Chaque point est tiré à partir de 10 000 populations aléatoires. On trace le taux de MC trouvé par l'algorithme exact à titre de référence.

La borne supérieure du taux de MC fournie en mode rapide est donnée par les cas où les tests préliminaires ont permis de prouver que le profil est non manipulable, par exemple lorsque le vainqueur sincère est un favori majoritaire. On constate que, dans ce type de culture, pour $C \geq 4$, les tests préliminaires ne sont capables de prouver la non-manipulabilité que dans une proportion des cas extrêmement faible[9].

En revanche, la borne inférieure du taux de MC fournie en mode rapide est très proche du taux exact. À titre de référence, on a aussi indiqué la manipulation triviale : en effet, un bon indice de performance est la proportion de manipulations qui ne sont pas trouvées par la MT mais le sont par l'algorithme rapide. Pour chaque valeur de C de cette figure, l'algorithme trouve toujours au moins 96 % des manipulations non trouvées par la MT (99 % des manipulations).

En résumé, le mode rapide permet de prouver la manipulabilité dans une proportion élevée des cas où c'est possible, grâce à l'algorithme rapide 2, mais permet rarement de prouver la non-manipulabilité.

9. Sur les expériences issues d'élections réelles du chapitre 9, nous verrons, cependant, que pour de nombreux modes de scrutin, les tests préliminaires permettent de conclure dans une proportion des cas non négligeable, ce qui motive leur implantation.

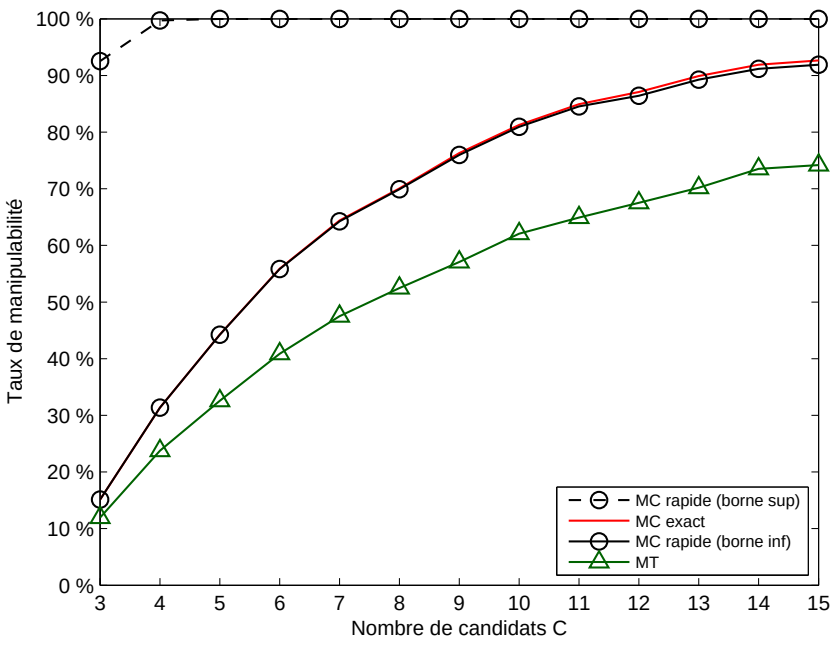


FIGURE 6.3 – SE : taux de manipulabilité trouvé par les différents algorithmes. Culture impartiale, $V = 33$.

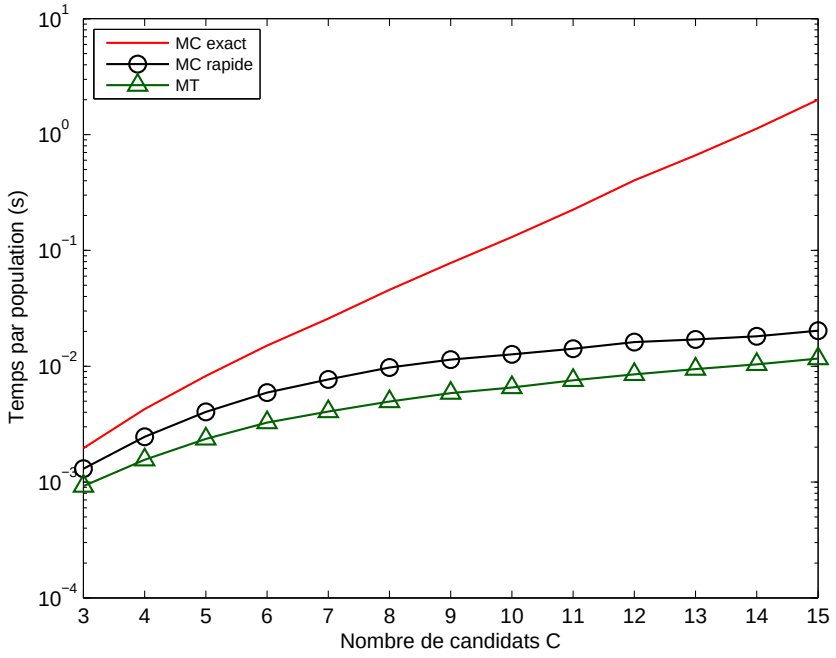


FIGURE 6.4 – SE : temps de calcul des différents algorithmes. Culture impartiale, $V = 33$.

La figure 6.4 permet de visualiser le gain de temps obtenu grâce à l'algorithme rapide. En pratique, la complexité moyenne observée sur ces courbes est $1{,}6 \times C$ ms/population pour le mode rapide, et $0{,}5 \times 1{,}75^C$ ms/population pour le mode exact de SWAMP [10]. On notera que, pour l'algorithme exact, la complexité n'est pas en 2^C, ce qui s'explique par plusieurs raisons : premièrement, le comportement théorique en 2^C exprime une complexité en pire cas, alors qu'on mesure la complexité moyenne. Deuxièmement, il n'y a pas le même taux de manipulabilité en fonction de C ; or, les configurations manipulables sont, d'habitude, plus rapides à traiter, notamment parce que l'algorithme rapide peut souvent conclure avant même qu'on fasse appel à l'algorithme exact. Troisièmement, SWAMP possède quelques optimisations supplémentaires, en particulier les tests préliminaires qui lui permettent de conclure rapidement dans un certain nombre de cas.

Ainsi, notre algorithme approché 2 pour la MC de SE est efficace pour prouver la manipulabilité et bien plus rapide que l'algorithme exact. Sa principale limite est qu'il est incapable de prouver la non-manipulabilité le cas échéant. En pratique, dans le mode rapide de SWAMP, l'algorithme 2 est complété par des tests préliminaires qui peuvent prouver la non-manipulabilité mais dont le taux de réussite peut être très faible dans certaines cultures. Pour des travaux futurs, il serait particulièrement intéressant d'avoir un algorithme en temps polynomial qui soit capable de prouver la non-manipulabilité avec un bon taux de succès, en complément de notre algorithme qui est relativement efficace pour traiter les cas de manipulabilité.

6.5.2 Algorithmes de MC pour VTI

Pour VTI, comme nous l'évoquions, le problème est plus complexe que pour SE : comme les manipulateurs sont immobilisés par leurs votes passés en faveur de candidats non encore éliminés, il est nécessaire de prendre en compte l'historique des tours précédents.

En mode exact, SWAMP utilise l'algorithme de Coleman et Teague (2007), dont la complexité est en $C!$.

En mode rapide, on utilise une approche gloutonne similaire à l'algorithme 2 : à chaque tour, parmi les candidats qu'on peut éliminer, on choisit celui qui maximise une fonction évaluant la situation obtenue. En l'occurrence, nous avons choisie la même fonction que pour SE.

Le calcul de MC pour VTI est donc basé sur deux algorithmes, rapide et exact. Mais en pratique, dans SWAMP, la MC de VTI possède trois options.

`fast` : Après les tests préliminaires, seul l'algorithme rapide est utilisé.

`slow` : En outre, on détermine si SE est manipulable en utilisant l'algorithme exact de SE. Si c'est faux, on conclut que VTI n'est pas manipulable. Si c'est vrai, on profite de l'occasion pour tester le chemin d'élimination permettant la manipulation en SE, au cas où il permettrait de manipuler également VTI.

`exact` : Si les procédés ci-dessus ne permettent pas de conclure, alors on utilise l'algorithme exact de VTI.

L'option `slow` offre donc une possibilité intermédiaire, plus rapide que l'option exacte, mais plus performante que l'option rapide, en particulier pour décider les cas de non-manipulabilité.

10. La machine utilisée est un Dell Precision M6600, Intel Core I7-2820QM à 2,30 GHz, 8 Mo de cache, 16 Go de RAM à 1,33 GHz en DDR3.

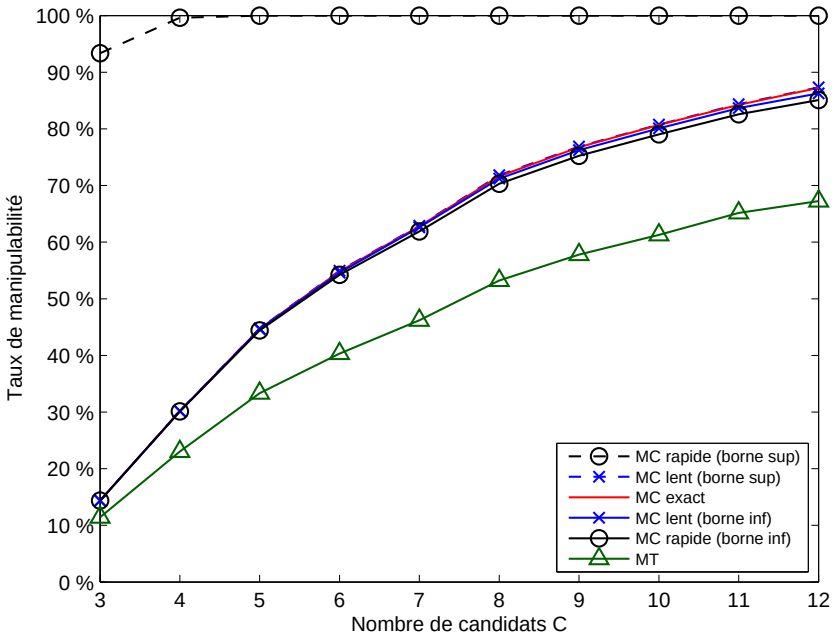


FIGURE 6.5 – VTI : taux de manipulabilité trouvé par les différents algorithmes. Culture impartiale, $V = 33$.

La figure 6.5 présente les performances des différentes options pour VTI, dans les mêmes conditions que la figure 6.3 pour SE. Pour chaque valeur de C de cette figure, le mode rapide trouve toujours au moins 88 % des manipulations non trouvées par la MT (97 % des manipulations). En revanche, comme pour SE, la borne supérieure trouvée en mode rapide est généralement proche de 1 : le mode rapide n'est pas très bien adapté pour prouver la non-manipulabilité.

En mode lent, on améliore la borne inférieure trouvée : pour chaque valeur de C de cette figure, le mode lent trouve toujours au moins 95 % des manipulations non trouvées par la MT (99 % des manipulations). Mais surtout, la borne supérieure fournie par le mode lent est très proche de la courbe de MC exacte, ce qui signifie qu'on est capable de prouver de nombreux cas de non-manipulabilité. En l'occurrence, pour chaque valeur de C, le mode lent trouve toujours au moins 99 % des cas de non-manipulabilité. Ainsi, l'écart entre les bornes inférieure et supérieure trouvées en mode lent est au maximum d'environ 1 %, ce qui signifie que la valeur conventionnelle `nan` n'a été renvoyée que dans environ 1 % des cas en mode lent.

Enfin, la figure 6.6 présente les temps de calculs pour les différents algorithmes dans le cas de VTI. En pratique, la complexité moyenne constatée sur cette figure est $2{,}0 \times C$ ms/population pour le mode rapide, $0{,}5 \times 1{,}75^C$ ms/population pour le mode lent et $4{,}1 \times (0{,}23C)^{0{,}57C}$ ms/population pour le mode exact [11]. Pour le mode lent, il s'agit du même comportement asymptotique que pour le mode exact de SE, ce qui est naturel vu notre façon de procéder : en effet, les calculs supplémentaires à effectuer pour l'adaptation à VTI sont négligeables devant ceux

11. Le comportement théorique en pire cas est de l'ordre de $C!$, c'est-à-dire de l'ordre de C^C en utilisant la formule de Stirling. En ajustant les paramètres, on obtient le comportement approximatif énoncé.

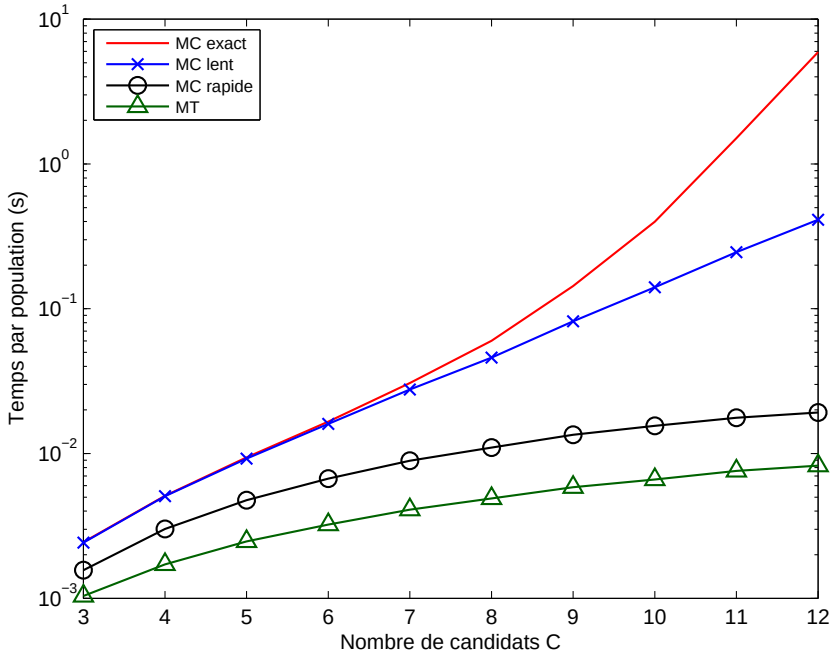


FIGURE 6.6 – VTI : temps de calcul des différents algorithmes. Culture impartiale, $V = 33$.

nécessaires pour l'algorithme exact de SE (qui demandent un temps exponentiel en C).

6.5.3 Algorithmes de MC pour CVTI

Pour CVTI, nous utilisons une heuristique que nous allons décrire brièvement.

Tout d'abord, pour chaque candidat $c \neq w$, on détermine si VTI est manipulable en utilisant l'une des options disponibles (rapide, lente ou exacte). Ensuite, en utilisant le chemin d'élimination obtenu, on essaie de mettre au point les bulletins des manipulateurs tout en empêchant qu'il y ait un vainqueur de Condorcet distinct de c. Si le profil initial est Condorcet, on sait d'ailleurs qu'il ne doit pas y avoir de vainqueur de Condorcet du tout (lemme 2.5).

Pour cela, on place, tout d'abord, le candidat c aussi haut que possible dans les bulletins, puis les candidats dont on sait déjà qu'ils possèdent une défaite dans la matrice des duels. Ce faisant, il est possible que d'autres candidats acquièrent des défaites et on peut itérer le processus. Pour plus de détails, on pourra se reporter au code de la classe correspondante dans SVAMP.

Si le processus réussit, c'est-à-dire s'il n'y a pas de vainqueur de Condorcet distinct de c à la fin de l'algorithme, alors MC est prouvée. Dans le cas contraire, on reste dans l'incertitude : en effet, on n'a testé qu'un chemin d'élimination parmi ceux qui permettent de manipuler VTI, et le processus employé pour empêcher l'apparition d'un vainqueur de Condorcet a fait l'objet de choix arbitraires. Ainsi, cet algorithme ne fournit qu'une borne inférieure de la manipulabilité.

Cependant, on dispose également d'une borne supérieure : si VTI n'est pas manipulable, alors CVTI ne l'est pas non plus, par le théorème faible de condorcification 2.9. Dans les chapitres suivants, nous verrons qu'en pratique, l'écart entre

la borne inférieure et la borne supérieure est en général très faible, ce qui assure, d'une part, que l'heuristique utilisée pour CVTI est relativement efficace, et montrera, d'autre part, que l'écart de manipulabilité entre VTI et CVTI, quoique non nul par le théorème fort de condorcification 2.20, est en fait assez faible.

Chapitre 7

Étude de cultures sphéroïdales

Dans ce chapitre et les deux suivants, nous étudions deux problématiques qui étaient au cœur de notre motivation pour développer SWAMP.

1. Alors que la manipulabilité est une nécessité théorique par le théorème de Gibbard (1973) et Satterthwaite (1975), est-elle un phénomène courant en pratique ?
2. Quelles sont les performances comparées des différents systèmes de vote à cet égard ?

Il s'agit des motivations principales pour ce chapitre et les deux suivants. Une motivation secondaire est d'illustrer les possibilités offertes par SWAMP.

En général, nous concentrons notre attention sur la manipulabilité par coalition (MC), qui est l'objet d'étude central de ce mémoire et que appelons aussi *manipulabilité* sans plus de précisions. Cependant, comme SWAMP ne dispose pas d'algorithmes exacts spécifiques pour la MC de chaque système de vote, nous utiliserons également la manipulation triviale (MT) comme point de comparaison entre les différents systèmes. Comme nous l'avons discuté dans la section 6.2.3, celle-ci présente l'avantage d'être facile à calculer (en temps polynomial, à temps de détermination du vainqueur près). En outre, quand la manipulation est possible, elle donne une mesure indicative des cas où celle-ci est relativement peu coûteuse à implanter en termes de communication.

Étant donné la variété des cultures étudiées, la lecture de ces chapitres 7 et 8 pourrait s'avérer fastidieuse pour le lecteur impatient, mais c'est un défaut que nous n'avons pas réussi à éviter. Pour une lecture plus rapide, nous suggérons de lire la section 7.1 qui définit les cultures utilisées dans ce chapitre, la section 7.2 qui présente notre scénario de référence, la section 7.3 qui permet de se familiariser avec les conventions utilisées pour tracer les courbes, la section 7.9 qui conclut ce chapitre, l'introduction du chapitre 8 et les sections 8.6 et 8.7 qui concluent celui-ci, avant de passer au chapitre 9 qui analyse des données issues d'élections réelles.

7.1 Présentation des cultures sphéroïdales

En théorie du vote, il est classique de considérer la culture impartiale (définition 1.14), qui est définie pour des préférences qui sont des ordres stricts totaux. Comme nous souhaitons étudier à la fois des systèmes ordinaux et cardinaux, la première question qui se pose est la façon dont on peut étendre ce modèle à des électeurs caractérisés par leurs utilités pour les candidats. Dans cette section, nous

décrivons brièvement la démarche utilisée, qui est détaillée et présentée de façon plus formelle dans l'annexe B et dans l'article de Durand *et al.* (2015).

Sans adopter toutes les hypothèses du modèle des utilités espérées de Von Neumann et Morgenstern (1944), nous allons toutefois nous en inspirer. Nous ne supposons pas nécessairement que les préférences des électeurs soient caractérisées par des utilités espérées, puisqu'il n'y a pas d'élément aléatoire une fois la population fixée : les systèmes de vote étudiés calculent le vainqueur de façon déterministe et les manipulateurs n'utilisent que des stratégies pures, et non des stratégies mixtes.

Cependant, nous conservons une propriété importante de ce modèle : dans un premier temps, le vecteur d'utilité $\mathbf{u}_v \in \mathbb{R}^C$ d'un électeur pour les différents candidats est défini à une constante additive et à une constante multiplicative strictement positive près, c'est-à-dire que le vecteur $\mathbf{u}_v$ représente les mêmes préférences qu'un vecteur $\alpha\mathbf{u}_v + \beta\mathbf{1}$, où α est un réel positif, β un réel quelconque et où $\mathbf{1}$ désigne le vecteur dont toutes les coordonnées sont égales à 1. Ce choix est motivé par le fait que chaque électeur ne possède aucune méthode canonique, ni pour fixer son point 0 d'utilité, ni pour choisir « l'unité de mesure » employée afin d'évaluer ses valeurs d'utilité : il est simplement capable de mesurer leurs intensités respectives.

Dans un second temps, afin d'avoir un modèle complet pour le vote par assentiment, nous ajoutons l'hypothèse que l'utilité 0 est une *limite d'approbation* (voir section B.7) : un électeur approuve les candidats qui ont une utilité positive et désapprouve les candidats qui ont une utilité strictement négative. Avec cette hypothèse supplémentaire, le vecteur d'utilité d'un électeur n'est plus défini qu'à une constante multiplicative α strictement positive près. L'espace des utilités est donc celui des demi-droites de $\mathbb{R}^C$ et tous les vecteurs d'une même demi-droite représentent les mêmes préférences en pratique, de façon indistinguable par l'expérience, même pour l'électeur concerné.

Intuitivement, une façon naturelle de représenter cet espace consiste à normaliser chaque vecteur d'utilité pour qu'il ait une norme euclidienne égale à 1. L'espace des utilités, c'est-à-dire l'ensemble des demi-droites, est alors représenté par la sphère unité de $\mathbb{R}^C$. Dans l'annexe B, nous prouvons, par des arguments de géométrie projective et de géométrie différentielle, que cette représentation est essentiellement la seule qui possède de bonnes propriétés [1].

On dispose alors d'une façon canonique de généraliser la culture impartiale : elle consiste à tirer un vecteur sur la surface de la sphère unité de $\mathbb{R}^C$ avec une loi de probabilité uniforme (au sens de la mesure euclidienne usuelle) : c'est ce que nous appelons la *culture sphérique*. Chaque électeur possède presque sûrement un ordre strict total de préférence, les électeurs sont indépendants et la culture est neutre : par conséquent, l'image ordinale de cette culture (définition 5.13) est bien la culture impartiale. Ainsi, pour les systèmes de vote ordinaux, cette culture est tout simplement équivalente à la culture impartiale ; en outre, elle permet d'étudier les systèmes ordinaux et cardinaux dans un cadre commun naturel.

Par la suite, il est intéressant d'enrichir ce modèle en introduisant une corrélation entre les électeurs. Pour cela, nous utilisons le modèle de Von Mises-Fisher ou

1. Plus exactement, dans le modèle avec limite d'approbation, les modèles convenables sont des ellipsoïdes qui sont l'image de la sphère unité par une dilatation le long de l'axe $\mathbf{1}$, qui constituent donc une famille à un paramètre réel. Le choix de ce paramètre n'a d'influence que pour le vote par assentiment : aux valeurs extrêmes du facteur de dilatation, on privilégie soit les vecteurs d'utilité qui ont beaucoup de composantes de même signe, soit ceux dont la somme des composantes est nulle et qui ont donc toujours des composantes des deux signes. Pour simplifier, nous considérons toujours le modèle sphérique qui est intermédiaire entre ces deux cas extrêmes.

VMF (Downs, 1966). Ce type de culture permet de modéliser des populations qui tendent à avoir des préférences similaires mais avec une certaine dispersion. Tout d'abord, on fixe un vecteur unitaire $\mathbf{n}$, appelé le *pôle* de la distribution. Ensuite, indépendamment pour chaque électeur, on tire un vecteur unité $\mathbf{u}$ dans $\mathbb{R}^C$ selon une distribution de VMF :

$$p(\mathbf{u}) = X_\kappa e^{\kappa \langle \mathbf{u} \,|\, \mathbf{n} \rangle},$$

où $\mathbf{n}$ est le pôle de la distribution, κ la concentration, X_κ une constante de normalisation et où $\langle \mathbf{u} \,|\, \mathbf{n} \rangle$ désigne le produit scalaire canonique.

Qualitativement, le modèle de VMF est similaire au modèle de Mallows (1957), qui est utilisé pour des préférences ordinales [2]. Outre le fait que le premier est mieux adapté aux préférences cardinales que le second, il possède d'autres avantages qui sont discutés dans la section B.5. En particulier, le modèle de VMF est caractérisé par la propriété suivante : parmi les distributions sur la sphère partageant une direction moyenne et une concentration données [3], la distribution de VMF maximise l'entropie tout comme, dans l'espace euclidien, la loi gaussienne maximise l'entropie parmi les lois possédant une moyenne et un écart-type donnés (Mardia, 1975). Par conséquent, en l'absence d'information supplémentaire, c'est la distribution « naturelle » qui devrait être utilisée.

Il s'agit bien d'une généralisation de la culture sphérique : en effet, pour $\kappa = 0$, le modèle de VMF est équivalent à la culture sphérique. Pour $\kappa = +\infty$, on obtient l'autre cas dégénéré de culture de VMF : un pic de Dirac sur le pôle $\mathbf{n}$.

Nous regroupons le modèle de VMF et la culture sphérique, qui en est un cas particulier, sous le nom de *sphéroïdales*, en raison du modèle sphérique sous-jacent employé pour l'espace des utilités.

7.2 Scénario de référence

Dans cette section, nous présentons notre méthodologie et notre scénario de référence. Il s'agit de la culture sphérique avec $V = 33$ électeurs et $C = 5$ candidats. Pour ce cas de base, on choisit un nombre d'électeurs impair pour éviter certaines questions d'égalité, par exemple les subtilités entre les différentes notions Condorcet : en particulier, les notions de vainqueur de Condorcet et de candidat Condorcet-admissible sont alors équivalentes. Pour choisir la valeur de V et de C, nous avons exploré les valeurs de $V = 2^k + 1$ pour $k \in [\![1, 10]\!]$ (donc pour V entre 3 et 1025) et les valeurs de $C \in [\![3, 15]\!]$. Dans les sections 7.3 et 7.4, nous verrons que, pour la plupart des systèmes de vote, en culture sphérique, le taux de manipulabilité semble croître et tendre vers 1 quand l'un ou l'autre de ces paramètres augmente. Notre choix de $V = 33$ et $C = 5$ permet d'avoir des taux de manipulabilité raisonnablement éloignés de 1 et donc de faciliter la comparaison entre les différents systèmes de vote étudiés.

2. Pour modéliser des populations qui possèdent une certaine homogénéité, on peut aussi utiliser les modèles d'urnes de Pólya-Eggenberger, comme le font par exemple Lepelley et Valognes (2003). L'inconvénient de ces modèles est qu'ils ne peuvent générer que des préférences ordinales et qu'il n'ont pas d'extension naturelle à l'espace des utilités.

3. Plus précisément, on considère les distributions possédant le même *vecteur résultant*. En statistiques sphériques, le vecteur résultant est défini comme la moyenne des vecteurs (unitaires) de la distribution. Par exemple, considérons la sphère unité de dimension 1, c'est-à-dire le cercle unité de $\mathbb{C}$. Si une distribution tire de façon sûre le vecteur d'angle polaire $\frac{\pi}{2}$, alors son vecteur résultant est $1 \cdot e^{i\frac{\pi}{2}}$. Mais si on tire une fois sur deux les vecteurs unités d'angles polaires $\frac{\pi}{6}$ et $\frac{5\pi}{6}$, alors le vecteur résultant est $\frac{1}{2} \cdot e^{i\frac{\pi}{2}}$: la direction moyenne est la même, mais la norme est moins élevée. De manière générale, la direction du vecteur résultant doit être vue comme la direction moyenne de la distribution, et sa norme comme une mesure de concentration (jouant le même rôle, qualitativement, que l'inverse d'un écart-type).

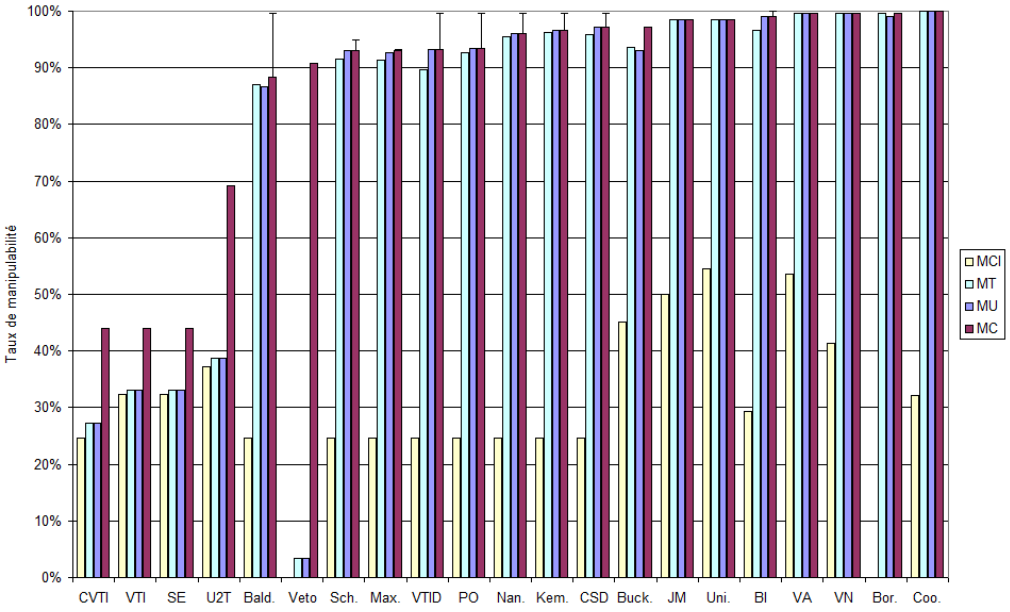


FIGURE 7.1 – Manipulation par coalition (MC) et variantes. Culture sphérique, $V = 33$, $C = 5$.

Dans la figure 7.1, on illustre les performances des divers systèmes de vote dans ce scénario de référence. On peut visualiser MC et ses variantes MT, MU et MCI. Pour cette figure, on a tiré 10 000 populations aléatoires avec la culture considérée.

La proportion de configurations MC (resp. MT, MU, MCI) pour un système de vote sur ces 10 000 tirages aléatoires fournit un estimateur du taux de MC (resp. MT, MU, MCI) pour la culture considérée. Ceci induit une incertitude que nous appellerons *statistique* de l'ordre de $\frac{1}{\sqrt{10\,000}} = 1\%$ sur le taux de MC (resp. MT, MU, MCI). Par convention, cette part de l'incertitude n'est représentée ni sur la figure 7.1 ni sur les suivantes.

Par ailleurs, pour certains systèmes de vote, l'algorithme utilisé pour déterminer MC n'est pas exact. Les histogrammes de la figure 7.1 représentent une borne inférieure de la proportion de configurations manipulables, qui est donnée par les cas où SWAMP a été capable de prouver MC. La barre d'incertitude (par exemple pour la méthode de Baldwin) indique une borne supérieure de cette proportion, qui est fournie par les cas où SWAMP a été capable de prouver que MC est impossible. Nous appellerons la différence entre ces deux bornes *l'incertitude d'origine algorithmique*.

Profitons-en pour examiner les performances des algorithmes implantés dans SWAMP pour MC.

Pour Maximin et les méthodes de Schulze et Borda, nous utilisons un algorithme approché mais il semble que la plupart du temps, SWAMP est en pratique capable de décider si ces systèmes sont manipulables ou pas : sur la figure 7.1, l'incertitude d'origine algorithmique est de 1,8 % pour la méthode de Schulze et elle est nulle pour Maximin et la méthode de Borda.

Pour Baldwin, VTID, PO, la méthode de Nanson, CSD, la méthode de Kemeny et BI[4], SWAMP utilise seulement ses méthodes de manipulation génériques, essentiellement constituées de tests préliminaires et de MT. Dans ce cas, on a essentiellement une borne inférieure de manipulabilité, qui suffira cependant à prouver que ces systèmes sont plus manipulables que d'autres dans certains cas. En l'occurrence, dans le scénario de référence de la figure 7.1, on a l'assurance que ces systèmes sont plus manipulables que CVTI et VTI par exemple.

Pour SE, nous utiliserons toujours l'algorithme exact. Pour VTI, nous utilisons l'algorithme exact également, sauf dans la section 7.3 où nous le remplaçons par l'option lente (section 6.5.2) quand le nombre de candidats est supérieur ou égal à 9 afin de limiter le temps de calcul. Pour CVTI, nous employons l'heuristique présentée en section 6.5.3 : bien qu'il ne s'agisse pas d'un algorithme exact, on constate dans le cas particulier de la figure 7.1 que l'incertitude d'origine algorithmique est nulle. Ceci signifie également qu'en pratique, CVTI a un taux de manipulabilité très proche de celui de VTI, même si nous savons qu'il est, en général, strictement inférieur par le théorème fort de condorcification 2.20.

Sur la figure 7.1, on peut cependant noter une différence non négligeable : CVTI est sensiblement moins vulnérable à MCI, MT et MU que VTI, même si son taux de MC est très proche. Ainsi, la manipulation semble plus difficile en CVTI qu'en VTI.

En revanche, VTI et SE présentent des performances très similaires, même si l'exemple de la table 1.1 (section 1.4.2) montre que SE est strictement plus manipulable que VTI (l'un et l'autre étant respectivement équivalents à U2T et U2TI dans cet exemple à 3 candidats).

Veto est le seul mode de scrutin pour lequel le taux de MT (3%) est très différent du taux de MC (91 %). Ceci s'explique par le fait que la stratégie triviale est très mauvaise en Veto : certes, elle est optimale pour éviter une victoire du vainqueur sincère w, mais les manipulateurs abandonnent tout contrôle sur le candidat vainqueur. Au contraire, la stratégie optimale consiste à se coordonner pour équilibrer le total de voix négatives sur tous les candidats, à part celui qu'on souhaite faire gagner.

Enfin, on constate que VTID est nettement plus manipulable que CVTI, VTI et SE. Pourtant, en apparence, il s'agit d'un système de vote dont le principe est similaire. On peut essayer d'en comprendre la raison avec un exemple à 3 candidats (éventuellement après que d'autres éliminations aient eu lieu). Si le vainqueur sincère w n'est pas vainqueur de Condorcet, alors le profil est, de toute façon, manipulable dans tout système vérifiant **CMInf**, donc l'exemple intéressant est celui où w est vainqueur de Condorcet. Notons c le candidat pour lequel on souhaite manipuler et d le troisième candidat.

Dans VTID, c ne doit jamais affronter w en duel, sinon il en sort perdant. Voici donc le seul scénario possible : w affronte d et perd, puis c affronte d et gagne. Ainsi, il faut que les candidats sélectionnés pour le premier duel électoral soient w et d : pour cela, la meilleure stratégie est de défendre c en le plaçant en

4. Dans ce chapitre et les suivants, nous n'incluons pas la méthode de Kim-Roush et VTIM, qui ont été implantés dans SWAMP après la réalisation de ces simulations et qui utilisent, eux aussi, les algorithmes génériques de manipulation. Pour des raisons de temps de calcul et pour éviter d'ajouter des coquilles de dernière minute, nous avons jugé plus sage de nous en tenir là, et nous prions le lecteur de nous en excuser. Ceci dit, d'après les premiers tests effectués, les performances de la méthode de Kim-Roush en matière de manipulabilité semblent assez médiocres. En ce qui concerne VTIM, les résultats sont au contraire prometteurs et nous y reviendrons en conclusion de ce mémoire. Enfin, nous excluons aussi la méthode de Copeland, pour les problèmes d'égalités exposés dans la section 1.6.

tête de bulletin. Par ailleurs, il faut que w perde contre d : pour cela, la meilleure stratégie est d'attaquer w en le plaçant en bas de bulletin. Ainsi, les différents impératifs, à savoir la défense de c et l'attaque du candidat dangereux w, sont parfaitement compatibles.

À présent, examinons la situation en SE ou en VTI. À chaque tour d'élimination, certains manipulateurs doivent défendre c pour l'empêcher d'être éliminé. De plus, tant que w n'est pas éliminé, aucun électeur sincère ne vote pour c : les manipulateurs ne peuvent donc compter que sur eux-mêmes pour cela. Par ailleurs, les manipulateurs doivent éviter un dernier tour entre c et w, pour les mêmes raisons que précédemment. Il leur faut donc attaquer w pour l'éliminer immédiatement. Mais pour cela, ils peuvent avoir besoin de répartir leurs voix entre c et d, ce qui est *a priori* plus difficile qu'en VTID où il suffisait de défendre c.

Ceci amène une autre remarque : bien que VTI soit la RPSI-ES de l'uninominal, il présente une similitude avec Veto (et, à l'inverse, la méthode de Coombs présente des similitudes avec l'uninominal). En effet, le choix du candidat éliminé en VTI se fait en Veto : chaque électeur, en votant pour un candidat, émet un veto contre son élimination et le candidat qui a le moins de vetos contre son élimination (c'est-à-dire possédant le moins de voix) est désigné (c'est-à-dire éliminé). Or nous avons vu que la famille maximale de Veto est, aux égalités près, la famille de seuil $V\frac{C-1}{C}$ (proposition 4.28) : il est donc particulièrement difficile de choisir le candidat éliminé en VTI, ce qui peut expliquer la faible manipulabilité de ce mode de scrutin, malgré le fait qu'il vérifie **CMInf** si on considère le système dans sa globalité et non chaque tour séparément.

7.3 Culture sphérique : nombre de candidats C

Dans cette section et les suivantes, nous allons considérer des variantes du scénario de référence, en faisant varier à chaque fois un paramètre particulier. Pour commencer, nous conservons la culture sphérique du scénario de référence, c'est-à-dire le cas particulier où la concentration κ est nulle, nous laissons constant le nombre d'électeurs $V = 33$ et nous représentons en figure 7.2 le taux de MC trouvé en fonction du nombre de candidats C.

Ces premières courbes sont l'occasion de préciser nos conventions graphiques pour la suite de ce chapitre et le suivant. Quand le nom d'un mode de scrutin est accompagné d'une astérisque dans la légende, cela signifie que, pour certains points de la courbe, l'algorithme approché utilisé par SVAMP a trouvé un écart entre la borne inférieure et la borne supérieure plus grand que 1 %. Dans ce cas, étant donné le nombre de systèmes de vote que nous considérons, nous traçons uniquement le taux de MC trouvé par SVAMP, c'est-à-dire la borne inférieure. Par exemple, c'est le cas de la méthode de Baldwin dans la figure 7.2.

Quand il n'y a pas d'astérisque dans la légende, cela ne signifie pas nécessairement que l'algorithme utilisé est exact mais que, pour chaque point de la courbe, SVAMP a trouvé une différence inférieure à 1 % entre la borne inférieure et la borne supérieure. Ainsi, pour CVTI ou Maximin, bien que les algorithmes utilisés soient approchés, l'incertitude d'origine algorithmique est inférieure à 1 % pour chaque point de la figure 7.2.

Pour faciliter la comparaison des systèmes de vote entre les différentes figures, on assigne à chaque système de vote un unique style de courbe. Ces styles sont regroupés par famille : les systèmes de vote basés sur des notes utilisent des tons de rouge avec des marqueurs en losange ; les RPS, des tons de vert avec des marqueurs en cercle ; les systèmes de Condorcet « naturels », des tons de brun

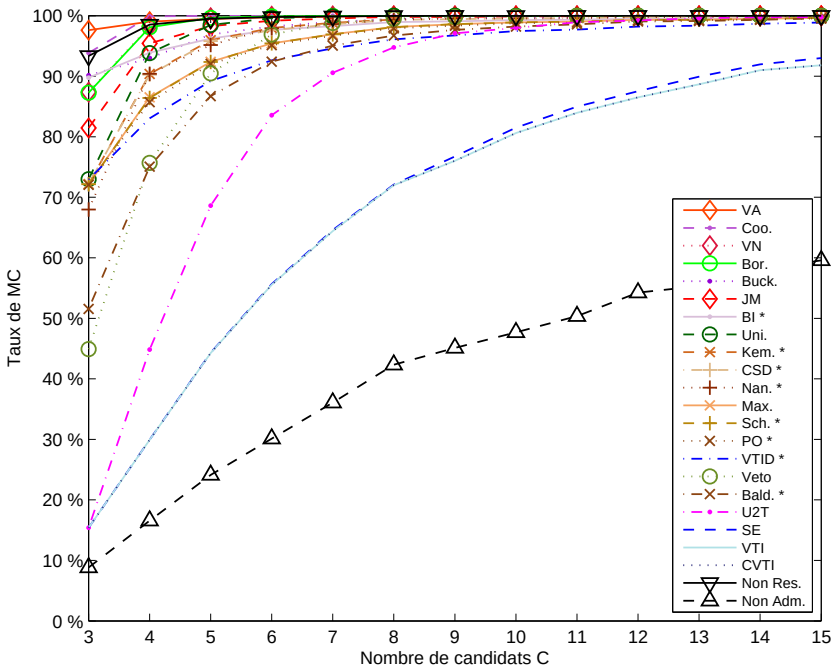


FIGURE 7.2 – Taux de MC en fonction du nombre de candidats C. Culture sphérique, $V = 33$.

avec des marqueurs en croix ; les systèmes de la famille VTI, des tons de bleu sans marqueur de point ; les autres systèmes, des tons de mauve et rose avec des marqueurs en gros points.

Afin de faciliter la correspondance avec les courbes, la légende affiche les systèmes de vote par ordre décroissant du taux de manipulabilité trouvé (en moyenne sur les points de la courbe).

À titre de référence, on représente la proportion de configurations non résistantes (c.à.d. où il n'existe pas de vainqueur de Condorcet résistant), en noir avec des triangles pointés vers le bas ; d'après le théorème 2.17, elle fournit une borne supérieure de manipulabilité pour tous les systèmes de vote qui respectent le critère de Condorcet. De même, on représente la proportion de configurations non admissibles (c.à.d. où il n'existe pas de candidat Condorcet-admissible), en noir avec des triangles pointés vers le haut, qui donne une borne inférieure de manipulabilité pour tous les modes de scrutin qui vérifient **CMInf** (lemme 2.7).

Quand le nombre d'électeurs est pair, on représentera aussi la proportion de situations non Condorcet (c.à.d. où il n'existe pas de vainqueur de Condorcet), en trait pointillé noir avec le même marqueur triangulaire. Cependant, quand le nombre de candidats est impair, comme dans la figure 7.2 et la plupart des suivantes, il est inutile de tracer cette courbe : comme nous considérons des cultures où les préférences sont presque sûrement des ordres stricts totaux, les notions de candidat Condorcet-admissible et de vainqueur de Condorcet sont alors équivalentes, ce qui n'est pas le cas pour un nombre pair d'électeurs.

La première observation issue de la figure 7.2 est la conjecture suivante.

Conjecture 7.1
En culture sphérique, pour $V \geq 3$, le taux de manipulabilité de chaque mode de scrutin considéré ici est une fonction croissante du nombre de candidats C.

Pour comprendre qualitativement ce fait, observons également que la probabilité d'avoir une configuration non admissible, c'est-à-dire non Condorcet puisque V est impair, semble croître avec le nombre de candidats.
Kelly (1974) a conjecturé que, pour $V = 3$ ou $V \geq 5$ électeurs, la probabilité qu'il existe un vainqueur de Condorcet faible (i.e. un candidat Condorcet-admissible) est une fonction décroissante de C ; ceci a été prouvé par Fishburn *et al.* (1979) pour $V = 3$ électeurs. De même, Kelly (1974) a conjecturé que, pour $V \geq 3$, la probabilité qu'il existe un vainqueur de Condorcet décroît ; ceci est prouvé pour $V \to +\infty$ (Gehrlein, 2006). À notre connaissance, les autres cas restent des conjectures. Dans le même ordre d'idées, la figure 7.2 nous amène à formuler la conjecture et la proposition suivante.

Conjecture 7.2
En culture sphérique, pour $V \geq 3$, la probabilité qu'il existe un vainqueur de Condorcet résistant est une fonction décroissante du nombre de candidats C.

Proposition 7.3
En culture sphérique, pour $V \geq 3$ et $C \to +\infty$, la probabilité qu'il existe un vainqueur de Condorcet résistant tend vers 0.

Il est facile de prouver cette proposition à partir d'un résultat antérieur : en effet, en culture impartiale et à nombre d'électeurs fixé, la probabilité qu'il existe un vainqueur de Condorcet tend vers 0 quand C tend vers $+\infty$ (May, 1971). Il en est donc de même pour le vainqueur de Condorcet résistant.

Sur la figure 7.2, on observe aussi le phénomène suivant, que nous allons pouvoir démontrer en partie.

Conjecture 7.4
En culture sphérique, pour $V \geq 3$ et $C \to +\infty$, le taux de manipulabilité de tout mode de scrutin vérifiant **CMInf** *tend vers* 1.

L'hypothèse **CMInf** concerne tous les modes de scrutin considérés ici, sauf Veto, sur lequel nous reviendrons dans un instant. Il est facile de démontrer la conjecture 7.4 pour un nombre d'électeurs impair. En effet, en culture impartiale et à nombre d'électeurs fixé, nous avons déjà rappelé que la probabilité qu'il existe un vainqueur de Condorcet tend vers 0 quand C tend vers $+\infty$ (May, 1971). Or, pour V impair, les notions de vainqueur de Condorcet et de candidat Condorcet-admissible sont équivalentes. Donc le vainqueur sincère w n'est pas Condorcet-admissible avec forte probabilité. Pour un système de vote vérifiant **CMInf**, le lemme 2.7 permet alors de conclure que la configuration est manipulable.
Pour un nombre pair d'électeurs, à notre connaissance, il n'est pas prouvé que la probabilité d'avoir un candidat Condorcet-admissible (c'est-à-dire un vainqueur de Condorcet faible, vu les hypothèses) tende vers 0 quand C tend vers $+\infty$. Si c'est vrai, alors notre preuve de la conjecture 7.4 pour tout système de vote vérifiant **CMInf** est également valide pour un nombre pair d'électeurs.

En revanche, bien que le taux de MC de Veto semble également tendre vers 1 sur la figure 7.2, on peut montrer que ce comportement limite n'est pas vrai pour Veto quand il est muni de la règle de départage utilisée dans SVAMP.

Proposition 7.5
En culture impartiale, pour V fixé et $C \to +\infty$, le taux de manipulabilité de

Veto (muni de la règle de départage par ordre d'indices sur les candidats) ne tend par vers 1.

Démonstration. Pour $C > V$, considérons la restriction de la culture impartiale pour C aux $V+1$ candidats de plus petit indice : il s'agit également de la culture impartiale. Par conséquent, avec une probabilité égale à $\frac{1}{((V+1)!)^V}$, tous les électeurs ont l'ordre de préférence suivant sur les $V+1$ premiers candidats : $(1 \succ 2 \succ \ldots \succ V+1)$. Leur relation de préférence sur les autres candidats n'a pas d'importance pour notre démonstration.

Avec de telles préférences, aucun électeur ne vote contre le candidat 1 donc celui-ci est élu. Il n'y a pas de manipulation possible pour les candidats $2, \ldots, V+1$ car aucun électeur n'est intéressé. Enfin, en raison de la règle de départage, aucun des candidats $V+2, \ldots, C$ ne peut être vainqueur : même si chacun des V premiers candidats reçoit un veto, c'est le candidat $V+1$ qui est élu. Ainsi, la manipulation est impossible.

Par conséquent, avec une probabilité au moins égale à $\frac{1}{((V+1)!)^V}$, on obtient un profil non manipulable. □

Sur la figure 7.2, on constate aussi que CVTI, VTI et SE sont significativement moins manipulables que les autres systèmes de vote et que l'uninominal à deux tours présente des performances intermédiaires entre ces trois variantes de VTI et les autres modes de scrutin, sauf pour de grandes valeurs de C. En effet, les résultats théoriques ci-dessus prouvent qu'U2T est (légèrement) plus manipulable que Veto pour un nombre de candidats suffisamment élevé.

On observe également que les courbes de CVTI, VTI et SE sont presque confondues. La différence de manipulabilité entre CVTI et VTI est indétectable sur cette figure : la différence de proportions manipulables observées est toujours nulle. Quant à celle entre VTI et SE, SWAMP trouve légèrement moins de manipulation pour VTI à partir de $C = 9$: cela correspond à la valeur à partir de laquelle on a utilisé l'algorithme approché pour VTI, alors qu'on a continué d'utiliser l'algorithme exact pour SE. L'incertitude d'origine algorithmique pour VTI, de l'ordre de 1 %, correspond précisément à la différence entre les courbes de VTI et SE. Il ne faut pas sur-interpréter cette différence : l'écart observé est une combinaison, dans des proportions inconnues, d'une (faible) différence de manipulabilité entre VTI et SE et d'une (légère) baisse de performance de l'algorithme approché de VTI quand le nombre de candidats croît.

Ainsi, même s'il est prouvé théoriquement que CVTI est strictement moins manipulable que VTI (par le corolaire 2.21 des théorèmes de condorcification) et que VTI est lui-même strictement moins manipulable que SE (section 1.4.2), ces différences de manipulabilité sont, en fait, très faibles, notamment quand on les compare avec les autres modes de scrutin étudiés ici.

Pour expliquer tous ces phénomènes, on peut avancer l'explication qu'ils sont essentiellement liés à la dégradation de la « qualité » de la population, à sa désunion croissante. En particulier, pour les modes de scrutin vérifiant **CMInf** (tous ceux étudiés ici sauf Veto), on sait, par le lemme 2.7, que les situations non admissibles sont condamnées à être manipulables. On peut donc se demander si l'augmentation du taux de configurations non admissibles est le seul facteur expliquant la croissance des taux de manipulabilité. Pour tester cette hypothèse, nous avons donc indiqué en figure 7.3 le *taux de MC normalisé*, c'est-à-dire rapporté à la proportion de configurations non admissibles. Pour les modes de scrutin vérifiant **CMInf**, il s'agit de leur taux de manipulabilité pour la restriction de la

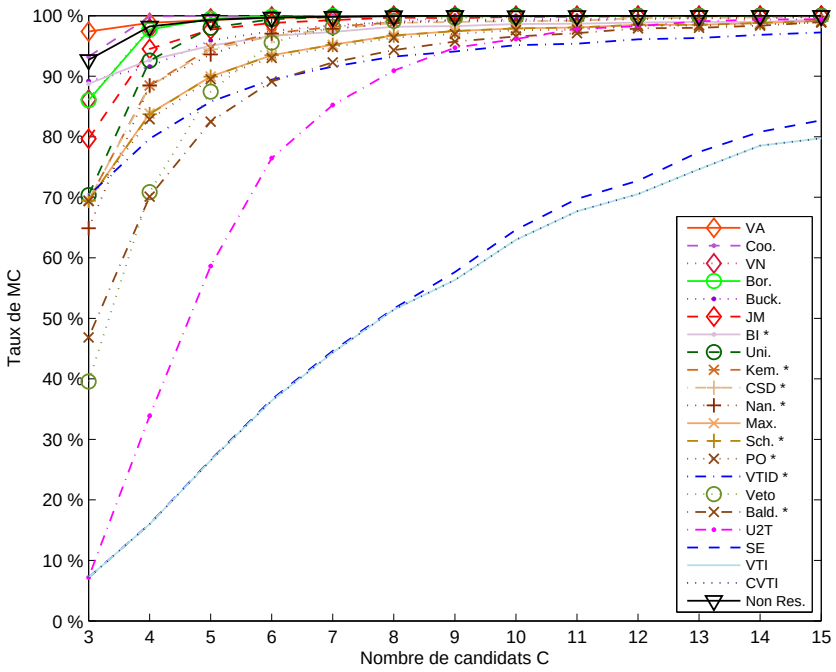


FIGURE 7.3 – Taux de MC normalisé (rapporté aux situations admissibles) en fonction du nombre de candidats C. Culture sphérique, $V = 33$.

culture impartiale aux situations admissibles[5]. La constatation intéressante est que, même ainsi, les phénomènes mentionnés précédemment restent vrais, ce qui nous amène à la conjecture suivante.

Conjecture 7.6

On considère la culture sphérique avec $V \geq 3$.

Le taux de manipulabilité normalisé de chaque mode de scrutin considéré ici est une fonction croissante du nombre de candidats C.

Pour $C \to +\infty$, *le taux de manipulabilité normalisé de chaque mode de scrutin considéré ici, hormis Veto, tend vers* 1.

Cette conjecture n'invalide pas notre explication intuitive concernant la croissance de la manipulabilité. Le fait qu'il y ait de moins en moins souvent de candidat Condorcet-admissible n'est, à notre avis, qu'un symptôme d'un phénomène plus général de désaccord croissant, qui donne une image intuitive expliquant que des modes de scrutin qui ne vérifient pas **CMInf**, comme Veto, soient également touchés.

5. Gehrlein a fait remarquer qu'idéalement, il serait intéressant de comparer les différents systèmes de vote en ignorant l'ensemble des configurations où ils sont tous manipulables : celles-ci ne font qu'augmenter tous les taux de manipulabilité, sans apporter de discrimination entre les différents systèmes. Cependant, un tel procédé dépend en général de l'ensemble des systèmes de votes étudiés.

La normalisation que nous proposons ici, basée sur les configurations admissibles, présente l'avantage d'être universelle pour tous les systèmes de vote vérifiant **CMInf** et, sur cette classe, elle réalise exactement l'objectif formulé par Gehrlein. En effet, toute configuration non-admissible est nécessairement manipulable. Et pour toute configuration admissible, il existe un mode de scrutin vérifiant **CMInf** (et même **Cond**) où elle est non manipulable : il suffit de choisir arbitrairement un de ses candidats Condorcet-admissibles a et de considérer le mode de scrutin Condorcet-doyen, où le doyen est a ; la non-manipulabilité découle alors du lemme 2.4.

7.4 Culture sphérique : nombre d'électeurs V

7.4.1 Nombre d'électeurs V impair

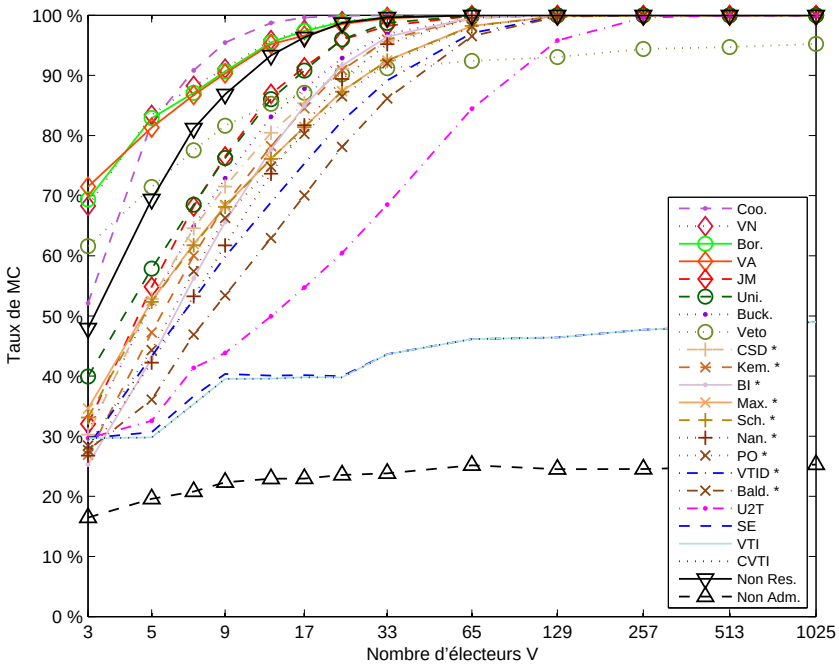


FIGURE 7.4 – Taux de MC en fonction du nombre d'électeurs V. Culture sphérique, $C = 5$, grandes valeurs impaires de V.

Dans la figure 7.4, on trace les taux de MC pour des nombres d'électeurs V impairs. La parité du nombre d'électeurs peut créer des phénomènes particuliers que nous examinerons dans la section 7.4.2.

Pour la plupart des systèmes de vote, le taux de MC semble être une fonction croissante du nombre d'électeurs V (sur l'ensemble des valeurs impaires). Dans le cas particulier de SE, la non-monotonie apparente entre $V = 9$ et $V = 13$ est inférieure à l'incertitude statistique de 1 %, donc même dans ce cas, il n'est pas exclu que le taux de MC soit une fonction croissante du nombre d'électeurs.

Pour CVTI, VTI et SE, on observe sur la figure 7.4 que les taux de MC sont, de nouveau, nettement plus faibles que pour les autres modes de scrutin et que les écarts entre ces trois modes de scrutin sont minimes. Cette fois, on a utilisé l'algorithme exact pour VTI. Comme on sait que VTI est au plus aussi manipulable que SE au sens ensembliste (section 1.4.2), la légère différence entre VTI et SE (au maximum légèrement supérieure à 1 %) n'est due qu'à une réelle différence de manipulabilité, c'est-à-dire à des profils où SE est manipulable mais où VTI ne l'est pas. Cet écart semble tendre vers 0 quand le nombre d'électeurs tend vers $+\infty$. Quant à la différence entre CVTI et VTI, elle est de nouveau extrêmement faible[6].

6. Notre algorithme approché pour CVTI n'a pas pu décider la manipulabilité pour 2 cas sur 10 000 pour $V = 5$ et 1 cas sur 10 000 pour $V = 13$ et $V = 17$. Dans tous les autres cas, la manipulabilité, vraie ou fausse, a pu être décidée et elle est égale à celle de VTI.

Ces modes de scrutin mis à part, l'uninominal à deux tours se comporte mieux que les autres modes de scrutin, sauf à partir d'une centaine d'électeurs où Veto devient moins manipulable.

Pour tous les systèmes de vote hormis CVTI, VTI, SE et Veto, le taux de MC semble tendre vers 1 quand le nombre d'électeurs V tend vers $+\infty$. Mais contrairement au cas $C \to +\infty$, la probabilité d'existence d'un vainqueur de Condorcet ne tend pas vers 0 quand V tend vers $+\infty$ (Gehrlein, 2006), donc on ne peut pas prouver cette observation par le même moyen. Il existe cependant des résultats théoriques sur la question. D'une part, Kim et Roush (1996) montrent les résultats suivant.

- Pour Veto, si $C = 3$ et $V \to +\infty$, le taux de MC tend vers $\frac{1}{2}$.
- Pour Veto, si $C > 3$ et $V \to +\infty$, le taux de MC tend vers une limite strictement comprise entre 0 et 1.
- Pour toutes les RPS sauf Veto, si $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers 1.
- Pour Maximin, si $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers 1.
- Pour la méthode de Coombs, si $C = 3$ et $V \to +\infty$, le taux de MC tend vers 1.

D'autre part, Lepelley et Valognes (1999) montrent que pour SE, si $C = 3$ et $V \to +\infty$, le taux de MC tend vers 0,16887 (valeur approchée) : en particulier, c'est une valeur strictement comprise entre 0 et 1.

À la lumière de ces résultats théoriques et des courbes de la figure 7.4, nous formulons les conjectures suivantes.

Conjecture 7.7

En culture impartiale, pour $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers 1 pour la méthode de Coombs, la méthode de Bucklin, CSD, la méthode de Kemeny, la méthode de Bucklin itérée, la méthode de Schulze, la méthode de Nanson, la méthode par paires ordonnées, VTID, la méthode de Baldwin et l'uninominal à deux tours.

En culture sphérique, il en est de même pour le vote par notation, le vote par assentiment et le jugement majoritaire.

Rappelons qu'un tel résultat est prouvé pour Maximin et pour toutes les RPS hormis Veto.

Conjecture 7.8

En culture impartiale, pour $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers une limite strictement comprise entre 0 et 1 pour CVTI, VTI et SE.

Rappelons qu'un tel résultat est déjà prouvé pour Veto.

En culture sphérique, il semblerait donc que CVTI, VTI, SE et Veto soient les seuls modes de scrutin, parmi ceux étudiés ici, dont le taux de manipulabilité ne tende pas vers 1 quand le nombre d'électeurs V tend vers $+\infty$.

7.4.2 Influence de la parité de V

Si on utilise des valeurs de V paires et impaires, on observe des phénomènes de non-monotonie pour le taux de MC de certains modes de scrutin. Dans la figure 7.5, on considère toutes les valeurs du nombre d'électeurs de 3 à 33 et on représente le taux de MC des modes de scrutin concernés ; les autres sont omis afin de ne pas alourdir la figure.

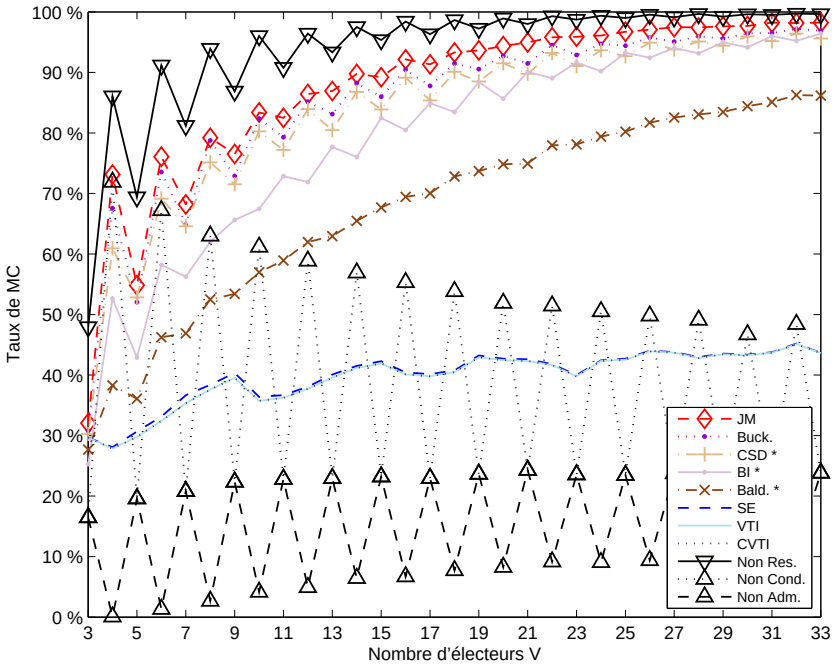


FIGURE 7.5 – Taux de MC en fonction du nombre d'électeurs V pour les modes de scrutin présentant une oscillation. Culture sphérique, $C = 5$, petites valeurs de V.

Tout d'abord, examinons le jugement majoritaire ou la méthode de Bucklin, pour lesquels le taux de MC présente des oscillations sans incertitude algorithmique entre les valeurs paires et impaires de V.

Ce comportement oscillatoire pourrait être lié au fait que les probabilités d'existence d'un vainqueur de Condorcet et d'un candidat Condorcet-admissible oscillent également entre les valeurs paires et impaires de V. Cependant, comme Gehrlein (2006) le prouve par des arguments théoriques, ces oscillations sont en opposition de phase. Quand V est impair, les deux notions sont équivalentes. Quand V passe d'une valeur impaire à une valeur paire adjacente, la probabilité d'avoir un candidat Condorcet-admissible augmente, alors que la probabilité d'avoir un vainqueur de Condorcet diminue. Sur ces courbes obtenues par simulation informatique, on constate qu'il y a aussi des oscillations pour la probabilité d'existence d'un vainqueur de Condorcet résistant, en phase avec le vainqueur de Condorcet [7]. Or, nous l'avons vu, la probabilité d'existence d'un candidat Condorcet-admissible, d'un vainqueur de Condorcet (ce qui était équivalent dans les courbes précédentes) ou d'un vainqueur de Condorcet résistant sont des indicateurs qui ont un impact sur la manipulabilité.

Cependant, si cette explication est correcte, il est surprenant que les autres systèmes de vote vérifiant **CMInf** ne présentent pas le même type de comportement oscillatoire. On peut formuler l'hypothèse qu'il y a deux effets concurrents :

7. On sait également (Gehrlein, 2006) qu'à parité fixée, la probabilité d'existence d'un vainqueur de Condorcet est monotone. En apparence, ceci ne semble pas vérifié dans la figure 7.5 entre $V = 28$, $V = 30$ et $V = 32$. Cependant, la différence constatée entre $V = 30$ et $V = 32$ est inférieure à 2 % et elle est uniquement d'origine statistique. Souvenons-nous que cette incertitude, qui est de l'ordre de 1 % pour une valeur et de $\sqrt{2} \times 1$ % pour la différence entre deux valeurs, n'est qu'un ordre de grandeur qui peut, parfois, être dépassé.

$\frac{V}{2}$ (sincères)	$\frac{V}{2}$ (manipulateurs)
Divers : 1	c : 1
w : 0,9	⋮
Divers : 0,5	Divers : 0,5
c : 0,2	⋮
Divers : 0	w : 0

TABLE 7.1 – Exemple explicatif pour le jugement majoritaire et la méthode de Bucklin.

la probabilité plus forte d'avoir un candidat Condorcet-admissible protège mieux contre la manipulation pour V pair, mais celle d'avoir un vainqueur de Condorcet voire un vainqueur de Condorcet résistant protège mieux contre la manipulation pour V impair. Ainsi, il est possible que les deux effets s'annulent en partie, ce qui pourrait expliquer que peu de modes de scrutin présentent des oscillations. On verra que ce phénomène oscillatoire est beaucoup plus répandu en culture de puits gaussien et nous expliquerons pourquoi (section 8.2.2).

Cette explication soulève une question supplémentaire. Pour les modes de scrutin vérifiant **CMInf**, la borne inférieure habituelle sur le taux de manipulabilité est donnée par la proportion de configurations non admissibles ; donc on pourrait penser que les oscillations de ces courbes seront en phase. Or c'est le contraire qui se produit : les oscillations du jugement majoritaire et de la méthode de Bucklin sont en phase avec la proportion de configurations Condorcet.

Pour ces raisons, nous proposons une autre piste d'explication, qui n'est pas exclusive de la précédente. Le jugement majoritaire et la méthode de Bucklin reposent sur la notion de médiane, qui a une définition légèrement différente selon que le nombre d'électeurs est pair ou impair. Plus précisément, examinons ce qui peut se passer avec un nombre pair d'électeurs, s'il y a exactement $\frac{V}{2}$ électeurs préférant un certain candidat c au vainqueur sincère w. En général, l'électeur sincère typique ne donnera à w ni le rang ni le score maximal ; et il ne donnera à c ni le rang ni le score minimal. Après manipulation, on aboutit au type de profil représenté de façon simplifiée dans la table 7.4.2.

En jugement majoritaire ou en méthode de Bucklin, par convention, on utilise la médiane défavorable, c'est-à-dire la note médiane inférieure en jugement majoritaire et le rang médian supérieur en méthode de Bucklin. Ainsi, la note médiane prise en compte par le jugement majoritaire (resp. le rang médian pris en compte par la méthode de Bucklin) pour w est 0 (resp. C), alors que la note médiane (resp. le rang médian) de c est 0,2 (resp. $C-1$ par exemple). Ainsi, une coalition formée d'un majorité d'électeurs, au sens large, peut généralement manipuler. En revanche, si on prend un type de profil comme l'exemple simplifié ci-dessus mais avec un nombre impair d'électeurs, il faut une majorité stricte d'électeurs pour faire gagner c. On peut ainsi expliquer que le taux de manipulabilité est plus élevé pour un nombre d'électeurs pair.

Pour CSD, le phénomène oscillatoire observé est à confirmer, puisque la courbe tracée n'est qu'une borne inférieure du taux de MC (en pratique, il s'agit du taux de MT). On peut tout de même remarquer que la règle de CSD présente un comportement qui dépend de la parité de V : en effet, la pénalité pour une défaite

de c contre d est $1 + D_{dc} - \frac{V}{2}$ quand V est pair, mais seulement $\frac{1}{2} + D_{dc} - \frac{V}{2}$ quand V est impair.

Pour CVTI, VTI et SE, on observe aussi des effets de non-monotonie, dont l'amplitude est supérieure à la marge d'erreur. Cependant, ils ne semblent pas avoir de pseudo-période évidente. Vu la façon dont ces systèmes de vote sont dépouillés, on peut formuler l'hypothèse que plusieurs effets concurrents se combinent, dépendant du modulo de V par tous les entiers de 2 à C.

Pour BI, comme pour CSD, la courbe ne fournit qu'une borne inférieure, mais elle peut être commentée en tant que courbe du taux de MT. Cette fois, il semble y avoir une transition autour de $V = 8$ ou $V = 9$: les oscillations demeurent mais la phase de la courbe s'inverse ! Pour des plus petites valeurs de V, la manipulabilité est plus grande pour V pair (que pour les valeurs impaires adjacentes) ; pour de plus grandes valeurs, elle est plus grande pour V impair. Comme c'est un mode de scrutin à plusieurs tours, on peut imaginer des effets de modulo comme pour CVTI, VTI et SE et des effets de parité dus à l'utilisation de la médiane comme pour la méthode de Bucklin. En tout cas, le fonctionnement complexe de ce mode de scrutin éloigne la perspective d'une explication simple.

Enfin, on remarque les phénomènes de non-monotonie observés diminuent en amplitude quand le nombre d'électeurs V croît. On peut aisément l'expliquer : quand le nombre d'électeurs est grand, les phénomènes d'égalité impliquant une exacte moitié des électeurs sont plus rares. De manière plaisante, on peut résumer cette observation par l'adage : « l'infini est impair ». Ou encore : « quand l'infini est pair, il ne s'en aperçoit pas ».

7.5 Culture monopolaire : concentration κ

Nous considérons maintenant des cultures de VMF non dégénérées, c'est-à-dire avec une concentration κ non nulle. Rappelons qu'on fixe un vecteur unitaire $\mathbf{n}$, appelé le *pôle* de la distribution. Ensuite, indépendamment pour chaque électeur, on tire un vecteur unité $\mathbf{u}$ dans $\mathbb{R}^C$ selon une distribution de VMF :

$$p(\mathbf{u}) = X_\kappa e^{\kappa \langle \mathbf{u} \,|\, \mathbf{n} \rangle},$$

où X_κ une constante de normalisation.

Maintenant que nous avons étudié l'impact du nombre de candidats ou du nombre d'électeurs, nous allons étudier l'effet de la concentration et de la position du pôle. Puis nous élargirons le modèle à plusieurs pôles afin de représenter plusieurs groupes sociaux possédant chacun une opinion typique.

Dans la figure 7.6, pour chaque population, le pôle $\mathbf{n}$ est tiré au hasard, puis les électeurs sont tirés avec une certaine concentration κ. On remarquera que les électeurs ne sont pas indépendants. Plus précisément, une fois que le pôle est fixé, ils sont indépendants ; mais, sur tout l'ensemble des populations possibles (avec toutes les positions possibles du pôle), les électeurs ne sont pas indépendants. En effet, *a priori*, un électeur a une probabilité uniforme sur la sphère unité de $\mathbb{R}^C$; mais si on sait, par exemple, que les $V - 1$ premiers électeurs sont proches d'un certain point de l'espace des utilités, alors le dernier électeur a une probabilité plus élevée d'être proche de ce point.

Comme on pouvait s'y attendre, plus la concentration κ est élevée, plus les taux de MC sont faibles, en tout cas pour la plupart des modes de scrutin. Dans le cas dégénéré $\kappa = +\infty$, tous les électeurs ont toujours le même vecteur d'utilité, donc aucun système de vote unanime (section 1.2.2) n'est manipulable : c'est le cas de tous les modes de scrutin étudiés ici, hormis Veto (voir ci-dessous).

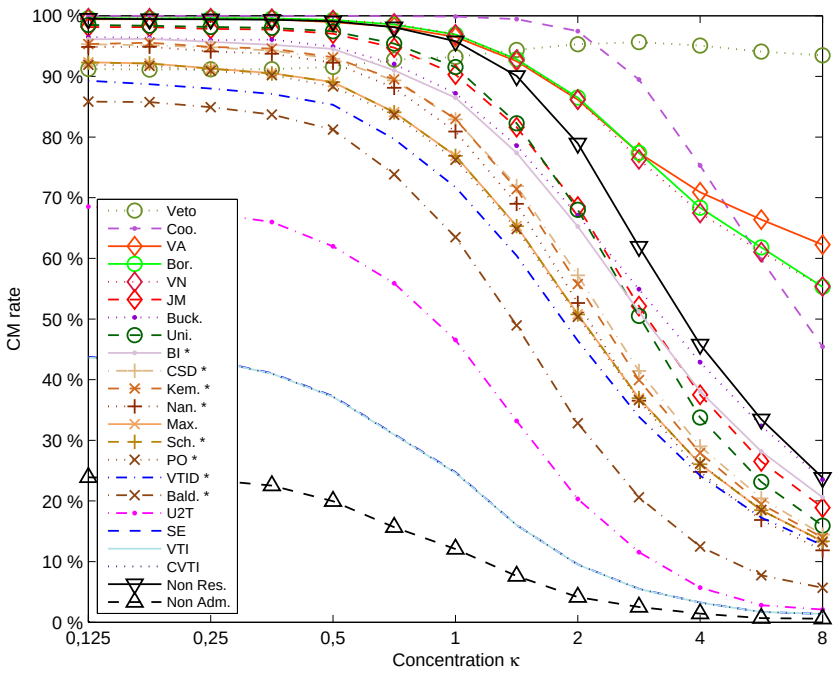


FIGURE 7.6 – Taux de MC en fonction de la concentration κ. Culture de VMF, $V = 33$, $C = 5$, un pôle de position aléatoire.

Cependant, certains systèmes de vote semblent significativement moins réactifs que les autres à une augmentation de la concentration κ : Veto, le vote par assentiment, la méthode de Borda, le vote par notation et, dans une moindre mesure, la méthode de Coombs. Cela signifie que, même quand les électeurs ont des opinions relativement similaires, ces modes de scrutin conservent une manipulabilité élevée.

Pour Veto, le taux de MC n'est même pas monotone. Et nous allons montrer qu'il ne tend pas vers 0, contrairement à ce qui se passe pour les modes de scrutin unanimes.

Proposition 7.9

On considère le mode de scrutin Veto, muni d'une règle de départage qui n'utilise aucune information sur les préférences autre que le candidat contre lequel vote chaque électeur.

On suppose $V \geq C - 1$.

En culture de VMF avec un pôle tiré uniformément, quand $\kappa \to +\infty$, *le taux de manipulabilité tend vers* $1 - \dfrac{1}{C-1}$.

L'hypothèse faite sur la règle de départage est en particulier vraie quand les candidats ex aequo sont départagés par ordre alphabétique ou par ordre d'indice, comme c'est le cas dans SVAMP. Par ailleurs, sur la figure 7.6, comme $C = 5$, le taux limite est $\dfrac{3}{4}$; mais il est loin d'être atteint : la convergence est donc relativement lente.

Démonstration. Dans la limite où tous les électeurs ont le même ordre de préférence, ils votent tous contre le même candidat et c'est seulement la règle de

départage qui désigne le vainqueur parmi les autres candidats. Comme le pôle est tiré de façon neutre, il y a une chance sur $C-1$ que le candidat élu soit celui préféré par les électeurs. Dans tous les autres cas, la totalité des électeurs peut former une coalition et manipuler pour faire gagner leur candidat préféré : en effet, comme $V \geq C-1$, il est possible qu'au moins un électeur vote contre chaque autre candidat. □

Encore une fois, on remarque les taux de MC faibles et extrêmement similaires de CVTI, VTI et SE, suivi des performances intermédiaires de l'uninominal à deux tours, lui-même moins manipulable que les autres modes de scrutin.

7.6 Culture monopolaire : position du pôle

À présent, nous allons travailler avec une concentration κ constante et faire varier la localisation des électeurs. Nous conservons le même ordre de préférence sur les candidats $1 \succ 2 \succ \ldots \succ 5$ mais nous faisons varier la forme de l'opinion de l'électeur typique (c'est-à-dire situé exactement au pôle de la distribution). Tout d'abord, nous considérons le cas limite où l'électeur typique préfère le candidat 1 et où il est indifférent entre les autres candidats ; et nous effectuons une transition vers un état relativement régulier où ses utilités sont les scores de Borda (figure 7.7). Enfin, nous déplaçons l'électeur typique d'un état de type Borda vers un autre cas limite où il déteste le candidat 5 et où il est indifférent entre les autres candidats (figure 7.8).

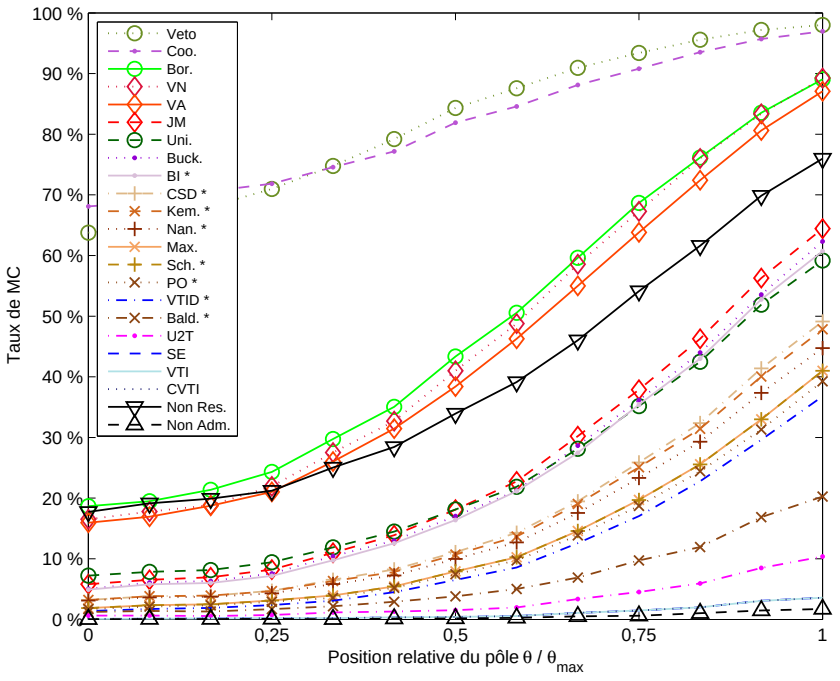


FIGURE 7.7 – Taux de MC en fonction de la position du pôle $\mathbf{n}$. Culture de VMF, $V = 33$, $C = 5$, un pôle de concentration $\kappa = 2$. Pour $\theta = 0$, on a $\mathbf{n} = \mathbf{n}_0 = \frac{1}{\sqrt{20}}(4,-1,-1,-1,-1)$. Pour $\theta = \theta_{\max}$, on a $\mathbf{n} = \mathbf{n}_1 = \frac{1}{\sqrt{10}}(2,1,0,-1,-2)$.

Dans la figure 7.7, la concentration $\kappa = 2$ est ainsi constante, mais la position du pôle $\mathbf{n}$ est imposée et non plus tirée au hasard. Considérons deux vecteurs

d'utilité unitaires $\mathbf{n}_0 = \frac{1}{\sqrt{20}}(4,-1,-1,-1,-1)$ et $\mathbf{n}_1 = \frac{1}{\sqrt{10}}(2,1,0,-1,-2)$. En notant $\theta_{\max}$ l'angle entre $\mathbf{n}_0$ et $\mathbf{n}_1$, on a $\theta_{\max} = 45°$. On explore la géodésique de la sphère unité allant de $\mathbf{n}_0$ à $\mathbf{n}_1$, en prenant comme paramètre l'angle polaire θ parcouru depuis $\mathbf{n}_0$. Quand $\theta = 0$, on a $\mathbf{n} = \mathbf{n}_0$: l'électeur typique préfère le candidat 1 et il est indifférent entre les autres candidats. Quand $\theta = \theta_{\max}$, on a $\mathbf{n} = \mathbf{n}_1$: l'électeur typique a un ordre de préférence bien établi $1 \succ 2 \succ 3 \succ 4 \succ 5$, avec des utilités qui sont les scores de Borda à normalisation près.

On observe que tous les systèmes de vote deviennent plus manipulables quand le pôle de la distribution s'approche de $\mathbf{n}_1$ (avec des amplitudes très différentes pour ce phénomène selon le système de vote). Nous proposons l'explication suivante. Avec une même valeur de κ, quand le pôle est $\mathbf{n}_0$, la population est fortement polarisée en faveur du candidat 1, donc il est assez improbable qu'un électeur préfère un autre candidat au candidat 1 ; mais, quand le pôle est $\mathbf{n}_1$, il est plus proche de points de l'hypersphère où les électeurs préfèrent le candidat 2 au candidat 1. Par conséquent, les coalitions en faveur du candidat 2 ont plus de membres et ont plus de chances de réussir une manipulation.

Comme précédemment, on note les bonnes performances de CVTI, VTI et SE, suivies de celles de l'uninominal à deux tours.

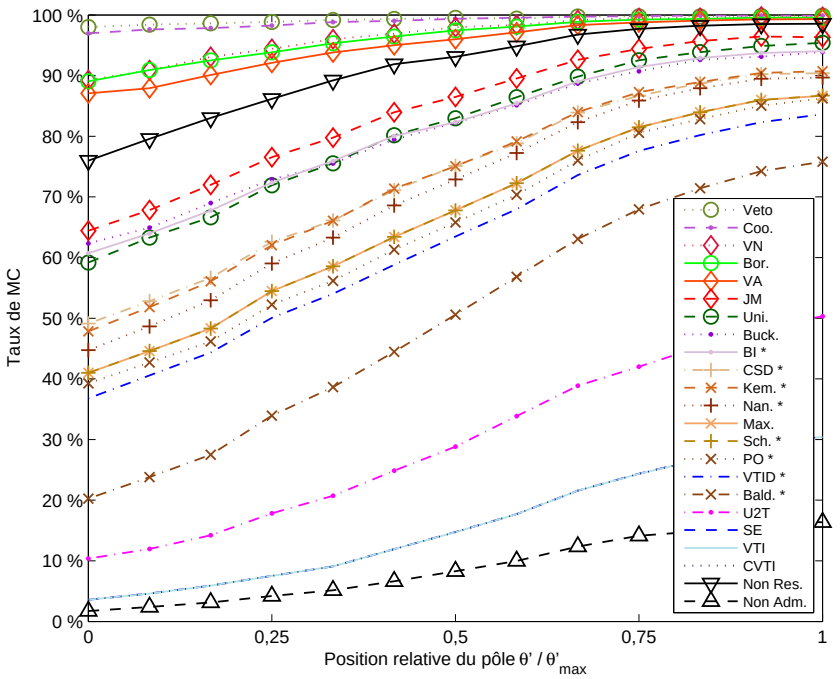


FIGURE 7.8 – Taux de MC en fonction de la position du pôle $\mathbf{n}$ (suite). Culture de VMF, $V = 33$, $C = 5$, un pôle de concentration $\kappa = 2$. Pour $\theta' = 0$, on a $\mathbf{n} = \mathbf{n}_1 = \frac{1}{\sqrt{10}}(2,1,0,-1,-2)$. Pour $\theta' = \theta'_{\max}$, on a $\mathbf{n} = \mathbf{n}_2 = \frac{1}{\sqrt{20}}(1,1,1,1,-4)$.

Dans la figure 7.8, on poursuit l'investigation en se dirigeant vers le pôle $\mathbf{n}_2 = \frac{1}{\sqrt{20}}(1,1,1,1,-4)$. En notant $\theta'_{\max}$ l'angle entre $\mathbf{n}_1$ et $\mathbf{n}_2$, on a de nouveau $\theta'_{max} = 45°$. On prend comme paramètre θ', l'angle polaire parcouru depuis $\mathbf{n}_1$. Pour $\theta' = 0$, on a $\mathbf{n} = \mathbf{n}_1$: l'électeur typique a un ordre de préférence bien établi $1 \succ 2 \succ 3 \succ 4 \succ 5$. Pour $\theta' = \theta'_{max}$, on a $\mathbf{n} = \mathbf{n}_2$: l'électeur typique a une aversion pour le candidat 5 et il est indifférent entre les autres candidats.

On remarquera que le trajet parcouru de $\mathbf{n}_0$ vers $\mathbf{n}_1$ puis de $\mathbf{n}_1$ vers $\mathbf{n}_2$ n'est pas une géodésique de la sphère[8] : en effet, comme on le montre dans l'annexe B, les géodésiques de l'espace des utilités sont les zones d'unanimité. Or en $\mathbf{n}_0$ et $\mathbf{n}_2$, l'électeur typique est indifférent entre les candidats 2, 3 et 4, alors que ce n'est pas le cas en $\mathbf{n}_1$. Le trajet parcouru par le pôle $\mathbf{n}$ sur la sphère entre la figure 7.7 et la figure 7.8 est donc une ligne brisée de $\mathbf{n}_0$ vers $\mathbf{n}_1$ puis de $\mathbf{n}_1$ vers $\mathbf{n}_2$.

Sur la figure 7.8, on constate que les taux de MC, de configurations non admissibles et de configurations non résistantes continuent d'augmenter. La culture devient semblable à une culture uniforme sur les quatre candidats crédibles (tous sauf le candidat 5), hormis le fait qu'il y a un candidat supplémentaire moins apprécié par les électeurs. Celui-ci n'a aucune chance de gagner dans tous les modes de scrutin raisonnables (même Veto) mais il peut perturber les modes de scrutin les plus sensibles à l'ajout d'un candidat non pertinent, par exemple la méthode de Borda : pour un manipulateur, ce candidat permet d'ajouter un point à la différence de score entre le candidat souhaité c et le vainqueur du vote sincère w, donc il augmente la possibilité de manipulation par rapport à une culture uniforme sur 4 candidats.

7.7 Culture multipolaire : nombre de pôles

Jusqu'à présent, nous avons considéré des cultures avec un seul pôle $\mathbf{n}$. Cependant, en pratique, une population d'agents n'a pas toujours des préférences centrées autour d'une opinion typique. On peut notamment envisager qu'il existe plusieurs groupes sociaux qui possèdent, chacun, une certaine cohésion. Nous employons ici l'expression *groupe social* dans une acception très large : par exemple, dans une élection politique, il peut s'agir d'une classe socio-professionnelle ou d'une communauté d'intérêt ; pour une élection dans d'autres contextes, comme une organisation professionnelle ou associative par exemple, il peut s'agir d'un groupe d'individus possédant une certaine cohésion, qu'elle soit fondée sur des bases idéologiques, personnelles ou autres.

Dans la figure 7.9, on élargit donc le modèle de VMF en considérant une culture multipolaire représentant une population constituée de plusieurs groupes sociaux. Formellement, on prend comme paramètre k, qui représente le nombre de groupes. Voici comment se déroule le tirage aléatoire d'une configuration. Tout d'abord, on tire k vecteurs unitaires $\mathbf{n}_1, \ldots, \mathbf{n}_k$ indépendamment et uniformément qui seront les pôles de la distribution, c'est-à-dire le vecteur d'utilité de l'électeur typique de chaque groupe social. Ensuite, pour chaque électeur indépendamment, on tire un entier i entre 1 et k de façon équiprobable qui représente son groupe social, puis on tire son vecteur d'utilité selon une distribution de VMF de pôle $\mathbf{n}_i$ et de concentration κ. Afin de simplifier l'étude, on considère donc que les groupes sociaux sont de taille comparables (puisque i est tiré de façon équiprobable) et de même concentration constante κ, même si SWAMP permet de faire également varier ces paramètres.

Comme on pouvait s'y attendre, on constate sur la figure 7.9 que les taux de MC croissent avec le nombre de pôles (sauf pour Veto).

Avec V et C fixés, quand le nombre de pôles k tend vers $+\infty$, la probabilité qu'il existe une paire d'électeurs qui sont dans le même groupe social tend vers 0, donc la distribution devient uniforme sur l'hypersphère : ainsi, la distribution limite est tout simplement la culture sphérique. Les courbes permettent de confirmer

8. À titre indicatif, si on poursuit l'arc de cercle de $\mathbf{n}_0$ vers $\mathbf{n}_1$ jusqu'à $\theta = 90°$, on obtient $\mathbf{n} = \frac{1}{\sqrt{20}}(0, 3, 1, -1, -3)$, ce qui ne représente pas le même ordre de préférence.

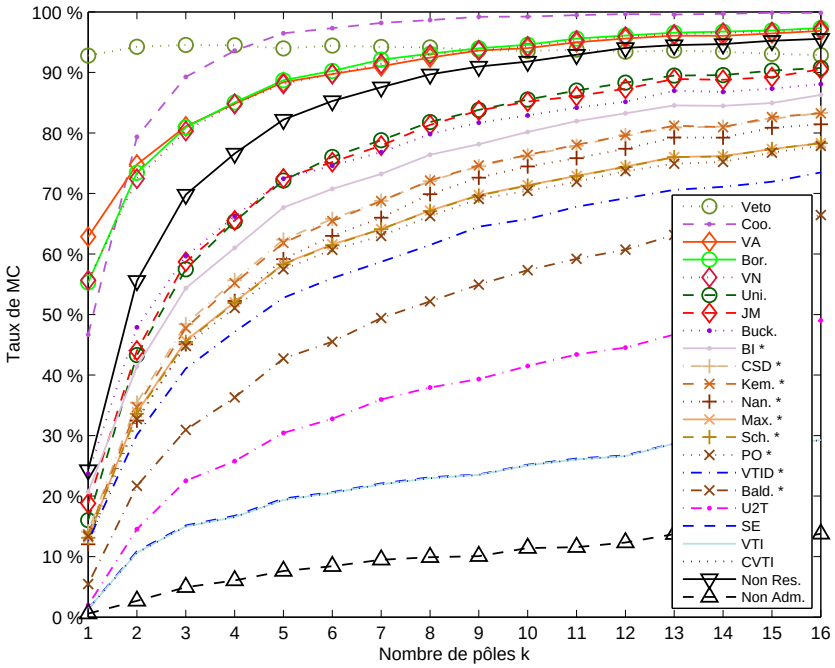


FIGURE 7.9 – Taux de MC en fonction du nombre de pôles. Culture de VMF, $V = 33$, $C = 5$, pôles équiprobables de positions aléatoires, $\kappa = 2$.

cette remarque théorique : pour $C = 16$, les taux de MC sont certes plus faibles mais déjà comparables à ceux observés pour une culture sphérique avec $V = 33$ et $C = 5$ (cf. figure 7.2 par exemple).

7.8 Culture multipolaire : positions relatives des pôles

Pour terminer, nous allons nous pencher sur la position relative des groupes sociaux dans l'espace des utilités. L'exemple le plus simple pour étudier ce phénomène consiste à considérer deux groupes sociaux de taille semblable (chaque électeur est équiprobablement dans l'un ou l'autre) et de même concentration κ.

Considérons les deux vecteurs d'utilité $\mathbf{n}_0 = \frac{1}{\sqrt{20}}(4,-1,-1,-1,-1)$ et $\mathbf{n}_1 = \frac{1}{\sqrt{20}}(-1,4,-1,-1,-1)$. En notant $\theta_{\max}$ l'angle entre $\mathbf{n}_0$ et $\mathbf{n}_1$, on a $\cos(\theta_{\max}) = -\frac{1}{4}$ donc $\theta_{max} \simeq 104°$.

Le premier pôle que nous prenons est toujours $\mathbf{n}_0$. Le second pôle explore la portion du cercle unité de $\mathbf{n}_0$ à $\mathbf{n}_1$, en utilisant l'angle θ comme paramètre. Quand $\theta = 0$, les deux pôles coïncident en $\mathbf{n}_0$: l'électeur typique préfère le candidat 1 et il est indifférent entre les autres candidats. Quand $\theta = \theta_{\max}$, le second pôle est $\mathbf{n}_1$; l'électeur typique du second groupe social préfère le candidat 2 et il est indifférent entre les autres candidats.

On observe sur la figure 7.10 que le taux de MC croît pour tous les systèmes de vote considérés quand θ croît. Ce n'est pas surprenant : quand le désaccord croît dans la population, la motivation pour manipuler croît également.

Encore une fois, on notera les performances remarquables de CVTI, VTI et SE, suivies de celles de l'uninominal à deux tours.

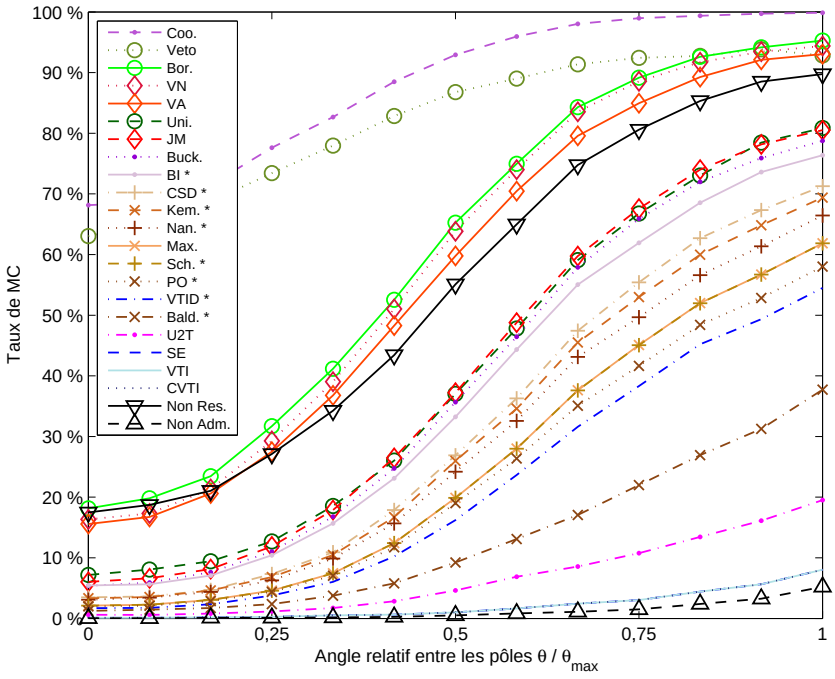


FIGURE 7.10 – Taux de MC en fonction de l'angle entre les pôles. Culture de VMF, $V = 33$, $C = 5$, deux pôles équiprobables, $\kappa = 2$. Le premier pôle est $\mathbf{n}_0 = \frac{1}{\sqrt{20}}(4, -1, -1, -1, -1)$. Pour $\theta = 0$, le second pôle est identique au premier. Pour $\theta = \theta_{max}$, le second pôle est $\mathbf{n}_1 = \frac{1}{\sqrt{20}}(-1, 4, -1, -1, -1)$.

7.9 Méta-analyse dans les cultures sphéroïdales

Afin de ne pas alourdir davantage cette étude, nous épargnons au lecteur méritant la variation d'autres paramètres : ainsi, on ne fera pas varier les tailles relatives typiques des groupes sociaux (c'est-à-dire les probabilités relatives qu'un électeur donné appartiennent à chacun d'entre eux) et on n'étudiera pas le cas où les concentrations κ sont différentes d'un groupe social à l'autre. Les résultats trouvés sont similaires et chacun peut reproduire de telles expériences en utilisant SWAMP.

De toutes les courbes précédentes, il semble se dégager une tendance générale. Certains modes de scrutin sont presque toujours moins manipulables que les autres : CVTI, VTI et SE. La seule exception a été obtenue pour $V = 3$ électeurs (figure 7.4), où la borne inférieure trouvée pour certains modes de scrutin est plus faible que le taux trouvé pour CVTI, VTI et SE. Cependant, on remarquera qu'avec 3 électeurs, la règle de départage joue un rôle si important qu'il est difficile d'en tirer des conclusions définitives. Nous choisirons donc d'ignorer le cas particulier $V = 3$ dans la suite de cette section.

Dans ce chapitre et les suivants, nous recourrons souvent à une *méta-analyse*, qui permet de présenter de façon compacte les résultats obtenus pour un ensemble de cultures qui peut être hétérogène et d'examiner les performances comparées des différents modes de scrutin.

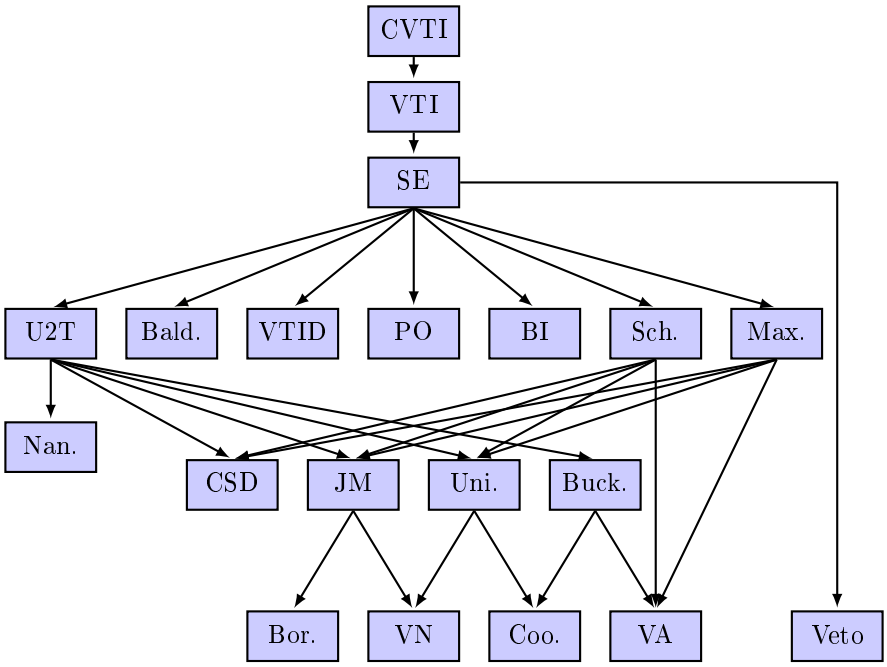


FIGURE 7.11 – Méta-analyse de MC dans les cultures sphéroïdales étudiées (sauf $V = 3$).

7.9.1 Méta-analyse de MC dans les cultures sphéroïdales

Le graphe de la figure 7.11 illustre cette méthode pour le taux de MC. Chaque sommet représente un mode de scrutin. Une arête d'un mode de scrutin f vers un mode de scrutin g signifie que, pour toutes les cultures observées dans ce chapitre (hormis $V = 3$), c'est-à-dire pour chaque point de chaque courbe de ce chapitre, le mode de scrutin f est prouvé au plus aussi manipulable que g.

Plus précisément, pour un mode de scrutin f et une culture donnée (c'est-à-dire avec un jeu de paramètres fixé), notons $\underline{\tau_{\mathrm{MC}}(f)}$ la borne inférieure trouvée par SWAMP et $\overline{\tau_{\mathrm{MC}}(f)}$ la borne supérieure. Pour deux modes de scrutin distincts f et g, on trace une arête de f vers g ssi on a toujours $\overline{\tau_{\mathrm{MC}}(f)} \leq \underline{\tau_{\mathrm{MC}}(g)}$.

Dans toutes nos méta-analyses, nous excluons l'incertitude statistique, c'est-à-dire que nous considérons la *proportion de configurations manipulables* sur chaque tirage de 10 000 expériences, et non le *taux de manipulabilité* exact sur la culture sous-jacente. La motivation en est la suivante. Imaginons qu'un mode de scrutin soit généralement beaucoup moins manipulable qu'un autre ; mais que pour une culture donnée, les deux modes de scrutin aient un taux de MC proche de 100 %. En raison de ce dernier cas, comme il y une incertitude statistique de l'ordre de 1 %, il est impossible de conclure que le premier mode de scrutin a toujours un taux de manipulabilité inférieur au second sur les cultures considérées. Or ce cas est assez courant, donc nous ne pourrions pas conclure pour une part importante des paires de modes de scrutin. En considérant la proportion de configurations manipulables dans les tirages effectués, il est plus souvent possible de conclure. Il faudra donc être conscient que la méta-analyse fournit des résultats sur un ensemble d'expériences aléatoires, et non sur un ensemble de cultures. Malgré

cette limitation, elle fournit une présentation synthétique des résultats qui fournit une indication qualitative de ce qui se passe dans les cultures étudiées.

L'autre avantage d'ignorer l'incertitude statistique est que le graphe de méta-analyse obtenu est nécessairement transitif (sinon, on pourrait avoir une relation binaire du même type que dans l'exemple 1.7). Dans un but de lisibilité, nous ne représentons que l'ensemble minimal d'arêtes qui permet de déduire l'ensemble du graphe par transitivité.

Pour certains modes de scrutin, comme la méthode de Baldwin, l'algorithme approché fournit essentiellement une borne inférieure de manipulabilité, donc il n'est pas possible d'établir que le mode de scrutin est au plus aussi manipulable qu'un autre. En revanche, il est possible de montrer qu'il est plus manipulable : c'est pour cette raison que la méthode de Baldwin, par exemple, possède des arêtes entrantes mais pas d'arête sortante, malgré le fait que la borne inférieure du taux de manipulabilité, représentée dans les courbes, soit généralement relativement faible. Pour Maximin et les méthodes de Schulze et Borda, la borne supérieure trouvée est suffisamment précise pour qu'on soit capable d'établir leur supériorité par rapport à certains autres modes de scrutin, même si le cas n'est pas observé en pratique pour la méthode de Borda.

La principale conclusion de la figure 7.11 est que CVTI, VTI et SE sont généralement moins manipulables que les autres modes de scrutin. Bien qu'on sache, par des arguments théoriques, que CVTI est strictement moins manipulable que VTI, lui-même strictement moins manipulable que SE, les différences constatées sont généralement très faibles : pour chaque tirage de 10 000 expériences effectué pour les courbes de ce chapitre, la différence dans les proportions de profils manipulables est au maximum 0,02 % entre CVTI et VTI et 1,3 % entre CVTI et SE.

Dans les simulations effectuées, ces modes de scrutin sont suivis de l'uninominal à deux tours, qui est généralement (mais pas toujours) moins manipulable que les autres modes de scrutin étudiés. Par exemple, la différence de manipulabilité en faveur de VTID contre l'uninominal à deux tours n'est jamais supérieure à 1,3 % ; inversement, la différence de manipulabilité en faveur de l'uninominal à deux tours par rapport à VTID peut dépasser 57 %. L'uninominal à deux tours ne peut être nettement dominé que par Veto, et seulement pour un nombre d'électeurs élevé. On peut ensuite distinguer informellement plusieurs groupes, regroupés par ligne dans la figure.

Pour la méthode de Baldwin, VTID, PO et BI, les manipulabilités trouvées sont en général assez comparables, mais il ne s'agit que de bornes inférieures qui ne permettent que d'établir leurs défaites face à SE et leur quasi-infériorité face à l'uninominal à deux tours. Quant à la méthode de Nanson, elle est toujours plus manipulable que l'uninominal à deux tours dans les courbes de ce chapitre. Maximin et la méthode de Schulze ont des performances très similaires (les différences constatées sont toujours inférieures à 1 %) et meilleures que celles de l'uninominal, de CSD et du jugement majoritaire. Elles sont également souvent meilleures que la méthode de Bucklin (l'écart de manipulabilité, quand il est en faveur de celle-ci, n'excède jamais 1 %).

L'algorithme approché utilisé pour CSD ne permet pas d'établir s'il est moins manipulable que les modes de scrutin du groupe inférieur ; on remarque simplement qu'il n'a pas été prouvé plus manipulable que l'uninominal, le jugement majoritaire ou la méthode de Bucklin, d'où sa position indicative dans la figure.

Les modes de scrutin présentant les pires performances sont la méthode de Borda (plus manipulable que le jugement majoritaire), le vote par notation (plus manipulable que le jugement majoritaire ou l'uninominal), la méthode de Coombs

(plus manipulable que l'uninominal ou la méthode de Bucklin) et le vote par assentiment (plus manipulable que la méthode de Bucklin). Dans de nombreuses figures de ce chapitre, les courbes de manipulabilité de ces quatre modes de scrutin sont au-dessus de la proportion de configurations non résistantes : pour les cas concernés, cela signifie que tout mode de scrutin respectant le critère de Condorcet est moins manipulable.

Veto peut être considéré comme un groupe en lui-même, puisqu'il se comporte de façon très différente selon les cas d'étude. D'un côté, il a un avantage rare : en culture impartiale, son taux de manipulabilité ne tend pas vers 1 quand le nombre d'électeurs V tend vers $+\infty$, contrairement à toutes les autres RPS et à Maximin. Sur la base des simulations effectuées, nous conjecturons même que, parmi les modes de scrutin étudiés ici, les seuls à partager cette propriété souhaitable sont CVTI, VTI et SE. De l'autre côté, Veto a un inconvénient très regrettable : dans une population où tous les électeurs ont les mêmes préférences, son taux de manipulabilité ne tend pas vers 0, contrairement à tous les modes de scrutin étudiés ici et plus généralement à tous ceux qui sont unanimes.

7.9.2 Méta-analyse de MT dans les cultures sphéroïdales

La figure 7.11 présente l'avantage de traiter la MC en toute généralité, mais avec l'inconvénient de ne pas pouvoir apporter de conclusions complètes concernant certains modes de scrutin pour lesquels on utilise un algorithme approché. Afin d'avoir une indication de manipulabilité complémentaire pour tous les modes de scrutin, on représente un graphe similaire pour la MT en figure 7.12. En outre, la MT fournit une mesure indicative des manipulations réalistes, qu'il est possible de réaliser avec un échange limité d'information. Pour la MT, nous faisons grâce au lecteur des courbes détaillées et nous passons directement à la méta-analyse.

Commençons par remarquer que la plupart des arêtes de la figure 7.11 (représentées graphiquement ou sous-entendues par transitivité) sont incluses dans celles de la figure 7.12 : ainsi, la MT fournit un indicateur raisonnable sur la hiérarchie de manipulabilité entre les modes de scrutin.

On observe que Veto forme une composante connexe isolée : il est parfois plus MT, parfois moins MT que tout autre mode de scrutin.

Pour la MT comme pour la MC, les modes de scrutin CVTI, VTI, SE, dans cet ordre, sont toujours moins manipulables que tous les autres modes de scrutin étudiés, hormis Veto. Ils sont toujours suivis de l'uninominal à deux tours, qui présente des performances intermédiaires. Ensuite, on peut, à titre indicatif, distinguer trois groupes, par MT croissante :

— Maximin et la méthode de Schulze (avec des performances très similaires), PO et la méthode de Baldwin ;
— CSD, BI, la méthode de Nanson et l'uninominal (même si la méthode de Nanson est moins MT que celui-ci) ;
— le vote par notation, le jugement majoritaire, le vote par assentiment, la méthode de Borda et la méthode de Coombs (tout en notant que le jugement majoritaire est moins MT que le vote par notation).

La méthode de Bucklin est un peu à part dans la figure : elle est toujours plus MT que l'uninominal à deux tours et toujours moins que le vote par notation, mais elle n'est comparable à aucun mode de scrutin des groupes intermédiaires. Le cas de VTID est similaire : toujours plus MT que l'uninominal à deux tours, moins MT que l'uninominal et que tous les modes de scrutin du dernier groupe, mais non comparable aux autres.

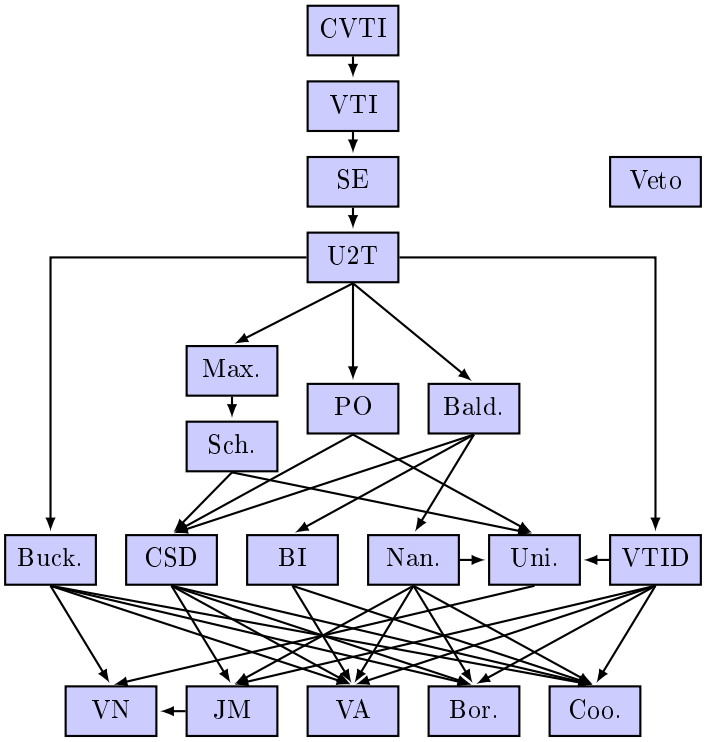


FIGURE 7.12 – Méta-analyse de MT dans les cultures sphéroïdales étudiées (sauf $V = 3$).

Chapitre 8

Étude de cultures basées sur un spectre politique

Dans ce chapitre, nous allons étudier les cultures basées sur un *spectre politique*, que nous avons simplement évoquées dans la section 6.1.3. Notre modèle de référence sera celui du *puits gaussien* et nous discuterons rapidement un autre modèle similaire, la *boîte euclidienne*, dans la section 8.5.

Dans ces modèles, on utilise un espace à n dimensions (pour un certain entier n) qu'on appelle le *spectre politique*. En puits gaussien, on prend comme paramètre un vecteur de réels positifs $(\sigma_1, \ldots, \sigma_n)$: chaque nombre σ_i est appelé la *longueur caractéristique* (en abrégé, *longueur*) du spectre politique selon l'axe i. Pour chaque électeur v (resp. chaque candidat c), on tire aléatoirement une position $\mathbf{x}_v = (x_v^1, \ldots, x_v^n)$ (resp. $\mathbf{y}_c = (y_c^1, \ldots, y_c^n)$). Chaque coordonnée x_v^i (resp. y_c^i) est tirée indépendamment selon une loi normale centrée d'écart-type σ_i. L'utilité d'un électeur v pour un candidat c est $A - \delta(\mathbf{x}_v, \mathbf{y}_c)$, où δ désigne la distance euclidienne usuelle et où A est une constante telle que la moyenne des utilités soit nulle. Notons que cette constante A n'a d'influence que sur le vote par assentiment.

Quand le spectre politique est unidimensionnel ($n = 1$), la culture est manifestement *single-peaked* (définition 1.11). Si le nombre d'électeurs est impair, il y a toujours un vainqueur de Condorcet qui est le candidat préféré de l'électeur médian (sur le spectre politique). Si le nombre d'électeurs est pair, les candidats Condorcet-admissibles sont les candidats préférés des deux électeurs médians ; s'il s'agit du même candidat, il est vainqueur de Condorcet. Par conséquent, la proportion de configurations non admissibles est nulle et ne peut nous fournir une borne inférieure de manipulabilité exploitable, contrairement au chapitre précédent. Pour mémoire et par souci d'homogénéité dans la présentation des résultats, nous continuerons, cependant, à la représenter dans les figures.

De façon imagée, dans un spectre politique unidimensionnel, nous appellerons les candidats Condorcet-admissibles des *centristes* [1]. De même, les candidats possédant une abscisse plus élevée (resp. moins élevée) dans le spectre politique seront appelés des candidats de *droite* ou de *gauche*.

Par ailleurs, dans le cas unidimensionnel, il est inutile de préciser la valeur de σ : en effet, à changement d'unité de longueur près dans le spectre politique, tous les modèles obtenus sont équivalents (car les vecteurs d'utilité sont définis à

1. Il ne s'agit que d'une convention de langage basée sur une analogie. Même si on accepte l'hypothèse que les préférences des électeurs dans une élection politique sont approximativement *single-peaked*, le vainqueur de Condorcet éventuel n'est pas forcément le candidat centriste, au sens politique habituel du terme. Par ailleurs, nous pourrons discuter la validité de cette hypothèse en examinant les résultat obtenus dans le chapitre 9 pour des expériences réelles.

une constante multiplicative près, voir annexe B). En revanche, dès qu'il y a deux dimensions, les longueurs relatives du spectre politique selon les différents axes sont importantes. La validité expérimentale de modèles multidimensionnels de ce type pour des élections politiques est démontrée notamment par Laslier (2004, 2006).

Enfin, rappelons que même quand les préférences sont toujours *single-peaked*, si l'ordre de référence P_{ref} n'est pas fixé *a priori*, la conclusion du théorème de Gibbard-Satterthwaite est toujours vraie : tout système non trivial est manipulable (section 1.3).

8.1 Puits gaussien unidimensionnel : nombre de candidats C

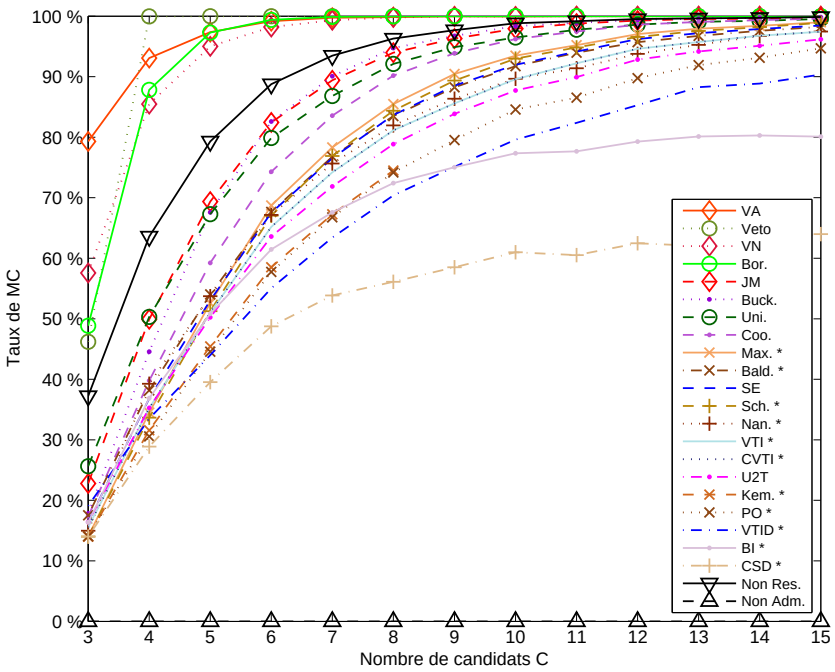


FIGURE 8.1 – Taux de MC en fonction du nombre de candidats C. Puits gaussien, $V = 33$, $n = 1$.

Dans la figure 8.1, on trace le taux de MC en fonction du nombre de candidats C dans un puits gaussien unidimensionnel.

La première conclusion de la figure 8.1 est que les taux de manipulabilité semblent tendre vers 1, tout comme en culture sphérique. En ce qui concerne CSD et BI, il est difficile de se prononcer sur la limite éventuelle, puisqu'on ne dispose que de bornes inférieures de manipulabilité qui ne tendent pas de manière évidente vers 1. De manière générale, les convergences observées vers un taux de manipulabilité égal à 1 semblent plus lentes qu'en culture sphérique. Dans le cas particulier de Veto, le taux de manipulabilité semble également tendre vers 1 mais on peut de nouveau prouver que c'est faux (avec la règle de départage par ordre d'indice) : en effet, on peut utiliser le même raisonnement que pour la proposition 7.5 car la restriction d'une culture de puits gaussien aux candidats $1, \ldots, V+1$ est également une culture de puits gaussien unidimensionnelle.

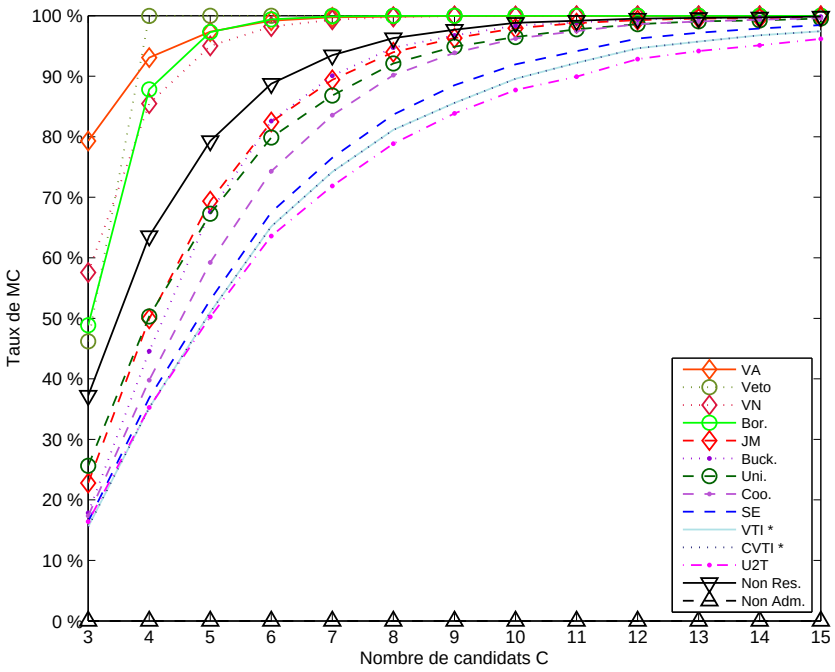


FIGURE 8.2 – Taux de MC en fonction du nombre de candidats C, pour une sélection de modes de scrutin. Puits gaussien, $V = 33$, $n = 1$.

On observe que les performances comparées des différents systèmes de vote ne sont pas les mêmes que ceux discutés dans le chapitre 7. En particulier, à présent, il n'est plus clair que CVTI, VTI et SE aient des taux de manipulabilité plus bas que les autres modes de scrutin. Afin de confirmer cette observation, la figure 8.2 présente les mêmes courbes, en se restreignant à ces trois modes de scrutin et à ceux pour lesquels la marge d'erreur est inférieure à 1 % en tout point (ce qui inclut la méthode de Borda, bien que l'algorithme utilisé ne soit pas exact).

Cette fois-ci, les marges d'erreur pour CVTI et VTI sont parfois supérieures à 1 % car on a utilisé l'option `fast` de SWAMP quand le nombre de candidats C est supérieur ou égal à 9 (d'où l'astérisque indiquée dans la légende de la figure 8.2). Cependant, on a utilisé l'algorithme exact pour SE, donc la courbe de SE fournit une borne supérieure pour celles de CVTI et VTI.

Sur la figure 8.2, on observe que l'uninominal à deux tours est (légèrement) meilleur que CVTI, VTI et SE, contrairement à ce qui se passait dans des cultures sphériques ou VMF. On peut en proposer une explication qualitative. En VTI, une fois que les candidats d'extrême-gauche et d'extrême-droite sont éliminés, leurs partisans votent pour des candidats de gauche et de droite modérés, et le centriste risque d'être éliminé parce qu'il bénéficie de peu de voix transférées. Comme VTI vérifie **CMInf**, cela signifie alors que la situation est manipulable. En revanche, en uninominal à deux tours, les candidats de gauche et droite modérées ne bénéficient pas de voix transférées au moment où la sélection pour le tour final (c.à.d. le second tour) est faite, donc le centriste a davantage de chances d'arriver au second tour.

On notera cependant sur la figure 8.2 que CVTI, VTI et SE restent moins manipulables que la méthode de Coombs, elle-même moins manipulable que l'uninominal, la méthode de Bucklin et le jugement majoritaire, eux-mêmes moins

manipulables que la méthode de Borda, le vote par notation, Veto et le vote par assentiment. Ainsi, la hiérarchie constatée précédemment sur ces modes de scrutin semble se confirmer, sauf pour la méthode de Coombs, qui semble mieux se comporter ici que dans les cultures sphériques et VMF. On peut avancer l'explication que ce bon comportement est en partie lié au fait que dans une culture *single-peaked*, la méthode de Coombs vérifie le critère de Condorcet (Grofman et Feld, 2004).

Pour la comparaison entre CVTI, VTI, SE et les autres modes de scrutin, la question reste ouverte. Dans des travaux futurs, il serait intéressant de développer des algorithmes plus performants pour ces modes de scrutin, afin d'établir s'ils sont moins manipulables que CVTI, VTI et SE dans une telle culture de puits gaussien unidimensionnel. Dans les méta-analyses des sections 8.6 et 8.7, l'utilisation de la MT nous permettra cependant d'en dire davantage.

8.2 Puits gaussien unidimensionnel : nombre d'électeurs V

8.2.1 Nombre d'électeurs V impair

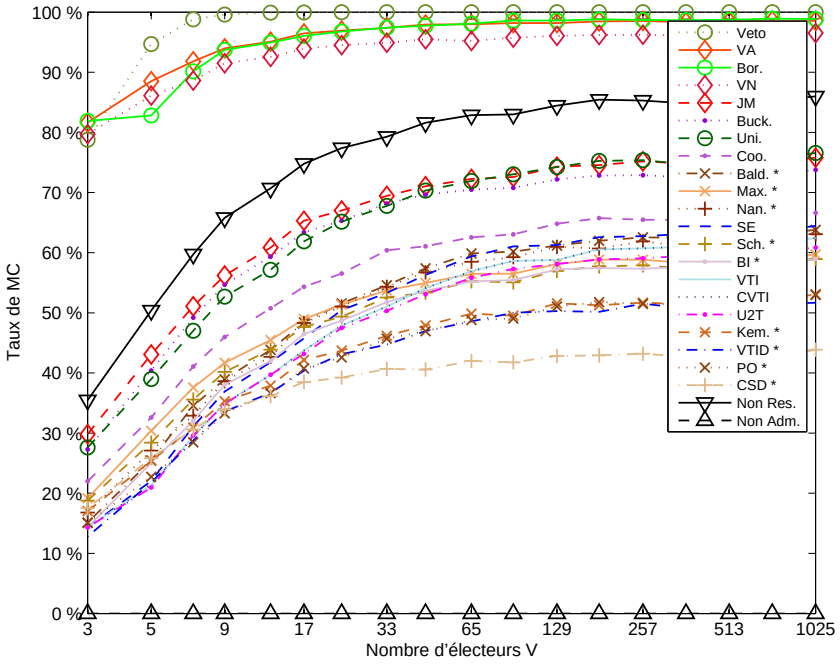


FIGURE 8.3 – Taux de MC en fonction du nombre d'électeurs V. Puits gaussien, $C = 5$, $n = 1$, grandes valeurs impaires de V.

De même que nous l'avions examiné en culture sphérique, la figure 8.3 représente le taux de MC en fonction du nombre d'électeurs V pour de grandes valeurs impaires de celui-ci, toujours pour un puits gaussien unidimensionnel. Comme en culture sphérique, les taux de manipulabilité des divers modes de scrutin semblent être des fonctions croissantes du nombre d'électeurs V (sur l'ensemble des valeurs impaires). Cependant, ils ne semblent pas tendre vers 1 pour la plupart des modes de scrutin, ce qui nous amène à la conjecture suivante.

Conjecture 8.1

En puits gaussien unidimensionnel, pour $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers 1 pour Veto.

En puits gaussien unidimensionnel, pour $C \geq 3$ et $V \to +\infty$, le taux de MC tend vers une limite strictement comprise entre 0 et 1 pour tous les autres modes de scrutin étudiés ici.

8.2.2 Influence de la parité de V

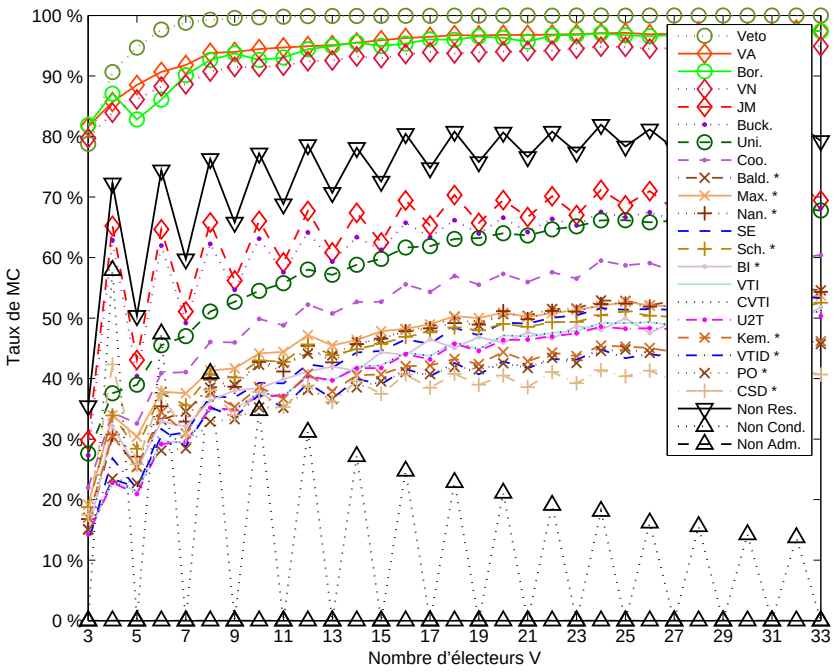


FIGURE 8.4 – Taux de MC en fonction du nombre d'électeurs V. Puits gaussien, $C = 5$, $n = 1$, petites valeurs de V.

En culture sphérique, nous avions vu dans la figure 7.5 que, pour certains modes de scrutin, le taux de manipulabilité présente des oscillations en fonction de la parité de V (et parfois des phénomènes de non-monotonie plus complexes). Afin d'examiner ce qu'il en est en puits gaussien unidimensionnel, la figure 8.4 représente le taux de MC en fonction du nombre d'électeurs V pour toutes les valeurs entre 3 et 33. Le phénomène d'oscillations est ici beaucoup plus général qu'en culture sphérique : il touche tous les modes de scrutin étudiés à part Veto, le vote par assentiment, la méthode de Borda et le vote par notation.

On peut l'expliquer par le fait que, contrairement au cas sphérique, il n'y a essentiellement plus d'effets opposés : les probabilités d'existence d'un vainqueur de Condorcet ou d'un vainqueur de Condorcet résistant, présentant des oscillations qui sont en phase, rendent les valeurs impaires de V moins propices à la manipulation. En particulier, pour V impair, rappelons qu'il y a toujours un vainqueur de Condorcet. Pour le jugement majoritaire et la méthode de Bucklin, les mécaniques propres à ces modes de scrutin ont un effet de parité qui va dans le même sens, comme nous l'avions expliqué dans la section 7.4.2. Quant à la probabilité d'existence d'un candidat Condorcet admissible, elle est constante (et égale à 1),

contrairement au cas sphérique où ses oscillations, en opposition de phase par rapport aux effets précédents, concurrençaient ceux-ci.

En culture *single-peaked*, on remarquera d'ailleurs que la notion de vainqueur de Condorcet résistant est équivalente à celle de favori majoritaire. En effet, si ce n'est pas un candidat extrême, il faut qu'une majorité d'électeurs le préfèrent simultanément aux deux candidats immédiatement à sa gauche à sa droite, donc également à tous les autres candidats. Si c'est un candidat d'extrême-gauche (par exemple), le simple fait d'être vainqueur de Condorcet exige qu'une majorité d'électeurs le préfèrent au candidat immédiatement à sa droite, donc à tout autre candidat.

Pour V pair, on observe aussi que la probabilité d'existence d'un vainqueur de Condorcet augmente avec V, ce qui est facile à expliquer : à nombre de candidats constant, plus les électeurs sont nombreux, moins il est probable que la limite entre deux candidats favoris se situe précisément entre les deux électeurs médians. De façon générale et peu surprenante, les effets de parité diminuent quand le nombre d'électeurs croît, comme en culture sphérique.

8.3 Puits gaussien unidimensionnel : décalage y_0

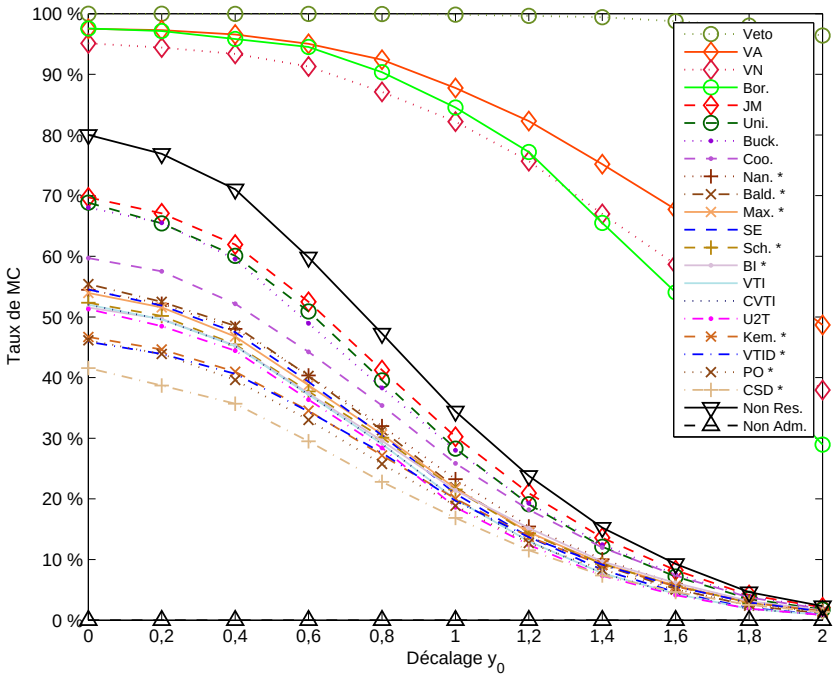


FIGURE 8.5 – Taux de MC en fonction du décalage entre électeurs et candidats. Puits gaussien, $V = 33$, $C = 5$, $n = 1$, $\sigma = 1$.

Dans la figure 8.5, on considère également un puits gaussien unidimensionnel de longueur caractéristique $\sigma = 1$. Mais nous ajoutons un décalage y_0 entre la distribution des candidats et celle des électeurs : la loi normale utilisée pour les candidats n'est plus centrée en 0 mais en y_0. Quand ce décalage tend vers $+\infty$, alors tout candidat est à droite de tout électeur. En particulier, le candidat le plus à gauche est préféré par tous les électeurs [2]. Une telle situation n'est pas

2. Il ne faut voir dans ce paragraphe aucun message caché. Nous aurions pu prendre la convention inverse, mais il fallait choisir.

très réaliste car, dans ce cas, certains candidats auraient intérêt à déplacer leur offre politique pour capter davantage d'électeurs ; tout en étant conscient de cette limitation, nous étudions ce cas afin d'explorer qualitativement l'influence des différents paramètres du modèle.

Nous explorons ce qui se passe pour un décalage variant de 0 à 2. Comme on pouvait s'y attendre, les systèmes de vote ont tendance à être moins manipulables quand le décalage augmente, c'est-à-dire quand les électeurs sont en meilleur accord.

Ce qui est plus intéressant, c'est la similarité avec ce que nous avons observé dans la figure 7.6, qui décrit une culture de VMF où la concentration κ croît. En particulier, certains systèmes de vote sont significativement moins réactifs à l'accord entre les électeurs, au sens où leur taux de manipulabilité décroît moins vite : le vote par assentiment, le vote par notation, la méthode de Borda et particulièrement Veto. Il s'agit des mêmes modes de scrutin que pour la figure 7.6, à part la méthode de Coombs qui se comporte mieux à présent.

D'un point de vue théorique, les arguments exposés en section 7.5 restent valables : comme la culture tend vers un accord parfait entre les électeurs, le taux de manipulabilité de chaque mode de scrutin unanime tend vers 0 et le taux de manipulabilité de Veto tend (lentement) vers $1 - \dfrac{1}{C-1}$.

8.4 Puits gaussien multidimensionnel : nombre de dimensions n

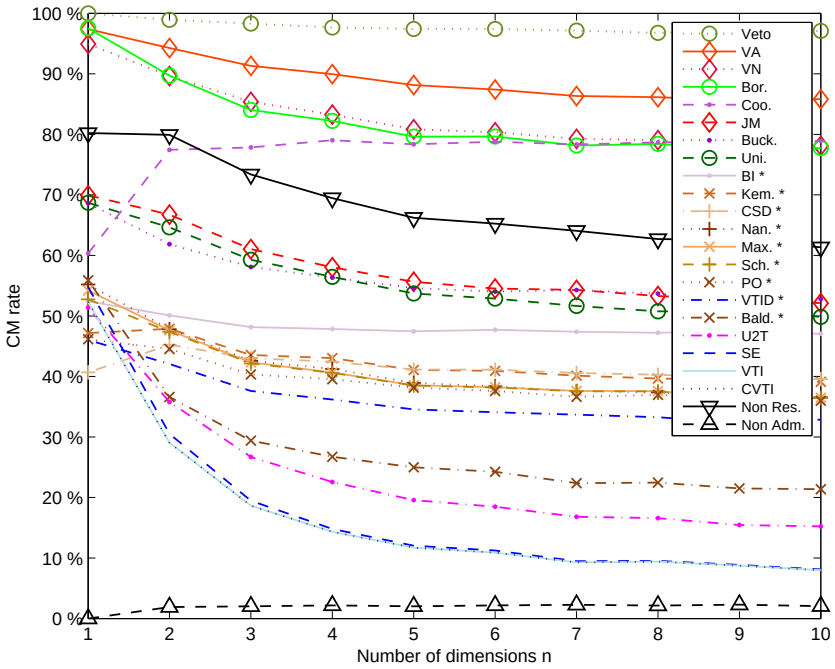


FIGURE 8.6 – Taux de MC en fonction du nombre de dimensions n. Puits gaussien, $V = 33$, $C = 5$, $\sigma = (1, \ldots, 1)$.

Jusqu'à présent, nous avions considéré des spectres politiques à une seule dimension. Dans la figure 8.6, nous généralisons le modèle en considérant un puits gaussien multidimensionnel, avec une dimension n variant de 1 à 10.

Au premier abord, on pourrait penser que si n croît, alors la culture tend vers une culture sphérique, comme pour une culture de VMF avec un nombre infini de pôles (cf. section 7.7). Dans ce cas, les taux de MC devraient augmenter. Cependant, un tel raisonnement serait erroné. En effet, même avec un spectre politique multidimensionnel, un candidat dont la position est proche de l'origine a plus de chances d'être préféré par les électeurs, donc ceux-ci ne sont pas indépendants, même dans la limite $n \to +\infty$: ceci suffit à prouver que la culture limite n'est pas la culture sphérique.

De fait, on observe sur la figure 8.6 que quand n croît, le taux de MC décroît pour la plupart des systèmes de vote (sauf pour la méthode de Coombs, ainsi que BI dont la variation est faible).

À partir de $n = 2$, CVTI, VTI et SE deviennent moins manipulables que les autres systèmes de vote. Par conséquent, il semble que la contre-performance (relative) de ces systèmes de vote est profondément liée au fait d'avoir un spectre politique à une seule dimension.

Ils peuvent même présenter des performances qui sont meilleures que celles observées dans les cultures sphéroïdales du chapitre 7. À titre de comparaison, dans le scénario de référence (figure 7.1), les taux de manipulabilité de ces modes de scrutin étaient de 44 %. Avec le même nombre d'électeurs et de candidats, en puits gaussien avec $n = 10$, les taux trouvés sont de 8 %.

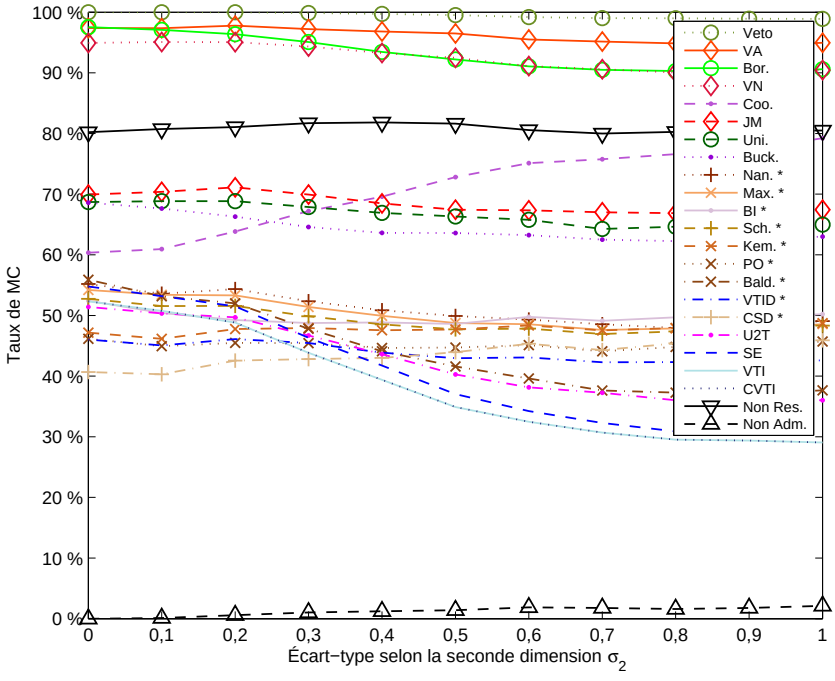


FIGURE 8.7 – Taux de MC en fonction de σ_2. Puits gaussien, $V = 33$, $C = 5$, $\sigma_1 = 1$.

Afin de compléter l'observation précédente, on explore dans la figure 8.7 la transition d'une culture unidimensionnelle vers une culture bidimensionnelle. Pour cela, on considère un puits gaussien à deux dimensions de longueurs caractéristiques $(1, \sigma_2)$. Pour $\sigma_2 = 0$, la culture est unidimensionnelle. Pour $\sigma_2 = 1$, on a une culture « carrée » : les deux dimensions ont la même importance.

Il est inutile de poursuivre la figure pour $\sigma_2 \in [1, +\infty]$ car on trouverait l'exact symétrique des courbes présentées. En effet, à changement d'unité de longueur et

à inversion des axes du spectre politique près, la culture $\sigma = (1, \sigma_2)$ est équivalente à la culture $\sigma' = (1, \frac{1}{\sigma_2})$.

Ces courbes semblent interpoler naturellement ce qui se passe entre les dimensions 1 et 2 dans la figure 8.6 : les taux de manipulabilité décroissent. Comme dans la figure 8.6, seul celui de la méthode de Coombs est clairement croissant. Pour $\sigma_2 \geq 0{,}4$ environ, CVTI, VTI et SE deviennent moins manipulables que les autres modes de scrutin, comme dans les cultures sphéroïdales. De même, pour $\sigma_2 \geq 0{,}5$ environ, l'uninominal à deux tours possède une manipulabilité intermédiaire entre ces trois modes de scrutin et les autres.

8.5 Comparaison avec une boîte euclidienne

On peut se demander si les phénomènes constatés en puits gaussien sont une conséquence qualitative d'une culture en spectre politique unidimensionnel ou s'ils dépendent de la distribution gaussienne utilisée pour les candidats et les électeurs.

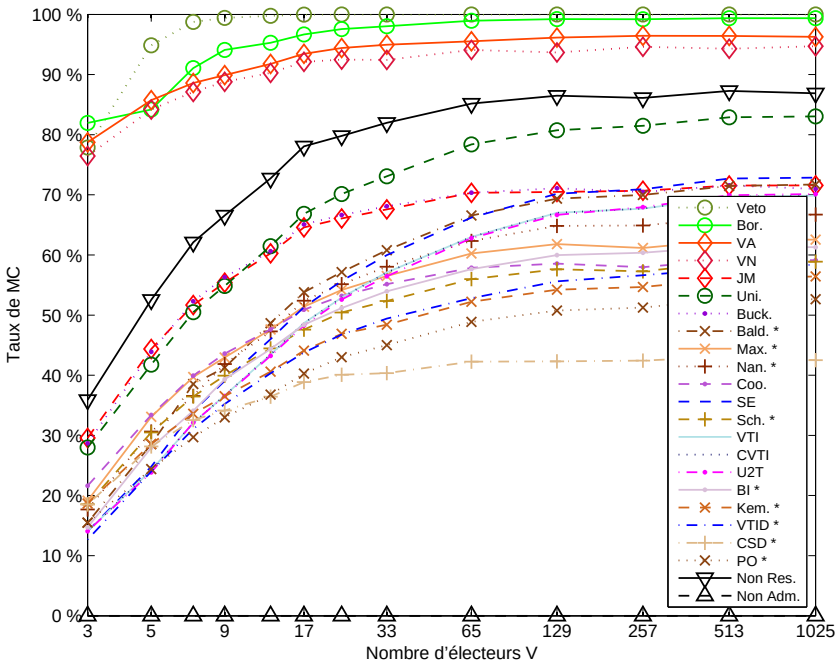


FIGURE 8.8 – Taux de MC en fonction du nombre d'électeurs V. Boîte euclidienne, $C = 5$, $n = 1$, grandes valeurs impaires de V.

À titre d'exemple, la figure 8.8 présente les taux de manipulabilité en boîte euclidienne unidimensionnelle, pour des valeurs impaires du nombre d'électeurs : au lieu d'utiliser une loi normale, les positions des électeurs et des candidats sont tirés uniformément sur un segment $[-1, 1]$.

On constate une grande similarité avec la figure 8.3, ce qui semble indiquer que nos constatations ont une certaine portée générale sur les cultures en spectre politique unidimensionnel. La différence principale concerne les performances relatives de U2T et VTI (ou CVTI), qui sont encore plus proches qu'en puits gaussien.

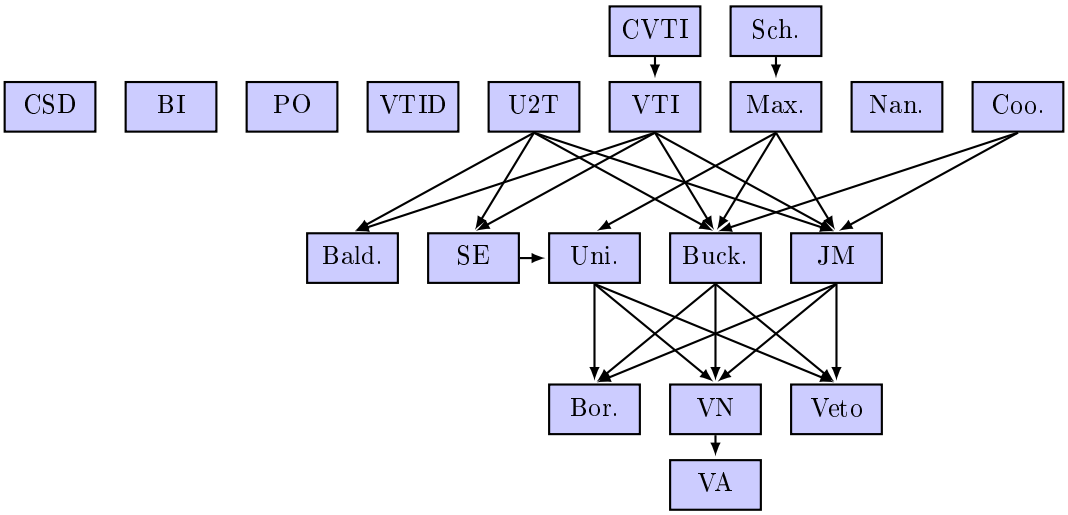


FIGURE 8.9 – Méta-analyse de MC dans les cultures unidimensionnelles étudiées.

8.6 Méta-analyse en culture unidimensionnelle

8.6.1 Méta-analyse de MC en culture unidimensionnelle

De même que la figure 7.11 pour les culture sphéroïdales, la figure 8.9 présente la comparaison de MC entre les modes de scrutin en culture de spectre politique unidimensionnel, que ce soit en puits gaussien ou en boîte euclidienne. Pour l'instant, nous nous concentrerons sur ce cas car nous avons vu en section 8.4 qu'en culture multidimensionnelle, les comportements observés sont relativement similaires à ceux observés dans le chapitre 7. Dans la section 8.7, nous ferons la synthèse de toutes les cultures étudiées, y compris celles du chapitre précédent.

Sur le graphe de la figure 8.9, on voit qu'en culture unidimensionnelle, l'uninominal à deux tours devient compétitif avec CVTI et VTI, et domine même SE dans les cultures que nous avons étudiées. En ce qui concerne CSD, BI, PO, VTID et les méthodes de Schulze et Nanson, les algorithmes approchés ne permettent pas toujours de conclure mais ces modes de scrutin présentent des résultats prometteurs. De nouveau, la méthode de Schulze et Maximin présentent des performances très proches (l'écart est au maximum de 2 %). La méthode de Baldwin est plus manipulable que l'uninominal à deux tours.

L'uninominal à deux tours, VTI, le scrutin exhaustif (lui-même plus manipulable que les deux précédents), Maximin (lui-même plus manipulable que la méthode de Schulze) et la méthode de Coombs servent de référence de manipulabilité pour comparer aux deux groupes suivants, par ordre croissant de manipulabilité :

— l'uninominal, la méthode de Bucklin et le jugement majoritaire ;
— la méthode de Borda, le vote par notation, Veto et le vote par assentiment (qui est lui-même plus manipulable que le vote par notation).

Comme en culture sphéroïdale, ces quatre derniers modes de scrutin ont souvent un taux de manipulabilité supérieur à la proportion de configuration non résistantes, ce qui les rend alors plus manipulables que tout mode de scrutin vérifiant le critère de Condorcet. La méthode de Coombs, qui semble mieux se comporter en spectre politique unidimensionnel qu'en culture sphéroïdale, n'a maintenant plus ce type de propriété. Ce n'est pas une évidence *a priori* : même si nous avons

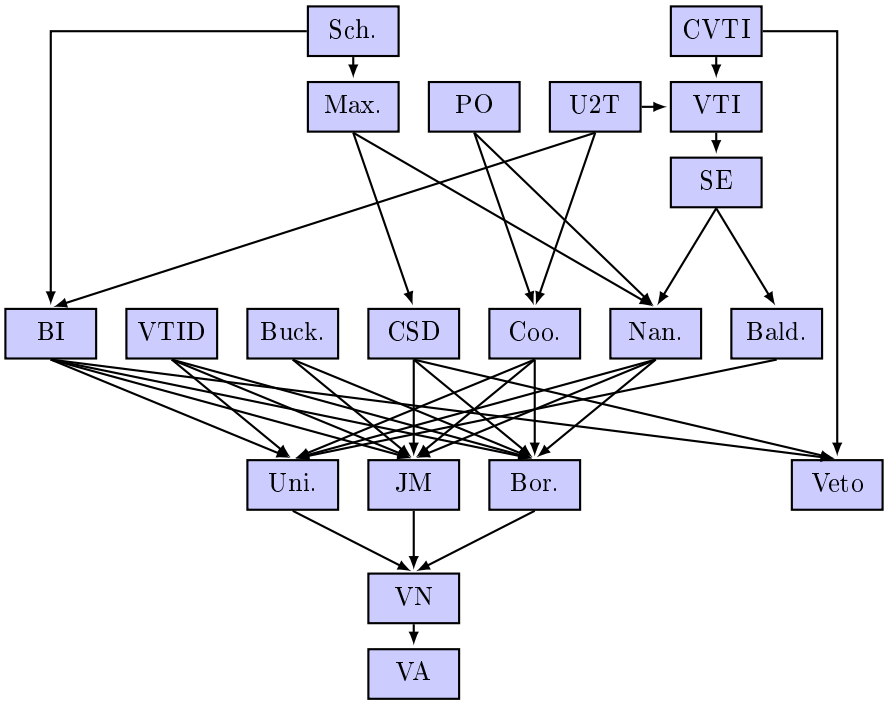


FIGURE 8.10 – Méta-analyse de MT dans les cultures unidimensionnelles étudiées

rappelé qu'elle vérifie le critère de Condorcet pour des préférences *single-peaked*, en particulier celles obtenues en vote sincère dans ce modèle, elle ne le vérifie pas nécessairement pour les configurations obtenues à l'issue d'une manipulation.

De manière générale, certains modes de scrutin Condorcet, en particulier la méthode de Schulze pour laquelle l'incertitude d'origine algorithmique est limitée, semblent mieux se comporter en puits gaussien unidimensionnel qu'en culture sphérique. Intuitivement, on peut en comprendre la raison : en culture sphérique, la probabilité d'existence d'un candidat Condorcet-admissible est relativement faible et on sait qu'elle tend même vers 0 pour V impair et $C \to \infty$ (Gehrlein, 2006). En revanche, en culture *single-peaked* et en particulier en puits gaussien unidimensionnelle, cette probabilité (égale à celle qu'il y ait un vainqueur de Condorcet pour V impair) est égale à 1. Un mode de scrutin non Condorcet part donc avec un handicap majeur : pour V impair, à chaque fois que le vainqueur est un candidat non Condorcet, la situation est manipulable.

8.6.2 Méta-analyse de MT en culture unidimensionnelle

Comme nous l'avions fait pour les cultures sphéroïdales, nous présentons en figure 8.10 un graphe similaire pour la MT. Comme il n'y a, alors, pas d'incertitude algorithmique, on peut conclure dans tous les cas. Ainsi, une absence d'arête (explicite ou sous-entendue par transitivité) entre deux modes de scrutin signifie que l'un est parfois moins MT que l'autre et que c'est parfois l'inverse.

Les modes de scrutin qui n'ont pas d'arête entrante, et sont donc les plus performants du point de vue de MT, sont la méthode de Schulze (avec des performances similaires à Maximin, quoique légèrement meilleures), PO, l'uninominal à deux tours, CVTI, VTID et la méthode de Bucklin.

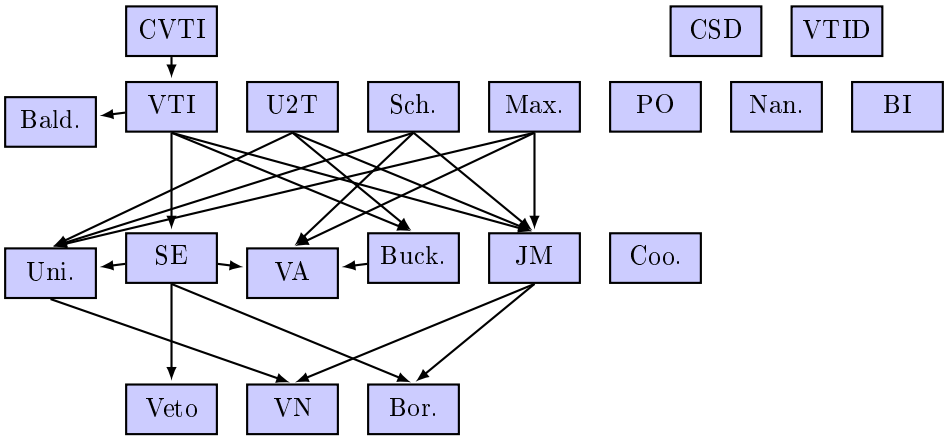


FIGURE 8.11 – Méta-analyse de MC dans l'ensemble des cultures étudiées aux chapitres 7 et 8 (sauf culture sphérique avec $V = 3$).

Les modes de scrutin qui présentent les moins bonnes performances sont l'uninominal, le jugement majoritaire et la méthode de Borda, eux-mêmes moins MT que le vote par notation, lui-même moins MT que le vote par assentiment.

Veto est de nouveau un cas à part : il est dominé par certains modes de scrutin, comme la méthode de Schulze, l'uninominal à deux tours ou CVTI, mais il est non comparable à de nombreux autres.

8.7 Méta-analyse dans l'ensemble des cultures étudiées

8.7.1 Méta-analyse de MC dans l'ensemble des cultures étudiées

Afin de réaliser la synthèse de ce chapitre et du précédent, on peut se demander si certains modes de scrutin sont moins manipulables que d'autres dans l'ensemble des cultures étudiées dans les chapitres 7 et 8. C'est l'objet du graphe de la figure 8.11, qui utilise toutes les courbes précédentes, excepté la culture sphérique pour $V = 3$, pour les raisons exposées en section 7.9.

Les modes de scrutin qui ne possèdent pas d'arête entrante sont CVTI, l'uninominal à deux tours, la méthode de Schulze et Maximin (qui présentent toujours des performances très similaires), la méthode de Coombs, CSD, VTID, PO, la méthode de Nanson et BI. En ce qui concerne ces cinq derniers, l'algorithme approché fournit essentiellement une borne inférieure de manipulabilité, ce qui les rend essentiellement incomparables sur l'ensemble des expériences. Nous compléterons, comme d'habitude, ces résultats par une méta-analyse de la MT.

À l'opposé, nous avons plusieurs fois remarqué les performances très médiocres de la méthode de Borda, du vote par notation et du vote par assentiment, dont le taux de manipulabilité est souvent supérieur à la proportion de configurations non résistantes. Dans les cas concernés, cela signifie que ces modes de scrutin sont plus manipulables que tout mode de scrutin vérifiant le critère de Condorcet.

Parmi les modes de scrutin dont l'utilisation est répandue en pratique, on remarquera que l'uninominal est dominé par des modes de scrutin comme CVTI,

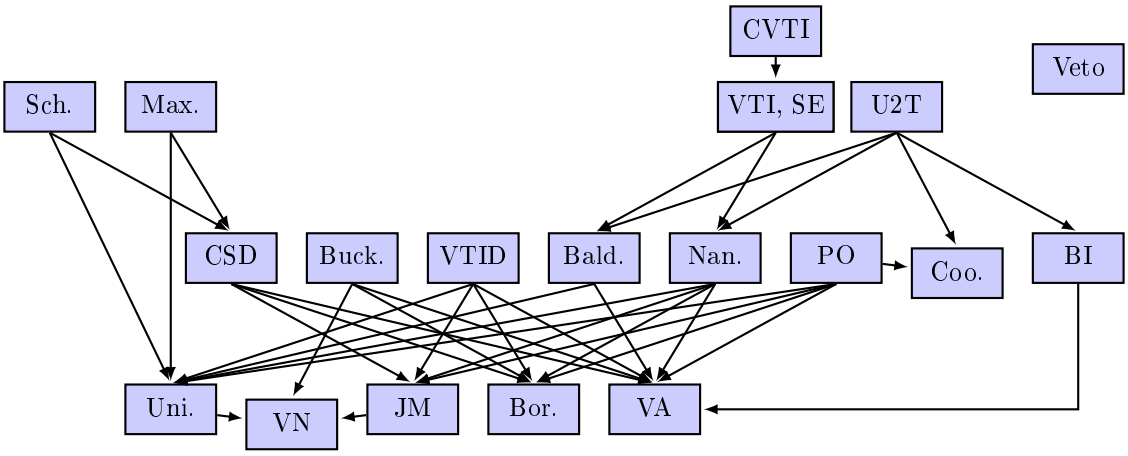


FIGURE 8.12 – Méta-analyse de MT dans l'ensemble des cultures étudiées aux chapitres 7 et 8 (sauf culture sphérique avec $V = 3$).

VTI, SE, l'uninominal à deux tours, la méthode de Schulze ou Maximin. Du point de vue de la manipulabilité, ceci semble plaider pour une utilisation limitée de l'uninominal dans des élections réelles et son remplacement par un de ces modes de scrutin.

8.7.2 Méta-analyse de MT dans l'ensemble des cultures étudiées

De même, la figure 8.12 permet de comparer la MT des modes de scrutin pour toutes les cultures de ce chapitre et du précédent (sauf la culture sphérique avec $V = 3$). Une arête d'un mode de scrutin vers un autre signifie donc que le premier est préférable du point de vue de la MT. De nouveau, on constate que les modes de scrutin présentant les meilleures performances sont CVTI, l'uninominal à deux tours, la méthode de Schulze, Maximin, la méthode de Bucklin. VTID n'a pas d'arête entrante non plus mais son écart de manipulabilité par rapport à l'uninominal à deux tours ne lui est jamais favorable par une amplitude importante dans les cultures étudiées (elle est toujours inférieure à 2 %), alors que nous avons vu que U2T est parfois beaucoup moins MT que VTID.

Tous les autres modes de scrutin, hormis Veto, sont dominés par au moins l'un de ces modes de scrutin. Veto constitue une composante connexe à lui tout seul, confirmant son caractère non comparable avec les autres modes de scrutin étudiés, amplifié par sa résistance à la manipulation triviale.

Enfin, on notera les performances médiocres des modes de scrutin suivants : la méthode de Coombs, l'uninominal, le vote par notation, le jugement majoritaire, la méthode de Borda et le vote par assentiment.

Cependant, comme nous le discutions en introduction de cette thèse, le vote par assentiment a l'avantage de proposer une stratégie assez naturelle en pratique, la *Leader rule* (Laslier, 2009), qui permet d'atteindre des équilibres avec un échange d'information limité. À l'inverse, la complexité d'établir un bulletin stratégique en VTI semble, expérimentalement, décourager les électeurs d'entreprendre un tel calcul et les inciter à voter sincèrement (Van der Straeten *et al.*, 2010), ce qu'on peut voir comme un avantage ou un inconvénient selon le point de vue.

Chapitre 9

Analyses de données expérimentales

Dans les chapitres précédents, nous avons étudié la manipulabilité des divers modes de scrutin dans des cultures artificielles sphéroïdales (chapitre 7) ou basées sur un spectre politique (chapitre 8).

Une conclusion importante de ces chapitres était la faible manipulabilité de CVTI, VTI et SE, sauf dans les cultures de spectre politique unidimensionnel où certains modes de scrutin, par exemple l'uninominal à deux tours et la méthode de Schulze, présentent des performances intéressantes. Il est donc naturel de se demander ce qu'il en est dans le monde réel, et c'est l'objet du présent chapitre.

Nous nous basons sur un corpus de 168 expériences issues de différents contextes. Afin de pouvoir utiliser l'algorithme exact pour VTI tout en conservant un temps de calcul raisonnable, nous nous limitons à des élections comportant 3 à 14 candidats. La méthode de Kemeny sera exclue des analyses de ce chapitre en raison du temps de calcul nécessaire pour déterminer le vainqueur.

9.1 Présentation des données expérimentales

9.1.1 Expériences réalisées

Les trois jeux de données suivants ont été obtenus selon des modalités relativement similaires, avec notre participation directe pour l'établissement des modalités de vote et la collecte des bulletins.

Algotel Lors de la conférence Algotel 2012 [1], le comité de programme a présélectionné 5 papiers, nommés ici arbitrairement A, B, C, D et E, pour l'élection du meilleur papier de la conférence. Pour les départager, il a été demandé à chaque participant de voter par écrit en donnant des notes entre 0 et 10 évaluant la qualité des articles et leur présentation. Il était possible de ne noter qu'une partie des papiers, l'absence de note valant 0. Sur 72 participants à la conférence, il y a eu 57 suffrages exprimés, 1 vote blanc et 2 votes nuls.

Les participants avaient été avertis que leurs bulletins seraient testés sur plusieurs modes de scrutin, mais ils ne savaient pas lequel serait finalement utilisé pour désigner le papier vainqueur. En cas de désaccord entre les modes de scrutin, nous aurions eu la possibilité de récompenser tous les papiers désignés par

1. http://algotel2012.ens-lyon.fr/

Note	Appréciation
5 étoiles	Culte !
4 étoiles	Franchement bien
3 étoiles	Pas mal
2 étoiles	Bof, sans plus
1 étoile	Vraiment pas aimé !

TABLE 9.1 – Échelle de notes et d'appréciations du site www.bdtheque.com.

au moins un mode de scrutin testé. Cependant, en pratique, on verra que tous les modes de scrutin désignaient le même vainqueur, donc la question ne s'est pas posée.

Les conclusions de cette expérience, qui sont en partie reprises dans ce chapitre, sont également disponibles dans l'article de Durand *et al.* (2014a).

Bordeaux Lors de la journée de l'École Doctorale de Mathématiques et d'Informatique de Bordeaux [2] en novembre 2014, un vote a été organisé pour élire le meilleur poster parmi 11 posters proposés par des doctorants en dernière année. Les modalités étaient similaires à Algotel 2012, mais les notes autorisées allaient de 0 à 20. Il y a eu 86 suffrages exprimés pour autant de participants.

Paris VII En avril 2015, un sondage interne a été organisé pour orienter le choix du nom du nouveau laboratoire d'informatique de l'université Paris VII–Paris Diderot [3]. 10 noms possibles étaient proposés. Cette fois, les électeurs n'attribuaient pas des notes mais des appréciations : *Bien*, *Plutôt bien*, *Ni bien ni mal*, *Plutôt mal* ou *Mal*. Il y a eu 95 suffrages exprimés pour autant de participants.

Nous remercions les organisateurs de ces événements d'avoir rendu ces expériences possibles.

9.1.2 Site internet www.bdtheque.com

Le site www.bdtheque.com est un site collaboratif consacré à la bande dessinée. Les internautes y sont invités à noter les séries de bandes dessinées selon l'échelle présentée dans le tableau 9.1, qui est, à la fois, une échelle de notes et une échelle d'appréciations.

En juin 2012, le webmestre du site, Alix Bergeret, que nous remercions, a eu l'amabilité de nous transmettre la base de données du site, à condition de l'utiliser sous une clause d'anonymat des internautes.

Nous en avons tiré 12 expériences de la façon suivante. Pour chaque entier $C \in [\![3, 14]\!]$, nous souhaitions choisir C séries candidates et sélectionner l'intersection des internautes les ayant toutes notées. En outre, notre but était d'avoir un nombre d'électeurs significatif voire de le maximiser si possible. Or ce problème, connu sous le nom de *Maximum Subset Intersection Problem*, a été récemment démontré $\mathcal{NP}$-complet par Xavier (2012).

En guise d'approximation, nous avons utilisé un algorithme glouton. On part de $C = 0$ (c'est-à-dire $\mathcal{C} = \varnothing$) et on initialise l'ensemble $\mathcal{V}$ à l'ensemble des internautes. À chaque incrémentation de C, on sélectionne la série ayant le plus

2. http://www.math.u-bordeaux1.fr/ED/ecole_doctorale/
3. http://www.univ-paris-diderot.fr/

C	3	4	5	6	7	8	9	10	11	12	13	14
V	33	24	21	19	18	15	14	14	13	13	12	12

TABLE 9.2 – Bdtheque : nombres de candidats et d'électeurs

de notes parmi les internautes encore présents dans $\mathcal{V}$ et on l'ajoute à l'ensemble des candidats $\mathcal{C}$, puis on élimine de $\mathcal{V}$ les internautes n'ayant pas noté la dernière série ajoutée. La table 9.2 montre la correspondance entre le nombre de séries candidates et le nombre d'internautes obtenu par cet algorithme [4].

La particularité de ce jeu de données est qu'au moment d'attribuer chaque note, l'internaute n'est pas placé dans un contexte d'élection : en effet, le but de l'opération n'est pas de sélectionner collectivement une unique option parmi un ensemble d'options candidates, comme c'est le cas quand il s'agit d'attribuer un prix ou d'élire une personne à un poste. De plus, cette évaluation n'a pas d'enjeu ou de conséquence pratique sur la vie des électeurs, contrairement à une élection politique par exemple. Enfin, il y a une façon de limiter et dans une certaine mesure de contrôler le vote tactique : en effet, la politique du site exige que toute note soit accompagnée d'un commentaire suffisamment étoffé pour la justifier. Si un internaute souhaite attribuer une note artificiellement basse ou haute à une série à des fins purement tactiques, il doit donc payer un coût cognitif et un coût en temps pour le faire. Pour toutes ces raisons, on peut espérer — mais on ne peut qu'espérer — que ces données sont relativement peu entachées de vote tactique.

9.1.3 Jugement de Paris

En 1976, 11 experts (9 français et 2 américains) se réunirent pour deux dégustations à l'aveugle : tout d'abord, 10 vins blancs de cépage Chardonnay, puis 10 vins rouges à base de Cabernet Sauvignon. Chacune de ces deux dégustations mena à une notation par les experts et à un palmarès.

Les notes allaient de 0 à 20, avec demi-point autorisé. Les experts savaient à l'avance que le vote par notation serait utilisé, ce qui est une différence avec les expériences que nous avons mentionnées ci-dessus. En effet, pour Algotel, Bordeaux et Paris VII, on avait informé les électeurs que leurs bulletins « seraient testés sur plusieurs modes de scrutin », sans davantage de précision. Et pour bdtheque, comme nous l'avons dit, il ne s'agit pas vraiment d'une situation d'élection.

À la surprise générale, dans chacune des deux catégories, ce fut un vin californien qui remporta la première place, malgré un jury en grande majorité français. Cette victoire amena à de nombreux commentaires dans la presse sur les modalités du vote et eut un impact important sur la réputation et le développement des vins américains.

Les données que nous utilisons pour le Jugement de Paris proviennent de la page `www.liquidasset.com/lindley.htm`.

4. Pour le lecteur curieux : les séries obtenues sont, dans l'ordre, *XIII*, *Lanfeust De Troy*, *Blacksad*, *La Quête de l'Oiseau du Temps*, *Universal War One*, *Le Troisième Testament*, *Largo Winch*, *De Cape et de Crocs*, *Les aventures de Tintin*, *Thorgal*, *Astérix*, *Peter Pan*, *Sillage* et *Lanfeust Des Étoiles*.

9.1.4 Base de données PrefLib

PrefLib (http://www.preflib.org/) est une base de données collaborative qui collecte des données de préférences collectives afin de les mettre à la disposition des spécialistes du choix social (Mattei et Walsh, 2013). À la différence des expériences présentées ci-dessus, il s'agit exclusivement de données ordinales : il peut s'agir d'ordres faibles (avec égalités) ou stricts (sans égalité), complets ou incomplets. Quand les ordres sont incomplets, nous supposons que tous les candidats non classés sont placés derrière les candidats classés [5].

Pour la présente étude, nous considérons *a priori* tous les fichiers d'élection de PrefLib, avec les exceptions suivantes.

— Afin de pouvoir utiliser l'algorithme exact de VTI tout en conservant un temps de calcul raisonnable, nous excluons les élections avec strictement plus de 14 candidats.
— Nous excluons les élections dont les préférences dérivent de préférences cardinales (notes ou approbations) puisque celles-ci ne sont pas fournies dans PrefLib. En effet, nous serions amenés à extrapoler arbitrairement des notes à partir des préférences ordinales, comme nous le verrons en section 9.2, et celles-ci ne seraient *a priori* pas conformes à l'expérience originale.
— Nous excluons les élections obtenues par échantillon aléatoire sur de grandes données (PrefLib 4, 11 et 15), considérant que la réalisation particulière de ce tirage aléatoire est propre à leurs auteurs.

Dans PrefLib, les fichiers sont organisés en jeux de données qui regroupent des élections tenues dans un contexte similaire et issues d'une même source. Par exemple, le jeu de données PrefLib 1 regroupe des élections politiques tenues à Dublin en 2002.

Dans cette étude, nous utilisons 151 expériences issues du monde politique (PrefLib 1, 5, 8, 16 à 23), du monde professionnel et associatif (PrefLib 2, 7, 9 et 12) ou d'expériences de cognition (PrefLib 24 et 25). Le jeu de données PrefLib 1 est une donation de Jeffrey O'Neill, qui gère le site http://www.openstv.org/. PrefLib 5, 8 et 16 à 23 proviennent de ce même site. PrefLib 2 provient de http://www.debian.org/vote/. PrefLib 7, 9 et 12 sont respectivement des donations de Nicolaus Tideman, Piotr Faliszewski et Carleton Coffrin. PrefLib 24 et 25 proviennent de Mao *et al.* (2013). Pour plus d'informations sur ces jeux de données, nous encourageons le lecteur à consulter le site http://www.preflib.org/.

La table 9.3 présente une vue d'ensemble de toutes les expériences utilisées dans ce chapitre. Au total, nous nous basons sur 168 expériences, de 10 à presque 300 000 électeurs et de 3 à 14 candidats.

9.2 Méthodologie

Considérons une expérience donnée, par exemple l'élection du meilleur papier Algotel. Comme l'illustre la figure 9.1, par rapport aux bulletins bruts, nous ajoutons aux notes un bruit aléatoire uniforme d'amplitude négligeable par rapport

5. Quand on n'a accès qu'à des préférences incomplètes, une autre possibilité étudiée par Konczak et Lang (2005) consiste à considérer l'ensemble des complétions possibles par des ordres complets. On peut, alors, déterminer l'ensemble des vainqueurs de Condorcet possibles, l'ensemble des vainqueurs possibles par un mode de scrutin donné, etc. Cependant, les auteurs montrent qu'on s'expose alors à de difficiles problèmes de complexité algorithmique, avec toutefois d'importantes exceptions comme la détermination des vainqueurs possibles d'une RPS ou des vainqueurs de Condorcet possibles.

Données	Expériences	V	C	Bulletins
Algotel	1	57	5	Notes 0–10
Bordeaux	1	86	11	Notes 0–20
Paris VII	1	95	10	Appréciations
Bdtheque	12	12–33	3–14	Notes 1–5
Jugement de Paris	2	11	10	Notes 0–20 *
PrefLib 1 : Dublin	3	29 988–64 081	9–14	OSI
PrefLib 2 : Debian	8	143–504	4–9	OSI
PrefLib 5 : Burlington	2	8 980–9 788	6	OFI
PrefLib 7 : ERS	75	32–3 419	3–14	OSI
PrefLib 8 : Glasgow	21	5 199–12 744	8–13	OSI
PrefLib 9 : AGH Course	2	146–153	7–9	OSC
PrefLib 12 : T-shirt	1	30	11	OSC
PrefLib 16 : Aspen	2	2 487–2 528	5–11	OFI
PrefLib 17 : Berkley	1	4 189	4	OFI
PrefLib 18 : Minneapolis	2	32 086–36 655	7–9	OSI
PrefLib 19 : Oakland	7	11 358–145 443	4–11	OFI
PrefLib 20 : Pierce	4	40 031–299 664	4–7	OFI
PrefLib 21 : San Francisco	11	24 180–184 046	4–10	OFI
PrefLib 22 : San Leandro	3	22 539–25 564	4–7	OFI
PrefLib 23 : Takoma	1	204	4	OFI
PrefLib 24 : MT Dots	4	794–800	4	OSC
PrefLib 25 : MT Puzzle	4	793–797	4	OSC
Total	168	10–299 664	3–14	

TABLE 9.3 – Récapitulatif des expériences analysées. OSC : ordres stricts complets. OSI : ordres stricts incomplets. OFI : ordres faibles incomplets. * Demi-points autorisés.

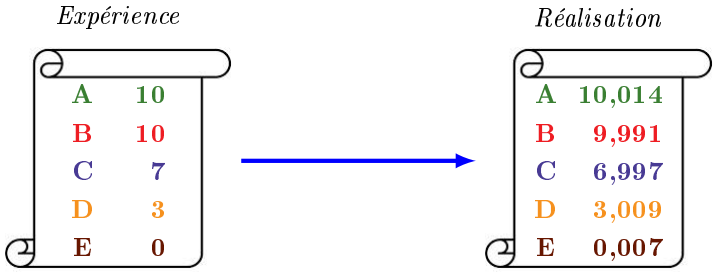


FIGURE 9.1 – Ajout d'un bruit aléatoire aux bulletins

aux écarts entre les notes. Une fois ce bruit ajouté, nous obtenons une configuration pour la population entière, que nous appelons une *réalisation*. Ainsi, si un électeur plaçait un candidat strictement devant un autre dans son bulletin original (par exemple D devant E dans la figure 9.1), cet ordre reste toujours respecté car l'amplitude du bruit est négligeable. En revanche, si un électeur mettait plusieurs candidats à égalité (par exemple A et B dans la figure 9.1), ceux-ci sont placés dans un ordre aléatoire après l'ajout du bruit.

Pour chaque expérience, nous tirons aléatoirement plusieurs réalisations. L'objectif est double : évaluer la cohérence des résultats sur un espace rendu plus riche que l'unique expérience initiale et casser les égalités pour simplifier l'analyse des modes de scrutin basés sur les ordres de préférence. Ainsi, quand nous mentionnons, par exemple, le taux de MC de l'uninominal pour l'expérience Algotel, il s'agit du taux de MC de l'uninominal pour une culture qui consiste à tirer une configuration aléatoire dans un petit voisinage de l'expérience Algotel.

Pour les données purement ordinales de chaque expérience PrefLib, nous effectuons une étape de pré-traitement supplémentaire : avant toute chose, les préférences ordinales sont converties en préférences cardinales en utilisant les scores de Borda. On notera que ce choix particulier n'a d'importance que pour le vote par assentiment et le vote par notation. En effet, pour les modes de scrutin ordinaux, notre technique revient de toute façon à conserver les préférences strictes des électeurs et à départager leurs indifférences de manière impartiale, c'est-à-dire symétrique par rapport aux candidats. Quant au jugement majoritaire, seul compte l'ordre topologique sur l'espace des notes ou des appréciations ; les valeurs numériques potentielles n'ont pas d'importance.

Pour l'expérience Paris VII, une étape de pré-traitement est également ajoutée. Chacune des cinq appréciations possibles est convertie en valeur entière comprise entre -2 et $+2$. Là encore, ce choix particulier n'a d'importance que pour le vote par assentiment et le vote par notation.

Dans ce chapitre, nous travaillons à deux niveaux :
- Nous utilisons l'élection du meilleur papier Algotel comme *exemple récurrent*. Quand nous analysons cet exemple, nous utilisons 10 000 réalisations.
- Nous menons également des *méta-analyses* sur l'ensemble des expériences : dans ce cas, nous nous basons sur 100 réalisations pour chacune des 168 expériences.

Dans les deux cas, l'aspect aléatoire du bruit induit une incertitude statistique de l'ordre du pourcent sur les taux mesurés ($\sqrt{1/10\,000}$ ou $\sqrt{1/16\,800}$).

Pour les méta-analyses, nous donnerons, d'une part, des histogrammes représentant des moyennes sur les 168 expériences réalisées. Ces résultats ne doivent pas être interprétés comme des vérités quantitatives précises mais plutôt comme des indications qualitatives : en effet, accorder le même poids aux 168 expériences considérées est un choix arbitraire en soi. En particulier, cela donne un poids relativement important (45 %) aux expériences du jeu de données PrefLib 7. À l'inverse, donner un poids égal à chaque jeu de données PrefLib conduirait à donner autant d'importance aux jeux de données qui ne contiennent qu'une expérience, comme PrefLib 12 ou PrefLib 17, qu'à PrefLib 7 dont on analyse 75 élections. Il n'y a donc pas de solution parfaite.

D'autre part, nous établirons aussi des graphes de méta-analyse (voir section 7.9). Ils indiqueront qu'un mode de scrutin est moins manipulable qu'un autre dans la totalité des expériences. ce qui est indépendant de toute pondération sur celles-ci. On verra alors que des tendances similaires se dégagent sur l'ensemble des expériences.

	A	B	C	D	E	Score de Borda	Têtes de bulletin
A	—	**35,5**	**36,0**	**36,5**	**39,0**	147,0	16,5
B	21,5	—	**29,5**	**33,0**	**33,0**	117,0	14,7
C	21,0	27,5	—	**32,0**	**31,5**	111,9	9,8
D	20,5	24,0	25,0	—	**29,5**	99,0	10,3
E	18,0	24,0	25,5	27,5	—	95,1	5,7

TABLE 9.4 – Algotel : matrice majoritaire moyenne, score de Borda et nombre de têtes de bulletin.

Enfin, il est impossible de déterminer *a posteriori* si un bulletin correspond à une opinion sincère ou à une stratégie plus élaborée. Nous sommes donc obligés de supposer que les participants ont joué le jeu et que tous les bulletins (bruités) sont sincères. Sans faire cette hypothèse, il est très difficile d'analyser quantitativement l'impact des manipulations. Ceci dit, comme nous l'avons noté, certains jeux de données sont plus vraisemblablement exempts de vote tactique, en particulier les expériences bdtheque.

9.3 Résultats « sincères » et notions Condorcet

Pour l'expérience Algotel, la table 9.4 donne la matrice des duels, le score de Borda et le nombre de bulletins où chaque candidat est classé en tête, en moyenne observée sur les bulletins bruités. Dans la matrice des duels, les victoires sont représentées en caractères gras. On observe, en particulier, que A remporte tous ses duels en moyenne : il est donc vainqueur de Condorcet moyen. En fait, on a même observé une propriété plus forte : A est vainqueur de Condorcet dans toutes les réalisations que nous avons testées.

En ce qui concerne les autres candidats, A est suivi de B qui ne perd que contre A, puis C qui ne perd que contre les deux premiers, et ainsi de suite. On a donc en moyenne un *ordre strict total de Condorcet* $A \succ B \succ C \succ D \succ E$. En pratique, cet ordre strict total a été observé dans 99 % des réalisations.

Il est donc tentant de dire que A devrait être élu, et c'est ce qui se passe dans la quasi-totalité des cas. La principale exception est le mode de scrutin uninominal, qui élit B dans 18% des cas : comme B est classé en tête presque autant de fois que A (table 9.4), le bruit aléatoire est susceptible de faire varier le résultat. De manière beaucoup plus marginale, les autres exceptions concernent les modes de scrutin Bucklin, uninominal à deux tours et VTI, qui peuvent élire B ou C avec une probabilité très faible (moins de 1‰ des réalisations testées).

Ces résultats sont relativement inattendus : comme c'était la première fois que nous menions une expérience de ce type, nous pensions initialement qu'il y aurait plus de diversité dans les vainqueurs possibles [6]. Ils montrent que si les bulletins sont sincères, A semble être un vainqueur canonique qui atteste un choix clair des participants.

6. Par exemple, en culture impartiale avec 3 candidats, Merlin *et al.* (2000) ont montré que la probabilité qu'une certaine collection de modes de scrutin classiques choisissent le même vainqueur était d'environ 50 %. Puisque que l'hypothèse $C = 3$ semble plutôt favorable de ce point de vue, la différence entre ce résultat théorique et l'observation montre à quel point la culture impartiale doit avant tout être vue comme un pire cas.

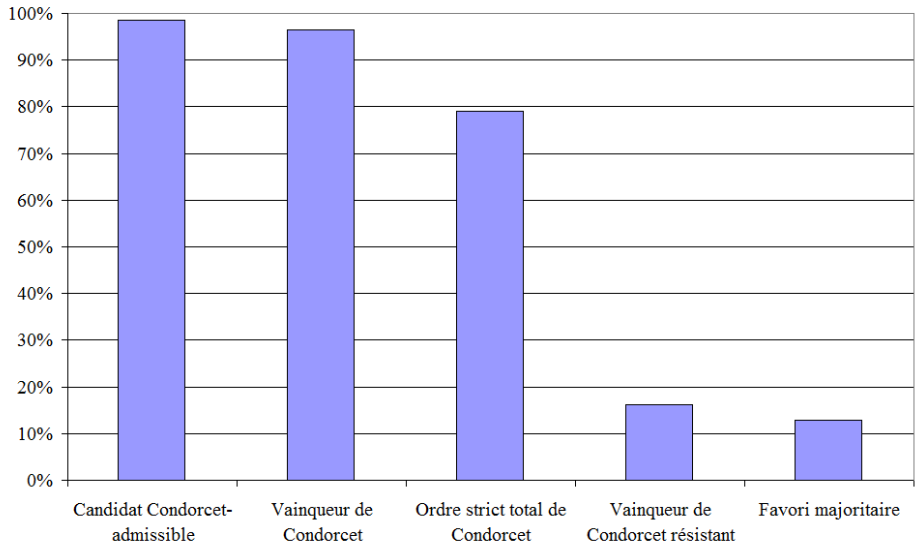


FIGURE 9.2 – Méta-analyse : taux d'apparition de certaines structures dans les préférences de la population.

On peut donc se demander si ces phénomènes sont exceptionnels. Pour y répondre, passons à la méta-analyse : pour rappel, il s'agit d'étudier les 168 expériences, chacune avec 100 réalisations. Comme le montre la figure 9.2, il y a, en fait, presque toujours au moins un candidat Condorcet-admissible (99 %) et il est très fréquent d'avoir un vainqueur de Condorcet (96 %). Ces résultats confirment et étendent ceux de Tideman (2006), basés sur des données qui sont dorénavant regroupées dans le jeu de données PrefLib 7 et qui sont donc incluses dans notre propre étude[7].

Nous avons aussi évalué le taux d'apparition de phénomènes plus forts (au sens où chacun d'entre eux implique l'existence d'un vainqueur de Condorcet). Ainsi, dans 79 % des réalisations, il y a un ordre strict total de Condorcet, ce qui en fait un phénomène assez fréquent malgré le caractère très structurant de cette propriété pour la population[8]. Le taux d'existence d'un vainqueur de Condorcet

7. Les résultats de Tideman sont réalisés avec une méthodologie différente, sans ajout de bruit. Nos résultats, menant à la même conclusion qualitative, sont donc complémentaires. À titre de comparaison, en excluant les données de PrefLib 7 afin d'obtenir une étude indépendante, les taux d'existence obtenus sont similaires : 98 % pour un candidat Condorcet-admissible et 95 % pour un vainqueur de Condorcet.

8. La probabilité d'existence de paradoxe d'intransitivité (absence de vainqueur de Condorcet, absence d'ordre total de Condorcet) a fait l'objet d'intenses études par la théorie ou par des simulations informatiques. Outre l'ouvrage de référence de Gehrlein (2006), on peut notamment citer Ruben (1954); Campbell et Tullock (1965); Garman et Kamien (1968); Niemi et Weisberg (1968); DeMeyer et Plott (1970); Pomeranz et Weil Jr (1970); Tullock et Campbell (1970); Gehrlein et Fishburn (1976); Fishburn *et al.* (1979); Gehrlein (1981); Jones *et al.* (1995); Gehrlein (1999); Maassen et Bezembinder (2002); Merlin *et al.* (2002); Tsetlin *et al.* (2003). À part dans les modèles single-peaked, ces travaux concluent généralement que l'absence de vainqueur de Condorcet arrive avec une probabilité relativement importante. L'étude de ces probabilités dans des expériences réelles, sans être totalement nouvelle, se développe actuellement grâce à un meilleur accès à des jeux de données et semble au contraire indiquer qu'il est très fréquent d'avoir un vainqueur de Condorcet, comme notre étude le confirme. Regenwetter *et al.* (2006) montrent également qu'avec des préférences réalistes, il est assez peu probable d'avoir un cycle dans la relation de préférence majoritaire.

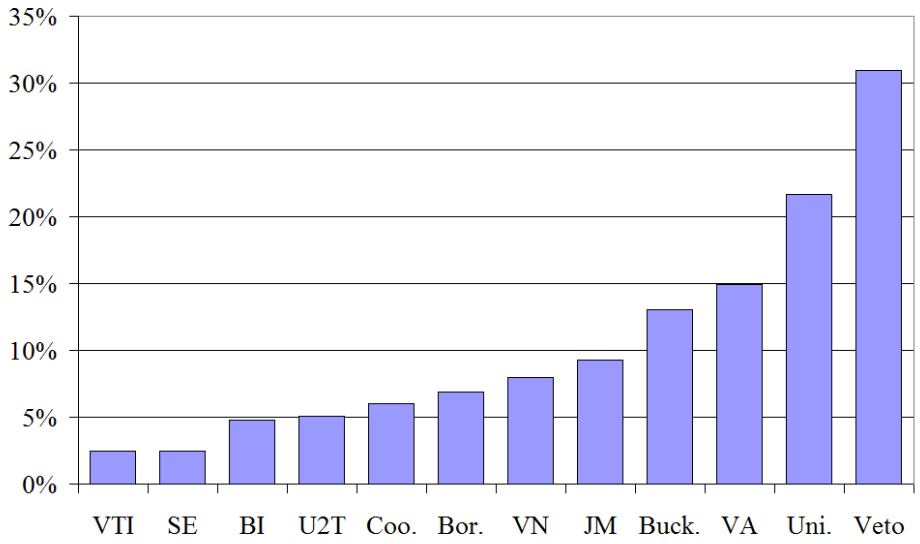


FIGURE 9.3 – Méta-analyse : taux de violation du critère de Condorcet

résistant, quant à lui, est non négligeable (16 %), ce qui assure que dans tous les cas concernés, tout mode de scrutin vérifiant le critère de Condorcet est non manipulable. Enfin, il y a un favori majoritaire dans 13 % des cas : la plupart des vainqueurs de Condorcet résistants observés en pratique sont donc des favoris majoritaires, mais pas tous.

Comme dans l'expérience Algotel, les partisans du critère de Condorcet estimeront, par conséquent, qu'il existe la plupart du temps un vainqueur canonique qui devrait être élu. On peut donc se demander dans quelle mesure chaque mode de scrutin est susceptible de violer le critère de Condorcet. La figure 9.3 montre la probabilité qu'un vainqueur de Condorcet existe mais qu'il ne soit pas élu par un mode de scrutin donné (en vote « sincère »). Comme ce taux de violation du critère de Condorcet est nécessairement nul pour les modes de scrutin Condorcet, ceux-ci ne sont pas représentés.

Parmi les modes de scrutin non Condorcet, on notera que VTI (équivalent à SE en vote sincère) présente les meilleures performances : sur le jeu de données utilisé, il viole le critère de Condorcet dans seulement 2 % des cas. Il est suivi par BI et l'uninominal à deux tours (5 %). Les moins bonnes performances sont atteintes par le vote par assentiment (15 %), l'uninominal (22 %) et Veto (31 %).

Pour tous les modes de scrutin présentés ici sauf Veto, chaque cas de violation du critère de Condorcet est aussi un cas de manipulabilité triviale en faveur du vainqueur de Condorcet : ainsi, on sait par exemple, d'ores et déjà, que le taux de MT d'uninominal est au moins de 22 %, ne serait-ce que pour cette raison.

9.4 Manipulation par coalition

9.4.1 Taux de MC moyens

Revenons à l'expérience Algotel. Sur la figure 9.4, on peut lire le taux de CM pour chaque mode de scrutin. Le graphique se lit ainsi : par exemple, l'uninominal

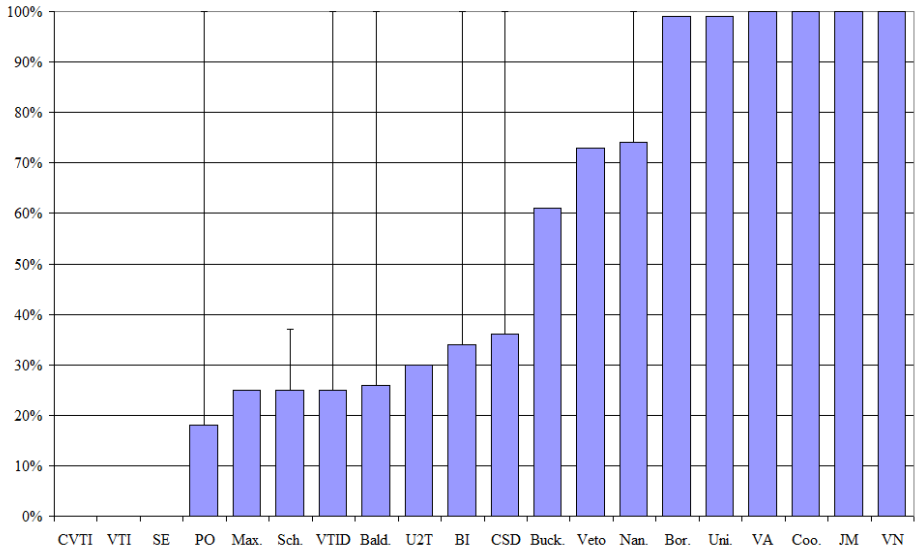


FIGURE 9.4 – Algotel : Taux de manipulabilité par coalition (MC)

à deux tours (U2T) est manipulable dans 30 % des réalisations. Comme nous le disions précédemment, il s'agit du taux de MC de ce mode de scrutin dans une culture qui consiste à tirer une configuration aléatoirement sur un petit voisinage autour de l'expérience réelle dans l'espace des préférences.

Comme certains algorithmes utilisés sont approchés, une barre d'erreur est indiquée pour les modes de scrutin concernés : par exemple, le taux de CM de la méthode de Schulze est compris entre 25 % et 37 %. Pour d'autres modes de scrutin, comme Maximin et la méthode de Borda, l'algorithme utilisé n'est pas exact en théorie[9], mais on constate que la barre d'erreur est de taille nulle : dans le cas particulier de cette expérience, l'algorithme a permis de décider la manipulabilité pour toutes les réalisations tirées.

Pour cette expérience Algotel, on voit que CVTI, VTI et SE sont meilleurs que tous les autres modes de scrutin, avec un taux de CM inférieur à 1 % (l'incertitude d'origine statistique). À l'inverse, les modes de scrutin présentant les pires performances avérées sont la méthode de Borda, l'uninominal, le vote par assentiment, la méthode de Coombs, le jugement majoritaire et le vote par notation, avec des taux de MC proches de 100 %.

Passons à présent à la méta-analyse. Les taux de MC moyens sont représentés dans la figure 9.5. Pour les modes de scrutin sans algorithme spécifique dédié (Baldwin, VTID, etc.), les bornes supérieures trouvées par SVAMP sont significativement inférieures à 100 %. Pour l'expliquer, rappelons qu'il y a un vainqueur de Condorcet résistant dans 16 % des réalisations : dans tous ces cas, SVAMP sait que les modes de scrutin Condorcet sont non manipulables (section 6.3.4). Pour la méthode de Bucklin itérée (BI), SVAMP utilise un résultat similaire avec le favori majoritaire. Remarquons que les algorithmes utilisés pour CVTI, Maximin et la méthode de Borda, bien qu'approchés en théorie, échouent à décider la manipulabilité dans moins de 1 % des situations étudiées ici.

9. Pour ces modes de scrutin, on a utilisé l'option `fast` de SVAMP.

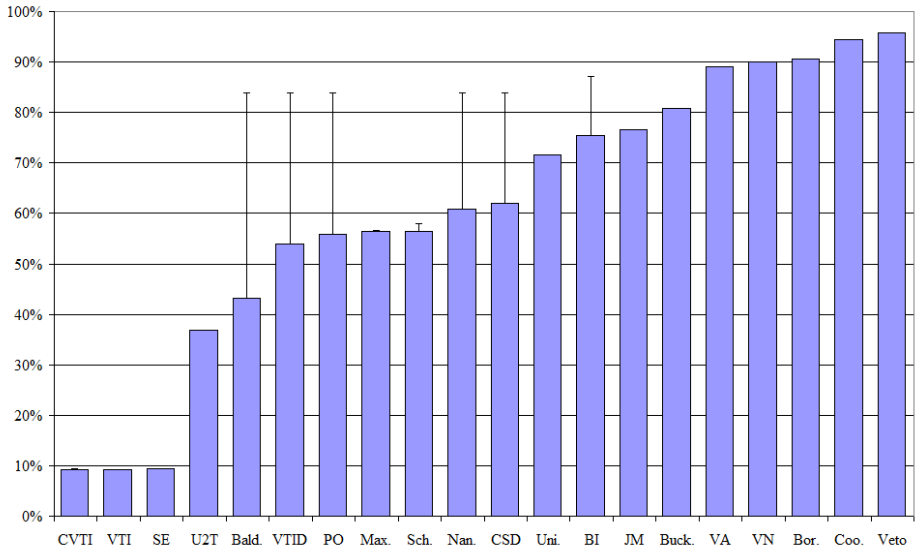


FIGURE 9.5 – Méta-analyse : Taux de manipulabilité par coalition (MC)

En ce qui concerne le taux de MC moyen, on est de nouveau frappé par la bonne performance de CVTI, VTI et SE (9%). Même si on sait, par la théorie, que CVTI est strictement moins manipulable que VTI (corolaire 2.21 des théorèmes de condorcification), qui est lui-même strictement moins manipulable que SE (section 1.4.2), la différence est inférieure à 1 % sur ce corpus d'expériences [10]. Ce résultat est similaire avec ce que nous avions constaté sur des populations artificielles dans les chapitres 7 et 8, hormis le fait que dans les cultures unidimensionnelles, on constatait plus souvent une légère différence entre VTI et SE.

Tous les autres modes de scrutin sont nettement plus manipulables, même l'uninominal à deux tours avec un taux de MC de 37 %. Qualitativement, cette différence est plus proche de ce qu'on constatait pour les cultures sphéroïdales ou de spectre politique multidimensionnel que pour les cultures unidimensionnelles.

Les plus manipulables en moyenne, avec des taux proches de 90 % voire supérieurs, sont le vote par assentiment, le vote par notation, la méthode de Borda, la méthode de Coombs et Veto. Ces résultats sont similaires avec ceux des chapitres 7 et 8, excepté la méthode de Coombs dont le mauvais comportement est plus proche des résultats obtenus dans les cultures sphéroïdales ou de spectre politique multidimensionnel que dans les cultures unidimensionnelles.

On sait que, dans tous les modes de scrutin considérés ici sauf Veto, si le gagnant n'est pas Condorcet-admissible (soit qu'il n'y en ait pas, soit que le mode de scrutin ne l'élise pas), la situation est forcément manipulable (lemme 2.7). Cela a amené certains auteurs (par exemple Lepelley et Valognes, 2003) à penser que les modes de scrutin Condorcet pourraient être moins manipulables. Si cela semble globalement se vérifier, il faut nuancer. En effet, il y une grande disparité : alors que CVTI a un taux de MC de 9%, Maximin et la méthode de Schulze ont un taux de CM compris entre 56 % et 58 %, ce qui les place significativement

10. Sur les réalisations utilisées, SVAMP trouve un taux de manipulabilité moyen de 9,32 % pour SE, 9,28 % pour IRV et compris entre 9,20 % et 9,28 % pour CVTI. Les différences sont donc détectables mais leurs amplitudes, inférieures à l'incertitude statistique, ne doivent être considérées que comme indicatives.

derrière VTI (9 %), qui ne respecte pourtant pas le critère de Condorcet. Sur ce point également, les résultats obtenus sont plus proches de ceux des cultures sphéroïdales et de spectre politique multidimensionnel que de ceux des cultures unidimensionnelles.

9.4.2 MC comparée des systèmes de vote

	CVTI	VTI	SE	U2T	Max.	Sch.	Bald.	VTID	PO	Nan.	Uni.	CSD	JM	BI	Buck.	VA	Bor.	Veto	VN	Coo.
CVTI	—	0 (4)	**0 (4)**	**40**	**62**	**62**	**49 (81)**	**60 (81)**	**61 (81)**	**63 (81)**	**70**	**64 (81)**	**76**	**78 (83)**	**80**	**86**	**88**	**92**	**86**	**90**
VTI	0	—	**3**	**40**	**62**	**62**	**49 (81)**	**60 (81)**	**61 (81)**	**63 (81)**	**70**	**64 (81)**	**76**	**78 (83)**	**80**	**86**	**88**	**92**	**86**	**90**
SE	0	0	—	**40**	**62**	**62**	**49 (81)**	**60 (81)**	**61 (81)**	**63 (81)**	**70**	**64 (81)**	**76**	**78 (83)**	**80**	**86**	**88**	**92**	**86**	**90**
U2T	0	0	0	—	**35**	**35 (37)**	**18 (55)**	**30 (55)**	**33 (55)**	**37 (55)**	**45**	**39 (55)**	**50**	**51 (57)**	**55**	**60**	**62**	**67**	**60**	**64**
Max.	0	0	0	5	—	0 (14)	1 (38)	0 (38)	1 (38)	**14 (38)**	**27**	**19 (38)**	**33**	**31 (40)**	**38**	**43**	**45**	**49**	**43**	**47**
Sch.	0	0	0	3 (5)	0 (4)	—	1 (38)	0 (38)	0 (38)	**10 (38)**	**26 (27)**	**13 (38)**	**32 (33)**	**27 (40)**	**33 (38)**	**42 (43)**	**44 (45)**	**49**	**42 (43)**	**46 (47)**
Bald.	0	0	0	1 (10)	0 (30)	0 (31)	—	0 (50)	0 (50)	0 (50)	2 (39)	0 (50)	4 (45)	0 (52)	3 (49)	**16 (55)**	**18 (57)**	**23 (61)**	**16 (55)**	**20 (59)**
VTID	0	0	0	1 (8)	0 (17)	0 (18)	0 (40)	—	0 (40)	0 (40)	2 (29)	0 (40)	4 (35)	0 (42)	3 (39)	**16 (45)**	**18 (46)**	**23 (51)**	**16 (45)**	**20 (49)**
PO	0	0	0	1 (7)	0 (11)	0 (15)	0 (38)	0 (38)	—	0 (38)	2 (27)	0 (38)	4 (33)	0 (40)	3 (38)	**16 (43)**	**18 (45)**	**23 (49)**	**16 (43)**	**20 (47)**
Nan.	0	0	0	1 (2)	0 (5)	0 (10)	0 (34)	0 (34)	0 (34)	—	2 (23)	0 (34)	4 (28)	0 (36)	3 (33)	**16 (39)**	**18 (40)**	**23 (45)**	**16 (39)**	**20 (43)**
Uni.	0	0	0	0	0	0 (2)	0 (23)	0 (23)	0 (23)	0 (23)	—	1 (23)	**15**	13 (24)	**15**	**26**	**29**	**34**	**27**	**32**
CSD	0	0	0	1	0	0 (6)	0 (33)	0 (33)	0 (33)	0 (33)	2 (20)	—	4 (27)	0 (35)	3 (31)	**16 (38)**	**17 (40)**	**23 (45)**	**16 (38)**	**20 (42)**
JM	0	0	0	1	0	0	0 (17)	0 (17)	0 (17)	1 (17)	4	1 (17)	—	9 (19)	**11**	**23**	**23**	**27**	**23**	**27**
BI	0	0	0	0 (3)	0 (3)	0 (7)	0 (21)	0 (21)	0 (21)	0 (21)	0 (16)	0 (21)	2 (17)	—	0 (19)	**10 (26)**	**8 (29)**	**14 (33)**	**10 (26)**	**11 (31)**
Buck.	0	0	0	1	0	0 (3)	0 (13)	0 (13)	0 (13)	1 (13)	10	2 (13)	**11**	1 (15)	—	**17**	**20**	**25**	**18**	**23**
VA	0	0	0	1	0	0	0 (6)	0 (6)	0 (6)	0 (6)	3	0 (6)	1	2 (7)	4	—	**8**	**11**	**5**	**10**
Bor.	0	0	0	1	0	0	0 (1)	0 (1)	0 (1)	0 (1)	1	0 (2)	2	0 (5)	0	5	—	**10**	**4**	**9**
Veto	0	0	0	2	4	4	3 (8)	4 (8)	4 (8)	5 (8)	5	4 (8)	8	5 (8)	5	10	9	—	**10**	**10**
VN	0	0	0	1	0	0	0 (2)	0 (2)	0 (2)	0 (2)	2	0 (2)	1	1 (3)	1	2	**4**	8	—	**6**
Coo.	0	0	0	0	0	0	0 (1)	0 (1)	0 (1)	0 (1)	0	0 (1)	1	0 (3)	1	5	1	7	4	—

TABLE 9.5 – Méta-analyse : Comparaison de la MC par paire de modes de scrutin (en pourcents)

Comme nous l'avons fait dans les méta-analyses des chapitres 7 et 8, nous allons, à présent, comparer les modes de scrutin par paire sur le corpus d'expériences. Comme il est difficile de donner des résultats détaillés pour toutes les expériences, contrairement aux chapitres précédents où des courbes permettaient

de visualiser les résultats comparés des différents modes de scrutin, nous fournissons la table 9.5 qui permet d'établir le graphe de méta-analyse.

Cette table se lit de la façon suivante. Dans la cellule (U2T, Sch.) par exemple, le premier chiffre signifie qu'on est *sûr* qu'U2T a un taux de MC strictement inférieur à celui de la méthode de Schulze dans 35 % des expériences (au sens où la borne supérieure pour le premier est strictement inférieure à la borne inférieure pour le second). Le chiffre indiqué entre parenthèses signifie qu'il est *possible* que ce soit le cas dans 37 % des expériences (au sens où la borne inférieure pour le premier est strictement inférieure à la borne supérieure pour le second).

À l'inverse, dans la cellule (Sch., U2T) de la table, on lit qu'il est sûr que la méthode de Schulze ait un taux de CM inférieur ou égal à celui d'U2T dans 3 % des expériences, et que c'est possible dans 5 % des expériences.

Avec une pondération uniforme sur les 168 expériences, on peut conclure que ce match est gagné pour U2T. En effet, même dans l'hypothèse la plus défavorable pour U2T, celui-ci a un taux de MC strictement inférieur à celui de la méthode de Schulze dans 35 % des expériences ; alors que la méthode de Schulze a un taux strictement inférieur à U2T dans 5 % des expériences. À titre indicatif, de telles victoires basées sur une pondération uniforme sont indiquées en caractères gras dans le tableau. En raison de l'incertitude sur certains modes de scrutin, nous ne sommes pas en mesure de fournir les résultats de tous les matchs et nous complèterons nos résultats par une méta-analyse de la MT.

Comme nous l'avions fait dans les chapitres précédents, nous présentons en figure 9.6 le graphe de méta-analyse, qui se base sur des victoires dans la totalité des expériences : une flèche d'un mode de scrutin f vers un mode de scrutin g signifie qu'il est sûr que g n'a un taux strictement inférieur à celui de f dans aucune expérience. Cette propriété est donc indépendante de la pondération choisie sur les 168 expériences.

La conclusion la plus frappante est que CVTI domine VTI, qui domine SE (ce que nous savions déjà par la théorie), et que ce dernier domine tous les autres modes de scrutin étudiés dans la totalité des expériences. Maximin et la méthode de Schulze, dont les algorithmes approchés permettent une incertitude limitée, fournissent un bon point de comparaison avec les autres modes de scrutin. On constate notamment que les performances les plus médiocres sont atteintes par les méthodes de Coombs et Borda, CSD, le jugement majoritaire, le vote par assentiment et le vote par notation. Les mauvaises performances de la méthode de Borda, du jugement majoritaire, du vote par assentiment et du vote par notation sont similaires à ceux des chapitres 7 et 8. Les points suivants sont plus proches des comportements constatés dans les cultures sphéroïdales ou de spectre politique multidimensionnel : la suprématie de CVTI, VTI et SE ; les performances intermédiaires de Maximin et la méthode de Schulze (alors qu'elles sont assez bonnes en culture unidimensionnelle) ; et les performances médiocres de la méthode de Coombs et de CSD (alors qu'elles sont correctes en culture unidimensionnelle).

9.4.3 MC par candidat

Pour terminer cette étude de la MC, nous allons nous intéresser à la manipulabilité par candidat pour l'expérience Algotel. En figure 9.7, on représente le taux de manipulabilité en faveur de chaque candidat.

Sans nous livrer à une analyse exhaustive, un certain nombre d'observations nous semblent pertinentes sur ce graphique. Tout d'abord, la faible manipulabilité pour A exprime juste le fait que A est quasiment toujours le vainqueur sincère et que ses partisans n'ont rien à faire. Les cas de manipulabilité pour A, notamment

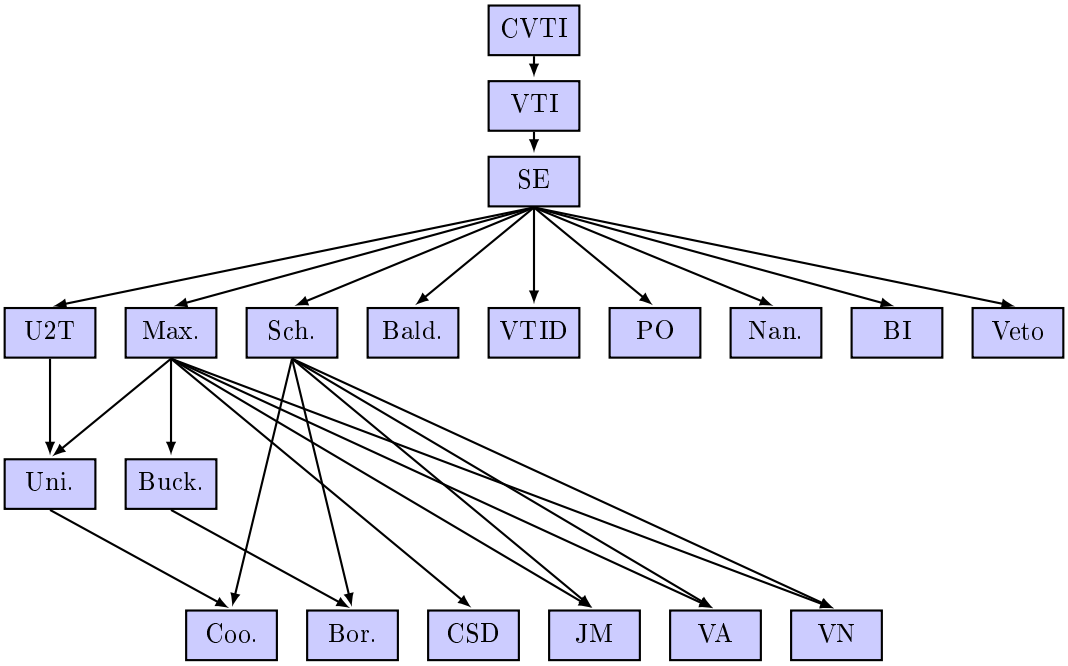


FIGURE 9.6 – Méta-analyse de MC dans les expériences.

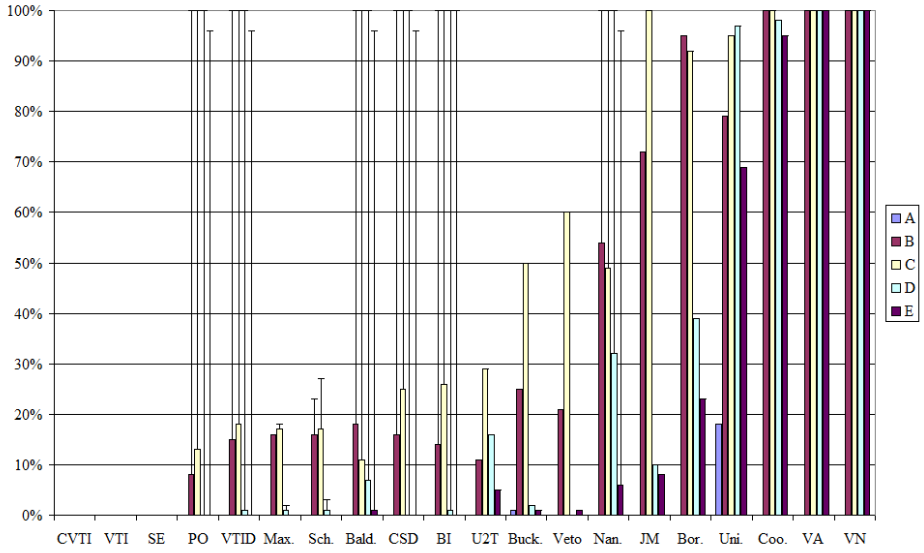


FIGURE 9.7 – Algotel : Taux de MC en faveur de chaque candidat

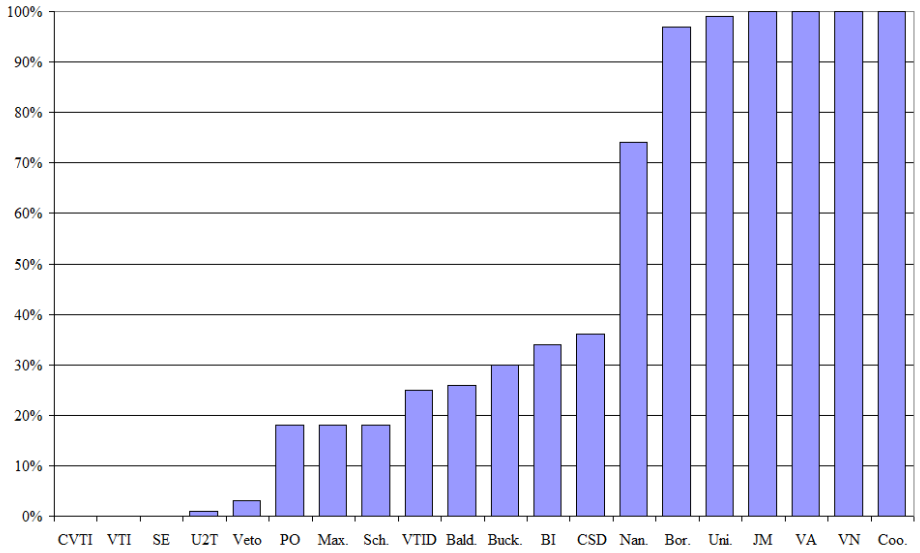


FIGURE 9.8 – Algotel : Taux de manipulabilité triviale (MT)

en uninominal, correspondent aux cas où B, voire C est le vainqueur sincère. Il existe alors toujours une manipulation qui élit A.

Les pires résultats sont obtenus pour deux modes de scrutin relativement naturels, le vote par assentiment et le vote par notation, qui sont manipulables en faveur des quatre rivaux de A dans toutes les réalisations ! Le jugement majoritaire, la méthode de Borda, l'uninominal et la méthode de Coombs présentent également un risque de manipulation fréquente pour une variété de candidats.

9.5 Manipulation triviale

Comme nous l'avons souvent évoqué, la manipulabilité triviale (MT) est un critère intéressant à double titre. D'une part, elle ne conserve que les cas où la manipulation est relativement facile à identifier et à coordonner, ce qui rend son occurrence plus crédible en pratique. D'autre part, sa simplicité algorithmique autorise des calculs exacts pour tous les modes de scrutin étudiés ici et permet donc de compléter la comparaison entre les modes de scrutin basée sur la MC.

9.5.1 Taux de MT moyens

Sur la figure 9.8, on a représenté le taux de MT pour chaque mode de scrutin dans l'expérience Algotel. Comme les conclusions des figures 9.8 et 9.9 sont qualitativement similaires, nous allons directement commenter la méta-analyse de la figure 9.9. Quatre modes de scrutin se distinguent nettement par des taux moyens inférieurs à 10 % : CVTI (3 %), VTI et SE [11] (5 %) et l'uninominal à deux tours (8 %). Ensuite viennent Veto (17 %) et la méthode de Baldwin (43 %), puis un grand nombre de modes de scrutin entre 50 % et 80 %. Enfin, quatre modes de scrutin ont un taux de MT qui avoisine ou dépasse 90 % : le vote par assentiment, le vote par notation, les méthodes de Borda et de Coombs.

11. Rappelons que VTI et SE sont équivalents du point de vue de la manipulation triviale.

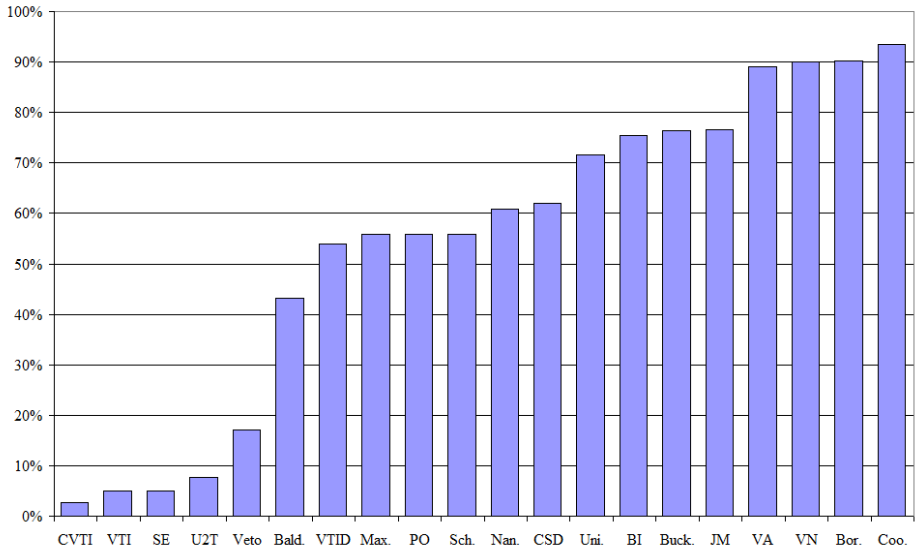


FIGURE 9.9 – Méta-analyse : Taux de manipulabilité triviale (MT)

Comme nous l'avons déjà remarqué et expliqué en section 7.2, Veto est le seul mode de scrutin pour lequel le taux de MT (17%) est très différent du taux de MC (96 %).

9.5.2 MT comparée des systèmes de vote

	CVTI	VTI	SE	U2T	Veto	Bald.	VTID	PO	Max.	Sch.	Nan.	CSD	BI	Uni.	Buck.	JM	VA	Bor.	Coo.	VN
CVTI	—	**14**	**14**	**16**	**38**	**53**	**64**	**65**	**66**	**66**	**67**	**68**	**82**	**74**	**82**	**80**	**90**	**92**	**94**	**90**
VTI	0	—	**0**	**12**	**38**	**51**	**63**	**64**	**64**	**64**	**65**	**67**	**80**	**73**	**80**	**78**	**88**	**90**	**92**	**88**
SE	0	**0**	—	**12**	**38**	**51**	**63**	**64**	**64**	**64**	**65**	**67**	**80**	**73**	**80**	**78**	**88**	**90**	**92**	**88**
U2T	0	1	1	—	**36**	**50**	**61**	**63**	**63**	**63**	**64**	**65**	**79**	**71**	**79**	**77**	**87**	**89**	**91**	**87**
Veto	5	8	8	10	—	**46**	**57**	**57**	**58**	**58**	**60**	**61**	**73**	**68**	**73**	**73**	**80**	**81**	**84**	**80**
Bald.	0	0	0	0	30	—	**25**	**28**	**29**	**29**	**32**	**33**	**46**	**39**	**47**	**45**	**55**	**57**	**59**	**55**
VTID	0	0	0	1	25	5	—	**15**	**16**	**16**	**20**	**21**	**37**	**29**	**38**	**35**	**45**	**46**	**49**	**45**
PO	0	0	0	1	25	1	3	—	**8**	**8**	**18**	**20**	**34**	**27**	**36**	**33**	**43**	**45**	**47**	**43**
Max.	0	0	0	1	24	2	1	5	—	**2**	**19**	**20**	**35**	**28**	**37**	**34**	**44**	**46**	**48**	**44**
Sch.	0	0	0	1	24	2	1	4	0	—	**18**	**19**	**34**	**27**	**36**	**33**	**43**	**45**	**47**	**43**
Nan.	0	0	0	0	23	0	1	0	2	2	—	**14**	**24**	**23**	**30**	**28**	**39**	**40**	**43**	**39**
CSD	0	0	0	0	23	0	0	0	0	0	2	—	**23**	**20**	**27**	**27**	**38**	**40**	**42**	**38**
BI	0	0	0	1	17	1	0	1	1	1	7	8	—	**16**	**17**	**17**	**26**	**29**	**31**	**26**
Uni.	0	0	0	0	20	0	0	0	0	0	0	1	13	—	**13**	**15**	**26**	**29**	**31**	**27**
Buck.	0	0	0	0	17	1	1	1	1	1	2	5	7	11	—	**13**	**21**	**24**	**27**	**23**
JM	0	0	0	0	17	0	0	0	0	0	1	1	9	4	8	—	**23**	**23**	**25**	**23**
VA	0	0	0	0	5	0	0	0	0	0	0	0	2	3	2	1	—	**8**	**10**	**5**
Bor.	0	0	0	0	7	0	0	0	0	0	0	0	0	1	0	2	6	—	**8**	**5**
Coo.	0	0	0	0	7	0	0	0	0	0	0	0	0	0	1	1	8	5	—	**7**
VN	0	0	0	0	5	0	0	0	0	0	0	0	1	2	0	1	2	4	6	—

TABLE 9.6 – Méta-analyse : Comparaison de la MT par paire de modes de scrutin (en pourcents)

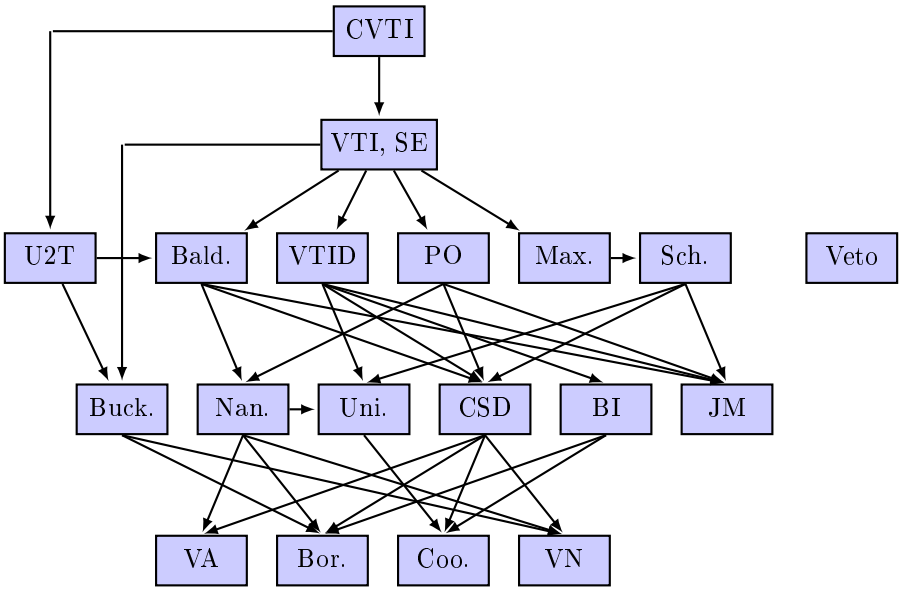


FIGURE 9.10 – Méta-analyse de MT dans les expériences.

De même que pour la MC, nous fournissons la table 9.6 de méta-analyse pour MT. Cette fois, il n'y a plus d'incertitude algorithmique, ce qui facilite la lecture de la table et fournit des résultats plus précis.

En particulier, le mode de scrutin le plus performant, CVTI, n'est jamais devancé par aucun autre mode de scrutin, hormis Veto. VTI et SE présentent des performances similaires, quoique devancés par l'uninominal à deux tours dans 1 % des expériences (ils sont cependant meilleurs dans 12 % des cas). Parmi les autres modes de scrutin, l'uninominal à deux tours est celui qui affiche les meilleurs résultats. Bien qu'il soit strictement moins bon que CVTI (du point de vue de MT) dans 16 % des expériences et strictement moins bon que VTI ou SE dans 12 % des expériences (contre 1 % pour l'inverse), il est strictement meilleur que Veto dans 36 % des expériences (contre 10 % pour l'inverse) et contre tout autre mode de scrutin étudié, il n'est devancé que dans 0 à 1 % des expériences.

Comme précédemment, le graphe de méta-analyse de la figure 9.10 est basé sur des suprématies dans la totalité des expériences et il est donc indépendant de toute pondération. On retiendra notamment que, sur le corpus d'expériences considéré, CVTI est strictement meilleur que tous les autres modes de scrutin étudiés, à l'exception de Veto qui constitue, à lui seul, une composante connexe isolée. En particulier, VTI, SE et U2T sont plus manipulables que CVTI. Ensuite, on peut, à titre indicatif, distinguer trois groupes, par ordre croissant de MT.

1. La méthode de Baldwin, VTID, PO, Maximin et la méthode de Schulze. De nouveau, ces deux derniers présentent des performances très similaires (avec toutefois un léger avantage pour Maximin).
2. Les méthodes de Bucklin et Nanson, l'uninominal, CSD, BI et le jugement majoritaire. Parmi eux, Nanson est moins MT que l'uninominal.
3. Le vote par assentiment, les méthodes de Borda et Coombs et le vote par notation.

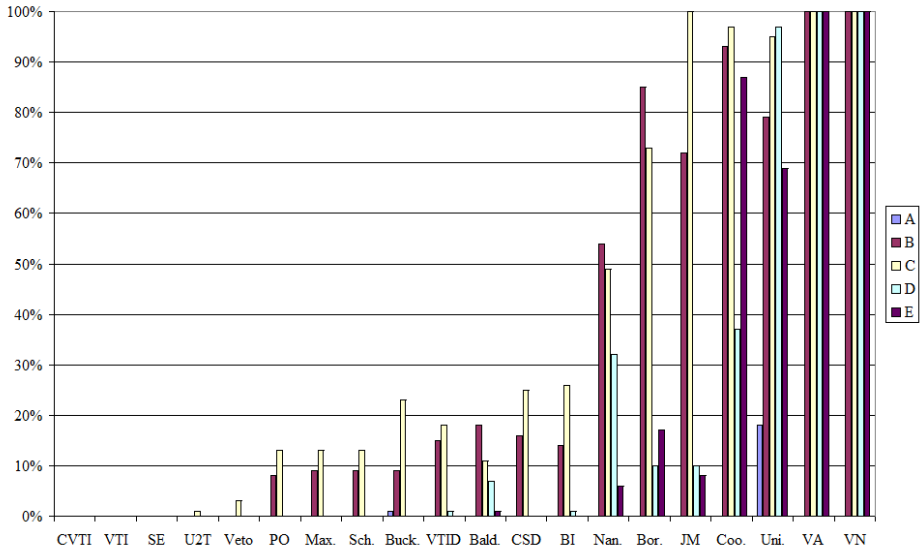


FIGURE 9.11 – Algotel : Taux de MT en faveur de chaque candidat

9.5.3 MT par candidat

Comme pour la MC, nous examinons dans la figure 9.11 la manipulabilité par candidat pour l'expérience Algotel. Globalement, les résultats sont similaires à ceux obtenus pour la MC dans la figure 9.7, en particulier la supériorité de CVTI, VTI et SE. On constate toutefois les différences suivantes.

Alors qu'U2T pouvait être parfois MC pour une variété de candidats, il n'est MT (et très rarement) qu'en faveur du candidat C. Ainsi, les manipulations en uninominal à deux tours demandent, en général, des stratégies relativement élaborées.

Veto est beaucoup moins MT que MC, ce que nous avons déjà remarqué et expliqué, puisque la stratégie triviale est très sous-optimale dans ce mode de scrutin.

Sans surprise, les résultats sont les mêmes en MC et en MT pour le jugement majoritaire, l'uninominal, le vote par assentiment et le vote par notation : pour ces modes de scrutin, la stratégie triviale est optimale au sens où si une manipulation est possible, alors cette stratégie permet de la réaliser. Comme nous l'avons discuté dès l'introduction de cette thèse, ceci peut être vu comme un défaut ou comme une qualité de ces systèmes de vote : le fait que la manipulation soit simple à réaliser permet de rechercher des équilibres avec un échange limité d'information. Cependant, la propriété d'être non manipulable nous semble encore plus intéressante (ce qui est le cas de CVTI, VTI et SE dans l'expérience Algotel, cf. figure 9.7 par exemple), puisqu'elle permet d'atteindre un équilibre avec un échange d'information nul, comme nous l'avons notamment discuté dans le chapitre 3.

9.6 Synthèse des résultats

Dans ce chapitre, nous avons analysé les résultats de 168 expériences issues du monde réel, en ajoutant un bruit aléatoire qui permet d'explorer l'espace des préférences au voisinage de chaque expérience.

Nous avons pu établir qu'il y a presque toujours un candidat Condorcet-admissible, très souvent un vainqueur de Condorcet et que la présence d'un ordre strict total de Condorcet, qui est pourtant une propriété très structurante pour une population, est, en fait, un phénomène relativement fréquent.

Parmi les modes de scrutin non Condorcet, VTI et SE sont ceux qui violent le moins souvent le critère de Condorcet, suivis de BI et de l'uninominal à deux tours. Ceux qui le violent le plus souvent sont le vote par assentiment, l'uninominal et Veto. Malgré ces différences, on constate que même les modes de scrutin non Condorcet élisent l'éventuel vainqueur de Condorcet dans une majorité des cas (de taille très variable en fonction du système utilisé).

Nous avons constaté que la manipulabilité n'est pas qu'un concept théorique, mais une réalité concrète susceptible d'apparaître même dans un contexte où le résultat semble *a priori* évident : en effet, une part importante des modes de scrutin étudiés sont manipulables dans une proportion importante des expériences.

Même si les différents modes de scrutin désignent relativement souvent le même vainqueur en vote « sincère » (le vainqueur de Condorcet quand il existe, même pour les modes de scrutin non Condorcet), le choix du mode de scrutin est donc également crucial pour une autre raison : les taux de MC observés vont de moins de 10 % à plus de 90 % en fonction du système utilisé.

CVTI, VTI et SE sont les moins manipulables. Par la théorie, on sait CVTI est strictement moins manipulable que VTI, qui est lui-même strictement moins manipulable que SE ; mais en pratique, la différence est très faible. Pour des cas pratiques, il est donc très envisageable de préférer VTI à CVTI, car le dépouillement de VTI est plus aisé si on ne dispose pas de moyens informatiques. Parmi les autres modes de scrutin, l'uninominal à deux tours est généralement le moins manipulable. Les modes de scrutin les plus manipulables sont le vote par assentiment, la méthode de Borda, la méthode de Coombs, le vote par notation et Veto.

Les résultats observés sont globalement plus proches de ceux obtenus dans les cultures sphéroïdales du chapitre 7 ou dans les cultures de spectre politique multidimensionnel du chapitre 8 que dans les cultures unidimensionnelles (chapitre 8 également), malgré le fait que notre corpus d'expériences inclut des élections issues du monde politique, où l'hypothèse de *single-peakedness* paraîtrait *a priori*, sinon réaliste, en tout cas une première approximation intéressante de la réalité. Sans valider totalement les modèles sphéroïdaux et les modèles de spectre politique multidimensionnel, cette conclusion encourage, en tout cas, à poursuivre leur étude dans des travaux futurs.

Chapitre 10

Modes de scrutin optimaux

Dans ce chapitre, nous nous intéressons aux système de vote dont le taux de manipulabilité est minimal au sein de l'ensemble de ceux qui vérifient **CMInf** : par commodité de langage, nous dirons qu'un tel système est *optimal*.

Grâce au corollaire 2.23 des théorèmes de condorcification, nous savons que pour chercher un tel mode de scrutin, on peut se restreindre aux systèmes Condorcet. Par le théorème d'optimalité 5.15, nous savons que si la culture est décomposable, en particulier si les électeurs sont indépendants, alors il existe un système optimal et on peut restreindre la recherche aux systèmes qui sont ordinaux et vérifient le critère de Condorcet.

Dans la section 10.1, nous présentons la technique employée. En section 10.1.1, nous définissons le *graphe d'opportunité* d'un espace électoral. Indépendant du système de vote utilisé, celui-ci indique si une configuration est susceptible d'être manipulable vers une autre, en fonction du candidat qui sera choisi comme vainqueur dans les deux configurations par le système de vote qu'on envisage d'utiliser. Dans cette représentation, une culture est simplement représentée par une pondération sur les sommets du graphe.

Dans les sections 10.1.2, 10.1.3 et 10.1.4, nous montrons que le problème de la recherche d'un optimum se simplifie si on se restreint aux systèmes Condorcet, si on suppose que les configurations semi-admissibles sont de probabilité nulle et si on se restreint aux systèmes ordinaux.

Dans la section 10.1.5, nous présentons un algorithme glouton qui permet de rechercher un optimum de façon approchée, mais dont nous verrons qu'il n'est pas exact. Dans la section 10.1.6, nous montrons que la recherche d'un optimum, avec les hypothèses précédentes, peut être ramenée à un problème d'optimisation en nombres entiers, que le logiciel dédié CPLEX nous permettra de traiter pour des valeurs modérées des paramètres.

Dans la section 10.2, nous nous restreignons à la culture impartiale avec un nombre impair d'électeurs, qui vérifie toutes les hypothèses précédentes. Nous discutons les solutions trouvées pour $V = 3$, 5 ou 7 et $C = 3$, ainsi que pour $V = 3$ et $C = 4$ (sections 10.2.1 à 10.2.4).

Dans la section 10.3, nous concluons ce chapitre en comparant le taux de manipulabilité minimal trouvé (au sein de **CMInf**) avec ceux des modes de scrutin classiques.

10.1 Graphe d'opportunité

10.1.1 Définition

Comme d'habitude, nous nous plaçons dans un espace électoral Ω. Le *graphe d'opportunité* que nous allons définir dans cette section est entièrement déterminé par Ω : il exprime une structure de l'espace électoral qui est indépendante du mode de scrutin utilisé par la suite. Dans ce cadre, nous verrons ensuite qu'une culture est simplement représentée par une pondération sur les sommets du graphe.

Pour définir le graphe d'opportunité, nous devons d'abord définir le (w, c)-*pointage*.

Définition 10.1

Soit ω et ψ deux configurations distinctes, w et c deux candidats distincts.

On dit que ω (w, c)-*pointe* vers ψ si et seulement si :

$$\forall v \in \mathrm{Sinc}_\omega(\mathrm{w} \to c), \omega_v = \psi_v.$$

Autrement dit :

$$\forall v \in \mathcal{V}, \omega_v \neq \psi_v \Rightarrow c \, \mathrm{P}_v(\omega_v) \, \mathrm{w}.$$

Si un électeur présente un état différent dans ω et ψ, il préfère c à w dans ω.

Intuitivement, cette propriété signifie que si w est vainqueur dans ω et si c est vainqueur dans ψ, alors ω est manipulable en faveur de c vers ψ : en effet, les électeurs qui préfèrent c à w peuvent changer leur bulletin afin de produire la configuration ψ.

Par exemple, considérons les configurations suivantes.

$$\omega : \begin{array}{c|c|c} a & b & c \\ b & c & b \\ c & a & a \end{array} \qquad \psi : \begin{array}{c|c|c} a & b & c \\ b & c & \mathbf{a} \\ c & a & \mathbf{b} \end{array}$$

La configuration ω (b, c)-pointe vers ψ : en effet, le seul électeur qui change d'état préfère c à b dans ω. Il serait donc néfaste pour un mode de scrutin de désigner comme vainqueur b dans ω et c dans ψ : en effet, comme ω (b, c)-pointe vers ψ, cela impliquerait que la configuration ω soit manipulable.

Dans l'exemple ci-dessus, la configuration ω (a, c)-pointe et (a, b)-pointe également vers ψ. Ce (a, b)-pointage nous donne l'occasion d'insister sur le fait que chaque électeur modifié (l'unique électeur modifié, en l'occurrence) préfère par hypothèse b à a dans la configuration de départ ω, mais pas nécessairement dans celle d'arrivée ψ. Ainsi, si un mode de scrutin désignait a dans ω et b dans ψ, le dernier électeur pourrait manipuler la configuration ω en faveur de b en faisant descendre b dans son ordre de préférence ! Une telle situation n'est pas du tout à exclure : certains modes de scrutin largement répandus, comme l'uninominal à deux tours ou VTI, présentent de tels défauts de monotonie.

Plus généralement, on n'a pas seulement une implication du pointage vers la manipulabilité mais l'équivalence suivante, qui est une simple relecture de la définition de la manipulabilité.

Proposition 10.2

Soit $f : \Omega \to \mathcal{C}$ un SVBE, ω et ψ deux configurations distinctes.

Le système f est manipulable en ω vers ψ si et seulement si ω $(f(\omega), f(\psi))$-pointe vers ψ.

Cette remarque va nous permettre de reformuler les questions de manipulabilité comme des problèmes de graphe.

Définition 10.3
Considérons le multigraphe étiqueté $(\Omega, E, e : E \to \mathcal{C}^2)$ défini de la façon suivante.

— Les sommets du graphe sont les états $\omega \in \Omega$.
— E est l'ensemble des arêtes (*edges*).
— e est une fonction qui, à chaque arête, associe une étiquette qui est un couple de candidats.
— Un sommet ω possède une arête étiquetée (w, c) vers un sommet ψ si et seulement si la configuration ω (w, c)-pointe vers ψ.

Le multigraphe étiqueté obtenu est appelé le *graphe d'opportunité* de l'espace électoral Ω.

Il s'agit d'un multigraphe : il est *a priori* tout à fait possible qu'un sommet ω possède plusieurs arêtes vers un sommet ψ avec des étiquettes différentes, comme nous l'avons déjà vu.

Un SVBE $f : \Omega \to \mathcal{C}$ est alors vu comme une fonction qui, à chaque sommet ω du graphe d'opportunité, associe un élément de $\mathcal{C}$. L'indicatrice de manipulabilité de f en ω, déjà notée $\mathrm{MC}_f(\omega)$, est une fonction booléenne sur les sommets qui vaut 1 en un sommet ω si et seulement s'il existe au moins un sommet ψ vers lequel le sommet ω possède l'arête étiquetée $(f(\omega), f(\psi))$.

Le graphe d'opportunité est une représentation qui permet d'étudier le taux de manipulabilité : si l'on se donne une loi de probabilité π sur l'espace électoral, il suffit de pondérer chaque sommet ω par un poids $\pi(\omega)$. Le taux de manipulabilité de f est alors le poids total des sommets manipulables selon f :

$$\tau_{\mathrm{MC}}(f) = \sum_{\omega \in \Omega} \mathrm{MC}_f(\omega)\pi(\omega).$$

Chercher un mode de scrutin de taux de manipulabilité minimal et respectant une certaine contrainte (par exemple **Cond**) signifie chercher une fonction f qui minimise le poids total des sommets manipulables parmi celles vérifiant cette contrainte.

Le problème se simplifie au moins dans les trois cas suivants, qui permettent de « nettoyer » le graphe, c'est-à-dire d'enlever des arêtes non pertinentes sans altérer sa validité pour la recherche d'un optimum.

— Si la contrainte choisie implique qu'une configuration ω ait pour vainqueur un certain w_0, alors pour $\mathrm{w} \neq \mathrm{w}_0$, il est inutile de considérer les arêtes (w, c) sortantes de ω.
— Si on sait qu'une configuration ω est nécessairement manipulable (à cause de la contrainte choisie), alors il est inutile de considérer les arêtes sortantes de ω.
— Si un événement $A \subset \Omega$ a une probabilité nulle, alors pour toute configuration $\omega \in A$, il est inutile de considérer ses arêtes sortantes : en effet, rendre ω manipulable n'augmente pas le taux de manipulabilité. En revanche, ses arêtes entrantes sont importantes car ω pourrait rendre manipulables certaines configurations ω' de probabilité non nulle qui pointent vers ω.

10.1.2 Restriction aux modes de scrutin Condorcet

Par le corollaire 2.23 des théorèmes de condorcification, on sait qu'il est intéressant de chercher un mode de scrutin qui, au sein de la classe de ceux qui

	ψ Condorcet	ψ semi-Condorcet	ψ non admissible
ω Condorcet (vainqueur w_0)	non	pour $w = w_0$	
ω semi-Condorcet	pour tout w		
ω non admissible	non		

TABLE 10.1 – Recherche d'un mode de scrutin Condorcet optimal : (w, c)-pointages à considérer de ω vers ψ.

vérifient **Cond**, possède une manipulabilité minimale. En effet, un tel mode de scrutin présentera un taux de manipulabilité minimal dans la classe plus large des modes de scrutin qui vérifient **CMInf**.

Pour cette contrainte particulière (**Cond**), les deux premières simplifications vues plus haut deviennent les suivantes.

— Si ω possède un vainqueur de Condorcet w_0, on sait que w_0 est nécessairement élu dans ω donc pour $w \neq w_0$, il est inutile de considérer les arêtes sortantes (w, c).

— Si ω est une configuration non admissible, on sait qu'elle est nécessairement manipulable, donc il est inutile de considérer ses (w, c)-pointages sortants.

De plus, si ω et ψ possèdent un vainqueur de Condorcet, alors ω ne peut pas être manipulable vers ψ, donc il est inutile de considérer les arêtes de ω vers ψ.

La table 10.1 récapitule quelles arêtes doivent être considérées pour rechercher un mode de scrutin Condorcet de manipulabilité minimale.

En résumé, les configurations Condorcet se simplifient parce qu'on connaît déjà le vainqueur ; et les configurations non admissibles, parce qu'on connaît déjà leur manipulabilité (égale à Vrai). On peut, d'ailleurs, remarquer que toute configuration possédant un vainqueur de Condorcet résistant est également un cas simple, puisqu'on connaît aussi sa manipulabilité (égale à Faux) ; en pratique, cela se traduit par le fait qu'une telle configuration n'a pas d'arête sortante, une fois le graphe nettoyé comme indiqué.

Les configurations semi-Condorcet (avec au moins un candidat Condorcet-admissible, mais sans vainqueur de Condorcet) sont les plus délicates : en effet, ni leur vainqueur ni leur manipulabilité ne sont connus *a priori*.

10.1.3 Configurations semi-Condorcet de probabilité nulle

À présent, nous ajoutons l'hypothèse que l'ensemble des configurations semi-Condorcet a une probabilité nulle. C'est le cas, en particulier, si le nombre d'électeurs est impair et si les préférences sont presque sûrement des ordres stricts totaux. Concernant cette dernière hypothèse, il peut en être ainsi soit parce que les autres relations binaires sont considérées comme impossibles dans l'espace électoral considéré, comme dans celui des ordres stricts totaux, soit parce qu'elles sont autorisées mais de probabilité nulle dans la culture étudiée : c'est notamment le cas si les préférences dérivent d'utilités et que la probabilité d'avoir deux utilités égales est nulle, comme dans toutes des cultures des chapitres 7 et 8. Il s'agit donc d'un cas particulier important.

Comme on le remarquait en section 10.1.1, cela implique qu'on peut ignorer les arêtes sortantes des configurations semi-Condorcet, puisque leur manipulabilité n'a aucune influence sur le taux de manipulabilité.

Cependant, *a priori*, une configuration semi-Condorcet ψ pourrait rendre manipulable une configuration ω de probabilité non nulle. Afin de contourner ce

	ψ Condorcet	ψ semi-Condorcet	ψ non admissible
ω Condorcet (vainqueur w_0)	non		pour $w = w_0$
ω semi-Condorcet	non		
ω non admissible	non		

TABLE 10.2 – Recherche d'un mode de scrutin Condorcet optimal : (w, c)-pointages à considérer de ω vers ψ, si les configurations semi-Condorcet ont une probabilité nulle. Sans altérer l'optimalité, on impose que le vainqueur de chaque configuration semi-Condorcet soit un candidat Condorcet-admissible arbitraire.

risque, nous ajoutons la contrainte suivante : pour toute configuration semi-Condorcet, le vainqueur est un candidat Condorcet-admissible arbitraire. Ainsi, une configuration semi-Condorcet ne peut rendre manipulable ni une configuration non admissible (qui est manipulable de toute façon) ni une configuration Condorcet (car le bénéficiaire d'une telle manipulation continuerait d'avoir une défaite contre le vainqueur de Condorcet, donc ne pourrait pas être Condorcet-admissible, lemme 2.6). En imposant cette propriété supplémentaire, on peut aussi ignorer les arêtes entrantes des configurations semi-Condorcet et finalement ignorer totalement ces configurations.

Pour la recherche d'un mode de scrutin de taux de manipulabilité minimal, cette hypothèse n'altère en rien l'optimalité du système trouvé : en effet, pour tout mode de scrutin optimal, les considérations précédentes montrent que si on le modifie en désignant un candidat Condorcet-admissible pour toute configuration semi-Condorcet, alors le système obtenu possède exactement le même taux de manipulabilité.

La table 10.2 résume les arêtes qu'on peut se contenter de considérer avec ces nouvelles hypothèses. Il s'agit maintenant d'un graphe étiqueté bipartite, puisque les arêtes proviennent uniquement des configurations Condorcet pour aller uniquement vers des configurations non admissibles.

Quand ω (w, c)-pointe vers ψ, il est à présent inutile de préciser w, puisque dans le graphe nettoyé, le candidat considéré est nécessairement le vainqueur de Condorcet dans ω. Par conséquent, on dira simplement que ω c-pointe vers ψ (ce qui sous-entend que ω est une configuration Condorcet et ψ une configuration).

La figure 10.1 montre une représentation d'un graphe d'opportunité de ce type. Il s'agit d'un exemple simplifié, qui ne correspond pas forcément à un espace électoral particulier. Chaque configuration ω^i est Condorcet, chaque configuration ψ^j est non-Condorcet. Les configurations semi-Condorcet ou résistantes ne sont pas représentées car elles n'ont aucune importance pour le problème à résoudre. Afin de représenter graphiquement l'arête étiquetée c d'une configuration ω vers une configuration ψ, on munit ψ^j d'une « prise » notée $\psi^j.c$ et on « branche » la configuration ω^i sur la prise $\psi^j.c$.

Le graphe se lit de la façon suivante : par exemple, dans la configuration ψ^2, si le candidat 2 est désigné comme vainqueur par un certain mode de scrutin, alors les configuration ω^4 et ω^5 deviennent manipulables : on dira aussi qu'elles sont *contaminées*. Le problème revient alors à choisir exactement une prise pour chaque configuration ψ tout en minimisant le poids total des configurations contaminées (ou plus simplement leur cardinal, si les configurations Condorcet sont en nombre fini et munies de la probabilité uniforme).

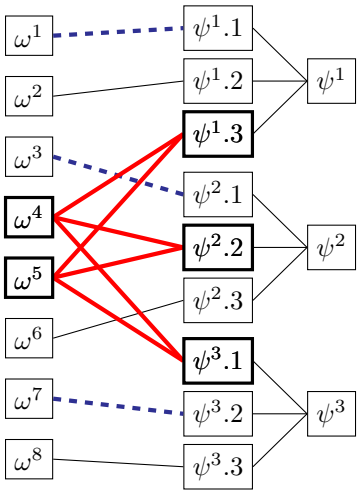


FIGURE 10.1 – Type de graphe utilisé pour la recherche d'un mode de scrutin Condorcet optimal, dans un espace électoral discret où les configurations semi-Condorcet sont de mesure nulle. Trait bleu discontinu : un résultat possible de l'algorithme glouton. Trait rouge épais : l'optimum global.

10.1.4 Restriction aux modes de scrutin ordinaux

Les théorèmes de tranchage (5.9 et 5.10) et leurs combinaisons avec les théorèmes de condorcification (5.12 et 5.15) nous suggèrent de chercher un optimum parmi les modes de scrutin qui, non seulement respectent le critère de Condorcet, mais sont également ordinaux (au sens où ils ne dépendent que des relations binaires de préférence, même si celles-ci ne sont pas des ordres). En effet, si la culture est décomposable, un tel optimum présentera un taux de manipulabilité minimal dans la classe plus large des modes de scrutin qui vérifient **CMInf** et qui ne sont pas nécessairement ordinaux (théorème d'optimalité 5.15). Nous pouvons donc ajouter l'hypothèse que le graphe d'opportunité est fini. Dans ce cas, il existe un nombre fini de modes de scrutin possibles, donc il existe nécessairement au moins un mode de scrutin dont le taux de manipulabilité est minimal, comme nous l'avons déjà remarqué dans la proposition 5.14.

10.1.5 Algorithme glouton

Il est facile de concevoir un algorithme de minimisation glouton qui procède par optimisation locale : pour chaque configuration non-Condorcet ψ^j, on choisit la prise $\psi^j.c$ qui possède le moins d'arêtes entrantes. Une telle solution approchée est représentée en trait bleu discontinu dans le figure 10.1. Dans cet exemple, on choisirait le candidat 1 ou 2 pour la configuration ψ^1 (ce qui contamine ω^1 ou ω^2), 1 ou 3 pour la configuration ψ^2 (ce qui contamine ω^3 ou ω^6) et 2 ou 3 pour la configuration ψ^3 (ce qui contamine ω^7 ou ω^8). Ainsi, 3 configurations seraient contaminées.

Cependant, cet algorithme n'est pas optimal pour un graphe général de ce type : sur la figure 10.1, on a représenté en rouge et en caractères gras la solution optimale, qui est strictement meilleure. En choisissant le candidat 3 pour ψ^1, 2 pour ψ^2 et 1 pour ψ^3, on ne contamine que 2 configurations, ω^4 et ω^5.

Dans le cas général d'un multigraphe étiqueté du type présenté dans la figure 10.1, il est probable que le problème à résoudre soit $\mathcal{NP}$-difficile. Cela n'exclut pas, *a priori*, que le problème puisse se simplifier, par exemple, pour le sous-ensemble des graphes obtenus comme graphes d'opportunité de chaque espace électoral des ordres stricts totaux (pour tout V impair et tout C). Mais il serait tout de même surprenant qu'on échappe à des situations pathologiques comme celle de la figure 10.1 et on peut d'ores et déjà s'attendre à ce que l'algorithme glouton ne soit pas optimal.

10.1.6 Approche exacte : optimisation linéaire en nombres entiers

Pour résumer, nous recherchons des modes de scrutin vérifiant **CMInf** et dont le taux de manipulabilité soit minimal. On peut restreindre la recherche aux modes de scrutin ordinaux (si la culture est décomposable) et vérifiant le critère de Condorcet. En outre, nous supposons que les configurations semi-Condorcet sont de mesure nulle et, dans celles-ci, nous imposons que le vainqueur soit un candidat Condorcet-admissible arbitraire, sans altérer l'optimalité du mode de scrutin obtenu.

Dans ce cadre, notre problème peut se ramener à un problème d'optimisation linéaire en nombres entiers de la façon suivante.

- Pour chaque couple (ψ, c), où ψ est une configuration non admissible et c un candidat, on déclare la variable entière $W(\psi, c)$ qui vaut 1 si c est choisi comme vainqueur pour ψ et 0 sinon.
- Pour chaque configuration Condorcet ω, on déclare la variable entière $\mathrm{MC}(\omega)$ qui vaut 1 si ω est manipulable et 0 sinon.
- Pour chaque configuration non admissible ψ, il y a un unique vainqueur, ce qu'on traduit par la contrainte $\sum_c W(\psi, c) = 1$.
- Pour chaque triplet (ω, ψ, c) tel que ω c-pointe vers ψ, on sait que si $W(\psi, c) = 1$, alors $\mathrm{MC}(\omega) = 1$, ce qu'on traduit par la contrainte $\mathrm{MC}(\omega) \geq W(\psi, c)$.
- L'objectif est de minimiser $\sum_\omega \pi(\omega)\mathrm{MC}(\omega)$.

L'avantage de cette formulation est qu'il existe des logiciels génériques qui implantent des algorithmes performants pour résoudre les problèmes d'optimisation linéaire en nombres entiers. À différents stades de notre travail, nous avons utilisé AIMMS puis IBM ILOG CPLEX Optimization Studio, qui utilisent tous deux le moteur CPLEX.

L'inconvénient de notre problème est qu'il nécessite un nombre de variables élevé. Par exemple, dans l'espace électoral des ordres stricts totaux avec V impair, il est compris entre $(C!)^V$ et $C \times (C!)^V$. En effet, il y a une variable pour chaque configuration Condorcet et C variables pour chaque configuration non admissible (il n'y a pas de configuration semi-Condorcet dans ce cas). Le problème n'est donc raisonnablement traitable que pour de très petites valeurs des paramètres. Cependant, nous allons voir que les cas exploitables en pratique sont déjà riches d'enseignements sur les modes de scrutin optimaux en général.

10.2 Modes de scrutin optimaux pour de faibles valeurs de V et C

À présent, nous considérons la culture impartiale : les électeurs ont donc presque sûrement un ordre strict total de préférence, ils sont indépendants, et

la culture est neutre (et anonyme). Cette culture vérifie donc bien les hypothèses précédentes. Dans ce cas particulier, il n'y a pas de configuration semi-Condorcet, donc les configurations non admissibles sont exactement les configurations non Condorcet. De plus, les configurations sont équiprobables, donc minimiser le poids total des configurations contaminées revient simplement à minimiser leur cardinal.

10.2.1 $V = 3$ et $C = 3$: des optima comme s'il en pleuvait

Pour $V = 3$ électeurs et $C = 3$ candidats, il y a $(C!)^V = 216$ configurations dans l'espace électoral et $C^{\mathrm{card}(\Omega)} \simeq 10^{103}$ modes de scrutin possibles au total (en incluant ceux qui ne vérifient pas le critère de Condorcet). Il est, cependant, possible de traiter le problème à la main, ce qui est intéressant pour forger l'intuition et va nous permettre de prouver un résultat plus général que la simple identification d'un optimum par la force brute de CPLEX.

Lemme 10.4
Sur les $(C!)^V = 216$ profils possibles, il y a 12 profils non-Condorcet. Il s'agit du profil suivant :

$$\psi : \begin{array}{c|c|c} a & b & c \\ b & c & a \\ c & a & b \end{array}$$

et de ses variantes par permutation des électeurs et/ou des candidats.

La configuration ψ est un exemple minimal du classique paradoxe de Condorcet : a bat b, qui bat c, qui bat a.

Démonstration. Pour qu'un profil soit non-Condorcet, il faut que les candidats en tête soient tous distincts, sinon le candidat placé en tête deux fois est vainqueur de Condorcet. Quitte à échanger b et c, supposons que l'électeur qui place a en tête préfère b à c. En comptant également l'électeur qui place b en tête, b est donc assuré d'une victoire contre c ; donc il est nécessaire d'avoir une victoire de a contre b et une victoire de c contre a. Il est alors facile de conclure qu'on est dans la configuration ψ. □

Pour chaque mode de scrutin vérifiant le critère de Condorcet, il s'agit uniquement de choisir le vainqueur dans chaque profil non Condorcet. Ainsi, il n'y a « que » $C^{12} = 531\,441$ modes de scrutin vérifiant le critère de Condorcet.

Avant d'aller plus loin sur le cas $V = 3$ et $C = 3$, on remarque le lemme facile suivant, qui est vrai en général, pas seulement pour cette valeur des paramètres.

Lemme 10.5
Pour qu'une configuration ω, de vainqueur de Condorcet w*, c-pointe vers un configuration ψ, il est nécessaire que* w *soit préféré à c dans ψ par une majorité stricte d'électeurs.*

Démonstration. C'est une simple reformulation du lemme 2.6. Dans ω, une majorité d'électeurs préfèrent w à c. Or, ces électeurs ne peuvent pas changer d'état dans ψ. □

Nous allons, à présent, procéder par une petite série de lemmes assez simples pour traiter le cas $V = 3$ et $C = 3$.

Lemme 10.6
La configuration ψ ci-dessus n'est a-pointée que par le profil suivant :

$$\omega : \begin{array}{c|c|c} a & b & c \\ c & c & a \\ b & a & b \end{array},$$

dont le vainqueur de Condorcet est c.

Démonstration. Soit ω un profil qui a-pointe vers ψ et w son vainqueur de Condorcet. En vertu du lemme 10.5, il faut que w possède une victoire contre a dans ψ : ce candidat w est donc nécessairement c. Dans ω, il y a donc au moins deux électeurs qui préfèrent c à a, et par a-pointage, leur bulletin ne change pas dans ψ. Comme seuls les deux derniers électeurs préfèrent c à a dans ψ, il s'agit nécessairement de ceux-là : ils ont donc le même bulletin dans ω. Reste à déterminer le bulletin du premier électeur, sachant que, par hypothèse, il préfère a à c. S'il place b en tête, c'est b qui est vainqueur de Condorcet (exclus) et s'il place b en milieu de liste, il s'agit de la configuration ψ. Il place donc nécessairement b en fin de bulletin. □

On en déduit immédiatement que ψ n'est b-pointée ou c-pointée que par un profil, obtenu à partir de ω par permutation des rôles.

Lemme 10.7
En reprenant les notations précédentes, la configuration ω, de vainqueur de Condorcet c, ne pointe que vers la configuration ψ.

Démonstration. Supposons que ω b-pointe vers une certaine configuration. Dans celle-ci, le premier et le dernier électeur, qui préfèrent c à b, sont conservés. Comme leurs bulletins ne sont pas des permutations circulaires l'un de l'autre, il est impossible d'obtenir un profil non-Condorcet ainsi (lemme 10.4).

Supposons, à présent, que ω a-pointe vers une certaine configuration. Dans celle-ci, les deux derniers électeurs, qui préfèrent c à a, sont conservés. Or on sait (lemme 10.4) que pour obtenir une configuration non-Condorcet, il faut que les bulletins soient des permutations circulaires les uns des autres, donc la seule configuration possible est ψ. □

Proposition 10.8
En culture impartiale, pour $V = 3$ et $C = 3$, tous les SVBE vérifiant le critère de Condorcet ont le même taux de manipulabilité : $\frac{24}{216} = \frac{1}{9} \simeq 11{,}11\ \%$.

Démonstration. La composante connexe du profil ψ contient exactement trois profils Condorcet, ω, ω' et ω'', comme on l'a représenté sur la figure 10.2. En effet, seuls ces trois profils pointent vers ψ (lemme 10.6) et aucun d'eux ne pointe vers un autre profil (lemme 10.7).

Par conséquent, le graphe contient exactement 12 composantes connexes non réduites à un singleton. Les autres composantes sont obtenues à partir de la figure 10.2 par permutation des électeurs et/ou des candidats.

Dans la composante représentée, quel que soit le vainqueur choisi pour le profil ψ, cela contamine un, et exactement un, profil Condorcet. Ainsi, les profils manipulables sont les 12 profils non-Condorcet et les 12 profils Condorcet contaminés : le taux de manipulabilité est $\frac{24}{216}$. □

Dans la section 2.8, nous avions affirmé de manière péremptoire que, dans l'espace électoral des ordres stricts totaux, il n'existe pas de SVBE Condorcet qui atteigne la borne supérieure de manipulabilité que nous avions donnée, c'est-à-dire

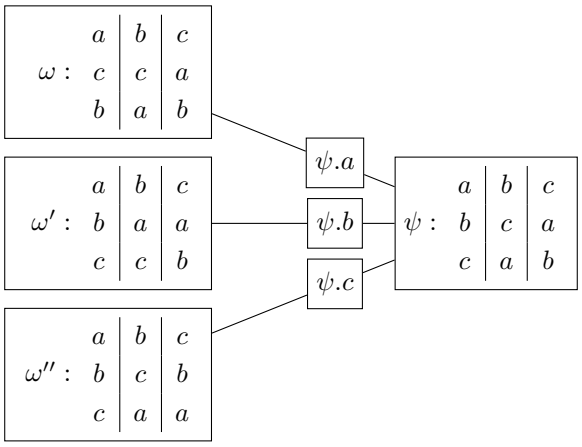


FIGURE 10.2 – Une composante connexe du graphe pour $V = 3$ et $C = 3$.

qui soit manipulable dans toutes les configurations non résistantes. La figure 10.2 permet de prouver ce fait : en effet, le profil ψ ne peut pas contaminer en même temps ω, ω' et ω''. Comme ces profils n'ont aucun autre moyen d'être contaminés (lemme 10.7), il est impossible qu'ils soient manipulables tous les trois. Il est, par ailleurs, facile de montrer que ces profils sont non-résistants, puisqu'en vertu du graphe de la figure 10.2, chacun d'entre eux peut être manipulable.

10.2.2 $V = 5$ et $C = 3$: CVTI et c'est tout !

Pour $V = 5$ électeurs et $C = 3$ candidats, il est difficile de certifier l'optimum global à la main, mais on peut, au moins, effectuer manuellement l'algorithme glouton. En fait, on va voir que, dans ce cas, le résultat est le même. L'approche manuelle va donc nous permettre d'avoir une meilleure compréhension du mode de scrutin optimal.

Parmi les $(C!)^V = 7\,776$ profils possibles, il est facile de montrer, par énumération manuelle à symétries près, que 540 profils sont non-Condorcet. D'une part, on a le profil suivant :

$$\psi : \begin{array}{|c|c|c|c|c|} a & a & b & b & c \\ b & b & c & c & a \\ c & c & a & a & b \end{array}$$

et ses 180 variantes par permutation des électeurs et/ou des candidats. D'autre part, on a le profil suivant :

$$\chi : \begin{array}{|c|c|c|c|c|} a & a & b & b & c \\ b & c & c & c & a \\ c & b & a & a & b \end{array}$$

et ses 360 variantes par permutation des électeurs et/ou des candidats. Le premier profil ψ possède deux fois moins de variantes que le second χ car il a une symétrie interne supplémentaire, celle consistant à échanger les deux premiers électeurs.

En appliquant les mêmes techniques que pour $V = 3$ électeurs, il est humainement possible de prouver qu'il existe un seul profil qui a-pointe vers ψ, un seul

profil qui a-pointe vers χ, et qu'il s'agit du même profil ω :

$$\omega : \begin{array}{c|c|c|c|c} a & a & b & b & c \\ c & c & c & c & a \\ b & b & a & a & b \end{array},$$

dont le vainqueur de Condorcet est c.

Il n'est guère plus difficile de montrer qu'en revanche, il y a strictement plus d'un profil qui b-pointe (ou qui c-pointe) vers ψ (ou vers χ). Ainsi, en utilisant l'algorithme glouton, on doit avoir $f(\psi) = f(\chi) = a$, ce qui rend manipulable uniquement le profil ω. On pourra, d'ailleurs, remarquer que ce profil ω est également rendu manipulable par le profil obtenu en échangeant les deux premiers électeurs de χ. Par conséquent, l'algorithme glouton contamine ω et ses variantes par permutation, c'est-à-dire 180 profils Condorcet. Au total, il y a donc $360 + 180 + 180 = 720$ profils manipulables et le taux de manipulabilité est $\frac{720}{7\,776} = \frac{5}{54} \simeq 9{,}26$ %.

L'utilisation de CPLEX permet d'établir qu'il s'agit, en fait, d'un optimum global. Pour savoir si cette optimum est unique, on demande ensuite à CPLEX de trouver l'optimum soit avec la contrainte supplémentaire $W(\psi, a) = 0$, soit avec la contrainte supplémentaire $W(\chi, a) = 0$, c'est-à-dire qu'on interdit au candidat a d'être vainqueur en ψ ou en χ. Dans les deux cas, on constate que le taux de manipulabilité optimal trouvé est strictement supérieur à $\frac{720}{7\,776} = \frac{5}{54}$. Il est donc obligatoire, pour un mode de scrutin optimal, d'élire le candidat a en ψ et en χ. Par argument de symétrie, il y a donc aussi unicité du vainqueur possible, dans un mode de scrutin optimal, pour tout profil obtenu à partir de ψ ou χ par permutation des électeurs et/ou des candidats. L'optimum exhibé précédemment est donc unique.

Par ailleurs, on constate que le mode de scrutin obtenu est CVTI : en effet, il suffit de vérifier que, dans les profils ψ et χ ci-dessus, le mode de scrutin donne le même résultat que VTI, ce qui est le cas. Pour $V = 5$ et $C = 3$, comme il n'y a jamais d'égalité entre plusieurs candidats en CVTI (ou en VTI), la question de la règle de départage utilisée ne se pose pas : CVTI définit donc parfaitement la solution de manière non ambiguë, et celle-ci est anonyme et neutre. On résume toutes ces observations par la proposition suivante et son corollaire immédiat.

Proposition 10.9

On considère l'espace électoral des ordres stricts totaux pour $V = 5$ et $C = 3$, muni de la culture impartiale.

Parmi les SVBE vérifiant le critère de Condorcet, CVTI est l'unique système possédant un taux de manipulabilité minimal.

Corollaire 10.10

On considère un espace électoral probabilisé avec $V = 5$ et $C = 3$. On suppose que la culture est décomposable (par exemple, parce que les électeurs sont indépendants) et que la loi de probabilité induite sur le profil P *est la culture impartiale.*

Parmi les modes de scrutin vérifiant **CMInf***, CVTI possède un taux de manipulabilité minimal, qui est égal à $\frac{720}{7\,776} = \frac{5}{54} = \simeq 9{,}26$ %. C'est le seul système qui possède cette propriété, est un SVBE et vérifie le critère de Condorcet. En outre, il est anonyme et neutre (cf. section 1.2.2).*

Par ailleurs, nous allons voir que le mode de scrutin optimal ne peut, dans ce cas, se reposer uniquement sur la matrice des duels. En effet, le profil χ vu

précédemment possède la matrice des duels suivante :

$D(\chi)$	a	b	c
a	–	3	2
b	2	–	3
c	3	2	–

Celle-ci est invariante quand on applique aux candidats la permutation circulaire $(a \to b \to c \to a)$. En particulier, c'est la même matrice majoritaire pour le profil :

$$\chi' : \begin{array}{c|c|c|c|c} b & b & c & c & a \\ a & c & a & a & b \\ c & a & b & b & c \end{array},$$

qui est obtenu à partir de χ par cette permutation.

Le mode de scrutin optimal, dont nous avons montré qu'il s'agit de CVTI pour ces paramètres, ne peut pas être basé uniquement sur la matrice majoritaire : il désigne comme vainqueur a dans χ et b dans χ'.

Plus formellement, considérons un mode de scrutin f dont le résultat ne dépend que de la matrice majoritaire. Ce mode de scrutin peut utiliser une règle de départage basée sur l'identité des candidats et/ou sur la matrice majoritaire, mais pas sur les préférences détaillées des électeurs : par exemple, il peut départager les candidats par ordre alphabétique ou par ordre d'indice, comme dans SVAMP. Un tel mode de scrutin f doit alors désigner le même vainqueur dans χ et dans χ' et il n'est donc pas optimal.

À nos yeux, une conséquence importante de cette étude de cas est que dans la classe **CMInf**, aucun mode de scrutin dont le résultat ne dépend que de la matrice des duels ne peut être optimal pour $V = 5$ et $C = 3$. En effet, nous venons de prouver cette affirmation pour les systèmes Condorcet. Pour un système non-Condorcet qui ne dépend que de la matrice majoritaire, il suffit de remarquer que sa condorcification (qu'on sait au plus aussi manipulable que l'original) ne dépend que de la matrice des duels également ; donc il ne saurait être optimal. Cette constatation exclut notamment Condorcet-doyen, PO, CSD, Maximin, ainsi que les méthodes de Baldwin, Borda, Black (Condorcet-Borda), Nanson, Kemeny et Schulze.

10.2.3 $V = 7$ et $C = 3$: bien choisir la règle de départage de CVTI

Pour $V = 7$ électeurs et $C = 3$ candidats, il devient extrêmement douloureux de gérer le problème à la main. Il y a $(C!)^V = 279\,936$ profils au total. Et surtout, il y a 7 types de profils non-Condorcet (à permutation des électeurs et/ou des candidats près), au lieu d'un ou deux types dans les cas précédents. Rappelons que, pour chacun de ces profils non-Condorcet, il faut considérer tous les profils Condorcet qui pointent vers lui.

Nous avons donc recours à des calculs par ordinateur. L'algorithme glouton trouve un mode de scrutin possédant un taux de manipulabilité de 13,46 %. CPLEX trouve une solution optimale avec un taux de $\frac{31\,920}{279\,936} = \frac{665}{5832} \simeq 11{,}40$ %. L'algorithme glouton n'est donc pas optimal dans ce cas.

Afin d'étudier les propriétés de l'ensemble des solutions, on peut utiliser CPLEX en imposant des contraintes supplémentaires. Par exemple, on peut imposer la valeur $W(\psi, c) = 0$ pour un certain ψ et un certain c : si l'optimum

n'est plus atteignable dans ce cas, cela signifie qu'il est nécessaire que c soit vainqueur en ψ.

On peut aussi chercher les systèmes de votes qui sont optimaux dans la classe plus restreinte de ceux qui sont anonymes, neutres ou les deux. Par exemple, pour trouver un optimum anonyme, on quotiente le graphe d'opportunité par la relation d'équivalence qui consiste, pour deux configurations, à être déduites l'une de l'autre par permutation des électeurs. Pour trouver un optimum neutre, on quotiente le graphe d'opportunité par permutation des candidats ; il faut alors prendre garde à modifier le c-pointage en conséquence.

Le taux de manipulabilité minimal trouvé par CPLEX pour un mode de scrutin anonyme et neutre est $\frac{31\,920}{279\,936} = \frac{665}{5832} \simeq 11{,}40$ % : il s'agit donc d'un optimum, même parmi les modes de scrutin qui ne sont pas forcément anonymes et/ou neutres.

La table 10.3 résume les principales conclusions de cette étude. Pour chacun des 7 types de profil non-Condorcet (à permutation des électeurs et/ou des candidats près), on donne une condition nécessaire sur le vainqueur pour que le mode de scrutin soit optimal, éventuellement en ajoutant la contrainte d'anonymat et de neutralité.

Prenons l'exemple du profil ψ_4. Pour un mode de scrutin optimal, on lit dans la colonne f de la table 10.3 qu'il est nécessaire que le vainqueur soit a ou b. Si, en outre, on exige que le mode de scrutin soit anonyme et neutre, alors on lit dans la colonne $f_{\text{a\&n}}$ de la table 10.3 que le vainqueur est nécessairement b. C'est loin d'être une évidence, puisque le profil ψ_4 ne possède aucune symétrie par rapport aux candidats. Cependant, si on impose a comme vainqueur pour un mode de scrutin anonyme et neutre, alors on constate, par utilisation de CPLEX, qu'il n'est plus possible d'atteindre le minimum de manipulabilité. Par ailleurs, nous verrons que les conditions de la colonne $f_{\text{a\&n}}$ sont, en fait, des conditions suffisantes pour obtenir un optimum.

Pour un optimum quelconque, les conditions indiquées dans la colonne f de la table 10.3 sont seulement nécessaires : il est possible que toutes ces conditions soient vérifiées, ainsi que leurs variantes par permutation des électeurs et/ou des candidats, mais que le mode de scrutin obtenu ne soit pas optimal. C'est le cas, par exemple, de CVTI muni de la règle de départage par ordre lexicographique sur les candidats utilisée dans SWAMP, comme nous le verrons dans la table 10.4 de la section 10.3. Ces conditions nécessaires sont ajustées, dans le sens suivant : pour chacune d'elles, il existe un optimum qui la vérifie. Par exemple, il existe au moins un optimum f tel que $f(\psi_4) = a$ et il en existe au moins un tel que $f(\psi_4) = b$. Ces conditions nécessaires permettent d'établir la proposition suivante.

Proposition 10.11

On considère l'espace électoral des ordres stricts totaux pour $V = 7$ et $C = 3$, muni de la culture impartiale.

Quelle que soit la règle de départage utilisée, les modes de scrutin suivants ne sont pas optimaux parmi les systèmes Condorcet : Condorcet-Bucklin, Condorcet-Borda (méthode de Black), Condorcet-Coombs et Condorcet-BI.

Aucun mode de scrutin basé uniquement sur la matrice des duels ne peut être optimal.

Tout mode de scrutin Condorcet optimal possède un taux de manipulabilité de $\frac{31\,920}{279\,936} = \frac{665}{5832} \simeq 11{,}40$ % et peut être vu comme CVTI, muni d'une règle de départage adéquate.

Profil non Condorcet	Matrice des duels	f	$f_{\text{a\&n}}$	CVTI
ψ_1 : $a\ a\ a\ b\ c\ c\ c$ / $c\ c\ c\ a\ b\ b\ b$ / $b\ b\ b\ c\ a\ a\ a$	$D(\psi_1)$: a: – 3 4; b: 4 – 1; c: 3 6 –	a	a	a
ψ_2 : $a\ a\ a\ b\ c\ c\ c$ / $b\ c\ c\ a\ b\ b\ b$ / $c\ b\ b\ c\ a\ a\ a$	$D(\psi_2)$: a: – 3 4; b: 4 – 2; c: 3 5 –	a	a	a
ψ_3 : $a\ a\ a\ b\ c\ c\ c$ / $b\ b\ c\ a\ b\ b\ b$ / $c\ c\ b\ c\ a\ a\ a$	$D(\psi_3)$: a: – 3 4; b: 4 – 3; c: 3 4 –	a	a	a
ψ_4 : $a\ a\ a\ b\ b\ c\ c$ / $c\ c\ c\ a\ c\ b\ b$ / $b\ b\ b\ c\ a\ a\ a$	$D(\psi_4)$: a: – 3 4; b: 4 – 2; c: 3 5 –	a, b	b	a, b
ψ_5 : $a\ a\ a\ b\ b\ c\ c$ / $b\ c\ c\ a\ c\ b\ b$ / $c\ b\ b\ c\ a\ a\ a$	$D(\psi_5)$: a: – 3 4; b: 4 – 3; c: 3 4 –	a, b	b	a, b
ψ_6 : $a\ a\ a\ b\ b\ c\ c$ / $c\ c\ c\ a\ a\ b\ b$ / $b\ b\ b\ c\ c\ a\ a$	$D(\psi_6)$: a: – 3 5; b: 4 – 2; c: 2 5 –	a, b	a, b	a, b
ψ_7 : $a\ a\ a\ b\ b\ c\ c$ / $b\ c\ c\ a\ a\ b\ b$ / $c\ b\ b\ c\ c\ a\ a$	$D(\psi_7)$: a: – 3 5; b: 4 – 3; c: 2 4 –	a, b	a, b	a, b

TABLE 10.3 – Solutions optimales pour $V = 7$ et $C = 3$. f : optimum quelconque. $f_{\text{a\&n}}$: optimum anonyme et neutre. Lecture : pour que f soit un optimum, il faut que $f(\psi_4)$ soit a ou b, mais les conditions présentées ne sont pas suffisantes. Pour que $f_{\text{a\&n}}$ soit un optimum anonyme et neutre, il faut et il suffit de vérifier les conditions de la colonne $f_{\text{a\&n}}$ (et leurs variantes par anonymat et neutralité).

Démonstration. D'après la table 10.3, pour que le mode de scrutin soit optimal, il faut que le candidat a soit élu en ψ_1. Or en Condorcet-Bucklin ou Condorcet-Borda (Black), c est vainqueur.

De même, il faut que a soit vainqueur en ψ_2. Or en Condorcet-Coombs ou en BI, c est vainqueur.

Toujours d'après la table 10.3, il est nécessaire que le candidat a soit vainqueur en ψ_3. Ce profil possède une matrice des duels symétrique par la permutation $(a \to b \to c \to a)$, mais le profil lui-même ne possède pas de symétrie par rapport aux candidats. Notons ψ_3' le profil obtenu en appliquant à ψ_3 cette permutation circulaire. Si le mode de scrutin ne dépend que de la matrice majoritaire, alors il doit aussi désigner a en ψ_3'. Mais s'il est optimal, alors le vainqueur doit être l'image de a par la permutation effectuée, c'est-à-dire b. Ces deux hypothèses sont donc incompatibles.

Dans la table 10.3, on constate que les vainqueurs autorisés pour un mode de scrutin optimal sont toujours exactement les vainqueurs possibles par CVTI, le vainqueur effectif dépendant de la règle de départage utilisée pour ce dernier. Ceci garantit que tout optimum est une variante de CVTI (mais pas la réciproque). □

Pour un optimum anonyme et neutre, CPLEX permet de s'assurer que les conditions présentées dans la colonne $f_{\text{a\&n}}$ de la table 10.3 sont non seulement nécessaires mais aussi suffisantes. On en déduit la proposition suivante.

Proposition 10.12

On considère l'espace électoral des ordres stricts totaux pour $V = 7$ et $C = 3$, muni de la culture impartiale.

Quelle que soit la règle de départage anonyme et neutre utilisée, les modes de scrutin suivants ne sont pas optimaux parmi les systèmes Condorcet : VTID, PO, CSD, Maximin, Condorcet-Uninominal, méthodes de Baldwin, Nanson, Schulze ou Kemeny.

Parmi les SVBE Condorcet, il en existe 4 qui sont optimaux, anonymes et neutres. Ils peuvent tous être vus comme CVTI, muni d'une règle de départage adéquate.

Démonstration. D'après la table 10.3, pour qu'un mode de scrutin anonyme et neutre soit optimal, il est nécessaire que b soit élu en ψ_4. Or ce n'est le cas ni en VTID (b affronte c dans le premier duel et perd), ni en PO (la première victoire validée est celle de c contre b), ni en CSD ou Maximin (b a le score le plus bas), ni en Condorcet-Uninominal (a est vainqueur), ni en méthode de Baldwin ou Nanson (b est éliminé au premier tour), ni en méthode de Schulze (c est *meilleur* que b), ni en méthode de Kemeny (b est nécessairement derrière c).

Pour définir un mode de scrutin optimal, anonyme et neutre, il est nécessaire et suffisant de choisir le vainqueur parmi deux possibilités pour les profils ψ_6 et ψ_7. Il y a donc 4 modes de scrutin possibles. Tous les autres profils non-Condorcet se déduisent de ceux de la table 10.3 par anonymat et neutralité. Le fait qu'ils soient tous une variante de CVTI est un simple cas particulier de la proposition 10.11, que nous rappelons par souci d'exhaustivité. □

Les propositions 10.11 et 10.12 excluent tous les modes de scrutin Condorcet que nous avons étudiés hormis CVTI, quelle que soit la règle de départage utilisée si celle-ci est anonyme et neutre. CVTI, s'il est muni d'une règle de départage anonyme et neutre, est optimal si et seulement la règle de départage vérifie les conditions indiquées dans la colonne $f_{\text{a\&n}}$ de la table 10.3.

Il serait intéressant de formuler une telle règle de départage, non pas par une énumération extensive des cas, mais par une propriété que nous pourrions

tenter d'appliquer à d'autres cas que $V = 7$ et $C = 3$ (définition intensive). Ceci permettrait peut-être d'identifier une version de CVTI particulièrement résistante à la manipulation dans le cas général.

Faisons un premier essai avec la règle suivante : en cas d'égalité entre deux candidats à un tour d'élimination, on organise un duel virtuel et on élimine le perdant [1]. Mais cette règle ne fonctionne pas. En effet, dans les profils ψ_4, ψ_5, ψ_6 et ψ_7, les candidats b et c sont en position d'élimination au premier tour et c'est b, le perdant du duel, qui est éliminé ; puis a est vainqueur. Par permutation des électeurs et/ou des candidats, on en déduit que cette règle de départage permet de fixer le vainqueur de façon unique dans toute situation non-Condorcet et que le mode de scrutin obtenu est anonyme et neutre. Mais, dans ce cas, le vainqueur en ψ_4 devrait être b et non a (table 10.3). Donc le système ne peut pas être optimal.

Considérons une autre règle de départage : si deux vainqueurs sont possibles par CVTI, alors on organise un duel virtuel et on élit le vainqueur [2]. Dans ψ_4, ψ_5, ψ_6 et ψ_7, le vainqueur obtenu est b, donc le mode de scrutin est effectivement optimal.

Cependant, cette solution n'est pas parfaitement satisfaisante. Pour qu'un mode de scrutin anonyme et neutre soit optimal, il est nécessaire que b soit vainqueur en ψ_4 et ψ_5 ; mais en ψ_6 ou ψ_7, choisir a convient également. La règle de départage exhibée est donc suffisante mais pas nécessaire : pour couvrir exactement les solutions anonymes et neutres, il faudrait identifier une règle de départage « naturelle » (c'est-à-dire définie de façon intensive et aussi concise que possible) qui choisisse le candidat b dans les profils ψ_4 et ψ_5 mais qui reste indécise entre a et b dans les profils ψ_6 et ψ_7. Nous laissons cette question pour des travaux futurs.

10.2.4 $V = 3$ et $C = 4$: un complexe ensemble d'optima

Pour $V = 3$ candidats et $C = 4$ candidats, on pourrait croire que le problème est plus simple car il n'y a « que » $(C!)^V = 13\,824$ profils au total. Cependant, il y a 12 types de profils non-Condorcet (à permutation des électeurs et/ou des candidats près), au lieu de 7 dans le cas précédent. En outre, nous allons voir qu'il n'y a aucun optimum anonyme et neutre et que l'espace des solutions est, par conséquent, plus difficile à explorer.

L'algorithme glouton trouve un mode de scrutin possédant un taux de manipulabilité de 21,09 %. CPLEX trouve une solution optimale avec un taux de $\frac{2\,688}{13\,824} = \frac{7}{36} \simeq 19{,}44$ %. L'algorithme glouton n'est donc de nouveau pas optimal dans ce cas. En étudiant de façon détaillée la solution que CPLEX exhibe par défaut, nous avons pu montrer qu'elle n'était ni anonyme ni neutre. Celui-ci est nécessairement non unique : en effet, tout mode de scrutin qui en est déduit par permutation des candidats en est distinct et il est également optimal.

Si on impose que la solution soit neutre, on trouve le même taux de manipulabilité. Si on impose que la solution soit anonyme, on obtient un taux de $\frac{2\,712}{13\,824} = \frac{113}{576} \simeq 19{,}62$ %. Si on impose que la solution soit anonyme et neutre, on obtient un taux de $\frac{3\,264}{13\,824} = \frac{17}{72} \simeq 23{,}61$ %.

1. En cas d'égalité entre davantage de candidats, on peut choisir le candidat éliminé en utilisant une règle de départage dont le duel électoral est un cas particulier : par exemple, on élimine le candidat ayant le plus faible score de Borda dans la matrice des duels restreinte aux candidats ex-aequo.

2. En cas d'égalité entre davantage de candidats, on peut choisir le candidat vainqueur en utilisant une règle de départage dont le duel électoral est un cas particulier : par exemple, on élit le candidat ayant le score de Borda le plus élevé dans la matrice des duels restreinte aux vainqueurs potentiels.

Il est difficile d'étudier l'espace des solutions quelconques dans ce cas. En effet, leur multiplicité entraîne qu'il y a généralement 2 voire 3 candidats vainqueurs possibles pour chaque profil non-Condorcet [3].

On peut, cependant, en dire davantage sur les solutions neutres. Pour cela, considérons les profils suivants.

$$\psi_1 : \begin{array}{c|c|c} a & c & b \\ b & a & c \\ c & b & a \\ d & d & d \end{array} \qquad \psi_2 : \begin{array}{c|c|c} a & b & c \\ b & c & a \\ c & a & b \\ d & d & d \end{array}$$

Par permutation des candidats, ces deux profils couvrent tous les cas où les électeurs placent le même candidat en dernier et réalisent un exemple minimal de paradoxe de Condorcet entre les trois autres candidats. Ces profils se déduisent l'un de l'autre par permutation des deux électeurs de droite.

En utilisant CPLEX, on constate qu'un optimum ne peut élire d en ψ_1, ce qui est assez intuitif puisque c'est le candidat placé en dernier par tous les électeurs. Par symétrie du profil, tout candidat a, b ou c peut être désigné par un mode de scrutin optimal. Cela revient à privilégier l'électeur 1, 2 ou 3 dans cette situation ψ_1 et celles qui s'en déduisent par permutation des candidats. Pour fixer les idées, supposons que a est désigné. Ceci désigne l'électeur 1 comme privilégié en ψ_1 et rompt également la symétrie entre les deux autres électeurs : l'électeur 2 (resp. 3) est celui pour lequel le candidat vainqueur est en seconde (resp. troisième) position.

Dans le profil ψ_2, l'utilisation de CPLEX nous informe qu'il n'y a alors plus que deux vainqueurs possibles : a ou b. Autrement dit, on favorise soit le même électeur que dans le profil ψ_1, soit l'électeur dont le second candidat préféré est élu en ψ_1. Quand on fixe l'une ou l'autre option, CPLEX permet de déterminer que l'optimum neutre est unique.

Comme il y a 3 choix possibles pour ψ_1 et ensuite 2 choix pour ψ_2, il y a 6 solutions possibles. En voyant ce chiffre, on pourrait penser qu'il s'agit d'un certain mode de scrutin et de ses variantes par les $3! = 6$ permutations des électeurs. Mais ce n'est pas le cas. En effet, considérons l'unique mode de scrutin optimal neutre f obtenu pour $f(\psi_1) = a$ et $f(\psi_2) = a$. Si on échange les deux derniers électeurs, alors le mode de scrutin optimal obtenu est le même. Ainsi, l'orbite de f par permutation des électeurs ne possède que 3 éléments et non 6. De même, si on considère l'unique mode de scrutin optimal neutre g obtenu pour $g(\psi_1) = a$ et $g(\psi_2) = b$, son orbite ne possède que 3 éléments par permutation des candidats. Il y a donc exactement deux solutions neutres distinctes à permutation près des électeurs.

Les profils ψ_1 et ψ_2 ci-dessus permettent également de comprendre intuitivement pourquoi l'optimum anonyme et neutre est sensiblement plus manipulable que l'optimum général. En effet, si on impose l'anonymat et la neutralité, il est nécessaire d'élire le candidat d dans ces deux profils, alors qu'il est placé dernier par tous les électeurs. On pourra rapprocher ce résultat du fait que, pour ces valeurs des paramètres, il n'existe pas de mode de scrutin neutre, anonyme et efficient (voir section 1.2.2).

La proposition suivante va nous permettre de résumer les constatations faites à ce stade.

3. Dans un mode de scrutin optimal, il y a 2 vainqueurs possibles dans 8 types de profils non-Condorcet et 3 vainqueurs possibles dans 4 types de profils non-Condorcet. Ceci reste vrai même si on impose la neutralité de la solution.

Proposition 10.13

On considère l'espace électoral des ordres stricts totaux pour $V = 3$ et $C = 4$, muni de la culture impartiale.

Tout mode de scrutin Condorcet optimal possède un taux de manipulabilité de $\frac{2\,688}{13\,824} = 19{,}44$ %. Il existe au moins un optimum non neutre (et il n'y a pas unicité).

Il existe exactement 6 modes de scrutin optimaux qui sont neutres. À permutation des électeurs près, il y a 2 solutions distinctes. Chacune possède 3 variantes par permutation des électeurs.

Il n'existe pas d'optimum qui soit anonyme.

À présent, nous allons présenter quelques résultats qui permettent de mettre les solutions en rapport avec les modes de scrutin classiques.

Proposition 10.14

On considère l'espace électoral des ordres stricts totaux pour $V = 3$ et $C = 4$, muni de la culture impartiale.

Quelle que soit la règle de départage neutre utilisée, les modes de scrutin suivants ne sont pas optimaux parmi les systèmes Condorcet : CVTI, Condorcet-Uninominal, CSD, Condorcet-Borda (méthode de Black), VTID, Condorcet-Coombs, méthodes de Baldwin, Nanson et Kemeny.

Démonstration. Comme ci-dessus, notons f le mode de scrutin optimal neutre obtenu pour $f(\psi_1) = a$ et $f(\psi_2) = a$ et g le mode de scrutin optimal neutre obtenu pour $g(\psi_1) = a$ et $g(\psi_2) = b$.

Considérons le profil suivant.

$$\psi : \begin{array}{c|c|c} b & a & d \\ c & c & a \\ d & b & b \\ a & d & c \end{array}$$

$D(\psi)$	a	b	c	d
a	–	2	2	1
b	1	–	2	2
c	1	1	–	2
d	2	1	1	–

Grâce à l'optimisation réalisée par CPLEX, nous avons pu déterminer que le vainqueur doit être c, aussi bien dans f que dans g.

Or il n'est élu ni en CVTI (c est éliminé au premier tour), ni en Condorcet-Uninominal (c n'a aucun vote), ni en CSD (c a le score le plus faible), ni en Condorcet-Borda ou en méthode de Nanson (c a le score de Borda le plus faible), ni en méthode de Kemeny (si on réarrange la matrice des duels avec c en première position, il y a deux défaites d'une amplitude d'une voix au-dessus de la diagonale, au lieu d'une seule défaite d'une seule voix si on dispose les candidats par ordre alphabétique comme ci-dessus).

En VTID, comme c ne possède aucune voix et que tous les autres en ont au moins une, c participe à tous les duels éliminatoires. Comme il n'est pas vainqueur de Condorcet, il est impossible qu'il gagne chacun d'eux. Donc c ne peut pas être élu.

En méthode de Baldwin ou en Condorcet-Coombs, pour que c soit vainqueur, il est nécessaire qu'il soit opposé à d au dernier tour car c'est le seul candidat qu'il peut vaincre en duel électoral. Au premier tour, il ne faut donc éliminer ni c ni d.

Or, en Condorcet-Coombs, il y a une égalité entre a, c et d au premier tour ; mais comme ni c ni d ne doit être éliminé, c'est a qui doit être éliminé. Dans ce cas, d est éliminé, puis c et c'est b qui est vainqueur. Donc c ne peut gagner.

En méthode de Baldwin, c'est précisément c ou d qui est éliminé au premier tour. Donc c ne peut pas être élu. □

Proposition 10.15

On considère l'espace électoral des ordres stricts totaux pour $V = 3$ et $C = 4$, muni de la culture impartiale.

Tout mode de scrutin optimal parmi les systèmes Condorcet peut être vu comme Condorcet-Bucklin, Condorcet-BI, la méthode de Schulze, PO ou Maximin, muni(e) d'une règle de départage adéquate.

Démonstration. Afin de ne pas alourdir l'exposé, nous donnons simplement la méthode suivie. Pour chacun des 12 types de profil non-Condorcet (à permutation des électeurs et/ou des candidats près), on détermine à l'aide de CPLEX quels sont les candidats qui peuvent être vainqueurs dans un mode de scrutin optimal, comme on l'avait fait dans la table 10.3 pour $V = 7$ et $C = 3$.

Dans chaque cas, on constate en pratique que l'ensemble des vainqueurs possibles est le même avec ou sans l'hypothèse de neutralité et qu'il est égal à l'ensemble des candidats qui peuvent gagner en Condorcet-Bucklin. On vérifie manuellement que cet ensemble est inclus, d'ailleurs strictement, dans l'ensemble des candidats qui peuvent gagner en Condorcet-BI (resp. méthode de Schulze, PO, Maximin). Nous insistons sur le fait que cela signifie que tout optimum est Condorcet-Bucklin muni d'une certaine règle de départage, mais pas nécessairement que toute règle de départage rende Condorcet-Bucklin optimal (comme nous le verrons dans la table 10.4). □

La proposition précédente peut être comprise comme un argument en faveur des cinq modes de scrutin cités. Cependant, parmi ceux-ci, certains entraînent plus d'égalités que d'autres. Ainsi, Condorcet-Bucklin laisse le moins de marge de manœuvre à la règle de départage : sur les 12 types de profils non-Condorcet (à permutation près des électeurs et/ou des candidats), elle cause une 3-égalité (c'est-à-dire une égalité entre 3 candidats) dans 4 profils et une 2-égalité dans 8 profils. Le système BI cause toujours une 3-égalité. La méthode de Schulze et PO causent une 4-égalité (c'est-à-dire une égalité entre tous les candidats !) dans 6 profils et une 3-égalité dans 6 profils. Enfin, Maximin cause une 4-égalité dans 7 profils et une 3-égalité dans 5 profils. Il n'est donc pas si étonnant que tout mode de scrutin optimal puisse être vu comme Maximin muni d'une règle de départage adéquate, puisque Maximin ne restreint que très peu l'ensemble des vainqueurs possibles dans le cas $V = 3$ et $C = 4$.

Ainsi, nous avons pu déterminer le taux de manipulabilité minimal pour un mode de scrutin Condorcet, et plus généralement pour un mode de scrutin vérifiant **CMInf**, en culture impartiale dans quatre cas : $V = 3, 5$ ou 7 électeurs et $C = 3$ candidats ; $V = 3$ électeurs et $C = 4$ candidats. Pour davantage d'électeurs ou de candidats, le problème est trop volumineux pour être traité sur la machine utilisée[4].

10.3 Comparaison entre l'optimum et les modes de scrutin classiques

À présent, nous allons comparer le taux optimum de manipulabilité (au sein de la classe **CMInf**) à celui de divers modes de scrutin. On remarquera que si on dispose du graphe d'opportunité, on peut déterminer facilement le taux de manipulabilité de n'importe quel mode de scrutin Condorcet : en effet, il suffit

4. Dell Precision M6600, Intel Core I7-2820QM à 2,30 GHz, 8 Mo de cache, 16 Go de RAM à 1,33 GHz en DDR3.

de reporter dans le graphe les vainqueurs des configurations non-Condorcet et de contaminer les configurations Condorcet correspondantes.

Dans la table 10.4, on a indiqué les taux de manipulabilité exacts d'une variété de modes de scrutin.

— L'optimum est obtenu par optimisation linéaire dans CPLEX, de même que l'optimum dans la classe restreinte des systèmes Condorcet qui sont anonymes, neutres ou les deux à la fois. Pour $V = 3$ et $C = 3$, il n'est pas possible d'avoir un mode de scrutin anonyme et neutre car le choix du vainqueur dans l'exemple minimal de paradoxe de Condorcet brise au moins une symétrie (voir proposition 1.16).

— L'optimum approché est obtenu par l'algorithme glouton présenté précédemment.

— Pour les modes de scrutin respectant le critère de Condorcet, on utilise le graphe d'opportunité pour déterminer exactement la manipulabilité. Pour ceux qui sont implantés dans SVAMP, on peut également utiliser celui-ci avec l'option `CM_option = 'exact'`, ce qui permet de vérifier la cohérence des résultats obtenus.

— Pour les autres modes de scrutin, par exemple Borda ou Veto, on utilise SVAMP avec l'option `CM_option = 'exact'`.

Dans cette table, il est important de remarquer que les règles de départage utilisées sont importantes, car on considère un faible nombre d'électeurs. Ainsi, il est abusif de dire qu'on a représenté le taux de manipulabilité de VTI : il s'agit de celui de VTI muni de la règle de départage utilisée dans SVAMP, c'est-à-dire par ordre d'indice sur les candidats. Ainsi, nous avons constaté que certains modes de scrutin connus (notamment CVTI pour $V = 7$ et $C = 3$; Condorcet-Bucklin pour $V = 3$ et $C = 4$) pouvaient être optimaux s'ils étaient munis d'une règle de départage adéquate, mais cela n'apparaît pas dans la table 10.4. Pour des travaux futurs, il serait intéressant de faire varier cette règle de départage.

Pour chaque valeur des paramètres, on a également indiqué le taux de profils non-Condorcet, ce qui fournissait une première borne inférieure du taux de manipulabilité pour les modes de scrutin vérifiant **CMInf** quand nous ne connaissions pas encore le taux optimal.

Chaque taux de manipulabilité est indiqué en caractères gras et verts s'il est égal au taux de manipulabilité minimal dans **CMInf**.

Pour $V = 3$ et $C = 3$, de nombreux modes de scrutin sont optimaux, en particulier tous ceux qui vérifient le critère de Condorcet, comme nous l'avons montré en section 10.2.1.

Pour $V = 5$ et $C = 3$, les modes de scrutin optimaux sont CVTI (qui coïncide avec le résultat de l'algorithme glouton et qui est le seul optimal Condorcet, comme nous l'avons montré en section 10.2.2), ainsi que VTI et SE (qui coïncide avec l'uninominal à deux tours puisqu'il y a 3 candidats). Ce n'est pas contradictoire avec ce que nous avons vu dans la section 10.2.2 : en effet, nous avons montré que, parmi les modes de scrutin vérifiant **CMInf**, CVTI est l'unique optimum qui est un SVBE Condorcet. Mais VTI et SE ne respectent pas le critère de Condorcet et SE n'est pas un SVBE. Il est donc impossible qu'ils soient moins manipulables que CVTI, mais il n'est pas interdit qu'ils aient la même manipulabilité. Par ailleurs, Green-Armytage *et al.* (2014) ont prouvé que, quand il y a 3 candidats, CVTI et VTI ont la même manipulabilité.

Pour $V = 7$ et $C = 3$, nous avons montré qu'il existe 4 modes de scrutin qui sont optimaux parmi les systèmes Condorcet et qui sont également anonymes et neutres. Ils peuvent tous être vus comme CVTI, muni d'une règle de départage adéquate.

Populations				
Nombre d'électeurs V	3	5	7	3
Nombre de candidats C	3	3	3	4
Situations non Condorcet (%)	5,56	6,94	7,50	11,11
Taux de MC (%)				
Optimum	**11,11**	**9,26**	**11,40**	**19,44**
Optimum neutre	**11,11**	**9,26**	**11,40**	**19,44**
Optimum anonyme	**11,11**	**9,26**	**11,40**	19,62
Optimum anonyme neutre	—	**9,26**	**11,40**	***23,61***
Optimum approché (glouton)	**11,11**	**9,26**	13,46	21,09
CVTI	**11,11**	**9,26**	12,38	21,35
VTI	**11,11**	**9,26**	12,38	21,35
SE	**11,11**	**9,26**	12,53	21,35
U2T	**11,11**	**9,26**	12,53	21,35
PO	**11,11**	18,52	26,56	21,09
Bald.	**11,11**	17,75	24,26	22,27
CBuck.	**11,11**	18,52	24,98	22,22
CDoy.	**11,11**	20,06	27,31	22,83
Nan.	**11,11**	***20,83***	***27,86***	***22,92***
CDict.	**11,11**	***20,37***	27,28	***23,96***
Kem.	**11,11**	***21,60***	***29,26***	***22,92***
CSD	**11,11**	***21,60***	***29,26***	***22,92***
Sch.	**11,11**	***21,60***	***29,26***	***23,09***
VTID	**11,11**	***22,38***	***30,76***	22,57
CCoo.	**11,11**	***22,38***	***36,16***	22,22
Max.	**11,11**	***21,60***	***29,26***	***24,05***
CUni.	**11,11**	***22,38***	***30,76***	***24,83***
BI	**11,11**	***22,38***	***36,16***	***23,26***
CBor.	**11,11**	***27,01***	***37,76***	***27,78***
Uni.	***16,67***	***26,23***	***33,01***	***30,21***
Buck.	***13,89***	***28,16***	***40,13***	***23,61***
Veto	***26,39***	***32,83***	***35,54***	***50,20***
Coo.	***22,22***	***38,58***	***53,41***	***39,84***
Bor.	***23,61***	***44,37***	***53,36***	***51,39***

TABLE 10.4 – Taux de MC exacts, en pourcents (culture impartiale). On utilise la règle de départage par ordre d'indice, qui est celle implantée dans SVAMP.

Pour $V = 3$ et $C = 4$, nous avons vu que CVTI ne peut pas être optimal, même avec une règle de départage bien choisie. En revanche, tout mode de scrutin optimal parmi les systèmes Condorcet peut être vu comme Condorcet-Bucklin, Condorcet-BI, la méthode de Schulze, PO ou Maximin, muni(e) d'une règle de départage adéquate.

Comme le mode de scrutin Condorcet-doyen est intuitivement le plus simple des modes de scrutin Condorcet (avec Condorcet-dictature), il nous semble fournir une bonne référence pour la manipulabilité : en effet, un mode de scrutin plus manipulable que celui-ci présente la propriété néfaste d'être moins performant que ce mode de scrutin, somme toute, assez peu subtil. À titre indicatif, les taux de manipulabilité supérieurs à celui de Condorcet-doyen sont représentés en caractères italiques et rouges dans la table. Il est notable que de nombreux modes de scrutin (Nanson, Condorcet-dictature, Kemeny, CSD, Schulze, VTID, Condorcet-Coombs, Maximin, Condorcet-Uninominal, Bucklin itéré, Condorcet-Borda, Uninominal, Bucklin, Veto, Coombs et Borda) sont généralement plus manipulables que Condorcet-doyen sur les valeurs considérées. Les seuls modes de scrutin qui ont toujours un taux de manipulabilité plus faible dans cette table sont celui obtenu par l'algorithme glouton, CVTI, VTI, SE, l'uninominal à deux tours, PO, la méthode de Baldwin et Condorcet-Bucklin.

Conclusion et perspectives

À l'heure de conclure ce mémoire, nous souhaitons revenir sur diverses notions rencontrées, avec une attention particulière pour les perspectives qu'elles suggèrent pour de futurs travaux.

Dans l'introduction de ce mémoire, nous avons rappelé ce qui nous semble être les paradoxes fondateurs de la théorie du vote. Au triptyque relativement traditionnel comprenant le paradoxe de Condorcet (1785), le théorème d'Arrow (1950) et le théorème de Gibbard (1973) et Satterthwaite (1975), il nous semble qu'il faut ajouter les divers résultats d'impossibilité sur la comparaison interpersonnelle des utilités, qui limitent une des seules échappatoires envisageables au théorème d'Arrow et, par là même, renforcent le résultat négatif qu'il énonce. Nous avons rappelé que le résultat de Gibbard (1973) ne s'applique pas seulement aux systèmes de vote ordinaux, contrairement à une formulation affaiblie du théorème qui est parfois présentée. Nous avons également évoqué le théorème de Gibbard (1977, 1978) non déterministe, qui caractérise entièrement la non-manipulabilité pour des systèmes de vote utilisant le hasard ; mais nous pensons que les systèmes obtenus (en particulier la *dictature aléatoire*) doivent être limités à des applications où les élections sont fréquentes et où le résultat a un impact modéré. À l'inverse, nos travaux concernent les systèmes de vote déterministes.

Nous avons insisté, comme le fait Gibbard (1973), sur le fait que le problème de fond lié à la manipulabilité est le défaut de *straightforwardness*, c'est-à-dire le fait que les électeurs ne peuvent pas toujours défendre leur opinion de façon optimale sans connaître les bulletins des autres électeurs. Ceci cause de nombreux doutes au moment de voter, une asymétrie *de facto* entre les électeurs informés et ceux qui ne le sont pas, des problèmes de légitimité du résultat et l'octroi d'un pouvoir discutable aux sources d'information comme les instituts de sondage. Mais nous avons aussi rappelé qu'une manipulabilité faible peut être vue comme une *straightforwardness* approchée, et que l'étude de la manipulabilité peut, par conséquent, être abordée comme un moyen technique d'aborder la *straightforwardness*.

Tout en respectant le point de vue qui condamne l'action de *manipulation*, nous avons défendu une vision peut-être moins orthodoxe qui consiste à défendre l'attitude stratégique des électeurs et à condamner la *manipulabilité*, non parce qu'on risque d'obtenir un résultat manipulé, mais parce que, par manque d'information ou de démarche stratégique des électeurs, on risque d'obtenir un résultat sincère qui n'est pas un équilibre de Nash fort (ENF) et qui impliquerait des regrets pour certains électeurs et un questionnement sur la légitimité du résultat. Nous avons ainsi avancé qu'une manipulation facile est plutôt plus souhaitable qu'une manipulation difficile ; mais que la « manipulation » la plus facile survient quand le vote sincère est stratégiquement optimal, c'est-à-dire quand la configuration est non manipulable.

Dans le chapitre 1, nous avons défini notre formalisme, en particulier la notion d'*espace électoral*, qui permet d'étudier en toute généralité des systèmes de vote basés sur des ordres mais aussi sur des valeurs d'approbation, des notes ou n'importe quel autre type d'information. Nous avons donné quelques exemples d'espaces électoraux afin de montrer la richesse d'un tel formalisme mais cette liste est loin d'être limitative. Nous avons défini les *systèmes de vote généraux*, inspirés du formalisme des *game forms* de Gibbard (1973). Nous avons concentré notre étude sur les *systèmes de vote basés sur l'état*, qui permettent de s'affranchir d'une éventuelle définition problématique de vote sincère canonique. En effet, nous avons montré que, dans l'optique de limiter la manipulabilité, on pouvait se limiter aux SVBE, généralisant ainsi un résultat de Moulin (1978) qui s'appliquait aux espaces électoraux ordinaux. Nous avons montré comment on pouvait faire l'aller-retour entre les systèmes généraux et les SVBE par la version basée sur l'état d'un système général et par l'implantation canonique d'un SVBE.

Nous avons profité de la section 1.2.2 pour caractériser les valeurs du couple (V, C) qui permettent d'avoir un système de vote anonyme et neutre dans l'espace électoral des ordres stricts totaux (proposition 1.16), offrant en cela une variation d'un résultat de Moulin (1978) qui traitait les systèmes de vote qui sont, en outre, *efficients*.

Dans la section 1.5, nous avons rappelé les notions usuelles de vainqueur de Condorcet relatif et de vainqueur de Condorcet faible, et nous avons défini les notions de vainqueur de Condorcet absolu et de candidat Condorcet-admissible, dont la suite de ce mémoire a montré la pertinence pour l'étude de la manipulabilité.

Dans la section 1.6, nous avons présenté les modes de scrutin étudiés dans cette thèse. Nous avons notamment défini la *méthode de Bucklin itérée* (BI). Le système *VTI à duels* (VTID), une variante du *vote à tours instantanés* (VTI), est une contribution de Laurent Viennot. Nous avons également défini les systèmes *Condorcet-doyen* et *Condorcet-dictature* qui, par leur simplicité, fournissent de bons exemples ou contre-exemples pour divers résultats théoriques.

Parmi les différentes notions introduites dans le chapitre 1, nous souhaitons revenir une dernière fois sur deux d'entre elles en particulier.

Premièrement, à travers l'anonymat (section 1.2.2), et plus tard le critère du bulletin majoritaire unisson, **BulMajUni** (section 3.1), nous avons vu que l'on pouvait définir des *actions du groupe des permutations sur un espace électoral* et des *isomorphismes de systèmes de vote*. D'une manière générale, il nous semble intéressant de définir des notions qui sont stables par isomorphisme afin d'avoir des propriétés intrinsèques et non liées à l'étiquetage des objets. Plusieurs applications pourraient être obtenues par une telle approche.

1. Par exemple, nous pensons que, pour certains systèmes de vote, mais pas tous, il est possible de définir un vote sincère canonique comme l'éventuelle stratégie qui est optimale dans toute culture neutre, c'est-à-dire stable par une certaine classe d'isomorphismes. Après avoir déminé les difficultés théoriques pour définir rigoureusement cette notion dans un espace électoral quelconque, il serait intéressant d'étudier quels systèmes de vote possèdent un tel vote sincère canonique et pour lesquels ce n'est pas le cas.

2. Par ailleurs, la notion d'isomorphisme pourrait être étendue à une notion plus générale de morphisme de mode de scrutin qui ne soit pas forcément un isomorphisme, c'est-à-dire pas forcément bijectif. Ceci ouvre des perspectives pour examiner des transformations de systèmes de vote ou pour établir des correspondances entre certains systèmes connus, qui pourraient expliquer certaines similitudes de comportement.

Deuxièmement, dans le contexte de ce mémoire, même si nous utilisons le mot *préférence* par concision, le sens exact de la relation binaire $c \, \mathrm{P}_v \, \mathrm{w}$ est le suivant : si un candidat w est vainqueur sincère, alors l'électeur v est susceptible d'agir pour que le candidat c soit vainqueur à la place. En autorisant des préférences qui ne sont pas antisymétriques mais des relations binaires quelconques, nous avons permis de modéliser une variété importante de comportements des électeurs. Nous avons principalement vu des exemples de préférences qui violent la transitivité voire la négative-transitivité (exemples 1.6, 1.7, 1.8). Un autre avantage possible de notre formalisme est de permettre de combiner la manipulabilité au sens habituel (où les électeurs manipulent pour que le résultat soit plus conforme à leurs opinions personnelles) et la corruption (où certains électeurs sont prêts à manipuler dans un sens ou son contraire, selon qu'ils reçoivent ou non un pot-de-vin). Nous n'avons évoqué cette possibilité que rapidement en section 1.3 mais il y a un champ d'étude beaucoup plus vaste dans cette direction. On peut imaginer, par exemple, que pour chaque électeur et chaque couple de candidats (w, c), il existe une somme d'argent minimale à partir de laquelle l'électeur accepte de participer à une manipulation pour c contre w. Ainsi, la matrice de sa relation binaire est remplacée par une matrice de coûts réels positifs ou nuls, où $+\infty$ correspond à **False** (jamais intéressé) et 0 correspond à **True** (toujours intéressé, même sans incitation financière). On peut, ensuite, examiner avec quel budget total il est possible de faire gagner un candidat donné, comme il est d'usage dans les problèmes de corruption.

Dans le chapitre 2, nous avons défini le *critère de la coalition majoritaire informée* (**CMInf**), qui est le plus faible des critères majoritaires que nous avons défini par la suite. Nous avons prouvé les théorèmes faible et fort de condorcification 2.9 et 2.20 : si un système de vote vérifie **CMInf**, alors sa condorcification est au plus aussi manipulable. Si, en outre, il ne vérifie pas le *critère de Condorcet-résistant* (**rCond**), alors sa condorcification est strictement moins manipulable. Comme nous l'avons vu, la version forte du théorème s'applique à une grande partie des systèmes de vote usuels.

Pour démontrer le théorème fort de condorcification 2.20, nous avons introduit et caractérisé la notion de vainqueur de Condorcet résistant. Cette notion fournit une borne supérieure (ensembliste) de manipulabilité pour les systèmes Condorcet, dont nous avons montré qu'elle était ajustée pour $C \geq 6$, mais dont nous avons vu plus tard qu'elle ne l'était pas pour $C = 3$ (chapitre 10). Il serait intéressant d'étudier les valeurs de C intermédiaires afin de disposer également d'une borne supérieure ajustée dans ce cas.

Nous avons montré que, dans le cas général, c'est la condorcification *absolue* qui permet de faire fonctionner les théorèmes de condorcification. Cependant, nous avons également vu qu'une condorcification *relative* peut, pour l'uninominal à un ou deux tours, diminuer la manipulabilité, y compris par rapport à la condorcification absolue. En revanche, la condorcification relative ne permet pas de diminuer la manipulabilité de VTI, du jugement majoritaire, du vote par assentiment ou du vote par notation. Il serait intéressant d'élargir cette investigation à l'ensemble des systèmes de vote usuels et si possible d'identifier une condition (s'appliquant à autant de systèmes que possible) suffisante pour que la condorcification relative soit au plus aussi manipulable que le système de vote initial, voire au plus aussi manipulable que la condorcification absolue.

Comme nous l'avons vu plus tard dans les simulations (chapitres 7 à 10), la diminution de manipulabilité offerte par la condorcification n'est pas nécessairement très importante quantitativement, notamment lorsque l'on compare VTI et sa condorcification CVTI. Il nous semble que les conséquences principales de ces

théorèmes de condorcification sont plutôt à voir dans les corollaires 2.22 et 2.23 : pour la recherche d'un système de vote de manipulabilité minimale au sein de la classe **CMInf**, on peut restreindre l'étude aux systèmes Condorcet.

Dans le chapitre 3, nous avons établi une hiérarchie de divers critères majoritaires pour un système de vote et nous avons montré leurs liens avec l'existence et une certaine forme d'unicité des équilibres de Nash forts. Nous avons notamment étendu des résultats de Sertel et Sanver (2004) et Brill et Conitzer (2015) en montrant que le critère **CMInf** est équivalent au *critère de restriction des ENF possibles aux candidats Condorcet-admissibles*, **RENFA**, selon lequel tout vainqueur d'un ENF est un candidat Condorcet-admissible (« unicité » de l'équilibre). Nous avons également étudié le *critère d'existence d'ENF pour un vainqueur de Condorcet*, **XENFC**, selon lequel il existe un ENF dans toute configuration de préférence Condorcet (existence d'un équilibre). Nous avons montré que la classe **XENFC** est incluse dans la classe **CMIgn** des systèmes vérifiant le *critère de la coalition majoritaire ignorante* et contient la classe **BulMaj** des systèmes vérifiant le *critère du bulletin majoritaire* (dans les espaces électoraux raisonnables) et que ces inclusions sont strictes en général. Il serait intéressant, à l'avenir, de posséder une caractérisation simple des systèmes de vote qui vérifient **XENFC**.

Par ailleurs, nous avons montré que le *critère de restriction des ENF possibles aux vainqueurs de Condorcet*, **RENFC**, et le *critère d'existence d'ENF pour un candidat Condorcet-admissible*, **XENFA**, sont moins « naturels » que leurs versions faibles, **RENFA** et **XENFC** respectivement, notamment parce qu'ils sont vérifiés par peu de systèmes de vote usuels, parce qu'ils n'entretiennent pas de relations d'inclusion simple avec les autres critères et parce que ces deux critères sont généralement incompatibles entre eux.

Nous avons examiné les critères vérifiés par les systèmes de vote usuels, en compilant des résultats classiques de la littérature et en présentant des résultats originaux, notamment, d'une part, sur les critères **CMInf** et **rCond** et, d'autre part, sur tous les critères pour la méthode de Bucklin itérée. Malgré nos efforts, quelques questions restent ouvertes : trouver une condition nécessaire et suffisante simple sous laquelle une règle positionnelle à score (RPS) vérifie **CMInf** et des conditions sous lesquelles les RPS itérées à élimination simple (RPSI-ES) et les RPSI itérées à élimination par la moyenne (RPSI-EM) vérifient **rCond**.

Dans la section 3.5, nous avons commenté de façon informelle les liens entre les critères majoritaires et la notion d'échange d'information, notamment pour atteindre des équilibres ou, en tout cas, pour trouver le même candidat vainqueur que dans un équilibre. Il y a, là encore, tout un champ d'investigation possible, en lien avec le domaine de l'algorithmique distribuée. Par exemple, on peut poser la question suivante : étant donné un mode de scrutin, existe-t-il un algorithme qui permette d'atteindre un ENF quand il en existe ? En exécution synchrone sans panne, c'est évident, puisqu'on peut émuler un comportement non distribué ; mais qu'en est-il en exécution asynchrone ? Est-il possible d'établir un algorithme qui soit en un sens résistant aux pannes, c'est-à-dire, dans ce contexte, à la non-collaboration de certains électeurs ? Et surtout, quelle est la meilleure complexité en échange d'information qu'un tel algorithme puisse avoir ? En pire cas ? En moyenne ?

Dans le chapitre 4, nous avons défini la *condorcification généralisée* en nous reposant sur la notion de *famille* qui est inspirée de la théorie des jeux simples. Ceci nous a permis, notamment, d'appliquer une transformation inspirée de la condorcification à des modes de scrutin qui ne vérifient pas de critère majoritaire particulier (théorème de condorcification généralisée 4.18). Nous avons comparé la manipulabilité des condorcifications effectuées avec deux familles différentes

(théorème de condorcification comparée 4.21). Enfin, nous avons défini la *famille maximale* d'un système de vote, c'est-à-dire la famille des coalitions qui peuvent manipuler de manière informée pour tel ou tel candidat. Nous avons montré que, sous certaines hypothèses, la condorcification utilisant cette famille maximale est la moins manipulable des condorcifications réalisées avec une famille $\mathcal{M}$ telle que le système vérifie $\mathcal{M}$**CInf** (théorème de condorcification maximale 4.25). Nous avons ainsi pu établir que, pour les systèmes usuels qui vérifient **CMInf**, la condorcification majoritaire est la condorcification maximale, et donc qu'elle est optimale en un certain sens. En particulier, nous avons montré que (pour un nombre d'électeurs impair et des préférences antisymétriques), les systèmes Condorcet sont leur propre condorcification maximale. Nous avons aussi donné des exemples de condorcification généralisée pour divers systèmes de vote violant de façon non marginale l'anonymat et/ou la neutralité, qui ont permis de montrer le large champ d'application des théorèmes de ce chapitre.

Pour étendre ces résultats, l'idéal serait de parvenir à un théorème de la forme suivante : la condorcification maximale du système f est le système de vote le moins manipulable (au sens large) parmi ceux qui partagent une certaine propriété bien choisie avec f. Le théorème de condorcification maximale dont nous disposons est, en quelque sorte, de cette forme, mais la propriété commune est d'être une condorcification généralisée de f par une famille $\mathcal{M}$ telle que f vérifie $\mathcal{M}$**CInf**. Il serait intéressant d'avoir une propriété plus simple et plus intuitive.

Par ailleurs, les familles que nous considérons sont indexées par un certain candidat c et décrivent les coalitions qui permettent de faire gagner c (en général de manière informée). Cette façon de faire présente l'avantage de permettre une traduction directe du raisonnement effectué pour le théorème faible de condorcification 2.9 vers le théorème de condorcification généralisée 4.18. Mais on pourrait également envisager de généraliser ce *vecteur* de collections de coalitions, indexées par c, à une *matrice* de collections de coalitions, indexées par le vainqueur à détrôner w et le candidat pour lequel on souhaite manipuler c.

Dans le chapitre 5, nous avons montré le théorème de tranchage 5.9 : si la culture est *décomposable*, alors pour tout système de vote, il existe une de ses tranches qui est au plus aussi manipulable que le système original (au sens probabiliste). En particulier, nous avons montré que c'est le cas lorsque les électeurs sont indépendants en prouvant un résultat plus général de probabilité, le lemme de la variable aléatoire complémentaire 5.16. Nous avons examiné les généralisations possibles de ce théorème et montré qu'on ne peut pas exiger une diminution de manipulabilité au sens ensembliste et qu'il est impossible de supprimer purement et simplement l'hypothèse de décomposabilité. La principale question ouverte que nous avons laissée est de savoir si on peut donner une hypothèse plus faible que la décomposabilité. En particulier, il serait intéressant de savoir si la condition (5.1) de la proposition 5.8, qui est nécessaire mais pas suffisante pour assurer la décomposabilité, est tout de même suffisante pour impliquer les mêmes conclusions que le théorème de tranchage. L'idéal serait même d'avoir un version ajustée du théorème, c'est-à-dire de trouver une condition sur la culture, non seulement suffisante, mais aussi nécessaire, pour que tout SVBE admette une tranche au plus aussi manipulable que le système original.

Le théorème de tranchage 5.9 n'est pas constructif, ce qui ne lui donne pas une application pratique immédiate. Mais, à notre avis, sa conséquence principale est le théorème d'optimalité 5.15 : si la culture est décomposable, alors un système qui est optimal (au sens probabiliste) parmi les systèmes de vote Condorcet ordinaux est également optimal dans la classe beaucoup plus large des systèmes qui

vérifient **CMInf** et qui peuvent ne pas être ordinaux. Si l'on se donne pour objectif de diminuer la manipulabilité, cela signifie que des systèmes comme le vote par notation, le jugement majoritaire ou le vote par assentiment sont *a priori* peu prometteurs, avant même d'examiner leurs résultats dans les simulations. Ceci dit, si on se donne pour objectif une certaine simplicité algorithmique dans l'identification d'un vote stratégique, le vote par assentiment présente des propriétés intéressantes, que nous avons déjà évoquées.

Dans le chapitre 6, nous avons présenté SWAMP, *Simulator of Various Voting Algorithms in Manipulating Populations*, un package Python dédié à l'étude des systèmes de vote et en particulier de leur manipulabilité. Les populations d'électeurs peuvent être caractérisées par des ordres stricts faibles ou par des utilités ; il est possible de les importer à partir de fichiers externes ou de les générer par une variété de modèles aléatoires. Ce simulateur implante la manipulation par coalition (MC) et diverses variantes (la manipulation par coalition informée MCI, la manipulation-unisson MU et la manipulation triviale MT), ainsi que la manipulation individuelle, l'indépendance des alternatives non pertinentes (IANP) et les notions Condorcet présentées dans ce mémoire.

Des méthodes génériques permettent d'implanter rapidement de nouveaux systèmes de vote et des algorithmes dédiés sont implantés pour certains d'entre eux. Nous avons réuni des algorithmes correspondant à l'état de l'art et nous avons développé des méthodes originales, notamment pour VTI et ses variantes, étant donné l'intérêt particulier de ces systèmes de vote pour obtenir une faible manipulabilité.

Un tel logiciel est par nature un chantier en perpétuelle évolution, où des développements sont toujours possibles. Concernant les populations, il serait envisageable d'implanter des modèles plus généraux, en particulier des préférences non transitives. Nous prévoyons également d'implanter les modèles de Mallows (1957) et de Pólya-Eggenberger (cf. Lepelley et Valognes, 2003), afin de comparer les résultats avec ceux obtenus dans le modèle de VMF. Pour les populations issues de données réelles, nous souhaitons aussi implanter des modules qui permettent de tester si une population donnée est proche d'un modèle donné (*single-peaked*, VMF, etc.) afin de tester le degré de réalisme de ces hypothèses.

Il serait également intéressant de permettre de faire varier la règle de départage utilisée pour chaque système de vote, mais il faut être conscient que cette possibilité entraîne des difficultés non négligeables pour les algorithmes de manipulation. Enfin, la principale source d'amélioration serait d'implanter des algorithmes de manipulation dédiés pour les systèmes de vote qui n'en possèdent pas encore et utilisent actuellement les algorithmes génériques. En ce qui concerne VTI, la principale amélioration serait de trouver une heuristique polynomiale qui permette de certifier la non-manipulabilité dans une proportion élevée des configurations non-manipulables, de la même façon que notre heuristique permet d'identifier une proportion importante des configurations manipulables.

Étant donné les bonnes performances de VTI en termes de manipulation, il serait intéressant d'implanter diverses variantes de VTI dans SWAMP afin de comparer leur manipulabilité. Nous avons déjà étudié VTID et bien sûr CVTI. À l'heure actuelle, deux autres variantes de VTI sont implantées dans SWAMP : VTIM, qui est la RPSI-EM associée à l'uninominal (sous le nom anglophone *IRV-Average*), ainsi qu'un autre système nommé *Instant-Condorcet Runoff Voting* (ICRV). Nous avons préféré ne pas inclure ces systèmes dans notre étude pour la raison suivante : comme ces systèmes sont généralement très peu MT (comme VTI), l'algorithme générique pour la MC donne généralement une incertitude algorithmique très élevée, ce qui ne permet pas de comparer la manipulabilité de ces

systèmes aux autres. À l'avenir, il serait intéressant d'implanter des algorithmes dédiés pour ces systèmes et d'ajouter d'autres variantes de VTI, par exemple celles mentionnées par Green-Armytage (2011), dont ICRV fait partie.

Un des objectifs de SVAMP à terme est de mesurer finement la manipulabilité, non plus d'une culture, mais d'un profil donné. Nous avons développé, pour cela, une méthodologie par ajout de bruit aléatoire mais il serait intéressant de développer d'autres métriques. Par exemple, une famille d'indicateurs de manipulabilité proposée dans la littérature se base sur la taille des coalitions (ce qui est déjà partiellement implanté dans SVAMP). Une autre approche nous semble intéressante, inspirée des problématiques de corruption et des réflexions autour de l'analyse des expériences réelles : la notion de *seuil de manipulation*. Pour chaque candidat c différent du vainqueur w et pour chaque réel ε, on peut se poser la question suivante : en considérant les électeurs pour lesquels c apporte une amélioration d'utilité supérieure à ε par rapport à w, est-ce que cette coalition est capable de faire gagner c ? En autorisant ε négatif, la coalition peut contenir tous les électeurs, donc il existe un seuil ε en-deçà duquel c'est effectivement possible [5]. Ce seuil contient non seulement l'information de la manipulabilité en faveur de c (qui est vraie ssi le seuil est positif) mais il permet également de mesurer la facilité de manipulation ou une forme de distance à une configuration manipulable.

Notons que SVAMP est conçu pour réaliser une large gamme de séries de simulation, dont seule une partie a été exploitée dans le présent mémoire. Les possibilités les plus importantes, que nous avons passées sous silence afin de nous consacrer à la manipulation par coalition, concernent la manipulation individuelle. En effet, des algorithmes dédiés sont implantés pour presque autant de systèmes de vote que dans le cas de la manipulation par coalition : c'est le cas pour le jugement majoritaire, Maximin, l'uninominal, l'uninominal à deux tours (U2T), Veto, le vote par assentiment, le vote par notation, le scrutin exhaustif (SE), VTI ainsi que les méthodes de Borda, Bucklin, Coombs et Schulze. Il serait donc intéressant d'utiliser SVAMP pour réaliser une étude de la manipulation individuelle.

Dans les chapitres 7 et 8, nous avons utilisé SVAMP pour étudier la manipulabilité de divers systèmes de vote dans des cultures sphéroïdales, en utilisant pour la première fois le modèle de Von Mises-Fisher, puis dans des cultures basées sur un spectre politique. Cela nous a permis de vérifier certains résultats connus dans la littérature et de proposer un certain nombre de conjectures sur la monotonie et la limite du taux de manipulabilité en fonction de certains paramètres, en particulier C et V. Tandis que nous avons prouvé certaines d'entre elles, d'autres restent des conjectures.

Parmi ces conjectures, celle qui nous semble le plus accessible est que le taux de manipulation des systèmes de vote raisonnables (**CMInf**) tend vers 1 quand $C \to \infty$ en culture sphérique (conjecture 7.4). Tandis que nous avons démontré ce résultat pour un nombre impair d'électeurs, il suffirait, pour étendre le résultat à V pair, de prouver qu'en culture impartiale, la probabilité d'avoir un candidat Condorcet-admissible (c'est-à-dire un vainqueur de Condorcet faible) tend vers 0 quand $C \to \infty$. Par ailleurs, nous avons prouvé que le taux de MC de Veto (muni de la règle de départage utilisée dans SVAMP) ne tend pas vers 1 quand $C \to +\infty$ en culture impartiale, même si les simulations suggèrent une limite proche de 1.

Il semble également intéressant d'examiner les taux limites de manipulabilité en culture sphérique quand le nombre d'électeurs est très grand (conjectures 7.7 et 7.8). Même si la culture sphérique n'a pas vocation à être un modèle descriptif,

5. On fait ici l'hypothèse implicite que le système de vote considéré est surjectif, c'est-à-dire que tout candidat est effectivement éligible. Cette condition est vérifiée par tous les systèmes de votes raisonnables (hormis éventuellement Veto si $C > V + 1$).

elle propose un scénario de référence normatif qui peut être vu comme un pire cas, puisque c'est le plus désordonné (l'entropie y est maximale). Ainsi, il est intéressant de savoir quels systèmes de vote n'ont pas un taux de manipulabilité tendant vers 1 quand $V \to \infty$ dans cette culture *a priori* plutôt défavorable [6]. Nous avons conjecturé que, parmi les systèmes de vote étudiés, seuls VTI, SE, CVTI et Veto ont un taux de MC qui tend vers une limite différente de 1 quand $V \to +\infty$ en culture impartiale.

Nous avons observé des phénomènes oscillatoires pour les taux de MC, qui sont plus prononcés en culture de spectre politique unidimensionnel qu'en culture sphéroïdales, et nous en avons proposé une explication qualitative. Pour le jugement majoritaire et la méthode de Bucklin, nous avons avancé l'hypothèse que ce comportement est amplifié par le fait de considérer une médiane défavorable, qui tend à rendre plus manipulables les cas où le nombre d'électeurs est pair. Pour de futurs travaux, il serait intéressant d'examiner des variantes de ces systèmes de vote où on utilise la médiane favorable pour voir si cela permet de diminuer la manipulabilité dans ce cas.

Quand la population tend à s'uniformiser (sections 7.5, 7.6, 7.8 et 8.3), nous avons constaté, sans surprise, que la plupart des systèmes de vote deviennent moins manipulables. Cependant, nous avons vu que certains systèmes sont nettement moins réactifs en termes de baisse de la manipulabilité dans ce cas, notamment Veto, le vote par assentiment, le vote par notation, la méthode de Borda et le jugement majoritaire. En ce qui concerne Veto (muni de la règle de départage utilisée dans SVAMP), nous avons même montré que sa manipulabilité ne tend pas vers 0 quand les électeurs ont des préférences identiques, contrairement à tous les systèmes de vote unanimes.

Par ailleurs, nous avons confirmé l'importance du vainqueur de Condorcet résistant pour la non-manipulabilité des systèmes de vote Condorcet, notamment grâce à la borne supérieure donnée en section 2.8. Pour cette raison, il serait intéressant de prouver la décroissance de sa probabilité d'existence en fonction de C en culture sphérique. De manière générale, il serait intéressant d'étudier la probabilité d'existence d'un vainqueur de Condorcet résistant dans divers contextes, comme l'ont fait Gehrlein (2006) et d'autres pour un vainqueur de Condorcet ou un vainqueur de Condorcet faible.

Enfin, ces simulations ont permis de donner une bonne vue d'ensemble des performances comparées des différents systèmes de vote. Dans les cultures sphéroïdales et dans les modèles de spectre politique multidimensionnel, CVTI, VTI et SE présentent les meilleures performances. Dans les cultures de spectre politique unidimensionnel, la méthode de Schulze, Maximin, l'uninominal à deux tours, la méthode de Bucklin et VTID présentent des résultats prometteurs, qui encouragent à approfondir l'étude de ces systèmes à l'avenir. De manière générale, le vote par notation, le vote par assentiment et la méthode de Borda sont généralement très manipulables ; en particulier, ils sont souvent plus manipulables que tous les systèmes Condorcet, comme on peut le constater en comparant avec le taux d'existence d'un vainqueur de Condorcet résistant.

Dans le chapitre 9, nous avons analysé les résultats de 168 expériences réelles, dont 17 proviennent d'expériences originales et les autres de la base de données PrefLib. Parmi ces jeux de données, il est notamment intéressant de considérer :

— ceux qui sont basés sur des données cardinales car elles permettent d'approfondir la comparaison entre les systèmes ordinaux et cardinaux et elles permettent aussi de considérer naturellement des problèmes basées sur la notion d'utilité, comme le seuil de manipulabilité décrit ci-dessus ;

6. À propos de cette idée, on pourra notamment consulter Tsetlin *et al.* (2003).

— ceux qui ne concernent pas un scénario d'élection et dont l'enjeu stratégique est *a priori* faible (comme les expériences bdtheque), car ce sont des jeux de données où l'on peut espérer avoir des opinions aussi sincères que possibles.

Pour des travaux futurs, il serait intéressant d'élargir ce corpus autant que possible.

Dans toutes les expériences analysées, CVTI, VTI et SE se distinguent par une MC toujours plus faible que celle des autres systèmes de vote. En MT, le même résultat est valable pour CVTI et presque toujours pour VTI et SE. Or nous avons vu, dans le chapitre 8, que les performances de ces systèmes de vote sont moins bonnes (comparées aux autres mode de scrutin) dans les cultures basées sur un spectre politique unidimensionnel que dans les autres. On voit donc une certaine opposition entre les performances des systèmes de type VTI dans nos jeux de données réelles et dans des spectres unidimensionnels. Il serait intéressant de réaliser davantage d'expériences sur des élections politiques à large échelle afin de tester la résistance de CVTI, VTI et SE à la manipulation dans ces cas d'application *a priori* plutôt défavorables et de vérifier à quel point l'hypothèse de spectre unidimensionnel se vérifie, en particulier du point de vue de la manipulabilité.

Il est troublant que parmi les systèmes de vote étudiés, les moins manipulables en pratique, c'est-à-dire VTI et ses variantes, soient également des systèmes qui violent d'autres propriétés importantes, en particulier la monotonie. Ceci amène à deux questions qui sont liées entre elles.

1. Est-il possible de définir la « monotonisation » d'un système de vote, comme nous avons défini la condorcification, de façon à obtenir un système monotone (et conservant certaines autres propriétés raisonnables, par exemple **CMInf**) sans rendre le système plus manipulable ? D'un point de vue technique, définir une telle monotonisation est loin d'être immédiat.
2. Y a-t-il obligatoirement un compromis à faire entre la faible manipulabilité et la monotonie ? On peut avancer l'argument qu'un système Condorcet, par exemple CVTI, est déjà monotone sur les configurations non-manipulables, puisque celles-ci sont incluses dans les configurations non-Condorcet.

Cette dernière remarque prend tout son sens quand on constate, comme nous l'avons fait, qu'il est très fréquent expérimentalement d'avoir un vainqueur de Condorcet (rejoignant, en cela, divers résultats antérieurs). En pratique, ceci limite la gravité du paradoxe de Condorcet, du théorème d'Arrow et des problèmes de non-monotonie. En effet, il existe une façon canonique de synthétiser les préférences des électeurs de manière anonyme, neutre et possédant de bonnes propriétés de monotonie : il s'agit de la relation de victoire dans la matrice des duels ; le seul défaut de cette relation est qu'elle n'est pas nécessairement transitive. Mais, dans les expériences, on constate qu'elle possède souvent un élément maximal (le vainqueur de Condorcet), ce qui en fait un vainqueur naturel et possédant de bonnes propriétés comme l'IANP. Pour cette raison, nous avançons l'idée que le comportement d'un système de vote dans les configurations non-Condorcet ne doit pas être examiné à l'aune de la pertinence du candidat élu quand une telle configuration survient en vote sincère (car il n'y a, dans ce cas, pas de solution parfaite), mais plutôt par l'impact qu'elle a sur les possibilités de manipulation des configurations Condorcet.

C'est précisément ce que nous avons fait dans le chapitre 10, où nous avons introduit le *graphe d'opportunité* d'un espace électoral et nous l'avons utilisé pour étudier les systèmes de vote optimaux, c'est-à-dire dont le taux de manipulabilité est minimal au sein d'une certaine classe. Nous avons considéré la classe **CMInf**

et nous avons donc pu nous restreindre aux systèmes vérifiant **Cond** grâce aux théorèmes de condorcification, mais le raisonnement effectué est valable en toute généralité.

Sur cet objet, il se pose plusieurs questions théoriques. Sur l'ensemble des multigraphes étiquetés et pondérés comme le graphe type de la figure 10.1, est-il vrai que résoudre le problème de la contamination minimale est $\mathcal{NP}$-difficile ? Si on se restreint aux graphes d'opportunité obtenus pour une certaine classe d'espaces électoraux, par exemple ceux des ordres stricts totaux pour toutes les valeurs de C et les valeurs impaires de V, est-ce que le problème reste $\mathcal{NP}$-difficile ? Peut-on élaborer un algorithme polynomial qui permette une meilleure approximation que l'algorithme glouton que nous avons présenté ?

Dans le cas de $C = 3$ candidats et pour $V = 3, 5$ ou 7 électeurs, nous avons constaté que CVTI, à condition d'être muni d'une règle de départage adéquate, était toujours optimal. Il est donc naturel de se demander s'il en est de même pour tout nombre d'électeurs et $C = 3$. Pour $C = 4$ candidats, nous savons d'ores et déjà que ce n'est pas le cas en général, puisque nous établi que ce n'est pas vrai pour $V = 3$ et $C = 4$.

Il serait intéressant d'étendre la recherche de systèmes optimaux à d'autres valeurs des paramètres. À nos yeux, le but principal de cette démarche n'est pas forcément d'utiliser ces systèmes en pratique : on ne pourra que rarement garantir à l'avance qu'on aura un nombre d'électeurs et un nombre de candidats pour lesquels on sait résoudre le problème. Il s'agit plutôt d'utiliser les quelques cas où l'on peut exhiber les systèmes optimaux pour mieux observer comment fonctionne la manipulabilité et mieux comprendre ce qui peut faire qu'un système est peu manipulable.

En résumé, nous avons introduit un formalisme unifié pour étudier les systèmes de vote ordinaux et non ordinaux. Nous avons proposé divers outils pour transformer un système de vote afin d'obtenir une manipulabilité plus faible : la réduction d'un système de vote à sa version basée sur l'état, la condorcification, la condorcification relative dans certains cas, la condorcification généralisée et le tranchage. Ces outils ont permis, notamment, d'établir des théorèmes d'optimalité qui suggèrent d'approfondir les recherches sur les systèmes ordinaux Condorcet. Par ailleurs, nous avons établi divers critères applicables à un système de vote et montré les liens profonds qui existent entre ces critères, la manipulabilité, les équilibres de Nash et la notion d'échange d'information. Nous avons posé les premières pierres de l'édifice des systèmes de vote Condorcet optimaux, permettant ainsi de disposer d'un mètre étalon pour la manipulabilité pour des valeurs modérées des paramètres et d'offrir une meilleure intuition des raisons qui permettent à un système d'être peu manipulable. En cela, nous avons accompli quelques pas sur le chemin nous conduisant « vers des modes de scrutin moins manipulables ».

Grâce à notre package logiciel SVAMP, nous avons également pu étudier la manipulabilité d'un point de vue quantitatif et proposer quelques réponses aux problématiques posées en introduction. En particulier : oui, la manipulabilité est un phénomène fréquent dans les cultures artificielles et dans le monde réel, ce n'est pas seulement une possibilité théorique. Elle pose des problèmes qui s'avèrent beaucoup plus fréquents que le paradoxe de Condorcet. Nous avons également pu mesurer les vulnérabilités très différentes des divers systèmes de vote à ce phénomène, et nous avons notamment constaté la spectaculaire suprématie de CVTI, VTI et SE dans toutes les expériences réelles étudiées, ce qui encourage à poursuivre l'étude d'autres systèmes de vote de la même famille.

Tous ces résultats n'ont pas forcément vocation à être utilisés immédiatement dans des élections à grande échelle, même si, dans le cas de VTI, on dispose d'un

système qui est déjà utilisé dans divers pays pour des élections politiques. Les progrès du vote électronique, notamment en matière de sécurité, permettront bientôt d'envisager l'utilisation de systèmes de vote au dépouillement plus complexe que ne le permet le vote manuel ; quand ce sera le cas, il nous semble important que les théoriciens du choix social soient capables d'offrir une vision aussi complète que possible des avantages et des inconvénients de chaque système afin de permettre des décisions éclairées. Par ailleurs, la profusion de structures associatives, professionnelles, de fondations, d'organisations sur Internet se jouant des frontières, offre un formidable terrain d'expérimentation qui permet à des groupes humains réels de bénéficier des résultats les plus récents du choix social et, éventuellement, de transmettre leur expérience en la matière à toutes les formes d'organisation humaine recherchant à améliorer leur fonctionnement démocratique.

Annexes

Annexe A

Espaces probabilisables et mesurabilité

Dans cette annexe, nous examinons la question de la *mesurabilité*, qui est sous-entendue dans toutes les notions probabilistes (par exemple la définition du taux de manipulabilité) et qui essentiellement utilisée dans le chapitre 5 sur le tranchage.

Afin de définir les notions probabilistes de manière rigoureuse, nous considérons des ensembles mesurables, qui sont constitués d'un ensemble E et d'une tribu Σ_E sur E. Un tel ensemble mesurable est noté (E, Σ_E), ou simplement E quand il n'y a pas d'ambiguïté. Un élément de Σ_E est appelé un *événement* sur E.

Nous munissons toujours l'ensemble $\mathcal{R}_\mathcal{C}$ des relations binaires sur $\mathcal{C}$ de sa tribu discrète, que nous notons $\Sigma_{\mathcal{R}_\mathcal{C}}$.

Quand nous considérons un produit cartésien E d'ensembles mesurables (E_v, Σ_{E_v}), nous le munissons toujours de sa tribu produit. En particulier, l'ensemble $\mathcal{R} = (\mathcal{R}_\mathcal{C})^V$ est muni de la tribu produit $\Sigma_\mathcal{R} = (\Sigma_{\mathcal{R}_\mathcal{C}})^V$, qui est simplement sa tribu discrète.

Nous disons que Ω est un *espace électoral mesurable* ssi chaque Ω_v est muni d'une tribu Σ_{Ω_v} et si chaque fonction $\mathrm{P}_v : \Omega_v \to \mathcal{R}_\mathcal{C}$ est mesurable. Quand nous considérons un espace électoral probabilisé (EEP), nous faisons toujours l'hypothèse préalable implicite qu'il s'agit d'un espace électoral mesurable.

Par exemple, considérons l'espace électoral de référence, où $\omega_v = (p_v, u_v, a_v)$. Munissons chaque Ω_v de la tribu produit de la tribu discrète sur $\mathcal{R}_\mathcal{C}$, la tribu de Lebesgue sur $[0,1]^C$ et la tribu discrète sur $\{0,1\}^C$. Alors chaque fonction P_v est manifestement mesurable. Par conséquent, Ω est un espace électoral mesurable.

Dans le contexte du chapitre 5, chaque ensemble $\mathcal{Y}_v$ des méthodes de tranchage est doté d'une tribu canonique. En effet, c'est le cas pour l'ensemble ${\Omega_v}^{\mathrm{P}_v(\Omega_v)}$ des fonctions de $\mathrm{P}_v(\Omega_v)$ vers Ω_v : en associant chaque fonction à la liste de ses valeurs, il suffit de considérer la tribu produit $\Sigma_{\Omega_v} \times \ldots \times \Sigma_{\Omega_v}$, avec un nombre de facteurs égal au cardinal de $\mathrm{P}_v(\Omega_v)$. L'espace $\mathcal{Y}_v$, en tant que sous-ensemble de ${\Omega_v}^{\mathrm{P}_v(\Omega_v)}$, hérite de cette tribu.

Le lemme suivant résout les questions de mesurabilité pour le théorème de tranchage 5.9 et le théorème d'optimalité 5.15.

Lemme A.1
Soit Ω un espace électoral mesurable, E un ensemble mesurable et $g : \Omega \to E$.

On suppose que g ne dépend que des relations binaires de préférence :

$$\forall(\omega,\psi)\in\Omega^2, \mathrm{P}(\omega)=\mathrm{P}(\psi)\Rightarrow g(\omega)=g(\psi).$$

Alors g est mesurable.

Démonstration. Puisque g ne dépend que de $\mathrm{P}(\omega)$, on peut définir $h:(\mathcal{R}_{\mathcal{C}})^V\to E$ de telle sorte que $g=h\circ\mathrm{P}$. Puisque $(\mathcal{R}_{\mathcal{C}})^V$ est muni de la mesure discrète, h est mesurable ; et par définition d'un espace électoral mesurable, P est mesurable. Par conséquent, $g=h\circ\mathrm{P}$ est mesurable. □

Dans le théorème de tranchage 5.9, ce lemme assure que, pour tout y, l'indicatrice de manipulabilité MC_{f_y} est mesurable. Par conséquent, $\tau^{\pi}_{\mathrm{MC}}(f_y)$ est bien défini.

De même, dans le théorème d'optimalité 5.15, ce lemme assure que $\mathrm{MC}_{g\circ\mathrm{P}}$ est mesurable. Par conséquent, $\tau^{\pi}_{\mathrm{MC}}(g\circ\mathrm{P})$ est bien défini.

Annexe B

Géométrie de l'espace des utilités

Dans cette annexe, qui résulte d'une collaboration avec Benoît Kloeckner, nous examinons les propriétés géométriques de l'espace des utilités espérées sur un nombre fini d'options, qui est couramment utilisé pour modéliser les préférences d'un agent en situation d'incertitude. Nous nous concentrons sur le cas où le modèle est neutre vis-à-vis des options disponibles, c'est-à-dire qu'il les traite, *a priori*, comme symétriques les unes des autres. En particulier, nous prouvons que la seule métrique riemannienne qui respecte les propriétés géométriques et les symétries naturelles de l'espace des utilités est la métrique ronde. Cette métrique canonique nous permet de définir une probabilité uniforme sur l'espace des utilités et de généraliser la culture impartiale à un modèle impliquant des utilités espérées. De manière générale, elle est à la base de toutes les cultures dites *sphéroïdales* que nous étudions au chapitre 7. Cette annexe reprend et développe les idées exposées par Durand *et al.* (2015).

Dans la littérature traditionnelle du choix social arrovien (Arrow, 1950), et particulièrement en théorie du vote, les préférences d'un agent sont souvent représentées par de l'information ordinale uniquement : un ordre strict total sur les options disponibles, ou parfois une relation binaire de préférence qui peut ne pas être un ordre strict total (par exemple, si l'indifférence est autorisée). Cependant, il peut être intéressant de considérer des préférences cardinales, au moins pour deux raisons.

D'une part, certains systèmes de vote ne sont pas seulement basés sur de l'information ordinale, comme le vote par assentiment ou le vote par notation.

D'autre part, les électeurs peuvent être dans une situation d'incertitude, soit parce que la règle du système de vote implique une part de hasard, soit parce que chaque électeur possède une information incomplète sur les préférences des autres électeurs et les bulletins qu'ils s'apprêtent à choisir. Afin d'exprimer les préférences dans une situation d'incertitude, un modèle classique et élégant est celui des utilités espérées (Von Neumann et Morgenstern, 1944; Fishburn, 1970; Kreps, 1990; Mas-Colell *et al.*, 1995). Pour chaque agent, une utilité numérique est associée à chaque option ; l'utilité d'une loterie sur les options est calculée comme une espérance. Par conséquent, le vecteur d'utilité $\mathbf{u}$ représentant les préférences d'un agent est un élément de $\mathbb{R}^C$, où C est le nombre d'options ou *candidats*.

Pour un large spectre d'applications en économie, les options considérées sont des récompenses financières ou des quantités d'un ou plusieurs biens économiques, ce qui a une conséquence importante : il existe une structure naturelle sur l'espace

des options. Par exemple, si les options sont des récompenses financières, alors, il y a un ordre naturel sur les montants possibles, qui est défini avant les préférences des agents.

Nous considérons le scénario opposé, où les options sont symétriques *a priori*. Cette hypothèse de symétrie est particulièrement pertinente en théorie du vote, par un principe normatif de neutralité, mais elle peut aussi s'appliquer à d'autres contextes de choix en situation d'incertitude quand il n'y a pas de structure naturelle préexistante sur l'espace des options disponibles.

La motivation de ce chapitre est venue des généralisations possibles pour la culture impartiale à des agents ayant des utilités espérées. La culture impartiale est un modèle probabiliste classique en théorie du vote où chaque agent tire indépendamment son ordre strict total de préférence avec une probabilité uniforme sur l'ensemble des ordres possibles.

La difficulté n'est pas de définir une loi de probabilité sur les utilités telle que sa projection sur l'information ordinale soit la culture impartiale. En effet, il est suffisant de définir une distribution où les électeurs sont indépendants et tous les candidats sont traités symétriquement. La problématique est d'en choisir une en particulier : une infinité de distributions vérifient ces conditions et on peut se demander si l'une de ces généralisations a des raisons canoniques d'être choisie plutôt que les autres.

Afin de répondre à cette question, nous avons besoin de prendre en compte un point technique important. Sous les hypothèses du théorème de Von Neumann–Morgenstern (que nous rappellerons), le vecteur d'utilité d'un agent est défini à deux constantes près, et choisir une normalisation particulière est arbitraire. En conséquence, l'espace des utilités est un espace quotient de $\mathbb{R}^C$ et, *a priori*, il n'y a pas de façon canonique de pousser une métrique de $\mathbb{R}^C$ vers cet espace quotient. Par conséquent, à première vue, il semble qu'il n'y ait pas de définition naturelle d'une distribution uniforme de probabilité sur cet espace.

Plus largement, chercher une généralisation naturelle de la culture impartiale à l'espace des utilités mène à explorer divers aspects de la géométrie de cet espace quotient et à mieux comprendre ses propriétés liées à l'algèbre, la topologie et la théorie de la mesure.

Le reste de cette annexe est organisé comme suit. En section B.1, nous présentons le formalisme de Von Neumann–Morgenstern et nous définissons l'espace des utilités. En section B.2, nous montrons que l'espace des utilités peut être vu comme le quotient du dual de l'espace des paires de loteries sur les candidats. En section B.3, nous définissons naturellement un opérateur opposé, qui correspond à inverser les préférences tout en préservant leurs intensités, et un opérateur somme, qui est caractérisé par le fait qu'il préserve les préférences unanimes.

Puisque l'espace des utilités est une variété, il est naturel de vouloir le munir d'une métrique riemannienne. En section B.4, nous prouvons que la seule représentation riemannienne qui préserve les propriétés projectives naturelles et la symétrie *a priori* entre les candidats est la métrique ronde. En section B.5, nous utilisons ce résultat pour donner une généralisation canonique de la culture impartiale et pour suggérer l'utilisation du modèle de Von Mises–Fisher pour représenter des cultures polarisées. En section B.6, nous résumons brièvement nos conclusions pour le modèle classique de Von Neumann-Morgenstern.

En section B.7, nous généralisons ces résultats à un modèle d'utilités avec *limite d'approbation*, c'est-à-dire que la valeur d'utilité 0 possède un sens particulier. Dans ce cas, le vecteur d'utilité d'un agent est défini à une constante multiplicative positive près, mais il n'y a plus de degré de liberté en translation. Nous montrons que, dans ce modèle, les seules représentations qui respectent les

propriétés naturelles de l'espace des utilités et la neutralité forment une famille d'ellipsoïdes de révolution à un paramètre réel.

B.1 Modèle de Von Neumann-Morgenstern

Dans cette section, nous définissons quelques notations et nous rappelons le théorème classique de Von Neumann-Morgenstern qui permet de représenter les préférences d'un agent sur des options probabilisées en introduisant la notion d'utilité espérée.

Soit $C \in \mathbb{N}\setminus\{0\}$. Nous considérerons C options mutuellement exclusives appelées *candidats*, chacune d'entre elles étant représentée par un indice appartenant à $\mathcal{C} = [\![1, C]\!]$.

Une *loterie* sur les candidats est un C-uplet $(L_1, \ldots, L_C) \in {\mathbb{R}_+}^C$ tel que $\sum_{c=1}^{C} L_c = 1$. L'ensemble des loteries est noté $\mathbb{L}$. Pour toute paire de loteries $L = (L_1, \ldots, L_C)$ et $M = (M_1, \ldots, M_C)$, pour tout $\lambda \in [0, 1]$, on définit naturellement leur *loterie composée réduite* comme un barycentre :

$$\lambda L + (1-\lambda)M = \Big(\lambda L_1 + (1-\lambda)M_1, \ldots, \lambda L_C + (1-\lambda)M_C\Big).$$

Les préférences d'un agent sur les loteries sont représentées par une relation binaire $\precsim$ sur $\mathbb{L}$. On note $\prec$ la relation stricte associée à $\precsim$, définie par : $L \prec M \Leftrightarrow L \precsim M$ et non $M \precsim L$.

Rappelons les hypothèses du théorème de Von Neumann–Morgenstern. On dit que la relation $\precsim$ est *complète* ssi $\forall (L, M) \in \mathbb{L}^2 : L \precsim M$ ou $M \precsim L$. Elle est *transitive* ssi $\forall (L, M, N) \in \mathbb{L}^3$, $L \precsim M$ et $M \precsim N \Rightarrow L \precsim N$. Elle est *archimédienne* ssi $\forall (L, M, N) \in \mathbb{L}^3$:

$$L \prec M \text{ et } M \prec N \Rightarrow \exists \varepsilon \in]0, 1[\text{ t.q. } (1-\varepsilon)L + \varepsilon N \prec M \prec \varepsilon L + (1-\varepsilon)N.$$

Et elle est *indépendante des alternatives non pertinentes* ssi $\forall (L, M, N) \in \mathbb{L}^3, \forall \lambda \in]0, 1]$:

$$L \prec M \Rightarrow \lambda L + (1-\lambda)N \prec \lambda M + (1-\lambda)N.$$

On dit que $\mathbf{u} = (u_1, \ldots, u_C) \in \mathbb{R}^C$ est un *vecteur d'utilité* représentant $\precsim$ ssi suivre la relation $\precsim$ est équivalent à maximiser l'utilité espérée au sens de $\mathbf{u}$; c'est-à-dire, ssi pour tout couple de loteries (L, M) :

$$L \precsim M \Leftrightarrow \sum_{c=1}^{C} L_c u_c \leq \sum_{c=1}^{C} M_c u_c.$$

Théorème B.1 (Von Neumann et Morgenstern)
Pour une relation binaire $\precsim$ sur $\mathbb{L}$, les conditions suivantes sont équivalentes.

1. *La relation $\precsim$ est complète, transitive, archimédienne et indépendante des alternatives non pertinentes.*
2. *Il existe un vecteur d'utilité $\mathbf{u} \in \mathbb{R}^C$ représentant $\precsim$.*

Quand elles sont vérifiées, $\mathbf{u}$ est défini à une constante additive et une constante multiplicative strictement positive près. Formellement, soit $\mathbf{u} \in \mathbb{R}^C$ un vecteur d'utilité représentant $\precsim$. Un vecteur $\mathbf{v} \in \mathbb{R}^C$ est aussi un vecteur d'utilité représentant $\precsim$ ssi $\exists \alpha \in]0, +\infty[, \exists \beta \in \mathbb{R}$ t.q. $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{1}$, où $\mathbf{1}$ est une notation pour le vecteur dont les C coordonnées sont égales à 1.

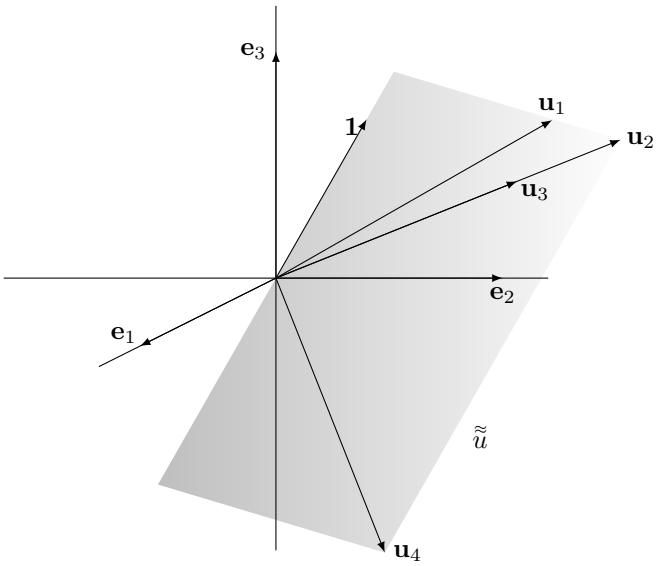


FIGURE B.1 – Espace $\mathbb{R}^3$ des vecteurs d'utilité pour 3 candidats (sans passage au quotient).

Ce n'est pas notre but ici de donner une preuve de ce théorème (Von Neumann *et al.*, 2007; Mas-Colell *et al.*, 1995; Kreps, 1990) ni de discuter si ses hypothèses sont valides expérimentalement (Fishburn, 1988; Mas-Colell *et al.*, 1995).

Afin de définir l'espace des utilités, tous les vecteurs représentant les mêmes préférences doivent être identifiés en un seul et même point. On note $\approx$ la relation d'équivalence définie par $\forall(\mathbf{u},\mathbf{v}) \in (\mathbb{R}^C)^2$:

$$\mathbf{u} \approx \mathbf{v} \Leftrightarrow \exists \alpha \in]0,+\infty[, \exists \beta \in \mathbb{R} \text{ t.q. } \mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{1}.$$

L'*espace des utilités* sur $\mathcal{C}$, noté $\mathbb{U}$, est défini comme l'ensemble quotient $\mathbb{R}^C/\approx$. On appelle *projection canonique* de $\mathbb{R}^C$ vers $\mathbb{U}$ la fonction :

$$\tilde{\tilde{\pi}} : \left| \begin{array}{ccc} \mathbb{R}^C & \to & \mathbb{U} \\ \mathbf{u} & \to & \tilde{\tilde{u}} = \{\mathbf{v} \in \mathbb{R}^C \text{ t.q. } \mathbf{v} \approx \mathbf{u}\}. \end{array} \right.$$

Pour tout $\mathbf{u} \in \mathbb{R}^C$, on note sans ambiguïté $\precsim_{\tilde{\tilde{u}}}$ la relation binaire sur $\mathbb{L}$ représentée par $\mathbf{u}$ au sens du théorème B.1 (Von Neumann–Morgenstern).

La figure B.1 représente l'espace $\mathbb{R}^3$ des vecteurs d'utilité pour 3 candidats, avant la projection sur l'espace quotient. La base canonique de $\mathbb{R}^3$ est notée $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Les vecteurs d'utilité $\mathbf{u}_1$ à $\mathbf{u}_4$ représentent les mêmes préférences que tout autre vecteur dans le demi-plan $\tilde{\tilde{u}}$, représenté en grisé. Plus généralement, tout point non-trivial $\tilde{\tilde{u}}$ de l'espace quotient correspond à un demi-plan de $\mathbb{R}^C$ délimité par la droite vect($\mathbf{1}$), le sous-espace vectoriel engendré par $\mathbf{1}$. La seule exception est le point de totale indifférence $\tilde{\tilde{0}}$. Dans $\mathbb{R}^C$, il ne correspond pas à un demi-plan mais à la droite vect($\mathbf{1}$) elle-même.

Afin de manipuler les utilités, en théorie et en pratique, il serait avantageux d'avoir un représentant canonique $\mathbf{u}$ pour chaque classe d'équivalence $\tilde{\tilde{u}}$. Dans la figure B.2, pour chaque $\tilde{\tilde{u}}$ non indifférent, nous choisissons son représentant vérifiant $\min(u_i) = -1$ et $\max(u_i) = 1$. L'espace des utilités $\mathbb{U}$ (hormis le point

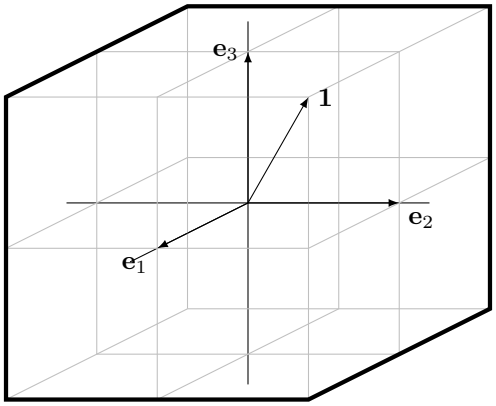


FIGURE B.2 – Avec 3 candidats, représentation de $\mathbb{U}$ comme des arêtes du cube unité dans $\mathbb{R}^3$.

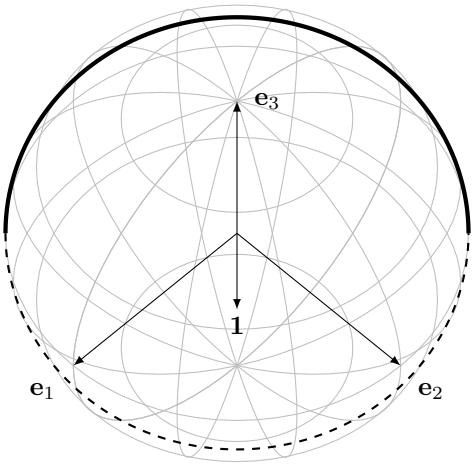


FIGURE B.3 – Avec 3 candidats, représentation de $\mathbb{U}$ comme un cercle de $\mathbb{R}^3$.

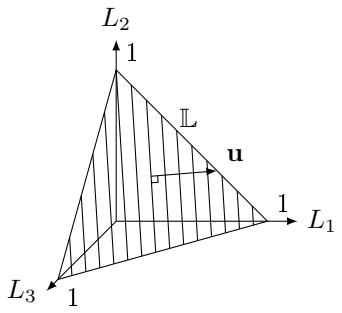


FIGURE B.4 – Espace $\mathbb{L}$ des loteries pour 3 candidats.

d'indifférence) est représenté dans $\mathbb{R}^3$ par six arêtes du cube unité. Dans la figure B.3, nous choisissons le représentant vérifiant $\sum u_i = 0$ et $\sum {u_i}^2 = 1$. Dans ce cas, $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$ est représenté dans $\mathbb{R}^3$ par le cercle unité du plan vectoriel qui est orthogonal à $\mathbf{1}$.

Si nous choisissons une telle représentation, l'espace quotient $\mathbb{U}$ peut hériter de la distance euclidienne de $\mathbb{R}^C$; par exemple, on peut évaluer les distances le long des arêtes du cube dans la figure B.2, ou le long du cercle unité dans la figure B.3. Mais il est clair que le résultat dépendra de la représentation choisie. Par conséquent, une question intéressante est de savoir si l'une de ces deux représentations, ou une autre encore, a des raisons canoniques d'être utilisée. Mais avant de répondre à cette question, nous avons besoin d'explorer plus généralement les propriétés géométriques de l'espace des utilités.

B.2 Dualité avec l'hyperplan tangent aux loteries

Dans cette section, nous remarquons que l'espace des utilités est un dual de l'espace des paires de loteries. Non seulement cela donne un point de vue différent sur l'espace des utilités (ce qui nous semble intéressant en soi), mais ce sera également utile pour prouver le théorème B.4, qui caractérise l'opérateur comme que nous définirons dans la section B.3.

Dans l'exemple représenté en figure B.4, on considère $C = 3$ candidats et $\mathbf{u} = (\frac{5}{3} ; -\frac{1}{3} ; -\frac{4}{3})$. Le grand triangle, ou simplexe, est l'espace des loteries $\mathbb{L}$. Les hachures sont des lignes d'indifférence de l'agent : celui-ci est indifférent entre tout couple de loteries se trouvant sur la même ligne d'indifférence (cf. Mas-Colell *et al.* (1995), section 6.B). Le vecteur d'utilité $\mathbf{u}$ représenté ici est dans le plan du simplexe mais ce n'est pas obligatoire : en effet, $\mathbf{u}$ peut être choisi arbitrairement dans sa classe d'équivalence $\tilde{\tilde{u}}$. Cependant, c'est un choix assez naturel, puisque la composante de $\mathbf{u}$ dans la direction $\mathbf{1}$ (orthogonale au simplexe) n'a aucune importance pour les préférences.

Le vecteur d'utilité $\mathbf{u}$ peut être vu comme un gradient de préférence [1] : en chaque point, il permet de savoir dans quelles directions trouver des loteries que l'agent préfère. Cependant, seuls la direction et le sens de $\mathbf{u}$ sont importants, tandis que sa norme n'a pas de signification particulière. Par conséquent, l'espace des utilités n'est pas tout à fait un dual mais un quotient de dual, comme nous allons le formaliser.

1. Plus précisément, il s'agit d'un gradient uniforme : il est le même en chaque point du simplexe.

Pour toute paire de loteries $L = (L_1, \dots, L_C)$ et $M = (M_1, \dots, M_C)$, on appelle *bipoint* de L à M le vecteur : $\overrightarrow{LM} = (M_1 - L_1, \dots, M_C - L_C)$. On appelle *polytope tangent* de $\mathbb{L}$ l'ensemble $\mathcal{T}$ des bipoints de $\mathbb{L}$.

On appelle *hyperplan tangent* de $\mathbb{L}$:

$$\mathcal{H} = \{(\Delta_1, \dots, \Delta_C) \in \mathbb{R}^C \text{ t.q. } \sum_{c=1}^{C} \Delta_c = 0\}.$$

Dans la figure B.4, le polytope tangent $\mathcal{T}$ est l'ensemble des bipoints du grand triangle, vu comme une partie d'espace vectoriel (alors que $\mathbb{L}$ est moralement affine). L'hyperplan $\mathcal{H}$ est tout le plan vectoriel qui contient $\mathcal{T}$.

Notons $\langle \mathbf{u} \,|\, \mathbf{v} \rangle$ le produit scalaire canonique de $\mathbf{u}$ et $\mathbf{v}$. On appelle *demi-hyperplan positif associé à* $\mathbf{u}$ l'ensemble $\mathbf{u}^{+\mathcal{H}} = \{\mathbf{\Delta} \in \mathcal{H} \text{ t.q. } \langle \mathbf{u} \,|\, \mathbf{\Delta} \rangle \geq 0\}$. Par définition, une loterie M est préférée à une loterie L ssi le bipoint $\overrightarrow{LM}$ appartient à ce demi-hyperplan positif :

$$L \precsim_{\widetilde{u}} M \Leftrightarrow \langle \mathbf{u} \,|\, \overrightarrow{LM} \rangle \geq 0 \Leftrightarrow \overrightarrow{LM} \in \mathbf{u}^{+\mathcal{H}}.$$

Soit $\mathcal{H}^\star$ l'espace dual de $\mathcal{H}$, c'est à dire l'ensemble des formes linéaires sur $\mathcal{H}$. Pour tout $\mathbf{u} \in \mathbb{R}^C$, on appelle *forme linéaire associée à* $\mathbf{u}$ l'élément suivant de $\mathcal{H}^\star$:

$$\langle \mathbf{u} \,| : \left| \begin{array}{ccc} \mathcal{H} & \to & \mathbb{R} \\ \mathbf{\Delta} & \to & \langle \mathbf{u} \,|\, \mathbf{\Delta} \rangle. \end{array} \right.$$

Nous avons déjà observé que le vecteur d'utilité peut être vu comme un gradient, à ceci près que seuls sa direction et son sens ont une importance, pas sa norme. Formalisons cette idée. Pour $(f, g) \in (\mathcal{H}^\star)^2$, on note $f \approx g$ ssi ces deux formes linéaires sont des multiples positifs l'une de l'autre, c'est-à-dire ssi $\exists \alpha \in]0, +\infty[$ t.q. $g = \alpha f$. On note $\widetilde{\widetilde{\pi}}(f) = \{g \in \mathcal{H}^\star \text{ t.q. } f \approx g\}$: c'est la classe d'équivalence de f à multiplication positive près.

Proposition B.2

Pour tout couple $(\mathbf{u}, \mathbf{v}) \in (\mathbb{R}^C)^2$*, on a :*

$$\mathbf{u} \approx \mathbf{v} \Leftrightarrow \langle \mathbf{u} \,| \approx \langle \mathbf{v} \,|.$$

L'application suivante est une bijection :

$$\Theta : \left| \begin{array}{ccc} \mathbb{U} & \to & \mathcal{H}^\star / \approx \\ \widetilde{\widetilde{\pi}}(\mathbf{u}) & \to & \widetilde{\widetilde{\pi}}(\langle \mathbf{u} \,|). \end{array} \right.$$

Démonstration. $\mathbf{u} \approx \mathbf{v}$

$\Leftrightarrow \exists \alpha \in (0, +\infty), \exists \beta \in \mathbb{R}$ t.q. $\mathbf{v} - \alpha \mathbf{u} = \beta \mathbf{1}$

$\Leftrightarrow \exists \alpha \in (0, +\infty)$ t.q. $\mathbf{v} - \alpha \mathbf{u}$ est orthogonal à $\mathcal{H}$

$\Leftrightarrow \exists \alpha \in (0, +\infty)$ t.q. $\forall \mathbf{\Delta} \in \mathcal{H}, \langle \mathbf{v} \,|\, \mathbf{\Delta} \rangle = \alpha \langle \mathbf{u} \,|\, \mathbf{\Delta} \rangle$

$\Leftrightarrow \langle \mathbf{u} \,| \approx \langle \mathbf{v} \,|.$

Le sens $\Rightarrow$ prouve que Θ est bien définie : en effet, si $\widetilde{\widetilde{\pi}}(\mathbf{u}) = \widetilde{\widetilde{\pi}}(\mathbf{v})$, alors $\widetilde{\widetilde{\pi}}(\langle \mathbf{u} \,|) = \widetilde{\widetilde{\pi}}(\langle \mathbf{v} \,|)$. Le sens $\Leftarrow$ assure que Θ est injective. Enfin, il est clair que Θ est surjective. □

L'espace des utilités peut donc être vu comme un quotient du dual $\mathcal{H}^\star$ de l'espace tangent $\mathcal{H}$ des loteries $\mathbb{L}$. Un vecteur d'utilité peut être vu, à une constante strictement positive près, comme un gradient uniforme, c'est-à-dire une forme linéaire sur $\mathcal{H}$ qui révèle, pour tout point de l'espace des loteries, dans quelles directions l'agent peut trouver des loteries qu'il préfère.

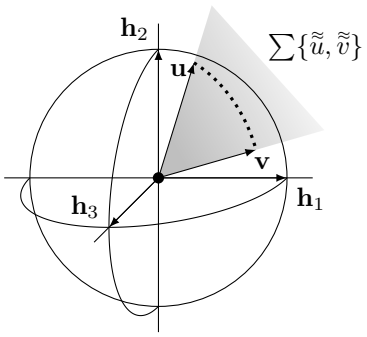


FIGURE B.5 – Somme de deux vecteurs d'utilité dans l'espace des utilités $\mathcal{U}_4$.

B.3 Opérateurs opposé et somme

En tant que quotient de $\mathbb{R}^C$, l'espace des utilités hérite d'opérations naturelles sur $\mathbb{R}^C$, la prise de l'opposé et la somme. Nous allons voir que ces deux opérations quotient ont une signification intuitive concernant les préférences. La somme permettra aussi de définir des *droites* dans la section B.4, ce qui sera une notion-clé pour le théorème B.6, caractérisant les métriques riemanniennes convenables pour l'espace des utilités.

On définit l'opérateur *opposé* de $\mathbb{U}$ par :

$$-: \left| \begin{array}{ccc} \mathbb{U} & \to & \mathbb{U} \\ \tilde{\tilde{\pi}}(\mathbf{u}) & \to & \tilde{\tilde{\pi}}(-\mathbf{u}). \end{array} \right.$$

Cet opérateur est correctement défini et c'est une bijection : en effet, on a par exemple $\tilde{\tilde{\pi}}(\mathbf{u}) = \tilde{\tilde{\pi}}(\mathbf{v})$ ssi $\tilde{\tilde{\pi}}(-\mathbf{u}) = \tilde{\tilde{\pi}}(-\mathbf{v})$. Considérer l'opposé correspond à renverser les préférences de l'agent, sans modifier leurs intensités relatives.

On souhaite, à présent, pousser l'opérateur somme de $\mathbb{R}^C$ vers le quotient $\mathbb{U}$. On utilise une méthode générique pour pousser un opérateur vers l'espace quotient : étant donné $\tilde{\tilde{u}}$ et $\tilde{\tilde{v}}$ dans $\mathbb{U}$, on prend leurs antécédents dans $\mathbb{R}^C$ grâce à $\tilde{\tilde{\pi}}^{-1}$, on calcule la somme dans $\mathbb{R}^C$, puis le résultat est converti dans l'autre sens vers l'espace quotient $\mathbb{U}$ grâce à $\tilde{\tilde{\pi}}$.

Cependant, il n'y a pas unicité du résultat. En effet, désignons arbitrairement des représentants $\mathbf{u} \in \tilde{\tilde{u}}$ et $\mathbf{v} \in \tilde{\tilde{v}}$. Pour calculer la somme, on peut envisager de prendre n'importe quels représentants. Donc les sommes possibles sont $\alpha\mathbf{u}+\alpha'\mathbf{v}+\beta+\beta'$, où α et α' sont strictement positifs et où $\beta+\beta'$ est un réel quelconque. En repassant au quotient, on obtiendra par exemple $\tilde{\tilde{\pi}}(2\mathbf{u}+\mathbf{v})$ et $\tilde{\tilde{\pi}}(\mathbf{u}+3\mathbf{v})$, qui sont généralement distincts. Par conséquent, le résultat n'est pas un point de l'espace des utilités $\mathbb{U}$ mais un ensemble de points, c'est-à-dire un élément de $\mathcal{P}(\mathbb{U})$.

Cet exemple montre comment on peut définir la somme de deux éléments $\tilde{\tilde{u}}$ et $\tilde{\tilde{v}}$. Pour plus de généralité, nous allons définir la somme d'un nombre quelconque d'élément de $\mathbb{U}$. Par conséquent, nous prendrons aussi $\mathcal{P}(\mathbb{U})$ comme ensemble de départ.

On définit l'opérateur *somme* de $\mathbb{U}$ par :

$$\sum: \left| \begin{array}{ccc} \mathcal{P}(\mathbb{U}) & \to & \mathcal{P}(\mathbb{U}) \\ A & \to & \left\{ \tilde{\tilde{\pi}} \left(\sum_{i=1}^{n} \mathbf{u}_i \right), n \in \mathbb{N}, (\mathbf{u}_1, \ldots, \mathbf{u}_n) \in \left(\tilde{\tilde{\pi}}^{-1}(A) \right)^n \right\}. \end{array} \right.$$

Exemple B.3

Considérons $\mathbb{U}$ pour $C = 4$ candidats. Dans la figure B.5, à des fins de visualisation, on représente sa projection dans $\mathcal{H}$, ce qui est rendu possible par le choix

de la constante de normalisation β. Comme $\mathcal{H}$ est un espace à 3 dimensions, on peut en considérer une base orthonormale $(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3)$.

Pour deux vecteurs d'utilité non triviaux $\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{v}}$, le choix des multiplicateurs de normalisation α permet de choisir des représentants $\mathbf{u}$ et $\mathbf{v}$ dont la norme euclidienne vaut 1.

Dans cette représentation, la somme $\sum\{\widetilde{\widetilde{u}}, \widetilde{\widetilde{v}}\}$ est constituée des utilités correspondant aux vecteurs $\alpha\mathbf{u}+\alpha'\mathbf{v}$, où α et α' sont des réels positifs ou nuls. En effet, on a pris une représentation dans $\mathcal{H}$, donc toutes les constantes de normalisation β disparaissent. De plus, α, α' ou les deux peuvent être nulles car notre définition autorise à ignorer $\mathbf{u}$, $\mathbf{v}$ ou les deux. Quitte à prendre des représentants de norme unitaire pour les vecteurs d'utilité non triviaux, remarquons que la somme $\sum\{\widetilde{u}, \widetilde{v}\}$ peut être représentée par la ligne pointillée et le point d'indifférence $\mathbf{0}$.

Géométriquement, la somme est le quotient de l'enveloppe convexe des arguments. On notera que cette enveloppe convexe est, en fait, un cône convexe. Nous verrons plus loin son interprétation en termes de préférences.

Étant donné notre définition de l'opérateur somme, on considère le cône fermé : par exemple, les arguments eux-mêmes (par exemple $\widetilde{\widetilde{u}}$) répondent à notre définition, de même que le point d'indifférence $\widetilde{\widetilde{\pi}}(\mathbf{0})$. Ce ne serait généralement pas le cas si on se contentait des $\widetilde{\widetilde{\pi}}(\alpha\mathbf{u}+\alpha'\mathbf{v}+\beta\mathbf{1})$, où $\alpha > 0$ et $\alpha' > 0$. Cette convention a pour but avoué d'obtenir un énoncé concis pour le théorème B.4 que nous allons énoncer.

Nous allons maintenant montrer que, si A est l'ensemble des vecteurs d'utilités d'une population, alors $\sum A$ est le lieu des vecteurs d'utilité qui respectent les préférences unanimes de la population.

Théorème B.4 (caractérisation de la somme)
Soit $A \in \mathcal{P}(\mathbb{U})$ et $\widetilde{\widetilde{v}} \in \mathbb{U}$.
Les conditions suivantes sont équivalentes.

1. $\widetilde{\widetilde{v}} \in \sum A$.
2. $\forall (L, M) \in \mathbb{L}^2 \ : \left(\forall \widetilde{\widetilde{u}} \in A, L \precsim_{\widetilde{\widetilde{u}}} M\right) \Rightarrow L \precsim_{\widetilde{\widetilde{v}}} M$.

Démonstration. Premièrement, remarquons que le polytope tangent $\mathcal{T}$ engendre l'hyperplan tangent $\mathcal{H}$ par multiplication positive. C'est-à-dire :

$$\forall \boldsymbol{\Delta} \in \mathcal{H}, \exists \overrightarrow{LM} \in \mathcal{T}, \exists \lambda \in]0, +\infty[\text{ t.q. } \boldsymbol{\Delta} = \lambda \overrightarrow{LM}.$$

En effet, $\mathcal{T}$ contient un voisinage de l'origine dans l'espace vectoriel $\mathcal{H}$.

Soit $\mathbf{v} \in \widetilde{\widetilde{\pi}}^{-1}(\widetilde{\widetilde{v}})$. On a les équivalences suivantes.

— $\forall (L, M) \in \mathbb{L}^2, \left(\forall \widetilde{\widetilde{u}} \in A, L \precsim_{\widetilde{\widetilde{u}}} M\right) \Rightarrow L \precsim_{\widetilde{\widetilde{v}}} M$,
— $\forall \overrightarrow{LM} \in \mathcal{T}, \left(\forall \mathbf{u} \in \widetilde{\widetilde{\pi}}^{-1}(A), \langle\, \mathbf{u} \,|\, \overrightarrow{LM} \,\rangle \geq 0\right) \Rightarrow \langle\, \mathbf{v} \,|\, \overrightarrow{LM} \,\rangle \geq 0$,
— $\forall \boldsymbol{\Delta} \in \mathcal{H}, \left(\forall \mathbf{u} \in \widetilde{\widetilde{\pi}}^{-1}(A), \langle\, \mathbf{u} \,|\, \boldsymbol{\Delta} \,\rangle \geq 0\right) \Rightarrow \langle\, \mathbf{v} \,|\, \boldsymbol{\Delta} \,\rangle \geq 0$ (car $\mathcal{T}$ engendre $\mathcal{H}$),
— $\bigcap_{\mathbf{u} \in \widetilde{\widetilde{\pi}}^{-1}(A)} \mathbf{u}^{+\mathcal{H}} \subset \mathbf{v}^{+\mathcal{H}}$,
— $\mathbf{v}$ est dans le cône convexe de $\widetilde{\widetilde{\pi}}^{-1}(A)$ (à cause de la dualité vue en section B.2),
— $\widetilde{\widetilde{v}} \in \sum A$.

□

Exemple B.5
Considérons un $\widetilde{\widetilde{u}}$ non indifférent et examinons le cas particulier de la somme de $\widetilde{\widetilde{u}}$ et de son opposé $-\widetilde{\widetilde{u}}$. Par application directe de la définition, on voit que la somme est constituée de $\widetilde{\widetilde{u}}$, $-\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{0}}$.

Or, nous venons de prouver que la somme est le lieu des vecteurs d'utilité qui conservent les préférences unanimes sur les loteries. Intuitivement, on pourrait penser que, puisque $\widetilde{\widetilde{u}}$ et $-\widetilde{\widetilde{u}}$ ne semblent jamais d'accord, tout vecteur d'utilité $\widetilde{\widetilde{v}}$ respecte l'ensemble vide de leurs préférences communes ; par conséquent, leur somme devrait être l'espace entier. Mais cette intuition n'est pas correcte.

En effet, examinons l'exemple de $\mathbf{u} = (1, 0, \dots, 0)$. Pour deux loteries L et M, les deux opinions opposées $\widetilde{\widetilde{u}}$ et $-\widetilde{\widetilde{u}}$ sont en accord si et seulement si $L_1 = M_1$: dans ce cas, toutes deux sont indifférentes entre L et M. Les seuls points de l'espace des utilités qui vérifient cette propriété commune sont $\widetilde{\widetilde{u}}$ et $-\widetilde{\widetilde{u}}$ eux-mêmes, ainsi que le point d'indifférence $\widetilde{\widetilde{0}}$.

B.4 Représentation riemannienne de l'espace des utilités

Comme l'espace des utilités est une variété, il est naturel de vouloir le munir d'une métrique riemannienne. Dans cette section, nous prouvons qu'il existe un choix limité de métriques qui soient cohérentes avec les propriétés naturelles de l'espace et la symétrie *a priori* entre les candidats.

Tout d'abord, notons que le point d'indifférence $\widetilde{\widetilde{0}}$ doit être exclus. En effet, son seul voisinage ouvert est $\mathbb{U}$ tout entier, et aucune distance n'est cohérente avec cette propriété [2]. En revanche, $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ possède la même topologie que la sphère de dimension $C - 2$, donc on peut le munir d'une distance.

Une propriété naturelle de la distance serait que ses géodésiques coïncident avec les segments d'unanimité définis par la somme. En effet, considérons un agent avec le vecteur d'utilité $\widetilde{\widetilde{u}}_0$ — disons $(-10, 0, 1, 2)$, qui est progressivement convaincu par un autre agent avec le vecteur d'utilité $\widetilde{\widetilde{u}}_1$ — disons $(10, 0, 1, 2)$, le long d'une courbe paramétrée $\widetilde{\widetilde{u}}_t$. Si les géodésiques coïncident avec les segments d'unanimité, cela signifie que si l'agent suit le plus court chemin pour la métrique choisie, ses préférences vont progressivement changer à propos de leurs points de désaccord, sans jamais dévier à propos des points sur lesquels ils sont déjà d'accord (dans l'exemple, toutes les paires de loteries qui ont la même probabilité pour le candidat 1). Une telle déviation pourrait se produire dans la vie réelle, mais peut raisonnablement ne pas être considérée comme un plus court chemin.

À présent, définissons la métrique ronde. Le quotient $\mathbb{R}^C / \operatorname{vect}(\mathbf{1})$ est identifié à $\mathcal{H}$ et muni du produit scalaire issu de celui canonique de $\mathbb{R}^C$. L'espace des utilités $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ est identifié à la sphère unitaire de $\mathcal{H}$ et muni de la structure riemannienne induite. On note ξ_0 cette métrique riemannienne sur $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$.

Pour s'en faire une représentation intuitive, on peut représenter toute position $\widetilde{\widetilde{u}}$ par un vecteur $\mathbf{u}$ qui vérifie $\sum u_i = 0$ et $\sum {u_i}^2 = 1$. On obtient la sphère unité de $\mathcal{H}$, de dimension $C - 2$, et on considère la métrique induite par la métrique euclidienne canonique de $\mathbb{R}^C$. C'est-à-dire qu'on mesure les distances sur la surface de la sphère en utilisant sur chaque espace tangent la restriction du produit scalaire canonique. Pour $C = 3$, une telle représentation a déjà été illustrée dans la figure B.3.

Avec cette représentation à l'esprit, on peut donner une formule pour calculer des distances au sens de ξ_0. On note I la matrice identité de taille C, J la matrice de taille $C \times C$ dont tous les coefficients valent 1 et D_0 la matrice de la projection

2. Techniquement, cette remarque prouve que $\mathbb{U}$ (muni de la topologie quotient) n'est pas un *espace* T_1 (Guénard et Lelièvre, 1985).

orthogonale sur $\mathcal{H}$:

$$D_0 = I - \frac{1}{C}J.$$

La norme euclidienne canonique de $\mathbf{u}$ est notée $\|\mathbf{u}\|$. Pour deux vecteurs d'utilités non indifférents $\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{v}}$, la distance entre $\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{v}}$ au sens de la métrique ξ_0 est :

$$\delta(\widetilde{\widetilde{u}}, \widetilde{\widetilde{v}}) = \arccos \left\langle \frac{D_0\mathbf{u}}{\|D_0\mathbf{u}\|} \,\middle|\, \frac{D_0\mathbf{v}}{\|D_0\mathbf{v}\|} \right\rangle.$$

Si $\mathbf{u}$ et $\mathbf{v}$ sont déjà des vecteurs unitaires de $\mathcal{H}$, c'est-à-dire des représentants canoniques de leurs classes d'équivalence $\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{v}}$, alors la formule se simplifie en $d(\widetilde{\widetilde{u}}, \widetilde{\widetilde{v}}) = \arccos\langle \mathbf{u} \,|\, \mathbf{v} \rangle$.

Nous allons, maintenant, prouver que pour $C \geq 4$, la représentation sphérique est la seule qui soit cohérente avec les propriétés naturelles de l'espace et qui respecte la symétrie *a priori* entre les candidats.

Théorème B.6 (représentation riemannienne de l'espace des utilités)
On suppose $C \geq 4$. Soit ξ une métrique riemannienne sur $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$.
Les conditions 1 et 2 suivantes sont équivalentes.

1. (a) *Pour tout couple de points non antipodaux $\widetilde{\widetilde{u}}, \widetilde{\widetilde{v}} \in \mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ (c.à.d. $\widetilde{\widetilde{v}} \neq -\widetilde{\widetilde{u}}$), l'ensemble $\sum\{\widetilde{\widetilde{u}}, \widetilde{\widetilde{v}}\}$ des éléments respectant les préférences unanimes de $\widetilde{\widetilde{u}}$ et $\widetilde{\widetilde{v}}$ est un segment de géodésique de ξ ; et*

   (b) *pour toute permutation σ de $[\![1, C]\!]$, l'action Φ_σ induite sur $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ par*

   $$(u_1, \ldots, u_C) \to (u_{\sigma(1)}, \ldots, u_{\sigma(C)})$$

   est une isométrie.

2. $\exists \lambda \in]0, +\infty[$ *t.q.* $\xi = \lambda \xi_0$.

Démonstration. Le sens 2 $\Rightarrow$ 1 étant évident, nous allons prouver 1 $\Rightarrow$ 2. Le résultat profond sous-jacent est un théorème classique de Beltrami, qui date du milieu du dix-neuvième siècle Beltrami (1866) (voir aussi Beltrami (1869)).

L'image d'un sous-espace de dimension 2 de $\mathcal{H}$ dans $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ par la projection canonique $\widetilde{\widetilde{\pi}}$ est appelée une *droite*[3] de $\mathbb{U}\setminus\{\widetilde{\widetilde{0}}\}$. Cette notion est profondément liée à l'opérateur somme : en effet, la somme de deux points non antipodaux de $\mathbb{U}\setminus\{\widetilde{\widetilde{0}}\}$ est un segment de la droite qui les relie. La condition 1a signifie précisément que les géodésiques de ξ sont les droites de $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$. Le théorème de Beltrami assure, ensuite, que $\mathbb{U}\setminus\{\widetilde{\widetilde{0}}\}$ a une courbure constante. Remarquons que ce résultat est, en fait, plus subtil en dimension 2 (c'est-à-dire pour $C = 4$) que dans les dimensions supérieures ; voir Spivak (1979a), théorème 1.18 et Spivak (1979b), théorème 7.2 pour des preuves.

Puisque $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ est une sphère topologique, cette courbure constante est nécessairement positive. Quitte à multiplier ξ par une constante, nous pouvons supposer que cette courbure constante vaut 1. Par conséquent, il existe une isométrie $\Psi : \mathcal{S}_{C-2} \to \mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$, où $\mathcal{S}_{C-2}$ est la sphère unité de $\mathbb{R}^{C-1}$ munie de sa métrique ronde usuelle. La fonction Ψ envoie manifestement les géodésiques sur des géodésiques. On va en déduire le lemme suivant.

Lemme B.7
Il existe une application linéaire $\Lambda : \mathbb{R}^{C-1} \to \mathcal{H}$ qui induit Ψ, c'est-à-dire telle que :

$$\Psi \circ \Pi = \Pi \circ \Lambda,$$

où Π désigne aussi bien la projection $\mathbb{R}^{C-1} \to \mathcal{S}_{C-2}$ que la projection $\mathcal{H} \to \mathbb{U}\setminus\{\widetilde{\widetilde{0}}\}$.

3. Le lecteur non expert en géométrie projective pourra s'y familiariser avec Kloeckner (2012).

Démonstration. Tout d'abord, Ψ envoie tout couple de points antipodaux de $\mathcal{S}_{C-2}$ vers un couple de points antipodaux de $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$: en effet, dans les deux cas, les couples antipodaux sont caractérisés par le fait qu'il existe plus d'une géodésique qui les contient tous les deux. Il s'ensuit que Ψ induit une fonction Ψ' de l'espace projectif $\mathbb{P}(\mathbb{R}^{C-1})$ (qui est l'ensemble des droites passant par l'origine dans $\mathbb{R}^{C-1}$, identifié avec l'ensemble des couples de points antipodaux de $\mathcal{S}_{C-2}$) vers l'espace projectif $\mathbb{P}(\mathcal{H})$ (qui est l'ensemble des droites passant par l'origine dans $\mathcal{H}$, identifié avec l'ensemble des couples de points antipodaux de $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$).

Le fait que Ψ envoie les géodésiques de $\mathcal{S}_{C-2}$ vers des géodésiques de $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$, combiné avec la condition 1a, implique que Ψ' envoie les droites projectives vers des droites projectives.

Or il est bien connu qu'un bijection $\mathbb{R}^n \to \mathbb{R}^n$ qui envoie les droites sur des droites est nécessairement une application affine ; un résultat similaire est vrai en géométrie projective, ce qui implique que Ψ' est une application projective. Voir par exemple Audin (2003) pour les deux résultats cités.

Le fait que Ψ' soit projective signifie précisément qu'il existe une application linéaire $\Lambda : \mathbb{R}^{C-1} \to \mathcal{H}$ qui induit Ψ', laquelle induit à son tour $\Psi : \mathcal{S}_{C-2} \to \mathbb{U} \setminus \{\tilde{\tilde{0}}\}$. □

En utilisant Λ pour pousser le produit scalaire canonique de $\mathbb{R}^{C-1}$, nous en déduisons qu'il existe un produit scalaire $(\mathbf{u}, \mathbf{v}) \to \phi(\mathbf{u}, \mathbf{v})$ sur $\mathcal{H}$ qui induit ξ, dans le sens que ξ est la métrique riemannienne obtenue en identifiant $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$ avec la sphère unité définie dans $\mathcal{H}$ par ϕ et en y restreignant ϕ.

Il nous reste à prouver que ϕ est le produit scalaire induit par le produit scalaire canonique de $\mathbb{R}^C$. Remarquons que l'hypothèse 1b est obligatoire, puisque n'importe quel produit scalaire sur $\mathcal{H}$ induit sur $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$ une métrique riemannienne qui satisfait 1a.

Chaque vecteur de la base canonique $\mathbf{e}_c = (0, \dots, 1, \dots, 0)$ définit un point de $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$ et une demi-droite ℓ_c de $\mathcal{H}$. La condition 1b garantit que ces demi-droites sont permutées par des isométries de $(\mathcal{H}, \phi)$. En particulier, il existe des vecteurs $\mathbf{u}_c \in \ell_c$ dont les distances deux à deux sont toutes égales (au sens de ϕ).

Lemme B.8

La famille $\mathbf{u}_1, \dots, \mathbf{u}_{C-1}$ est, à multiplication près par un scalaire, la seule base de $\mathcal{H}$ telle que $\mathbf{u}_c \in \ell_c$ et $\sum_{c<C} \mathbf{u}_c \in -\ell_C$.

Démonstration. Tout d'abord, par définition des $\mathbf{u}_c$, ces vecteurs forment un simplexe régulier et $\sum_c \mathbf{u}_c = \mathbf{0}$. Il s'ensuit que $\mathbf{u}_1, \dots, \mathbf{u}_{C-1}$ a la propriété désirée et il nous reste à montrer l'unicité.

Supposons que $\mathbf{v}_1, \dots, \mathbf{v}_{C-1}$ est une base de $\mathcal{H}$ telle que $\mathbf{v}_c \in \ell_c$ et $\sum_{c<C} \mathbf{v}_c \in -\ell_c$. Alors il existe des scalaires positifs $a_1, \dots, a_C$ tels que $\mathbf{v}_c = a_c \mathbf{u}_c$ pour tout $c < C$, et $\sum_{c<C} \mathbf{v}_c = a_C \sum_{c<C} \mathbf{u}_c$.

Alors $\sum_{c<C} a_c \mathbf{u}_c = \sum_{c<C} a_c \mathbf{u}_c$, et comme les u_c forment une base, on a nécessairement $a_c = a_C$ pour tout c. □

À présent, considérons le produit scalaire canonique ϕ_0 sur $\mathcal{H}$ induit par le produit scalaire canonique de $\mathbb{R}^C$. Puisque les permutations de coordonnées sont des isométries, nous déduisons que les vecteurs $\mathbf{v_c} = \Pi(\mathbf{e}_c)$ (où Π désigne maintenant la projection orthogonale de $\mathbb{R}^C$ sur $\mathcal{H}$) forment un simplexe régulier pour ϕ_0, tel que $\sum_c \mathbf{v_c} = \mathbf{0}$. Par conséquent, $\mathbf{u}_c = \lambda \mathbf{v_c}$ pour un certain $\lambda > 0$ et pour tout c. On en déduit que les vecteurs $\mathbf{u}_c$ forment un simplexe régulier à la fois pour ϕ et ϕ_0, qui sont donc multiples l'un de l'autre. Ceci achève la preuve du théorème B.6. □

En revanche, l'implication 1 ⇒ 2 du théorème est fausse pour $C = 3$. Pour chaque vecteur d'utilité non indifférent, considérons son représentant qui vérifie $\min(u_i) = 0$ et $\max(u_i) = 1$. On identifie ainsi $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ à des arêtes du cube unité de $\mathbb{R}^3$, comme dans la figure B.2. On utilise cette identification pour munir $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ de la métrique induite sur ces arêtes par le produit scalaire canonique de $\mathbb{R}^3$. Alors les conditions 1a et 1b du théorème sont vérifiées, mais pas la condition 2.

Ce théorème amène une autre remarque importante. Comme nous avons un représentant canonique $\mathbf{u}$ pour chaque classe d'équivalence $\widetilde{\widetilde{u}}$ qui est un vecteur unité de $\mathcal{H}$, nous pourrions être tentés de l'utiliser pour comparer les utilités de différents agents.

Nous insistons sur le fait que cette représentation ne peut pas être utilisée pour des comparaisons interpersonnelles de différences d'utilités.

Par exemple, pour deux agents, considérons les représentants suivants :

$$\begin{cases} \mathbf{u} = (0,00; 0,71; -0,71)\,, \\ \mathbf{v} = (0,57; 0,22; -0,79)\,. \end{cases}$$

Le fait que $v_3 < u_3$ ne signifie pas qu'un agent avec les préférences $\widetilde{\widetilde{v}}$ a davantage d'aversion pour le candidat 3 qu'un agent avec des préférences $\widetilde{\widetilde{u}}$. De même, quand on passe du candidat 1 au candidat 2, le gain pour l'agent $\widetilde{\widetilde{u}}$ (+0.71) ne peut pas être comparé à la perte pour l'agent $\widetilde{\widetilde{v}}$ (−0.35).

Le théorème B.6 ne contient aucun message pour la comparaison interpersonnelle des utilités. En effet, des utilités appartenant à deux agents sont essentiellement incomparables en l'absence d'hypothèses supplémentaires (Hammond, 1991). Prendre des représentants canoniques dans la sphère à $(C-2)$ dimensions est seulement utilisé pour calculer des distances entre deux points de l'espace des utilités.

B.5 Application : mesures de probabilité sur l'espace des utilités

Une fois l'espace muni d'une métrique, il est muni d'une mesure de probabilité naturelle : la mesure uniforme au sens de cette métrique (ce qui est possible parce que l'espace a une mesure totale finie). Nous noterons μ_0 cette mesure, qui est donc le volume riemannien normalisé défini par la métrique ξ_0.

En pratique, pour tirer des vecteurs selon une probabilité uniforme sur $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$, il suffit d'utiliser une loi uniforme sur la sphère unité de $\mathcal{H}$. Autrement dit, une fois qu'on a identifié $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ avec la sphère unité de $\mathcal{H}$, alors μ_0 est exactement la mesure uniforme usuelle.

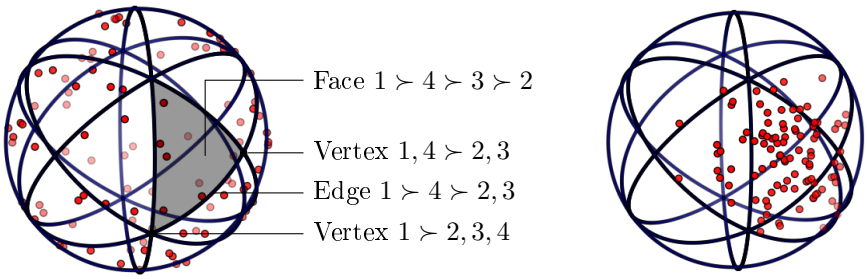
En l'occurrence, le fait que la sphère ronde possède de nombreuses symétries implique des propriétés agréables supplémentaires pour μ_0, que nous pouvons résumer en une proposition.

Proposition B.9

Soit μ une mesure de probabilité sur $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$.

1. *Supposons que pour tout $r > 0$, μ donne la même probabilité à toutes les boules de $\mathbb{U} \setminus \{\widetilde{\widetilde{0}}\}$ de rayon r (selon la métrique ξ_0) ; alors $\mu = \mu_0$.*
2. *Si μ est invariante par toute isométrie de la métrique ξ_0, alors $\mu = \mu_0$.*

Les deux caractérisations sont classiques. La première est (une version forte de) la définition du volume riemannien. La seconde découle de la première et du



(a) Uniforme (culture impartiale). (b) Von Mises–Fisher.

FIGURE B.6 – Deux distributions de 100 agents sur $\mathbb{U}$ avec 4 candidats.

fait que deux points quelconques de la sphère ronde peuvent être envoyés l'un vers l'autre par une isométrie.

Dans la figure B.6(a), réalisée avec SVAMP, on représente une distribution de 100 agents tirés uniformément et indépendamment sur la sphère, avec 4 candidats. Comme pour la figure B.5, on a représenté uniquement la sphère unité de $\mathcal{H}$.

Les lignes bleues forment le *permutoèdre*, une figure géométrique représentant l'aspect ordinal de ces préférences. Chaque *face* est constituée de tous les points qui partagent le même ordre strict total de préférence. Un vecteur d'utilité appartient à une *arête* s'il a seulement 3 utilités distinctes : par exemple, si un agent préfère le candidat 1 à 4, 4 à 2 et 3, mais est indifférent entre les candidats 2 et 3. Enfin, un point est un *sommet* s'il a seulement deux utilités distinctes : par exemple, si l'agent préfère le candidat 1 à tous les autres mais est indifférent entre ceux-ci.

Dans cette distribution, chaque agent a presque sûrement un ordre strict total de préférence. Chaque ordre a la même probabilité et les agents sont indépendants, donc cette distribution est une généralisation naturelle de la culture impartiale quand on considère des utilités espérées.

Puisque le point $\tilde{\tilde{0}}$ est une singularité géométrique, il est difficile de l'inclure naturellement dans une telle mesure. Si on souhaite le prendre en compte, la solution la plus simple consiste à le tirer avec une certaine probabilité et à utiliser la distribution sur $\mathbb{U} \setminus \{\tilde{\tilde{0}}\}$ dans les autres cas. Ceci dit, nous venons de remarquer que tous les autres ordres stricts non totaux ont une mesure égale à 0 ; donc pour un modèle théorique canonique, un choix naturel est d'attribuer également une mesure 0 au point d'indifférence.

Avoir une distance, et donc une mesure uniforme, permet également de définir d'autres mesures par leur densité par rapport à la mesure uniforme. Voici un exemple de loi définie par sa densité. Étant donné un vecteur $\mathbf{u}_0$ de la sphère unité de $\mathcal{H}$ et κ un réel positif ou nul, la distribution de Von Mises–Fisher (VMF) de pôle $\mathbf{n}$ et de concentration κ est définie par la densité de probabilité suivante par rapport à la loi uniforme de la sphère unité de $\mathcal{H}$:

$$p(\mathbf{u}) = X_\kappa e^{\kappa \langle\, \mathbf{u} \,|\, \mathbf{u}_0 \,\rangle},$$

où X_κ est une constante de normalisation. Étant donné le vecteur résultant d'une distribution sur la sphère, la distribution de VMF maximise l'entropie tout comme, dans l'espace euclidien, la loi gaussienne maximise l'entropie parmi les lois possédant une moyenne et un écart-type donnés (Mardia, 1975). Par conséquent, en l'absence d'information supplémentaire, c'est la distribution « naturelle » qui

devrait être utilisée. Cette culture est implantée dans SWAMP (section 6.1.3) et elle est étudiée en détail au chapitre 7. La figure B.6(b) représente une telle distribution, avec les mêmes conventions que dans la figure B.6(a). Pour tirer une distribution de VMF, on utilise l'algorithme d'Ulrich révisé par Wood (Ulrich, 1984; Wood, 1994).

Qualitativement, le modèle de VMF est similaire au modèle de Mallows, qui est utilisé pour des préférences ordinales (Mallows, 1957). Dans ce dernier, la probabilité d'un ordre de préférence σ est :

$$p(\sigma) = X'_\kappa e^{-\kappa\delta(\sigma,\sigma_0)},$$

où σ_0 est le mode de la distribution, $\delta(\sigma, \sigma_0)$ une distance entre σ et σ_0 (typiquement la distance tau de Kendall), κ un réel positif ou nul (concentration) et X'_k une constante de normalisation. Les modèles de VMF et de Mallows décrivent tous deux une culture où la population est polarisée, c'est-à-dire dispersée autour d'un point central, avec plus ou moins de concentration.

Cependant, il y a plusieurs différences.

— La distribution de VMF permet de tirer un point particulier de l'espace des utilités, alors que celle de Mallows choisit seulement une face du permutoèdre.

— En particulier, le pôle d'une distribution de VMF peut être situé en un point quelconque de ce continuum. Par exemple, même si le pôle est sur la face $1 \succ 4 \succ 3 \succ 2$, il peut être plus proche de la face $1 \succ 4 \succ \mathbf{2} \succ \mathbf{3}$ que de la face $\mathbf{4} \succ \mathbf{1} \succ 3 \succ 2$. Une telle nuance n'est pas possible dans le modèle de Mallows.

— Au voisinage du pôle, la probabilité de VMF décroît comme l'exponentielle du carré de la distance (car le produit scalaire est le cosinus de la distance sur la sphère), alors que la probabilité de Mallows décroît comme l'exponentielle de la distance (sans le carré).

— VMF est la distribution d'entropie maximale, étant donné la moyenne sphérique et la dispersion (tout comme une distribution gaussienne dans un espace euclidien).

L'existence d'une mesure canonique permet de définir d'autres mesures de probabilité facilement, en plus des deux mesures que nous venons de décrire. De telles mesures peuvent générer des populations artificielles d'agent à des fins de simulations. Elles peuvent aussi être utilisées pour faire correspondre des données d'expériences réelles à un modèle théorique et servir de point de comparaison neutre pour de telles données.

Pour développer ce dernier point, insistons sur le fait qu'étant donné une distribution (raisonnablement régulière) μ sur un espace tel que $\mathbb{U} \setminus \{\widetilde{0}\}$, il n'y a *a priori* aucune façon de définir ce que signifie, pour un élément $\widetilde{u}$, d'être plus probable qu'un autre élément $\widetilde{v}$. En effet, les deux ont une probabilité nulle et ce qui aurait un sens, c'est de comparer les probabilités d'être proche $\widetilde{u}$ ou de $\widetilde{v}$. Mais il faut alors comparer des voisinages de la même taille, ce qui n'a de sens que si on dispose d'une métrique. En d'autres termes, si on possède une distribution de référence comme μ_0, alors il est censé de considérer la densité $p = \frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}$, qui est une fonction (disons continue) sur $\mathbb{U} \setminus \{\widetilde{0}\}$. Alors, on peut comparer $p(\widetilde{u})$ et $p(\widetilde{v})$ pour dire si l'un de ces éléments est plus probable que l'autre selon μ. On remarquera qu'en l'occurrence, comparer la probabilité de r-voisinages pour la métrique ξ_0 ou la densité relativement à μ_0 donne le même résultat dans la limite $r \to 0$, ce qui est précisément la définition du volume riemannien.

B.6 Conclusion

Nous avons étudié les propriétés géométriques du modèle classique des utilités espérées introduit par Von Neumann et Morgenstern, quand les candidats proposés sont *a priori* considérés comme symétriques. Nous avons remarqué que l'espace des utilités peut être vu comme un dual de l'espace des loteries, que les opérateurs opposé et somme hérités de $\mathbb{R}^C$ ont une interprétation naturelle en termes de préférences et que l'espace a une topologie sphérique quand on supprime le point d'indifférence.

Nous avons prouvé que la seule représentation riemannienne qui respecte les droites projectives définies naturellement par l'opérateur somme et la symétrie entre les candidats est une sphère ronde.

Toutes nos considérations se reposent sur le principe consistant à ajouter le minimum d'information dans le système, notamment en respectant la symétrie *a priori* entre les candidats. Cela n'implique pas que la représentation sphérique de l'espace des utilités $\mathbb{U}$ soit la plus adaptée à l'étude de telle ou telle situation particulière. En effet, dès qu'on dispose d'une information supplémentaire (par exemple, un modèle de placement des candidats sur un spectre politique), il paraît naturel de la prendre en compte dans le modèle choisi. Cependant, si on souhaite, par exemple, étudier un mode de scrutin en toute généralité, sans se concentrer sur son application dans tel ou tel domaine, il paraît naturel de considérer un espace d'utilité avec une métrique aussi neutre que possible, comme celle définie dans ce chapitre par la représentation sphérique.

B.7 Extension : espace des utilités avec limite d'approbation

Le modèle que nous venons d'étudier correspond à une situation où il existe deux degrés de liberté dans le choix de la représentation des utilités. Dans cette section, qui n'a pas encore été publiée, nous présentons une variation légèrement différente, où la valeur d'utilité 0 possède un sens particulier, ce qui supprime le degré de liberté en translation. La motivation principale de ce modèle est de se donner un cadre naturel dans lequel on puisse étudier non seulement les systèmes de vote ordinaux et le vote par notation mais également le vote par assentiment.

B.7.1 Modèles avec options virtuelles

Étant donné la position $\widetilde{\widetilde{u}}$ d'un agent dans l'espace des utilités, on peut choisir indifféremment un vecteur $\mathbf{u} \in \widetilde{\widetilde{u}}$ pour représenter ses préférences. Cette liberté offre l'opportunité, en choisissant un vecteur $\mathbf{u}$ particulier, de stocker de l'information supplémentaire qu'on estime pertinente. Nous allons donner plusieurs méthodes : dans certaines applications particulières, elles permettent de choisir un représentant spécifique de $\widetilde{\widetilde{u}}$ d'une manière qui est pertinente dans le contexte. Ceci nous amènera à la notion *d'options virtuelles* et finalement au modèle avec *limite d'approbation*.

Avant de formaliser ces concepts, nous présentons quelques exemples où le choix d'un vecteur d'utilité spécifique permet de stocker une information supplémentaire.

Exemple B.10 (fête foraine)
À un stand de fête foraine, on peut ne rien gagner ($\varnothing$), gagner une peluche de

lynx (L) ou d'ornithorynque (O), ou les deux (LO). L'ensemble des candidats est $\mathcal{C} = \{\varnothing, \mathrm{L}, \mathrm{O}, \mathrm{LO}\}$.

Si on écrit le vecteur d'utilité d'un joueur, quitte à ajuster les paramètres α et β du théorème B.1 et en supposant que tout joueur préfère gagner les deux peluches plutôt que rien, on peut adopter la convention que l'utilité de ne rien gagner sera 0 et que celle de gagner les deux peluches sera 1. Par exemple, on peut avoir pour un joueur en particulier :

$$(u_\varnothing, u_\mathrm{L}, u_\mathrm{O}, u_\mathrm{LO}) = (0\,;0{,}3\,;0{,}9\,;1)\,.$$

Dans ce cas, on a choisi une convention de normalisation en utilisant deux options réelles (disponibles dans le jeu) et fixées : si on exprimait les préférences d'un deuxième joueur, on utiliserait ces mêmes options réelles comme points de référence.

Exemple B.11 (carrefour routier)

Alice et Bob envisagent de voyager ensemble. Ils se trouvent à un croisement routier où les directions possibles sont les points cardinaux Nord-Est (NE), Sud-Est (SE) et Ouest (O).

Dans l'absolu, la meilleure option pour Alice serait Sud (S) et la pire serait Nord (N). Si les hypothèses du théorème B.1 sont vérifiées, alors nous pouvons l'utiliser pour représenter ses préférences sur l'ensemble $\{\mathrm{NE}, \mathrm{SE}, \mathrm{O}, \mathrm{S}, \mathrm{N}\}$. Comme les utilités sont définies à deux constantes près, nous pouvons prendre pour convention que l'utilité de la pire option (N) sera 0 et que l'utilité de la meilleure option (S) sera 1. Par exemple :

$$(u_\mathrm{NE}, u_\mathrm{SE}, u_\mathrm{O}, u_\mathrm{S}, u_\mathrm{N}) = \left(\frac{1}{4}\,;\frac{3}{4}\,;\frac{1}{2}\,;1\,;0\right).$$

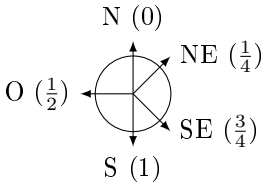


FIGURE B.7 – Carrefour routier : utilités d'Alice.

Finalement, les préférences d'Alice sur $\{\mathrm{NE}, \mathrm{SE}, \mathrm{O}\}$ peuvent être représentées dans $[0,1]^3$ par le vecteur $(u_\mathrm{NE}, u_\mathrm{SE}, u_\mathrm{O}) = (\frac{1}{4}\,;\frac{3}{4}\,;\frac{1}{2})$, comme on le représente sur la figure B.7.

Dans ce cas, les utilités comportent des informations supplémentaires par rapport à la situation du théorème B.1. En effet, dans le cadre du théorème B.1, un vecteur d'utilité comme $(u_\mathrm{NE}, u_\mathrm{SE}, u_\mathrm{O}) = (0\,;1\,;\frac{1}{2})$ aurait exactement le même sens que $(\frac{1}{4}\,;\frac{3}{4}\,;\frac{1}{2})$: dans les deux cas, l'agent préfère SE puis O puis NE et elle est indifférente entre O et une combinaison équiprobable de SE et NE. Mais, grâce à notre choix particulier, nous avons aussi enregistré, par exemple, qu'il est indifférent entre NE et une combinaison probabiliste de sa meilleure option imaginable avec probabilité $\frac{1}{4}$ et de sa pire option imaginable avec probabilité $\frac{3}{4}$. En particulier, Nord-Est est sa pire option *disponible*, mais pas sa pire option *imaginable*.

Du pur point de vue de la théorie des jeux, cela revient au même : qu'on prenne la représentation $(\frac{1}{4}\,;\frac{3}{4}\,;\frac{1}{2})$ ou $(0\,;1\,;\frac{1}{2})$, les préférences d'Alice sur les loteries de $\{\mathrm{NE}, \mathrm{SE}, \mathrm{O}\}$ sont identiques, tout comme ses intérêts stratégiques dans un jeu dont les résultats possibles sont $\{\mathrm{NE}, \mathrm{SE}, \mathrm{O}\}$.

Cependant, l'information que nous avons ajoutée prend sens dans un cadre étendu, ne se limitant pas aux seules options proposées et aux éventuels comportements stratégiquement optimaux. Par exemple, elle peut se révéler utile à la modélisation si nous estimons que, dans le problème étudié, deux personnes de vecteurs d'utilité respectifs $(\frac{1}{4}\,;\frac{3}{4}\,;\frac{1}{2})$ et $(0\,;1\,;\frac{1}{2})$, avec la convention ci-dessus, ne se comportent pas de la même façon, même si leurs intérêts stratégiques sont les mêmes.

Exemple B.12 (carrefour routier, suite)
Poursuivons l'exemple précédent. Imaginons, à présent, que pour Bob, la meilleure option imaginable soit Ouest (O) et que la pire soit Est (E). Si on utilise le même procédé que précédemment, il est possible qu'on obtienne le vecteur d'utilité $(v_{\mathrm{NE}}, v_{\mathrm{SE}}, v_{\mathrm{O}}) = (\frac{1}{4}\,;\frac{1}{4}\,;1)$, mais ces valeurs ne sont plus relatives à Sud et Nord (comme pour Alice), mais à Ouest et Est.

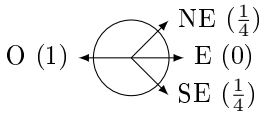


FIGURE B.8 – Carrefour routier : utilités de Bob.

Ce double exemple illustre ce que nous appelons une normalisation avec *options virtuelles meilleure et pire*. Plutôt que prendre pour références deux options concrètes et communes comme dans l'exemple B.10, cette normalisation suppose que tous les joueurs peuvent comprendre et appliquer des notions abstraites qui permettent d'établir une unité de mesure pour chacun qui est subjective mais repose sur des concepts concertés. Chaque agent considère les meilleure et pire options imaginables, même si elle ne sont pas disponibles, et normalise ses utilités par rapport à ces deux options virtuelles et subjectives. Bien sûr, la notion « d'imaginable » n'est pas une propriété mathématique : avant d'utiliser ce modèle pour un problème particulier, il faut définir ce qu'on considère « imaginable » par chaque joueur. De plus, déterminer de telles options virtuelles peut ne pas être suffisant pour une normalisation cohérente, comme le montre l'exemple suivant.

Exemple B.13 (jeu avec récompense financière)
On propose un jeu où l'on peut perdre 5 pièces ou gagner 10 ou 20 pièces : $\mathcal{C} = \{-5\,;10\,;20\}$. On s'intéresse à des agents qui jouent toujours selon l'espérance de gain financier. Ces agents ont donc une utilité pour chaque option qui est exactement (une fonction linéaire positive de) la récompense monétaire [4].

Considérons que, pour un agent, la meilleure option imaginable est de gagner une infinité de pièces et la pire, d'en perdre une infinité. Par conséquent, si nous voulons utiliser une normalisation avec options virtuelles meilleure et pire, nous devons avoir $u_{+\infty} = 1$ et $u_{-\infty} = 0$. Mais alors, il est manifestement impossible de normaliser les utilités de -5, 10 et 20. En effet, les différences d'utilité entre -5, 10 et 20 sont infinitésimales par rapport à celle entre $+\infty$ et $-\infty$, de sorte

4. Au moins, ils se comportent comme si c'était le cas, ce qui revient au même de notre point de vue d'observateur.

qu'on aurait la même utilité pour ces trois options : les préférences de l'agent se seraient pas représentées correctement. Le cœur du problème est que les préférences de l'agent, bien qu'archimédiennes sur $\{-5\,;10\,;20\}$, ne le sont pas sur $\{-5\,;10\,;20\,;+\infty\,;-\infty\}$.

Même si l'infinité n'est pas « imaginable », la capacité de concevoir une quantité arbitrairement grande d'argent (gagné ou perdu) mène à la même impossibilité d'utiliser un modèle avec options virtuelles meilleure et pire.

Afin de résumer les leçons des exemples précédents, une normalisation naturelle apparaît à chaque fois qu'il est pertinent de définir deux options virtuelles spéciales pour chaque agent (correspondant au meilleur et au pire dans nos exemples, mais ce n'est pas obligatoire) telles que :

— l'agent préfère strictement une option spéciale à l'autre ;
— même en incluant ces deux options spéciales, les préférences respectent encore les hypothèses du théorème B.1.

Alors, on peut choisir un vecteur d'utilité sur l'ensemble des options réelles et spéciales de telle sorte que ces dernières reçoivent les valeurs 0 et 1. Ceci donne une représentation normalisée des utilités des options réelles dans $\mathbb{R}^C$ (ou simplement $[0,1]^C$ quand les options spéciales correspondent à des options virtuelles meilleure et pire). Cet espace d'utilité $\mathbb{R}^C$ est doté d'une géométrie naturelle évidente, c'est pourquoi nous ne l'étudierons pas davantage, préférant nous concentrer sur le modèle où il y a une seule option spéciale, que nous allons introduire dans l'exemple suivant.

Exemple B.14 (élection et limite d'approbation)

À une élection, plusieurs candidats sont proposés : $\mathcal{C} = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$. Supposons qu'un électeur ait pour vecteur d'utilité $\mathbf{u} = (1\,;\frac{1}{3}\,;0)$ à normalisation près.

Pour chaque candidat, on pose à l'électeur la question suivante : « vous considérez-vous globalement favorable ou défavorable à ce candidat ? ». Quitte à choisir un vecteur particulier de $\widetilde{\widetilde{u}}$ pour représenter les préférences de l'électeur, il peut être intéressant de garder trace de la réponse à cette question, par exemple si notre but est d'étudier le mode de scrutin par assentiment. Nous pouvons envisager une convention simple : l'utilité est positive si et seulement si la réponse est positive. Par exemple, si l'électeur répond oui uniquement pour le candidat A, nous pouvons choisir comme vecteur d'utilité $(1\,;-1\,;-2)$.

On peut aller plus loin : supposons que l'électeur soit capable d'imaginer un candidat-type imaginaire S (candidat spécial) pour lequel il hésiterait entre une réponse positive ou négative. Si ses préférences sur $\{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{S}\}$ respectent les hypothèses du théorème B.1, on peut écrire des utilités et ajuster le β du théorème pour fixer par convention l'utilité du candidat spécial à 0. Par exemple :

$$(u_\mathrm{A}, u_\mathrm{B}, u_\mathrm{C}, u_\mathrm{S}) = (3\,;-1\,;-3\,;0).$$

La dernière coordonnée n'apporte aucune information puisqu'elle est purement conventionnelle : on peut simplement écrire $(u_\mathrm{A}, u_\mathrm{B}, u_\mathrm{C}) = (3\,;-1\,;-3)$. Remarquons qu'avec cette convention, ce vecteur n'a plus le même sens que $(1\,;-1\,;-2)$: par exemple, avec $(3\,;-1\,;-3)$, l'électeur préfère strictement une combinaison équiprobable de A et B plutôt que le candidat spécial S.

Nous appelons cette convention : normalisation avec pour option virtuelle la *limite d'approbation*. Dans ce cadre, le vecteur d'utilité est défini à une constante multiplicative strictement positive près, comme nous allons le formaliser ci-dessous. Là encore, l'information enregistrée par notre choix ne change rien aux stratégies optimales du joueur, mais on peut l'utiliser comme information comportementale dans un contexte d'application particulier.

De manière générale, on peut étendre ces idées à toute situation où on estime pertinent de définir une option virtuelle spéciale pour chaque agent, qui peut être subjective, telle qu'en incluant cette option spéciale, ses préférences respectent encore les hypothèses du théorème B.1.

En résumé, les méthodes présentées dans cette section ont pour but de prendre en compte divers facteurs additionnels, dont voici quelques exemples.

— En-dehors de son résultat principal, comme le candidat élu dans une élection, le jeu peut avoir un résultat secondaire, par exemple le « message » transmis par les électeurs, comme une satisfaction ou un mécontentement global à l'égard des candidats. Certes, on pourrait prendre comme espace des résultats le produit cartésien de l'espace des candidats et de celui des messages et appliquer le théorème B.1, mais cela complexifierait le modèle de façon importante. Plus simplement, on peut faire la différence entre un électeur $(-1\,;-0{,}9\,;-0{,}8)$, souhaitant exprimer son mécontentement, et un électeur $(0{,}8\,;0{,}9\,;1)$, globalement satisfait, même si leurs préférences sur les loteries de candidats et leurs stratégies optimales sont les mêmes.

— Les joueurs peuvent savoir que d'autres options auraient pu être proposées par les instances qui organisent le jeu. Cette connaissance peut avoir une influence sur leur façon de jouer, indépendamment de la nature de leurs stratégies optimales.

— Les joueurs peuvent avoir une variable comportementale, par exemple bonne ou mauvaise humeur, qui influence également leur façon de jouer, indépendamment de leurs préférences sur les loteries.

Ainsi, les normalisations proposées permettent d'ajouter dans la modélisation des agents des informations supplémentaires, notamment contextuelles ou comportementales, en exploitant les deux degrés de liberté laissés par le théorème B.1.

Pour conclure cette section, nous allons, à présent, donner un cadre mathématique du modèle avec limite d'approbation.

On note $\mathcal{C}' = [\![0, C]\!]$. Les éléments de $\mathcal{C} = [\![1, C]\!]$ sont toujours appelés *candidats*. L'élément 0 est appelé *candidat spécial*.

On note $\mathbb{L}' = \{(L_0, \dots, L_C) \in {\mathbb{R}_+}^{(C+1)} \text{ t.q. } \sum_{c=0}^{C} L_c = 1\}$. Les éléments de $\mathbb{L}$ sont identifiés à ceux de $\mathbb{L}'$ dont la coordonnée d'indice 0 est nulle. Le mélange probabiliste de deux loteries de $\mathbb{L}'$ est défini de façon naturelle.

Pour une relation binaire $\precsim$ sur $\mathbb{L}'$, les notions précédemment définies de complétude, transitivité, caractère archimédien et indépendance des alternatives non pertinentes s'étendent de manière évidente.

On dit que $\mathbf{u} = (u_1, \dots, u_C) \in \mathbb{R}^C$ est un *vecteur d'utilité avec limite d'approbation* représentant la relation binaire $\precsim$ si et seulement si, pour toute paire de loteries $L = (L_0, \dots, L_C)$ et $M = (M_0, \dots, M_C)$ de $\mathbb{L}'$:

$$L \precsim M \Leftrightarrow \sum_{c=1}^{C} L_c u_c \leq \sum_{c=1}^{C} M_c u_c.$$

On notera que les sommes ne prennent pas l'indice 0, ce qui revient à fixer conventionnellement l'utilité du candidat spécial à 0.

Nous allons, à présent, traduire le théorème de Von Neumann-Morgenstern B.1 dans ce nouveau modèle.

Proposition B.15

Soit $\precsim$ une relation binaire sur $\mathbb{L}'$.

Les conditions suivantes sont équivalentes.

1. *La relation binaire $\precsim$ est complète, transitive, archimédienne et indépendante des alternatives non pertinentes.*
2. *Il existe $\mathbf{u} \in \mathbb{R}^C$ un vecteur d'utilité avec limite d'approbation représentant $\precsim$.*

Quand elles sont vérifiées, $\mathbf{u}$ est défini à une constante multiplicative strictement positive près : si $\mathbf{u} \in \mathbb{R}^C$ est un vecteur d'utilité avec limite d'approbation représentant $\precsim$, alors $\mathbf{v} \in \mathbb{R}^C$ est un vecteur d'utilité avec limite d'approbation représentant $\precsim$ si et seulement si $\exists \alpha \in]0, +\infty[$ t.q. $\mathbf{v} = \alpha \mathbf{u}$.

Démonstration. C'est une conséquence du théorème B.1 : fixer l'utilité du candidat spécial à 0 est possible grâce au choix de β, mais on perd le degré de liberté en translation. □

De manière similaire à la relation d'équivalence $\approx$ dans le modèle classique, on définit, à présent, la relation d'équivalence $\sim$ par $\forall (\mathbf{u}, \mathbf{v}) \in (\mathbb{R}^C)^2$:

$$\mathbf{u} \sim \mathbf{v} \Leftrightarrow \exists \alpha \in]0, +\infty[\text{ t.q. } \mathbf{v} = \alpha \mathbf{u}.$$

On appelle *espace des utilités sur C candidats avec limite d'approbation*, et on note $\mathbb{V}$, l'ensemble quotient $\mathbb{R}^C / \sim$. Il s'agit de l'espace des demi-droites de $\mathbb{R}^C$ (auxquelles on ajoute le point d'indifférence), qui est similaire au modèle de *géométrie projective orientée* présenté par exemple par Stolfi (1987); Laveau et Faugeras (1996).

On appelle *projection canonique* de $\mathbb{R}^C$ vers $\mathbb{V}$ la fonction :

$$\widetilde{\pi} : \left| \begin{array}{ccc} \mathbb{R}^C & \rightarrow & \mathbb{V} \\ \mathbf{u} & \rightarrow & \{\mathbf{v} \in \mathbb{R}^C \text{ t.q. } \mathbf{v} \sim \mathbf{u}\}. \end{array} \right.$$

Pour $\mathbf{u} \in \mathbb{R}^C$ et $\widetilde{u} = \widetilde{\pi}(\mathbf{u})$, on note sans ambiguïté $\precsim_{\widetilde{u}}$ la relation binaire sur $\mathbb{L}'$ représentée par $\mathbf{u}$ au sens de la proposition B.15.

B.7.2 Correspondance entre les utilités avec limite d'approbation et les formes affines

De même que l'espace des utilités classiques correspond aux formes linéaires sur des bipoints de loteries (section B.2), nous allons montrer que l'espace des utilités avec limite d'approbation correspond aux formes affines sur les loteries.

Pour toute paire de loteries $L = (L_0, L_1, \ldots, L_C)$ et $M = (M_0, M_1, \ldots, M_C)$, on appelle *bipoint* de L à M le vecteur de $\mathbb{R}^C$:

$$\overrightarrow{LM} = (M_1 - L_1, \ldots, M_C - L_C).$$

Nous avons choisi de le définir comme un vecteur de $\mathbb{R}^C$ et non de $\mathbb{R}^{C+1}$. En effet, comme $\sum_{c=0}^{C} (M_c - L_c) = 0$, la coordonnée d'indice 0 n'apporterait aucune information supplémentaire.

En reprenant la figure B.4, pour deux loteries quelconques L et M de $\mathbb{L}'$, les points L et M sont identifiés aux points $(L_1, \ldots, L_C)$ et $(M_1, \ldots, M_C)$ du tétraèdre plein qui est délimité par le grand triangle (simplexe) et les axes du repère. Une loterie appartient à $\mathbb{L}$ si et seulement si le point correspondant appartient au simplexe.

On appelle *polytope tangent* de $\mathbb{L}'$ l'ensemble $\mathcal{T}'$ des bipoints de $\mathbb{L}'$:

$$\mathcal{T}' = \{\overrightarrow{LM}, (L, M) \in \mathbb{L}'^2\}.$$

Le polytope tangent $\mathcal{T}'$ est l'ensemble des bipoints du tétraèdre, vu comme une partie d'espace vectoriel.

En notant que le polytope $\mathcal{T}'$ contient un voisinage de l'origine dans l'espace vectoriel qu'il engendre, on obtient les deux observations suivantes.

Lemme B.16 ($\mathcal{T}'$ génère $\mathbb{R}^C$)
$\forall \mathbf{\Delta} \in \mathbb{R}^C, \exists \overrightarrow{LM} \in \mathcal{T}', \exists \lambda \in]0, +\infty[$ *t.q.* $\mathbf{\Delta} = \lambda \overrightarrow{LM}$.

Lemme B.17 ($\mathbb{L}$ génère $\mathbb{R}^C$)
$\mathbb{L}$ *étant vu comme une partie de l'espace vectoriel* $\mathbb{R}^C$, *on a* $vect(\mathbb{L}) = \mathbb{R}^C$.

De même que nous avons noté $\mathcal{H}^\star$ l'ensemble des formes linéaires sur l'espace vectoriel $\mathcal{H}$, notons $\mathbb{L}^\star$ l'ensemble des formes affines [5] sur $\mathbb{L}$.

Pour $\mathbf{u} = (u_1, \ldots, u_C) \in \mathbb{R}^C$, on appelle *forme affine associée à* $\mathbf{u}$ l'élément suivant [6] de $\mathbb{L}^\star$:

$$\langle \mathbf{u} | : \left| \begin{array}{ccc} \mathbb{L} & \to & \mathbb{R} \\ L = (L_1, \ldots, L_C) & \to & \langle \mathbf{u} | L \rangle = \sum_{c=1}^{C} u_c L_c. \end{array} \right.$$

On appelle *demi-espace positif associé à* $\mathbf{u}$:

$$\mathbf{u}^+ = \{\mathbf{\Delta} \in \mathbb{R}^C \text{ t.q. } \langle \mathbf{u} | \mathbf{\Delta} \rangle \geq 0\}.$$

Par simple relecture des définitions, on notera que pour $(L, M) \in \mathbb{L}^2$ et $\mathbf{u} \in \mathbb{R}^C$, on a les propriétés suivantes.

1. $L \precsim_{\widetilde{u}} M \Leftrightarrow \langle \mathbf{u} | \overrightarrow{LM} \rangle \geq 0 \Leftrightarrow \overrightarrow{LM} \in \mathbf{u}^{+\mathcal{H}}$.
2. $L \succsim_{\widetilde{u}} (0, \ldots, 0) \Leftrightarrow \langle \mathbf{u} | L \rangle \geq 0$.

Nous pouvons, à présent, énoncer la proposition suivante, qui présente un résultat de dualité similaire à celui donné par la proposition B.2 pour le modèle classique sans limite d'approbation.

Proposition B.18
Pour $(f, g) \in (\mathbb{L}^\star)^2$, *on note* $f \sim g$ *si et seulement si* $\exists \alpha \in]0, +\infty[$ *t.q.* $g = \alpha f$. *On note* $\widetilde{\pi}(f) = \{g \in \mathbb{L}^\star, f \sim g\}$.
Pour $(\mathbf{u}, \mathbf{v}) \in (\mathbb{R}^C)^2$, *on a :*

$$\mathbf{u} \sim \mathbf{v} \Leftrightarrow \langle \mathbf{u} | \sim \langle \mathbf{v} |.$$

L'application suivante est une bijection :

$$\Theta : \left| \begin{array}{ccc} \mathbb{V} & \to & \mathbb{L}^\star / \sim \\ \widetilde{\pi}(\mathbf{u}) & \to & \widetilde{\pi}(\langle \mathbf{u} |). \end{array} \right.$$

Démonstration. $\mathbf{u} \sim \mathbf{v}$
$\Leftrightarrow \exists \alpha \in]0, +\infty[$ t.q. $\mathbf{v} - \alpha \mathbf{u} = \mathbf{0}$
$\Leftrightarrow \exists \alpha \in]0, +\infty[$ t.q. $\forall L \in \mathbb{L}, \langle \mathbf{v} - \alpha \mathbf{u} | L \rangle = 0$ (cf. lemme B.17)
$\Leftrightarrow \langle \mathbf{u} | \sim \langle \mathbf{v} |$.
L'implication $\Rightarrow$ prouve que Θ est bien définie : en effet, si $\widetilde{\pi}(\mathbf{u}) = \widetilde{\pi}(\mathbf{v})$, alors $\widetilde{\pi}(\langle \mathbf{u} |) = \widetilde{\pi}(\langle \mathbf{v} |)$. L'implication $\Leftarrow$ assure que Θ est injective. Enfin, il est clair que Θ est surjective. □

5. Nous faisons un léger abus de vocabulaire pour simplifier : $\mathbb{L}^\star$ est l'ensemble des restrictions à $\mathbb{L}$ des formes de l'espace affine engendré par $\mathbb{L}$. Comme les formes affines et leur restrictions sont en bijection évidente, il n'y a pas de risque de confusion.

6. Nous utilisons la même notation que pour la forme linéaire associée à $\mathbf{u}$. Comme leurs expressions sont identiques, il n'y a pas de risque de confusion.

L'espace des utilités avec limite d'approbation peut donc être vu comme un quotient de l'espace $\mathbb{L}^\star$ des formes affines sur les loteries $\mathbb{L}$. Un vecteur d'utilité avec limite d'approbation peut être vu, à une constante strictement positive près, comme une forme affine déterminant, pour tout point de l'espace des loteries, avec quelle intensité il est apprécié par l'agent (par rapport à un niveau 0 qui est le candidat spécial).

B.7.3 Opérateurs opposé et somme dans l'espace des utilités avec limite d'approbation

En tant que quotient de $\mathbb{R}^C$, l'espace des utilités hérite d'opérations naturelles sur $\mathbb{R}^C$, la prise de l'opposé et la somme. Nous allons voir que, comme pour le modèle sans limite d'approbation, ces deux opérations quotient ont une signification intuitive concernant les préférences.

Dans $\mathbb{V}$, on définit l'opérateur *opposé* par :

$$-: \left| \begin{array}{ccc} \mathbb{V} & \to & \mathbb{V} \\ \widetilde{\pi}(\mathbf{u}) & \to & \widetilde{\pi}(-\mathbf{u}). \end{array} \right.$$

L'opérateur opposé est correctement défini et c'est une bijection. En effet, on a $\widetilde{\widetilde{\pi}}(\mathbf{u}) = \widetilde{\widetilde{\pi}}(\mathbf{v})$ si et seulement si $\widetilde{\widetilde{\pi}}(-\mathbf{u}) = \widetilde{\widetilde{\pi}}(-\mathbf{v})$. Comme dans l'espace des utilités habituel $\mathbb{U}$, considérer l'opposé correspond à renverser les préférences de l'agent, sans modifier leurs intensités relatives. Dans $\mathbb{V}$, en outre, l'opposé préserve les intensités relatives par rapport à la limite d'approbation.

Sur les parties de $\mathbb{V}$, on définit l'opérateur *somme* par :

$$\sum: \left| \begin{array}{ccc} \mathcal{P}(\mathbb{V}) & \to & \mathcal{P}(\mathbb{V}) \\ A & \to & \left\{ \widetilde{\pi}\left(\sum_{i=1}^n \mathbf{u}_i\right), n \in \mathbb{N}, (\mathbf{u}_1, \ldots, \mathbf{u}_n) \in \left(\widetilde{\pi}^{-1}(A)\right)^n \right\}. \end{array} \right.$$

De même que dans le modèle sans limite d'approbation, nous allons montrer que la somme est le lieu des vecteurs d'utilité qui préservent les préférences unanimes de la population.

Théorème B.19 (caractérisation de la somme dans l'espace des utilités avec limite d'approbation)
Soit $A \in \mathcal{P}(\mathbb{V})$ et $\widetilde{v} \in \mathbb{V}$.
Les conditions suivantes sont équivalentes.

1. $\widetilde{v} \in \sum A$.
2. $\forall (L, M) \in \mathbb{L'}^2 : (\forall \widetilde{u} \in A, L \precsim_{\widetilde{u}} M) \Rightarrow L \precsim_{\widetilde{v}} M$.

Démonstration. Soit $\mathbf{v} \in \widetilde{\pi}^{-1}(\widetilde{v})$. On a les équivalences suivantes.
— $\forall (L, M) \in \mathbb{L'}^2, (\forall \widetilde{u} \in A, L \precsim_{\widetilde{u}} M) \Rightarrow L \precsim_{\widetilde{v}} M$,
— $\forall \overrightarrow{LM} \in \mathcal{T}', \left(\forall \mathbf{u} \in \widetilde{\pi}^{-1}(A), \langle\, \mathbf{u} \,|\, \overrightarrow{LM} \,\rangle \geq 0\right) \Rightarrow \langle\, \mathbf{v} \,|\, \overrightarrow{LM} \,\rangle \geq 0$,
— $\forall \mathbf{\Delta} \in \mathbb{R}^C, \left(\forall \mathbf{u} \in \widetilde{\pi}^{-1}(A), \langle\, \mathbf{u} \,|\, \mathbf{\Delta} \,\rangle \geq 0\right) \Rightarrow \langle\, \mathbf{v} \,|\, \mathbf{\Delta} \,\rangle \geq 0$ (cf. lemme B.16),
— $\bigcap_{\mathbf{u} \in \widetilde{\pi}^{-1}(A)} \mathbf{u}^+ \subset \mathbf{v}^+$,
— $\mathbf{v}$ est dans le cône convexe de $\widetilde{\pi}^{-1}(A)$,
— $\widetilde{v} \in \sum A$.

□

B.7.4 Représentation riemannienne de l'espace des utilités avec limite d'approbation

Tout comme pour le modèle classique, nous allons exhiber les représentations riemanniennes qui sont cohérentes avec les propriétés naturelles de l'espace des utilités avec limite d approbation et avec la symétrie *a priori* entre les candidats (hormis le candidat spécial).

Pour les mêmes raisons que précédemment, le point d'indifférence $\widetilde{0}$ doit être exclu : en effet, son seul voisinage ouvert est $\mathbb{V}$ tout entier et aucune distance n'est compatible avec cette propriété. En revanche, $\mathbb{V}$ possède la même topologie que la sphère de dimension $C-1$, donc on peut le munir d'une distance.

La situation étant, ici, moins symétrique que dans le cas de $\mathbb{U}\setminus\{\widetilde{\widetilde{0}}\}$, nous avons besoin d'une famille de métriques plus diverses. Plutôt que de ne considérer que les produits scalaires proportionnels au produit canonique, on va considérer tous ceux qui :

— rendent $\mathcal{H}$ orthogonal à la direction $\mathbf{1}$;
— ont une restriction à $\mathcal{H}$ proportionnelle à la restriction du produit canonique ;

c'est-à-dire les $\lambda\psi_\rho$, où ψ_ρ est le produit scalaire qui coïncide sur $\mathcal{H}$ avec le produit canonique, tel que $\mathbf{1}$ soit orthogonal à $\mathcal{H}$ et que $\psi_\rho(\mathbf{1},\mathbf{1}) = \rho^2$.

On note ζ_ρ la métrique sur $\mathbb{V}\setminus\{\widetilde{0}\}$ induite par ψ_ρ.

Pour s'en faire une image, il suffit d'imaginer que les vecteurs d'utilité sont normalisés de façon à appartenir à l'ellipsoïde qui est l'image de la sphère unité par l'affinité orthogonale de facteur ρ le long de la direction $\mathbf{1}$. Ensuite, on mesure la distance entre deux points, non pas en utilisant la métrique induite par le produit scalaire canonique de $\mathbb{R}^C$, mais la métrique induite par le produit scalaire pour lequel cet ellipsoïde est la sphère de diamètre λ.

Par exemple, la distance entre deux points diamétralement opposés sera toujours $\lambda\pi$ (où π désigne l'habituel nombre pi, c'est-à-dire le demi-périmètre d'un cercle unité), ce qui ne serait pas le cas avec la métrique issue du produit scalaire canonique. Nous reviendrons sur cette construction dans la section B.7.5.

Le réel $\rho \in]0,+\infty[$ est donc le facteur de distorsion qui définit l'élongation ou l'aplatissement de l'ellipsoïde dans la direction $\mathbf{1}$. Le réel $\lambda \in]0,+\infty[$ est un facteur de dilatation global qui multiplie toutes les distances.

Théorème B.20 (représentation riemannienne de l'espace des utilités avec limite d'approbation)

On suppose $C \geq 3$. Soit ζ une métrique riemannienne sur $\mathbb{V}\setminus\{\widetilde{0}\}$.
Les conditions 1 et 2 suivantes sont équivalentes.

1. (a) *Pour tout couple de points non antipodaux $\widetilde{u},\widetilde{v} \in \mathbb{V}\setminus\{\widetilde{0}\}$ (c.à.d. $\widetilde{v} \neq -\widetilde{u}$), l'ensemble $\sum\{\widetilde{u},\widetilde{v}\}$ des éléments respectant les préférences unanimes de $\widetilde{u}$ et $\widetilde{v}$ est un segment de géodésique de ζ ; et*

   (b) *pour toute permutation σ de $[\![1,C]\!]$, l'action Φ_σ induite sur $\mathbb{V}\setminus\{\widetilde{0}\}$ par*

$$(u_1,\ldots,u_C) \to (u_{\sigma(1)},\ldots,u_{\sigma(C)})$$

   est une isométrie.

2. $\exists\rho \in]0,+\infty[, \exists\lambda \in]0,+\infty[$ *t.q.* $\zeta = \lambda\zeta_\rho$.

Démonstration. Comme dans le théorème B.6, on déduit de la condition 1a l'existence d'un produit scalaire ψ sur $\mathbb{R}^C$ induisant ζ.

Par la condition 1b, la direction $\mathbf{1}$ est orthogonale (au sens de ψ) à $\mathcal{H}$. En effet, le vecteur $(1, -1, 0, \dots, 0)$ est envoyé par la transposition $\Phi_{(12)}$ sur son opposé, et celle-ci étant une isométrie, l'angle entre $\mathbf{1}$ et $(1, -1, 0, \dots, 0)$ est droit. Le même argument s'applique à toutes les permutations de ce vecteur, lesquelles engendrent $\mathcal{H}$.

En particulier, les projections orthogonales sur $\mathcal{H}$ au sens de ψ et du produit scalaire canonique de $\mathbb{R}^C$ coïncident. Il s'ensuit que les Φ_σ permutent les projections des vecteurs de la base canonique sur $\mathcal{H}$. On peut donc raisonner comme dans la démonstration du théorème B.6 pour conclure que la restriction de ψ à $\mathcal{H}$ est proportionnelle au produit canonique. □

B.7.5 La représentation riemannienne en pratique

Pour tout réel $\rho \neq 0$, on note D_ρ la matrice de l'affinité orthogonale qui dilate d'un facteur ρ dans la direction de $\mathbf{1}$:

$$D_\rho = D_0 + \rho \frac{1}{C} J = I + \frac{\rho - 1}{C} J.$$

La représentation ellipsoïdale du théorème B.20 est alors l'image de la sphère unité par D_ρ (dont la bijection réciproque est $D_{(\rho^{-1})}$).

Pour $(\mathbf{u}, \mathbf{v}) \in \left(\mathbb{R}^C \setminus \{\mathbf{0}\}\right)^2$, la distance entre $\tilde{u}$ et $\tilde{v}$ au sens de la métrique $\lambda\zeta_\rho$ du théorème B.20 est :

$$d(\tilde{u}, \tilde{v}) = \lambda \arccos \left\langle \frac{D_{(\rho^{-1})}\mathbf{u}}{\|D_{(\rho^{-1})}\mathbf{u}\|} \,\middle|\, \frac{D_{(\rho^{-1})}\mathbf{v}}{\|D_{(\rho^{-1})}\mathbf{v}\|} \right\rangle.$$

En effet, on a deux possibilités :

— calculer le représentant de $\tilde{u}$ qui appartient à l'ellipsoïde puis le ramener sur la sphère par l'affinité $D_{(\rho^{-1})}$;

— calculer l'image $\mathbf{u}'$ de $\mathbf{u}$ par l'affinité $D_{(\rho^{-1})}$ puis le normaliser, c'est-à-dire calculer l'intersection de la demi-droite définie par $\mathbf{u}'$ avec la sphère unité.

Comme la sphère unité est l'image de l'ellipsoïde par l'affinité $D_{(\rho^{-1})}$, ces deux possibilités sont équivalentes. Choisir la seconde prouve immédiatement la formule proposée.

Si on veut tirer des vecteurs suivant une loi de probabilité uniforme sur $\mathbb{V} \setminus \{\tilde{0}\}$ (muni de la métrique $\lambda\zeta_\rho$), il suffit de tirer $\mathbf{u}$ selon une loi uniforme sur la sphère unité de $\mathbb{R}^C$ puis de renvoyer $D_\rho \mathbf{u}$.

Afin de nous forger une intuition correcte de cette représentation ellipsoïdale, la principale difficulté est qu'il ne faut pas considérer la métrique euclidienne canonique sur l'ellipsoïde en question, mais une métrique pour laquelle cette ellipse est une sphère, comme nous allons le voir sur la page suivante.

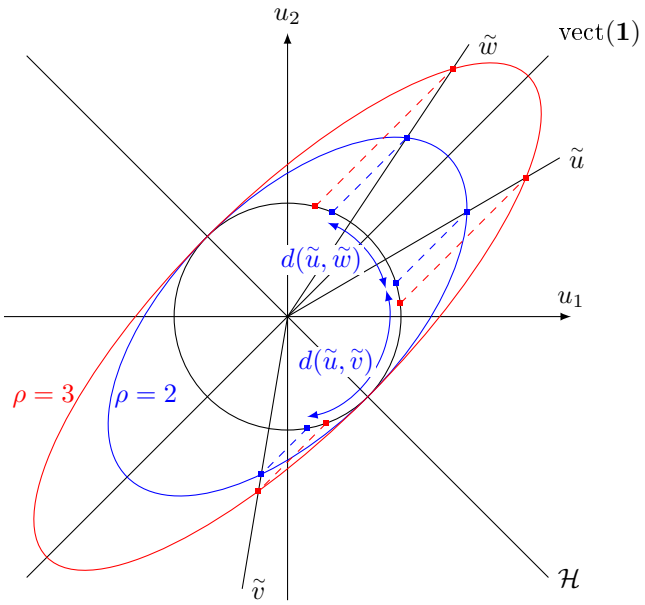


FIGURE B.9 – Représentations ellipsoïdales de l'espace des utilités avec limite d'approbation.

Dans la figure B.9, on montre comment la distance est calculée, par exemple entre $\widetilde{u}$ et $\widetilde{v}$. Le cas $\rho = 2$ est en bleu et $\rho = 3$ en rouge. Dans chaque cas, on prend l'intersection de la demi-droite $\widetilde{u}$ avec l'ellipsoïde, puis on utilise l'affinité $D_{(\rho^{-1})}$ pour ramener ce point sur la sphère. On fait de même avec $\widetilde{v}$ puis on mesure la distance à la surface de la sphère.

Quand ρ tend vers l'infini, on voit que $d(\widetilde{u},\widetilde{v})$ tend vers 0 alors que $d(\widetilde{u},\widetilde{w})$ tend vers un réel strictement positif (qui est $\lambda\pi$). Dans ce cas dégénéré, la distance entre deux points ne dépend que des utilités relatives sur les candidats, sans considérer la position par rapport à la limite d'approbation : de ce point de vue, cela ressemble au modèle sphérique à $C-2$ dimensions. Cependant, cela signifie également qu'une distance $\lambda\pi$ se trouve concentrée entre $\widetilde{u}$ et $\widetilde{w}$, et de manière générale entre tout couple de demi-droites situées de part et d'autre de vect(**1**) : finalement, si on utilise la loi de probabilité uniforme, on obtient presque sûrement $-\mathbf{1}$ ou $\mathbf{1}$.

Quand ρ tend vers 0, on obtient le comportement inverse : la distance entre deux points ne dépend que de l'approbation ou de la désapprobation globale (c'est-à-dire dans quel demi-espace on est par rapport à l'hyperplan $\mathcal{H}$), pas des appréciations relatives entre candidats. Cependant, si on utilise la loi de probabilité uniforme, on obtient presque sûrement un vecteur appartenant à $\mathcal{H}$.

Dans les applications pratiques du chapitre 7, nous choisissons toujours $\rho = 1$, c'est-à-dire une représentation sphérique. Il s'agit d'un cas intermédiaire entre les deux cas limites présentés ci-dessus, qui favorisent soit les utilités de même signe, soit des vecteurs d'utilités de somme nulle (donc avec presque sûrement des utilités des deux signes).

Annexe C

Systèmes de vote appliqués aux réseaux de télécommunication

Nous considérons un modèle simplifié où des nœuds d'un réseau veulent établir un chemin pour transmettre des paquets, avec un revenu partagé également entre les participants et un coût qui dépend du chemin utilisé. Nous examinons la possibilité que les nœuds utilisent un système de vote pour déterminer le chemin suivi.

Si les nœuds votent sincèrement, le vote par notation donne par définition l'optimum économique. Mais les nœuds ont intérêt à voter tactiquement : en vote par notation, nous montrons que les performances s'effondrent alors. En revanche, si on utilise VTI, le système est bien moins manipulable et les chemins choisis sont tout de même proches de l'optimum économique.

Cette annexe reprend pour l'essentiel les travaux présentés par Durand *et al.* (2013, 2014b).

Internet est devenu un écosystème économique où interviennent de nombreux acteurs concurrents. Afin de prendre des décisions pertinentes, tout en empêchant les acteurs économiques de manipuler le résultat naturel du processus de décision, la théorie des jeux est un cadre naturel. Les systèmes de vote représentent une alternative intéressante qui, à notre connaissance, n'a pas encore été considérée. Ils permettent à des entités concurrentes de décider entre différentes options. Dans ce chapitre, nous examinons leur utilisation pour la sélection d'un chemin d'un point à un autre dans un réseau qui fait intervenir plusieurs opérateurs, et nous analysons leur manipulabilité par vote tactique et leur efficacité économique. Nous montrons que VTI est bien plus efficace et résistant au vote tactique que le système de vote naturel pour trouver l'optimum économique, c'est-à-dire le vote par notation.

L'émergence de nouveaux services, couplée à la croissance constante des services existants, a mené à une explosion du trafic internet. Le réseaux des réseaux est devenu un énorme écosystème économique dont de nombreuses entreprises tirent des revenus. Ceci inclut des opérateurs de réseau, qui interconnectent leurs infrastructures pour constituer l'Internet, mais aussi des fournisseurs de service, qui monnaient leurs prestations au sein d'Internet. Dans ce contexte, il est important d'assurer la justesse des décisions qui impliquent de nombreux acteurs concurrents, avec le but d'atteindre une forme d'optimum économique global.

En effet, les participants d'un réseau décentralisé ont souvent à prendre des décisions globales basées sur des intérêts particuliers. Par exemple, les routes internet s'étendent sur plusieurs Systèmes Autonomes (AS) alors qu'elles résultent de décisions locales et en partie arbitraires. De nombreux champs de la théorie des jeux ont été proposés afin de mieux comprendre la prise de décision distribuée (Nisam *et al.*, 2007). Cependant, à notre connaissance, un de ces champs reste à explorer : les systèmes de vote, qui permettent à des entités concurrentes de prendre une décision entre plusieurs options.

Le but principal de ce chapitre est d'examiner l'utilisation de systèmes de vote au sein de l'écosystème économique d'Internet sur un cas d'usage donné. En particulier, nous nous concentrons sur la question de la manipulabilité, qui est cruciale dans un contexte de prise de décision. Nous savons qu'à part des cas dégénérés, tous les systèmes de vote sont manipulables (Gibbard, 1973; Satterthwaite, 1975). Cependant, les modes de scrutin ne sont pas égaux devant la manipulabilité.

La section C.1 définit et modélise le cas d'utilisation de cette annexe : l'établissement d'un chemin dans un réseau à plusieurs opérateurs. Il s'agit d'un modèle simplifié initialement élaboré par Ludovic Noirie, dont le but est avant tout d'étudier qualitativement les comportements des différents systèmes de vote sans s'appesantir sur la complexité technologique du système. Enfin, en section C.2, nous analysons les résultats obtenus, en illustrant comment et pourquoi les systèmes de votes peuvent être très intéressants pour une utilisation pratique dans un écosystème économique tel que celui d'un réseau à plusieurs opérateurs.

C.1 Présentation du modèle d'établissement d'un chemin par plusieurs opérateurs

Dans cette section, on définit le problème que nous souhaitons résoudre : l'établissement de chemin d'un point à un autre par plusieurs opérateurs. Nous expliquons comment les systèmes de vote peuvent être utilisés à cette fin, et comment les chemins candidats sont pré-sélectionnés. Puis nous définissons la modélisation des coûts et des gains et nous présentons comment les systèmes de vote sont utilisés et comment la manipulabilité est calculée.

C.1.1 Modélisation sous forme d'élection

Le problème à résoudre est le suivant : dans un réseau à plusieurs opérateurs, étant donné un nœud de départ (ingress) et un nœud d'arrivée (egress), quel chemin devrait être choisi si on prend en compte les préférences des opérateurs ?

La figure C.1 représente un exemple [1] d'interconnexion pour un réseau à plusieurs opérateurs, avec un opérateur par pays européen connecté aux membres d'un certain voisinage géographique.

Nous utilisons des systèmes de vote en utilisant le modèle suivant :

— Les électeurs sont les opérateurs.
— Les candidats sont les chemins de routage possibles pour une certaine demande ou un sous-ensemble d'entre eux.

1. Cet exemple n'est pas représentatif d'un véritable réseau à plusieurs opérateurs. Par exemple, une telle topologie plate de connexion diffère de la topologie hiérarchique BGP historique d'Internet. Néanmoins, notre but ici est de donner un modèle simple où les coûts dérivent d'une forme de métrique sous-jacente. C'est pourquoi se baser sur la géographie du réseau est un choix naturel. On notera que certaines études récentes montrent une évolution vers des topologies plus plates que dans le passé (Labovitz *et al.*, 2010; Dhamdhere et Dovrolis, 2010).

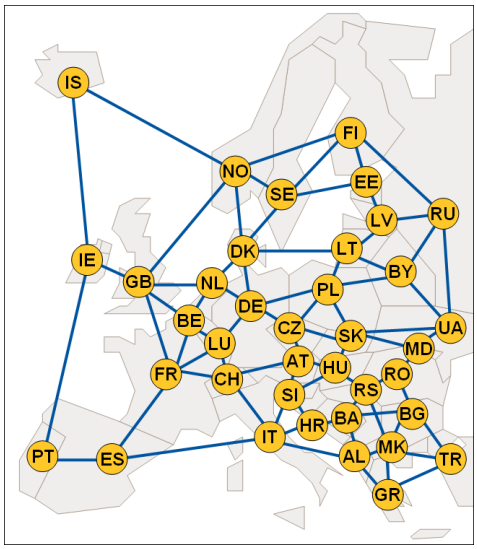


FIGURE C.1 – Exemple de réseau de connections entre plusieurs opérateurs

— Les préférences de chaque opérateur sont représentées par un vecteur d'utilité, à partir duquel un bulletin est défini.
— Le résultat de l'élection est le chemin sélectionné pour la requête client. N'importe quel système de vote peut être utilisé. Ici, nous supposons qu'il existe une entité fiable et indépendante, appelée le *superviseur*, en charge du processus de l'élection. Le superviseur collecte les bulletins de tous les opérateurs et calcule la fonction de dépouillement pour décider le chemin vainqueur (mais d'autres options sont possibles ; par exemple, les opérateurs peuvent participer coopérativement à l'ensemble du dépouillement).

Il est possible que certains opérateurs possèdent de l'information sur les utilités de leurs concurrents :

— par connaissance publique des utilités ;
— par coopération entre certains opérateurs (coalition) ;
— par inférence à partir des scrutins précédents (apprentissage) ;
— par espionnage ou interception d'information.

Les opérateurs peuvent utiliser leur connaissance pour mentir à propos de leurs propres préférences, de façon à améliorer leurs bénéfices, potentiellement au détriment de l'intérêt général. Par conséquent, étudier la manipulabilité des systèmes de vote est important dans ce contexte.

C.1.2 Sélection des chemins candidats

Dans notre cas d'étude, nous considérons l'interconnexion de $V = 38$ opérateurs de la figure C.1.

Une *demande* est une requête de connexion avec la seule contrainte de démarrer à un opérateur donné (ingress) et de terminer à un autre opérateur donné (egress).

Pour chaque demande, si l'on considère l'ensemble de tous les chemins possibles sans boucle, la taille de cet ensemble croît exponentiellement avec le nombre de nœuds dans le réseau. Pour pallier ce problème, nous appliquons la règle suivante pour limiter le nombre de chemins proposés. Pour appliquer cette règle, seule la connaissance de la topologie du réseau est nécessaire.

— Le superviseur fixe un nombre minimal de chemin candidats $C_{\min}$ et un seuil initial $\varepsilon_{\min}$.
— Il connaît la topologie des liens (mais pas les coûts). Il calcule le nombre minimal de sauts $h_{\min}$ nécessaires pour satisfaire la demande.
— Il considère tous les chemins sans boucle possédant un nombre de sauts inférieur ou égal à $h_{\min} + \varepsilon$, où ε est le plus petit seuil supérieur ou égal à $\varepsilon_{\min}$ qui sélectionne au moins $C_{\min}$ candidats.

Il est important de noter que :

— Les chemins candidats sont complètement déterminés par la demande et par les paramètres $(C_{\min}, \varepsilon_{\min})$. Leur nombre exact dépend de la demande.
— Pour chaque chemin candidat c, seul un sous-ensemble $\mathcal{V}_c$ des opérateurs $v \in \mathcal{V}$ est concerné par le chemin candidat.
— Parmi les opérateurs, certains peuvent être concernés par seulement un sous-ensemble des candidats.

Pour les évaluations numériques, nous utilisons deux jeux de paramètres : $(C_{\min}, \varepsilon_{\min}) \in \{(5,0);(10,1)\}$. Le premier réglage donne une moyenne de $9,94$ chemins candidats par requête (minimum 5, maximum 43), et le second réglage donne une moyenne de $21,25$ chemins par requête (minimum 10, maximum 127).

C.1.3 Modèle des coûts et des gains pour les opérateurs

Nous définissons les utilités de chaque opérateur comme la différence entre ses gains et ses coûts pour chaque chemin possible. Divers modèles de coûts et de gains pourraient être définis, mais le but principal de cette étude est d'évaluer la manipulabilité des systèmes de vote dans ce contexte, donc nous utilisons les modèles simplifiés suivants. Si on souhaitait tester des modèles plus proches de la réalité, il n'y aurait aucun obstacle technique à le faire. Selon notre intuition, cela n'apporterait que des changements mineurs aux tendances observées, mais il serait intéressant de valider cette intuition dans des travaux futurs.

Coût

Le coût pour un opérateur v de $\mathcal{V}_c$ pour porter le chemin c est noté $\alpha_{v,c}$. Nous définissons $\alpha_{v,c}$ comme la somme de la moitié du coût de la connexion entrante (nulle pour l'opérateur ingress) et de la moitié du coût de la connexion sortante (nulle pour l'opérateur egress). Pour le coût d'un connexion entre deux opérateurs adjacents v et v', nous choisissons une fonction linéaire $X_0 + \delta(v,v')/\delta_0$ de la distance géographique $\delta(v,v')$ entre v et v' (v et v' étant les capitales des pays dans notre exemple de réseau). Dans notre étude numérique, nous considérons trois options pour le coût : dominé par le coût constant X_0 ($X_0 = 1$ et $\delta_0 = 100 \times \max(d(v,v'))$), purement linéaire ($X_0 = 0$, $\delta_0 = \max(\delta(v,v'))$) et intermédiaire ($X_0 = 1$, $\delta_0 = \max(\delta(v,v'))/3$).

Gain

Concernant le gain pour les opérateurs, nous considérons que le client paie un montant fixé A pour une demande donnée (*flat fare*). Si le chemin c est sélectionné, ce montant A est distribué également entre les opérateurs concernés (c'est-à-dire les membres de $\mathcal{V}_c$). Nous avons fixé la valeur de A du prix utilisateur de telle sorte que, quand on considère le chemin de moindre coût pour chaque requête, le revenu global moyen représente 140% du coût global moyen (ce qui se traduit par un bénéfice de 40%).

Utilité

La valeur sincère d'utilité pour l'opérateur v pour porter le chemin candidat c est définie comme le revenu net (positif ou négatif) pour l'opérateur dans le cas où ce chemin candidat est sélectionné

- $u_{v,c} = \frac{A}{card(\mathcal{V}_c)} - \alpha_{v,c}$ si $v \in \mathcal{V}_c$,
- $u_{v,c} = 0$ si $v \notin \mathcal{V}_c$ (un opérateur est donc indifférent entre les chemins auxquels il ne participe pas).

Revenu total

Le revenu total d'un opérateur donné (par unité de temps) dépend de la distribution des demandes. Dans notre étude, nous considérons que les demandes sont uniformément distribuées sur les paires d'opérateurs ingress et egress.

C.1.4 Systèmes de vote étudiés

Parmi les systèmes de vote présentés dans le chapitre 1, nous en considérons deux en particulier, que nous rappelons tout en indiquant les spécificités de leur application à l'établissement d'un chemin dans le réseau.

Vote par notation (VN)

Avec les utilités définies précédemment, le système de vote le plus naturel est celui qui maximise le revenu global, c'est-à-dire la somme des utilités de tous les opérateurs qui contribuent au chemin sélectionné. Ceci correspond exactement au vote par notation, c'est pourquoi celui-ci sera le système de référence naturel. En détails :

- Les opérateurs donnent leurs utilités au superviseur ;
- Le superviseur calcule la somme des notes pour chaque chemin candidat et sélectionne le chemin avec la valeur maximale.

Un des inconvénients de VN est que les opérateurs doivent donner toute l'information à propos de leurs coûts au superviseur. Même s'il s'agit d'une entité fiable et indépendante, il est possible qu'ils préfèrent éviter divulguer ce type d'information.

VTI

De nombreux autres systèmes de vote peuvent être appliqués. Des travaux précédents (par exemple Chamberlin *et al.* (1984) ou Walsh (2010a)), ainsi que ceux de cette thèse, suggèrent que VTI fait partie des systèmes de vote connus les moins manipulables. C'est pourquoi nous utilisons VTI comme deuxième système de vote pour notre cas d'étude[2]. La sélection de chemin avec *VTI* fonctionne de la façon suivante.

- Chaque électeur (opérateur) donne son ordre de préférence sur les chemins candidats qui le concernent et, pour chacun d'entre eux, si elle l'approuve (gain financier) ou le désapprouve (perte financière) de façon à les placer par rapport aux chemins auxquels il ne participe pas. De cette façon, chaque opérateur est forcé d'exprimer une indifférence entre les chemins auxquels il ne participe pas.

2. En réalité, nous avons aussi essayé d'autre systèmes de vote dans des expériences qui ne sont pas présentées ici, et vérifié que VTI était le système de vote le plus prometteur en termes de manipulabilité, à part notamment Condorcet-VTI qui est plus complexe et dont les performances sont similaires.

— Le superviseur calcule le dépouillement selon la règle habituelle de VTI, avec la règle suivante en cas d'indifférence entre plusieurs candidats : lors de chaque tour de VTI, pour un opérateur donné, son vote est divisé également entre les candidats encore en lice qui sont à égalité en tête de son bulletin. À chaque tour, le superviseur élimine le chemin candidat avec le moins de votes (le nombre de votes n'est pas nécessairement entier). En cas d'égalité, on élimine le chemin candidat de plus grand indice.

C.1.5 Algorithmes de manipulation

Tout d'abord, pour les deux systèmes de vote, la manipulation du scrutin par un opérateur est limitée aux chemins candidats qui le concernent : l'opérateur ne peut pas prétendre aimer ou ne pas aimer un chemin candidat pour lequel il devrait manifestement être indifférent.

VN

Quand les électeurs préférant c à w essaient de faire gagner c, leur meilleure stratégie est manifestement la stratégie triviale : donner la note maximale à c et la note minimale aux autres candidats, sauf pour les chemins auxquels ils ne participent pas, pour lesquels ils sont contraints de donner une note égale à 0. Pour cette raison, la manipulation par coalition (MC) et la manipulation triviale (MT) sont équivalentes. Dans notre étude, la note maximale est fixée à A et la note minimale autorisée est $-A$.

VTI

Décider s'il existe une façon de manipuler VTI est bien plus difficile. En fait, le problème est connu pour être $\mathcal{NP}$-difficile (Bartholdi et Orlin, 1991). Par conséquent, nous utilisons des méthodes simples pour tester la manipulabilité d'une demande donnée. Quand les tests ne sont pas conclusifs pour une demande donnée, on peut seulement répondre *peut-être*, mais quand on considère toutes les demandes, cela permet de donner des bornes inférieure et supérieure de la manipulabilité.

Afin de prouver que VTI est manipulable, nous essayons uniquement la manipulation triviale. Ceci donne une borne inférieure de la manipulabilité.

Afin de prouver que VTI n'est pas manipulable, notre algorithme est une adaptation de Coleman et Teague (2007). La difficulté provient du fait que les électeurs ont des égalités imposées (pour les chemins candidats auxquelles ils ne participent pas) et qu'il est possible que leur voix soit divisée entre plusieurs candidats à un tour de dépouillement donné. L'idée de notre adaptation est d'utiliser une variante du système de vote qui donne davantage de pouvoir aux manipulateurs et pour laquelle MC peut être calculée exactement. Si le système modifié ne peut pas être manipulé, alors VTI ne peut pas être manipulé non plus. En détails :

— À chaque tour, on autorise les manipulateurs à changer leur vote (comme dans le scrutin exhaustif). L'intérêt est qu'éliminer le candidat c puis d ou d puis c mène à la même situation, quel que soit l'historique pour y parvenir. Ceci facilite le fait d'avoir une approche itérative au lieu d'une approche récursive.

— À chaque tour, chaque manipulateur peut diviser son vote entre plusieurs candidats, y compris de manière non égale, par exemple $\frac{1}{3}$ vote pour un candidat et $\frac{2}{3}$ vote pour un autre. Ceci permet une approche en *water-filling* et évite les problèmes de type sac à dos.

— On autorise les électeurs à mentir, y compris à propos des chemins auxquels ils ne participent pas. Par conséquent, tous les manipulateurs sont égaux en droit et on n'a pas besoin de les gérer individuellement.

Avec ces modifications, on peut gérer le groupe des manipulateurs globalement : à chaque tour, on peut diviser leurs votes comme on le souhaite entre les candidats, de façon à éliminer le candidat voulu.

Quand la manipulation est impossible avec ces règles adaptées, elle est aussi impossible avec les règles données en C.1.4 : ceci fournit une borne supérieure pour la manipulabilité de VTI.

Comme nous le verrons dans la prochaine section, les bornes inférieure et supérieure ont tendance à être proches l'une de l'autre, ce qui donne une bonne estimation de la manipulabilité.

Remarque En fait, l'algorithme précédent reste très coûteux (potentiellement d'ordre 2^C). Quand il y a plus de 25 candidats et que la manipulation triviale ne fonctionne pas, nous n'essayons pas de prouver l'impossibilité et nous considérons directement le test non conclusif.

C.2 Manipulabilité et efficacité des systèmes de vote étudiés

Dans cette section, nous analysons les résultats obtenus sur le cas de réseaux à plusieurs opérateurs décrit dans la section C.1.

Aussi bien pour *VN* que pour *VTI*, nous nous concentrons d'abord sur un scénario spécifique qui servira de référence, puis nous étendons les résultats en changeant divers paramètres. Nous observons la manipulabilité et l'efficacité économique, avec des bulletins sincères et non-sincères.

C.2.1 Scénario de référence

Notre scénario de référence est le cas avec $(C_{\min}, \varepsilon_{\min}) = (5, 0)$ (en moyenne $9,94$ chemins candidats par demande) et le modèle de coût intermédiaire ($X_0 = 1$, $\delta_0 = \max(d(v, v'))/3$). Pour ce scénario, on mesure :

— *Efficacité sincère* : revenu net global, en pourcentage du revenu net global optimal.

— *Manipulabilité* : proportion des demandes qui sont MC. Pour le vote par notation, le calcul est exact. Pour VTI, nous utilisons MT comme borne inférieure et la variante présentée en section C.1.5 pour donner une borne supérieure.

— *Efficacité manipulée* : efficacité du système quand des manipulations sont permises. Pour une demande donnée, plusieurs manipulations peuvent se produire. Nous pouvons mesure l'efficacité manipulée moyenne, quand un chemin bénéficiant d'une manipulation possible est choisi au hasard, ou la pire situation si pour chaque demande manipulable, on choisit la manipulation qui minimise le revenu global net (on se souviendra cependant que par définition, le revenu net de chaque manipulateur est nécessairement augmenté).

Les résultats sont présentés dans la table C.1. VTI présente de nombreux avantages par rapport au vote par notation :

— En vote sincère, VN donne l'optimum économique par définition. Bien que VTI n'ait pas pour cible intrinsèque l'optimum économique, ce système

Système de vote	VN	VTI
Efficacité sincère	100%	95 %
Manipulabilité	96%	$< 20\%$ (MT : 18%)
Efficacité manipulée (moyenne)	37%	90%
Efficacité manipulée (pire cas)	-75%	89%

TABLE C.1 – Résultats principaux pour le scénario de référence ($(C_{\min}, \varepsilon_{\min}) = (5, 0), (C_0, \delta_0) = (1, \max(d(v, v'))/3)$)

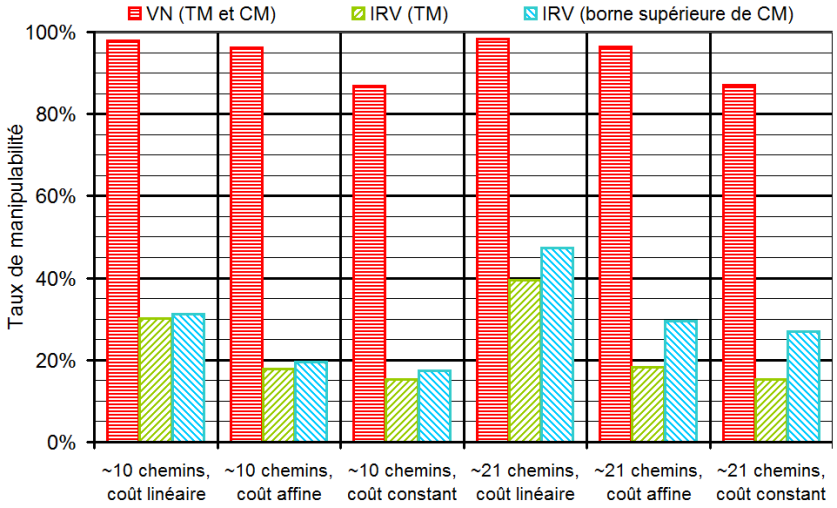


FIGURE C.2 – Manipulabilité de VTI comparée au vote par notation

parvient à atteindre 95% de l'optimum. Cette légère perte d'efficacité peut être vue comme le prix de la robustesse, comme nous allons le voir.

— Presque toutes les demandes (96%) peuvent être manipulées si le vote par notation est utilisé, contre moins de 20% pour VTI.

— En ce qui concerne l'impact sur l'efficacité, la dégradation est très élevée pour le vote par notation : elle descend à 37% de l'optimum en moyenne, et à -75% dans le pire scénario. De son côté, VTI maintient un robuste 90% en moyenne (89% dans le pire scénario).

Par conséquent, on retiendra que VTI est bien moins sujet à manipulation que le vote par notation, et que, même quand une manipulation existe, son impact sur le bien-être global est supportable.

C.2.2 Influence des paramètres

Nous avons proposé en section C.1 deux paramètres de limitation des chemins candidats et trois modèles de coût, donc nous avons six paramétrages possibles.

Manipulabilité

La figure C.2 indique la manipulabilité des scénarios considérés.

Pour les deux systèmes de vote, plus il y a de candidats, plus c'est manipulable. Ceci ne nous surprend guère, puisque plus de candidats signifie qu'il y a plus

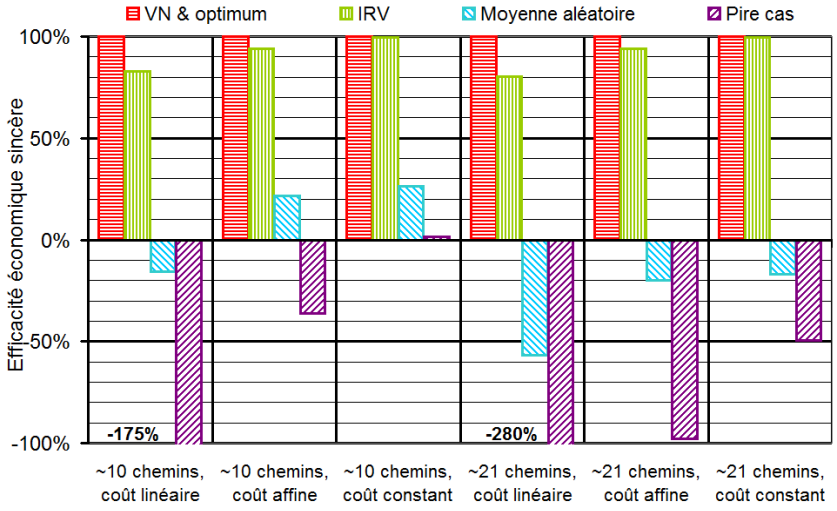


FIGURE C.3 – Efficacité de VTI et du vote par notation en vote sincère

de chemins pouvant bénéficier de manipulation et aussi plus d'espace d'expression dans les bulletins. Par conséquent, il est recommandable que le superviseur limite le nombre de candidats proposés afin de diminuer la manipulabilité, tout en conservant assez de candidats pour permettre la sélection juste et pertinente d'un chemin.

En ce qui concerne le modèle de coût, on observe que plus il est plat, plus la manipulabilité est basse (pour les deux systèmes de vote). Mais le résultat le plus intéressant est que la manipulabilité de VTI est bien plus basse que celle du vote par notation dans tous les cas. Alors que la manipulation de VN est toujours supérieure à 85%, celle de VTI est aux environs de 30% ou moins sauf pour un scénario (entre 40% et 48% pour le plus grand nombre de chemins candidats et avec un coût purement linéaire).

On notera que pour VTI, les bornes inférieure et supérieure de manipulabilité sont relativement proches. La différence croît avec le nombre de chemins candidats, mais c'est dû à la méthode utilisée pour calculer la borne supérieure (voir la section C.1), qui consiste à abandonner l'évaluation pour les demandes qui donnent lieu à trop de chemins candidats.

Efficacité économique sincère

La figure C.3 donne l'efficacité sincère de chaque système de vote pour les scénarios considérés. Par définition, VN donne 100% d'efficacité économique. Mais VTI donne une efficacité économique proche de cet optimum. Que ce soit pour un nombre de chemins candidats faible ou élevé, cette efficacité est d'environ 80% pour le modèle de coût purement linéaire, environ 95% pour le modèle de coût intermédiaire, et plus de 99% pour le modèle de coût constant. Ceci confirme que, même s'il est possible que VTI donne un chemin qui n'est pas optimal pour le revenu économique global, le chemin sélectionné est relativement proche du choix optimal.

À titre de référence, la figure C.3 indique également l'efficacité obtenue quand le chemin est choisi au hasard parmi les candidats, et quand le pire candidat est choisi.

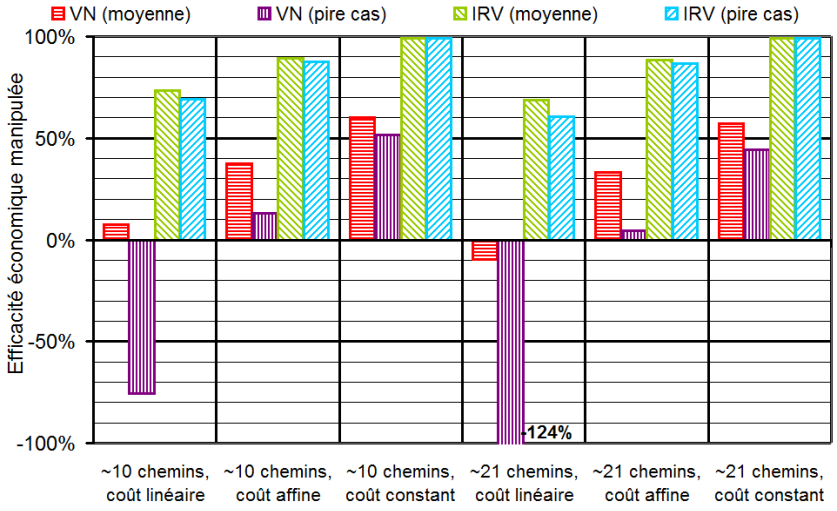


FIGURE C.4 – Efficacité de VTI et du vote par notation avec manipulation

Efficacité manipulée

La figure C.4 montre les efficacités manipulées. Pour tous les scénarios, on observe un gain spectaculaire en choisissant VTI plutôt que le vote par notation. Pour VTI, l'efficacité économique n'est que légèrement dégradée par rapport au cas sincère, et le pire cas n'est jamais loin du cas moyen. Pour le vote par notation, l'efficacité économique est largement dégradée par rapport au vote sincère : alors que l'efficacité économique sincère est par définition de 100%, la manipulation la fait descendre à environ 0% en moyenne et de grandes valeurs négatives en pire cas pour le modèle de coût purement linéaire ; environ 35% en moyenne et environ 8% en pire cas pour le modèle de coût intermédiaire ; et à environ 60% en moyenne et 50% en pire cas pour le modèle de coût constant.

Tous ces résultats confirment que VTI est vraiment plus sûr pour préserver les bénéfices économiques que le vote par notation.

C.3 Conclusion

Les résultats théoriques existants montrent que, sauf pour quelques cas dégénérés, tout système de vote est susceptible, dans certains scénarios, d'être manipulé, même par un seul électeur. Dans ce chapitre, nous avons proposé de quantifier la manipulabilité dans des scénarios pratiques et de mesurer ses effets par rapport au bien-être global. Nous nous sommes concentrés sur l'établissement d'un chemin d'un point à un autre dans un réseau à plusieurs opérateurs. Nous avons comparé deux systèmes de vote :

— *Vote par notation* : VN maximise le revenu net global du système, à condition que les opérateurs (les électeurs) donnent leurs utilités sincères (leurs revenus nets) sur les chemins candidats proposés.

— *VTI* : basé sur des ordres faibles de préférence avec des tours d'élimination successifs, VTI a la réputation d'être moins manipulable (Chamberlin *et al.*, 1984; Walsh, 2010a), ce que nous avons pu confirmer dans nos travaux.

Notre étude souligne l'intérêt des systèmes de vote dans le contexte de l'écosystème économique d'Internet où de nombreux joueurs concurrents sont impliqués, avec le cas d'usage pratique d'une sélection de chemin d'un point à un autre.

Nous avons aussi observé que tous les systèmes de vote ne sont pas équivalents, loin de là. Quand les opérateurs peuvent mentir sur leurs préférences, VTI est nettement meilleur que le vote par notation pour maximiser le gain économique :

— Le taux de manipulabilité de VTI peut être aussi bas que 20% alors que, dans le même scénario, VN est proche de 100%.
— Avec VTI, les opérateurs n'ont pas besoin de donner toute l'information pour que le scrutin soit dépouillé, l'ordre de préférence avec limite d'approbation suffit, alors que pour VN, ils doivent donner toute l'information sur leurs coûts.
— Bien que VTI ne soit pas conçu intrinsèquement pour viser le choix économiquement optimal, il en est très proche en pratique. Le prix de la non-manipulabilité est faible (de l'ordre de 5% dans notre étude).
— En ce qui concerne les manipulations, la dégradation de l'efficacité économique est limitée avec VTI, alors qu'elle est très importante avec VN.

Travaux futurs

Les travaux présentés dans cette annexe ouvrent de nombreuses opportunités pour des travaux futurs :

— Analyse de la sensibilité des résultats par rapport à divers paramètres, divers modèles de coûts et de revenu.
— Analyse d'autres systèmes de vote.
— Identification d'autres cas d'usage dans le contexte d'Internet, avec d'autres types d'écosystèmes.

En effet, le cadre des systèmes de vote peut être utile dans toute situation où on peut identifier des électeurs qui ont besoin de se décider entre plusieurs options. Nous avons montré dans notre cas d'étude qu'en choisissant correctement le système de vote, on peut limiter la manipulabilité par vote tactique d'une coalition des électeurs et préserver le revenu pour l'écosystème économique global.

Glossaire des notations

Symboles non alphabétiques

$[\alpha, \beta[$	Intervalle réel de α inclus à β exclu.
$[\![j, k]\!]$	Intervalle entier de j à k inclus.
$\lfloor \alpha \rfloor$	Partie entière du réel α par défaut.
$\lceil \alpha \rceil$	Partie entière du réel α par excès.
$\lvert \mathcal{A}(v) \rvert$	Cardinal des électeurs v qui vérifient l'assertion $\mathcal{A}(v)$.
$\pi(A \mid B)$	Probabilité conditionnelle de l'événement A sachant B.

Alphabet grec

μ	La loi de la variable P (sauf mention explicite du contraire).
π	Une *culture* sur l'espace électoral Ω. Plus généralement, une mesure de probabilité.
$\tau^{\pi}_{\mathrm{MC}}(f)$	Taux de manipulabilité par coalition du système de vote f en culture π.
Ω	Ensemble $\prod_{v \in \mathcal{V}} \Omega_v$ des configurations ω possibles. Aussi utilisé comme raccourci de notation pour un espace électoral $(V, C, \Omega, \mathrm{P})$.
Ω_M	Ensemble des états ω_M possibles pour les électeurs d'un ensemble M.
Ω_v	Ensemble des états ω_v possibles pour l'électeur v.

Alphabet latin

$C \in \mathbb{N} \setminus \{0\}$	Nombre de candidats.
$\mathcal{C}$	Ensemble $[\![1, C]\!]$ des indices des candidats.
$\mathrm{card}(E)$	Cardinal de l'ensemble E.

$D(\omega)$	Matrice des duels en ω. Le coefficient d'indices c et d est noté $D_{cd}(\omega)$ ou, en abrégé, D_{cd}.
$\mathcal{F}_{\mathcal{C}}$	Ensemble des ordres stricts faibles sur $\mathcal{C}$.
f	Un système de vote basé sur l'état (SVBE), c'est-à-dire une fonction $\Omega \to \mathcal{C}$. Dans le cas d'un système de vote général, f désigne sa fonction de dépouillement $\mathcal{S}_1 \times \ldots \times \mathcal{S}_V \to \mathcal{C}$.
f^*	Condorcification de f.
$f^{\text{adm}}, f^{!\text{adm}}$	Variantes de condorcification de f basées sur la notion de candidat Condorcet-admissible.
$f^{\text{faible}}, f^{!\text{faible}}$	Variantes de condorcification de f basées sur la notion de vainqueur de Condorcet faible.
f^{rel}	Condorcification relative de f.
$f^{\mathcal{M}}$	$\mathcal{M}$-condorcification de f.
f_y	Tranche de f par une méthode de tranchage y.
$c\ \mathrm{I}_v\ d$	L'électeur v est indifférent entre c et d.
Id	La fonction identité (le contexte précise dans quel ensemble).
$\mathcal{L}_{\mathcal{C}}$	Ensemble des ordres stricts totaux sur $\mathcal{C}$.
$\mathcal{M}$	Une famille de collections de coalitions.
$\mathcal{M}_c \in \mathcal{P}(\mathcal{P}(\mathcal{V}))$	Une collection de coalitions dites *gagnantes* pour le candidat c.
$\mathrm{Manip}_{\omega}(\mathrm{w} \to c)$	Ensemble des électeurs qui préfèrent c à w. En abrégé, $\mathrm{Manip}(\mathrm{w} \to c)$.
MC_f	Ensemble des configurations ω où f est manipulable (ou fonction indicatrice de cet ensemble).
$\mathrm{moy}(x_1, \ldots, x_k)$	Moyenne arithmétique de $x_1, \ldots, x_k$.
P	Fonction $\Omega \to \mathcal{R}$ qui, à l'état ω de la population, associe le profil $\mathrm{P}(\omega) = (\mathrm{P}_1(\omega_1), \ldots, \mathrm{P}_V(\omega_V))$.
$c\ \mathrm{P}_v\ d$	L'électeur v préfère c à d.
$c\ \mathrm{P}_{\text{abs}}\ d$	c a une victoire absolue contre d : $\lvert c\ \mathrm{P}_v\ d\rvert > \frac{V}{2}$.
$c\ \mathrm{P}_{\text{rel}}\ d$	c a une victoire relative contre d : $\lvert c\ \mathrm{P}_v\ d\rvert > \lvert d\ \mathrm{P}_v\ c\rvert$.
$c\ \mathrm{P}_{\mathcal{M}}\ d$	c a une $\mathcal{M}$-victoire contre d : $\{v \text{ t.q. } c\ \mathrm{P}_v\ d\} \in \mathcal{M}_c$.
$c\ \mathrm{PM}_v\ d$	L'électeur v préfère c à d et vice versa (impossible si P_v est antisymétrique).
$c\ \mathrm{PP}_v\ d$	L'électeur v préfère c à d mais pas d à c (synonyme de $c\ \mathrm{P}_v\ d$ si P_v est antisymétrique).
$\mathcal{R}$	Ensemble ${\mathcal{R}_{\mathcal{C}}}^V$ dont un élément (profil) représente les relations binaires de préférence de toute la population des électeurs.
$\mathcal{R}_{\mathcal{C}}$	Ensemble des relations binaires sur $\mathcal{C}$.
$\mathrm{Sinc}_{\omega}(\mathrm{w} \to c)$	Ensemble des électeurs qui ne préfèrent pas c à w. En abrégé, $\mathrm{Sinc}(\mathrm{w} \to c)$.
$V \in \mathbb{N} \setminus \{0\}$	Nombre d'électeurs.
$(V, C, \Omega, \mathrm{P})$	Un espace électoral. En abrégé, Ω.
$\mathcal{V}$	Ensemble $[\![1, V]\!]$ des indices des électeurs.

$\mathrm{vect}(E)$	Sous-espace vectoriel engendré par E, où E est une partie d'espace vectoriel.
$\mathcal{Y}$	Ensemble $\prod_{v\in\mathcal{V}} \mathcal{Y}_v$ des méthodes de tranchage y pour toute la population d'électeurs.
$\mathcal{Y}_v$	Ensemble $\{y_v : \mathrm{P}(\Omega_v) \to \Omega_v \text{ t.q. } \mathrm{P}_v \circ y_v = \mathrm{Id}\}$ des méthodes de tranchage y_v pour l'électeur v.

Sigles et abréviations

Bald.	Méthode de Baldwin.
BI	Méthode de Bucklin itérée.
Bor.	Méthode de Borda.
Buck.	Méthode de Bucklin.
BulMaj	Critère du bulletin majoritaire.
BulMajUni	Critère du bulletin majoritaire unisson.
Cond	Critère de Condorcet.
Coo.	Méthode de Coombs.
Cop.	Méthode de Copeland.
CMIgn	Critère de la coalition majoritaire ignorante.
CMInf	Critère de la coalition majoritaire informée.
CSD	Méthode de Condorcet avec somme des défaites.
FavMaj	Critère du favori majoritaire.
CVTI	Condorcification de VTI.
IANP	Indépendance des alternatives non pertinentes.
Kem.	Méthode de Kemeny.
KR	Méthode de Kim-Roush.
JM	Jugement majoritaire.
Max.	Maximin.
MC	Manipulation / manipulable par coalition.
MCI	Manipulation / manipulable par coalition ignorante.
MI	Manipulation individuelle / manipulable individuellement.
MT	Manipulation triviale / manipulable trivialement.
MU	Manipulation / manipulable-unisson.
Nan.	Méthode de Nanson.
PO	Méthode par paires ordonnées.
ssi	Si et seulement si.
t.q.	Tel(le) que.
Sch.	Méthode de Schulze.
SE	Scrutin exhaustif.

SVBE	Système de vote basé sur l'état.
SVAMP	Simulator of Various Voting Algorithms in Manipulating Populations.
U2T	Uninominal à deux tours.
U2TI	Uninominal à deux tours instantané.
Uni.	Uninominal (à un tour).
VA	Vote par assentiment.
VMF	(Modèle de) Von Mises–Fisher.
VN	Vote par notation.
VTI	Vote à tours instantanés.
VTID	Vote à tours instantanés à duels.
VTIM	Vote à tours instantanés basé sur la moyenne.

Bibliographie

Fuad Aleskerov, Daniel Karabekyan, Remzi Sanver et Vyacheslav Yakuba : Computing the degree of manipulability in the case of multiple choice. *Computational Social Choice (COMSOC-2008)*, page 27, 2008.

Fuad Aleskerov et Eldeniz Kurbanov : Degree of manipulability of social choice procedures. *In* Society for the Advancement of Economic Theory, éditeur : *Current trends in economics : theory and applications*, Studies in economic theory, pages 13–27. Springer, 1999.

Kenneth Arrow : A difficulty in the concept of social welfare. *The Journal of Political Economy*, 58(4):328–346, 1950.

Navin Aswal, Shurojit Chatterji et Arunava Sen : Dictatorial domains. *Economic Theory*, 22(1):45–62, 2003.

Michèle Audin : *Geometry*. Universitext (Berlin. Print). Springer Berlin Heidelberg, 2003.

Haris Aziz, Serge Gaspers, Nicholas Mattei, Nina Narodytska et Toby Walsh : Ties matter : Complexity of manipulation when tie-breaking with a random vote. *In AAAI*, 2013.

Michel Balinski et Rida Laraki : *Majority Judgment : Measuring, Ranking, and Electing*. MIT Press, 2010.

Salvador Barberá : Strategy-proofness and pivotal voters : A direct proof of the Gibbard-Satterthwaite theorem. *International Economic Review*, 24(2):413–417, 1983.

Salvador Barberà : An introduction to strategy-proof social choice functions. *Social Choice and Welfare*, 18:619–653, 2001.

Salvador Barberá et Bezalel Peleg : Strategy-proof voting schemes with continuous preferences. *Social Choice and Welfare*, 7:31–38, 1990.

John Bartholdi et James Orlin : Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8:341–354, 1991.

John Bartholdi, Craig Tovey et Michael Trick : The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989a.

John Bartholdi, Craig Tovey et Michael Trick : Voting schemes for which it can be difficult to tell who won the election. *Social Choice and welfare*, 6 (2):157–165, 1989b.

Eugenio Beltrami : Résolution du problème de reporter les points d'une surface sur un plan, de manière que les lignes géodésiques soient représentée par des lignes droites. *Annali di Matematica*, 1866.

Eugenio Beltrami : Essai d'interprétation de la géométrie noneuclidéenne. Trad. par J. Hoüel. *Ann. Sci. École Norm. Sup.*, 6:251–288, 1869.

Jean-Pierre Benoît : The Gibbard-Satterthwaite theorem : a simple proof. *Economics Letters*, 69(3):319–322, 2000.

Douglas Bernheim, Bezalel Peleg et Michael Whinston : Coalition-proof Nash equilibria I. Concepts. *Journal of Economic Theory*, 42(1):1–12, 1987.

Nadja Betzler, Jiong Guo et Rolf Niedermeier : Parameterized computational complexity of Dodgson and Young elections. *Information and Computation*, 208 (2):165–177, 2010.

Nadja Betzler, Rolf Niedermeier et Gerhard Woeginger : Unweighted coalitional manipulation under the Borda rule is NP-hard. *In Proceedings of the 22th International Joint Conference on Artificial Intelligence (IJCAI '11)*, 2011.

Duncan Black : *The theory of committees and elections.* University Press, 1958.

Jean-Marie Blin et Mark Satterthwaite : Strategy-proofness and single-peakedness. *Public Choice*, 26:51–58, 1976.

Kim Border et J.S. Jordan : Straightforward elections, unanimity and phantom voters. *The Review of Economic Studies*, 50(1):153–170, 1983.

Steven Brams : Voting procedures. *In* R.J. Aumann et S. Hart, éditeurs : *Handbook of Game Theory with Economic Applications*, volume 2 de *Handbook of Game Theory with Economic Applications*, chapitre 30, pages 1055–1089. Elsevier, 1994.

Steven Brams : Approval voting. *In* C.K. Rowley et F. Schneider, éditeurs : *The Encyclopedia of Public Choice*, pages 344–346. Springer US, 2003.

Steven Brams et Peter Fishburn : Approval voting. *American Political Science Review*, 72:831–847, 1978.

Felix Brandt et Markus Brill : Necessary and sufficient conditions for the strategyproofness of irresolute social choice functions. *In Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 136–142. ACM, 2011.

Simina Brânzei, Ioannis Caragiannis, Jamie Morgenstern et Ariel Procaccia : How bad is selfish voting? *In Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

Markus Brill et Vincent Conitzer : Strategic voting and strategic candidacy. *In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, USA*, 2015.

Markus Brill et Felix Fischer : The price of neutrality for the ranked pairs method. *COMSOC 2012*, page 95, 2012.

Colin Campbell et Gordon Tullock : A measure of the importance of cyclical majorities. *The Economic Journal*, 75(300):853–857, 1965.

Ioannis Caragiannis, Jason Covey, Michal Feldman, Christopher Homan, Christos Kaklamanis, Nikos Karanikolas, Ariel Procaccia et Jeffrey Rosenschein : On the approximability of Dodgson and Young elections. *In Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1058–1067. Society for Industrial and Applied Mathematics, 2009.

Ioannis Caragiannis, Christos Kaklamanis, Nikos Karanikolas et Ariel Procaccia : Socially desirable approximations for dodgson's voting rule. *ACM Transactions on Algorithms (TALG)*, 10(2):6, 2014.

John Chamberlin, Jerry Cohen et Clyde Coombs : Social choice observed : Five presidential elections of the american psychological association. *The Journal of Politics*, 46:479–502, 1984.

Graciela Chichilnisky : Von Neumann-Morgenstern utilities and cardinal preferences. *Mathematics of Operations Research*, 10(4):633–641, 1985.

Tom Coleman et Vanessa Teague : On the complexity of manipulating elections. *In* Joachim Gudmundsson et Barry Jay, éditeurs : *Thirteenth Computing : The Australasian Theory Symposium (CATS2007)*, volume 65 de *CRPIT*, pages 25–33. ACS, 2007.

Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet : *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie royale, 1785.

Vincent Conitzer, Jérôme Lang et Tuomas Sandholm : How many candidates are needed to make elections hard to manipulate ? *In Proceedings of the 9th conference on Theoretical aspects of rationality and knowledge*, pages 201–214. ACM, 2003.

Vincent Conitzer et Tuomas Sandholm : Complexity of manipulating elections with few candidates. *In Eighteenth national conference on Artificial intelligence*, pages 314–319. American Association for Artificial Intelligence, 2002.

Vincent Conitzer et Tuomas Sandholm : Universal voting protocol tweaks to make manipulation hard. *In Proceedings of the 18th international joint conference on Artificial intelligence*, pages 781–788. Morgan Kaufmann Publishers Inc., 2003.

Vincent Conitzer et Tuomas Sandholm : Nonexistence of voting rules that are usually hard to manipulate. *In Proceedings of the 21st national conference on Artificial intelligence - Volume 1*, pages 627–634. AAAI Press, 2006.

Vincent Conitzer, Tuomas Sandholm et Jérôme Lang : When are elections with few candidates hard to manipulate ? *J. ACM*, 54, juin 2007.

Jessica Davies, George Katsirelos, Nina Narodytska et Toby Walsh : Complexity of and algorithms for Borda manipulation. *AAAI*, 11:657–662, août 2011.

Jessica Davies, George Katsirelos, Nina Narodytska, Toby Walsh et Lirong Xia : Complexity of and algorithms for the manipulation of Borda, Nanson's and Baldwin's voting rules. *Artificial Intelligence*, 217:20–42, 2014.

Frank DeMeyer et Charles Plott : The probability of a cyclical majority. *Econometrica*, 38(2):345–354, 1970.

Amogh Dhamdhere et Constantine Dovrolis : The Internet is flat : modeling the transition from a transit hierarchy to a peering mesh. *In Proceedings of the 6th International Conference*, Co-NEXT '10, page 21. ACM, 2010.

Thomas Downs : Some relationships among the Von Mises distributions of different dimensions. *Biometrika*, 53(1/2):269–272, 1966.

John Duggan et Thomas Schwartz : Strategic manipulability without resoluteness or shared beliefs : Gibbard-Satterthwaite generalized. *Social Choice and Welfare*, 17:85–93, 2000.

François Durand, Benoît Kloeckner, Fabien Mathieu et Ludovic Noirie : Geometry on the utility sphere. *In Proceedings of the 4th International Conference on Algorithmic Decision Theory (ADT)*, 2015.

François Durand, Fabien Mathieu et Ludovic Noirie : Manipulability of voting systems. Groupe de travail Displexity, http://www.liafa.univ-paris-diderot.fr/~displexity/docpub/6mois/votes.pdf, 2012.

François Durand, Fabien Mathieu et Ludovic Noirie : On the manipulability of voting systems : application to multi-operator networks. *In Proceedings of the 9th International Conference on Network and Service Management (CNSM)*, pages 292–297. IEEE, 2013.

François Durand, Fabien Mathieu et Ludovic Noirie : Élection du best paper AlgoTel 2012 : étude de la manipulabilité. *In AlgoTel 2014 – 16èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications*, 2014a.

François Durand, Fabien Mathieu et Ludovic Noirie : Élection d'un chemin dans un réseau : étude de la manipulabilité. *In AlgoTel 2014 – 16èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications*, 2014b.

François Durand, Fabien Mathieu et Ludovic Noirie : Making a voting system depend only on orders of preference reduces its manipulability rate. https://hal.inria.fr/hal-01009136, 2014c.

François Durand, Fabien Mathieu et Ludovic Noirie : Making most voting systems meet the Condorcet criterion reduces their manipulability. https://hal.inria.fr/hal-01009134, 2014d.

François Durand, Fabien Mathieu et Ludovic Noirie : Reducing manipulability. Poster présenté au 5th International Workshop on Computational Social Choice (COMSOC), 2014e.

Bhaskar Dutta, Matthew Jackson et Michel Le Breton : Strategic candidacy and voting procedures. *Econometrica*, 69(4):1013–1037, 2001.

Bhaskar Dutta, Matthew Jackson et Michel Le Breton : Voting by successive elimination and strategic candidacy. *Journal of Economic Theory*, 103(1):190–218, 2002.

Lars Ehlers, Hans Peters et Ton Storcken : Threshold strategy-proofness : on manipulability in large voting problems. *Games and Economic Behavior*, 49 (1):103–116, 2004.

Edith Elkind et Helger Lipmaa : Hybrid voting protocols and hardness of manipulation. *In* Xiaotie Deng et Ding-Zhu Du, éditeurs : *Algorithms and Computation*, volume 3827 de *Lecture Notes in Computer Science*, pages 206–215. Springer Berlin / Heidelberg, 2005a.

Edith Elkind et Helger Lipmaa : Small coalitions cannot manipulate voting. *In* Andrew Patrick et Moti Yung, éditeurs : *Financial Cryptography and Data Security*, volume 3570 de *Lecture Notes in Computer Science*, pages 578–578. Springer Berlin / Heidelberg, 2005b.

Piotr Faliszewski et Ariel Procaccia : AI's war on manipulation : Are we winning ? *AI Magazine*, 31(4):53–64, 2010.

Pierre Favardin et Dominique Lepelley : Some further results on the manipulability of social choice rules. *Social Choice and Welfare*, 26:485–509, 2006.

Pierre Favardin, Dominique Lepelley et Jérôme Serais : Borda rule, Copeland method and strategic manipulation. *Review of Economic Design*, 7:213–228, 2002.

Allan Feldman : *Welfare economics and social choice theory.* Kluwer Nijhoff Publishing. Martinus Nijhoff Pub., 1980.

Peter Fishburn : *Utility Theory for Decision Making.* Wiley, New York, 1970.

Peter Fishburn : *Nonlinear preference and utility theory.* Johns Hopkins series in the mathematical sciences. Johns Hopkins University Press, 1988.

Peter Fishburn et Steven Brams : Efficacy, power and equity under approval voting. *Public Choice*, 37(3):425–434, 1981.

Peter Fishburn et William Gehrlein : Borda's rule, positional voting, and Condorcet's simple majority principle. *Public Choice*, 28(1):79–88, 1976.

Peter Fishburn, William Gehrlein et Eric Maskin : Condorcet proportions and Kelly's conjectures. *Discrete Applied Mathematics*, 1(4):229–252, 1979.

Mark Garman et Morton Kamien : The paradox of voting : Probability calculations. *Behavioral Science*, 13:306–316, 1968.

Serge Gaspers, Thomas Kalinowski, Nina Narodytska et Toby Walsh : Coalitional manipulation for Schulze's rule. *In Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 431–438. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

John Geanakoplos : Three brief proofs of Arrow's impossibility theorem. *Economic Theory*, 26(1):211–215, 2005.

William Gehrlein : The expected probability of Condorcet's paradox. *Economics Letters*, 7(1):33–37, 1981.

William Gehrlein : Approximating the probability that a Condorcet winner exists. 1999.

William Gehrlein : *Condorcet's Paradox.* Theory and Decision Library C. Springer, 2006.

William Gehrlein et Peter Fishburn : The probability of the paradox of voting : A computable solution. *Journal of Economic Theory*, 13(1):14–25, 1976.

Allan Gibbard : Manipulation of voting schemes : A general result. *Econometrica*, 41(4):587–601, 1973.

Allan Gibbard : Manipulation of schemes that mix voting with chance. *Econometrica*, 45(3):665–681, 1977.

Allan Gibbard : Straightforwardness of game forms with lotteries as outcomes. *Econometrica*, 46(3):595–614, 1978.

Allan Gibbard : Social choice and the Arrow conditions. *Economics and Philosophy*, 30(03):269–284, 2014.

Peter Gärdenfors : Manipulation of social choice functions. *Journal of Economic Theory*, 13:217–228, 1976.

Peter Gärdenfors : A concise proof of theorem on manipulation of social choice functions. *Public Choice*, 32:137–142, 1977.

James Green-Armytage : Four Condorcet-Hare hybrid methods for single-winner elections. *Voting matters*, 29:1–14, 2011.

James Green-Armytage : Strategic voting and nomination. *Social Choice and Welfare*, 42(1):111–138, 2014.

James Green-Armytage, Nicolaus Tideman et Rafael Cosman : Statistical evaluation of voting rules. 2014.

Bernard Grofman et Scott Feld : If you like the alternative vote (a.k.a. the instant runoff), then you ought to know about the Coombs rule. *Electoral Studies*, 23(4):641–659, 2004.

François Guénard et Gilbert Lelièvre : *Compléments d'analyse*. Numéro 1 *in* Compléments d'analyse. E.N.S., 1985.

Peter Hammond : Interpersonal comparisons of utility : Why and how they are and should be made. *In Interpersonal Comparisons of Well-Being*, pages 200–254. University Press, 1991.

Lane Hemaspaandra, Rahman Lavaee et Curtis Menton : Schulze and ranked-pairs voting are fixed-parameter tractable to bribe, manipulate, and control. *In Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 1345–1346. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

Andrew Jennings : *Monotonicity and Manipulability of Ordinal and Cardinal Social Choice Functions*. BiblioBazaar, 2011.

Bradford Jones, Benjamin Radcliff, Charles Taber et Richard Timpone : Condorcet winners and the paradox of voting : Probability calculations for weak preference orders. *The American Political Science Review*, 89(1):137–144, mars 1995.

Nathan Keller : A tight quantitative version of Arrow's impossibility theorem. *Journal of the European Mathematical Society*, 14(5):1331–1355, 2012.

Jerry Kelly : Voting anomalies, the number of voters, and the number of alternatives. *Econometrica*, pages 239–251, 1974.

Jerry KELLY : Almost all social choice rules are highly manipulable, but a few aren't. *Social Choice and Welfare*, 10:161–175, 1993.

John KEMENY : Mathematics without numbers. *Daedalus*, 88:575–591, 1959.

K.H. KIM et F.W. ROUSH : Statistical manipulability of social choice functions. *Group Decision and Negotiation*, 5:263–282, 1996.

Benoît KLOECKNER : *Un bref aperçu de la géométrie projective*. Calvage & Mounet, 2012.

Kathrin KONCZAK et Jérôme LANG : Voting procedures with incomplete preferences. *In Proc. IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling*, 2005.

David KREPS : *A Course in Microeconomic Theory*. Princeton University Press, 1990.

Craig LABOVITZ, Scott IEKEL-JOHNSON, Danny MCPHERSON, Jon OBERHEIDE et Farnam JAHANIAN : Internet inter-domain traffic. *In Proceedings of the ACM SIGCOMM 2010 conference*, pages 75–86, 2010.

Jérome LANG, Nicolas MAUDET et Maria POLUKAROV : New results on equilibria in strategic candidacy. *In Algorithmic Game Theory*, pages 13–25. Springer, 2013.

Jean-Francois LASLIER : Spatial approval voting. *Political Analysis*, 14:160–185(26), 2006.

Jean-François LASLIER : The leader rule : A model of strategic approval voting in a large electorate. *Journal of Theoretical Politics*, 21(1):113–136, 2009.

Jean-François LASLIER et Karine Van der STRAETEN : A live experiment on approval voting. *Experimental Economics*, 11(1):97–105, 2008.

Jean-François LASLIER : *Le vote et la règle majoritaire*. CNRS science politique. CNRS Editions, 2004.

Stephane LAVEAU et Olivier FAUGERAS : Oriented projective geometry for computer vision. *In ECCV96*, pages 147–156. Springer-Verlag, 1996.

Dominique LEPELLEY, Ahmed LOUICHI et Hatem SMAOUI : On Ehrhart polynomials and probability calculations in voting theory. *Social Choice and Welfare*, 30:363–383, 2008.

Dominique LEPELLEY et Boniface MBIH : The proportion of coalitionally unstable situations under the plurality rule. *Economics Letters*, 24(4):311–315, 1987.

Dominique LEPELLEY et Boniface MBIH : The vulnerability of four social choice functions to coalitional manipulation of preferences. *Social Choice and Welfare*, 11:253–265, 1994.

Dominique LEPELLEY et Vincent MERLIN : Choix social positionnel et principe majoritaire. *Annales d'Economie et de Statistique*, pages 29–48, 1998.

Dominique LEPELLEY et Vincent MERLIN : Scoring run-off paradoxes for variable electorates. *Economic Theory*, 17(1):53–80, 2001.

Dominique Lepelley et Fabrice Valognes : On the Kim and Roush voting procedure. *Group Decision and Negotiation*, 8:109–123, 1999.

Dominique Lepelley et Fabrice Valognes : Voting rules, manipulability and social homogeneity. *Public Choice*, 116:165–184, 2003.

Ramon Llull : *De arte electionis*. 1299.

Ramon Llull : *Blanquerna*. c. 1285.

Hans Maassen et Thom Bezembinder : Generating random weak orders and the probability of a Condorcet winner. *Social Choice and Welfare*, 19(3):517–532, 2002.

Colin Mallows : Non-null ranking models. *Biometrika*, pages 114–130, 1957.

Andrew Mao, Ariel Procaccia et Yiling Chen : Better human computation through principled voting. *In Proceedings of of the 27th Conference on Artificial Intelligence (AAAI'13)*, 2013.

K. V. Mardia : Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(3):349–393, 1975.

Andreu Mas-Colell, Michael Whinston et Jerry Green : *Microeconomic Theory*. Oxford University Press, 1995.

Nicholas Mattei et Toby Walsh : Preflib : A library of preference data. *In Proceedings of Third International Conference on Algorithmic Decision Theory (ADT 2013)*, Lecture Notes in Artificial Intelligence. Springer, 2013.

Kenneth May : A set of independent, necessary and sufficient conditions for simple majority decision. *Econometrica*, 20(4):680–684, 1952.

Robert May : Some mathematical remarks on the paradox of voting. *Behavioral Science*, 16:143–151, 1971.

Iain McLean : The Borda and Condorcet principles : three medieval applications. *Social Choice and Welfare*, 7(2):99–108, 1990.

Vincent Merlin : The axiomatic characterizations of majority voting and scoring rules. *Mathématiques et sciences humaines*, (163), 2003.

Vincent Merlin, Maria Tataru et Fabrice Valognes : On the probability that all decision rules select the same winner. *Journal of Mathematical Economics*, 33(2):183–207, 2000.

Vincent Merlin, Monica Tataru et Fabrice Valognes : On the likelihood of Condorcet's profiles. *Social Choice and Welfare*, 19:193–206, 2002.

Elchanan Mossel : A quantitative Arrow theorem. *Probability Theory and Related Fields*, 154(1-2):49–88, 2012.

Hervé Moulin : *La Stratégie du vote*. Cahiers du CEREMADE. Université Paris IX-Dauphine, Centre de recherche de mathématiques de la décision, 1978.

Hervé Moulin : On strategy-proofness and single peakedness. *Public Choice*, 35 (4):437–455, 1980.

Hervé Moulin : Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53–64, juin 1988.

Eitan Muller et Mark Satterthwaite : The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, 14(2):412–418, 1977.

Roger Myerson : Theoretical comparisons of electoral systems. *European Economic Review*, 43(4):671–697, 1999.

Kenjiro Nakamura : The vetoers in a simple game with ordinal preferences. *International Journal of Game Theory*, 8(1):55–61, 1979.

Nina Narodytska, Toby Walsh et Lirong Xia : Manipulation of Nanson's and Baldwin's rules. *In Workshop on Social Choice and Artificial Intelligence*, page 64, 2011.

Richard Niemi et Herbert Weisberg : A mathematical solution for the probability of the paradox of voting. *Behavioral Science*, 13(4):317–323, 1968.

Noam Nisam, Tim Roughgarden, Éva Tardos et Vijay Vazirani : *Algorithmic Game Theory*. Cambridge University Press, 2007.

Shmuel Nitzan : The vulnerability of point-voting schemes to preference variation and strategic manipulation. *Public Choice*, 47:349–370, 1985.

Svetlana Obraztsova et Edith Elkind : On the complexity of voting manipulation under randomized tie-breaking. *COMSOC 2012*, page 347, 2012.

Svetlana Obraztsova, Edith Elkind et Noam Hazon : Ties matter : Complexity of voting manipulation revisited. *In The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 71–78. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

David Parkes et Lirong Xia : A complexity-of-strategic-behavior comparison between Schulze's rule and ranked pairs. *In Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI'12)*. American Association for Artificial Intelligence, 2012.

Bezalel Peleg : *Game theoretic analysis of voting in committees*. Numéro 7 *in* Econometric Society monographs in pure theory. Cambridge Univ. Press, 1984.

Elizabeth Penn, John Patty et Sean Gailmard : Manipulation and single-peaknedness : A general result. *American Journal of Political Science*, 55:436–449, 2011.

Joaquín Pérez : The strong no show paradoxes are a common flaw in Condorcet voting correspondences. *Social Choice and Welfare*, 18(3):601–616, 2001.

John Pomeranz et Roman Weil Jr : The cyclical majority problem. *Commun. ACM*, 13:251–254, avril 1970.

Geoffrey Pritchard et Arkadii Slinko : On the average minimum size of a manipulating coalition. *Social Choice and Welfare*, 27:263–277, 2006.

Geoffrey Pritchard et Mark Wilson : Exact results on manipulability of positional voting rules. *Social Choice and Welfare*, 29:487–513, 2007.

Ariel Procaccia et Jeffrey Rosenschein : Junta distributions and the average-case complexity of manipulating elections. *In Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, AAMAS '06, pages 497–504. ACM, 2006.

Ariel Procaccia et Jeffrey Rosenschein : Average-case tractability of manipulation in voting via the fraction of manipulators. *In Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, AAMAS '07, page 105. ACM, 2007.

Michel Regenwetter, Bernard Grofman, Anthony Marley et Ilia Tsetlin : Behavioral social choice. *Cambridge University Press*, 13:58–68, 2006.

Philip Reny : Arrow's theorem and the Gibbard-Satterthwaite theorem : a unified approach. *Economics Letters*, 70:99–105, janvier 2001.

Reyhaneh Reyhani : *Strategic manipulation in voting systems.* Thèse de doctorat, 2013.

Reyhaneh Reyhani, Geoffrey Pritchard et Mark Wilson : A new measure of the difficulty of manipulation of voting rules, 2009.

Harold Ruben : On the moments of order statistics in samples from normal populations. *Biometrika*, 41(1/2):200–227, juin 1954.

Donald Saari : Susceptibility to manipulation. *Public Choice*, 64(1):21–41, 1990.

Donald Saari : *Geometry of voting*, volume 3. Springer Science & Business Media, 2012.

Donald Saari et Vincent Merlin : A geometric examination of Kemeny's rule. *Social Choice and Welfare*, 17(3):403–438, 2000.

Mark Satterthwaite : Strategy-proofness and Arrow's conditions : Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

Markus Schulze : A new monotonic, clone-independent, reversal symmetric, and Condorcet-consistent single-winner election method. *Social Choice and Welfare*, 36:267–303, 2011.

Arunava Sen : Another direct proof of the Gibbard-Satterthwaite theorem. *Economics Letters*, 70(3):381–385, 2001.

Murat Sertel et Remzi Sanver : Strong equilibrium outcomes of voting games are the generalized Condorcet winners. *Social Choice and Welfare*, 22:331–347, 2004.

Arkadii Slinko : How large should a coalition be to manipulate an election? *Mathematical Social Sciences*, 47(3):289–293, 2004.

Arkadii Slinko et Shaun White : Nondictatorial social choice rules are safely manipulable. *In COMSOC'08*, pages 403–413, 2008.

David Smith : Manipulability measures of common social choice functions. *Social Choice and Welfare*, 16:639–661, 1999.

John Smith : Aggregation of preferences with variable electorate. *Econometrica : Journal of the Econometric Society*, pages 1027–1041, 1973.

Michael Spivak : *A comprehensive introduction to differential geometry. Vol. III.* Publish or Perish Inc., second édition, 1979a.

Michael Spivak : *A comprehensive introduction to differential geometry. Vol. IV.* Publish or Perish Inc., second édition, 1979b.

Jorge Stolfi : Oriented projective geometry. *In Proceedings of the third annual symposium on Computational geometry*, SCG '87, pages 76–85. ACM, 1987.

Alan Taylor : *Social choice and the mathematics of manipulation.* Outlooks Series. Cambridge University Press, 2005.

Nicolaus Tideman : Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4:185–206, 1987.

Nicolaus Tideman : *Collective Decisions And Voting : The Potential for Public Choice.* Ashgate, 2006.

Ilia Tsetlin, Michel Regenwetter et Bernard Grofman : The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare*, 21:387–398, 2003.

Gordon Tullock et Colin Campbell : Computer simulation of a small voting system. *The Economic Journal*, 80(317):97–104, 1970.

Gary Ulrich : Computer generation of distributions on the m-sphere. *Applied Statistics*, pages 158–163, 1984.

Karine Van der Straeten, Jean-François Laslier, Nicolas Sauger et André Blais : Strategic, sincere, and heuristic voting under four election rules : an experimental study. *Social Choice and Welfare*, 35(3):435–472, 2010.

John Von Neumann et Oskar Morgenstern : *Theory of games and economic behavior.* Princeton University Press, 1944.

John Von Neumann, Oskar Morgenstern, Harold Kuhn et Ariel Rubinstein : *Theory of Games and Economic Behavior (Commemorative Edition).* Princeton Classic Editions. Princeton University Press, 2007.

Toby Walsh : An empirical study of the manipulability of single transferable voting. *In ECAI*, volume 10, pages 257–262, 2010a.

Toby Walsh : Manipulability of single transferable vote. *In* F. Brandt, V. Conitzer, L. Hemaspaandra, J.-F. Laslier et W. Zwicker, éditeurs : *Computational Foundations of Social Choice*, numéro 10101 *in* Dagstuhl Seminar Proceedings. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2010b.

Tiance Wang, Paul Cuff et Sanjeev Kulkarni : Condorcet methods are less susceptible to strategic voting. 2014.

G.S. Watson et E.J. Williams : On the construction of significance tests on the circle and the sphere. *Biometrika*, 43(3/4):344–352, 1956.

Tjark Weber : Alternatives vs. outcomes : A note on the Gibbard-Satterthwaite theorem. Rapport technique, 2009.

Andrew Wood : Simulation of the Von Mises Fisher distribution. *Communications in Statistics-Simulation and Computation*, 23(1):157–164, 1994.

Eduardo Xavier : A note on a maximum k-subset intersection problem. *Information Processing Letters*, 112(12):471–472, 2012.

Lirong Xia et Vincent Conitzer : Generalized scoring rules and the frequency of coalitional manipulability. *In Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 109–118. ACM, 2008.

Lirong Xia, Michael Zuckerman, Ariel Procaccia, Vincent Conitzer et Jeffrey Rosenschein : Complexity of unweighted coalitional manipulation under some common voting rules. *In International Joint Conference on Artificial Intelligence*, pages 348–353, 2009.

Peyton Young et Arthur Levenglick : A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, 35(2), 1978.

Michael Zuckerman, Omer Lev et Jeffrey Rosenschein : An algorithm for the coalitional manipulation problem under Maximin. *In The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '11, pages 845–852, 2011.

Michael Zuckerman, Ariel Procaccia et Jeffrey Rosenschein : Algorithms for the coalitional manipulation problem. *Artificial Intelligence*, 173(2):392–412, 2009.