Optimal search strategies and intermittent random walk : from restriction enzymes to the albatross flight

Claude Loverdo

▶ **To cite this version:**

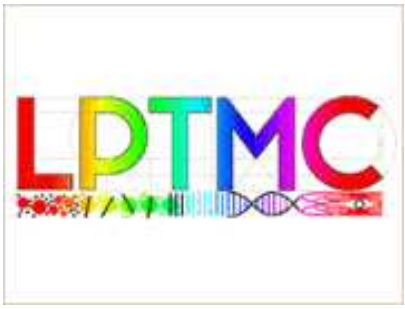
Claude Loverdo. Optimal search strategies and intermittent random walk : from restriction enzymes to the albatross flight. Biological Physics [physics.bio-ph]. Université Pierre et Marie Curie - Paris VI, 2009. English. NNT : . tel-00445482

HAL Id: tel-00445482
https://theses.hal.science/tel-00445482

Submitted on 8 Jan 2010

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 : PHYSIQUE

(École Doctorale de Physique de la Région Parisienne – ED 107)

(Laboratoire de physique théorique de la matière condensée – UMR 7600)

Présentée par

Claude Loverdo

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Stratégies de recherche optimales et marches aléatoires intermittentes

De l'enzyme de restriction au vol de l'albatros

Soutenance prévue le jeudi 10 décembre 2009 devant un jury composé de :

Olivier Bénichou (directeur de thèse)
Didier Chatenay
Pierre Desbiolles (invité)
Jean-François Joanny
Pierre Levitz (rapporteur)
Ralf Metzler (rapporteur)
Emmanuel Trizac

Contents

1 Résumé en français / Summary in French **5**
1.1 Introduction . . . 5
1.2 Stratégies de recherche intermittentes à l'échelle macroscopique . . . 7
1.3 Stratégies de recherche intermittentes à l'échelle microscopique . . . 14
1.4 Les stratégies intermittentes : des stratégies robustes . . . 24
1.5 Extensions et perspectives . . . 30
1.6 Conclusion . . . 36

2 Introduction **37**
2.1 Foreword and outline . . . 37
2.2 General framework . . . 40

3 Intermittent search strategies at the macroscopic scale **43**
3.1 Lévy strategies . . . 44
3.2 A simple model based on intermittence . . . 45
3.3 Influence of the target distribution on the search time . . . 47
3.4 Minimal model of intermittent search in two dimensions . . . 52
3.5 Are Lévy strategies really so advantageous? . . . 57
3.6 Conclusion on animal foraging . . . 59

4 Intermittent search strategies at the microscopic scale **61**
4.1 Protein/DNA interaction . . . 61
4.2 Active transport of vesicles in cells . . . 92

5 Intermittent search : a robust strategy **101**
5.1 Introduction . . . 101
5.2 Model, notations and simulations methods . . . 102
5.3 One dimension . . . 105
5.4 Two dimensions . . . 119
5.5 Three dimensions . . . 129
5.6 Discussion and conclusion . . . 140

6 Extensions and perspectives **143**
6.1 Beyond the first moment . . . 144
6.2 Taking into account partial correlations in ballistic phases . . . 150
6.3 Other distributions of phases duration . . . 155
6.4 Other models of intermittent search . . . 159
6.5 Designing efficient searches . . . 163
6.6 Comparison with experimental data . . . 165

7 Conclusion **171**

8 Appendices **173**
8.1 Complements to hops and jumps . . . 173
8.2 Complements to the generic intermittent search model . . . 182
8.3 Complements to the extensions and perspectives . . . 193

8.4 An example of the importance of transport in biology 200

References **213**

Publications linked to this thesis **223**

Remerciements **225**

Résumé/Abstract **230**

1 Résumé en français / Summary in French

Sommaire

1.1 Introduction 5
1.2 Stratégies de recherche intermittentes à l'échelle macroscopique 7
1.2.1 Introduction 7
1.2.2 Modèle 8
1.2.3 Influence de la distribution de cibles sur le temps de recherche 9
1.2.4 Modèle de recherche intermittente à 2 dimensions : une alternative aux stratégies de Lévy 12
1.2.5 Conclusion : les animaux ont-ils vraiment intérêt à suivre les stratégies de Lévy ? 14
1.3 Stratégies de recherche intermittentes à l'échelle microscopique 14
1.3.1 Recherche par une protéine d'une séquence spécifique sur l'ADN 14
1.3.2 Transport actif de vésicules en milieu cellulaire 22
1.4 Les stratégies intermittentes : des stratégies robustes . . 24
1.4.1 Motivations 24
1.4.2 Ingrédients du modèle générique 24
1.4.3 Méthodes 26
1.4.4 Résultats principaux 28
1.5 Extensions et perspectives 30
1.5.1 Distribution complète du temps de recherche 30
1.5.2 Mémoire temporelle 31
1.5.3 Mémoire spatiale 32
1.5.4 Comparaison avec des expériences 33
1.6 Conclusion 36

1.1 Introduction

Quelle stratégie adopter pour trouver un objet caché ? Si vous avez déjà perdu vos clés, vous vous êtes posé cette question.

Au-delà de cet exemple de la vie de tous les jours, les situations dans lesquelles un chercheur – que ce chercheur soit une personne, un animal ou n'importe quel organisme ou particule – doit trouver au plus vite une cible se rencontrent dans de nombreux contextes, de la recherche de naufragés en mer aux réactions chimiques dont le premier pas est la rencontre des réactifs.

Dans cette thèse, nous étudions quantitativement l'efficacité d'une nouvelle classe de stratégies de recherche, dénommées "recherches intermittentes", et introduites au LPTMC en 2004 juste avant mon arrivée en thèse. L'idée centrale de cette classe de stratégies peut facilement se comprendre en reprenant la situation de la vie quotidienne. Imaginez que vous avez perdu une clé sur une plage. Quand vous cherchez

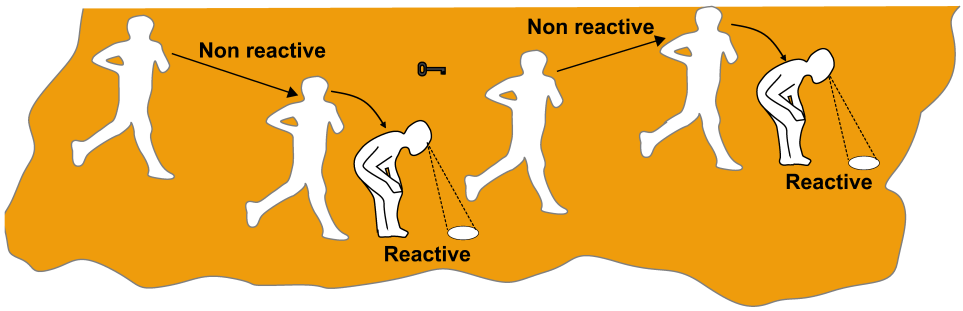


FIG. 1: Recherche intermittente : le chercheur alterne des phases d'inspection minutieuse pour détecter la clé, et des phases de déplacement rapide pendant lesquelles il va trop vite pour pouvoir détecter la cible.

sans aucun indice la clé perdue dans le sable (voir Fig.1), vous pouvez vous accroupir, et inspecter soigneusement le sable. De cette manière, si la clé est près de vous, vous la retrouverez. Mais dans cette position, il est difficile de se déplacer. Vous pouvez donc décider de vous lever, et de courir vers un autre point de la plage. Pendant la phase de course, vous allez trop vite pour voir ou sentir la clé. Donc d'une part, vous utilisez du temps qui aurait pu être passé dans la phase de détection. Mais d'autre part, vous déplacer vite peut vous permettre de mieux explorer la plage. Le bilan net d'une telle stratégie intermittente n'est donc pas clair.

Deux types de questions peuvent alors être posées :
- Ces stratégies de recherche sont-elles finalement efficaces, et si oui, comment répartir au mieux le temps de recherche entre les phases de recherche minutieuse et les phases de déplacement ?
- Ces stratégies de recherche sont-elles pertinentes vis-à-vis de la description de situations réelles ?

Ces stratégies sont en fait observées à des échelles variées. Les premiers modèles de recherche intermittente introduits par l'équipe qui m'a accueillie au LPTMC concernent d'une part la recherche de nourriture par certains animaux [Bénichou et al., 2005a] (échelle macroscopique) et d'autre part la recherche par une protéine de sa séquence cible sur l'ADN [Coppey et al., 2004] (échelle microscopique). Ces deux approches mettent en jeu des déplacements à une dimension dans la phase de recherche, ce qui correspond aux situations modélisées.

Au cours de cette thèse, j'ai d'une part étendu ces modèles au cas plus général de déplacements à 2 ou 3 dimensions, pertinents dans un grand nombre de situations. D'autre part, j'ai démontré la robustesse de ces stratégies de recherche, suggérant ainsi qu'elles peuvent constituer un mode de recherche efficace générique. Les parties 1.2 et 1.3 présentent des travaux de modélisation en lien avec des études expérimentales, alors que la partie 1.4, plus technique, présente des résultats sur les temps de premier passage de marches aléatoires intermittentes.

Plus précisément, dans le cadre de la modélisation de trajectoires animales, j'ai développé un modèle de stratégies de recherche intermittentes bidimensionnel qui apparaît comme une alternative à un modèle très célèbre, dit des "stratégies de Lévy",

très souvent utilisé jusqu'alors dans ce contexte. Ces situations macroscopiques font l'objet de la première partie de mon manuscrit.

Je me suis également intéressée à plusieurs situations à l'échelle microscopique (deuxième partie du manuscrit). La première concerne le problème de la recherche par une enzyme d'une séquence spécifique sur l'ADN. Cette recherche, qui met en jeu des phases de diffusion 1D le long de la molécule d'ADN et des excursions 3D en volume, peut en fait être considérée comme intermittente au sens général défini précédemment. J'ai été amenée au cours de cette thèse à étendre les premiers modèles de recherche intermittente en calculant la distribution de la taille des sauts le long de l'ADN à la suite d'une excursion 3D. J'ai montré comment ce calcul théorique a pu être adapté aux spécificités d'une expérience à l'échelle de la molécule unique, développée dans l'équipe de Pierre Desbiolles (LKB-ENS), pour rendre compte des trajectoires observées. La deuxième situation concerne le transport de vésicules en milieu cellulaire. Le transport en milieu cellulaire est en effet souvent une combinaison de phases de diffusion "passive" et de phases de transport balistique "actif" le long des filaments du cytosquelette assuré par des moteurs moléculaires. J'ai proposé un modèle analytique de réactions limitées par le transport dans de tels milieux actifs et montré quantitativement comment le transport actif peut accélérer les réactions en milieu cellulaire.

Les stratégies de recherche intermittentes étant observées à des échelles variées, j'ai alors émis l'hypothèse qu'elles pourraient constituer un mécanisme de recherche générique. La troisième partie de cette thèse, plus technique, est ainsi consacrée à l'étude de la robustesse des marches aléatoires intermittentes, dans le cadre d'un modèle théorique assez général. La dernière partie de ce manuscrit rassemble des extensions et des perspectives relatives aux modèles étudiés précédemment, dont certaines sont en cours de réalisation.

1.2 Stratégies de recherche intermittentes à l'échelle macroscopique

1.2.1 Introduction

La recherche de nourriture par des animaux est un exemple intéressant de recherche à l'échelle macroscopique, qui a fait l'objet de nombreuses études en écologie comportementale, mais aussi plus récemment en physique. Un très grand nombre de trajectoires animales ont été décrites à l'aide du modèle dit des "marches de Lévy". Dans ce modèle proposé par Viswanathan et al. [1999], le chercheur peut trouver sa cible tout le long de sa trajectoire, si elle est à une distance inférieure au rayon de détection a. Ce chercheur avance à vitesse constante sur une distance l distribuée selon une loi de Lévy $p(l) \propto l^{-\mu}$, puis se réoriente aléatoirement. Les auteurs s'intéressent à deux types de cibles très différentes. Dans le cas de cibles "revisitables" (*i.e.* régénérées au même endroit après avoir été découvertes), ce type de stratégies est optimal pour $\mu \simeq 2$. Pour le cas très souvent plus pertinent de cibles qui ne peuvent pas être revisitées, c'est-à-dire qui sont détruites quand elles sont trouvées, l'optimum théorique est obtenu pour $\mu \to 1$. Cet optimum correspond en fait à la stratégie très intuitive de la ligne droite, sans changement de direction, qui est la meilleure façon d'explorer l'espace.

Comme indiqué précédemment, ce modèle des marches de Lévy suppose que le chercheur est capable de se déplacer et de trouver les cibles en même temps. Une question naturelle est celle de savoir comment décrire les trajectoires animales quand ces deux activités sont incompatibles. Il est en effet couramment admis que, très souvent, la vitesse diminue les capacités de détection [Kramer and McLaughlin, 2001]. Cette question est d'autant plus importante que de nombreuses observations, concernant des lézards, des poissons ou encore des oiseaux [O'Brien et al., 1990, Kramer and McLaughlin, 2001] font effectivement état de l'existence de deux phases de déplacement : une phase de déplacement rapide et une phase plus lente d'inspection de l'environnement. Un premier modèle unidimensionnel de ces comportements de recherche animaux a été proposé en 2005 [Bénichou et al., 2005a,c] et a montré l'efficacité de ces stratégies de recherche intermittentes.

Nous rappelons d'abord les résultats principaux de ce modèle, avant de présenter deux extensions réalisées au cours de cette thèse : (i) l'influence de la répartition des cibles sur l'efficacité de la recherche ; (ii) le développement d'un modèle "minimal" bidimensionnel de recherche intermittente, qui montre l'efficacité et même l'optimalité de telles stratégies. Ce dernier modèle apparaît ainsi comme une alternative au modèle des marches de Lévy.

1.2.2 Modèle

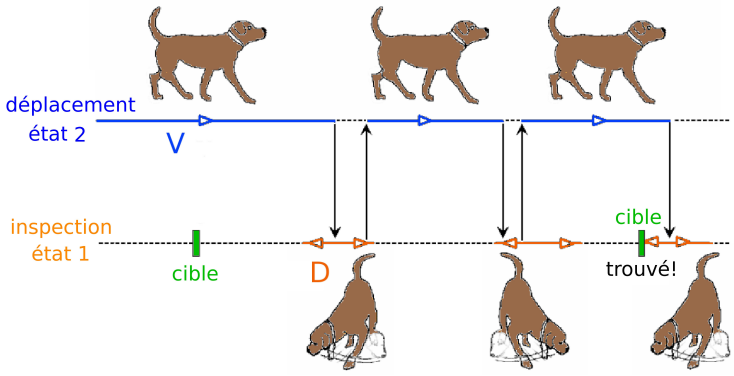


FIG. 2: Modèle simple d'intermittence pour les animaux à la recherche de nourriture.

La caractéristique centrale de ce modèle est l'alternance de deux phases (voir Fig.2) :

- Une phase d'inspection, notée 1, pendant laquelle les organes sensoriels de l'animal inspectent son environnement immédiat. Cette phase est modélisée par une diffusion de coefficient D. La cible est trouvée dès que le chercheur l'atteint. Comme le processus de détection requiert un temps minimum, la durée moyenne de cette phase est supposée bornée inférieurement par τ_1^{min}.
- Une phase de déplacement, notée 2, pendant laquelle le chercheur avance à la vitesse V toujours dans la même direction, et pendant laquelle la cible ne peut pas être détectée. Cette hypothèse est justifiée par le fait que, dans le cas de certains animaux, des corrélations significatives sont observées entre

les directions des phases balistiques successives [O'Brien et al., 1990]. Elle correspond donc au cas limite de corrélation parfaite, qui rend le problème équivalent à un problème à une dimension, avec toujours la même direction du déplacement balistique.

La probabilité de passer d'un état à l'autre est de plus supposée constante, ce qui conduit à des distributions exponentielles des durées des phases d'inspection et de déplacement, en accord avec de nombreuses observations expérimentales (par exemple : Fujiwara et al. [2002], Pierce-Shimonura et al. [1999], Hill et al. [2000], Li et al. [2008]). Par la suite, la durée moyenne d'une phase i est notée τ_i. Les proies sont supposées immobiles, espacées d'une distance L.

Sous ces hypothèses, le temps moyen exact de détection de la cible est calculé analytiquement (voir partie 3.2). Un point important est que ce temps moyen est proportionnel à L, alors qu'avec la phase d'inspection seule, il est proportionnel à L^2 : cela montre que pour une distance entre cibles suffisamment grande, l'intermittence diminue le temps moyen de détection de manière significative.

Le minimum du temps moyen de détection est obtenu pour τ_1 aussi petit que possible ($\tau_1 = \tau_1^{min}$), et à τ_1 fixé, pour une durée moyenne optimale de la phase de déplacement $\tau_2 = \tau_2^{opt}$. Dans le régime où $\tau_1 \gg \tau$ où $\tau = D/V^2$ est une échelle de temps qui dépend seulement des caractéristiques du chercheur, le temps optimal à passer dans la phase balistique est $\tau_2^{opt} = \left(\frac{3\tau\tau_1^2}{4}\right)^{1/3}$. Dans le régime opposé $\tau_1 \ll \tau$, $\tau_2^{opt} = \left(\frac{\tau^2\tau_1^3}{8}\right)^{1/5}$. Cette stratégie optimale rend raisonnablement bien compte des observations de O'Brien et al. [1990], Kramer and McLaughlin [2001] portant sur les animaux au comportement de recherche intermittent.

1.2.3 Influence de la distribution de cibles sur le temps de recherche

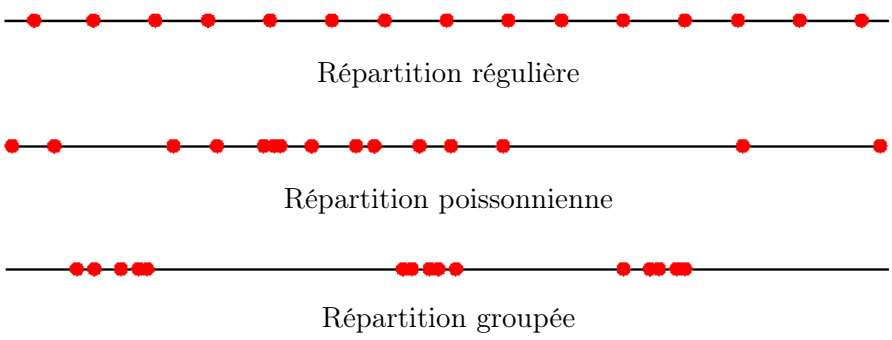


FIG. 3: Exemples de répartitions de cibles.

Les répartitions de cibles réelles sont principalement décrites comme régulièrement espacées, groupées ou aléatoires [Bell, 1991] (voir Fig.3). Dans le modèle présenté précédemment, les cibles sont régulièrement espacées. C'est une répartition représentative du cas où les cibles tendent à s'éloigner au maximum les unes des autres. Il s'agit aussi d'une approximation de champ moyen d'autres types de répartitions. Le deuxième type de répartition est celui où les cibles sont groupées.

Dans ce cas, quand une cible est trouvée, il est probable qu'il y en ait d'autres dans le voisinage. La recherche se décompose donc en deux phases, l'une étant la recherche d'un groupe de cibles, et l'autre l'exploitation du groupe de cibles. La première étape est dans ce cas équivalente à la recherche d'une cible, en prenant pour L la distance entre groupes, et est donc couverte par le modèle précédent. Nous considérons ici le troisième cas représentatif, celui d'une répartition poissonnienne de cibles. Les résultats correspondant ont été publiés dans [Moreau et al., 2007b, 2009].

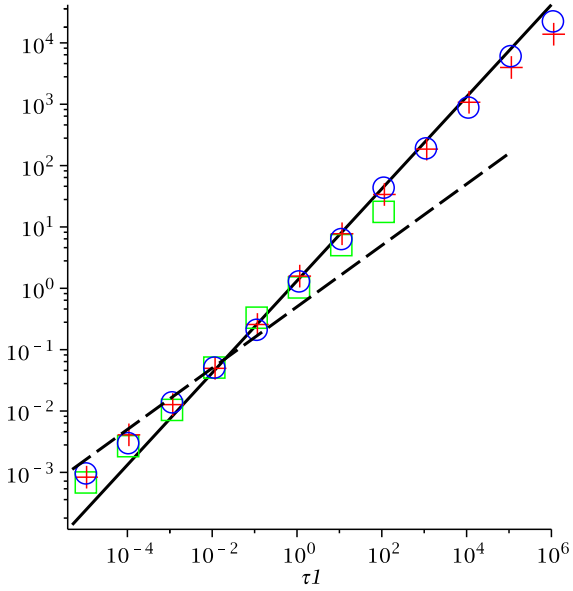


FIG. 4: $ln(\tau_2^{opt})$ en fonction de $ln(\tau_1)$. Prédiction analytique pour $\tau_1 \ll \tau$ (4) (pointillés noirs). Prédiction analytique pour $\tau_1 \gg \tau$ (1) (ligne noire). Simulations (points), pour $L = 10$ (□), $L = 10^3$ (+), $L = 10^5$ (○). $D = 1$, $V = 1$.

Nous reprenons le même modèle que précédemment, mais avec une répartition poissonnienne de cibles, L désignant désormais la distance moyenne entre cibles. L'expression approchée du temps moyen de détection de la cible (donnée dans la partie 3.3) permet de conclure que comme précédemment, l'intermittence est favorable pour L assez grand. La stratégie optimale consiste encore à prendre $\tau_1 = \tau_1^{min}$, mais la relation entre τ_2^{opt} et τ_1 peut prendre deux formes :

– Dans le régime $\tau_1 \gg \tau$:

$$\frac{\tau_2^{opt}}{\tau} \sim \sqrt{\frac{7}{4}} \left(\frac{\tau_1}{\tau}\right)^{3/4}. \tag{1}$$

À l'optimum, le temps moyen de détection de la cible vaut :

$$\langle t \rangle^{opt} \sim \frac{L}{2V} \sqrt{\frac{\tau_1}{\tau}}. \tag{2}$$

Le gain G de l'intermittence par rapport à la phase d'inspection seule (le temps moyen de détection de la cible sans intermittence divisé par le temps moyen de détection de la cible avec l'intermittence, en prenant τ_1 et τ_2 optimaux) est

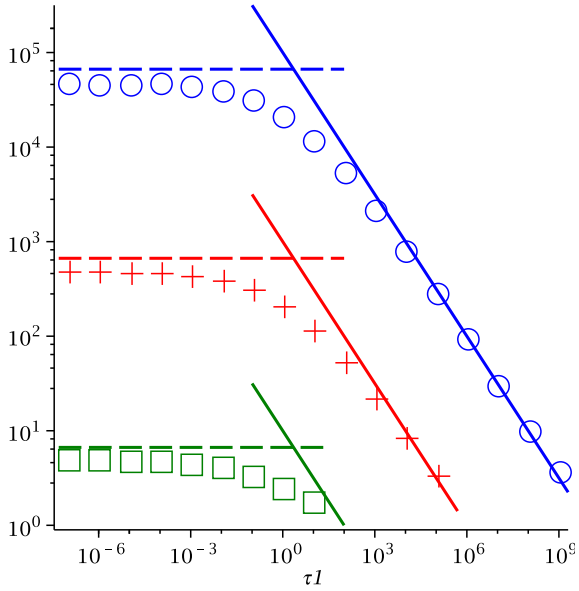


FIG. 5: $ln(G)$ en fonction de $ln(\tau_1)$ (τ_2 pris optimal). Prédiction analytique pour $\tau_1 \ll \tau$ (6) (pointillés). Prédiction analytique pour $\tau_1 \gg \tau$ (3) (ligne). simulations (points). $L = 10$ ($\square$, vert), $L = 10^3$ (+, rouge), $L = 10^5$ ($\circ$, bleu). $D = 1$, $V = 1$.

alors :

$$G \sim \frac{L}{\sqrt{D\tau_1}}. \tag{3}$$

– Dans le régime $\tau_1 \ll \tau$:

$$\frac{\tau_2^{opt}}{\tau} \sim \frac{1}{2}\sqrt{\frac{\tau_1}{\tau}}. \tag{4}$$

À l'optimum, le temps moyen de détection de la cible est :

$$\langle t \rangle^{opt} \sim \frac{3L}{4V}, \tag{5}$$

le gain G de l'intermittence par rapport à la phase d'inspection seule est :

$$G \sim \frac{2LV}{3D}. \tag{6}$$

Les simulations montrent que l'approximation utilisée dans les calculs analytiques est satisfaisante. Si la valeur de τ_2^{opt} obtenue à petit τ_1 (équation (4)) n'est pas tout à fait celle des simulations, la valeur de τ_2^{opt} obtenue à grand τ_1 (équation (1)) est en revanche en bon accord avec les simulations (voir Fig.4). Par ailleurs, le gain estimé est lui aussi en bon accord avec les simulations (voir Fig.5).

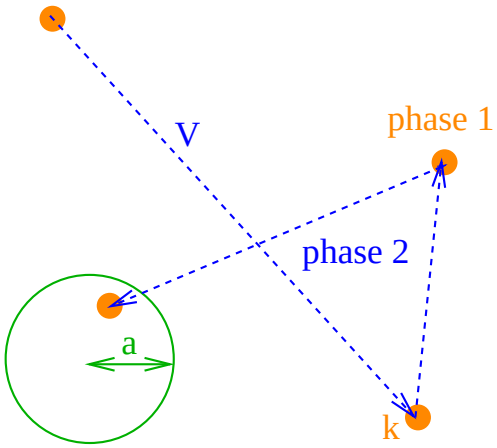
Ainsi, comme dans le cas de la répartition régulière de cibles, l'intermittence est favorable quand L est grand, avec à l'optimum $\tau_1 = \tau_1^{min}$. La dépendance de τ_2^{opt} avec τ_1 est un peu différente (loi de puissance, d'exposant 3/4 pour la répartition poissonnienne au lieu de 2/3 pour la répartition régulière).

1.2.4 Modèle de recherche intermittente à 2 dimensions : une alternative aux stratégies de Lévy

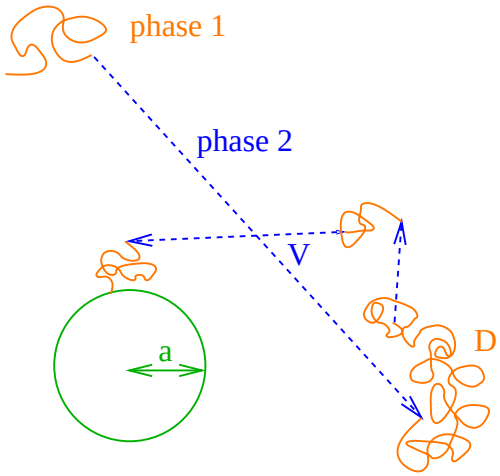
Motivation

Jusqu'ici, la recherche intermittente a été étudiée dans le cas unidimensionnel où la direction des phases balistiques est toujours la même. Comme la plupart des chercheurs "terrestres" se déplacent à 2 dimensions, nous présentons un modèle de recherche intermittente en deux dimensions, qui recouvre un champ d'applications beaucoup plus large. Sur la base de ce modèle aux ingrédients minimaux, nous montrons que les stratégies de recherche intermittentes à 2 dimensions permettent bel et bien de minimiser le temps de recherche dans le cas de cibles non revisitables, contrairement aux stratégies de Lévy, qui dans ce cas conduisent à l'optimum trivial donné par la trajectoire balistique (voir partie précédente). De ce point de vue, ce modèle fournit une alternative aux stratégies de Lévy. Nos résultats ont été publiés dans Bénichou et al. [2006], complétés par Bénichou et al. [2007], et ont fait l'objet d'un "*News and views*" de la revue *Nature* par Shlesinger [2006].

Modèle



Mode statique. Pendant la phase de détection, le chercheur est immobile et réagit avec la cible si celle-ci est à une distance inférieure à a avec une probabilité k par unité de temps.



Mode diffusif. Pendant la phase de détection, le chercheur a un mouvement diffusif (coefficient D), et réagit immédiatement avec la cible si elle est à une distance inférieure à a.

FIG. 6: Modèle de recherche intermittente : le chercheur alterne des phases de détection (phases 1) de durée moyenne τ_1 (distribution exponentielle), et des phases "aveugles" balistiques (vitesse V, direction aléatoire) (phases 2) de durée moyenne τ_2 (distribution exponentielle).

Le chercheur intermittent considéré dans ce modèle alterne comme précédemment des phases de détection (phases 1) de durée moyenne τ_1, distribuées exponentiellement, et des phases "aveugles" balistiques de vitesse V et de direction aléatoire (phases 2), de durée moyenne τ_2 distribuée exponentiellement. (voir Fig.6). Deux modélisations alternatives simples du mode de détection sont envisagées : (i) un premier mode "statique", pour lequel, pendant la phase de détection, le chercheur

est immobile et réagit avec la cible si elle est à une distance inférieure à a avec une probabilité k par unité de temps ; (ii) un deuxième mode "diffusif", pour lequel, pendant la phase de détection, le chercheur a un mouvement diffusif (coefficient D), et réagit immédiatement avec la cible si elle est à une distance inférieure à a. Chacun des deux modes (statique et diffusif) est pertinent par lui même, et permet de rendre compte de différentes situations réelles.

Méthodes

Le calcul se fait dans la géométrie suivante : un disque de rayon b et de bords réfléchissants, avec la cible au centre. Cette géométrie représente à la fois le cas d'une cible unique dans un domaine fini et celui d'un réseau régulier de cibles dans un espace infini.

Nous notons t_i le temps moyen qu'un chercheur met à trouver la cible quand il part de la position $\mathbf{r}$ dans l'état i (et avec la vitesse $\mathbf{V}$ pour l'état 2). Nous montrons que ces temps moyens vérifient des équations de type Fokker-Planck vers le passé [Redner, 2001] :

$$D\nabla^2_{\mathbf{r}} t_1 + \frac{1}{2\pi\tau_1}\int_0^{2\pi}(t_2 - t_1)d\theta_{\mathbf{V}} - k\mathrm{I}_a(\mathbf{r})t_1 = -1, \tag{7}$$

$$\mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 - \frac{1}{\tau_2}(t_2 - t_1) = -1, \tag{8}$$

où $\mathrm{I}_a(\mathbf{r})$ vaut 1 sur la cible ($r < a$), et 0 ailleurs. Ces intégrations intégro-différentielles ne semblent pas pouvoir être résolues analytiquement suivant des techniques standards. Nous avons donc eu recourt à une approximation, consistant essentiellement à supposer que la direction initiale de la vitesse n'a que peu d'influence sur le temps de recherche dans la limite de faible densité de cibles $b \gg a$. Cette approximation est validée par des simulations numériques.

Optimisation du temps de recherche

Le résultat principal de cette étude est que, dans la limite de faible densité, le temps de recherche admet un minimum global comme fonction des temps moyens τ_1 et τ_2 passés dans chacune des deux phases. La stratégie optimale est déterminée analytiquement, et prend la forme suivante : pour le mode diffusif (avec $D/V \ll a \ll b$) :

$$\tau_1^{opt} \sim \frac{D}{2V^2}\frac{\ln^2(b/a)}{2\ln(b/a) - 1},\ \tau_2^{opt} \sim \frac{a}{V}(\ln(b/a) - 1/2)^{1/2}, \tag{9}$$

alors que pour le mode statique ($a \ll b$) :

$$\tau_1^{opt} = \left(\frac{a}{Vk}\right)^{1/2}\left(\frac{2\ln(b/a) - 1}{8}\right)^{1/4},\ \tau_2^{opt} = \frac{a}{V}\left(\ln(b/a) - 1/2\right)^{1/2}. \tag{10}$$

Un point important est que la durée moyenne optimale de la phase rapide 2 est identique pour ces deux modes de détection. Ce résultat suggère que l'optimisation est assez largement indépendante du détail de la modélisation de la phase de détection. Nous verrons par la suite qu'il s'agit en fait d'une propriété assez générale des processus intermittents. Ces points sont développés de manière systématique dans la partie 1.4.

Conclusion

Ce modèle bidimensionnel a un statut de modèle minimal, puisqu'il ne met en jeu que les ingrédients essentiels caractérisant l'intermittence. En utilisant une approximation vérifiée numériquement, nous avons montré que ce modèle donne lieu à une minimisation du temps de recherche comme fonction des temps moyens passés dans chacune des phases. Les stratégies de recherche intermittentes sont donc de véritables stratégies optimales, contrairement aux stratégies de Lévy, qui elles ne sont optimales que dans le cas très spécifique de cibles revisitables précédemment discuté. En cela, ce modèle apparaît comme une alternative à ces stratégies de Lévy.

1.2.5 Conclusion : les animaux ont-ils vraiment intérêt à suivre les stratégies de Lévy ?

L'optimisation des stratégies de recherche dans le contexte de l'écologie comportementale est une question qui a suscité des polémiques récentes. Comme nous l'avons rappelé, nous avons remis en cause en 2006 d'un point de vue théorique le statut des marches de Lévy comme stratégies optimales de recherche, la meilleure stratégie obtenue dans ce cadre étant trivialement la ligne droite. Ce modèle des marches de Lévy a depuis été remis en cause d'un point de vue expérimental. En effet, une étude récente [Edwards et al., 2007] a montré que l'essentiel des données qui avaient conduit à l'identification des marches de Lévy dans les trajectoires animales avaient été en fait mal interprétées, et conduisaient plus vraisemblablement à des lois de type exponentiel ou loi gamma. Notre étude montre finalement que contrairement à une idée communément admise, il n'y a aucun paradoxe [Travis, 2007] à ce que les animaux ne suivent pas les stratégies de Lévy, puisqu'elles ne sont pas optimales. En outre, elle met en avant les stratégies intermittentes, qui elles sont réellement optimisables, comme modèle alternatif possible de trajectoires animales.

1.3 Stratégies de recherche intermittentes à l'échelle microscopique

Comme vu précédemment, les stratégies de recherche intermittentes sont observées à l'échelle macroscopique. Dans cette partie, nous allons voir sur deux exemples que les stratégies de recherche intermittentes sont aussi observées à l'échelle microscopique.

1.3.1 Recherche par une protéine d'une séquence spécifique sur l'ADN

Dans cette partie, nous définissons d'abord le concept de diffusion facilitée introduit par Berg et al. [1981], et montrons que ce mécanisme à deux états peut en fait s'interpréter comme un processus de recherche intermittent, au sens général défini en introduction : pour trouver plus vite leur cible, une séquence spécifique d'ADN, certaines protéines alternent des phases liées à l'ADN et des phases de diffusion à trois dimensions. Comme nous allons l'expliquer, connaître la distribution de la taille de ces excursions à trois dimensions est nécessaire pour quantifier la cinétique de ces réactions. Nous la calculons d'abord dans un cas très simple et adaptons ensuite ces

résultats pour interpréter quantitativement une expérience en molécule unique. Enfin, nous étendons ces résultats au cas d'un milieu encombré, comme l'est le milieu cellulaire.

Diffusion facilitée

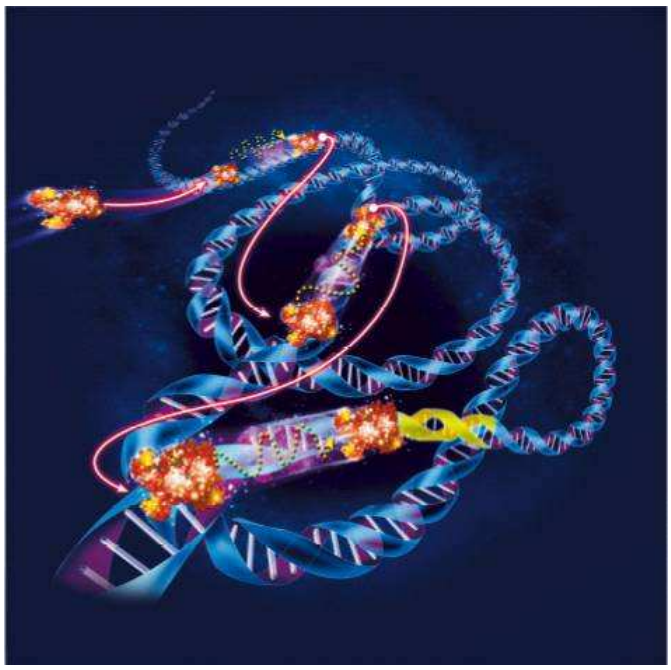


FIG. 7: Vue d'artiste d'une enzyme qui combine des diffusions à 1 et 3 dimensions pour trouver sa cible. Dessin de Virginie Denis/Pour la Science n°352, février 2007.

Si les virus qui menacent les humains bénéficient d'une grande couverture médiatique, il existe également de nombreux virus attaquant les bactéries. Pour se défendre, elles ne peuvent pas utiliser des cellules du système immunitaire comme nous le faisons, mais elles ont cependant leur propre système de défense. Elles produisent des enzymes qui reconnaissent une séquence spécifique sur l'ADN, suffisamment courte (4-8 paires de bases) pour être présente sur n'importe quel ADN viral (typiquement 5.10^4 paires de bases pour un bactériophage). Quand une telle enzyme trouve sa séquence cible, elle coupe l'ADN, inactivant ainsi le virus. Sur l'ADN de la bactérie cette séquence est protégée par une autre enzyme qui empêche ainsi la bactérie de détruire son propre ADN. En revanche, quand un virus pénètre la bactérie, une course contre la montre commence : l'enzyme doit trouver sa cible avant que le virus n'ait eu le temps de détourner la machinerie cellulaire à son profit.

D'une manière plus générale, de nombreuses protéines agissant sur des séquences spécifiques d'ADN sont connues pour trouver leur cible très vite. Par exemple, Riggs et al. [1970] ont montré que le répresseur *lac* agit sur sa cible à une vitesse plusieurs ordres de grandeur au-dessus de ce qui est prévu pour les réactions limitées par la diffusion classique. Dans une série de papiers fondateurs, Berg et al. [1981], Winter and Von Hippel [1981], Winter et al. [1981] ont proposé qu'en plus de la diffusion à 3 dimensions, les enzymes puissent se lier à l'ADN non-spécifique et diffuser à une dimension le long de la séquence (voir Fig.7). Il s'agit en fait d'une stratégie de

recherche intermittente dans le sens général défini dans l'introduction : la diffusion à 3 dimensions est relativement rapide, mais l'enzyme ne peut pas réagir avec sa cible, alors qu'à une dimension l'enzyme peut se lier à sa cible, mais la diffusion est beaucoup plus lente. Des modèles très simples (par exemple celui de Coppey et al. [2004]) montrent qu'effectivement le temps moyen de détection de la cible peut être minimisé par une telle stratégie intermittente. Cependant, des modèles plus réalistes sont nécessaires comme nous le détaillons dans la partie 4.1.1.

Un point particulièrement important concerne la description détaillée des excursions à 3 dimensions. En effet, suivant la conformation de l'ADN, un point proche dans l'espace à 3 dimensions peut être en fait très loin sur la séquence linéaire de l'ADN. On appelle *jumps* (que nous traduisons par *bonds*) les excursions à 3 dimensions qui reviennent sur l'ADN à un point de la séquence non corrélé avec le point de départ. Les excursions à 3 dimensions corrélées sont appelées *hops* (que nous traduisons par *sauts*) (voir Fig.8). Que dire de la statistique des sauts et des bonds, et de leur importance relative dans le mécanisme de diffusion facilitée ? Dans cette partie, la distribution des excursions à 3 dimensions est calculée dans un cas simple. Cette distribution est ensuite adaptée pour interpréter des expériences en molécule unique réalisées par Pierre Desbiolles, Andreas Biebriecher, Isabelle Bonnet et Natacha Porté de l'équipe optique et biologie du LKB ENS, Paris. Enfin, le cas d'un milieu encombré est discuté. Ces travaux ont donné lieu à deux publications [Bonnet et al., 2008, Loverdo et al., 2009c].

Distribution de la longueur des sauts

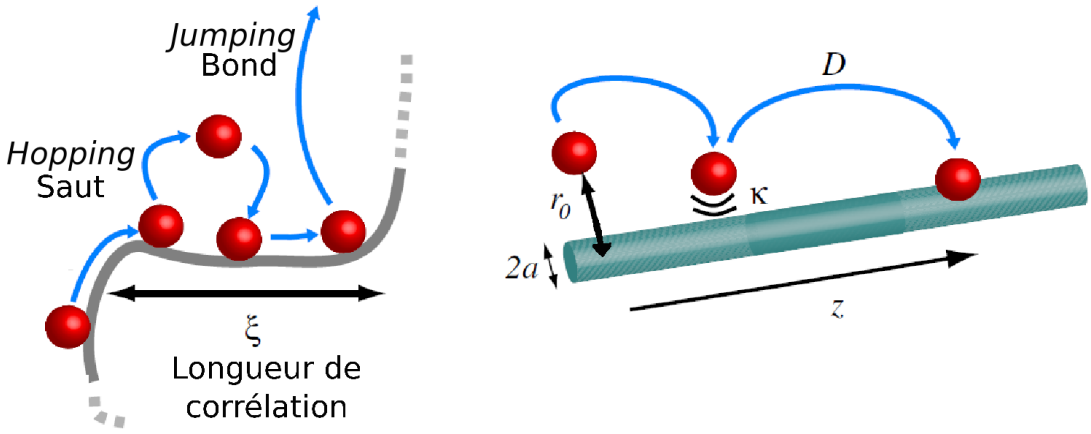


FIG. 8: Distribution des sauts. *Gauche :* Sauts et bonds. *Droite :* Paramètres du modèle.

Le modèle adopté pour les excursions à trois dimensions, très simple, consiste à représenter l'enzyme par un point qui diffuse librement en volume, et l'ADN par un cylindre infini de rayon a (la somme des rayons de l'ADN et de l'enzyme), le tout dans un espace infini (voir Fig.8). La position de l'enzyme à l'instant t est $\mathbf{r}$, z désignant la longueur selon l'axe du cylindre, et r la distance par rapport à l'axe du cylindre. L'enzyme part de $\mathbf{r_0}$ ($z = 0$ et $r = r_0$). La surface du cylindre est considérée comme semi-réfléchissante, ce qui se traduit par la condition au bord en

$r = a$ [Redner, 2001] :

$$\frac{\partial P(r,z,t|\mathbf{r_0})}{\partial r} = \kappa P(r,t|\mathbf{r_0}), \tag{11}$$

avec $P(r,z,t|\mathbf{r_0})$ la probabilité qu'une enzyme partant à $t = 0$ de $z = 0$ et $r = r_0$ soit à r et z à t. La cas limite $\kappa \to 0$ correspond à un cylindre réfléchissant, alors que $\kappa \to \infty$ correspond à un cylindre absorbant. Cette modélisation permet aussi de prendre le point de départ de l'enzyme à $r_0 = a$ (dans le cas contraire d'un cylindre parfaitement absorbant, l'enzyme se rattacherait aussitôt).

$P(r,z,t|\mathbf{r_0})$ satisfait l'équation de diffusion suivante :

$$\partial_t P(\mathbf{r},t|\mathbf{r_0}) = D\Delta_{\mathbf{r}} P(\mathbf{r},t|\mathbf{r_0}). \tag{12}$$

Au final (les calculs étant donnés dans la partie 4.1.2), la probabilité que l'enzyme se lie au cylindre en z est :

$$P(z|r_0) = \frac{1}{\pi}\int_0^\infty \cos(kz)\frac{K_0(kr_0)}{\frac{k}{\kappa}K_1(ka) + K_0(ka)}dk, \tag{13}$$

où $K_i(u)$ est une fonction de Bessel modifiée de deuxième espèce. Le comportement à grands z est donné par $P(z|r_0) \sim (\ln(r_0/a) + (\kappa a)^{-1})/(2z\ln^2(z/a))$. Cette distribution est ainsi très large, et la distance moyenne de retour est infinie. La valeur typique de la distance des retours peut néanmoins être estimée à partir de la médiane. Dans le cas où $r_0 = a$, nous montrons que celle-ci est de l'ordre de κ^{-1} quand $\kappa^{-1} \ll a$, et de l'ordre de $a\exp(1/(\kappa a))$ quand $\kappa^{-1} \gg a$. Cela donne une autre interprétation physique de κ, initialement défini par l'équation (11).

Cette distribution donne en fait accès à différentes quantités relatives au problème *in vivo*. D'une part, l'équation (13) donne la distribution analytique des sauts pour $|z| < \xi$ (ξ étant la longueur de corrélation de l'ADN), échelle à laquelle représenter l'ADN comme un cylindre droit est légitime. D'autre part, les retours à $|z| > \xi$ étant des bonds, la distribution cumulative complémentaire donne directement la proportion de bonds :

$$\bar{C}(|z = \xi|) = \int_{|z|>\xi} P(z|r_0)dz \sim \frac{\ln(r_0/a) + 1/\kappa a}{\ln(\xi/a)}. \tag{14}$$

Quand les cibles sont plus rapidement trouvées avec des bonds qu'avec des sauts, diminuer la longueur de persistance accélère le processus de recherche, comme trouvé par exemple par van den Broek et al. [2008]. Cependant, notre résultat montre que la dépendance dans cette longueur de persistance est seulement logarithmique.

Interprétation d'expériences

Nous adaptons désormais le calcul précédent aux spécificités d'une expérience de molécule unique. L'expérience de l'équipe optique et biologie du LKB consiste à observer l'interaction d'une enzyme de restriction, EcoRV, modifiée pour être fluorescente, avec des molécules d'ADN attachées par leurs extrémités à une surface (voir partie 4.1.3). Cette expérience permet de suivre comment une enzyme interagit avec une molécule d'ADN. On appelle "interactions effectives" les périodes durant lesquelles l'observation indique une colocalisation de l'enzyme avec l'ADN. Du fait

de la résolution temporelle et spatiale finie de l'expérience, ces interactions effectives peuvent en fait être composées de véritables interactions avec l'ADN, et de petits sauts que l'on ne peut pas distinguer. En pratique, une interaction effective s'arrête quand l'enzyme disparaît pendant un temps d'observation supérieur à $t_{obs} = 40$ ms, ou bien si pendant t_{obs} l'enzyme s'est déplacée de plus de $z_m = 200$ nm.

Deux éléments principaux sont observés. D'une part, l'interaction effective à une dimension, le long de l'ADN, peut être considérée comme diffusive ($D_{1D}^{\text{eff}} \simeq 10^{-2}$ µm^2.s^{-1} : à titre de comparaison, $D_{3D} \simeq 50$ µm^2.s^{-1}). D'autre part, il arrive qu'en $t_{obs} = 40$ ms, l'enzyme se déplace de plus de 200 nm, et ceci beaucoup plus fréquemment que la diffusion à une dimension ne peut l'expliquer. Deux questions se posent. Ces grands déplacements sont-ils compatibles avec de la diffusion à 3 dimensions, c'est-à-dire avec des sauts ? Si c'est le cas, des sauts plus petits que 200 nm doivent aussi exister, et ainsi influencer les propriétés de la diffusion effective à une dimension : pouvons-nous en rendre compte ? Pour répondre à ces questions, le modèle précédent est adapté.

Tout d'abord, la distribution des sauts est donnée précédemment pour un temps d'attente des retours infini, alors qu'ici seuls les retours avant t_{obs} doivent être pris en compte (voir partie 4.1.3), ce qui donne :

$$P_{t_{obs}}(z) = \int_0^{t_{obs}} P_{\parallel}(z,t) F_{\perp}(t|a) \mathrm{d}t, \tag{15}$$

avec $P_{\parallel}(z,t) = (4\pi D_{3D} t)^{-1/2} \exp(-z^2/4D_{3D}t)$ le propagateur correspondant à la diffusion le long de l'axe du cylindre. $F_{\perp}(t|a)$ désigne la distribution du temps de premier retour sur l'ADN dans le plan orthogonal à l'ADN, dont la transformée de Laplace (inversée ensuite numériquement) est donnée par [Redner, 2001] :

$$\widehat{F}_{\perp}(s|a) = \frac{K_0(x)}{K_0(x) + (\kappa a)^{-1} K_1(x)}. \tag{16}$$

Ensuite, la géométrie de l'expérience doit aussi être prise en compte. La molécule d'ADN est finie, et attachée à une surface (qui peut être considérée comme réfléchissante, car les enzymes interagissent faiblement avec celle-ci). Comme la géométrie réelle de l'expérience ne se prête pas à des développements analytiques, nous adoptons une géométrie effective, qui est celle d'un cylindre entre deux plans réfléchissants (voir Fig.9). La méthode des images permet alors d'obtenir la distribution $P^*(z)$ dans cette géométrie effective (voir partie 4.1.3 pour plus de détails) :

$$P^*(z) = \frac{1}{L} \sum_{n=0}^{\infty} Y_n \tag{17}$$

avec :

$$Y_0 = \int_z^{\infty} P(x)dx + (L - v)P(z) \tag{18}$$

$$Y_{n>0} = (L - z)\left(P(2nL + z) + P(2nL - z)\right) - \int_{2nL-z}^{2nL+z} P(x)dx. \tag{19}$$

Cette distribution permet d'interpréter la statistique des grands sauts observés expérimentalement. Les résultats expérimentaux montrent tout d'abord que cette

FIG. 9: Comparaison entre les géométries expérimentale et modèle. En haut : la géométrie expérimentale : une molécule d'ADN étirée à 70% et attachée par ses extrémités sur une surface réfléchissante. En bas : la géométrie utilisée dans le modèle : un cylindre entre deux surfaces réfléchissantes.

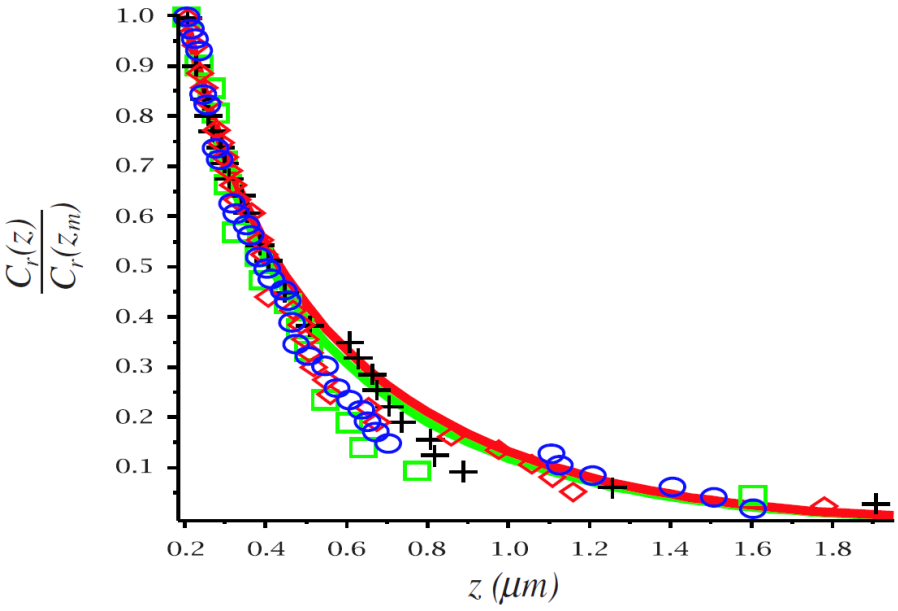


FIG. 10: Probabilité de faire un saut plus grand que z en fonction de z (µm), normalisé par le nombre de sauts plus grands que $z_m = 200$ nm. Data (dans la solution PIPES) : $+$ = 10 mM NaCl, $\diamond$ = 20 mM NaCl, $\square$ = 40 mM NaCl, $\circ$ = 60 mM NaCl. Lignes : distribution théorique pour $\kappa^{-1} = 0,5$ nm (vert), $\kappa^{-1} = 20$ nm (rouge).

distribution ne dépend que faiblement de la concentration en sel [NaCl], en accord avec des études biochimiques qui ont montré que l'association de l'enzyme avec l'ADN dépend très peu de la concentration [NaCl] [Lohman, 1986]. De plus, notre

modèle rend bien compte de la dépendance en z de la distribution des sauts observés (voir Fig.10 et partie 4.1.3).

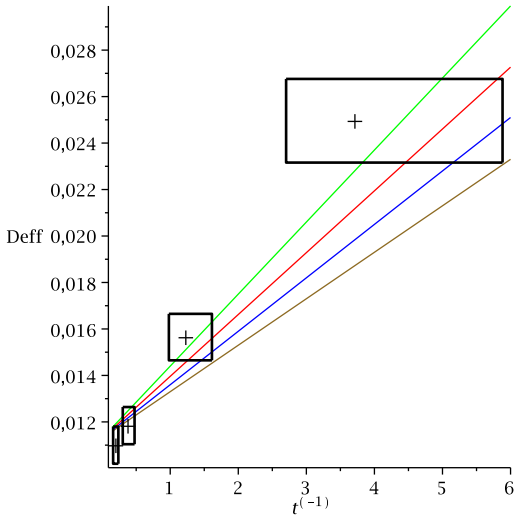


FIG. 11: D_{1D}^{eff} ($\mu\text{m}^2.\text{s}^{-1}$) en fonction de $1/\tau$ (s^{-1}) (τ est la durée moyenne d'une interaction effective). Données expérimentales (points noirs) avec les barres d'erreur (boîtes noires). Le point à 10 mM de sel est en bas à gauche, le point à 60 mM de sel est en haut à droite. Formule (20) pour $\kappa^{-1} = 0,5$ nm (vert), $\kappa^{-1} = 2$ nm (rouge), $\kappa^{-1} = 3,3$ nm (bleu), $\kappa^{-1} = 5$ nm (marron).

Les interactions effectives observées combinent alors de la diffusion véritablement à une dimension le long de l'ADN, et des sauts trop petits pour être distingués. Ces sauts vont influer sur le coefficient de diffusion D_{1D}^{eff}. Comme vu précédemment, la concentration en sel ne change pas les sauts. En revanche, plus il y a de sel, plus l'enzyme se détache vite de l'ADN, et plus l'effet des sauts va se faire sentir sur D_{1D}^{eff} : le temps d'interaction effectif diminue quand la concentration en sel augmente (voir Fig.11). On obtient (voir partie 4.1.3) :

$$\frac{D_{1D}^{\text{eff}}}{D_{1D}} = 1 + \frac{(N-1)\langle l_{3D}^2 \rangle}{2D_{1D}\tau}, \tag{20}$$

où τ est la durée moyenne d'une interaction effective, $N-1$ le nombre moyen de sauts dans une interaction effective (le saut N étant celui où l'enzyme se "perd"), D_{1D} le coefficient réel de diffusion à une dimension, et $\langle l_{3D}^2 \rangle$ la distance quadratique moyenne des sauts trop petits pour être distingués. On peut rendre compte des points expérimentaux avec cette équation pour $\kappa^{-1} \leq 5$ nm (voir Fig.11), ce qui donne une borne supérieure à ce paramètre.

Tous ces éléments montrent que l'enzyme combine bien une diffusion à une dimension et une diffusion libre à 3 dimensions.

Effet de l'encombrement cellulaire

L'expérience ci-dessus est réalisée *in vitro*, dans un milieu dans lequel la diffusion est effectivement normale. Le milieu cellulaire est cependant décrit comme très encombré. Le transport observé est alors souvent sous-diffusif [Tolic-Norrelykke et al.,

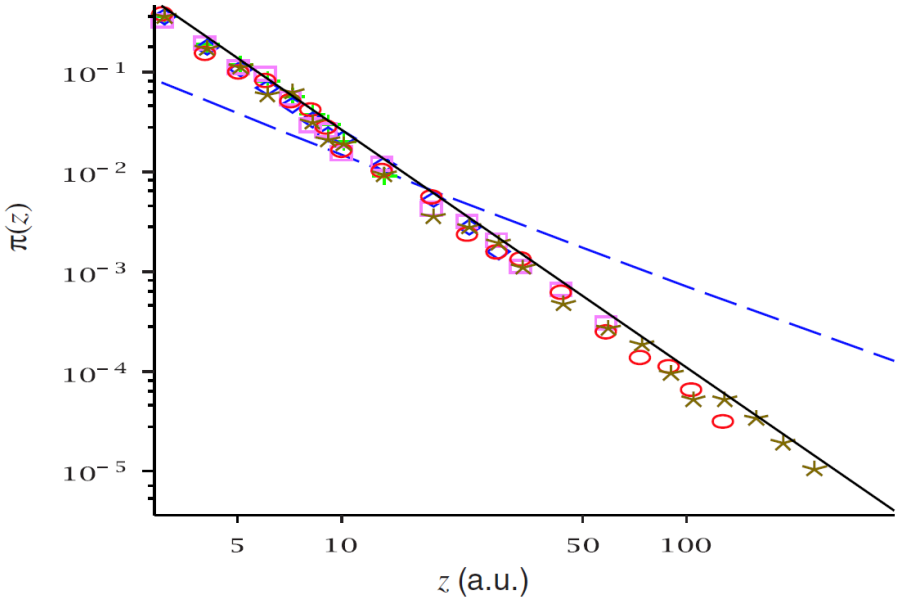


FIG. 12: Distribution de la taille des sauts $P(z|\mathbf{r_0})$ pour des excursions sur un amas de percolation critique dans un réseau cubique à 3 dimensions, pour $r_0 = 1$ et $a = 0$. Des simulations (normalisées pour $z \geq 2$) sont faites pour différentes tailles de système (points), et se confondent avec la courbe théorique (ligne) obtenue par l'équation (21). Ici $d_w^{\|} = d_w^{\perp} = 3,88...$, $d_f^{\perp} = 2$ et $d_f^{\|} = 1$ puisque $d_f = 2,53...$ [Ben-Avraham and Havlin, 2000]. À grands z, $P(z|\mathbf{r_0}) \sim 1/z^{2,88...}$, et est comparé avec la diffusion normale pour $r_0 = 1$ et $a = 0,1$ (pointillés).

2004, Golding and Cox, 2006], c'est-à-dire caractérisé par un déplacement quadratique moyen vérifiant : $\langle r^2(t)\rangle \propto t^{2/d_w}$, où $d_w > 2$ est la dimension de la marche [Metzler and Klafter, 2000]. Plusieurs mécanismes microscopiques peuvent induire un tel effet. Une première possibilité est d'avoir des temps de piégeage très longs pendant la marche (modèles de type "continuous time random walks" [Metzler and Klafter, 2000]). Dans ce cas, une excursion à 3 dimensions prend plus de temps, mais la trajectoire n'est pas modifiée. la distribution des sauts obtenue pour un temps d'observation infini sera donc inchangée, puisqu'il s'agit d'un résultat purement géométrique.

Une deuxième possibilité est que le milieu soit fractal. Dans ce dernier cas, la distribution des sauts va changer. Comme précédemment, nous supposons que la distribution des sauts peut s'écrire en décomposant la diffusion selon l'axe du cylindre et selon le plan perpendiculaire. Au final, nous obtenons (voir partie 4.1.4 pour les détails) :

$$P(z|\mathbf{r_0}) = \alpha \int_0^{\infty} u^{(\nu_\perp - \nu_\|)/2} J_{\nu_\perp}(\beta\sqrt{u}) K_{\nu_\|}(\gamma\sqrt{u}) \mathrm{d}u, \tag{21}$$

avec $\gamma = 2z^{d_w/2}/d_w$, $\beta = 2r_0^{d_w/2}/d_w$, $\nu_i = 1 - d_f^i/d_w$ (ici $i = \perp, \|$) et

$$\alpha = \frac{2 d_w^{\nu_\| - \nu_\perp - 1} r_0^{\nu_\perp d_w/2} z^{(d_f^\| + d_w)/2 - 1}}{\Gamma(\nu_\perp)\,\Gamma\left(1 - \nu_\|\right)}. \tag{22}$$

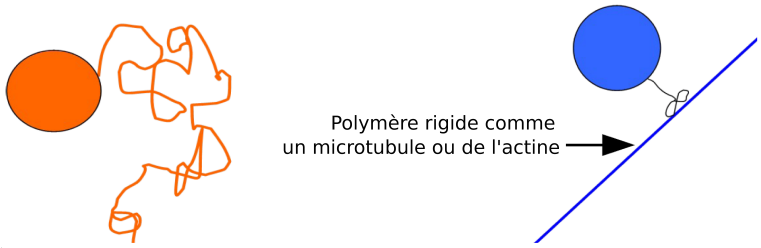
d_f est la dimension de la fractale, $d_f^{\|}$ est la dimension de cette fractale projetée

sur l'axe du cylindre, et $d_f^{\perp}$ est la dimension de cette fractale projetée sur le plan orthogonal au cylindre. Ce résultat est confirmé par des simulations numériques (voir Fig.12). La proportion de bonds est dans ce cas proportionnelle à $\bar{C}(\xi) \sim \xi^{-d_w+d_f^{\perp}}$: elle est donc beaucoup plus petite que dans le cas de la diffusion normale, ce qui montre que l'encombrement cellulaire peut changer les caractéristiques de la recherche.

1.3.2 Transport actif de vésicules en milieu cellulaire

À l'échelle microscopique, les stratégies intermittentes sont également observées dans le cas du transport actif de vésicules (ou d'autres types de traceurs) en milieu cellulaire. Nous avons proposé une théorie des réactions limitées par un tel type de transport intermittent. Ce travail a été publié dans Loverdo et al. [2008], et a été commenté par Mirny [2008] dans la rubrique "news and views" de Nature Physics.

Vésicules



Sans ATP
Distance$^2 \propto Dt$ (ou moins)
Le coefficient de diffusion diminue quand la taille de l'objet augmente

Consomme de l'ATP
Distance $= Vt$
Vitesse ~ 1 µm.s^{-1}, presque indépendante de la taille de l'objet (voir par exemple [Block et al., 2003])

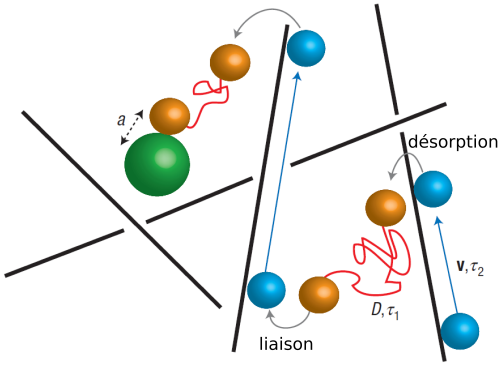
FIG. 13: Options pour le transport de vésicules à l'intérieur de la cellule.

Du fait de la nature intrinsèquement hors d'équilibre du milieu cellulaire, le transport d'un traceur dans une cellule peut être très différent de la diffusion d'origine thermique ordinaire. En effet, dans une cellule, un traceur type comme une vésicule peut diffuser (voire sous-diffuser), mais aussi s'associer à des moteurs moléculaires qui, en consommant de l'ATP (l'énergie chimique cellulaire), "marchent" le long des filaments du cytosquelette, induisant un déplacement balistique [Alberts, 2002] (voir Fig.13).

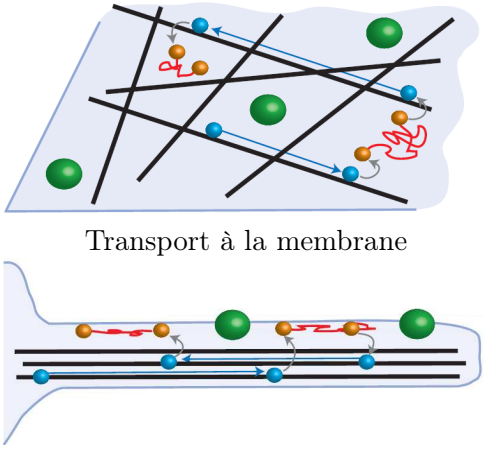
Ce transport dit "actif" est bien intermittent au sens général défini en introduction. Il met en effet en jeu une phase lente, ici la diffusion thermique, et une phase rapide de transport balistique induit par les moteurs. Par ailleurs, on s'intéressera ici à des situations dans lesquelles le traceur ne peut pas réagir quand il est associé aux moteurs, comme par exemple dans le cas où les cibles sont des protéines membranaires. De telles trajectoires intermittentes ont notamment été observées à l'échelle de la vésicule unique par Huet et al. [2006]. Ces expériences semblent in-

diquer qu'effectivement, la réaction (dans ce cas de l'exocitose) est nettement plus favorable dans la phase diffusive que dans la phase balistique.

Ce transport actif est important pour la localisation finale des protéines transportées par les vésicules, comme souligné dans la partie 8.4. De manière générale, on s'intéresse ici à l'impact de ce type de transport sur la cinétique réactionnelle en milieu cellulaire. Pour cela, nous estimons la constante cinétique pour une réaction de type premier ordre par l'inverse du temps moyen de premier passage sur la cible. Techniquement, les calculs apparaissent comme un cas particulier de l'étude systématique présentée dans la partie 1.4.



Transport de vésicules dans le volume de la cellule



Transport à la membrane

Transport dans des structures tubulaires comme les dendrites ou les axones

FIG. 14: Transport des vésicules dans les cellules : différentes géométries pertinentes.

Optimisation de la constante cinétique

Dans le cas d'un traceur se déplaçant en 3 dimensions, par exemple dans le volume du cytoplasme (voir Fig.14 gauche), nous montrons que le transport intermittent est plus efficace qu'un simple transport diffusif et qu'il permet de maximiser la constante de réaction dès que le rayon de réaction vérifie : $a > a_c \simeq 6.5\frac{D}{V}$ (voir partie 5.5.2). Dans des conditions cellulaires normales, D va de 10^{-2} $\mu m^2.s^{-1}$ pour des vésicules à 10 $\mu m^2.s^{-1}$ pour de petites protéines, et la vitesse typique d'un moteur est de l'ordre de $V \sim 1$ $\mu m.s^{-1}$. Le rayon de réaction critique obtenu va alors de 10 nm, ce qui est plus petit que n'importe quelle organelle cellulaire, jusqu'à 10 μm, ce qui est comparable à la taille de la cellule tout entière. Pour des objets assez gros comme des vésicules, le transport par les moteurs permet donc de mieux explorer l'espace, alors que pour des objets plus petits, des petites protéines par exemple, la diffusion seule est plus efficace. De plus, dans ces mêmes conditions standards, le temps moyen d'interaction avec les moteurs correspondant à la stratégie optimale est de l'ordre de 0,1 s pour un rayon de réaction de 0,1 μm. Cette valeur est compatible avec les observations expérimentales [Alberts, 2002], ce qui suggère que le transport cellulaire pourrait être proche de l'optimum. Enfin, pour une vésicule de rayon de réaction $a \gtrsim 0,1$ μm, toujours dans des conditions cellulaires standards, on peut montrer que la constante de réaction peut être jusqu'à 10 fois plus grande que dans le cas d'une diffusion seule.

Dans le cas de traceurs évoluant à 2 dimensions, typiquement sur des membranes, ou à 1 dimension dans des structures tubulaires comme les axones ou les dendrites (voir Fig.14 droite), la constante cinétique peut également être optimisée. Cette fois, la constante cinétique optimale peut devenir beaucoup plus grande. À une dimension, pour de faibles concentrations de cibles, elle est peut être 100 fois plus grande que dans le cas d'une diffusion seule, car contrairement au cas à trois dimensions, le gain dépend de la concentration de cibles.

1.4 Les stratégies intermittentes : des stratégies robustes

1.4.1 Motivations

Les stratégies de recherche intermittentes sont finalement observées aussi bien à l'échelle microscopique que macroscopique. Les modèles utilisés pour interpréter les trajectoires d'animaux (partie 1.2.4) et pour les vésicules (partie 1.3.2) sont très semblables. Cette similarité appelle au développement d'un modèle générique qui généralise ces aspects. Nous étudions dans cette partie ce modèle de marches aléatoires intermittentes de manière systématique à 1, 2 et 3 dimensions, et pour trois descriptions différentes de la phase de détection, d'une part pour mieux rendre compte de la diversité des situations réelles, et d'autre part pour tester la robustesse des résultats. Cette étude systématique a été publiée dans Loverdo et al. [2009a].

La quantité minimisée est le temps moyen de premier passage sur la cible. Est-ce que l'intermittence peut permettre de diminuer le temps moyen de détection de la cible par rapport à la phase de détection seule ? Et si c'est le cas, existe-t-il une manière optimale de répartir le temps de recherche entre les deux phases ?

1.4.2 Ingrédients du modèle générique

Dimensions de l'espace

Dans les exemples biologiques, l'espace de recherche est à une, deux ou trois dimensions (voir Fig.15). De nombreux modèles de recherche sont à une dimension (par exemple Oshanin et al. [2007], Bressloff and Newby [2009], Rojo et al. [2009], Reynolds [2006]). Au-delà du fait que les calculs sont souvent plus simples à une dimension, certains problèmes réels de recherche sont effectivement à une dimension. À l'échelle macroscopique, on peut penser aux fourmis, qui suivent souvent des bords, ou des pistes [Dussutour et al., 2005]. Le modèle à deux dimensions est particulièrement pertinent pour des animaux vivant à la surface du sol. D'autres animaux utilisent les 3 dimensions de l'espace, comme les poissons, le plancton [Bartumeus et al., 2003], ou le ver *C.elegans* dans son habitat naturel (le sol) [Kiontke and Sudhaus, 2005]. À l'échelle microscopique, comme nous l'avons vu dans la partie 1.3.2, les espaces à 1, 2 et 3 dimensions sont pertinents pour décrire le trafic intracellulaire [Alberts, 2002].

Descriptions de la phase de détection

La phase 1 est une phase de déplacement lent durant laquelle la détection est possible. Le rayon de détection a est la distance maximale au-delà de laquelle le

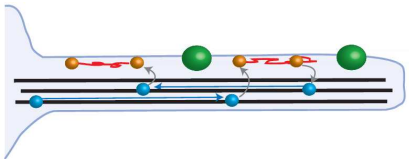
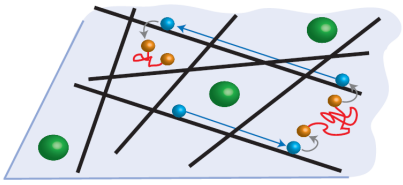
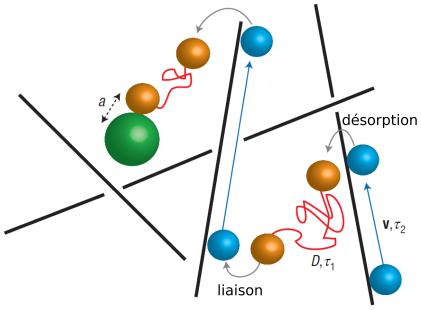
Microscopique		Macroscopique
	1D	
Vésicules dans des structures cellulaires tubulaires comme les dendrites ou les axones.		La ligne sombre est une piste de fourmis.
	2D	
Trafic de vésicules à une membrane.		Pêcheurs de coques : les coques sont cachées sous la surface. Photo de Michel Sokolowski [Lepolard, 2007].
	3D	
Vésicules dans le cytoplasme.		Les poissons utilisent les 3 dimensions de l'espace.

FIG. 15: Illustrations microscopiques et macroscopiques à 1, 2 et 3 dimensions.

chercheur ne peut jamais trouver d'information sur la position de la cible. Trois modélisations différentes de cette phase sont proposées (voir Fig.16) :

- *Mode statique* : le chercheur est immobile, et si la cible est à une distance du chercheur inférieure à a, il la détecte avec une probabilité k par unité de temps.
- *Mode diffusif* : le mouvement du chercheur est diffusif, de coefficient de diffusion D, et la cible est immédiatement détectée si elle est à une distance inférieure à a du chercheur.
- *Mode balistique* : le mouvement du chercheur est balistique (vitesse v_l) dans une direction aléatoire. La cible est détectée immédiatement si elle est à une distance inférieure à a. On peut remarquer que Viswanathan et al. [1999] ont proposé un mode de détection similaire, mais sans autre phase disponible. S'il

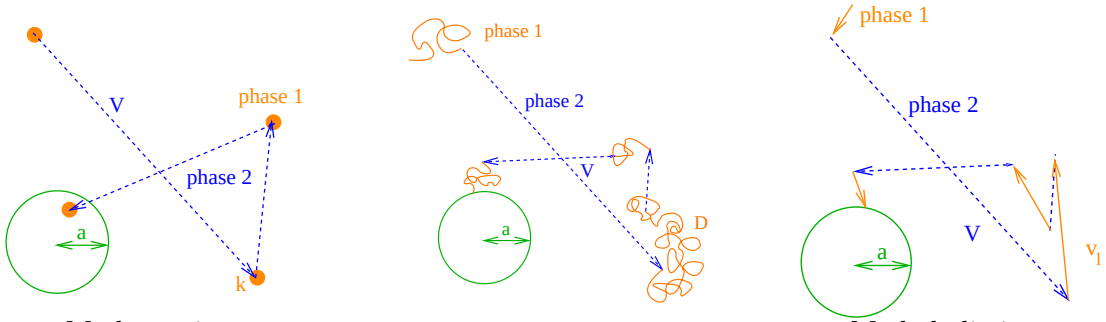


FIG. 16: Les différentes descriptions de la phase 1, la phase où la cible peut être détectée (ici représentés à deux dimensions).

n'y a pas d'autre phase disponible, et si la cible est détruite quand elle est trouvée (on ne peut pas gagner à repasser sur la même cible), la meilleure stratégie est la ligne droite, sans jamais se réorienter (voir partie 1.2.1). Nous allons voir dans la suite que si le chercheur a accès à une phase plus rapide (même aveugle), le temps moyen de détection de la cible peut être encore plus court.

Ces différents modes permettent de représenter des situations réelles très variées. Par exemple, quand des animaux cherchant de la nourriture sont observés [Bell, 1991, O'Brien et al., 1990], les phases de détection sont décrites comme presque immobiles (statiques), aléatoires (diffusives), ou avec une vitesse faible (balistique). Certaines situations réelles combinent plusieurs modes. Par exemple, quand les vésicules ne sont pas liées à un moteur, elles diffusent, et peuvent aussi ne pas réagir parfaitement avec leur cible : c'est donc une combinaison des modes statique et diffusif. Pour des raisons de simplicité, nous traitons ces trois modes séparément. En combinant ces modes schématiques, une grande partie des déplacements imaginables peut être décrite, depuis la sous-diffusion (voire même l'immobilité), jusqu'à la super-diffusion (voire même un déplacement balistique).

Mémoire minimale

La phase 2 est une phase de déplacement rapide pendant laquelle le chercheur ne peut pas détecter sa cible. Il se déplace à une vitesse constante V, dans une direction tirée au hasard à chaque nouvelle phase 2, sans corrélation avec la phase précédente : le chercheur n'a pas de mémoire spatiale.

La probabilité de passer d'une phase à l'autre est fixe, indépendante du temps. Le chercheur n'a donc pas non plus de mémoire temporelle. La distribution de la durée des phases est donc exponentielle, de moyenne τ_i pour la phase i.

1.4.3 Méthodes

Géométrie

La géométrie utilisée dans les calculs consiste en une sphère réfléchissante de dimension d, de rayon b, avec une cible immobile en son centre. Pour tenir compte du fait que le chercheur ne connaît pas initialement la position de la cible, le temps

moyen de détection de la cible est moyenné sur la position de départ, prise aléatoirement uniformément dans la sphère. Cette géométrie représente bien deux situations : une cible dans un domaine fini, ou des cibles régulièrement espacées dans un espace infini. C'est aussi une approximation de champ moyen d'autres distributions. Comme vu précédemment dans la partie 1.2.3, une distribution aléatoire de cibles peut modifier significativement les résultats à une dimension. Cependant, nous nous attendons à ce que cette différence soit moindre pour de plus grandes dimensions.

Équations de base

t_1 est défini comme le temps moyen de premier passage sur la cible en partant de la phase 1 à la position $\mathbf{r}$, et t_2 comme le temps moyen de premier passage sur la cible en partant de la phase 2 à la position $\mathbf{r}$ et avec la vitesse $\mathbf{V}$. L'angle solide de la sphère à d dimensions est noté Ω_d. Nous écrivons les équations vers le passé satisfaites par t_1 et t_2 [Redner, 2001].

Pour le mode statique, avec I_a une fonction qui vaut 1 si $\mathbf{r}$ est dans la cible, 0 ailleurs, nous obtenons :

$$\begin{cases} \dfrac{1}{\tau_1}\displaystyle\int (t_2 - t_1)\frac{d\omega_{\mathbf{V}}}{\Omega_d} - k\mathrm{I}_a(\mathbf{r})t_1 = -1 \\ \mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 - \dfrac{1}{\tau_2}(t_2 - t_1) = -1 \end{cases} \tag{23}$$

Pour le mode diffusif, ces équations s'écrivent :

$$\begin{cases} D\Delta_{\mathbf{r}} t_1 + \dfrac{1}{\tau_1}\displaystyle\int (t_2 - t_1)\frac{d\omega_{\mathbf{V}}}{\Omega_d} = -1 \\ \mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 - \dfrac{1}{\tau_2}(t_2 - t_1) = -1 \\ t_1(r < a) = 0 \end{cases} \tag{24}$$

Pour le mode balistique, avec t_1 qui est ici le temps moyen de premier passage sur la cible en partant de la phase 1 à la position $\mathbf{r}$ et avec la vitesse $\mathbf{v_l}$, nous écrivons :

$$\begin{cases} \mathbf{v_l}\cdot\nabla_{\mathbf{r}} t_1 + \dfrac{1}{\tau_1}\displaystyle\int (t_2 - t_1)\frac{d\omega_{\mathbf{V}}}{\Omega_d} = -1 \\ \mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 + \dfrac{1}{\tau_2}\displaystyle\int (t_1 - t_2)\frac{d\omega_{\mathbf{v_l}}}{\Omega_d} = -1 \\ t_1(r < a) = 0 \end{cases} \tag{25}$$

Résolution

Comme ces équations sont intégro-différentielles, elles ne peuvent pas être résolues exactement dans le cas général. À une dimension, elles peuvent être résolues exactement pour les trois modes de détection. À deux et trois dimensions, l'étude numérique du mode balistique montre que l'optimum est soit pour $\tau_2 \to 0, \tau_1 \to \infty$ (pas d'intermittence), soit pour $\tau_1 \to 0$ (intermittence) : nous avons ensuite étudié ces deux cas analytiquement. Pour les modes statique et diffusif, nous utilisons une approximation valable dès que $b \gg a$, vérifiée par simulation numérique.

1.4.4 Résultats principaux

La base de ce modèle est le partage du temps de recherche entre deux phases, l'une qui permet de détecter la cible, mais avec un déplacement qui ne permet pas une exploration efficace de l'espace, l'autre pendant laquelle le déplacement est plus efficace mais sans détection de la cible. L'étude systématique de ce modèle pour 3 modes différents de description de la phase de détection, et à 1, 2 et 3 dimensions, permet d'une part de couvrir un grand nombre de situations réelles, et d'autre part de tirer des conclusions générales sur les propriétés des marches aléatoires intermittentes.

Plus précisément, le temps moyen de premier passage sur la cible a été calculé, et minimisé en fonction des durées moyennes des deux phases. Le tableau 1 récapitule les résultats. En particulier, cette étude montre qu'il y a toujours des régimes (que nous avons identifiés) dans lesquels l'intermittence est favorable. Il existe un minimum global du temps moyen de détection de la cible en fonction des durées moyennes de chacune des deux phases. Les durées optimales des phases, ainsi que le gain par rapport à la phase de détection seule, dépendent de la densité de cibles a/b en 1 dimension. Cette dépendance n'est plus que logarithmique à 2 dimensions, et disparaît totalement à 3 dimensions. L'intérêt accru de l'intermittence en faible dimension peut se comprendre. En effet, pour des temps longs, les trajectoires intermittentes que nous proposons sont équivalentes à des marches aléatoires dont la dimension est 2. À une dimension, la marche est récurrente, c'est-à-dire que le chercheur repasse souvent à des endroits déjà inspectés, et les long déplacements balistiques sont donc utiles pour aller vers des zones encore vierges. À trois dimensions, la marche est transitoire, c'est-à-dire qu'il y a toujours des zones proches inexplorées, donc les phases balistiques longues sont moins utiles.

Enfin, il y a des régimes dans lesquels la durée optimale de la phase balistique ne dépend pas du mode de détection. Cela montre que les stratégies optimales de recherche intermittentes sont robustes. Leur robustesse et leur efficacité pourrait expliquer pourquoi elles sont observées si souvent, et dans des contextes si différents.

	Mode statique	Mode diffusif			Mode balistique	
1D	toujours intermittence	$b < \frac{D}{V}$	$b > \frac{D}{V}$, $a \ll \sqrt{\frac{b}{a}}\frac{D}{V}$	$b > \frac{D}{V}$, $a \gg \sqrt{\frac{b}{a}}\frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq \frac{V}{2}\sqrt{\frac{3a}{b}}$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt{\frac{a}{2Vk}}\sqrt[4]{\frac{b}{3a}}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \simeq \left(\frac{b^2 D}{36V^4}\right)^{\frac{1}{3}}$	$\tau_1^{opt} \simeq \frac{Db}{48V^2 a}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \left(\frac{2b^2 D}{9V^4}\right)^{\frac{1}{3}}$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$
	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}\left(1+\sqrt{\frac{V}{2ka}}\sqrt[4]{\frac{3a}{b}}\right)^2$	$t_m^{opt} \simeq \frac{b^2}{3D}$	$t_m^{opt} \simeq \left(\frac{3^5 b^4}{2^4 DV^2}\right)^{\frac{1}{3}}$	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}$	$t_m^{opt} \simeq \frac{b}{v_l}$	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}$
2D	toujours intermittence	$b < \frac{D}{V}$	$b \gg \frac{D}{V} \gg a$	$b \gg a \gg \frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq \pi V / \left(4\sqrt{\ln\left(\frac{b}{a}\right)}\right)$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt{\frac{a}{2Vk}}\sqrt[4]{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \simeq \frac{b^2}{D}\frac{4\ln w - 5 + c}{w^2(4\ln w - 7 + c)}$	$\tau_1^{opt} \simeq \frac{D}{2V^2}\frac{\left(\ln\left(\frac{b}{a}\right)\right)^2}{2\ln\left(\frac{b}{a}\right)-1}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{b}{V}\frac{\sqrt{4\ln w - 5 + c}}{w}$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$
	$t_m^{opt} \simeq \frac{2b^2}{aV}\left(\sqrt[4]{\ln\left(\frac{b}{a}\right)} + \sqrt{\frac{V}{2ak}}\right)^2$	$t_m^{opt} \simeq \frac{b^2}{2D}\ln\left(\frac{b}{a}\right)$	t_m^{opt}, c, w : partie 5.4.2	$t_m^{opt} \simeq \frac{2b^2}{aV}\sqrt{\ln\left(\frac{b}{a}\right)}$	$t_m^{opt} \simeq \frac{\pi b^2}{2av_l}$	$t_m^{opt} \simeq \frac{2b^2}{aV}\sqrt{\ln\left(\frac{b}{a}\right)}$
3D	toujours intermittence	$a \lesssim 6\frac{D}{V}$		$b \gg a \gtrsim 6\frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq 0,6V$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt[4]{\frac{3}{10}}\sqrt{\frac{a}{Vk}}$	$\tau_1^{opt} \to \infty$		$\tau_1^{opt} \simeq \frac{6D}{V^2}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq 1,1\frac{a}{V}$	$\tau_2^{opt} \to 0$		$\tau_2^{opt} \simeq 1,1\frac{a}{V}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq 1,1\frac{a}{V}$
	$t_m^{opt} \simeq \frac{b^3}{Va^2}\left(\sqrt[4]{\frac{24}{5}} + \sqrt{\frac{V}{ak}}\right)^2$	$t_m^{opt} \simeq \frac{b^3}{3Da}$		$t_m^{opt} \simeq 2,2\frac{b^3}{Va^2}$	$t_m^{opt} \simeq \frac{4b^3}{3a^2 v_l}$	$t_m^{opt} \simeq 2,2\frac{b^3}{Va^2}$

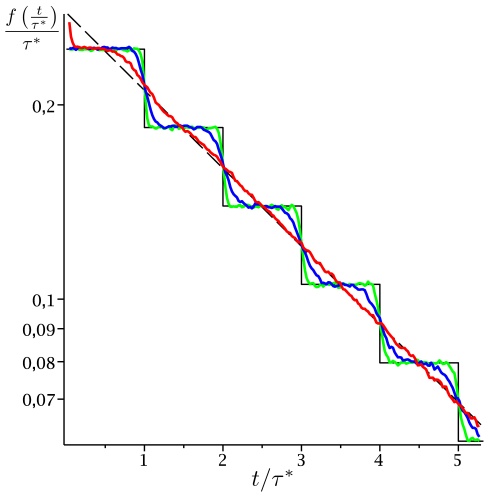
TAB. 1 – Récapitulation des résultats principaux du modèle générique : les stratégies minimisant le temps moyen de premier passage. Dans chaque case, régime de validité, τ_1 à l'optimum, τ_2 à l'optimum, t_m au minimum (t_m calculé pour $\tau_i = \tau_i^{opt}$). Le fond jaune souligne la valeur de τ_2^{opt} indépendante de la description de la phase de détection. Les résultats sont donnés dans la limite où $b \gg a$.

1.5 Extensions et perspectives

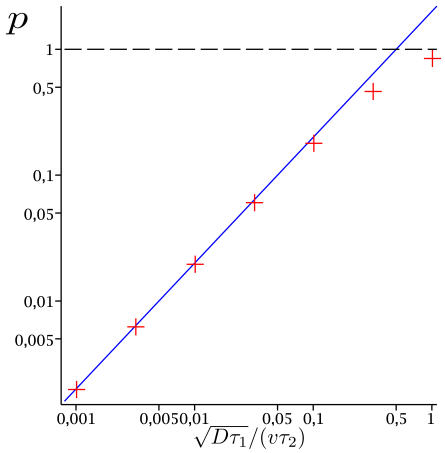
Nous présentons dans cette partie trois axes de développement du modèle générique détaillé dans la partie précédente, dont la finalisation est en cours. D'abord, nous n'avons regardé dans la partie précédente que le temps moyen de premier passage, alors qu'il pourrait ne pas être représentatif du temps typique. La distribution complète du temps de premier passage sur la cible est nécessaire pour vérifier que le choix du temps moyen comme quantité à minimiser est le bon. Ensuite, dans ce modèle, la mémoire du chercheur était minimale, et l'effet de mémoire temporelle ou spatiale est à étudier. Enfin, nous présentons des donnés expérimentales de recherche de cibles à l'échelle humaine.

1.5.1 Distribution complète du temps de recherche

Dans la partie précédente, nous avons calculé et minimisé le temps moyen de premier passage sur la cible. La distribution complète du temps de premier passage sur la cible donne beaucoup plus d'informations, mais son calcul analytique est évidemment beaucoup plus difficile.



Distribution du temps de premier passage sur la cible renormalisé $f(t/\tau^*)/\tau^*$ en fonction de t/τ^* (τ^* défini éq. (26)). Prévision analytique (27) avec p ajusté (=0,244) (ligne noire). Simulations (couleurs) avec $L = 2.10^5$ (rouge), $L = 2.10^6$ (bleu), $L = 2.10^7$ (vert). Distribution exponentielle (pointillés noirs). $\tau_1 = 10^6$, $\tau_2 = 7100$.



Probabilité p de trouver la prochaine cible en fonction de $\sqrt{D\tau_1}/(V\tau_2)$. Simulations (croix rouges), prédiction théorique $p = 2\sqrt{D\tau_1}/(V\tau_2)$ (28) (ligne bleue), borne supérieure $p = 1$ (pointillés noirs). $L = 10^5$, $\tau_2 = 10^3$.

FIG. 17: Distribution du temps de premier passage sur la cible dans le cas de cibles régulièrement espacées en une dimension avec les phases balistiques toujours dans la même direction : comparaison entre simulations et expressions analytiques. $D = 1, V = 1$.

Nous présentons ici une première distribution, celle obtenue dans une limite d'un cas unidimensionnel, plus précisément celui de la partie 1.2.2 : la phase de détection est diffusive, et la phase sans détection est balistique, toujours dans la

même direction. Les cibles sont réparties de manière régulière. τ^* est défini comme le temps typique mis pour aller d'une cible à la suivante :

$$\tau^* = \frac{L}{V\tau_2}\left(\tau_1 + \tau_2\right). \tag{26}$$

La distribution des temps de premier passage est en "marches d'escalier" si $L \gg V\tau_2, \sqrt{D\tau_1}$, sinon elle tend vers une exponentielle. Dans ce premier régime, pour $t \in [n\tau^*; (n+1)\tau^*]$ $(n \in \mathbb{N})$:

$$f(t) = \frac{p(1-p)^n}{\tau^*} = p(1-p)^n \frac{V\tau_2}{L\left(\tau_1 + \tau_2\right)}, \tag{27}$$

où p est la probabilité que le chercheur détecte la cible suivante. Quand $L \gg V\tau_2 \gg \sqrt{D\tau_1}$, p peut être estimée :

$$p = 2\frac{\sqrt{D\tau_1}}{V\tau_2}. \tag{28}$$

Ces deux expressions (27) et (28) sont en bon accord avec les simulations (voir Fig.17).

Avec la distribution complète des temps de premier passage, on peut obtenir tous les moments de la distribution :

$$\frac{\langle t^m \rangle}{(\tau^*)^m} = \frac{p^2}{m+1}\left[\frac{\partial}{\partial p}(p-1)\right]^{m+1}\frac{1}{p} = \frac{p^2}{m+1}\frac{\partial}{\partial p}(p-1)\frac{m}{p^2}\frac{\langle t^{m-1}\rangle}{(\tau^*)^{m-1}}. \tag{29}$$

En particulier, $\langle t\rangle/\tau^* = (2-p)/(2p)$, et $\langle t^2\rangle/\tau^{*2} = (6-6p+p^2)/(3p^2)$.

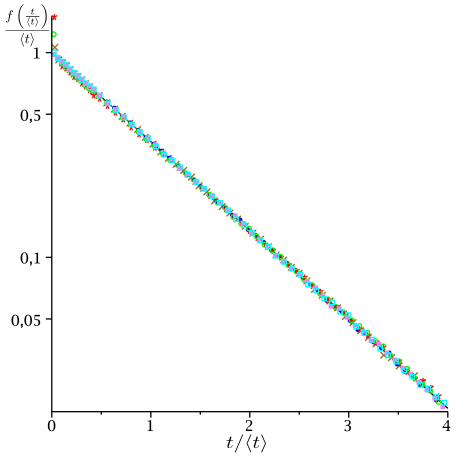
Cette distribution des temps de premier passage en marches d'escaliers est un cas très particulier : elle n'est obtenue qu'à une dimension, avec les phases balistiques toujours dans la même direction, et pour une répartition régulière de cibles. Dès qu'une contrainte est levée, la distribution tend vers une exponentielle, comme par exemple si les cibles sont distribuées de manière poissonnienne, ou encore quand les phases balistiques ne sont plus corrélées (voir Fig.18).

Que la distribution soit en escalier ou exponentielle, le temps moyen de recherche est bien représentatif du temps typique de recherche, puisqu'une seule échelle de temps est mise en jeu dans la distribution. Ces résultats sont importants car ils montrent que le premier moment du temps de recherche étudié tout au long de cette thèse est représentatif.

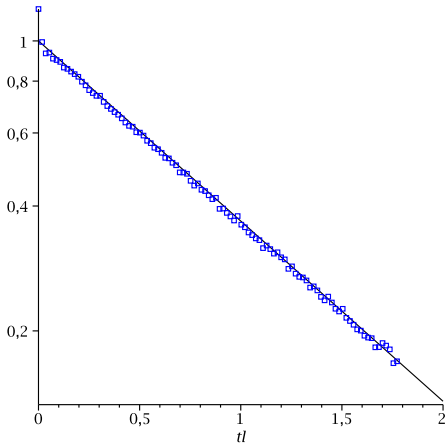
1.5.2 Mémoire temporelle

L'idée de stratégie de recherche intermittente présentée ici a suscité de l'intérêt dans la communauté scientifique, et elle a été reprise depuis dans d'autres travaux (voir partie 6.3, et partie 6.4 pour une liste qui complète ce qui suit).

Notre modèle générique suppose des distributions exponentielles des durées des phases. Une première possibilité d'extension est d'étudier l'effet de distributions non exponentielles, qui correspondent à des processus non markoviens, et donc à des chercheurs disposant d'une certaine mémoire temporelle. Par exemple, si le chercheur change de phase à τ_i exactement (temps d'attente déterministes), on peut montrer



Cas parfaitement corrélé (dimension effective 1), pour des cibles réparties aléatoirement.
$L = 100$: $\tau_1 = 1$, $\tau_2 = 1$ (⋆) ;
$\tau_1 = 10^{-2}$, $\tau_2 = 0,05$ (×).
$L = 10^4$: $\tau_1 = 10^4$, $\tau_2 = 10^3$ (○) ;
$\tau_1 = 10^2$, $\tau_2 = 32$ (□) ;
$\tau_1 = 1$, $\tau_2 = 1$ (●) ; $\tau_1 = 10^{-2}$, $\tau_2 = 0,05$ (■).



Cas sans corrélation à trois dimensions (mais le résultat est similaire à 2 dimensions) avec des cibles réparties régulièrement.
$a = 10$, $b = 100$, $\tau_1 = 6$, $\tau_2 = 11$.

FIG. 18: $f(t/\langle t\rangle)/\langle t\rangle)$ en fonction de $t/\langle t\rangle$. Distribution exponentielle (ligne noire), simulations (points). $D = 1$, $V = 1$.

que le temps moyen de détection de la cible diminue, mais reste du même ordre de grandeur (voir partie 6.3.1 pour plus de détails). Par ailleurs, Lomholt et al. [2008] ont montré qu'à une dimension (pour un mode de détection diffusif), si la distance entre cibles L est assez grande, et si les phases balistiques sont de durée distribuée selon une loi de Lévy ($p(t) \propto t^{-\alpha-1}$, avec $\alpha \in]1, 2[$) le temps moyen de détection de la cible obtenu est également plus petit que dans le cas exponentiel. Cependant, cette stratégie est moins efficace qu'une simple distribution de Lévy tronquée (voir partie 6.3.2).

1.5.3 Mémoire spatiale

La mémoire peut être temporelle, comme nous venons de le voir, mais elle peut aussi être spatiale. En effet, dans le modèle générique (partie 1.4), il n'y a aucune corrélation entre deux phases balistiques successives, c'est-à-dire aucune mémoire spatiale. Le cas inverse de corrélation parfaite (c'est-à-dire où la direction de la phase balistique est toujours la même) est traité partie 1.2. Comme de nombreuses trajectoires d'animaux présentent en fait des niveaux intermédiaires de corrélations [O'Brien et al., 1990], il est important de tenir compte de ces effets dans nos modèles. C'est aussi une question intéressante du point de vue théorique. En effet, si on prend par exemple le mode de détection statique à une dimension (le plus simple à traiter), et en définissant p comme la probabilité qu'une phase balistique soit dans la même direction que la phase balistique précédente, nous montrons que :

– Dans le cas d'une corrélation parfaite ($p = 1$), l'optimum est obtenu pour τ_2 tendant vers 0, et τ_1 tel que $\tau_1 = \alpha\tau_2$, où α ne dépend que de ak/V et a/b

(explicitement donné dans la partie 6.2) ;

– Dans le cas où il n'y a aucune corrélation ($p = 0,5$), le temps moyen de détection de la cible a un minimum global en fonction de τ_1 et τ_2, pour $\tau_1^{opt} = \sqrt{\frac{a}{Vk}}\left(\frac{b}{12a}\right)^{1/4}$ et $\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}$ (voir tableau 1).

Quand le niveau de corrélation est intermédiaire ($0,5 < p < 1$), le temps moyen de détection de la cible est-il minimisé pour des durées de phase finies (comme pour $p = 0,5$) ou tendant vers 0 (comme pour $p = 1$) ? Où se situe la transition ?

Ayant calculé le temps moyen de détection de la cible (dans la partie 6.2), nous montrons que dès que $p < 1$, les durées optimum des deux phases sont finies (même si elles peuvent être très petites).

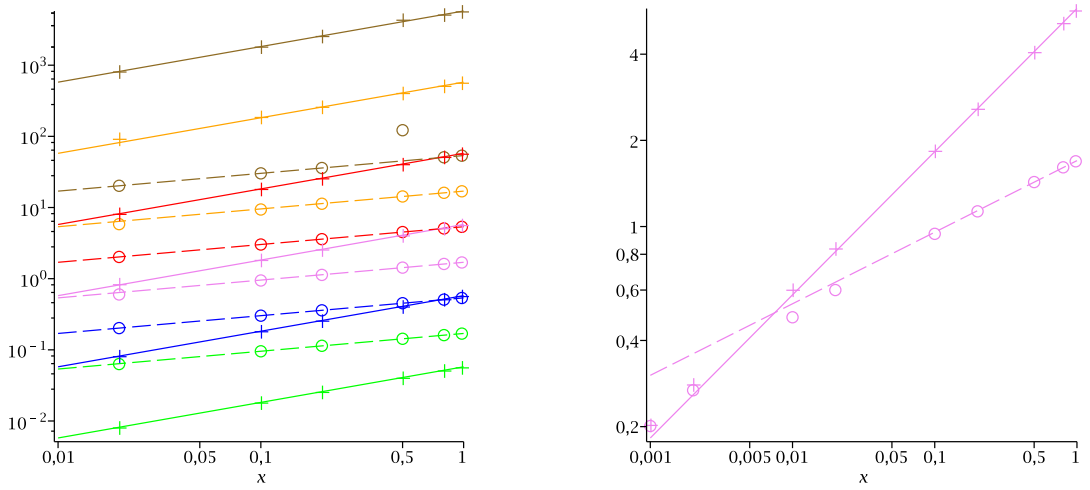


FIG. 19: τ_i^{opt} en fonction de $2(1-p)$. Prédiction analytique pour τ_1^{opt} (30) (pointillés) et prédiction analytique pour τ_2^{opt} (31) (lignes), comparées à τ_1^{opt} (∘) et τ_2^{opt} (□) obtenus par la minimisation numérique de l'expression exacte du temps moyen de détection de la cible (équation (334)). $a = 0,01$, $b = 1$ (vert), $a = 0,01$, $b = 100$ (bleu), $a = 1$, $b = 100$ (violet), $a = 1$, $b = 10^4$ (rouge), $a = 100$, $b = 10^4$ (orange), $a = 100$, $b = 10^6$ (marron). $V = 1$, $D = 1$.

En particulier, pour $(1-p) \gtrsim a/b$, le temps moyen de détection de la cible est minimisé pour des expressions de τ_i dans la continuité de celles du cas sans aucune corrélation (voir Fig.19) :

$$\tau_1^{opt} = \sqrt{\frac{a}{Vk}}\left(\frac{b}{a}\frac{1-p}{6}\right)^{1/4}, \tag{30}$$

$$\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{a}\frac{2(1-p)}{3}}. \tag{31}$$

Ces résultats, intéressants du point de vue théorique aussi bien que pour la modélisation de trajectoires animales, sont à étendre, au moins numériquement, aux autres modes de détection, et à 2 et 3 dimensions.

1.5.4 Comparaison avec des expériences

Une autre perspective est la comparaison avec des données expérimentales. Une première collaboration est en cours avec Vincent Fourcassié (CRCA, Université Paul

Sabatier, Toulouse) sur des trajectoires individuelles de fourmis. Une deuxième collaboration a été initiée avec Michel Sokolowski (chercheur en psychologie à l'université de Picardie - Jules Verne, Amiens), en vue d'étudier le comportement de recherche de cible cachée à l'échelle humaine.

La caméra, en hauteur, filme un sujet qui cherche des cibles dans l'arène.

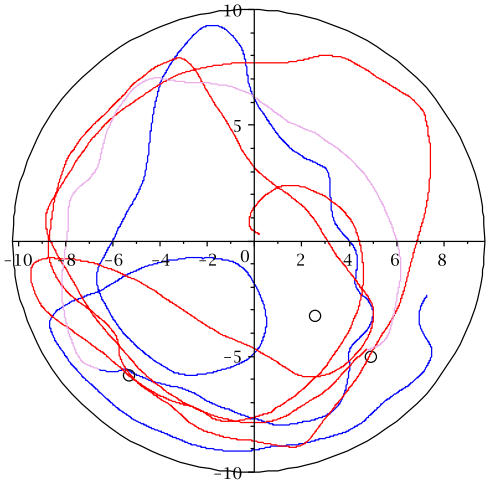
FIG. 20: Expérience de recherche de cible par des humains.

Le protocole est le suivant : un sujet doit trouver de petites cibles sur le sol d'une arène de 10 mètres de rayon, et les sujets les plus rapides obtiennent une récompense. Ils sont filmés (voir Fig.20), et les images sont ensuite traitées, jusqu'à obtenir les trajectoires qui sont ensuite analysées (pour plus de détails, se référer à la partie 6.6). Les résultats préliminaires peuvent être classés dans deux catégories.

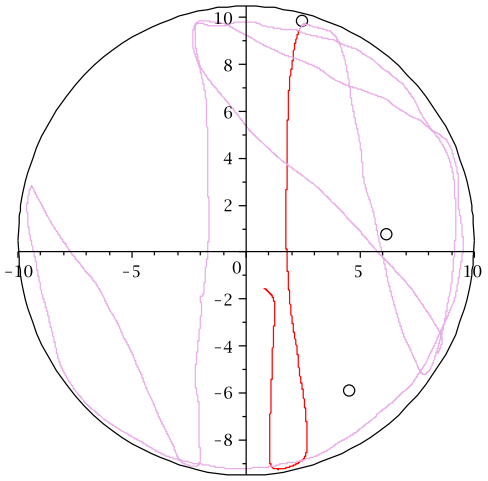
Si les sujets ont un champ de vision normal, ou s'ils ont un chapeau qui leur cache les repères extérieurs, mais qui leur permet cependant de voir jusqu'à un mètre devant eux, les sujets tentent de suivre des stratégies soit du type "spirale", soit du type "Pearson" (voir Fig.21). Ils essaient d'aller suffisamment lentement pour pouvoir détecter la cible tout au long de la trajectoire : ils n'adoptent pas de stratégies intermittentes.

En revanche, quand les yeux des sujets sont bandés, les trajectoires observées sont intermittentes, car se déplacer en même temps qu'explorer le sol est difficile (voir Fig.22). Les directions des phases de déplacement successives sont fortement corrélées.

Ces données sont en cours d'exploitation.

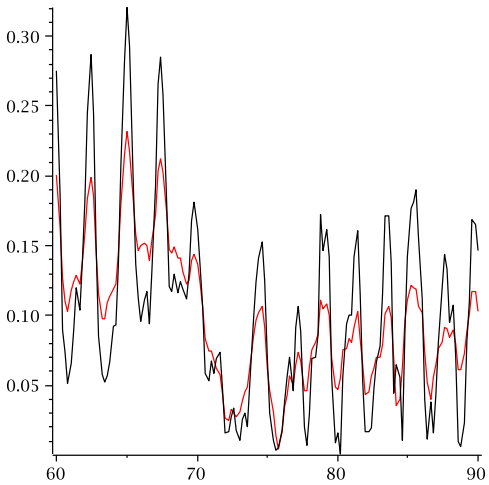


Trajectoire en spirale

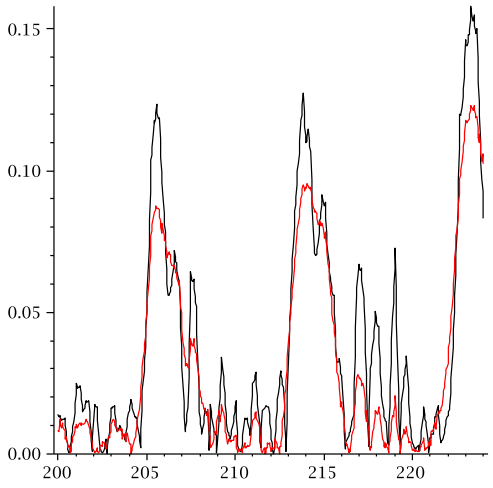


Trajectoire “Pearson”

FIG. 21: Exemples représentatifs de trajectoires de 7 minutes avec un champ de vision normal ou presque. Le grand cercle noir représente l’arène (10 mètres de rayon, l’échelle est en mètres). Les petits cercles noirs montrent où sont les cibles. les trajectoires sont les lignes colorées dont la couleur change quand une cible a été détectée.



Exploration rapide. Un genou est déplacé, une main explore la zone proche, puis symétriquement, et ainsi de suite.



Exploration lente. Les deux genoux sont avancés, puis la surface proche est explorée avec les deux mains, et ainsi de suite.

FIG. 22: Vitesse (m/s) en fonction du temps (s) sur des trajectoires où le sujet a les yeux bandés. Les lignes noires et rouges correspondent à des lissages plus ou moins forts de la trajectoire.

1.6 Conclusion

Les stratégies de recherche intermittentes reposent sur une hypothèse simple : le chercheur alterne entre deux phases de déplacement, l'une lente durant laquelle la cible peut être détectée, et l'autre rapide mais ne permettant pas la détection de la cible.

Au cours de cette thèse, nous avons montré que de telles stratégies de recherche sont effectivement observées à des échelles très variées. À l'échelle macroscopique, c'est notamment le cas de certains animaux à la recherche de nourriture. Motivés par cet exemple, nous avons proposé un modèle minimal bidimensionnel de stratégies de recherche intermittentes. Nous avons montré analytiquement que le temps moyen de recherche d'une cible peut être minimisé comme fonction des durées moyennes de chaque phase. Il existe ainsi une manière unique de répartir son temps entre les deux phases pour trouver une cible au plus vite. Cette intermittence constitue donc une stratégie de recherche optimale. En cela, ce modèle constitue une alternative au célèbre modèle des stratégies de Lévy qui, elles, ne sont optimales que dans des cas très spécifiques.

Ces stratégies de recherche intermittentes sont également observées à l'échelle microscopique. Un premier exemple est donné par la recherche par des protéines de cibles spécifiques sur l'ADN. Comme suggéré par Berg et al. [1981], les trajectoires réactionnelles combinent des phases de diffusion 1D le long de l'ADN et des phases de diffusion en volume. Ce mécanisme de recherche, dont l'efficacité a été discutée par différents groupes, est ainsi par essence intermittent. Dans cette thèse, nous avons calculé analytiquement la distribution de la distance parcourue le long de l'ADN lors d'une excursion 3D, qui joue un rôle important dans la détermination du temps de recherche. Nous l'avons adaptée aux spécificités d'une expérience à l'échelle d'une molécule unique, et avons montré que les trajectoires observées étaient bien décrites par une alternance de phase de diffusion à 1D et 3D. Un deuxième exemple de trajectoire intermittente à l'échelle microscopique est donné par le transport actif en milieu cellulaire, qui met en jeu des phases de diffusion thermique et des phases de déplacement balistique assurées par des moteurs moléculaires. Nous avons proposé un modèle général de réaction chimique du premier ordre limitée par ce type de transport. Sous l'hypothèse que la réaction n'est possible que dans la phase diffusive, nous avons montré qu'il était possible d'optimiser la constante de réaction. Le transport actif permet ainsi d'optimiser la cinétique réactionnelle en milieu cellulaire. Cet effet est particulièrement marqué pour des structures cellulaires de basse dimension, comme les membranes ou les structures tubulaires de type axone.

Ces stratégies intermittentes étant observées à des échelles variées, nous avons émis l'hypothèse qu'elles pourraient constituer un mécanisme de recherche générique. Nous avons ainsi, dans une dernière partie plus technique, étudié de manière systématique l'influence d'une part de la modélisation de la phase de détection des cibles, et d'autre part de la dimension de l'espace sur les temps moyens de premier passage de marches aléatoires intermittentes. Cette étude montre que l'optimalité des stratégies intermittentes est un résultat très largement robuste. Finalement, si les stratégies intermittentes sont si largement observées dans la nature, c'est probablement parce qu'elles constituent un mécanisme de recherche efficace.

2 Introduction

Contents

2.1 Foreword and outline 37
2.2 General framework 40
2.2.1 Effect of cues 40
2.2.2 Searching without cues : systematic *vs.* random strategies 40
2.2.3 Framework of this thesis 41

2.1 Foreword and outline

Figure 23: What is the best strategy to adopt when searching for a key?

What is the best strategy to find a missing object? If you have ever lost your keys, you have faced this problem (see figure 23). This every day life situation is a prototypical example of a search problem, which in its simplest form involves a searcher (either a person, an animal, or any kind of organism or particle) who is in general able to move across the search domain, and one or several targets. Even if very schematic, the search problem as stated turns out to be a quite universal question, which pops up at different scales and in various fields, and has generated an increasing number of works in recent years.

Theoretical studies of search strategies can be traced back to World War II, during which the US navy tried to most efficiently hunt for submarines and developed rationalized search procedures [Champagne et al., 2003]. Similar search algorithms have since been developed and utilized in the context of castaway rescue operations [Frost and Stone, 2001], or even for the recovery of Scorpion, a nuclear submarine lost near the Azores in 1968 [Richardson and Stone, 1971]. Another important and widely studied example of search processes at the macroscopic scale relates to animals searching for mates, food or a shelter [Charnov, 1976, O'Brien et al., 1990, Bell, 1991, Viswanathan et al., 1999, Bénichou et al., 2006, Shlesinger, 2006,

Edwards et al., 2007] which we discuss in more details in this thesis. At the microscopic scale, search processes naturally occur in the context of chemical reactions, for which the encounter of reactive molecules – or in other words the fact that one searcher molecule finds a reactive site – is a required first step. One should obviously mention the theory of diffusion-controlled reactions, initiated many years ago by the celebrated theory of von Smoluchowski [1917], developed by innumerable researchers (see for instance the review of Hanggi et al. [1990]). More recently, there has been a renewed interest in this field, in the context of biochemical reactions in cells where the number of reactive molecules is sometimes very small, making the first step – the search for a reaction partner – crucial for the kinetics. A prototypical example is the search for specific DNA sequences by transcription factors [Berg et al., 1981, Von Hippel, 2007, Gorman and Greene, 2008, Bonnet et al., 2008, Mirny, 2008].

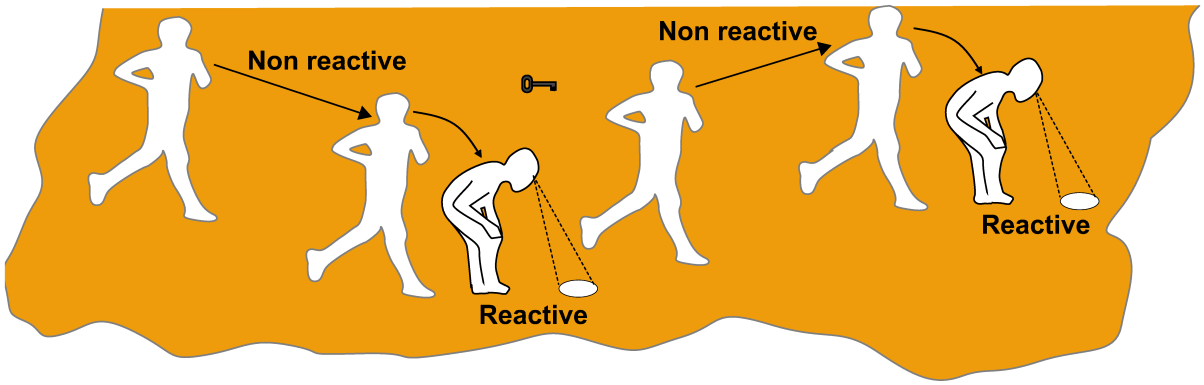


Figure 24: Intermittent search : the searcher either carefully inspects the surroundings, or runs to reach unvisited areas.

In all these examples, the time needed to discover a target is a limiting quantity, and consequently minimizing this search time often appears as essential. In order to get an intuition of what could be an efficient search strategy on general grounds, let us go back to the everyday-life example mentioned above. Imagine that you have lost a tiny object – let us say a key – in a large sandy beach, and that you have no *a priori* information on the position of the key. The key is so small that it cannot be detected if you pass by too fast. What is then the best strategy for finding the key as fast as possible? A first strategy consists in a slow and careful exploration (to make sure that the key will be detected upon encounter) of the sand all along the beach. In the case of a very large beach, the search time can then be very long. An alternative strategy you can think of consists in interrupting the slow and careful exploration of the sand by mere displacement phases, during which you run in order to relocate on the beach very fast, but without even trying to detect the key (see figure 24). We will call hereafter "intermittent search" such strategies that combine two distinct phases : a phase of slow displacement which enables target detection, and a phase of faster motion during which the target cannot be detected (note that Bartumeus [2009] has used the same word more recently, but with a broader definition, covering any intrinsic change of motion in the searcher's trajectory). The efficiency of such intermittent strategies results from a trade–off between speed and detection and can

be qualitatively discussed. Intuitively, the advantage of the fast relocation phases is to reach unvisited regions. The drawback however is that during these phases time is consumed without any chance of detecting the target. The net efficiency of this strategy is therefore not trivial, and the main goal of my thesis is to explicitly answer the following questions : (i) Can phases of fast motion which disable detection make the global search more efficient? (ii) If so, is there an optimal way for the searcher to share the time between the two phases? (iii) Are these intermittent search patterns relevant to the description of real situations?

Models of intermittent search strategies, as previously defined, have been introduced by my advisors in 2004. The first models of intermittent search that they have introduced relate to foraging animals at the macroscopic scale [Bénichou et al., 2005a], and to proteins searching for a target sequence on DNA at the microscopic scale [Coppey et al., 2004]. Both models are developed in one dimension, which is the relevant geometry to these real-life examples.

During my PhD, I have first extended these models, in particular to displacements in 2 and 3 dimensions, which cover a much broader range of real situations. I have also studied the robustness of these search strategies, and have come to the conclusion that they could constitute genuine efficient search strategies. This thesis consists on the one hand of modeling aspects in close connection with experimental studies (section 3 and 4), and on the other hand of theoretical developments and more technical results on first passage time for intermittent random walks (see section 5).

More precisely, for the case of foraging animals trajectories, I have developed a model of bidimensional intermittent search strategies, which appears as an alternative to a famous model, named "Lévy strategies", very often used in this context. These macroscopic situations are the object of section 3. I have also studied two examples of search problems at the microscopic scale (section 4). The first relates to proteins searching for a specific sequence on DNA. Since generally these proteins alternate between 1-dimensional phases of motion along DNA and 3D excursions in the bulk, this search process can be qualified as intermittent, in the general meaning previously defined. During my PhD, I have extended the existing models by calculating analytically the distribution of the length traveled along the DNA during a 3D excursion. I have then adapted this calculation to the quantitative interpretation of a single-molecule experiment, developed by the group of Pierre Desbiolles (LKB-ENS). The second microscopic example I have studied relates to vesicle transport in cellular medium. Vesicles can either diffuse passively, or bind to molecular motors performing an active ballistic motion. I have proposed an analytical model of reactions limited by transport in such active media and shown quantitatively how active transport can speed up reaction kinetics. Given that intermittent search strategies are observed at various scales, I have suggested that they could constitute a generic search mechanism. In the more technical section 5, I study the robustness of intermittent random walks, in a quite general theoretical framework. In the final section (section 6) gathers extensions and perspectives related to the previous models, some of them being currently under investigation.

2.2 General framework

The search problem can take multiple forms ; in this section we define more precisely the framework of this thesis – namely the random intermittent search strategies – and introduce the main hypothesis which will be made.

2.2.1 Effect of cues

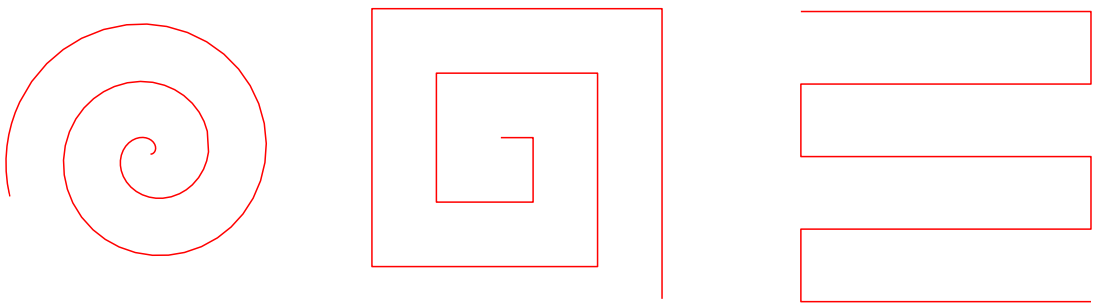
Although in essence in a search problem the target location is unknown and cannot be found from a rapid inspection of the search domain, in practical cases there are often cues which restrict the territory to explore, or give indications on how to explore it. We can quote the classic example of chemotaxis [Berg, 2004], which keeps arising interest in the biological and physical communities (see for example Yuzbasyan et al. [2003], Kafri and Da Silveira [2008]). Bacteria like E.coli swim with a succession of "runs" (approximately straight moves) and "tumbles" (random changes of direction). When they sense a gradient of chemical concentration, they swim up or down the gradient by adjusting their tumbling rate : when the environment is improving, they tumble less, whereas when the environment worsens, they tumble more, which results in a bias towards favorable regions. Chemical cues can be as varied as salts, glucose, amino-acids, oxygen, etc., but this behavior can also be triggered by other kinds of gradients : temperature gradients [Maeda et al., 1976, Salman et al., 2006, Salman and Libchaber, 2007], light intensity [Sprenger et al., 1993], etc. Chemotactic search requires a well defined gradient of chemoattractant, and is therefore applicable only when the concentration of cues is sufficient. On the contrary at low concentrations cues can be sparse, or even discrete signals which do not allow a gradient based strategy. For example, this is the case for animals sensing odors in air or water where the mixing due to the flow turbulence distribute odors in random and disconnected patches of high concentration. Vergassola et al. [2007] proposed a search algorithm, which they called "infotaxis", designed to work when cues are sporadic and information partial. It locally maximizes the expected rate of information gain. Among the trajectories produced by infotaxis, there are "zigzagging" and "casting" paths similar to those observed in the flight of moths.

In this thesis we focus on the case where no cue is present, or before finding the first cue. This assumption applies to targets are qualitatively "hard to find", that is targets that can be detected only if the searcher is within a given detection radius a that is much smaller than the typical extension of the search domain. In particular this assumption clearly covers the case of search problems at the scale of chemical reactions, and more generally the case of searchers whose motion is independent of any exterior cue that could be emitted by the target.

2.2.2 Searching without cues : systematic *vs.* random strategies

Whatever the scale, the behavior of the searcher strongly relies on his ability, or inability, to keep memories of his past explorations. Depending on the searcher and on the space to explore, there is more or less capacity for spatial memory.

In an extreme case, the searcher has a mental map of the exploration space and thus performs a systematic search. Figure 25 presents several systematic patterns



Spiral Expanding square Lawn-mower

Figure 25: Examples of patterns for systematic exploration of space.

which are designed to avoid oversampling : spiral, expanding square and lawn-mower (for more patterns, see for example Champagne et al. [2003]).

In many search problems however the searcher has no memory of his past exploration. It does not mean than there is no strategy to adopt. For example, when you are searching for a key, if you do not move, search will be unsuccessful, and the rule "walk straight ahead" would be a better strategy for you. In particular, rules can rely on stochastic processes, *i.e.* the sampling of probability distributions. This is obviously the case for "molecular" searchers at the microscopic scale, but also at larger scales of animals with low memory skills. Search trajectories can in this case be qualified as random, and the theory of stochastic processes provides powerful tools for their quantitative analysis. This thesis is focused on random search problems.

2.2.3 Framework of this thesis

To summarize, in this thesis we shall focus on intermittent search strategies for targets which emit no cue. The searchers will be assumed to have no (or low) memory skills, resulting in their trajectories being random walks. To assess the efficiency of such strategies from a kinetic point of view, we shall mainly calculate the mean first passage time to a target, and study the minimization of this quantity. And when possible, we shall investigate the complete distribution of first passage time, which provides more information (see section 6.1).

3 Intermittent search strategies at the macroscopic scale

Contents

3.1 Lévy strategies 44

3.2 A simple model based on intermittence 45

3.2.1 Model 45

3.2.2 Results 46

3.3 Influence of the target distribution on the search time . 47

3.3.1 How real targets are distributed? 47

3.3.2 Analytical results in the case of a Poissonian distribution of targets 48

3.3.3 Simulations in the case of Poissonian targets 49

3.3.4 Conclusion 51

3.4 Minimal model of intermittent search in two dimensions 52

3.4.1 Motivation 52

3.4.2 Model 52

3.4.3 Methods 53

3.4.4 Results for the diffusive mode 54

3.4.5 Results for the static mode 55

3.4.6 Conclusion 56

3.5 Are Lévy strategies really so advantageous? 57

3.5.1 The albatross story 57

3.5.2 Optimizing the encounter rate with targets with Lévy walks : how and when? 57

3.5.3 Do animals really perform Lévy walks? 58

3.6 Conclusion on animal foraging 59

The case of foraging animals is an interesting example of search at the macroscopic scale. There is a vast literature on animal displacements. Trajectories of numerous animals have been recorded, in particular when they are in search for food, shelter or mate. The observed search trajectories are often described as a sequence of ballistic segments (or relatively straight motion), interspersed by much slower phases, or even stops [O'Brien et al., 1990, Bell, 1991]. We first review two models that have been proposed to interpret the experimental data : Lévy walks and a simple model based on intermittence. For this thesis project, we have extended this last model in two directions. (i) We have studied the influence of the target distribution. (ii) More importantly, we have extended this model to two dimensions. After presenting these extensions, we discuss the relevance of Lévy walks and of these models based on intermittence for the description of animal trajectories.

3.1 Lévy strategies

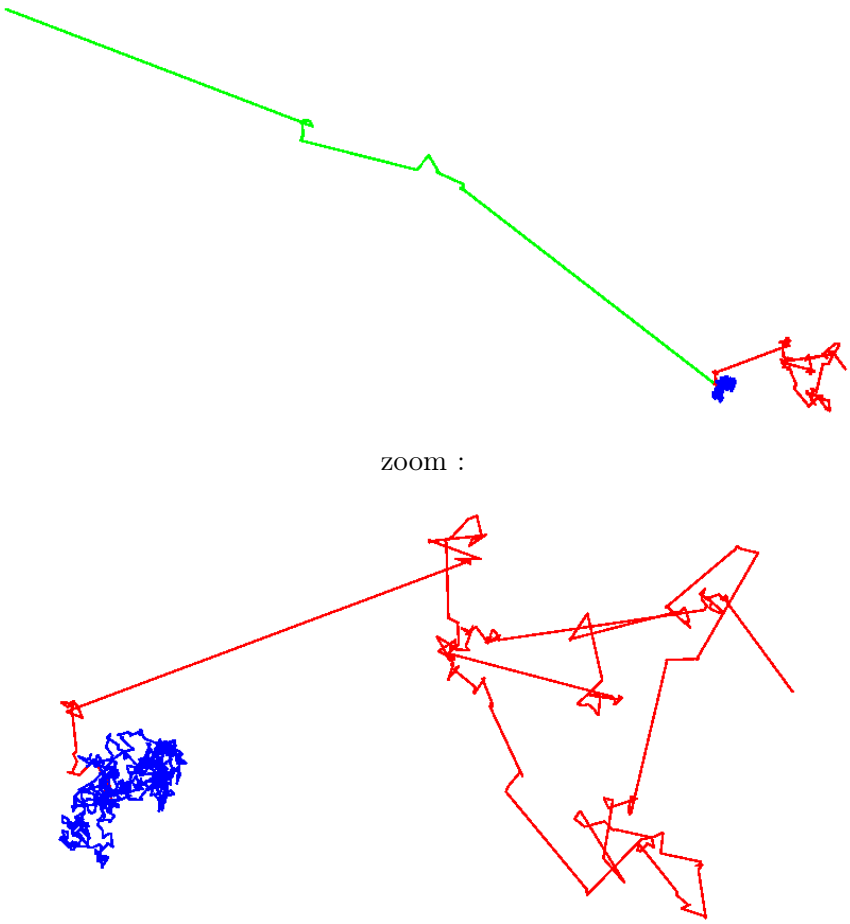


Figure 26: Example of Lévy walks, with $\mu = 1.5$ (green, not present on the zoom), $\mu = 2$ (red), $\mu = 3$ (blue). The total path length is the same for the 3 examples.

The ballistic phases interspersed with turns of animal trajectories mentioned above have often been interpreted as Lévy walks [Viswanathan et al., 1999, 2008]. Shlesinger and Klafter [1986] were the first to report that, due to their weak oversampling properties (see figure 26), Lévy walks could be an efficient way to explore space and could be used to model trajectories of foraging animals in particular. This observation led Viswanathan et al. [1999] to propose the following Lévy search model, in the presence of fixed targets randomly and sparsely distributed : they consider a searcher performing ballistic flights at constant speed, and detecting targets closer than r_v. A target is found when the searcher encounters it for the first time. The flight lengths are drawn from a Lévy distribution $p(l) \propto l^{-\mu}$, with $1 < \mu < 3$. For $\mu = 1$, the probability distribution is not defined. For $1 < \mu \leq 2$, the distribution has no mean and no variance. For $2 < \mu < 3$, the distribution has a mean but no variance. For $\mu \geq 3$, the distribution has both a mean and a variance, and thus obeys the central limit theorem : after sufficiently many flights, the probability distribution of the distance covered from the starting point is a Gaussian whose

variance scales linearly with time, as in the case of diffusion. They found that in the case of revisitable targets, $\mu^{opt} \simeq 2$, whereas in the case of non revisitable targets (that are destructed when encountered) $\mu^{opt} \to 1$, which in fact is simply a ballistic motion without reorientations.

3.2 A simple model based on intermittence

In the Lévy walks model, the searcher is assumed to be able to detect its target all along its trajectory. However, as it was the case with the example of the lost key, it is evident that in some situations speed degrades perception. In some studies on animal behavior, the reported slow phases are clearly aimed at sensing the environment in order to detect the targets [Kramer and McLaughlin, 2001]. Is this alternation of slow and fast phases beneficial for the search? To answer this question, a simple model relying on intermittence was proposed in 2005 [Bénichou et al., 2005a,c]. We first review how this model is constructed and its results, before extending it.

3.2.1 Model

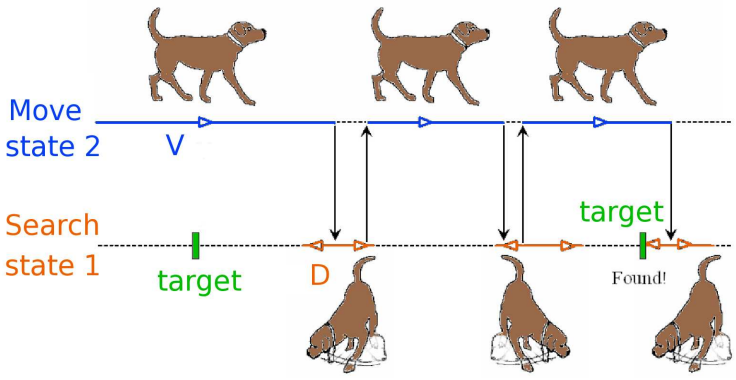


Figure 27: Simple model for intermittent search.

The central point of this schematic model is that it relies on the explicit description of search trajectories as intermittent. The searcher is assumed to display alternatively two distinct attitudes (see figure 27) :

- a scanning phase, named phase 1, during which the sensory organs of the searcher explore its immediate vicinity. This phase is modeled as a "slow" diffusive movement (a continuous random walk with diffusion coefficient D). The target is found when this movement reaches the target location for the first time. As focusing and processing the information received by sensory organs require a minimum time, the scan phase cannot be too short, which implies a minimal mean time spent in this phase, noted τ_1^{min}.

- a motion phase, named phase 2, during which the searcher moves "fast" and is unable to detect the target. These repositioning moves are characterized by a ballistic motion (at constant velocity V). In the case of animals, there

are usually correlations in the angles between two successive ballistic phases [O'Brien et al., 1990]. Here the case studied is the limit of high correlations, making the problem effectively 1-dimensional for both phases, with the velocity in phase 2 always in the same direction.

Next, the searcher is assumed to switch randomly from phase 1 (respectively, 2) to phase 2 (respectively, 1) with a fixed rate per unit time λ_1 (respectively, λ_2), which corresponds to a searcher without time memory. It leads to a mean duration $\tau_i = 1/\lambda_i$ of phase i, and the phase durations are distributed exponentially, as observed in numerous experimental studies (see for example Fujiwara et al. [2002], Pierce-Shimonura et al. [1999], Hill et al. [2000], Li et al. [2008], but this list is far from exhaustive). Preys are assumed to be immobile. The assumed geometry is one centered target in a domain of length L with reflective boundaries. This geometry also mimics an array of regularly spaced targets on an infinite segment.

3.2.2 Results

The mean first passage time to the target, uniformly averaged over the starting position of the searcher, is [Bénichou et al., 2005a] :

$$\langle t \rangle = (\tau_2 + \tau_1) \left(\frac{L}{2} \frac{\left(\mathrm{e}^{\alpha+\beta} - 1\right) \sqrt{1+4r} + (1+2r)\left(\mathrm{e}^{\beta} - \mathrm{e}^{\alpha}\right)}{\sqrt{1+4r}\,(\mathrm{e}^{\beta}-1)\,(\mathrm{e}^{\alpha}-1)\,\tau_2 V} - \frac{1}{r} - 1 \right), \tag{32}$$

with :

$$r = \tau_2^2 V^2/(D\tau_1), \tag{33}$$

$$\alpha = \frac{L}{2}\left(-\frac{1}{\tau_2 V} + \sqrt{\frac{1}{{\tau_2}^2 V^2} + 4\,\frac{1}{D\,\tau_1}} \right), \tag{34}$$

$$\beta = -\frac{L}{2}\left(\frac{1}{\tau_2 V} + \sqrt{\frac{1}{{\tau_2}^2 V^2} + 4\,\frac{1}{D\,\tau_1}} \right). \tag{35}$$

In the limit of $L \gg V\tau_2, \sqrt{D\tau_1}, D\tau_1/(V\tau_2)$, this expression simplifies :

$$\langle t \rangle \simeq \frac{L\,(\tau_2+\tau_1)\,(D\,\tau_1 + 2\,{\tau_2}^2 V^2)}{2\tau_2\,V\,\sqrt{D\,\tau_1}\sqrt{D\,\tau_1 + 4\,{\tau_2}^2 V^2}}. \tag{36}$$

Note that thanks to intermittence, $\langle t \rangle \propto L$, whereas for diffusion alone the mean detection time is $\tau_{\mathrm{diff}} \propto L^2$. Intermittence is thus favorable, at least for L large enough. The gain of intermittence compared to diffusion alone is :

$$gain = \frac{\tau_{\mathrm{diff}}}{\langle t \rangle} = \frac{L\left(\tau_2\,V\,\sqrt{D\,\tau_1}\sqrt{D\,\tau_1 + 4\,{\tau_2}^2 V^2}\right)}{6D(\tau_1+\tau_2)\,(D\,\tau_1 + 2\,{\tau_2}^2 V^2)}. \tag{37}$$

Intermittence is favorable in particular when τ_1 and τ_2 are taken as optimal. The mean search time is minimized for $\tau_1^{opt} = \tau_1^{min}$, and τ_2^{opt} satisfying the relation :

$$\tau_1^3 + 6\frac{\tau_1^2\tau_2^2}{\tau} - 8\frac{\tau_2^5}{\tau^2} = 0, \tag{38}$$

where $\tau = D/V^2$ is an extra characteristic time, depending on the searcher's characteristics. This minimum takes a simple form in two different regimes.

- If $\tau_1 \gg \tau$, the minimum of the search time is for $\tau_1 = \tau_1^{min}$, and :

$$\tau_2^{opt} = \left(\frac{3\tau\tau_1^2}{4}\right)^{1/3}. \tag{39}$$

  In this regime, $\tau_1 > \tau_2$: the searcher spends more time scanning than moving.

- If $\tau_1 \ll \tau$, the minimum of the search time is for $\tau_1 = \tau_1^{min}$, and :

$$\tau_2^{opt} = \left(\frac{\tau^2\tau_1^3}{8}\right)^{1/5}. \tag{40}$$

  In this regime, $\tau_1 < \tau_2$, which means that the searcher spends more time moving.

These results have been compared to experimental data from O'Brien et al. [1990] and Kramer and McLaughlin [2001], who provide the average duration of detection and ballistic phases, characterizing the saltatory behavior of 18 different species, as various as planktivorous fish, ground foraging birds, or lizards. The optimal strategy obtained above has been shown to account reasonably well for these data.

Now that this first simple model has been defined, we present two extensions that we have considered for this thesis project. On the one hand, we study the influence of the target distribution on the search time. On the other hand, we extend this model to a two-dimensional space, which is relevant in many search problems at the macroscopic scale.

3.3 Influence of the target distribution on the search time

We first study the influence of target distribution on the previous results.

3.3.1 How real targets are distributed?

Target distributions are often described as regular, random or patched [Bell, 1991] (see figure 28). In the model presented above, the chosen geometry can be interpreted as one target in a finite domain, or as an infinite array of regularly spaced targets. The regular distribution is representative of the real-life case of targets that repel each other, thus being as far as they can from each other. This distribution is also a mean-field approximation of other distributions. As the regular distribution has already been studied, let us discuss the other representative distributions.

If targets are in patches, when a target is found it is likely that there are other targets in the immediate surroundings. Thus a simple strategy is to switch behavior when a target is encountered, as proposed for example by Benhamou [1992]. The search is then in two steps : finding a patch, and exploiting it. For the first step, if patches are regularly spaced, previous results are still valid, except for the density of targets which has to be replaced by the density of patches. The patches could also be randomly located.

The last case of Poissonian targets corresponds to situations of non-interacting targets. The following results have been published in [Moreau et al., 2007b, 2009].

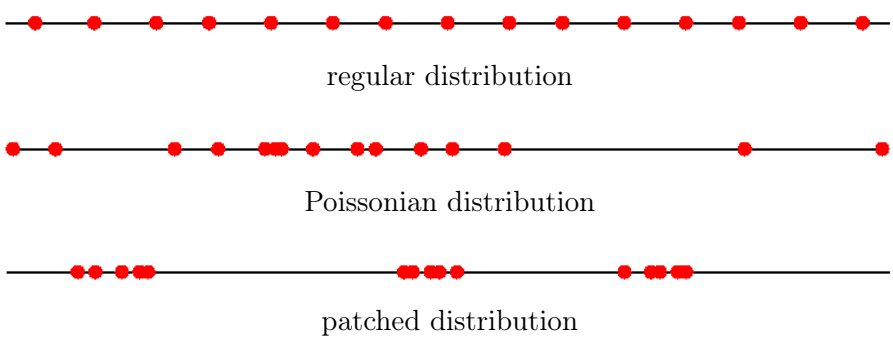


Figure 28: Examples of target distributions.

3.3.2 Analytical results in the case of a Poissonian distribution of targets

In the case of a Poissonian distribution of targets, the distance between two consecutive targets is exponentially distributed. Except for this change, the other parameters remain as defined in the previous model (see section 3.2).

The mean search time is in general hard to calculate for Poissonian targets, which can be seen as frozen disorder. However, estimates (for $L \gg D/V$) can be given in 3 regimes (see Moreau et al. [2007b, 2009] for details) :

- In the large ballistic displacements limit (when $V\tau_2 \gg \sqrt{D\tau_1}$), two successive diffusive phases can be considered as non-overlapping. It can be shown that in this regime :
$$\langle t \rangle \simeq L\frac{\tau_1 + \tau_2}{2\sqrt{D\tau_1}}. \tag{41}$$

- In the small ballistic displacements limit (when $V\tau_2 \ll \sqrt{D\tau_1}$), it can be considered that successive diffusive phases considerably overlap. It leads to :
$$\langle t \rangle \simeq L\frac{\tau_1 + \tau_2}{V\tau_2}. \tag{42}$$

- The most interesting situation is the intermediary regime. Indeed, in the first case (large ballistic displacements), relocations are too long and overshoot the target; and in the second case (small ballistic displacements), there are often repetitive scans of the same areas. In the intermediary regime, the mean first passage time to the target can be approximated by :
$$\langle t \rangle \simeq L\frac{\tau_1 + \tau_2}{V\tau_2}\frac{(1+\theta)^2(1+\epsilon\theta)}{(1+4\theta+2\epsilon\theta^2)}, \tag{43}$$
with $\theta = V\tau_2/\sqrt{D\tau_1}$ and $\epsilon = \sqrt{D\tau_1}/L$.

This last regime enables us to discuss the efficiency of the intermittent search. We discuss the efficiency comparing $\langle t \rangle^{opt}$, the mean search time with intermittence at the minimum with $\tau_{\text{diff}} = L^2/(2D)$, the mean search time with diffusion alone. We define $\tau_{bal} = L/V$ as the typical time scale needed to travel in the ballistic mode

the distance between two consecutive targets. Intermittence is found to decrease the search time. To minimize the mean search time, τ_1 should be taken as small as possible. The optimization with respect to τ_2 leads to two regimes, depending of the value of τ_1 compared to the previously introduced timescale $\tau = D/V^2$, characteristic of the searcher. :

- When $\tau_1 \gg \tau$:

$$\frac{\tau_2^{opt}}{\tau} \sim \sqrt{\frac{7}{4}} \left(\frac{\tau_1}{\tau}\right)^{3/4}. \tag{44}$$

At the optimum, the mean search time is :

$$\langle t \rangle^{opt} \sim \frac{L}{2V}\sqrt{\frac{\tau_1}{\tau}}, \tag{45}$$

and the gain is :

$$G \sim \frac{L}{\sqrt{D\tau_1}} = \sqrt{\frac{2\tau_{\text{diff}}}{\tau_1}}. \tag{46}$$

As we shall see in the following, in this regime the approximations are very accurate.

- When $\tau_1 \ll \tau$:

$$\frac{\tau_2^{opt}}{\tau} \sim \frac{1}{2}\sqrt{\frac{\tau_1}{\tau}}. \tag{47}$$

At the optimum, the mean search time is :

$$\langle t \rangle^{opt} \sim \frac{3L}{4V} = \frac{3}{4}\tau_{bal}, \tag{48}$$

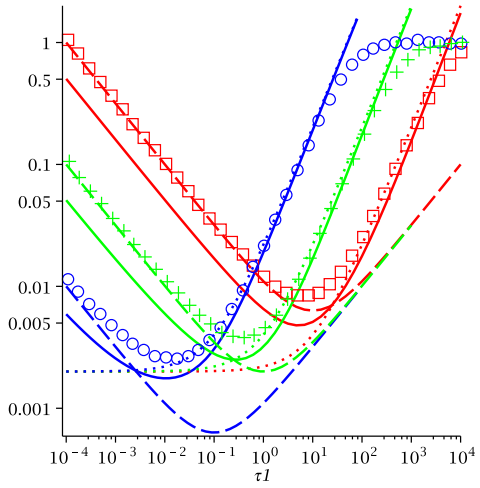
and the gain is :

$$G \sim \frac{2LV}{3D} = \frac{4\tau_{\text{diff}}}{3\tau_{bal}}. \tag{49}$$

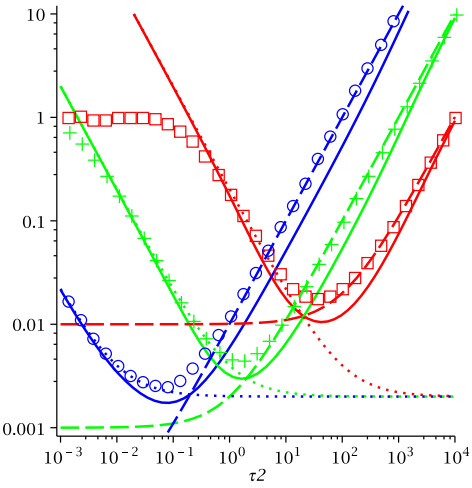
As we shall see in the following, the approximations are qualitatively good in this regime, but not as precise as in the other regime. Indeed, the gain obtained here would mean that the mean first passage time to the target is smaller than τ_{bal}, which is the minimal mean time to travel to the target (except if $\tau_{bal} > \tau_{\text{diff}}$). In fact, as can be seen in figure 31, simulations show that $\langle t \rangle^{opt} \to \tau_{bal}$. It means that very fast intermittence enables the searcher to retain the best of the two phases : reactivity of phase 1 and motion of phase 2.

3.3.3 Simulations in the case of Poissonian targets

Simulations are needed to check the approximations. Generating an array of randomly spaced targets is very easy. It is enough to know the position of the nearest target on the left and of the nearest target on the right. When the searcher is in the diffusive mode when it passes on one of these two targets, the search is over. When the searcher is in the ballistic mode, it can only go past the right target, as its motion is always in the same direction. When the searcher moves past its right



$ln(\langle T\rangle/T_{\text{diff}})$ as a function of $ln(\tau_1)$. $\tau_2 = 10$ (red, □), $\tau_2 = 1$ (green, +), $\tau_2 = 0.1$ (blue, ∘).



$ln(\langle T\rangle/T_{\text{diff}})$ as a function of $ln(\tau_2)$. $\tau_1 = 100$ (red, □), $\tau_1 = 1$ (green, +), $\tau_1 = 0.01$ (blue, ∘).

Figure 29: Validity of the approximations. Mean first passage time to the target, renormalized by the mean first passage time without intermittence. Small ballistic displacements approximation (42) (dashed line). Large ballistic displacements approximation (41) (dotted line). Intermediary approximation (43) (line). Numerical simulations (symbols). $D = 1$, $V = 1$, $L = 10^3$.

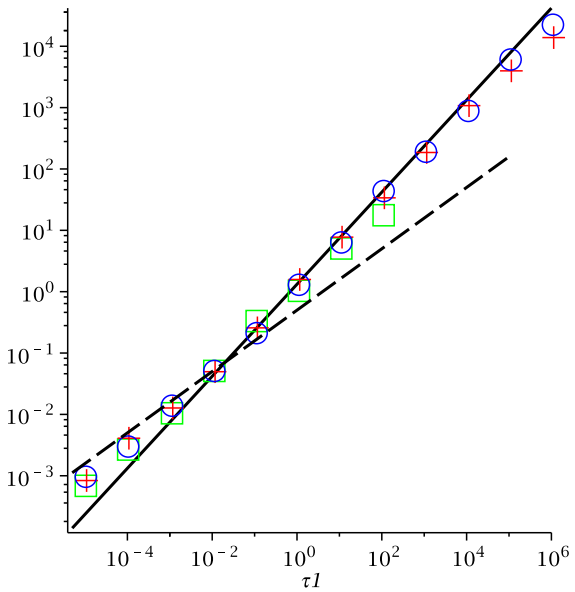


Figure 30: $ln(\tau_2^{opt})$ as a function of $ln(\tau_1)$. Small τ_1 analytical prediction (47) (dashed black line). Large τ_1 analytical prediction (44) (solid black line). Numerical values (symbols), for $L = 10$ (green □), $L = 10^3$ (red +), $L = 10^5$ (blue ∘). $D = 1$, $V = 1$.

target, it is sufficient to randomize a new target, the right target becoming the left target and the new target becoming the target on the right. These steps can be repeated until the new position after a ballistic phase is between the left and right

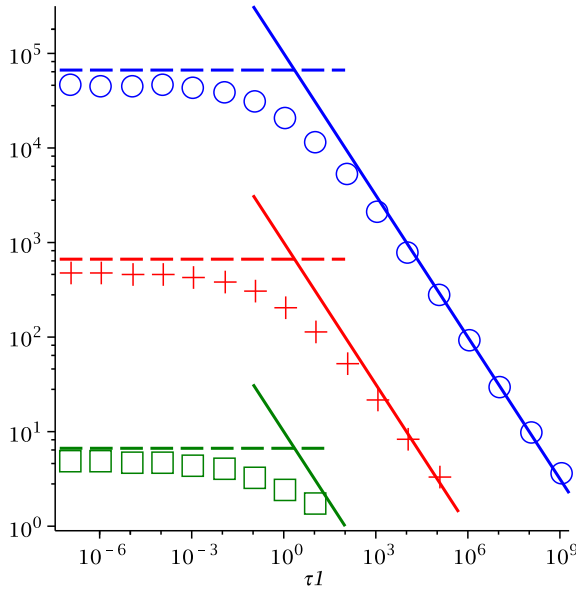


Figure 31: $ln(G)$ as a function of $ln(\tau_1)$ (τ_2 taken optimal). Small τ_1 analytical prediction (49) (dotted line). Large τ_1 analytical prediction (46) (line). Numerical simulations (points). $L = 10^1$ (green, $\square$), $L = 10^3$ (red, $+$), $L = 10^5$ (blue, $\circ$). $D = 1$, $V = 1$.

targets. Only two positions are to be kept in memory. To gain time without losing accuracy, the length of diffusion steps is tuned depending on the distance to the target, as proposed by Berezhkovskii et al. [1998].

Figure 29 represents the mean search time $\langle t \rangle$ as a function of τ_1 and τ_2 for typical values of the other parameters. They allow comparing the numerical results with the approximations (41) and (42), and with the intermediary approximation (43). It shows that the approximations of large and small ballistic displacements are valid in the expected conditions, and the intermediary approximation (43) correctly reproduces the existence and the position of the minimum of $\langle t \rangle$. Figure 30 supports the scaling laws relating τ_1 and the corresponding optimal waiting time τ_2 at the optimum. The exponent $3/4$ of the theoretical scaling law (44) for $\tau \ll \tau_1$ is very well confirmed by the simulations. This is not the case for the law (47) for $\tau \gg \tau_1$, which indicates that the approximations should be handled with care for short waiting times τ_1, τ_2, although their results are qualitatively correct. Figure 31 shows the gain as a function of τ_1 in different possible conditions. It supports the conclusions of the theoretical study, and indeed confirms that the gain due to intermittence can be very important if $\tau_{\text{diff}} \gg \tau_{bal}$.

3.3.4 Conclusion

In the case of a Poissonian distribution of targets, we have shown that the intermittence remains valid as a strategy minimizing the search time. The optimal strategy still consists in taking τ_1 as small as possible. However, τ_2^{opt} is different from the case of regularly spaced targets. The optimal mean duration of ballistic flights scales as $\sqrt{\frac{7}{4}}(\tau_1^3 \tau)^{1/4}$ in the limit $\tau_1 \gg \tau = D/V^2$. In this regime, at the opti-

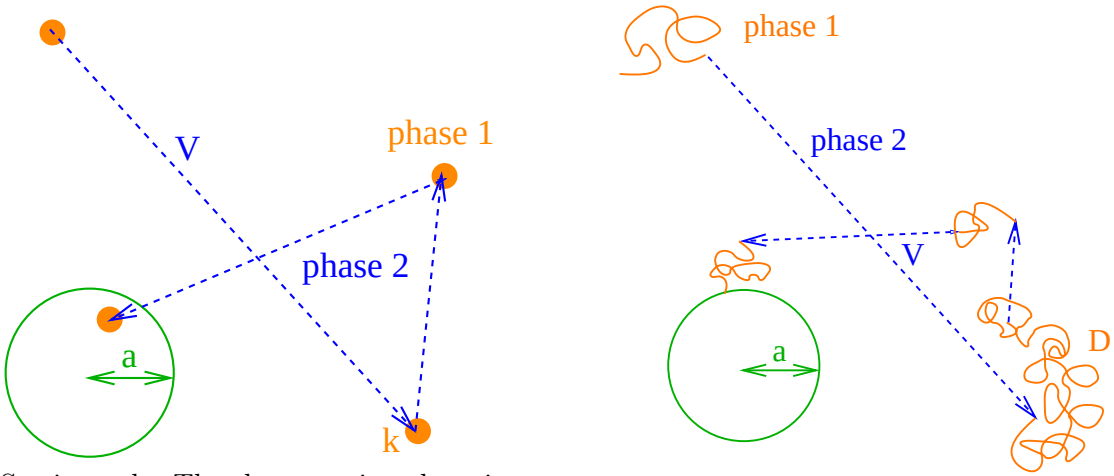
mum, $\langle t \rangle \simeq \frac{1}{2}\frac{L}{V}\sqrt{\frac{\tau_1}{\tau_2}}$, with a gain compared to diffusion alone $\propto L/\sqrt{D\tau_1}$. As τ_2^{opt} scales differently with τ_1 (exponent 3/4 instead of 2/3 when targets are regularly distributed), if accurate experimental data were available, the two cases could be distinguished.

3.4 Minimal model of intermittent search in two dimensions

3.4.1 Motivation

The models of intermittent search presented previously are one-dimensional, with ballistic phases infinitely correlated, in the sense that the direction taken is always the same. Here we develop a model of intermittent search strategies in two dimensions without correlation, which encompasses a much broader field of applications, in particular for animal or human searchers (published in [Bénichou et al., 2006], completed in [Bénichou et al., 2007] and highlighted by Shlesinger [2006]). On the basis of this model with minimal ingredients, we show that bidimensional intermittent search strategies *do optimize* the search time for non revisitable targets, *i.e.* targets that are destructed when found. We explicitly determine the optimal way to share the time between the phases of non reactive displacement and of reactive search. Our approach relies on an approximate analytical solution based on a decoupling hypothesis, which proves to reproduce quantitatively our numerical simulations over a wide range of parameters. The main results of this model are discussed in this section, while its technical analysis is left for section 5 – since this model appears as a special case of a general model presented in detail in section 5.

3.4.2 Model



Static mode. The slow reactive phase is static and detection takes place with a finite rate k.

Diffusive mode. The slow reactive phase is diffusive and detection is infinitely efficient.

Figure 32: Two models of intermittent search: The searcher alternates between slow reactive phases (regime 1) of mean duration τ_1, and fast non reactive ballistic phases (regime 2) of mean duration τ_2.

Following the previous model, we consider a two-state searcher (see figure 32) of position $\mathbf{r}$ that performs slow reactive phases (denoted 1), randomly interrupted by fast relocating ballistic flights of constant velocity V and random direction (phases 2). We assume the duration of each phase i to be exponentially distributed with mean τ_i. As fast motion usually strongly degrades perception abilities [O'Brien et al., 1990, Kramer and McLaughlin, 2001], we consider that the searcher is able to find a target only during reactive phases 1. The detection phase involves complex biological processes that we do not aim at modeling accurately here. However, we put forward two modes of detection. The first one, referred to in the following as the "diffusive mode", corresponds to a diffusive modeling (with diffusion coefficient D) of the search phase like in the previous one-dimensional model in agreement with observations for vision [Huey, 1968], tactile sense or olfaction [Bell, 1991]. The detection is assumed to be infinitely efficient in this mode : a target is found as soon as the searcher-target distance is smaller than the reaction radius a. On the contrary, in the second mode, denoted as the "static mode", the reaction takes place with a finite rate k, but the searcher is immobile during search phases. Note that this description is commonly adopted in reaction-diffusion systems [Rice, 1985] or operational research [Frost and Stone, 2001]. A more realistic description is obtained by combining both modes and considering a diffusive searcher with diffusion coefficient D and finite reaction rate k. In order to reduce the number of parameters and to extract the main features of each mode, we study them separately by taking successively the limits $k \to \infty$ and $D \to 0$ of this general case. More precisely, in these two limiting cases, we address the following questions : what is the mean time it takes the searcher to find a target? Can this search time be minimized? And if so for which values of the average durations τ_i of each phase?

3.4.3 Methods

We now present the basic equations combining the two search modes introduced above in the case of a point-like target centered in a spherical domain of radius b with reflexive boundary. Note that this geometry mimics both relevant situations of a single target and of infinitely many regularly spaced non revisitable targets. For this process, the mean first passage time to a target satisfies the following backward equations (for derivation of these equations, see section 5 and [Redner, 2001]) :

$$D\nabla^2_{\mathbf{r}} t_1 + \frac{1}{2\pi\tau_1}\int_0^{2\pi}(t_2 - t_1)d\theta_{\mathbf{V}} - k\mathrm{I}_a(\mathbf{r})t_1 = -1, \tag{50}$$

$$\mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 - \frac{1}{\tau_2}(t_2 - t_1) = -1, \tag{51}$$

where t_1 stands for the mean first passage time starting from state 1 at position $\mathbf{r}$, and t_2 for the mean first passage time starting from state 2 at position $\mathbf{r}$ with velocity $\mathbf{V}$. $\mathrm{I}_a(\mathbf{r}) = 1$ if $|\mathbf{r}| \leq a$ and $\mathrm{I}_a(\mathbf{r}) = 0$ if $|\mathbf{r}| > a$. In the present form, these integro-differential equations do not seem to allow for an exact resolution with standard methods. If the searcher initially starts in phase 2, and if the target is close, its initial direction matters. But as soon as the initial position is far from the target, there are numerous reorientations before finding the target, implying that

the initial direction does not matter. We take into account that the searcher initially does not know where the target is by averaging the mean search time over the initial starting position, uniformly distributed in the disk. Consequently, if $b \gg a$ and once the mean search time has been averaged over the starting position, the effect of the initial direction can be neglected. It allows us to make an approximation and solve the system (for more technical details, see section 5.4.1 for the static mode and section 5.4.2 for the diffusive mode).

3.4.4 Results for the diffusive mode

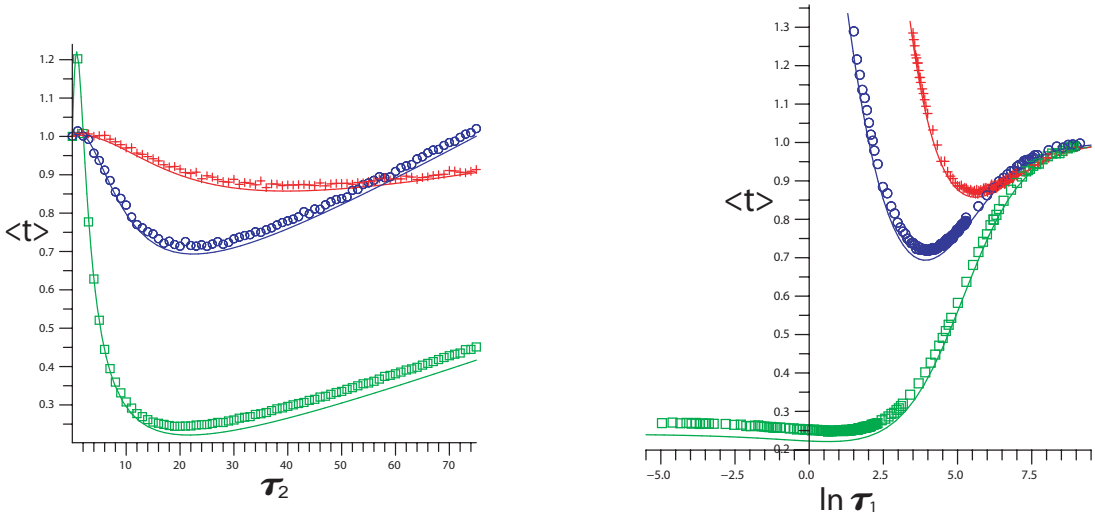


Figure 33: Simulations (symbols) versus analytical approximate (equation 52) (line) of the search time in the diffusive mode : the search time rescaled by the value in absence of intermittence as a function of τ_2 (left) and $\ln \tau_1$ (right) (the logarithmic scale has been used due to the flatness of the minimum), for $D = 1$, $V = 1$, $b = 451$ and $a = 10$ (green $\square$), $a = 1$ (blue $\circ$) and $a = 0.1$ (red $+$).

For the diffusive mode ($k \to \infty$), the mean search time $\langle t \rangle$ estimate is :

$$\langle t \rangle = (\tau_1+\tau_2)\frac{1-a^2/b^2}{(\alpha^2 D\tau_1)^2}\left\{a\alpha(b^2/a^2-1)\frac{M}{2L_+} - \frac{L_-}{L_+} - \frac{\alpha^2 D\tau_1}{8\widetilde{D}\tau_2}\frac{(3-4\ln(b/a))b^4 - 4a^2b^2 + a^4}{b^2-a^2}\right\}, \tag{52}$$

$$\begin{aligned}\text{with } L_\pm =& \mathrm{I}_0\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_1(a\alpha) - \mathrm{I}_1(a\alpha)\mathrm{K}_1(b\alpha)) \\ &\pm \alpha\sqrt{\widetilde{D}\tau_2}\,\mathrm{I}_1\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_0(a\alpha) + \mathrm{I}_0(a\alpha)\mathrm{K}_1(b\alpha))\,,\end{aligned} \tag{53}$$

and :

$$\begin{aligned} M =& \mathrm{I}_0\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)\left(\mathrm{I}_1(b\alpha)\mathrm{K}_0(a\alpha)+\mathrm{I}_0(a\alpha)\mathrm{K}_1(b\alpha)\right) \\ & -4\frac{a^2\sqrt{\widetilde{D}\tau_2}}{\alpha(b^2-a^2)^2}\mathrm{I}_1\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)\left(\mathrm{I}_1(b\alpha)\mathrm{K}_1(a\alpha)-\mathrm{I}_1(a\alpha)\mathrm{K}_1(b\alpha)\right), \end{aligned} \tag{54}$$

where $\alpha = (1/(D\tau_1) + 1/(\widetilde{D}\tau_2))^{1/2}$, and I_i and K_i are modified Bessel functions. This expression (52) has proved to be in good agreement with numerical simulations for a wide range of parameters (see figure 33). The optimization of the explicit expression (52) leads to simple forms in the following situations, depending on the relative magnitude of the three characteristic lengths of the problem a, b, and D/V. We limit ourselves to the case of low target density ($a \ll b$), which is the most relevant for hidden target search problems. Three regimes then arise. In the first regime $a \ll b \ll D/V$, the relocating phases are not efficient and intermittence is useless. In the second regime $a \ll D/V \ll b$, it can be shown (see section 5.4.2) that the intermittence can significantly speed up the search (typically by a factor 2), but that it does not change the order of magnitude of the search time. On the contrary, in the last regime $D/V \ll a \ll b$, the optimal strategy, obtained for

$$\tau_1^{opt} \sim \frac{D}{2V^2}\frac{\ln^2(b/a)}{2\ln(b/a)-1},\ \tau_2^{opt} \sim \frac{a}{V}(\ln(b/a)-1/2)^{1/2}, \tag{55}$$

leads to a search time arbitrarily smaller than the non intermittent search time when $V \to \infty$:

$$gain = \frac{t_{\mathrm{diff}}}{t_m^{opt}} \simeq \frac{\sqrt{2}aV}{8D}\left(\frac{1}{4\ln(b/a)-3}\frac{\mathrm{I}_0\left(2/\sqrt{2\ln(b/a)-1}\right)}{\mathrm{I}_1\left(2/\sqrt{2\ln(b/a)-1}\right)}+\frac{1}{2\sqrt{2\ln(b/a)-1}}\right)^{-1}. \tag{56}$$

This optimal strategy corresponds to a scaling law

$$\frac{\tau_1^{opt}}{\tau_2^{opt}} \sim \frac{D}{a^2}\frac{1}{(2-1/\ln(b/a))^2}, \tag{57}$$

which does not depend on V.

3.4.5 Results for the static mode

We now turn to the static mode ($D \to 0$), which leads to the following estimation for the search time

$$\begin{aligned} \langle t\rangle =& \frac{\tau_1+\tau_2}{2k\tau_1 y^2}\left[\frac{1}{x}(1+k\tau_1)(y^2-x^2)^2\frac{\mathrm{I}_0(x)}{\mathrm{I}_1(x)}\right. \\ & \left.+\frac{1}{4}\left(8y^2+(1+k\tau_1)\left(4y^4\ln(y/x)+(y^2-x^2)(x^2-3y^2+8)\right)\right)\right], \end{aligned} \tag{58}$$

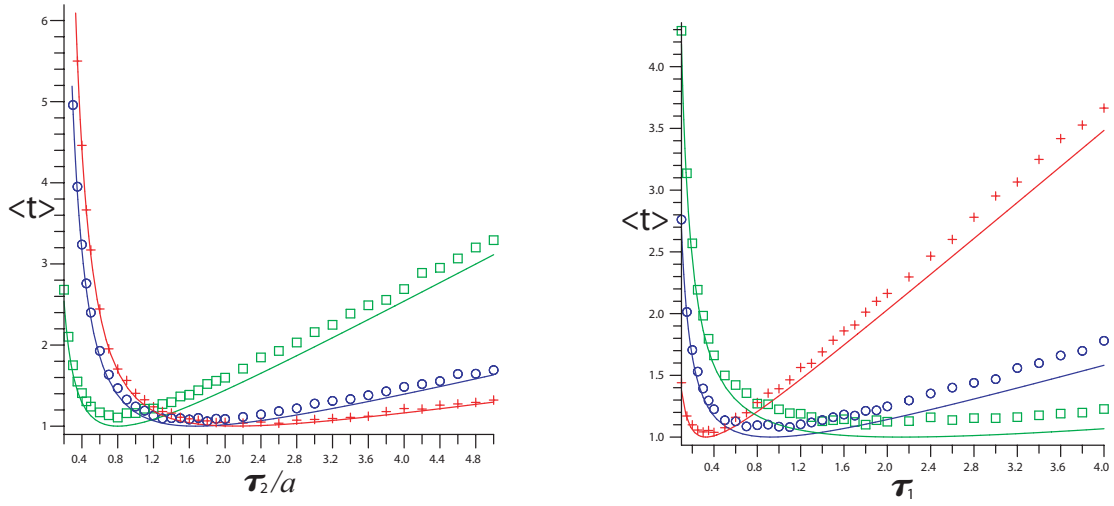


Figure 34: Simulations (points) versus analytical approximation (line) of the search time in the static mode : the search time rescaled by the optimal value as a function of τ_2/a (left) and τ_1 (right), for $k = 1$, $V = 1$, $b = 28$ and $a = 10$ (green $\square$), $a = 1$ (blue $\circ$) and $a = 0.1$ (red $+$).

$$\text{where } x = \sqrt{\frac{2k\tau_1}{1+k\tau_1}}\frac{a}{V\tau_2} \text{ and } y = \sqrt{\frac{2k\tau_1}{1+k\tau_1}}\frac{b}{V\tau_2}. \tag{59}$$

Here again, this expression (58) is in very good agreement with numerical simulations for a wide range of the parameters (see figure 34)). In this case, intermittence is trivially necessary to find the target, and the optimization of the search time (58) leads for $b \gg a$ to :

$$\tau_{1,\min} = \left(\frac{a}{Vk}\right)^{1/2}\left(\frac{2\ln(b/a)-1}{8}\right)^{1/4}, \tag{60}$$

$$\tau_{2,\min} = \frac{a}{V}\left(\ln(b/a) - 1/2\right)^{1/2}, \tag{61}$$

which corresponds to the scaling law $\tau_{2,\min} = 2k\tau_{1,\min}^2$, which still does not depend on V.

3.4.6 Conclusion

We have proposed a two state model of search processes for non-revisitable targets, which closely relies on the experimentally observed intermittent strategies adopted by foraging animals. Using a decoupling approximation numerically validated, we have studied analytically the physically relevant bidimensional geometry, allowing us to draw conclusions. (i) The mean search time $\langle t \rangle$ presents a global minimum for finite values of the τ_i, which means that intermittence is an optimal strategy. (ii) The optimal τ_1^{opt} are different and depend explicitly on D and k, leading to different scaling laws which are susceptible to discriminate between the two search modes . (iii) A very striking and non intuitive feature is that both modes of search studied lead to the *same optimal value* of τ_2^{opt}. As this optimal time does not depend on the specific characteristics D and k of the search mode, it seems to constitute a property

of intermittent search strategies, as discussed in detail in the framework of a more general model studied in section 5.

3.5 Are Lévy strategies really so advantageous?

As seen before, intermittent strategies are an alternative to Lévy walks (defined in section 3.1) for interpreting trajectories of foraging animals. However, the Lévy walks are often thought as optimal and widespread in nature. Is it really true?

3.5.1 The albatross story

Many foraging animals, including albatrosses, deers and bumblebees to name a few, have been thought for long to adopt Lévy strategies described in the pioneering work of Viswanathan et al. [1999]. These foraging behaviors were repeatedly accounted for by a simple model stating in the more general framework of search processes that Lévy walks are optimal search strategies, as they constitute the best way to explore space. Recently, Edwards et al. [2007] reanalyzed these data, completed by newly gathered data on foraging albatrosses, and showed that in fact there was no experimental evidence for the Lévy flight behavior. This study challenges the interpretation of several experimental works, but also raises a new important and puzzling question : why animals do not adopt the Lévy flight strategy which has however been reported to be an optimal search strategy [Travis, 2007]? Here we clarify this apparently paradoxical situation.

3.5.2 Optimizing the encounter rate with targets with Lévy walks : how and when?

Viswanathan et al. [1999] actually consider in their model (see section 3.1 above) two very different types of targets, which lead to two very different optimal strategies (*i.e.* maximizing the number of targets detected at large time) :

- In the first case of what they call "revisitable targets" - meaning that , as soon as detected, a target reappears *at the same location* - they find that the encounter rate is optimized for a Lévy exponent $\mu \simeq 2$.

- In the second case of "destructive search" where each target can be found only once, or in the case of a single available target, the optimal strategy proposed in Viswanathan et al. [1999] is not anymore of Lévy type, but reduces to a simple linear ballistic motion.

Since then, further studies have completed these results.

Bartumeus et al. [2002] studied the case of non-revisitable moving targets. They showed that a Lévy strategy with $\mu = 2$ is often better than a "Brownian" one ($\mu \geq 3$). However, James et al. [2008] extended the study to ballistic motion, that outperformed these Lévy strategies.

An intermediate situation has been studied by Raposo et al. [2003], Santos et al. [2004], where the immobile target is destructed upon encounter, but regenerates after a time τ at the same place. There are two regimes. When τ is large ($> \tau_c$,

a critical time), ballistic motion remains the best strategy. When $\tau < \tau_c$, the best μ is $1 < \mu^{opt} < 2$. However, if τ is smaller than the time needed to travel between two targets, waiting for renewal of the target will outperform searching for an hypothetical other target. When τ_c is explicitly calculated or simulated in Raposo et al. [2003], Santos et al. [2004], it is smaller than the typical time spent to travel from one target to another.

In Bartumeus and Levin [2008], targets are in patches or Lévy distributed. Even if the targets are destroyed upon encounter, finding a target means that the presence of other targets in the vicinity is likely, which is close to the case of revisitable targets. Unsurprisingly, optimum is achieved for a Lévy distribution, with $\mu \simeq 2$.

In Reynolds and Bartumeus [2009], the optimum for destructive targets is $\mu \to 1$ except in two cases (where $1 < \mu^{opt} \leq 2$). On the one hand the optimum is not ballistic when the searcher can fail in capturing a detected target. On the other hand, for targets destroyed upon encounter, and for the very specific one-dimensional case, as the measure of efficiency is the number of targets captured during a long time, the searcher is after some time in a situation with a target close on one side, but the next target on the other side very far away : pure ballistic motion is not favored because it can take the wrong direction.

Finally, in the case of revisitable targets and the related cases (regenerating targets, patches, failed capture), the Lévy strategy $\mu \simeq 2$ emerges only as a compromise between trajectories returning to one and the same ever target zone, and straight ballistic motion which is indeed the best way to explore space. However, it does not mean that Lévy strategy is the most relevant here. Indeed, once a target is found, the animal could switch behavior as suggested by Benhamou [1992]. In the case of non-destructed or regenerating target, it would be enough to stop and wait. In the case of a failed capture, or of patched preys, "Brownian" search ($\mu \geq 3$), or other strategies which enable the searcher to thoroughly explore its surroundings where target(s) lie, would be efficient. For allowing the searcher to use strategies better than Lévy, it is enough that the internal state of the animal changes when it detects a target, which is a very minimal form of memory. Thus Lévy walks are optimal only in restrictive conditions.

3.5.3 Do animals really perform Lévy walks?

As the optimality of Lévy strategies crucially requires conditions on the targets (regenerating at the same place, patched or hard to capture) and conditions on the searcher (no switch when a target is found, which is a very simple form of memory), it cannot be taken as a general rule even if realistic for certain species. On the contrary, we argue that the general question of determining the best strategy for finding a single hidden target belongs to the situation of destructive search, where in the framework of the model of Viswanathan et al. [1999], the most efficient way to find a randomly hidden target is simply a linear ballistic motion and *not* a Lévy strategy. As a consequence, there is no paradox : the reason why Lévy walks are not observed in the work of Edwards et al. [2007] is probably because they do not constitute robust optimal search strategies.

And what about other experimental observations? Among experimental studies analyzing organisms trajectories as a succession of segments interspersed with turns,

many of them report that the times between turns are distributed exponentially (a list of examples far from exhaustive : *C.Elegans* worm [Fujiwara et al., 2002, Pierce-Shimonura et al., 1999], fish [Hill et al., 2000], plankton in a part of the conditions studied by Bartumeus et al. [2003], amoebae [Li et al., 2008], etc.). However, apart from the controversial albatrosses study [Edwards et al., 2007], there is a boom in publications claiming that Lévy behavior is observed for some animals. A part of them can be dismissed as a convincing proof of Lévy behavior. On the one hand, as explained in details in Edwards et al. [2007], due to experimental limitations, most data span over a very limited range (typically 1-2 decades), and only a few studies go beyond a linear fit with logarithmic scale which cannot convincingly rule out other decreasing functions (exponential or Gamma laws for example). On the other hand, patterns and processes should not be confused, as underlined by Benhamou [2007]. The same observed patterns can often be explained by different models. It is not because a trajectory is similar to a Lévy walk that the underlying process is necessarily a Lévy walk. For example, a composite classical random walk can look very similar to a Lévy walk for a time short enough (see figure 35). Nonetheless, not all the studies could be discarded (see Viswanathan et al. [2008] for a review), but it does not mean that Lévy walks are an efficient search strategy. Indeed, as underlined by Viswanathan et al. [2008], other selection pressures could be predominant. For example, when targets location is known and when exploitation is optimized instead of search, Lévy walks can emerge from the interaction between the environment and the searcher [Boyer et al., 2006, Santos et al., 2007, Jiang et al., 2009].

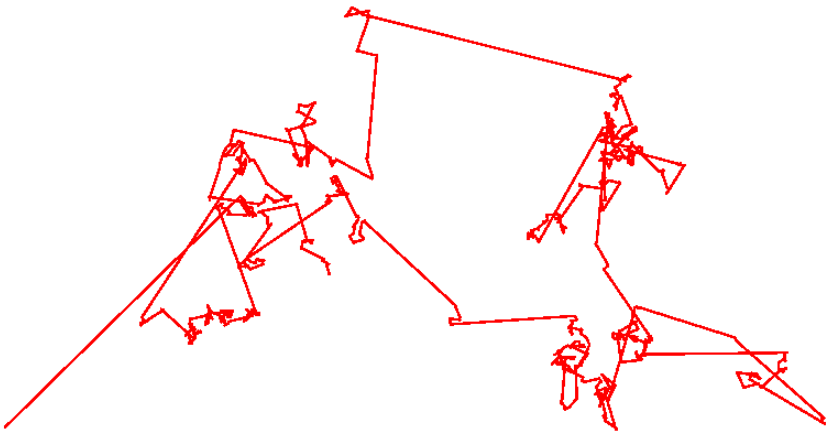
3.6 Conclusion on animal foraging

Lévy walks are a fashionable model for interpreting trajectories of foraging animals. Two main restrictions should be kept in mind. On the one hand, there is a controversy about at least a part of the experimental data which were thought to support Lévy walks. On the other hand, the conditions in which Lévy walks are optimal are very restrictive. However, this does not rule out any contribution from Lévy walks. For example, as discussed by Lomholt et al. [2008] and in section 6.3.2, intrinsically intermittent search models could be advantageously combined with Lévy walks.

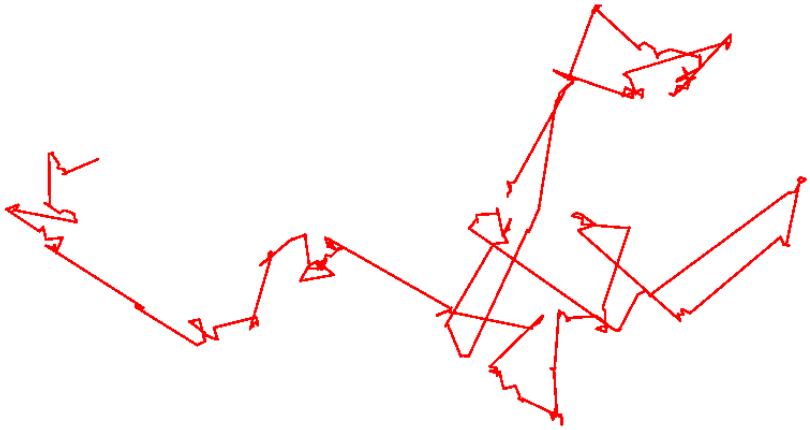
In this context, we argue that some animals cannot detect their target when they are moving ballistically, and in fact alternate between these fast but blind phases with slow detection phases.

A simple model of intermittent search has been extended for this thesis project. The influence of target distribution on the mean search time has been studied. The relevant case of a searcher in a two-dimensional space without correlation in the direction of successive fast "blind" phases is another important extension. We have shown that the mean search time for such intermittent behavior can be minimized by tuning the mean durations of each phase. In particular, in the latter case, the optimal mean duration of the fast blind phase does not depend on the description chosen for the slow detection phase.

Intermittent search strategies, because they rely on experimental observations, and because they prove to be efficient and robust, appear as a good alternative for interpreting animals trajectories.



Lévy walk with $\mu = 2$



Composite random walk : alternation of about 10 short steps (mean 1, distributed exponentially), and one large step (mean 10, distributed exponentially).

Figure 35: Comparison between a Lévy walk and a composite random walk : they look similar at short times.

4 Intermittent search strategies at the microscopic scale

Contents

4.1 Protein/DNA interaction 61
4.1.1 Short review of facilitated diffusion 62
4.1.2 Hops and jumps : normal diffusion 70
4.1.3 Hops and jumps : comparison with single molecule experiments 77
4.1.4 Hops and jumps : subdiffusion 88
4.1.5 Conclusion on DNA enzyme interaction 90
4.2 Active transport of vesicles in cells 92
4.2.1 Active transport in cells 92
4.2.2 Model 93
4.2.3 Methods 93
4.2.4 Active transport in the cytoplasm 94
4.2.5 Active transport at membranes 96
4.2.6 Active transport in tubular structures 97
4.2.7 Conclusion 98

We have just seen that intermittent search strategies are observed at the macroscopic scale. They are also observed at the microscopic scale. In what follows, we will focus on two examples : the localization by a protein of a specific DNA sequence; and active transport towards reactive targets inside cells. In the first example, it has already been shown that the search is intermittent and that this intermittence does speed up the search. For this thesis project, an important aspect of these modelings has been completed, namely the distribution of length traveled on DNA for a 3-dimensional excursion. To describe the case of vesicle transport, we build a completely new model based on intermittent search.

4.1 Protein/DNA interaction

The first example of microscopic intermittent search we present is the search by a protein for a specific sequence on DNA.

First, we introduce the concept of facilitated diffusion and review recent models. In particular, the searching protein alternates between 1-dimensional and 3-dimensional phases. The 3-dimensional excursions can be separated in "jumps" and "hops" that we define in the following. We explain why the hop distribution is needed (section 4.1.1). After this introduction, we analytically give the hop distribution in the simple case of normal diffusion in an infinite space (section 4.1.2). Then, these results are adapted to interpret a single-molecule experiment, done in a finite observation time and in a specific geometry (section 4.1.3). Eventually, the results are extended to anomalous diffusion, relevant to in vivo situations (section 4.1.4).

The work presented in this section has been done in collaboration with Pierre Desbiolles, along with his post-doc Andreas Biebriecher and his PhD students Isabelle Bonnet and Natacha Porté, of the optic and biology team, LKB (ENS, Paris). Christophe Escudé (MNHN) helped also for the DNA construct and showed me some molecular biology. This work is reported in two publications : Bonnet et al. [2008] and Loverdo et al. [2009c].

4.1.1 Short review of facilitated diffusion

Biological context

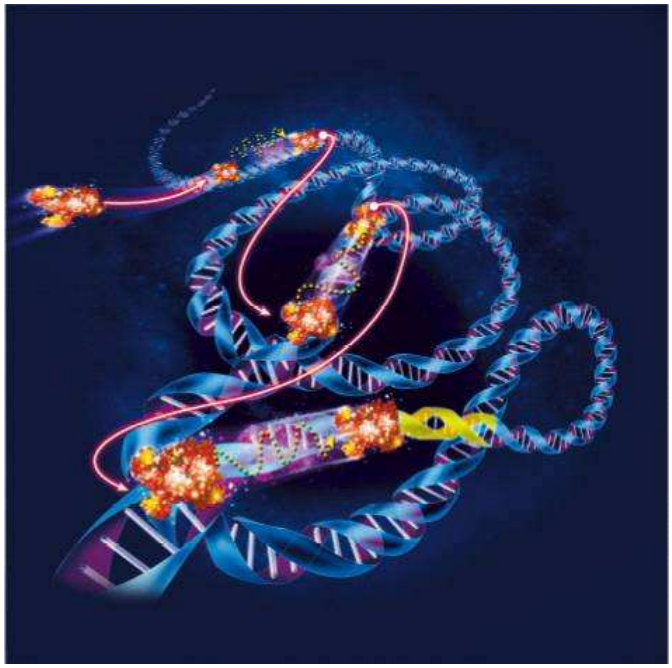


Figure 36: Artist view of a DNA/protein interaction, which combines 1–dimensional sliding phases and 3–dimensional relocation phases. Picture by Virginie Denis/Pour la Science n°352, February 2007.

Various functions of living cells - and therefore at larger scales of living organisms - are regulated by chemical reactions between different molecules. The first step before reaction is the encounter between these different molecules. The importance of the kinetics of such biochemical reactions is illustrated by the bacterial restriction and modification system [Wilson and Murray, 1991], that involves couples of methyltransferase and restriction enzymes recognizing the same sequence on DNA (for example *Eco*RV recognizes the sequence $GATATC$ [Taylor and Halford, 1989]). Methyltransferase enzymes methylate this specific sequence on the bacterial DNA in order to protect it from restriction enzymes. When an unmethylated viral DNA enters the cell, it is very likely to contain the target sequence. Indeed, this sequence, typically 4-8 base pairs, is short compared to the viral genome, which, depending on the virus, is made of $10^3 - 10^6$ base pairs (typically 5.10^4 for bacteriophages) (additionally, for a better efficiency, there are different restriction enzymes). The infected bacteria then faces a vital search problem : restriction enzymes must find

their target sequence on the viral DNA to inactivate the virus before it exploits the bacteria machinery and kills it.

More generally, it is well established that some sequence-specific proteins find their target site in a remarkable short time. For the lac repressor for example, Riggs et al. [1970] measured association rates orders of magnitude larger than those expected for reactions limited by the classical three-dimensional diffusion (results confirmed by Hsieh and Brenowitz [1997] at different salt concentrations, ruling out electrostatic effects as the only explanation). Halford [2009] argues that in fact only a few enzymes react significantly faster than the 3-dimensional diffusion limit. But he underlines that there are many enzymes reacting at rates close to the diffusion limit, and that this observation is still impressive. Indeed, enzymes have to find their target in experiments with a considerable excess of DNA, and, as they have to probe the target to be sure it is the right one, they are expected to waste a lot of time on bad sequences. In a series of seminal articles, Berg et al. [1981], Winter and Von Hippel [1981], Winter et al. [1981] proposed that 3D diffusion ("hopping"/"jumping") was not the only motion available to the protein, even if no energy is consumed (unlike some enzymes consuming energy to scan DNA processively). In some cases, proteins have several sites that can associate to DNA, and thus they could do "intersegmental transfer". We shall consider in the following the case of proteins with only one site of association with DNA, ruling out this possibility. Another possibility is a weak electrostatic interaction between these proteins and DNA, enabling the proteins to diffuse along the DNA, a process called "sliding" (see Von Hippel [2007] and Dahirel et al. [2009] for more details on the weak electrostatic interaction). Berg and Von Hippel proposed that the combination of sliding and 3D diffusion, *i.e.* facilitated diffusion, makes the search for a sequence faster than 3D diffusion alone.

Actually, this search mechanism can be classified as intermittent, in the general meaning defined in section 2.1. Indeed, on the one hand, 3-dimensional diffusion off the DNA is fast, but does not allow for target detection. On the other hand, sliding is a phase of motion along DNA, which therefore enables target detection. The enzyme motion during sliding is however much slower (there is much more friction, and in some models, a protein of size R is assumed to follow the helical groove, leading to a diffusion coefficient scaling as R^3 instead of R for classical diffusion [Schurr, 1979]).

Such intermittent trajectories can reduce the mean search time, as shown by different models (see for example Coppey et al. [2004], Slutsky and Mirny [2004] for simple approaches). We present below a simple argument for the efficiency of intermittence using a simple model. Sliding is taken as diffusive, with its duration exponentially distributed, of mean τ_1. The mean span of such a diffusive phase is $L_1 = 2\sqrt{D\tau_1}$. 3D excursions are taken as "teleportation", meaning that after a 3D excursion, the new position on the DNA sequence is taken at random, uncorrelated with the previous position. These excursions are supposed to have a mean duration τ_2. The mean search time is then $\langle t \rangle \simeq (\tau_1 + \tau_2)N$, where N is the mean number of 3D excursions needed to find the target. The average probability of finding the target during one sliding phase is L_1/ℓ, ℓ being the DNA length. Consequently $\langle t \rangle \simeq (\tau_1 + \tau_2)\frac{\ell}{2\sqrt{D\tau_1}}$. Here, the mean search time scales linearly with the DNA length, whereas with sliding alone, the mean search time scales as ℓ^2. More precisely,

the mean search time is minimized for $\tau_1 = \tau_2$. It is what is found by Slutsky and Mirny [2004] and is refined by Coppey et al. [2004].

The pioneering studies on facilitated diffusion [Riggs et al., 1970, Berg et al., 1981, Winter and Von Hippel, 1981, Winter et al., 1981] are based on ensemble measurements, which were for a long time the only way to experimentally access to protein/DNA interaction. Recently developed techniques make possible the observation of this interaction at the level of a single molecule, with a resolution in space and time still improving (for a recent review on the experimental results of such techniques, see Gorman and Greene [2008]). It is now confirmed directly that many proteins searching for a specific sequence on DNA combine "hopping/jumping" and "sliding" (see figure 36). Sliding phases have been clearly identified (both in vitro [Kabata et al., 1993] and in vivo [Elf et al., 2007]), as well as hopping/jumping phases [Gowers et al., 2005, van den Broek et al., 2008, Komazin-Meredith et al., 2008, Bonnet et al., 2008].

With these new experiments, theoretical studies have bloomed too. In what follows, we review recent models (we do not treat intersegmental transfer, but only motions accessible for proteins with a single DNA binding site). First, we define the methods and the observables used, then the descriptions of the 1D phase and the 3D phase that have been proposed.

Approaches

We first classify the main approaches used.

Elementary interactions

Some models (see for example Florescu and Joyeux [2009]) propose to go back to the electrostatic potential created by DNA and to follow dynamics by molecular simulations. When they simulate the trajectory of a protein subject to these potentials, they do observe sliding and 3D excursions. However, answers are only found numerically, and hypotheses on the interaction energies are not more justified than hypotheses directly on the 1D and 3D motions properties.

Scaling laws

Some other models are based on scaling arguments. Halford and Marko [2004] for example propose an optimization of the sliding length by roughly estimating the time for the protein to find the coil, then the time to find the target inside the coil. They do find an optimal sliding length. Hu et al. [2006] give scaling relations using equality between fluxes. They propose that there is an "antenna length", *i.e.* a typical scale below which the dominant transport is sliding. They see the DNA as a collection of beads of antenna length size. They balance the 3D flux to DNA, and the 1D flux to the target. The latter flux is dependent on the probability that an enzyme is in the bead containing the target, and on the typical time and probability of finding the target when in the right bead.

Stochastic modeling

Another way to construct models is to make assumptions on the 1D and 3D motions, follow what can happen to a single protein and with which probability,

attempt to obtain mathematically the probability distribution of the search time, rich in information, and exploit it. This method is used by Coppey et al. [2004], Lomholt et al. [2005, 2007, 2009], Eliazar et al. [2007, 2008], Meroz et al. [2009], Bénichou et al. [2009]. Meroz et al. [2009] underline that this stochastic approach gives similar results as "kinetics" approach (based on concentrations and reaction rates), but allows numerous extensions.

Parameters relevant to the optimization of the search

Mean search time

Minimizing the mean search time is the optimization procedure the most often used (see for example Coppey et al. [2004]). However, obtaining the entire distribution of the search time is more informative than the mean search time alone. Indeed, the relevant parameter for the cell is often the mean search time (as it is at the first order the inverse of the reaction rate), but not always, as will be illustrated.

Variability

Wunderlich and Mirny [2008] partly focus on variability. A target which is close to the starting point of the protein can be found by sliding alone, and with a low variability of the search time. In contrast, if the target is far away from the starting point of the protein, it is found after numerous 3D excursions. The mean search time is longer, and the spread of the distribution of the search time is larger, even relatively to the median of the distribution. This spread induces more noise. As noise could sometimes be problematic, reliability of the signal may be more important for some functions than kinetics.

Typical time

Another outcome of variability is that the mean time is not necessarily representative of the typical time. Bénichou et al. [2009] propose a model in which enzymes can be stuck forever at the wrong place. The mean search time is then infinite. But, depending on the parameters, the probability p of such a catastrophic event can be very low, and what matters will be the typical time. They underline that this effect is particularly important when there are $n > 1$ searchers. In this case, the catastrophic event happens only when all the n enzymes are stuck, event of probability p^n, decreasing rapidly with n.

Mean search time for several searchers

The influence of the number of searchers is also discussed in Sokolov et al. [2005], Eliazar et al. [2007, 2008], Meroz et al. [2009]. As before, the mean time for the target to be found by any of the searchers is not simply the mean time divided by the number of searchers, it depends on the distribution of the search time for a single searcher. Moreover, if the concentration of searchers increases considerably, searchers cannot be considered as independent anymore. Proteins sliding on DNA will act as "roadblocks" for each other.

Large number of searchers

These roadblocks will decrease the effective sliding length, and may also hide the target. This can have negative consequences. Li et al. [2009] see it as a trade-off. On the one hand, the more the proteins, the more the searchers for the target, the quicker the search. On the other hand, the more the proteins, the more the crowding, the less efficient the search of a single protein. They predict that the optimum is $10^4 - 10^5$ DNA binding proteins for E.Coli, close to the experimental value of 30 000 proteins.

Colocalization

These roadblocks can however be beneficial if the target is close, as they prevent the sliding searcher from overshooting the target. Kolesov et al. [2007], Wunderlich and Mirny [2008] study transcription factors, proteins binding on DNA and regulating how often a gene sequence is read, thus regulating gene expression. They argue that a sequence coding for a transcription factor should be colocalized with its target sequence (see also Bénichou et al. [2008b] for a further optimization of this colocalization effect with respect to the diffusion coefficient of the protein). It is indeed what is observed in real prokaryotes genomes. It only works in prokaryotes, where there is no cell compartments separating protein production from DNA. But other mechanisms could be imagined in eukaryotes.

Eukaryotes

Indeed, in eukaryotes, DNA is packed inside the nucleus. Not all the DNA is equivalent. A first heterogeneity comes from having DNA close to the nucleus pores, and DNA buried deep inside the nucleus. Moreover, eukaryote DNA is packed in what is called chromatin (a complex system of proteins binding to the DNA and controlling the 3D structure). Some DNA regions could be more or less accessible depending on the chromatin configuration. The 3D regular structure of chromatin could also often bring together the same sequences in the 3D space, a sequence far away in the linear sequence could consequently help locating the target sequence. Kampmann [2005] argues qualitatively that proteins binding to DNA could take advantage of these heterogeneities : depending on the searched sequence, the optimal strategy is not the same.

Specificity

If different strategies are adopted depending on the target, it will also help for specificity. For example, in the case of the restriction-modification system, the difficulty is methylating the sequence on bacterial DNA while cutting the same sequence on viral DNA. As bacterial and viral DNA do not have the same characteristics (bacterial DNA is supercoiled, longer, with more bound proteins), the properties of the enzymes search could have been selectioned by evolution to ensure maximal inactivation rate for viral DNA, while not plaguing bacteria with auto-immune damage [Nardone et al., 1986].

Finally, various observables can be optimized, but the first step of modeling always consists in defining the different motions. Facilitated diffusion consists of 1D and 3D excursions. Different modelings of these two phases have been proposed, which we review in the next sections.

Descriptions of the 1D phase (sliding and recognition)

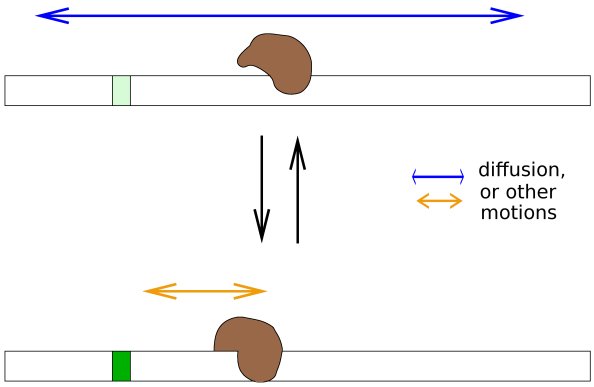


Figure 37: Sliding is often represented by diffusion with perfect reactivity on the target. This figure shows the two main directions for a more realistic description of the 1D phase : on the one hand, the sliding is not necessarily diffusive, and on the other hand the 1D phase could be in fact a combination of 2 phases, one fast but with low recognition, and another slow (or immobile), but with high recognition of the target.

Simplest description

The phase of one-dimensional interaction with DNA, "sliding", is often described as Brownian diffusion with perfect reaction when passing on the target (see for example Coppey et al. [2004] or Bénichou et al. [2008b]).

How the motion could depart from diffusion

A first limitation of this simple description is when proteins binding to DNA are numerous and create traffic jams [Sokolov et al., 2005]. But even in the case of a single protein, it should be kept in mind that the DNA is not homogeneous. Indeed, the DNA function is to encode information, implying an heterogeneous sequence. Then, the DNA/protein interaction energy will vary with the sequence. The energy distribution is often assumed to be Gaussian [Barbi et al., 2004, Wunderlich and Mirny, 2008]. Barbi et al. [2004] show that in this case sliding is not purely diffusive : the protein will be trapped for a longer time on some sequences. However, with realistic numbers, they argue that for sliding longer than a hundred base pairs, diffusive behavior is recovered because of averaging.

How the motion influences results

As sliding could depart from Brownian diffusion, the influence of the type of sliding motion has to be investigated. Coppey et al. [2004] propose a simple model describing 3D phases as "teleportation", and 1D motion as Brownian motion, with

exponential durations of both phases. The exact analytical expression of the Laplace transform of the first passage time distribution is calculated. Intermittence is found to decrease the mean search time if the length of DNA ℓ is large enough. Indeed, with the 3D excursions, the mean search time scales as ℓ, whereas with sliding alone, the mean search time scales as ℓ^2. This simple model is extended by Eliazar et al. [2007] (see also the shorter articles Eliazar et al. [2008], Meroz et al. [2009]). They rewrite the Laplace transform of the first passage time distribution, and discuss results for several non Brownian 1D motions (ballistic, self-similar, with halts (in particular when halts durations are widely distributed, leading to a subdiffusive behavior)). They find that there are always regimes in which intermittence is favorable. In particular, in the case of 3D excursions with finite mean durations, the mean search time with arbitrary scanning mechanism cases remains of order $\propto \ell$. Consequently, for long enough DNA, intermittence is favorable compared to sliding alone, for a wide range of sliding motions.

Target recognition

Another simplification often used is to assume perfect reactivity with the target, *i.e.* when sliding, the searcher detects immediately its target when it passes on it. Slutsky and Mirny [2004] discuss the influence of sequence roughness. If σ, the typical energy scale of interaction energy between protein and DNA, is of order k_BT, the sliding diffusion is fast, but the enzyme would not stop at its target. If the typical energy scale σ is of order $5k_BT$, recognition is at the contrary high, but sliding diffusion coefficient is very low, leading to a huge search time. They propose that the enzyme changes conformation while sliding, alternating a search state (low σ) and a recognition state (high σ). If the two energy landscapes are highly correlated, it is possible to conciliate high speed and high reliability. Hu et al. [2008] have extended their previous scaling arguments [Hu et al., 2006] in the case of a similar 2-state enzyme. More quantitatively, Bénichou et al. [2009] study a search strategy where three modes are available for the enzyme : 3-dimensional excursions, modeled as teleportation; 1-dimensional "search" mode : diffusive sliding, but without target detection; and 1-dimensional "recognition" mode : immobile, but allowing the protein to form a stable complex with the target. The energetic profile of the barrier from the "search" to the "recognition" state is taken Gaussian, the target being the lowest point of this profile. In the extreme case with no return from the recognition phase, there are catastrophic events when all the enzymes are stuck at a wrong sequence. However, depending on the parameters, the probability of such events and the typical search time can be low.

Descriptions of the 3D phase (jumping/hopping)

We now review the different descriptions of the phase in 3 dimensions.

Time distribution

In the simple model of Coppey et al. [2004], the distribution of the duration of the 3D phase is assumed to be exponential. The Laplace transform of the distribution of the first passage time to the target for this model is extended for any distribution of 3D excursion duration by Eliazar et al. [2007] (see also Eliazar et al. [2008],

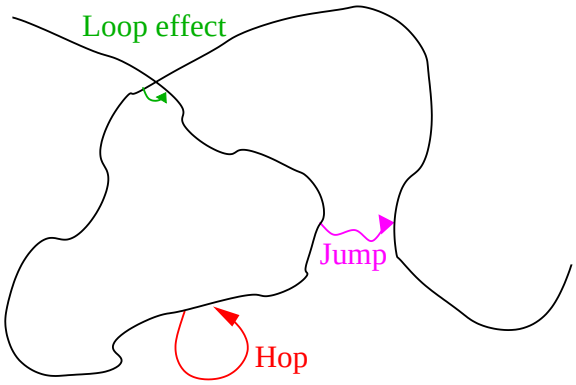


Figure 38: Description of 3D excursions : there are uncorrelated jumps, that can be considered as "teleportation"; but loops formed by the DNA can induce Lévy-distributed relocations; and 3D excursions can also return on DNA at short scales, leading to correlations in the position on the sequence of the starting and ending point (hops).

Meroz et al. [2009]). The expression of the mean search time does not depend on the duration distribution of the 3D phase, but only of its mean duration, as Coppey et al. [2004] have already noticed. Eliazar et al. additionally discuss what would happen with infinite mean relocation time. Lomholt et al. [2007] propose an explanation for such a large distribution. They explore the effect of a crowded environment with subdiffusion $\langle r^2(t)\rangle \propto t^\alpha$ $(0 < \alpha < 1)$ caused by waiting times distributed as $p(t) \sim \tau^\alpha / t^{1+\alpha}$. They argue that because of these waiting times, the probability that the protein has not yet left the DNA at time t and the probability that an unbound protein has not yet bound to DNA after a time t both scale as $1/t^{1+\alpha}$. Their results have two main practical implications. On the one hand, in an experiment, as enzymes can remain stuck for very long times, ensemble averages do not lead to the same results as time averages. On the other hand, as enzymes would slide for a longer time and as it would take them a very long time to return to DNA, the genes coding for transcription factors should be close to their target sequences, as already underlined by Wunderlich and Mirny [2008].

Jumps

Another widespread assumption (used for example by Hu et al. [2006] at scales larger then the antenna size, Coppey et al. [2004], Eliazar et al. [2007], Bénichou et al. [2009], etc.) is that the 3D excursions are "teleportations" : no matter how long an excursion lasts, the new position on DNA is chosen at random, uniformly over the DNA . The justification for such an approximation is that in biological systems, DNA is coiled. When the protein has a span (end-to-end length of the 3D excursion) greater than the DNA correlation length (also called persistence length, and noted ξ), the protein "sees" the coil. Thus, when performing a 3D excursion, the enzyme could bind on DNA at a location close in 3D, but very far away in the sequence. Randomness is justified in these relocations. We shall call "jumps" these 3D excursions, whose starting and ending points are uncorrelated in the 1D sequence. We shall call "hops" the correlated 3D excursions. van den Broek et al. [2008], Lomholt et al. [2009] underline that coiling is important, notably because

it changes the local DNA concentration : the more packed the DNA, the higher the local DNA concentration, the higher the probability of jumps compared to hops (enabling the protein to explore previously unscanned areas), and the less the time spent in 3D.

Loops

In complement to these "teleportation" jumps, Lomholt et al. [2005] suggested that as polymers form loops with length between contact points distributed as $|x|^{-1-\alpha}$ (for instance $\alpha = 0.5$ for Gaussian chains, $\alpha \simeq 1.2$ for self-avoiding walks chains), it adds relocations following this law : at the end of a 3D excursion, even proteins with a single binding site are likely to land on regions of DNA that are close in the 3D space. They obtain a rich behavior and optimization depending on α, for an annealed description of the loops (relevant as the DNA conformation fluctuates). Li et al. [2009] suggest that sites susceptible to be close to the target by looping could have been designed to help searching for the target.

Hops

Hopping is often included in an effective sliding length (Hu et al. [2006] for example). It is not always satisfying. On the one hand, this effective sliding length could be hard to estimate (because it is not related only to the binding energy). On the other hand, hopping has different properties from sliding : a target is very unlikely to be found with a hop, and a hop could enable a protein to bypass obstacles. Wunderlich and Mirny [2008] present a composite model : sliding as 1D diffusion; teleportation jumps; and hops as 3D excursions rebinding to the same DNA strand, evaluated mainly numerically. They use numerical simulations to evaluate the distributions of hops (where the protein goes back) and the probability that a 3D excursion is a jump instead of a hop (found to be 0.1675). Their results depend strongly on the size of the nodes chosen (here 1 nm). A simple way to understand why the node size matters is to perform a random walk in one dimension : a target is located at $x = 0$, and an absorbing wall is at $x = 4$ (analogue to the limit length beyond which 3D excursions are considered as hops). If you make steps of $\delta x = 2$, and you start one step away from your target, you find your target half of the time. If you make steps of $\delta x = 1$, and you start one step away from your target, you find the target $3/4$ of the time. Wunderlich and Mirny [2008] also try to evaluate analytically the jump probability (for an extension of their formula, see equation (78)). In the following section, we give a much more comprehensive analytical account of hops.

4.1.2 Hops and jumps : normal diffusion

Motivation

An analytic description of hops (*i.e.* 3D excursions with a span shorter than ξ, the DNA correlation length) is needed for more realistic models of facilitated diffusion. Indeed, two questions are to be solved. On the one hand, the hop distribution is important in itself, as it shows how far a hop goes, and how hopping could complete the other correlated motion, sliding. On the other hand, the proportion of hops and jumps among 3D excursions is important, as jumps enable the protein to

scan sequences far away from the starting point. We first give analytically the hop distribution for a simple configuration.

Note that Chechkin et al. [2009] and Levitz et al. [2008] study very similar problems of diffusion close to an absorbing cylinder, but in the limit of perfect absorption. They analyze the time of first return to the cylinder, whereas, as detailed in the following, we convert the time to the span of a 3D excursion. The conversion of time into distance is simple for normal diffusion, as diffusion in different directions are independent. The return time to the cylinder is calculated by considering only the plane perpendicular to the cylinder, whereas the quantity of interest is the distance traveled along the axis of the cylinder.

Model and methods

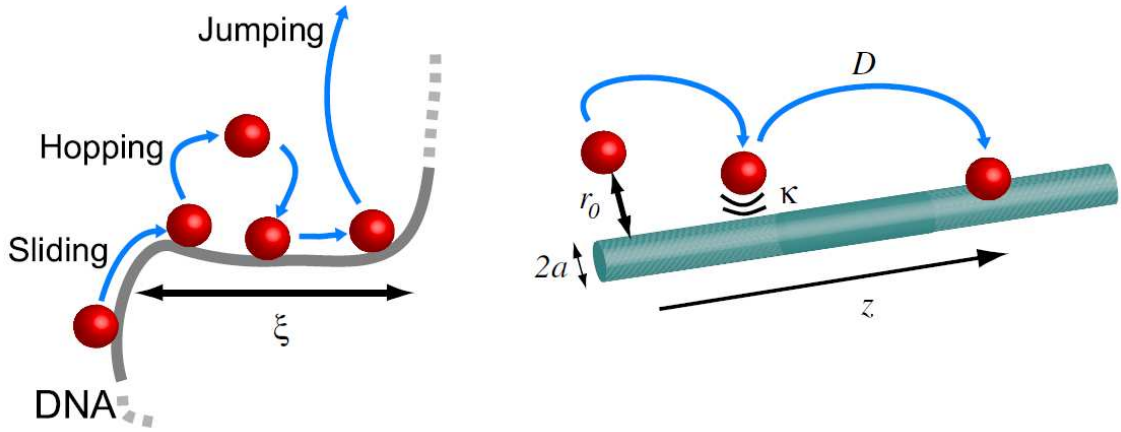


Figure 39: Facilitated diffusion of a protein on DNA. *Left:* schematic definition of sliding, hopping and jumping. *Right:* model parameters.

The simplest modeling of 3D excursions is normal 3D diffusion, with diffusion coefficient D. We model the DNA as an infinite cylinder of radius $a = R_{DNA} + R_{protein}$, and consider a point-like protein (modeling equivalent to a cylinder of radius R_{DNA} and a spherical protein of radius $R_{protein}$) (see Fig. 39 right).

We denote $c = c(r, z, \theta, t|r_0, z_0, t = 0)$ as the probability that a protein, starting at $t = 0$ from the point $r = r_0$, $z = z_0$, with uniform probability on θ, will be at (r, z, θ) at t. Because of the symmetry in θ,

$$\frac{\partial c}{\partial t} = D\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)c. \tag{62}$$

The boundary conditions are :

$$c(r, z, t = 0) = \frac{1}{2\pi r_0}\delta(r - r_0)\delta(z - z_0), \tag{63}$$

$$\left.\frac{\partial c}{\partial r}\right|_{r=a} = \kappa c(r = a). \tag{64}$$

This latter condition, a radiative boundary condition, traduces the idea that the cylinder is semi-absorbing [Redner, 2001]. If $\kappa \to \infty$, it is equivalent to a perfectly

absorbing cylinder. If $\kappa \to 0$, it is equivalent to a perfectly reflective cylinder. These equations are rewritten for $g = g(r, z|, r_0, z_0) = \int_0^\infty c dt$:

$$D\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right) g = -\frac{1}{\pi r_0}\delta(r - r_0)\delta(z - z_0), \tag{65}$$

$$\left.\frac{\partial g}{\partial r}\right|_{r=a} = \kappa g(r = a). \tag{66}$$

The probability that the protein comes back to the cylinder at z (starting from r_0 and $z_0 = 0$) is :

$$P(z|r_0) = 2\pi a \int_0^\infty D \left.\frac{\partial c}{\partial r}\right|_{r=a} dt = 2\pi a D \left.\frac{\partial g}{\partial r}\right|_{r=a} = 2\pi a D \kappa g(r = a). \tag{67}$$

g is symmetric in respect to z, as we have chosen $z_0 = 0$. It is convenient to decompose g into Fourier components :

$$g = \frac{1}{2\pi}\int_0^\infty \cos(kz) h_k(r) dk. \tag{68}$$

Then :

$$P(z|r_0) = aD\kappa \int_0^\infty \cos(kz) h_k(r = a) dk, \tag{69}$$

with h_k solution of :

$$\frac{d^2 h_k}{dr^2} + \frac{1}{r}\frac{dh_k}{dr} - k^2 h_k = -\frac{\delta(r - r_0)}{2\pi r_0 D}. \tag{70}$$

The solutions of the homogeneous equation are $K_0(kr)$ and $I_0(kr)$ (modified Bessel functions). $h_<(r)$ is the solution for $r \in [a, r_0]$, $h_>(r)$ is the solution for $r \in [r_0, \infty[$. The boundary conditions are :

$$h_>(r \to \infty) \to 0, \tag{71}$$

$$\left.\frac{\partial h_<}{\partial r}\right|_{r=a} = \kappa h_<(r = a), \tag{72}$$

$$h_<(r = r_0) = h_>(r = r_0), \text{ and} \tag{73}$$

$$\left.\frac{\partial h_>}{\partial r}\right|_{r=r_0} - \left.\frac{\partial h_<}{\partial r}\right|_{r=r_0} = -\frac{1}{2\pi r_0 D}. \tag{74}$$

We solve and obtain (noting $r_m = min(r, r_0)$ and $r_M = max(r, r_0)$, and using the Wronskian relation $K_0(x)I_1(x) + K_1(x)I_0(x) = 1/x$) :

$$h_k(r) = \frac{1}{2\pi D}\frac{K_0(kr_M)}{kK_1(ka) + \kappa K_0(ka)}\left(I_0(kr_m)\left(kK_1(ka) + \kappa K_0(ka)\right) + K_0(kr_m)\left(kI_1(ka) - \kappa I_0(ka)\right)\right). \tag{75}$$

In particular, for $r = a$:

$$h_k(a) = \frac{1}{2\pi D a}\frac{K_0(kr_0)}{kK_1(ka) + \kappa K_0(ka)}. \tag{76}$$

Results

Consequently :

$$P(z|r_0) = \frac{1}{\pi}\int_0^\infty cos(kz)\frac{K_0(kr_0)}{\frac{k}{\kappa}K_1(ka)+K_0(ka)}dk. \tag{77}$$

The behavior at large z is given by $P(z|r_0) \sim (\ln(r_0/a) + (\kappa a)^{-1})/(2z\ln^2(z/a))$. As underlined by Levitz et al. [2008], the tail of the distribution of returns to a cylinder is very wide, with an infinite mean return time, which translates here in an infinite mean distance $|z|$ between starting and ending points.

Going back to the original problem of DNA hopping/jumping in vivo, when the DNA is randomly coiled, this distribution answers two questions. On the one hand, Eq.(77) gives the analytical distribution of hops, as for $z < \xi$ (ξ being the DNA persistence length), modeling the DNA as a straight cylinder is legitimate. On the other hand, all returns with $z > \xi$ will be jumps, thus we have directly access to the proportion of jumps with the complementary cumulative distribution:

$$\bar{C}(|z = \xi|) = \int_{|z|>\xi} P(z|r_0)dz \sim \frac{\ln(r_0/a) + 1/\kappa a}{\ln(\xi/a)}. \tag{78}$$

Wunderlich and Mirny [2008] have also evaluated analytically the jump probability. They found $p_{jump} \sim \ln(r_0/a)/\ln(R_+/a)$. with r_0 and a defined similarly to our model, and R_+ is an upper limit, beyond which the protein is likely to bind to another DNA strand. They have indeed considered an infinitely absorbing cylinder, and their expression is coherent with ours. They have studied one asymptote (κ finite and $r_0 \to a$) whereas we completely solve the problem. In section 4.1.3, when we compare our model with experimental data, we choose the other limit regime (κ finite and $r_0 \to a$) because we think that this description is more relevant than the other limit regime for describing the interaction of an enzyme with DNA. Indeed, it is improbable that DNA could be perfectly absorbing, because, for example, the enzyme can present its wrong face to the DNA, which would not allow for the binding. Moreover, the enzyme starts a 3D excursion just after a sliding phase, very close to the DNA. It should be conceded that our model is continuous, and that it possible that a length $r_0 - a$ exists in the microscopic system. But this length is hard to evaluate, and if we describe the interaction with κ finite and $r_0 > a$, in the very likely regime $(r_0 - a) \ll \kappa^{-1}$, the contribution of r_0 to the jump proportion is negligible.

The equation (78) means that as soon as the protein departs from the DNA further than its diameter, the jump probability is high. Returning to the search problem, when jumping is favorable, decreasing the persistence length speeds up the search process, as found by van den Broek et al. [2008].

Median of the distribution

As seen above from equation (77), the mean distance of return is infinite. The relevant quantity to look at is therefore the typical return distance, given for instance by the median. In what follows, we study the median of the length of return (length

along the axis of the cylinder). We first study the two limits ($r_0 = a$ and κ finite first, then $\kappa \to \infty$ and $r_0 > a$) before deriving the full expression. Besides giving the typical return distance, these results give another physical interpretation of κ.

Regime $r_0 = a$

We first study the case of $r_0 = a$. In this regime, the probability of return at a distance z is :

$$P(z) = P(z|r_0 = a) = \frac{1}{\pi}\int_0^\infty \cos(kz)\frac{K_0(ka)}{\frac{k}{\kappa}K_1(ka) + K_0(ka)}dk. \tag{79}$$

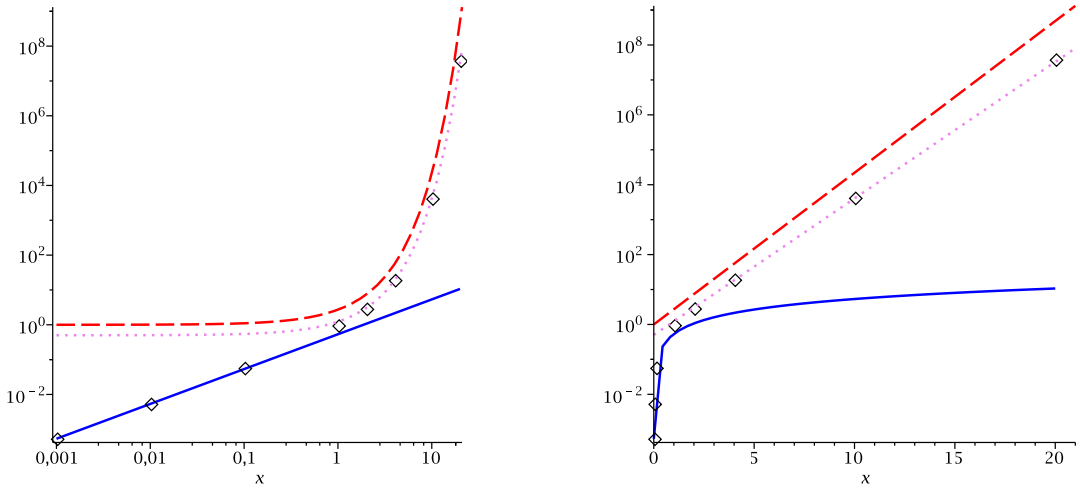
If $\kappa \to \infty$ (the cylinder is infinitely absorbing), it tends to $\delta(z)$, as expected.

The distribution of the size of the hops ($|z|$ instead of z) is :

$$P(|z|) = \frac{2}{\pi}\int_0^\infty \cos(kz)\frac{K_0(ka)}{\frac{k}{\kappa}K_1(ka) + K_0(ka)}dk. \tag{80}$$

The probability of making a hop smaller than z writes :

$$C(|z|) = \int_0^z P(|z'|)dz' = \frac{2}{\pi}\int_0^\infty \frac{\sin(kz)}{k}\frac{K_0(ka)}{\frac{k}{\kappa}K_1(ka) + K_0(ka)}dk. \tag{81}$$



$log\left(\frac{z_{med}}{a}\right)$ as a function of $log\left(\frac{1}{\kappa a}\right)$ $log\left(\frac{z_{med}}{a}\right)$ as a function of $\frac{1}{\kappa a}$

Figure 40: Median of returns as a function of κ^{-1} (with $r_0 = a$). Blue solid line : analytical expression for $\kappa^{-1} \ll a$: $z_{med} \simeq 0.535\kappa^{-1}$. Red dashed line : analytical expression for $\kappa^{-1} \gg a$: $z_{med} = a\exp\left(\frac{1}{\kappa a}\right)$. Violet dotted line : fit $z_{med} = \frac{a}{2}exp\left(\frac{0.9}{\kappa a}\right)$. $\diamond$: numerical estimate for t finite, with more than 95% of returns. $a = 0.005$ µm (everything can be renormalized by this length), $\beta = 1.05$ (parameter of the discretization of the integral (see appendix 8.1.1)) (except $\kappa^{-1} = 0.1$ µm : $\beta = 1.1$), α (see appendix 8.1.1) such as $f(\alpha)\alpha \simeq 10^{-5}$.

In the regime $\kappa^{-1} \ll a$, we assume that the median is $z_{med} \ll a$. We can then make a large a expansion :

$$C(|z|) = \frac{2}{\pi}\int_0^\infty \frac{sin(kz)}{k}\frac{\kappa}{\kappa + k}dk. \tag{82}$$

The median z_{med} being defined by :

$$\frac{2}{\pi}\int_0^\infty \frac{sin(kz_{med})}{k}\frac{\kappa}{\kappa+k}dk = \frac{1}{2}, \tag{83}$$

it leads to $z_{med} = \kappa^{-1}\exp(g)$, with g solution of $4ge^g - 4e^g + 4\gamma e^g - \pi e^{2g} + \pi = 0$ (γ being the Euler constant). Therefore, $z_{med} \simeq 0.535\kappa^{-1}$. Since $\kappa^{-1} \ll a$, the assumption $z_{med} \ll a$ is verified. This expression works very well (see figure 40).

We now study the regime $\kappa^{-1} \gg a$. We define $u = kz$, $z^* = z/a$ and $\kappa^* = \kappa a$. The probability that a return is smaller than z is

$$C(|z|) = \frac{2}{\pi}\int_0^\infty \frac{sin(u)}{u} f(u)du, \tag{84}$$

with f defined as follows :

$$f(u) = \frac{1}{1 + \frac{u}{\kappa^* z^*}\frac{K_1\left(\frac{u}{z^*}\right)}{K_0\left(\frac{u}{z^*}\right)}}. \tag{85}$$

We suppose that the median is large compared to a, thus we can restrict the study to $z^* \gg 1$, and additionally $\kappa^* \ll 1$. Under these assumptions :

$$\frac{K_1\left(\frac{u}{z^*}\right)}{K_0\left(\frac{u}{z^*}\right)} \simeq \frac{u}{z^*}\frac{1}{\ln(2z^*/u) - \gamma} \simeq \frac{u}{z^*\ln(z^*)}. \tag{86}$$

Consequently :

$$\frac{u}{\kappa^* z^*}\frac{K_1\left(\frac{u}{z^*}\right)}{K_0\left(\frac{u}{z^*}\right)} \simeq \frac{1}{\kappa^*\ln(z^*)}, \tag{87}$$

$$C(|z|) = \frac{2}{\pi}\int_0^\infty \frac{sin(u)}{u} f(u)du \simeq \frac{1}{(\kappa^*\ln(z^*))^{-1} + 1}. \tag{88}$$

The median (defined by $C(|z_{med}|) = 1/2$) is $\kappa^*\ln(z^*_{med}) = 1$, which is equivalent to :

$$z_{med} \simeq a\exp\left(\frac{1}{\kappa a}\right). \tag{89}$$

This expression does not work perfectly well (see figure 40), but the scaling is indeed exponential. Problems in the estimate comes from the numeric solution which is calculated for up to 95% of the returns. We shall see in the following (see equation 95) the generalization of equation 89 : if the proportion p of the hops are longer than z_p, $z_p \propto a\exp\left(\frac{\frac{1}{p}-1}{\kappa a}\right)$. Assuming that the returns taking the longest time are also the larger returns, $p = 0.525$, leading to $(\frac{1}{p} - 1) \simeq 0.9$: it explains the observed shift.

To summarize, when the enzyme starts from the DNA ($r_0 = a$), the hop length median is $\simeq 0.53\kappa^{-1}$ when $\kappa^{-1} \ll a$, and $\propto a\exp(1/(\kappa a))$ when $\kappa^{-1} \gg a$. It gives another physical interpretation of the parameter κ originally defined by equation (64).

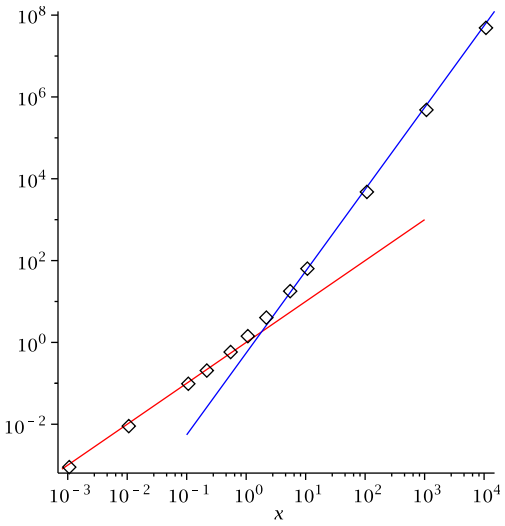


Figure 41: Median in the regime $\kappa^{-1} \to 0$, $r_0 > a$: $log(z_{med}/a)$ as a function of $log(r_0/a - 1)$. Numerical calculation ($\diamond$), fit $z_{med} = r_0 - a$ (red line), fit $z_{med}/a = 0.55(r_0/a)^2$ (blue line).

Regime $\kappa^{-1} \to 0$ and $r_0 > a$

When $\kappa^{-1} \to 0$ and $r_0 > a$, the probability that a return is smaller than z is :

$$C(|z|) = \frac{2}{\pi} \int_0^\infty du \frac{sin(u)}{u} \frac{K_0\left(u\frac{r^*}{z^*}\right)}{K_0\left(\frac{u}{z^*}\right)}, \tag{90}$$

with $z^* = z/a$ and $r^* = r_0/a$.

When $(r_0 - a) \ll a$:

$$C(|z|) \simeq \frac{2}{\pi} \int_0^\infty du \frac{sin(u)}{u} \exp\left(-u\frac{r_0 - a}{z}\right). \tag{91}$$

When $z = z_{med}$, $C(z_{med}) = 1/2$, leading to $z_{med} \simeq (r_0 - a)$. It is logical because for small distances, the enzyme does not see that the target surface is a cylinder and not a plane, and thus the typical traveled distance is $(r_0 - a)$ when returning to the cylinder.

When $r_0 \gg a$, we develop $C(z)$ assuming $z^* \gg r^* \gg 1$. We obtain : $C(|z|) = 1 - \frac{log(r^*)}{log(z^*)}$. z_n is the length with only $1/n$ of the returns larger than z_n (z_2 is the median) : $z_n/a = (r_0/a)^n$.

These expressions are in good agreement with numerical data (see figure 41).

Full expression : κ finite and $r_0 > a$

We start from : $C(|z|) = \frac{2}{\pi} \int_0^\infty \frac{sin(u)}{u} f(u)$, with $f(u) = \frac{K_0(ur_0/z)}{K_1(ua/z)u/(\kappa z) + K_0(ua/z)}$.

- $z_{med} \ll a$ if $r_0 - a$ and κ^{-1} are small compared to a. If $r_0 - a \ll \kappa^{-1}$, the typical hop length will be κ^{-1}, and reversely.

- If $z_{med} \gg a$, we can make the assumption $z \gg a$ to simplify $f(u)$:

$$f(u) \simeq \frac{log\left(\frac{z}{r_0}\right)}{\frac{1}{\kappa a} + log\left(\frac{z}{a}\right)}, \tag{92}$$

and $f(u)$ does not depend on u anymore.

$$C(z) \simeq \frac{log\left(\frac{z}{r_0}\right)}{\frac{1}{\kappa a} + log\left(\frac{z}{a}\right)}, \tag{93}$$

and as z_{med} is defined as $C(z_{med}) = 1/2$:

$$z_{med} = \frac{r_0^2}{a} exp\left(\frac{1}{\kappa a}\right). \tag{94}$$

Results are consistent with the limits $\kappa \to \infty$ or $r_0 \to a$. z_n, the length such as $1/n$ of the returns are larger than this length is :

$$\frac{z_n}{a} = \left(\frac{r_0}{a}\right)^n exp\left(\frac{n-1}{\kappa a}\right). \tag{95}$$

To summarize, the mean distance of return is infinite, but we have given the median of the distribution of return length, that is more representative of the typical returns. Doing so, we have found another physical interpretation of κ.

4.1.3 Hops and jumps : comparison with single molecule experiments

We now adapt the previous hop distribution (calculated for an infinite cylinder in an infinite space and for an infinite observation time) to the specificities of single-molecule experiments. These experiments are made by the optics and biology team, LKB, ENS, Paris (Pierre Desbiolles, Andreas Biebricher, Isabelle Bonnet and Natacha Porté). We first present the experiments, their main results and the reason why a theoretical model for hops is needed. Then we describe how the previous model for hop distribution is adapted. Finally, we compare with experimental results and show that the observations are indeed compatible with a combination of sliding and hopping.

Experiment

The experiment is aimed at observing directly the interaction between an enzyme and DNA molecules. This direct observation is made possible by single-molecule techniques.

Experimental set-up

EcoRV restriction enzymes are modified so as to be fluorescent. DNA molecules are biotinylated at their extremities, to bind to the surface, which is coated with streptavidin. Indeed, biotin and streptadivin form strong bonds. DNA is introduced

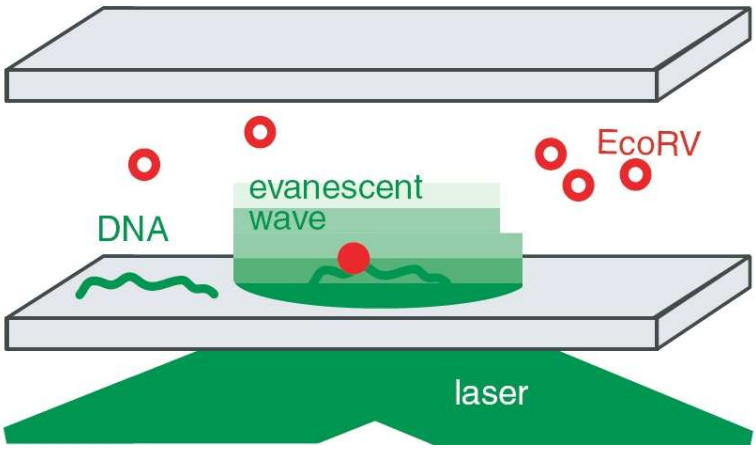


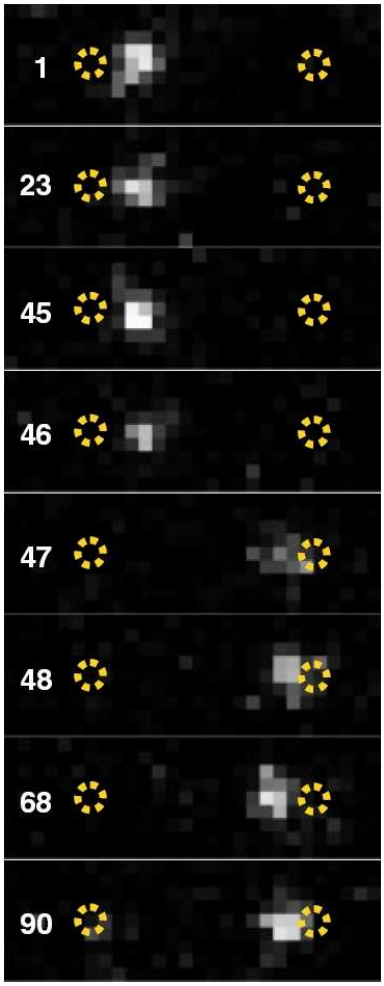
Figure 42: Experimental set-up. Biotinylated DNA molecules are attached at both ends to a streptavidin-coated surface. The molecules are in an elongated conformation, but free to fluctuate. Fluorescent proteins are visualized using Total Internal Reflection Fluorescence Microscopy (TIRFM).

in the DNA chamber first without flow, for having DNA molecules bound by one of their extremities to the surface. Then, a flow is applied, and the velocity of the flow determines the typical distance at which the other end of the DNA will attach. The DNA extension rate is about 70%, so that a DNA molecule is elongated enough for observations, but at the same time free to fluctuate, and with not too important influence of the surface. The distance between the attached ends is $L = 2.2$ µm on average. Due to a surface treatment by a blocking reagent, the interaction between enzymes and the surface is weak, so that the surface can be considered as reflective. To visualize proteins, a laser light is pointed to the surface at the reflection angle, so that only evanescent light enters the sample (see figure 42). Only enzymes close to the surface (distance of the order of magnitude of the wavelength (532 nm)) can be seen, thus enzymes in interaction with DNA are observed without the fluorescent background of all the enzymes in solution. Images are acquired at a rate of 20 ms, and the position of the enzyme is fitted with a Gaussian, leading to a maximum precision of ~ 30 nm.

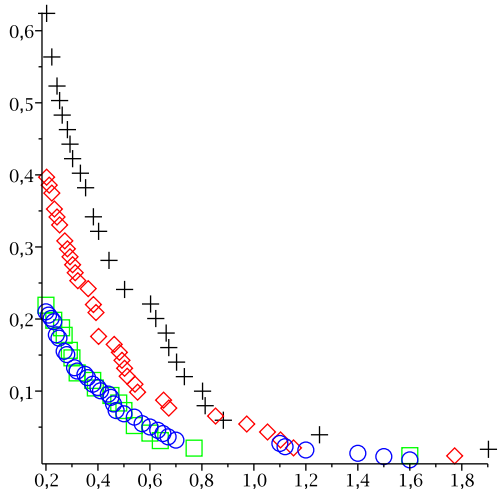
Results

Before presenting the results, we shall define the terms used afterwards. The enzyme is observed where the DNA is located during several frames. It is what we shall call "interactions", or more precisely "effective interactions", as the enzyme may combine sliding, when the enzyme is really in interaction with DNA, and small hops, with no interaction with DNA, but that cannot be distinguished from sliding due to the limited spatial resolution. The effective interaction, for example when studying the large changes in position, is ended if the observed enzyme is not seen during two frames (it ensures that the probability of confounding another enzyme with the initial one is very low). When studying the effective diffusion coefficient, an effective interaction is also ended if the enzyme's position changes more than $z_m = 200$ nm in $t_{obs} = 40$ ms (two frames). z_m is chosen to be 200 nm, because this length is

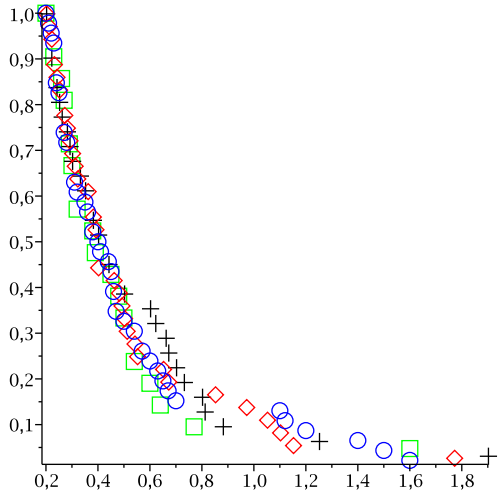
large compared to the typical length traveled by the effective 1-dimensional diffusion during t_{obs} (30 nm), and is also large compared to DNA fluctuations (70 nm). These large relocations (change of position larger than z_m) have to be explained by something else, as we shall detail in the following. We assume in our model that enzymes alternate between sliding and 3-dimensional excursions. When using the number of 3-dimensional excursions by effective interaction, we also count the last excursion which ends the effective interaction. Now that we have defined the terms used, we can describe what is observed in the experiment.



Images of the experiment. Frames are numerated, and this number is proportional to time. The white disk is the enzyme. The yellow circle represents the extremities of DNA. The distance between the extremities is $L = 2.2$ µm.



Number of relocations larger than z during an effective interaction as a function of z (µm).



Number of relocations larger than z renormalized by the number of relocations larger than $z_m = 200$ nm as a function of z (µm).

Figure 43: Experimental results : large relocations. *Left* Between frame 46 and 47 there is a large relocation. *Right* Distribution of the relocation larger than $z_m = 200$ nm in 1 or 2 frames (maximum observation time $t_{obs} = 40$ ms). Data (in PIPES buffer): $+$ = 10 mM NaCl, $\diamond$ = 20 mM NaCl, $\square$ = 40 mM NaCl, $\circ$ = 60 mM NaCl.

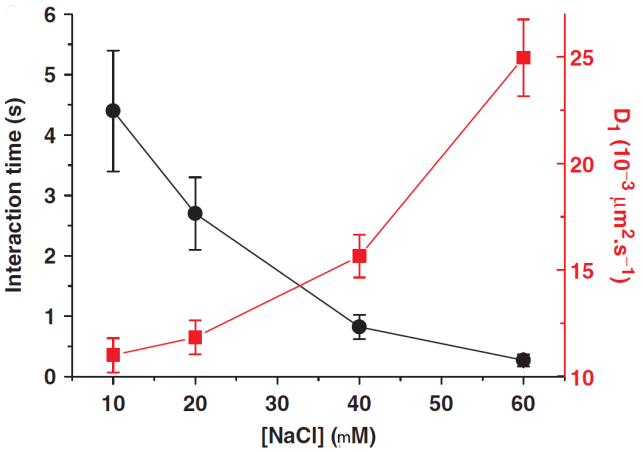


Figure 44: In PIPES buffer, measurements are made at different salt concentration. The efficient 1D diffusion coefficient (■) was measured, along with the mean effective interaction time (●).

Two main observations are made in these experiments. On the one hand, enzymes interact with DNA. The durations of effective interactions are exponentially distributed. Motion observed is compatible with 1D diffusion along the DNA, with an effective coefficient of typical magnitude $D_{1D}^{\text{eff}} \simeq 10^{-2}$ µm^2.s^{-1} (see figure 44), much smaller than the 3D diffusion coefficient $D_{3D} \simeq 50$ µm^2.s^{-1} (measured separately by fluctuation correlation spectroscopy). On the other hand, relocations on the DNA larger than z_m (for example between frames 46 and 47 in the left of figure 43) are observed within t_{obs}, in a proportion much higher than expected with the 1D diffusion alone. Indeed, except for these large relations, the 1D steps are Gaussian distributed, with a variance $\sigma = 30$ nm, and $z_m > 6\sigma$, which for a Gaussian distribution results in an extremely low probability. The concentration of enzymes is low enough, so that the event of a second enzyme rebinding within t_{obs} after the first enzyme leaves is highly improbable, and cannot explain these relatively frequent large relocations. The distribution of these large relocations is determined (figure 43 right). Additionally, experiments are made with different salt concentrations : increasing the salt concentration decreases the interaction time, and increases the diffusion coefficient (see figure 44).

Questions raised

During t_{obs}, the typical distance traveled along the DNA axis for 1D diffusion is $\sqrt{2D_{1D}^{\text{eff}} t_{obs}} \simeq 30$ nm, whereas the typical distance covered by free 3D diffusion is $\sqrt{2D_{3D} t_{obs}} \simeq 2$ µm. The first question question is to determine whether the observed distribution of large relocations is compatible with free 3D diffusion. The second question is to determine to which extent 3D relocations which are too short to be distinguished impact on the effective interactions, which in this picture are actually a combination of real 1D interaction (sliding), and of small 3D relocations.

Distribution of the durations of the effective 1D interactions

The observed distribution of the durations of these effective interactions is exponential. If one dimensional effective interactions are a combination of real 1D diffusion (sliding) and of small hops that cannot be distinguished, the first step is to check that an exponential distribution of durations is compatible with a combination of sliding and hopping.

Assuming standard bimolecular association, the distribution of sliding durations is exponential : $p_{1D}(t) = \lambda_1 \exp(-t\lambda_1)$, with the mean duration of a sliding phase $\tau_{1D} = 1/\lambda_1$. For each 3D excursion, the probability that the enzyme does not go back on DNA before t_{obs} is q, assumed to be fixed, *i.e.* not to vary between excursions. The probability P_n that an effective interaction has n hops (including the hop when the enzyme does not return before t_{obs}) is :

$$P_n = q(1-q)^{n-1}. \tag{96}$$

The mean duration of a 3D excursion τ_{3D} is assumed to be much smaller than the mean duration of a sliding phase τ_{1D}. Indeed, $D_{3D} \simeq 50\ \mu\text{m}^2.\text{s}^{-1} \gg D_{1D}^{\text{eff}} \simeq 0.01\ \mu\text{m}^2.\text{s}^{-1}$: if τ_{1D} and τ_{3D} were of the same order of magnitude, it would not be possible to have such a low D_{1D}^{eff}. The Fourier transform of p_{1D} is:

$$\varphi_{1D}(u) = \int_{\Re} e^{iut} p_{1D}(t)dt = \frac{1}{1-\frac{iu}{\lambda_1}}. \tag{97}$$

The Fourier transform of the probability that an effective interaction having n hops lasts t is :

$$\varphi(u,n) = (\varphi_{1D}(u))^n. \tag{98}$$

The Fourier transform of the probability that an effective interaction lasts t is :

$$\varphi(u) = \sum_{n=0}^{\infty} P_n \varphi(u,n) = \sum_{n=0}^{\infty} q(1-q)^{n-1} \left(\frac{1}{1-\frac{iu}{\lambda_1}}\right)^n = q\frac{1}{1-\frac{1-q}{1-\frac{iu}{\lambda_1}}} = \frac{1}{1-q}\left(\frac{q}{q-\frac{iu}{\lambda_1}} - \frac{\frac{iu}{\lambda_1}}{1-\frac{iu}{q\lambda_1}}\right). \tag{99}$$

Inverting the Fourier transform, the probability distribution of duration of effective interactions writes :

$$P(t) = \lambda_1 \frac{q}{1-q} e^{-\lambda_1 qt} + \frac{1}{\lambda_1}\frac{d}{dt}\left(\lambda_1 \frac{q}{1-q} e^{-\lambda_1 qt}\right) = \lambda_1 q e^{-\lambda_1 qt} \tag{100}$$

This distribution is exponential, as observed experimentally. However, it is not very discriminating : having a fixed loss probability by hop is a sufficient condition, and does not say anything about the hop length distribution. The distribution of effective interaction durations is compatible with a combination of sliding and hopping, but we have to go further. We next adapt our model of hops to the finite time and specific geometry of the experiment.

Finite time

The relocations observed in the experiment are hops (and not jumps), because in the set-up the DNA is elongated : the starting point always matters. We will

take $r_0 = a$, as the enzyme starts a hop by unbinding from DNA. However, we have to adapt our model of hops to the experiment.

First, we have calculated the hop distribution waiting an infinite time for returns, whereas here the observation time t_{obs} is finite. We write the hop length distribution for a finite time :

$$P_{t_{obs}}(z) = \int_0^{t_{obs}} P_{\parallel}(z,t) F_{\perp}(t|\mathbf{a}) \mathrm{d}t, \tag{101}$$

where $P_{\parallel}(z,t) = (4\pi D_{3D} t)^{-1/2} \exp(-z^2/4D_{3D}t)$ represent the 3D diffusion along the DNA axis; $F_{\perp}(t|\mathbf{a})$ is the first passage density to DNA in the orthogonal plane, obtained by the numerical inversion of its Laplace transform, known analytically (with $x = a\sqrt{s/D}$) [Redner, 2001] :

$$\widehat{F}_{\perp}(s|\mathbf{a}) = \frac{K_0(x)}{K_0(x) + (\kappa a)^{-1} K_1(x)}, \tag{102}$$

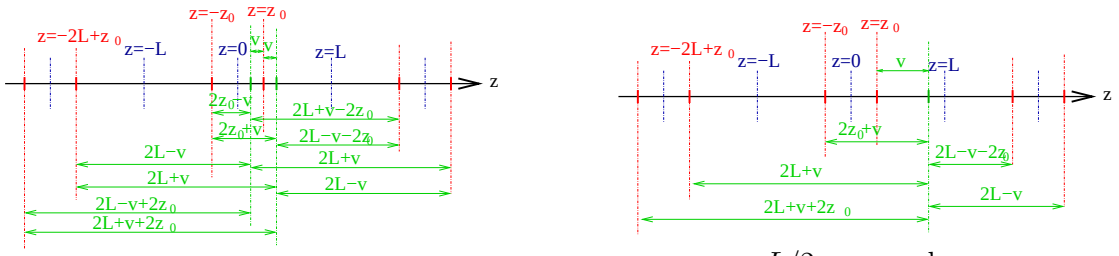
For details on the numerical inversion, see appendix 8.1.1. Note that the distribution given equation 101 is not normalized : $1 - \int_{-\infty}^{\infty} P_{t_{obs}}(z) dz$ give the proportion of 3D excursions that are "losses" (in the sense that they end effective interactions).

Geometry

Figure 45: Comparison between experimental and model geometries. Top : experimental geometry : DNA elongated at 70%, bound by its extremities to a reflective surface. Bottom : model geometry : a cylinder between reflective planes.

Second, we have previously considered an infinite cylinder in an infinite space, whereas here elongated DNA is finite, stuck by its ends on a surface (DNA ends are separated by $L \simeq 2.2$ µm). This surface can be considered as reflective, as non-specific interactions with the proteins are weak. To simplify the problem, we

use an effective geometry which retains the main features of the real geometry, while making analytical calculations tractable. We model the DNA as a cylinder of radius a and length L between two reflective planes perpendicular to the cylinder axis (see figure 45) (for other possible choices and why they are not adequate, see appendix 8.1.2). At the extremities, this geometry locally represents well that the ends are stuck to a surface. As t_{obs} is small enough, multiple reflections on the planes are unlikely ($L_{typ} = \sqrt{2D_{3D}t_{obs}} \simeq 1.4$ µm, $L = 2.2$ µm), thus we expect this approximation to be valid.



case $L/2 > z_0 > v$ | case $L/2 > z_0$ and $v > z_0$

Figure 46: Methods of images : a finite cylinder between reflective planes with one source (in $z = z_0$) is equivalent to an infinite cylinder with an infinity of mirrored sources.

Reflective boundaries at $z = 0$ and $z = L$ with a source (the starting point) at $z = z_0$ are equivalent to an infinite cylinder with sources in $z = z_0 + 2nL$ and in $z = -z_0 + 2nL$, with $n \in \mathbb{Z}$ (methods of images) (see figure 46).

We note $P(v)$ the probability of a hop of size v with an infinite cylinder, and $P^*(v|z_0)$ the probability of a hop of size v starting from z_0 with reflective boundaries. The protein has a probability $1/2$ of making a hop on the left and $1/2$ of making a hop on the right. When $L/2 > z_0 > v$ (cf. figure 46 left) :

$$\begin{aligned} P^*(v|z_0) =& P(v) + \sum_{n=1}^{\infty} \left(P(2nL - v) + P(2nL + v)\right) + \frac{1}{2}\left(P(2z_0 + v) + P(2z_0 - v)\right) \\ &+ \frac{1}{2}\sum_{n=1}^{\infty}[P(2nL + v - 2z_0) + P(2nL - v - 2z_0) \\ &+ P(2nL + v + 2z_0) + P(2nL - v + 2z_0)]. \end{aligned} \tag{103}$$

When $L/2 > z_0$ but $v > z_0$ (cf. figure 46 right) :

$$\begin{aligned} P^*(v|z_0) =& \frac{1}{2}\Big[P(v) + P(2z_0 + v) + \\ &\sum_{n=1}^{\infty}\left(P(2nL - v) + P(2nL + v) + P(2nL - v - 2z_0) + P(2nL + v + 2z_0)\right)\Big]. \end{aligned} \tag{104}$$

$P^*(v) = \frac{1}{L}\int_0^L P^*(v|z_0)dz_0 = \frac{2}{L}\int_0^{L/2} P^*(v|z_0)dz_0$ is the probability of making a hop of size v starting from a random position of the cylinder, with uniform probability.

Because of symmetry, starting from a random point $z_0 \in [0, L/2]$ is equivalent. $P^*(v)$ can be written as a infinite sum :

$$P^*(v) = \frac{1}{L}\sum_{n=0}^{\infty} W_n, \tag{105}$$

with :

$$W_0 = \frac{L}{2}P(v) + \left(\frac{L}{2} - v\right)P(v) + \int_0^{L/2} P(2z_0 + v)dz_0 + \int_v^{L/2} P(2z_0 - v)dz_0, \tag{106}$$

$$W_0 = (L - v)P(v) + \frac{1}{2}\int_v^{L+v} P(x)dx + \frac{1}{2}\int_v^{L-v} P(x)dx, \tag{107}$$

$$W_0 = (L - v)P(v) + \int_v^{L-v} P(x)dx + \frac{1}{2}\int_{L-v}^{L+v} P(x)dx, \tag{108}$$

and :

$$\begin{aligned} W_n = \; & \tfrac{L}{2}\left(P(2nL - v) + P(2nL + v)\right) + \left(\tfrac{L}{2} - v\right)\left(P(2nL - v) + P(2nL + v)\right) \\ & + \int_0^{L/2}\left(P(2nL - v - 2z_0) + P(2nL + v + 2z_0)\right)dz_0 \\ & + \int_v^{L/2}\left(P(2nL + v - 2z_0) + P(2nL - v + 2z_0)\right)dz_0, \end{aligned} \tag{109}$$

$$\begin{aligned} W_n = \; & (L - v)\left(P(2nL - v) + P(2nL + v)\right) \\ & + \tfrac{1}{2}\left(\int_{(2n-1)L-v}^{2nL-v} P(x)dx + \int_{2nL+v}^{(2n+1)L+v} P(x)dx + \int_{(2n-1)L+v}^{2nL-v} P(x)dx + \int_{2nL+v}^{(2n+1)L-v} P(x)dx\right). \end{aligned} \tag{110}$$

It can be noticed that :

$$\begin{aligned} \sum_{n=1}^{\infty} W_n = & (L - v)\sum_{n=1}^{\infty}\left(P(2nL - v) + P(2nL + v)\right) + \int_{L-v}^{\infty} P(x)dx \\ & - \frac{1}{2}\int_{L-v}^{L+v} P(x)dx - \sum_{n=1}^{\infty}\left(\int_{-v}^{v} P(2nL + x)dx\right), \end{aligned} \tag{111}$$

leading to :

$$P^*(v) = \frac{1}{L}\left(\int_v^{\infty} P(x)dx + (L - v)\left(\sum_{n=0}^{\infty} P(2nL + v) + \sum_{n=1}^{\infty} P(2nL - v)\right) - \sum_{n=1}^{\infty}\int_{-v}^{v} P(2nL + x)dx\right) \tag{112}$$

Consequently, $P^*(v)$ can also be written as another infinite sum :

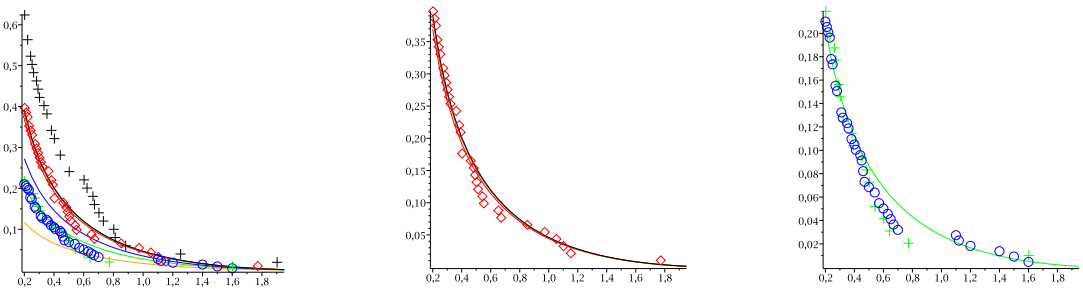
$$P^*(v) = \frac{1}{L}\sum_{n=0}^{\infty} Y_n, \tag{113}$$

with :

$$Y_0 = \int_v^{\infty} P(x)dx + (L - v)P(v), \tag{114}$$

$$Y_{n>0} = (L - v)\left(P(2nL + v) + P(2nL - v)\right) - \int_{2nL-v}^{2nL+v} P(x)dx. \tag{115}$$

In practice, the formula (113) is more convenient that the form (105), and for both the first 2 terms of the sum are sufficient (see appendix 8.1.2).



all [NaCl]=20 mM [NaCl]=40-60 mM

Figure 47: Probability of observing a hop larger than z during an effective interaction as a function of z (µm). Experimental data (in PIPES buffer) (symbols) : + = 10 mM NaCl, ◇ = 20 mM NaCl, □ = 40 mM NaCl, ○ = 60 mM NaCl. Solid lines: distribution from our model for $\kappa^{-1} = 0.5$ nm (black), $\kappa^{-1} = 1$ nm (sienna), $\kappa^{-1} = 2$ nm (red), $\kappa^{-1} = 10$ nm (blue), $\kappa^{-1} = 20$ nm (green), $\kappa^{-1} = 50$ nm (orange).

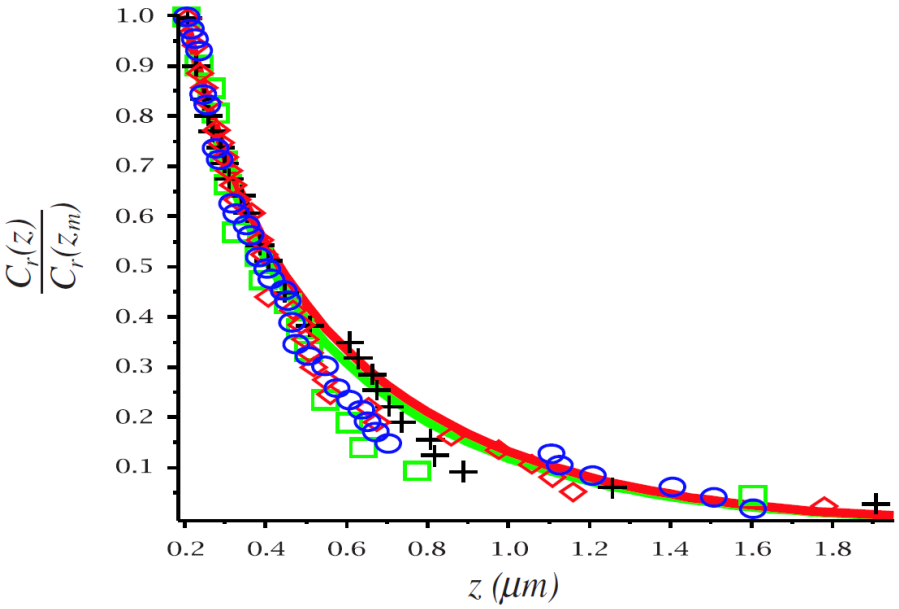


Figure 48: Cumulative distribution of hop length z (μm) normalized by the number of hops larger than $z_m = 200$ nm. Data (in PIPES buffer): + = 10 mM NaCl, ◇ = 20 mM NaCl, □ = 40 mM NaCl, ○ = 60 mM NaCl. Solid lines: distribution from our model for $\kappa^{-1} = 0.5$ nm (green), $\kappa^{-1} = 20$ nm (red).

Distribution of large hops

Now we have adapted the distribution both to a finite observation time and to the specific experimental geometry, we can compare the experimental distribution of large hops to what we expect from 3D diffusion.

The large hops distribution have been obtained for different salt concentrations in the PIPES buffer. Biochemical studies reported only a slight dependence on [NaCl] of the binding rate to DNA [Lohman, 1986]. Thus, the hop distribution should depend only weakly on [NaCl]. The data at 40 mM and 60 mM cannot be distinguished, but the number of large relocations by interaction is higher with lower salt concentration (see figure 47). To fit these results, we use $C^*_{t_{obs}}(z) = \int_z^\infty P^*_{t_{obs}}(z)dz$.

The number of large relocations is measured for an effective interaction (not for a 3D excursion), thus we have to renormalize $C^*_{t_{obs}}(z)$ by the mean number of 3D excursions by effective interaction, $N = (1 - C^*_{t_{obs}}(0))^{-1}$: we use $NC^*_{t_{obs}}(z)$ to fit the curves of figure 47. We fit varying the parameter κ : $\kappa^{-1} \simeq 20$ nm for the 40 mM and 60 mM salt points and with $\kappa^{-1} \leq 2$ nm (the values for $\kappa^{-1} \leq 2$ nm are confounded) for the 20 mM salt points. The 10 mM salt value could not be fitted. However on the one hand, experimental curves are renormalized. Indeed, the fluorophore bleaches, *i.e.* it enters a non-fluorescent state, with a probability per unit time proportional to light intensity. Consequently, a part of the trajectories ending is not caused by the enzyme leaving, but by the enzyme becoming invisible. The curves are renormalized for this effect, and this correction is only roughly estimated in particular for the 10 mM case, as the interaction time is longer. On the other hand, we have an effective geometry which may not give a completely accurate estimate of N. Thus, a completely quantitative fit seems out of reach. However, the order of magnitude of the number of jumps by effective interaction is consistent with the model.

Furthermore, the z dependence of the relocation length distribution can be precisely accounted for by the model by considering the normalized distribution $C^*_{t_{obs}}(z)/C^*_{t_{obs}}(z_m)$. The shape of the distribution is very good. Indeed, after normalization by the total number of observed relocations, the cumulative distributions of the relocation lengths at different salt concentrations collapse (see figure 48). We compare the experimental curve with the theoretical prediction $C^*_{t_{obs}}(z)/C^*_{t_{obs}}(z_m)$. Since the latter depends weakly on κ, it is not possible to evaluate quantitatively this parameter. As the adjustable parameter κ is not really adjustable, we obtain a good agreement between data and the model with no real fitting parameter (see figure 48).

Effective diffusion coefficient

We have shown that the large relocations are compatible with free 3D diffusion. Thus we now suppose that the observed diffusion coefficient D^{eff}_{1D} is a combination of real 1D diffusion (sliding), and hops too small to be distinguished ($< z_m$). The question is how these hops quantitatively influence the observed diffusion coefficient. This coefficient is measured at different salt concentration. As we have just seen, the large relocation distribution depends only weakly on [NaCl], as was expected [Lohman, 1986]. We will then suppose that the hop distribution does not change with salt concentration. On the contrary, the time spent bound to DNA (sliding), decreases with the salt concentration. Indeed, the mean effective interaction time τ decreases when the salt concentration increases.

Let us suppose there are $n-1$ hops during a trajectory (n 3D excursions, leading to $n-1$ returns and 1 loss) :

$$2D^{\text{eff}}_{1D}\left(\sum_{i=1}^{n} t_i^{1D} + \sum_{j=1}^{n-1} t_j^{3D}\right) = \left(\sum_{i=1}^{n} \epsilon_i l_i^{1D} + \sum_{j=1}^{n-1} \epsilon_j l_j^{3D}\right)^2, \tag{116}$$

with $\epsilon_k = \pm 1$. Averaging over the different trajectories, and with N the mean number of 3D excursions by effective interaction, we can write :

$$2D^{\text{eff}}_{1D}\left(N\tau^{1D} + (N-1)\tau^{3D}\right) = N\left\langle l^2_{1D}\right\rangle + (N-1)\left\langle l^2_{3D}\right\rangle, \tag{117}$$

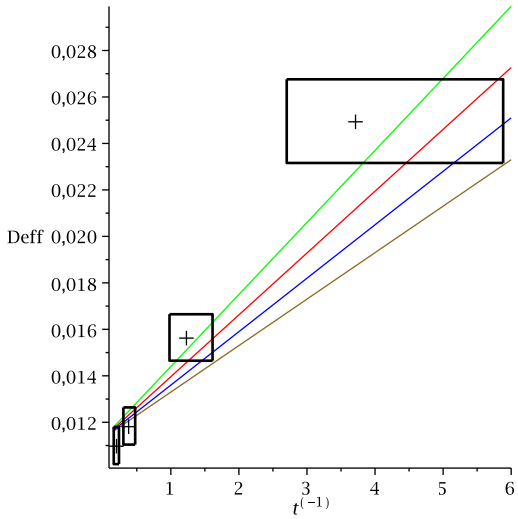


Figure 49: D_{1D}^{eff} ($\mu m^2.s^{-1}$) as a function of $1/\tau$ (s^{-1}) (τ is the mean effective interaction time). Experimental data (+) with error bars (□). 10 mM point is bottom left and 60 mM is top right. Formula (121) for $\kappa^{-1} = 0.5$ nm (green), $\kappa^{-1} = 2$ nm (red), $\kappa^{-1} = 5$ nm (blue), $\kappa^{-1} = 5$ nm (sienna).

with τ^{iD} the mean duration of a phase in i dimension. It is reasonable to assume $\tau^{3D} \ll \tau^{1D}$. Indeed $D_{3D} \sim 50\ \mu m^2.s^{-1} \gg 0.01\ \mu m^2.s^{-1} \sim D_{1D}$, and if $\langle l_{1D}^2 \rangle$ is larger or of the same order of magnitude than $\langle l_{3D}^2 \rangle$, then it is sure that $\tau^{3D} \ll \tau^{1D}$. We obtain :

$$2D_{1D}^{\text{eff}}\tau^{1D} = 2D_{1D}\tau^{1D} + \left(1 - \frac{1}{N}\right)\langle l_{3D}^2 \rangle, \tag{118}$$

and assuming $N \gg 1$:

$$D_{1D}^{\text{eff}} \simeq D_{1D} + \frac{\langle l_{3D}^2 \rangle}{2\tau^{1D}}, \tag{119}$$

$$\frac{D_{1D}^{\text{eff}}}{D_{1D}} = 1 + \frac{\langle l_{3D}^2 \rangle}{2D_{1D}\tau^{1D}}. \tag{120}$$

In terms of τ the mean effective interaction time, and $N-1$ the estimated number of hops returning at < 200nm during the observation time (the N^{th} is the hop that ends the effective interaction), we obtain :

$$\frac{D_{1D}^{\text{eff}}}{D_{1D}} = 1 + \frac{(N-1)\langle l_{3D}^2 \rangle}{2D_{1D}\tau}. \tag{121}$$

We use this expression for figure 49 : D_{1D} is a fit parameter, and for a given κ we calculate from our hop model $(N-1)\langle l_{3D}^2 \rangle$. We can fit the diffusion coefficients for $\kappa^{-1} \leq 5$ nm. It gives an upper bound for this parameter. With this expression, we can explain the two-fold variation of the effective diffusion coefficient with the salt concentration.

Conclusion

To summarize, to interpret quantitatively single-molecule experiments, we have adapted a simple model to both a finite time and a specific geometry. First, the large relocations are compatible with 3-dimensional free diffusion excursions. Then, the distribution of effective interaction durations is compatible with a combination of hop and sliding. Finally, the dependence of the effective diffusion coefficient on the effective interaction duration is also accounted for by a combination of sliding and hopping. All these elements strongly support that EcoRV combines sliding and hopping, and that hopping description as free 3D diffusion is relevant.

4.1.4 Hops and jumps : subdiffusion

Motivation

The previously studied experiment is in vitro, in water with few other components, where the 3D excursions can be considered as Brownian diffusion. However, in vivo, the cellular environment is crowded, and the dynamics of tracers molecules often departs from usual diffusion [Tolic-Norrelykke et al., 2004, Golding and Cox, 2006]. Anomalous diffusion modifies importantly the kinetics on transport limited reactions [Metzler and Klafter, 2000, Lomholt et al., 2007, Condamin et al., 2007, 2008]. We present here how it also affects the hop distribution.

Subdiffusion behavior is characterized by a mean square displacement that scales as $\langle \Delta \mathbf{r}^2 \rangle \sim t^{2/d_w}$, where $d_w > 2$ is defined as the walk dimension [Metzler and Klafter, 2000]. There are several microscopic mechanisms that can explain such a scaling. We present here two possibilities :

- subdiffusion could come from waiting times. Continuous time random-walks (CTRWs) models [Metzler and Klafter, 2000], in which at each step the walker can be trapped for a long period of time, mimic the cage effect in crowded environments
- subdiffusion could come from spatial inhomogeneities. Bancaud et al. [2009] describe chromatin as a fractal medium. For example, on critical percolation clusters [Ben-Avraham and Havlin, 2000], the numerous dead-ends make random walks subdiffusive.

Which mechanisms are at play inside biological cells is still discussed [Condamin et al., 2008].

These different models of microscopic mechanisms lead to the same mean square displacement, but not to the same hop length distribution. In the case of subdiffusion stemming from CTRW, it is only the time which is affected, not the geometry of the trajectories. Thus the relocation length distribution will not change : Eq. (77) and its extensions remain valid in this case. On the contrary, in the case of fractal type subdiffusion, the geometry of trajectories is changed by crowding, as we shall show quantitatively in the following.

Fractal crowding

After dissociating from DNA, a protein walks on a fractal embedded in 3D space. We assume again that transport along the DNA axis is independent from the transport in the direction perpendicular to the cylinder. Similarly to Eq. (101), the distribution of relocation length writes $P(z|\mathbf{r_0}) = \int_0^\infty P_\parallel(z,t)F_\perp(t|\mathbf{r_0})\mathrm{d}t$, where $P_\parallel$ stands for the longitudinal propagator, characterized by d_w the walk dimension and $d_f^\parallel$ the dimension of the projection of the fractal on the axis parallel to the cylinder, and $F_\perp$ is the transverse first passage time density to DNA, characterized by dimensions d_w and $d_f^\perp$.

Following O'Shaughnessy and Procaccia [1985], we assume that the diffusion current obeys a generalized Ficks's law. We thus write an effective transport operator, such as any diffusing quantity c is solution of :

$$\partial_t c = K\Delta_r c(r) \equiv \frac{K}{r^{d_f-1}}\frac{d}{dr}\left(r^{d_f-d_w+1}\frac{d}{dr}c(r)\right), \tag{122}$$

where K is the generalized diffusion coefficient. Note that for subdiffusion $d_w > 2 \geq d_f^\perp$: the exploration is termed compact [de Gennes, 1976] and we can take the limit $a \to 0$. After some algebra the first-passage-time density $F_\perp$ reads:

$$F_\perp(t|\mathbf{r_0}) = \frac{d_w^{-\nu_\perp} K r_0^{\nu_\perp d_w/2}}{\Gamma(\nu_\perp)}\int_0^\infty e^{-Kut}u^{\nu_\perp/2}J_{\nu_\perp}(\beta\sqrt{u})\mathrm{d}u, \tag{123}$$

where $\beta = 2r_0^{d_w/2}/d_w$ and $\nu_i = 1 - d_f^i/d_w$ (here $i = \perp, \parallel$). Using the propagator of O'Shaughnessy and Procaccia [1985], we get :

$$P(z|\mathbf{r_0}) = \alpha\int_0^\infty u^{(\nu_\perp-\nu_\parallel)/2}J_{\nu_\perp}(\beta\sqrt{u})K_{\nu_\parallel}(\gamma\sqrt{u})\mathrm{d}u, \tag{124}$$

with $\gamma = 2z^{d_w/2}/d_w$, and

$$\alpha = \frac{2d_w^{\nu_\parallel-\nu_\perp-1}r_0^{\nu_\perp d_w/2}z^{(d_f^\parallel+d_w)/2-1}}{\Gamma(\nu_\perp)\Gamma\left(1-\nu_\parallel\right)}. \tag{125}$$

Comparison with numerical simulations

To check that our decoupling assumption of the longitudinal and transverse projections of motion is valid, we compare the analytical expression (124) of the relocation length distribution for subdiffusion in fractals to numerical simulations in the representative example of a percolation cluster embedded in 3D space (see figure 50) (for more details and other examples, see appendix 8.1.3). The large z behavior (using 124) is $P(z|\mathbf{r_0}) \sim r_0^{d_w-d_f^\perp}/z^{1+d_w-d_f^\perp}$, which always decays faster than the diffusion case ($\sim 1/z\ln^2(z/a)$) as illustrated in Fig. 50. The proportion of jumps in the case of random conformation of DNA in the fractal type crowding scales like $C(\xi) \sim \xi^{-d_w+d_f^\perp}$: it is much smaller than for regular diffusion, thus fractal crowding changes the overall search. If the enzyme and its target are colocalized (as suggested by Wunderlich and Mirny [2008], see section 4.1.1), our results further suggest that in this case crowding could be beneficial for target location since it enhances the local scanning of the DNA by reducing the proportion of jumps.

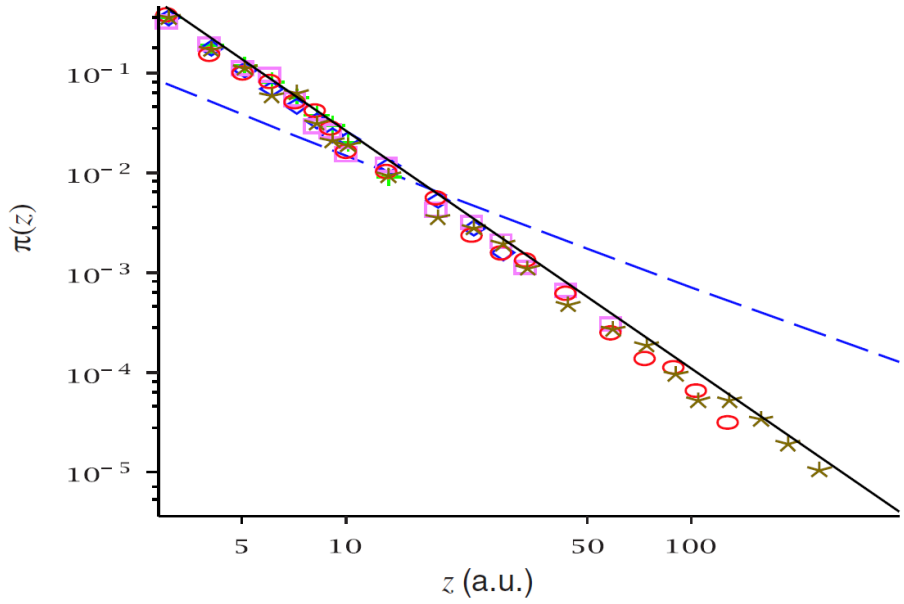


Figure 50: Distribution of the relocation length $P(z|\mathbf{r_0})$ for 3D excursions on a critical percolation cluster embedded in a 3D cubic lattice, for $r_0 = 1$ and $a = 0$. Simulations (normalized for $z \geq 2$) are performed for different system sizes to rule out finite size effects (symbols), and collapse on the theoretical curve (plain line) obtained from Eq.(124). Here $d_w^{\|} = d_w^{\perp} = 3.88...$, $d_f^{\perp} = 2$ and $d_f^{\|} = 1$ since $d_f = 2.53...$ [Ben-Avraham and Havlin, 2000]. The large z scaling follows $P(z|\mathbf{r_0}) \sim 1/z^{2.88...}$, and is compared to the case of normal diffusion (dashed line) with $r_0 = 1$ and $a = 0.1$.

4.1.5 Conclusion on DNA enzyme interaction

When searching for a specific DNA sequence, enzymes perform facilitated diffusion. Facilitated diffusion is intermittent : it is a combination of motion *(i)* in one dimension, in close interaction with DNA, called "sliding", which is the state in which the target is found; and *(ii)* 3D excursions, that change the global transport of the searcher. Development of new experimental techniques to study the classical facilitated diffusion raises new questions, that are addressed by recent models that we have reviewed. The relevant quantity to optimize is not necessarily the mean search time, as other quantities may matter more, depending on what is vital for the cell. Models with calculation of the whole probability distribution of the search time make possible to extract more information relevant to biology. To describe accurately what happens in biological cells, several modelings of the 1D and the 3D phases are proposed. The 1D phase ("sliding") is not necessarily diffusive. There is also a trade-off between motion and complex stability : a two-state sliding model have been proposed, with different protein conformations corresponding to search and recognition states. The 3D phase can be mainly separated between "jumping" (3D excursions with starting and ending positions uncorrelated in the DNA sequence) and "hopping" (correlated 3D excursions). As they have different properties for the search, knowing what proportion of 3D excursions are hops is important, along with the distribution of hop size.

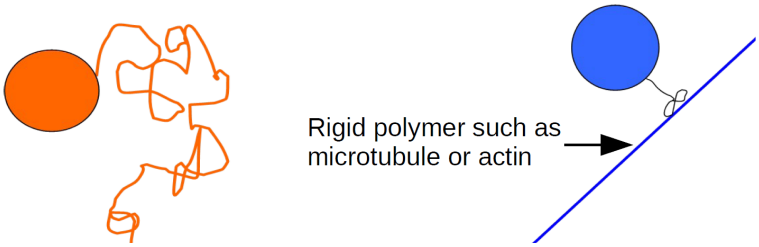
We first answer this question by calculating analytically the distribution of hops

for an infinite cylinder, for normal diffusion and an infinite observation time. These results can be applied to a coiled DNA. On the one hand, as DNA can be considered as a cylinder at small scales (smaller than the persistence length) the hop distribution is valid in this regime. On the other hand, as hops larger than the persistence length will actually be jumps if the DNA is coiled, we also obtain the jump proportion. We then adapt our results to interpret quantitatively a single molecule experiment : we obtain the distribution for a finite observation time, and in an effective geometry. Our results support that the interaction observed in the experiment is a combination of sliding and hopping, that can actually be modeled as free 3D normal diffusion. We extend the model to the case of subdiffusion. Indeed, cellular crowding can affect the results and has to be taken into account.

4.2 Active transport of vesicles in cells

After this first microscopic example of intermittent search, we turn to another example : active transport of vesicles reacting at specific locations in cells.

4.2.1 Active transport in cells



Does not need ATP
Distance$^2 \propto Dt$ (or less)
Diffusion coefficient decreases when the size of the object increases

Consumes ATP
Distance $= Vt$
Velocity ~ 1 µm.s^{-1}, almost independent from the object size (see for example Block et al. [2003])

Figure 51: Transport options for vesicles inside cells.

Various motor proteins such as kinesins or myosins are able to convert the chemical fuel provided by ATP into mechanical work by interacting with the semiflexible oriented filaments (mainly F–actin and microtubules) of the cytoskeleton [Alberts, 2002]. As many molecules or larger cellular organelles like vesicles, lysosomes or mitochondria, hereafter referred to as tracer particles, can randomly bind and unbind to motors, the overall transport of a tracer in the cell can be described as alternating phases of standard diffusive transport, and phases of active directed transport powered by motor proteins [Alberts, 2002, Salman et al., 2005] (see figure 51). In particular, Huet et al. [2006] studied the rate of transitions between ballistic, diffusive and "on the target" states of vesicles, and found that the vesicles studied are much more likely to react in the free diffusive phase than when bound to motors. Active transport is therefore clearly a further example of intermittent behavior. Active transport in cells has been extensively studied both experimentally, for instance by single particle tracking methods [Sheetz and Spudich, 1983, Howard et al., 1989], and theoretically by evaluating the mean displacement of a tracer [Shlesinger and Klafter, 1989, Ajdari, 1995], or stationary concentration profiles [Nedelec et al., 2001]. This transport is important for example for dynamically regulating the distribution of proteins such as membrane receptors. Appendix 8.4 shows an instance where activity of the receptor and transport are linked, a variation in the activity leading to changes in the receptor distribution , which has important functionality consequences.

Most of cell functions are regulated by coordinated chemical reactions which involve low concentrations of reactants (such as ribosomes or vesicles carrying targeted proteins), and which are therefore limited by transport. However, a general quantitative analysis of the impact of active transport on reaction kinetics in cells, and

more generally in generic active media, was missing, even if a few specific examples had been tackled (see for instance Holcman [2007]). We propose [Loverdo et al., 2008] (highlighted in a "News and views" of Nature Physics by Mirny [2008]) an analytical model based on the idea of intermittence which allows us to determine the kinetic constant of transport limited reactions in active media, and show that it can be optimized. We give the main results of the model, and postpone the technical details for section 5 (where they appear as a part of a more general model).

4.2.2 Model

The model relies on the idea of intermittent search strategies and has important similarities with the model of section 3.4, similarities discussed in section 5. We consider a tracer particle evolving in a d–dimensional space (in practice $d = 1, 2, 3$) which performs thermal diffusion phases of diffusion coefficient D (denoted phases 1), randomly interrupted by ballistic excursions bound to motors (referred to as phases 2) of constant velocity V and direction pointing in the solid angle $\omega_{\mathbf{V}}$ (see figure 52). The distribution of the filaments orientation is denoted by $\rho(\omega_{\mathbf{V}})$, and will be taken as either disordered or polarized (see figures 52, 54, 56), which schematically reproduces the different states of the cytoskeleton [Alberts, 2002]. The random duration of each phase i is assumed to be exponentially distributed with mean τ_i. The tracer T can react with reactants R (supposed immobile) during free diffusion phases 1 only, as T is assumed to be inactive when bound to motors. Reaction occurs with a finite probability per unit of time k when the tracer-reactant distance is smaller than a given reaction radius a. In what follows we explicitly determine the kinetic constant K of the reaction $T + R \to R$.

4.2.3 Methods

We now present the basic equations in the case of a reactant centered in a spherical domain of radius b with reflecting boundary. This geometry both mimics the relevant situation of a single target and provides a mean field approximation of the general case of randomly located reactants with concentration $c = a^d/b^d$, where b is the typical distance between reactants. We start from a mean field approximation of the first order reaction constant [Berg and Blomberg, 1976] and write $K = 1/\langle t \rangle$, where the key quantity of our approach is the reaction time $\langle t \rangle$ which is defined as the mean first passage time [Redner, 2001, Condamin et al., 2007] of the tracer at a reactant position uniformly averaged over its initial position. t_1 is defined as the mean reaction time if the tracer starts in phase 1 at position $\mathbf{r}$, and t_2 is defined as the mean reaction time if the tracer starts in phase 2 at position $\mathbf{r}$ with velocity $\mathbf{v}$. For the active intermittent dynamics defined above, t_1 and t_2 satisfy the following backward equations [Redner, 2001] (see section 5 for derivation) :

$$\begin{cases} D\Delta_{\mathbf{r}} t_1 + \dfrac{1}{\tau_1}\displaystyle\int (t_2 - t_1)\rho(\omega_{\mathbf{V}})d\omega_{\mathbf{V}} - k\mathrm{I}_a(\mathbf{r})t_1 = -1 \\ \mathbf{V}\cdot\nabla_{\mathbf{r}} t_2 - \dfrac{1}{\tau_2}(t_2 - t_1) = -1 \end{cases} \tag{126}$$

where I_a is the indicator function of the sphere of radius a. As these equations (126) are of integro-differential type, standard methods of resolution are not available for

a general distribution ρ.

However, in the case of a *disordered* distribution of filaments ($\rho(\omega_v) = 1/\Omega_d$, where Ω_d is the solid angle of the d–dimensional sphere), these equations can be solved exactly in one dimension. In 2 and 3 dimensions, we make the approximation that the first direction taken does not matter, and as the mean search time is averaged over the starting position, it leads to good results when $b \gg a$. The details of calculation are given in section 5, where this model is generalized. We present here simplified expressions of the resulting kinetic constant by taking alternatively the limit $k \to \infty$, which corresponds to the ideal case of perfect reaction, and the limit $D \to 0$ which allows us to isolate the k dependence.

4.2.4 Active transport in the cytoplasm

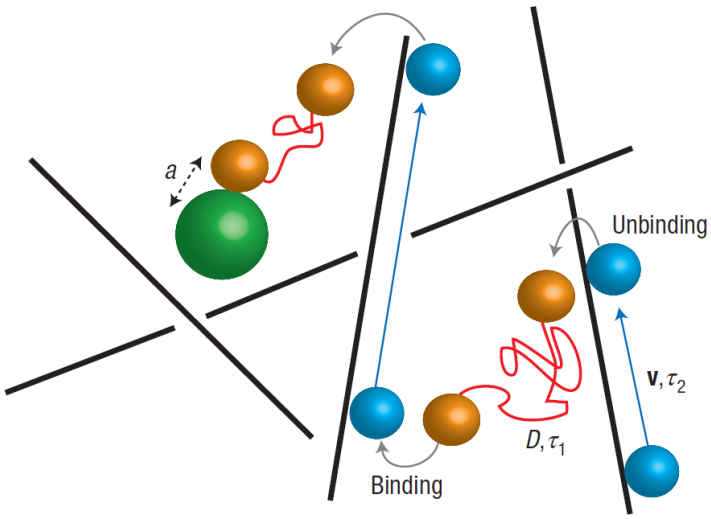


Figure 52: Vesicle transport in the bulk (3 dimensions).

We first discuss the $d = 3$ disordered case (see figure 52), which provides a general description of the actin cytoskeleton of a cell in non polarized conditions, or of a generic in vitro active solution. An analytical form of the mean first passage time $\langle t \rangle = 1/K_{3d}$ is given section 5.5.2, and plotted in figure 53. Strikingly, K_{3d} can be maximized as soon as the reaction radius exceeds a threshold $a_c \simeq 6D/V$ for the following value of the mean interaction time with motors:

$$\tau_{2,3d}^{\text{opt}} = \frac{\sqrt{3}a}{V x_0} \simeq 1.078 \frac{a}{V}, \tag{127}$$

where x_0 is the solution of $2\tanh(x) - 2x + x\tanh(x)^2 = 0$. The τ_1 dependence is very weak, but one can roughly estimate the optimal value by $\tau_{1,3d}^{\text{opt}} \simeq 6D/V^2$. This gives in turn the maximal reaction rate

$$K_{3d}^m \simeq \frac{cV}{a} \frac{\sqrt{3}\,(x_0 - \tanh(x_0))}{x_0^2}, \tag{128}$$

so that the gain with respect to the reaction rate K_{3d}^p in a passive medium is $G_{3d} = K_{3d}^m/K_{3d}^p \simeq CaV/D$ with $C \simeq 0.26$.

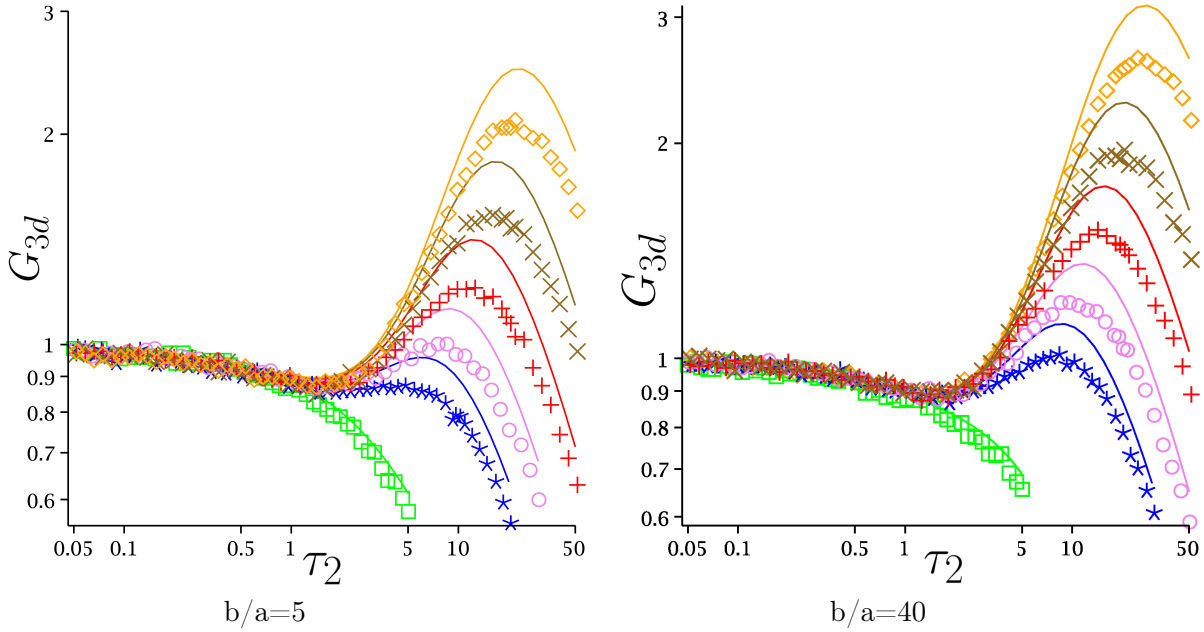


Figure 53: Optimization of the reaction rate for intermittent active transport. Gain of reactivity due to active transport in 3 dimensions as a function of τ_2 for different values of the ratio b/a (logarithmic scale). The analytical form (the mean detection time with diffusion alone (288) divided by the mean detection time with intermittence (280)) (plain lines) is plotted against numerical simulations (symbols) for the following values of the parameters (arbitrary units): $a = 1$ (green, $\square$), $a = 5$ (blue, $\star$), $a = 7$ (purple, $\circ$), $a = 10$ (red, $+$), $a = 14$ (brown, $\times$), $a = 20$ (orange, $\diamond$), with $\tau_1 = 6$, $V = 1$, $D = 1$. G_{3d} presents a maximum only for $a > a_c \simeq 4$.

Several comments are in order. (i) First, $\tau_{2,3d}^{\rm opt}$ neither depends on D, nor on the reactant concentration. A similar analysis for k finite (in the $D \to 0$ limit) shows that this optimal value does not depend on k either (see section 5.5.1), which proves that the optimal mean interaction time with motors is widely independent of the parameters characterizing the diffusion phase 1. (ii) Second, the value a_c should be discussed. In standard cellular conditions D ranges from $\simeq 10^{-2}$ μm^2.s^{-1} for vesicles to $\simeq 10$ μm^2.s^{-1} for small proteins, whereas the typical velocity of a motor protein is $V \simeq 1$ μm.s^{-1}, value which is widely independent of the size of the cargo [Alberts, 2002]. This gives a critical reaction radius a_c ranging from $\simeq 10$ nm for vesicles, which is smaller than any cellular organelle, to $\simeq 10$ μm for single molecules, which is comparable to the whole cell dimension. Hence, this shows that in such 3–dimensional disordered case, active transport can optimize reactivity for sufficiently large tracers like vesicles, as motor mediated motion permits a fast relocation to unexplored regions, whereas it is inefficient for standard molecular reaction kinetics, mainly because at the cell scale molecular free diffusion is faster than motor mediated motion. This could help justifying that many molecular species in cells are transported in vesicles. Interestingly, in standard cellular conditions $\tau_{2,3d}^{\rm opt}$ is of order 0.1 s for a typical reaction radius of order 0.1 μm. This value is compatible

with experimental observations [Alberts, 2002], and suggests that cellular transport is close to optimum. (iii) Last, the typical gain for a vesicle of reaction radius $a \gtrsim 0.1$ µm in standard cellular conditions is $G_{3d} \gtrsim 2.5$ (see figure 53) and can reach $G_{3d} \simeq 10$ for the fastest types of molecular motors ($V \simeq 4$ µm.s^{-1}, see refs [Alberts, 2002, Sheetz and Spudich, 1983]), independently of the reactant concentration c. As we shall see below the gain will be significantly higher in lower dimensional structures such as axons.

4.2.5 Active transport at membranes

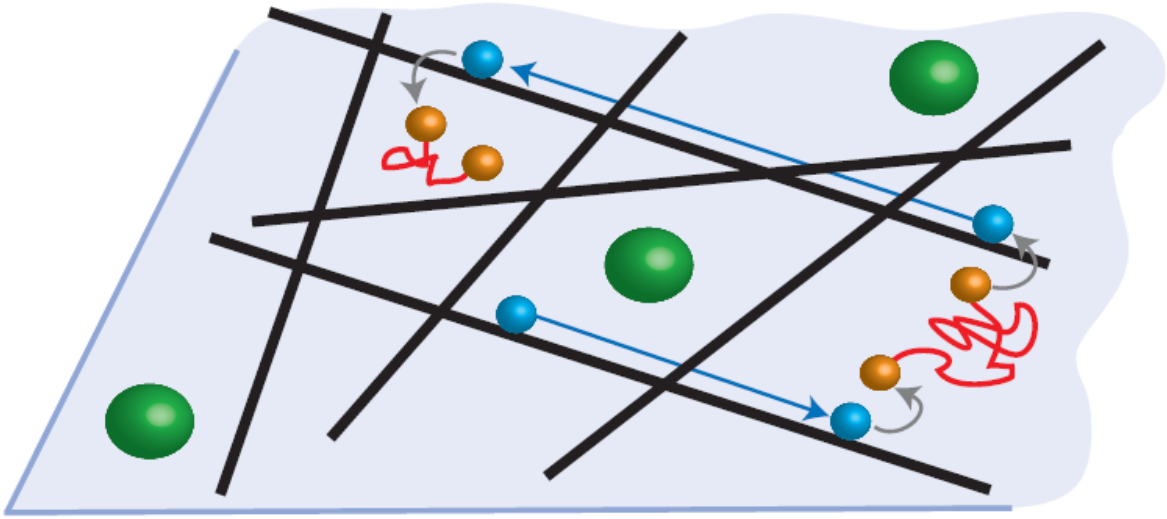


Figure 54: Planar structures such as membranes and lamellipodia ($d = 2$).

We now come to the $d = 2$ disordered case (see figure 54). Striking examples in cells are given by the cytoplasmic membrane, which is closely coupled to the network of cortical actin filaments, or the lamellipodium of adhering cells [Alberts, 2002]. In many cases the orientation of filaments can be assumed to be random. It can be shown that as for $d = 3$ (see section 5.4.2), the reaction rate K_{2d} can be optimized in the regime $D/V \ll a \ll b$. Remarkably, the optimal interaction time $\tau_{2,2d}^{\rm opt}$ takes one and the same value in the two limits $k \to \infty$ and $D \to 0$:

$$\tau_{2,2d}^{\rm opt} \simeq \frac{a}{V\sqrt{2}}(\ln(1/c) - 1)^{1/2}, \tag{129}$$

which indicates that again $\tau_{2,2d}^{\rm opt}$ does not depend on the parameters of the thermal diffusion phase, neither through D nor k. In the limit $k \to \infty$ one has $\tau_{1,2d}^{\rm opt} = \frac{D}{8V^2}\frac{\ln^2(1/c)}{\ln(1/c)-1}$, and the maximal reaction rate can then be obtained :

$$K_{2d}^m \simeq \frac{cV}{a\sqrt{2\ln(1/c)}}. \tag{130}$$

Comparing this expression to the case of passive transport yields a gain $G_{2d} = K_{2d}^m/K_{2d}^p \simeq av\sqrt{\ln(1/c)}/(4D\sqrt{2})$. As in the $d = 3$ case, this proves that active

transport enhances reactivity for large enough tracers (with a critical reaction radius $a_c \simeq D/V$ of the same order as in the $d = 3$ case) such as vesicles. However, here the gain G_{2d} depends on the reactant concentration c, and can be more significant : with the same values of D, V and a as given above in standard cellular conditions, and for low concentrated reactants (like specific membrane receptors) with a typical distance between reactants $b \gtrsim 10$ µm, the typical gain is $G_{2d} \gtrsim 8$, and reaches 10 for single reactants (like examples of signaling molecules) (see figure 55).

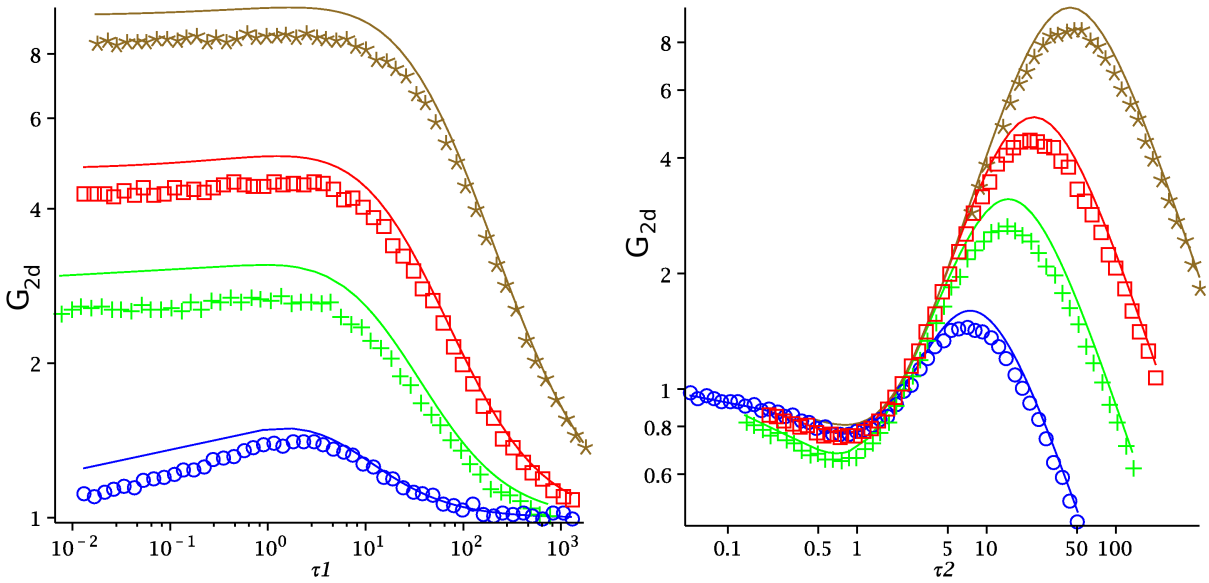


Figure 55: Optimization of the reaction rate for intermittent active transport. Gain of reactivity due to active transport G_{2d} in two dimensions as a function of τ_1 or τ_2 (logarithmic scale). The analytical form (the mean detection time with diffusion alone (234) divided by the mean detection time with intermittence (230)) (plain lines) is plotted against numerical simulations (symbols) for the following values of the parameters (arbitrary units): $a = 20$, $b = 2000$ (brown, $\star$), $a = 10$, $b = 1000$ (red, $\square$), $a = 10$, $b = 100$ (green, $+$), $a = 2.5$, $b = 250$ (blue, $\circ$) with $V = 1$, $D = 1$. These curves represent standard cellular conditions (as discussed in the text).

4.2.6 Active transport in tubular structures

The case of *nematic order* of the cytoskeletal filaments, which depicts for instance the situation of a polarized cell [Alberts, 2002], can be shown to be equivalent in a first approximation to the 1–dimensional case, which is exactly solvable (see figure 56) (for calculations, see section 5.3.2). The $d = 1$ case is also important on its own in cell biology as many 1–dimensional active structures such as axons, dendrites, or stress fibers are present in living cells [Alberts, 2002]. As an illustration, we take the example of an axon, filled with parallel microtubules pointing their plus end in a direction $\mathbf{e}$. We consider a tracer particle interacting with both kinesins ("+" end directed motors, of average velocity $V\mathbf{e}$) and dyneins ("-" end directed motors, of average velocity $-V\mathbf{e}$) with the same characteristic interaction time τ_2 (see figure

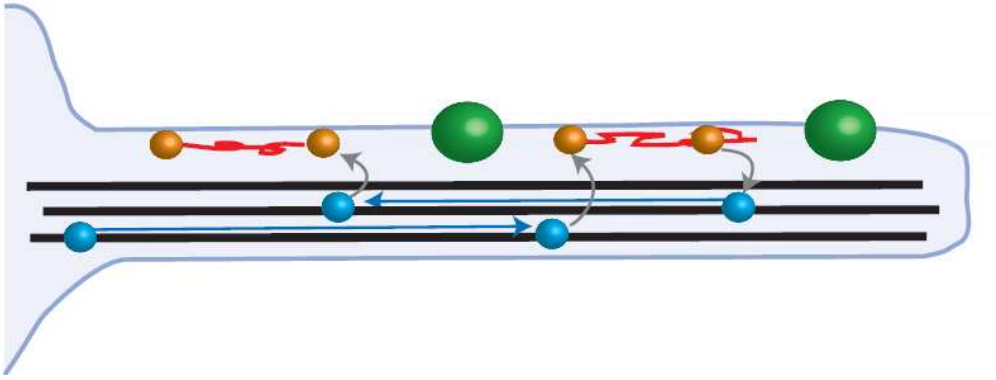


Figure 56: Tubular structures in cells such as axons and dendrites ($d = 1$).

1b). For this type of tracer, the mean first passage time satisfies equations (126) with an effective nematic distribution of filaments $\rho(\omega_{\mathbf{V}}) = \frac{1}{2}(\delta(\mathbf{V}-\mathbf{e}) + \delta(\mathbf{V}+\mathbf{e}))$. The reaction rate K_{1d} is maximized in the regime $D/V \ll a \ll b$ for the following values of the characteristic times (see figure 57) :

$$\tau_{1,1d}^{\text{opt}} = \frac{1}{48}\frac{D}{V^2 c}, \; \tau_{2,1d}^{\text{opt}} = \frac{1}{\sqrt{3}}\frac{a}{V c^{1/2}}, \tag{131}$$

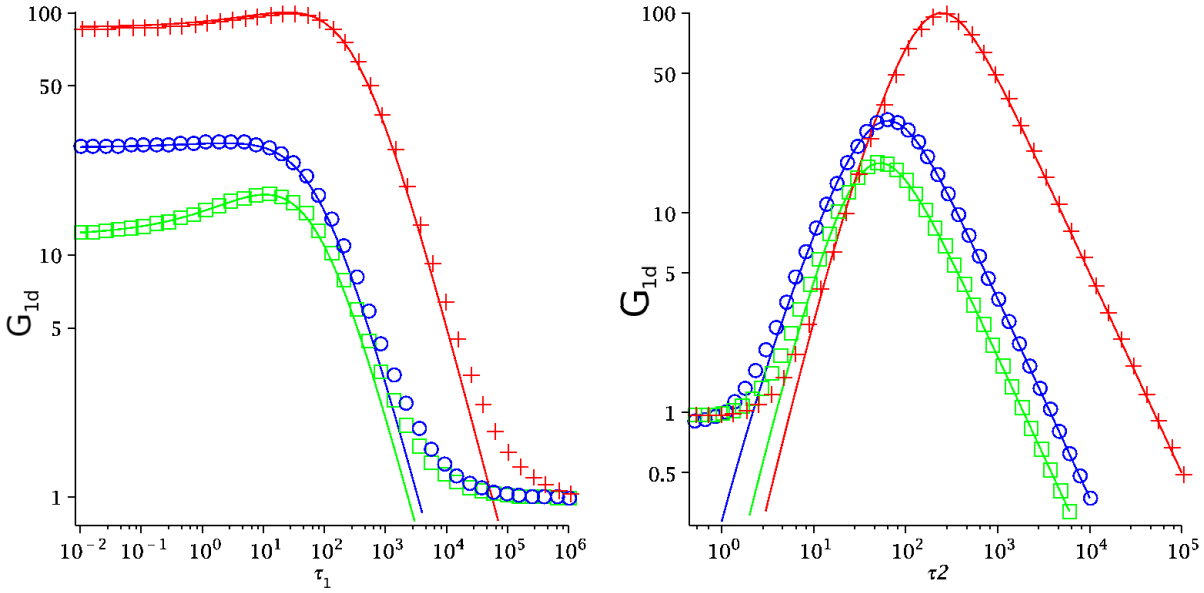
for $k \to \infty$. The maximal reaction rate K_{1d}^m is then given by

$$K_{1d}^m \simeq \frac{\sqrt{3} V c^{3/2}}{2a}, \tag{132}$$

and the gain is $G_{1d} = K_{1d}^m / K_{1d}^p \simeq aV/(2\sqrt{3}Dc^{1/2})$, which proves that active transport can optimize reactivity as in higher dimensions. Very interestingly the c dependence of the gain is much more important than for $d = 2, 3$, which shows that the efficiency of active transport is strongly enhanced in 1-dimensional or nematic structures at low concentration. Indeed, with the same values of D, V and a as given above in standard cellular conditions, and for a typical distance between reactants $b \gtrsim 100$ µm (like low concentrated axonal receptors), one obtains a typical gain $G_{1d} \gtrsim 100$ (see figure 57). In the limit of finite reactivity (k finite and $D \to 0$) one has $\tau_{1,1d}^{\text{opt}} = \sqrt{\frac{a}{Vk}} \left(\frac{2\ln(1/c)-1}{8}\right)^{1/4}$ and the same optimal value (131) of $\tau_{2,1d}^{\text{opt}}$. As in higher dimensions $\tau_{2,1d}^{\text{opt}}$ depends neither on the thermal diffusion coefficient D of phases 1, nor on the association constant k, which shows that the optimal interaction time with motors τ_2^{opt} presents remarkable universal features. Furthermore, our approach permits an estimate of τ_2^{opt} compatible with observations in standard cellular conditions, which suggests that cellular transport could be close to optimum.

4.2.7 Conclusion

Starting from the observation of vesicles alternating free diffusion and phases bound to motors performing ballistic motion, and from the observation that (at least in some cases), vesicles can only react in the free phase, we have proposed a model for



$\tau_2 \simeq \tau_2^{opt}$. $\tau_1 \simeq \tau_1^{opt}$.

Figure 57: Optimization of the reaction rate for intermittent active transport. Gain of reactivity due to active transport G_{1d} in one dimension as a function of τ_1 or τ_2 (logarithmic scale). The analytical form (the mean detection time with diffusion alone (177) divided by the mean detection time with intermittence (179)) is plotted against the exact solution (361) (symbols), for the following values of the parameters (arbitrary units): $D = 1, V = 1$ for all curves and $a = 10, b = 10^4$ (red, $+$), $a = 10, b = 10^3$ (blue, $\circ$), $a = 2.5, b = 10^3$ (green, $\square$). Standard cellular conditions (as discussed in the text) correspond to blue and red curves.

intermittent active transport. We have explicitly calculated the reaction rate, which can be approximated by the inverse of the mean first passage time. We have explored it for various cellular geometries (bulk cytoplasm, membranes, tubular structures), with random orientations of cytoskeletal filaments. We have shown that intermittent transport can indeed increase reaction rates, in particular for large objects such as vesicles, and in particular in low dimensions. The model for the reactive phase is either diffusive or static (with a reaction rate), and both lead to the same optimal duration of the ballistic phase. The latter point is investigated in more details in next section.

5 Intermittent search : a robust strategy

Contents

5.1 Introduction 101
5.2 Model, notations and simulations methods 102
5.2.1 Model 102
5.2.2 Methods 104
5.2.3 Methods of simulation 105
5.3 One dimension 105
5.3.1 Static mode 106
5.3.2 Diffusive mode 110
5.3.3 Ballistic mode 114
5.3.4 Conclusion in one dimension 119
5.4 Two dimensions 119
5.4.1 Static mode 119
5.4.2 Diffusive mode 121
5.4.3 Ballistic mode 125
5.4.4 Conclusion in two dimensions 129
5.5 Three dimensions 129
5.5.1 Static mode 130
5.5.2 Diffusive mode 133
5.5.3 Ballistic mode 137
5.5.4 Conclusion in three dimensions 140
5.6 Discussion and conclusion 140

As shown in previous sections, intermittent search strategies are observed at the macroscopic scale (foraging animals) as well as at the microscopic scale (localization of a DNA sequence by a protein, vesicle transport within cells). The models we have used to interpret these findings, in particular in sections 3.4 and 4.2, present similar general features.

In this section, we present a generic model of intermittent search [Loverdo et al., 2009a] based on these general features. We study it systematically in 1, 2, and 3 dimensions, and for three different modelings of the detection phase. General conclusions on intermittent random walks can be drawn from this systematic study, and are summarized in table 3.

5.1 Introduction

The generic model presented here relies on a succession of slow phases with detection, and ballistic phases without detection, without direction correlation between ballistic phases. This model is minimal is the sense that the searcher has low memory skills.

Indeed, without correlations between ballistic phases, there is no spatial memory. We also assume a Markovian searcher, *i.e.* that does not have temporal memory. As illustrated in the previous examples the search time is often a limiting quantity whose optimization can be very beneficial for the system – be it an animal or a single cell. In the case of intermittent search strategies, the minimization of the search time can be qualitatively discussed : on the one hand, the fast but non-reactive phases can appear as a waste of time, since they do not give any chance of target detection. On the other hand, such fast phases can provide an efficient way to relocate and explore space. This puts forward the following questions : is it beneficial for the search to perform such fast but non reactive phases? Is it possible, by properly tuning the kinetic parameters of trajectories (such as the durations of each of the two phases) to minimize the search time? We develop in what follows a systematic analytical study of intermittent random walks in one, two and three dimensions and fully characterize the optimal regimes. Overall, this systematic approach allows us to identify robust features of intermittent search strategies. In particular, the slow phase that enables detection is often hard to characterize experimentally. Here we propose and study three distinct modelings for this phase, which allows us to assess to which extent our results are robust and model independent. Our analysis covers in details intermittent search problems in one, two and three dimensions and is aimed at giving a quantitative basis – as complete as possible – to model real search problems involving intermittent searchers.

We first define our model and explain our methods. Then we systematically examine each case, studying the search problem in one, two and three dimensions, where for each dimension different types of motion in the slow phase are considered. Each case is ended by a short summary, and we highlight the main results for each dimension. Eventually we synthesize the results in the table 3 where all cases, their differences and similarities are gathered. This table finally leads us to draw general conclusions.

5.2 Model, notations and simulations methods

5.2.1 Model

The general framework of the model again relies on intermittent trajectories. We consider a searcher that switches between two phases. The switching rate λ_1 (resp. λ_2) from phase 1 to phase 2 (resp. from phase 2 to phase 1) is time-independent, which assumes that the searcher has no temporal memory and implies an exponential distribution of durations of each phase i of mean $\tau_i = 1/\lambda_i$.

Phase 1 denotes the phase of slow motion, during which the target can be detected if it lies within a distance from the searcher which is smaller than a given detection radius a, which is the maximum distance within which the searcher can get information about target location. We propose 3 different modelings of this phase, in order to cover various real life situations (see figure 58).

- In the first modeling of phase 1, hereafter referred to as the "static mode", the searcher is immobile, and detects the target with probability per unit time k if it lies at a distance less than a.

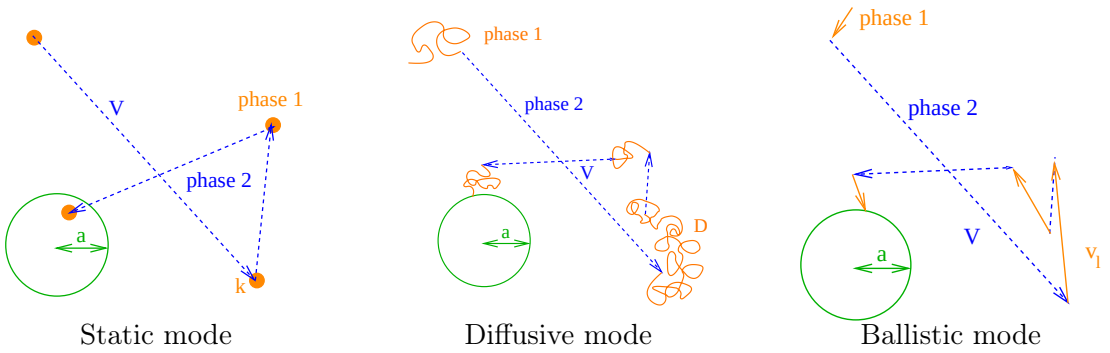


Figure 58: The three different descriptions of phase 1 (the phase with detection), here represented in two dimensions.

- In the second modeling, called the "diffusive mode", the searcher performs a continuous diffusive motion, with diffusion coefficient D, and finds immediately the target if it lies at a distance less than a.
- In the last modeling, called the "ballistic mode", the searcher moves ballistically in a random direction with constant speed v_l and reacts immediately with the target if it lies at a distance less than a. We note that this mode is equivalent to the model of Lévy walks searches proposed by Viswanathan et al. [1999], except for the law of the time between reorientations (see section 3.1). They showed that for destructed targets, *i.e.* targets that cannot be revisited, the optimal strategy is obtained for a straight ballistic motion, without reorientations (see section 3.5). We show here that if another motion, "blind" (*i.e.* without detection) but with higher velocity is available, there are regimes outperforming the straight line strategy.

Some comments on these different modelings of the slow phase 1 are to be made. First, these 3 modes schematically cover experimental observations of the behavior of animals searching for food [Bell, 1991, O'Brien et al., 1990], where the slow phases of detection are often described as static, random or with slow velocity. Several real situations are also likely to involve a combination of two modes. For instance the motion of a reactive particle in a cell not bound to motors can be described by a combination of the diffusive and static modes. For the sake of simplicity, here we treat these modes independently, and our approach can therefore be considered as a limit of more realistic models. Finally, combining these three schematic modes covers a wide range of possible motions, from subdiffusive (even static), diffusive, to superdiffusive (even ballistic). Beyond the modeling of real-life systems, studying different detection modes enables us to assess the robustness of our results.

The phase 2 denotes the fast phase during which the target cannot be found. In this phase, the searcher performs a ballistic motion at a constant speed V in a random direction, redrawn each new phase 2, independently of previous phases. In real examples correlations between successive ballistic phases could exist, as observed for foraging animals [O'Brien et al., 1990]. If correlations are very high, it is close to a 1-dimensional problem with all the phases 2 in the same direction, a different problem already treated section 3.2. We consider here the limit of low correlation, that is of a searcher with no memory skills.

We assume that the searcher evolves in a d-dimensional spherical domain of radius b, with reflective boundaries and with one centered immobile target of radius a. As the searcher does not initially know the target's location, we start the walk from a random point of the d-dimensional sphere, and average the mean target detection time over the initial position. This geometry models the case of a single target in a finite domain, and also provides a good approximation of an infinite space with regularly spaced targets. Such regular array of targets corresponds to a mean-field approximation of random distributions of targets, which can be more realistic in some experimental situations. We note that in the 1 dimensional case, we have shown that a Poissonian distribution of targets can lead to significantly different results from the regular distribution (see section 3.3). We expect this difference to be less in dimension 2 and 3, and we limit ourselves here to the mean field treatment for the sake of simplicity.

5.2.2 Methods

We explain here the general methods, and introduce the notations.

We define $s_i(\mathbf{r},t)$ the probability that the searcher has not yet found the target at t, *starting* from $\mathbf{r}$ in state i, where state $i=1$ is the slow motion phase with detection and state $i=2$ is the fast motion phase without target detection. Note that in dimension 1, the space coordinate will be denoted by x, and in the case of a ballistic mode for phase 1, the upper index in $t_i^{\pm}$ stands for ballistic motion with direction $\pm x$. To find the equations $s_i(\mathbf{r},t)$ is solution of, we enumerate all that could have happened in the beginning of the walk. (i) If the searcher is on the target ($r<a$) and if it is in the detection phase, the searcher can detect the target. In the diffusive and the ballistic modes, $s_1(r<a,t)=0$. In the static mode, as the searcher finds its target with probability k per unit of time, it leads to the term $-ks_1(\mathbf{r},t)I_a(\mathbf{r})$, with $I_a(\mathbf{r})=1$ when $r<a$, and 0 else. (ii) The searcher has a motion characterized by the adjoint operator L [Gardiner, 1996]. For example, for diffusion, $Ls_i(\mathbf{r},t)=D\Delta_{\mathbf{r}}s_i(\mathbf{r},t)$, and for ballistic motion with velocity $\mathbf{v}$, $Ls_i(\mathbf{r},t)=\mathbf{v}.\nabla_{\mathbf{r}}s_i(\mathbf{r},t)$. (iii) The searcher has a probability $1/\tau_i$ per unit of time of switching to the other state j, and in this case, its survival probability will be s_j instead of s_i. Finally, it leads to the backward equation (with $-ks_i(\mathbf{r},t)I_a(\mathbf{r})\delta(i-1)$ only in the case of the static mode) :

$$Ls_i(\mathbf{r},t)+\frac{1}{\tau_i}\left(s_j(\mathbf{r},t)-s_i(\mathbf{r},t)\right)-ks_i(\mathbf{r},t)I_a(\mathbf{r})\delta(i-1)=\frac{\partial s_i(\mathbf{r},t)}{\partial t}. \tag{133}$$

Now we define $t_i(\mathbf{r})$ as the mean first passage time to the target, for a searcher *starting* in the phase i from point $\mathbf{r}$. $t_i(\mathbf{r})$ and $s_i(\mathbf{r},t)$ are related as follows :

$$t_i(\mathbf{r})=-\int_0^\infty t\frac{\partial s_i(\mathbf{r},t)}{\partial t}dt. \tag{134}$$

It can be noticed that :

$$-\int_0^\infty t\frac{\partial^2 s_i(\mathbf{r},t)}{\partial t^2}dt=-\left[t\frac{\partial s_i(\mathbf{r},t)}{\partial t}\right]_0^\infty+\int_0^\infty\frac{\partial s_i(\mathbf{r},t)}{\partial t}dt=0+[s_i(\mathbf{r},t)]_0^\infty=0-1=-1. \tag{135}$$

Consequently, $t_i(\mathbf{r})$ is solution of :

$$Lt_i(\mathbf{r}) + \frac{1}{\tau_i}\left(t_j(\mathbf{r}) - t_i(\mathbf{r})\right) - kt_i(\mathbf{r})I_a(\mathbf{r})\delta(i-1) = -1. \tag{136}$$

For each case, the corresponding equations are explicitly written. These equations are then solved (exactly in one dimension, with an approximation in 2 and 3 dimensions). We assume that the searcher starts in phase 1, and to take into account the fact that it does not initially know the target's location, we average the mean detection time over the starting point, leading to the following definition of the mean search time :

$$t_m = \frac{1}{V(\Omega_d)}\int_{\Omega_d} t_1(\mathbf{r})d\mathbf{r}, \tag{137}$$

with Ω_d the d-dimensional sphere of radius b and $V(\Omega_d)$ its volume. Unless specified, we will consider the low target density limit $a \ll b$.

Our general aim is to minimize t_m as a function of the mean durations τ_1, τ_2 of each phase, and in particular to determine under which conditions an intermittent strategy (with finite τ_2) is faster than a usual single state search in the phase 1 only, which is given by the limit $\tau_1 \to \infty$. In the static mode, intermittence is necessary for the searcher to move, and is therefore always favorable. In the diffusive mode, we will compare the mean search time with intermittence t_m to the mean search time for a single state diffusive searcher t_{diff}, and define the gain as $gain = t_{\text{diff}}/t_m$. Similarly in the ballistic mode, we will compare t_m to the mean search time for a single state ballistic searcher t_{bal} and define the gain as $gain = t_{bal}/t_m$.

The upper index "opt" is used to denote the value of a parameter or variable at the minimum of t_m .

5.2.3 Methods of simulation

As we will show, in 1 dimension, calculations are exact; but in 2 and 3 dimensions, there are approximations in the analytical calculations, that we have checked through simulations.

To save time in the simulations of diffusion, we use variable step length, as in Berezhkovskii et al. [1998] (see figure 59).

Another method to save time in all the simulations is to use squares (2d) or cubes (3d) with reflective boundaries instead of disks or spheres of the same surface or volume. We numerically test that there is no difference in the results obtained through the two geometries, as soon as $b \gg a$ (see figure 60).

5.3 One dimension

Besides the fact that it involves more tractable calculations, the 1-dimensional case is also interesting to model real search problems (see figure 61). At the microscopic scale, tubular structures of cells such as axons or dendrites in neurons can be considered as 1-dimensional [Alberts, 2002]. The active transport of reactive particles, which alternate between diffusion phases and ballistic phases when bound to molecular motors, can be schematically captured by our model with diffusive mode

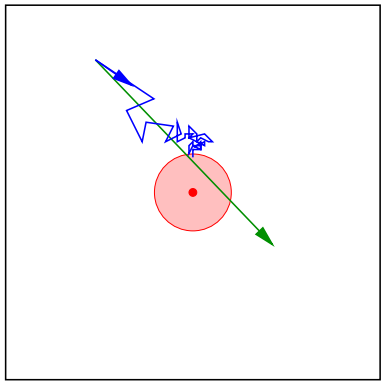
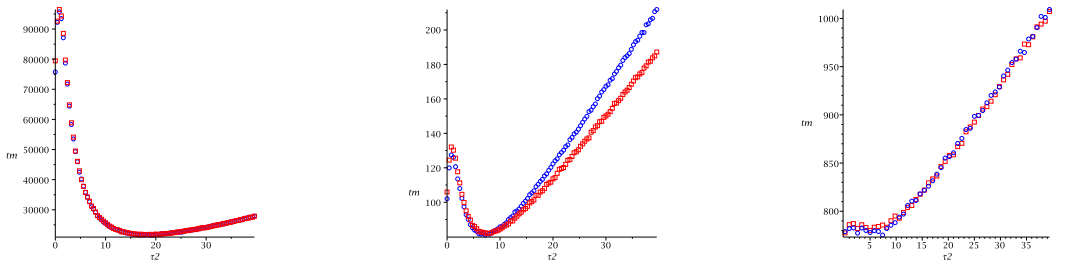


Figure 59: A diffusive phase. Red : the target of detection radius a. Green : a trajectory if the diffusion step is taken equal to the duration of the phase. It artificially decreases the probability of finding the target. Blue : a diffusive step the way it is simulated : step length decreases when approaching the target.



$a = 10$, $b = 250$, τ_1=1.65 | $a = 10$, $b = 25$, τ_1=1.74 | $a = 1$, $b = 25$, τ_1=75.5

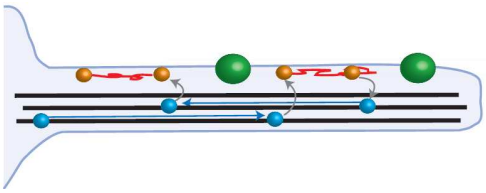
$a = 10$, $b = 250$, τ_2=19.3 | $a = 10$, $b = 25$, τ_2=8.14 | $a = 1$, $b = 25$, τ_2=6.77

Figure 60: Comparison of simulations made in a square (□) and in a disk (○), for the diffusive mode in 2 dimensions. Mean search time is represented as a function of τ_1 (with $\tau_2 \simeq \tau_2^{opt}$) or τ_2 (with $\tau_1 \simeq \tau_1^{opt}$). $D = 1$, $V = 1$. b is the radius of the disk, the side of the square being $L = \sqrt{\pi} b$.

[Loverdo et al., 2008]. At the macroscopic scale, one could cite animals like ants [Dussutour et al., 2005] which tend to follow tracks or one-dimensional boundaries.

5.3.1 Static mode

In this section we assume that the detection phase is modeled by the static mode. Hence the searcher does not move during the reactive phase 1, and has a fixed reaction rate k per unit time with the target if it lies within its detection radius a (see figure 62). It is the limit of a very slow searcher in the reactive phase.



Microscopic scale : vesicles in a tubular cell structure such as dendrites or axons.

Macroscopic scale : ants following a track.

Figure 61: Examples of biological systems which dimension is effectively one.

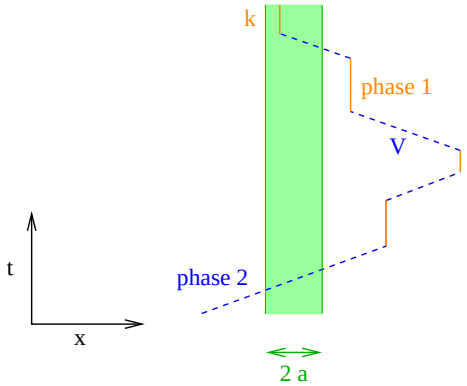


Figure 62: Static mode in one dimension.

Equations

Outside the target (for $x > a$), we have the following backward equations for the mean first-passage time :

$$V\frac{dt_2^+}{dx} + \frac{1}{\tau_2}(t_1 - t_2^+) = -1, \tag{138}$$

$$-V\frac{dt_2^-}{dx} + \frac{1}{\tau_2}(t_1 - t_2^-) = -1, \text{ and} \tag{139}$$

$$\frac{1}{\tau_1}\left(\frac{t_2^+ + t_2^-}{2} - t_1\right) = -1. \tag{140}$$

Inside the target ($x \leq a$), the first two equations are identical, but the third one is written :

$$\frac{1}{\tau_1}\frac{t_2^+ + t_2^-}{2} - \left(\frac{1}{\tau_1} + k\right)t_1 = -1. \tag{141}$$

We introduce $t_2 = \frac{t_2^+ + t_2^-}{2}$ and $t_2^d = \frac{t_2^+ - t_2^-}{2}$. Then outside the target we have the following equations :

$$V\frac{dt_2}{dx} - \frac{1}{\tau_2}t_2^d = 0, \tag{142}$$

$$V^2\tau_2\frac{d^2t_2}{dx^2}+\frac{1}{\tau_2}(t_1-t_2)=0, \tag{143}$$

$$\frac{1}{\tau_1}(t_2-t_1)=-1. \tag{144}$$

Inside the target the first two equations are identical, but the last one writes :

$$\frac{1}{\tau_1}t_2-\left(\frac{1}{\tau_1}+k\right)t_1=-1. \tag{145}$$

Due to the symmetry $x \leftrightarrow -x$, we can restrict the study to the part $x \in [0,a]$ and the part $x \in [a,b]$. This symmetry also implies :

$$\left.\frac{dt_2^{in}}{dx}\right|_{x=0}=0, \tag{146}$$

$$\left.\frac{dt_2^{out}}{dx}\right|_{x=b}=0. \tag{147}$$

In addition, continuity at $x=a$ for t_2^+ and t_2^- gives:

$$t_2^{in}(x=a)=t_2^{out}(x=a), \tag{148}$$

$$t_2^{d,in}(x=a)=t_2^{d,out}(x=a). \tag{149}$$

This set of linear equations enables us to explicitly determine t_1, t_2, t_2^d inside and outside the target.

Results

An exact analytical expression of the mean first passage time to the target is then given by :

$$t_m=\frac{\tau_1+\tau_2}{b}\left(\frac{b}{k\tau_1}+\frac{(b-a)^3}{3V^2\tau_2^2}+\frac{\beta(b-a)^2}{V\tau_2}coth\left(\frac{a}{V\tau_2\beta}\right)+(b-a)\beta^2\right), \tag{150}$$

where $\beta=\sqrt{(k\tau_1)^{-1}+1}$.

In order to determine the optimal strategy, we need to simplify this expression, by expanding (150) in the regime $b \gg a$:

$$t_m=(\tau_1+\tau_2)\left(\frac{1}{k\tau_1}+\frac{b^2}{3V^2\tau_2^2}+\frac{\beta b}{V\tau_2}coth\left(\frac{a}{V\tau_2\beta}\right)+\beta^2\right). \tag{151}$$

We make the further assumption $\frac{a}{V\tau_2} \ll 1$ and obtain, using $\beta > 1$:

$$t_m=(\tau_1+\tau_2)\left(\frac{1}{k\tau_1}+\frac{b^2}{3V^2\tau_2^2}+\beta^2\frac{b}{a}+\beta^2\right). \tag{152}$$

Since $\beta > 1$ and $\beta > 1/(k\tau_1)$, we obtain in the limit $b \gg a$:

$$t_m=(\tau_1+\tau_2)\left(\frac{b^2}{3V^2\tau_2^2}+\left(\frac{1}{k\tau_1}+1\right)\frac{b}{a}\right). \tag{153}$$

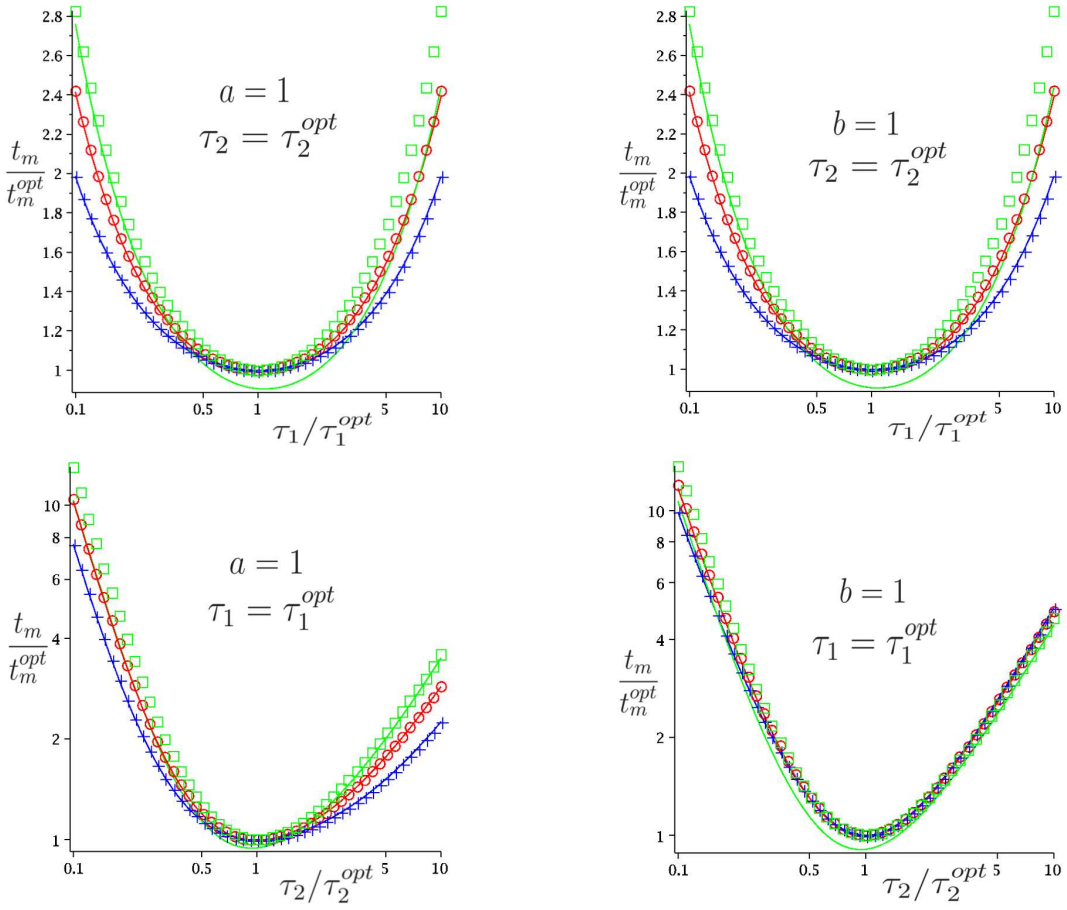


Figure 63: Static mode in one dimension. Exact expression of t_m (150) (lines) compared to the approximation of t_m (153) (symbols), both rescaled by t_m^{opt} (156). τ_1^{opt} from (154), τ_2^{opt} from (155). $V = 1$, $k = 1$. $b/a = 10$ (green, $\square$), $b/a = 100$ (red, $\circ$), $b/a = 1000$ (blue, $+$).

This simple expression gives a very good and convenient approximation of the mean first passage time to the target as shown in figure 63.

We use this approximation (153) to find τ_1 and τ_2 values minimizing t_m :

$$\tau_1^{opt} = \sqrt{\frac{a}{Vk}} \left(\frac{b}{12a}\right)^{1/4}, \tag{154}$$

$$\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}. \tag{155}$$

It can be noticed than τ_2^{opt} does not depend on k. Then the expression of the minimal value of the search time t_m (153) with $\tau_1 = \tau_1^{opt}$ and $\tau_2 = \tau_2^{opt}$ is :

$$t_m^{opt} = \frac{b}{ak}\left(\frac{b}{3a}\right)^{1/4}\left(\sqrt{\frac{2bk}{3V}}\left(\frac{3a}{b}\right)^{1/4} + 1\right)\left(\sqrt{\frac{2ka}{V}} + \left(\frac{3a}{b}\right)^{1/4}\right). \tag{156}$$

Summary

For the static modeling of the detection phase in one dimension, in the $b \gg a$ limit, the mean detection time is :

$$t_m = (\tau_1 + \tau_2)\left(\frac{b^2}{3V^2\tau_2^2} + \left(\frac{1}{k\tau_1} + 1\right)\frac{b}{a}\right). \tag{157}$$

Intermittence is always favorable, and the optimal strategy is realized when $\tau_1^{opt} = \sqrt{\frac{a}{Vk}}\left(\frac{b}{12a}\right)^{1/4}$ and $\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}$. Importantly, the optimal duration of the relocation phase does not depend on k, *i.e.* on the description of the detection phase.

5.3.2 Diffusive mode

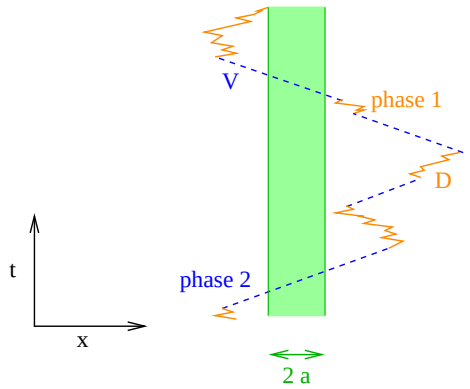


Figure 64: Diffusive mode in one dimension.

We now turn to the diffusive modeling of the detection phase. The detection phase 1 is now diffusive, with immediate detection of the target if it is within a radius a from the searcher (see figure 64).

Equations

Along the same lines, the backward equations for the mean first-passage time read outside the target $(x > a)$:

$$V\frac{dt_2^+}{dx} + \frac{1}{\tau_2}(t_1 - t_2^+) = -1, \tag{158}$$

$$-V\frac{dt_2^-}{dx} + \frac{1}{\tau_2}(t_1 - t_2^-) = -1, \tag{159}$$

$$D\frac{d^2t_1}{dx^2} + \frac{1}{\tau_1}\left(\frac{t_2^+}{2} + \frac{t_2^-}{2} - t_1\right) = -1, \tag{160}$$

and inside the target $(x \leq a)$:

$$V\frac{dt_2^+}{dx} - \frac{1}{\tau_2}t_2^+ = -1, \tag{161}$$

$$-V\frac{dt_2^-}{dx} - \frac{1}{\tau_2}t_2^- = -1, \text{ and} \tag{162}$$

$$t_1 = 0. \tag{163}$$

We introduce the variables $t_2 = \frac{t_2^+ + t_2^-}{2}$ and $t_2^d = \frac{t_2^+ - t_2^-}{2}$. This leads to the following system outside the target ($x > a$) :

$$V\frac{dt_2}{dx} = \frac{1}{\tau_2}t_2^d, \tag{164}$$

$$V^2\tau_2\frac{dt_2}{dx} + \frac{1}{\tau_2}(t_1 - t_2) = -1, \tag{165}$$

$$D\frac{d^2t_1}{dx^2} + \frac{1}{\tau_1}(t_2 - t_1) = -1, \tag{166}$$

and inside the target ($x \leq a$) :

$$V\frac{dt_{2,in}}{dx} = \frac{1}{\tau_2}t_{2,in}^d, \tag{167}$$

$$V^2\tau_2\frac{dt_{2,in}}{dx} - \frac{1}{\tau_2}t_{2,in} = -1, \text{ and} \tag{168}$$

$$t_1 = 0. \tag{169}$$

Interestingly, this system is exactly of the same type that what would be obtained with 2 diffusive phases, with $D_2^{\text{eff}} = V^2\tau_2$ in phase 2. Boundary conditions result from continuity and symmetry :

$$t_1(a) = 0, \tag{170}$$

$$t_2^+(a) = t_{2,in}^+(a), \tag{171}$$

$$t_2^-(a) = t_{2,in}^-(a), \tag{172}$$

$$\left.\frac{dt_2}{dx}\right|_{x=b} = 0, \tag{173}$$

$$\left.\frac{dt_1}{dx}\right|_{x=b} = 0 \text{ and} \tag{174}$$

$$\left.\frac{dt_{2,in}}{dx}\right|_{x=0} = 0. \tag{175}$$

Results

Standard but lengthy calculations lead to an exact expression of the mean first detection time of the target t_m given in appendix 8.2.1. We first studied numerically the minimum of t_m in appendix 8.2.1, and identified 3 regimes. In the first regime ($b < \frac{D}{V}$) intermittence is not favorable. For $b > \frac{D}{V}$ intermittence is favorable and two regimes ($\frac{bD^2}{a^3V^2} < 1$ and $\frac{bD^2}{a^3V^2} > 1$) should be distinguished . We now study analytically each of these regimes.

Regime where intermittence is not favorable : $b < \frac{D}{V}$

If $b < \frac{D}{V}$, the time spent to explore the search space is smaller in the diffusive phase than in the ballistic phase. Intermittence cannot be favorable in this regime, as confirmed by the numerical study in appendix 8.2.1.

Without intermittence, the searcher only performs diffusive motion and the problem can be solved straightforwardly. The backward equations read $t_{\text{diff}} = 0$ inside the target ($x \leq a$), and outside the target ($x > a$) :

$$D\frac{d^2 t_{\text{diff}}}{dx^2} = -1. \tag{176}$$

Since $t_{\text{diff}}(x = a) = 0$ and $\frac{dt_{\text{diff}}}{dx}\big|_{x=b} = 0$, we obtain $t_{\text{diff}}(x) = \frac{1}{2D}((b-a)^2 - (b-x)^2)$. The mean first passage time to the target then reads :

$$t_{\text{diff}} = \frac{(b-a)^3}{3Db}, \tag{177}$$

which in the limit $b \gg a$ leads to :

$$t_{\text{diff}} \simeq \frac{b^2}{3D}. \tag{178}$$

Optimization in the first regime where intermittence is favorable : $b > \frac{D}{V}$ **and** $\frac{bD^2}{a^3V^2} \gg 1$

As explained in details in appendix 8.2.1, we use the approximation of low target density ($b \gg a$), and we use assumptions on the dependence of τ_1^{opt} and τ_2^{opt} on b and a. These assumptions lead to the following approximation of the mean first passage time :

$$t_m = (\tau_1 + \tau_2)\, b \left(\frac{b}{3V^2\tau_2^2} + \frac{1}{\sqrt{D\tau_1}} \right). \tag{179}$$

We checked numerically that this expression gives a good approximation of t_m in this regime, in particular around the optimum ((see figure 65)).

The simplified t_m expression (179) is minimized for :

$$\tau_1^{opt} = \frac{1}{2}\sqrt[3]{\frac{2b^2D}{9V^4}}, \tag{180}$$

$$\tau_2^{opt} = \sqrt[3]{\frac{2b^2D}{9V^4}}, \tag{181}$$

$$t_m^{opt} \simeq \sqrt[3]{\frac{3^5}{2^4}\frac{b^4}{DV^2}}. \tag{182}$$

This compares to the case without intermittence (177) according to :

$$gain^{opt} = \frac{t_{\text{diff}}}{t_m^{opt}} \simeq \sqrt[3]{\frac{2^4}{3^8}} \left(\frac{bV}{D}\right)^{\frac{2}{3}} \simeq 0.13 \left(\frac{bV}{D}\right)^{\frac{2}{3}}. \tag{183}$$

These results are in agreement with numerical minimization of the exact t_m (Table 4 in appendix 8.2.1).

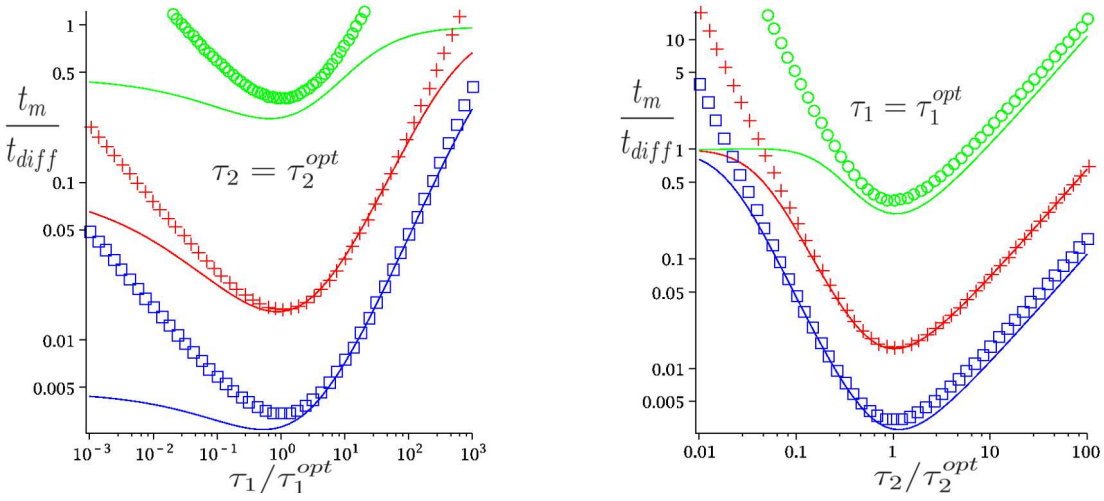


Figure 65: Diffusive mode in one dimension. $\frac{t_m}{t_{\text{diff}}}$, t_{diff} from (178), and t_m exact expression (361) (line), approximation in the regime of favorable intermittence and $\frac{bD^2}{a^3V^2} \gg 1$ (179) (symbols). $a = 1$ and $b = 100$ (green, $\circ$), $a = 1$, $b = 10^4$ (red, $+$), $a = 10$, $b = 10^5$ (blue, $\square$). $D = 1$, $V = 1$. τ_1^{opt} is from expression (180), τ_2^{opt} is from expression (181).

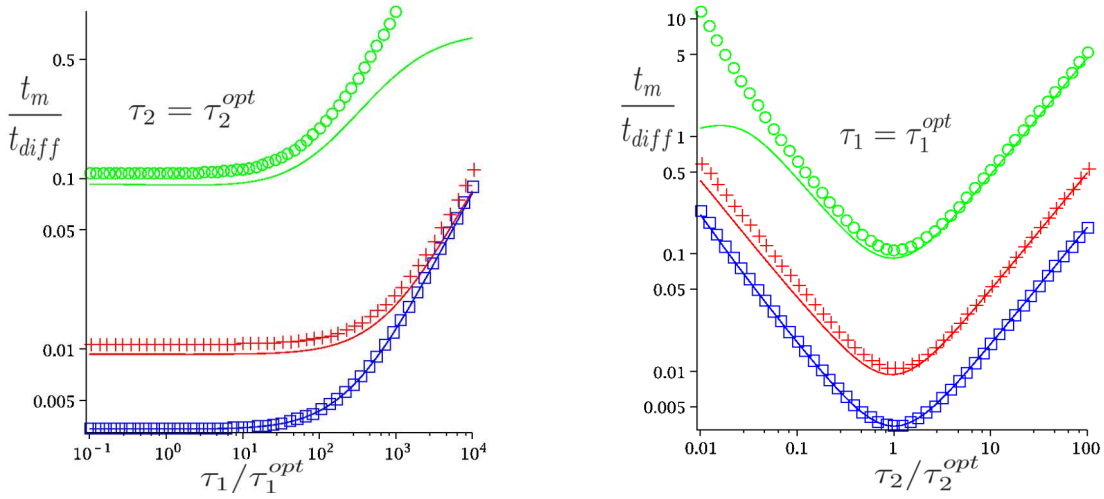


Figure 66: Diffusive mode in 1 dimension. $\frac{t_m}{t_{\text{diff}}}$, t_{diff} from (178), and t_m exact expression (361) (line), approximation in the regime of favorable intermittence and $\frac{bD^2}{a^3V^2} \ll 1$ (184) (symbols). $a = 10$ and $b = 100$ (green, $\circ$), $a = 10$, $b = 1000$ (red, $+$), $a = 100$, $b = 10^4$ (blue, $\square$). $D = 1$, $V = 1$. τ_1^{opt} is from expression (185), τ_2^{opt} is from expression (186).

Optimization in the second regime where intermittence is favorable : $b > \frac{D}{V}$ and $1 \gg \frac{bD^2}{a^3V^2}$

We start from the exact expression of t_m (361). As detailed in appendix 8.2.1, we make assumptions on the dependence of τ_1^{opt} and τ_2^{opt} with b and a, and use the assumptions that $b \gg a$ and $1 \gg \frac{bD^2}{a^3V^2}$. It leads to :

$$t_m \simeq \frac{b}{a}(\tau_1 + \tau_2)\left(\frac{a}{a + \sqrt{D\tau_1}} + \frac{ab}{3V^2\tau_2^2}\right). \tag{184}$$

This expression gives a good approximation of t_m, at least around the optimum ((see

figure 66)), which is characterized by:

$$\tau_1^{opt} = \frac{Db}{48V^2a}, \tag{185}$$

$$\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}, \tag{186}$$

$$t_m^{opt} \simeq \frac{2a}{V\sqrt{3}}\left(\frac{b}{a}\right)^{3/2}, \tag{187}$$

$$gain \simeq \frac{1}{2\sqrt{3}}\frac{aV}{D}\sqrt{\frac{b}{a}}. \tag{188}$$

These results are in very good agreement with numerical data (Table 4 in appendix 8.2.1). Note that the gain can be very large at low target density.

Summary

We calculated explicitly the mean first passage time t_m in the case where the detection phase is modeled by the diffusive mode. We minimized t_m as a function of τ_1 and τ_2, the mean phases durations, with the assumption $a \ll b$. There are three regimes:

- when $b < \frac{D}{V}$, intermittence is not favorable, *i.e.* $\tau_1^{opt} \to \infty$, $\tau_2^{opt} \to 0$, $t_m^{opt} = t_{\text{diff}} \simeq \frac{b^2}{3D}$
- when $b > \frac{D}{V}$ and $\frac{bD^2}{a^3V^2} \gg 1$, intermittence is favorable, with $\tau_2^{opt} = 2\tau_1^{opt} = \sqrt[3]{\frac{2b^2D}{9V^4}}$, and $t_m^{opt} \simeq \sqrt[3]{\frac{3^5}{2^4}\frac{b^4}{DV^2}}$
- when $b > \frac{D}{V}$ and $\frac{bD^2}{a^3V^2} \ll 1$, intermittence is favorable, with $\tau_1^{opt} = \frac{Db}{48V^2a}$, $\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}$, $t_m^{opt} \simeq \frac{2a}{v\sqrt{3}}\left(\frac{b}{a}\right)^{3/2}$.

This last regime is of particular interest, since the value obtained for τ_2^{opt} is the same as in the static mode (see section 5.3.1).

5.3.3 Ballistic mode

We now treat the case where the detection phase 1 is modeled by the ballistic mode (see figure 67). This model schematically accounts for the general observation that speed often degrades perception abilities. Our model corresponds to the extreme case where only two modes are available : either the motion is slow and the target can be found, or the motion is fast and the target cannot be found. Note that this model can be compared to the model of Viswanathan et al. [1999], where there is only the detection phase.

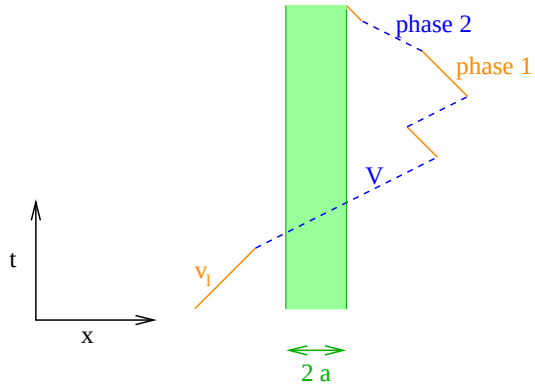


Figure 67: Ballistic mode in one dimension.

Equations

The backward equations read outside the target ($x > a$) :

$$v_l \frac{dt_1^+}{dx} + \frac{1}{\tau_1}\left(\frac{t_2^+}{2} + \frac{t_2^-}{2} - t_1^+\right) = -1, \tag{189}$$

$$-v_l \frac{dt_1^-}{dx} + \frac{1}{\tau_1}\left(\frac{t_2^+}{2} + \frac{t_2^-}{2} - t_1^-\right) = -1, \tag{190}$$

$$V \frac{dt_2^+}{dx} + \frac{1}{\tau_2}\left(\frac{t_1^+}{2} + \frac{t_1^-}{2} - t_2^+\right) = -1 \text{ and} \tag{191}$$

$$-V \frac{dt_2^-}{dx} + \frac{1}{\tau_2}\left(\frac{t_1^+}{2} + \frac{t_1^-}{2} - t_2^-\right) = -1. \tag{192}$$

Defining $t_i^d = \frac{t_i^+ - t_i^-}{2}$ and $t_i = \frac{t_i^+ + t_i^-}{2}$, we get the following equations (and similar expressions with $v_l \to V$, $t_1 \to t_2$, $t_2 \to t_1$) :

$$v_l \frac{dt_1^d}{dx} + \frac{1}{\tau_1}(t_2 - t_1) = -1, \tag{193}$$

$$v_l \frac{dt_1}{dx} - \frac{1}{\tau_1} t_1^d = 0, \tag{194}$$

which eventually lead to the following system :

$$v_l^2 \tau_1 \frac{d^2 t_1}{dx^2} + \frac{1}{\tau_1}(t_2 - t_1) = -1 \tag{195}$$

$$V^2 \tau_2 \frac{d^2 t_2}{dx^2} + \frac{1}{\tau_2}(t_1 - t_2) = -1, \tag{196}$$

together with :

$$t_1^d = v_l \tau_1 \frac{dt_1}{dx}, \tag{197}$$

$$t_2^d = V \tau_2 \frac{dt_2}{dx}. \tag{198}$$

Inside the target ($x \leq a$), one has $t_1^{+,in}(x) = t_1^{-,in}(x) = 0$, and :

$$V\frac{dt_2^{+,in}}{dx} - \frac{1}{\tau_2}t_2^{+,in} = -1, \tag{199}$$

$$-V\frac{dt_2^{-,in}}{dx} - \frac{1}{\tau_2}t_2^{-,in} = -1. \tag{200}$$

Finally, the boundary conditions read :

$$\left.\frac{dt_2}{dx}\right|_{x=b} = 0, \tag{201}$$

$$\left.\frac{dt_1}{dx}\right|_{x=b} = 0, \tag{202}$$

$$t_2^+(a) = t_{2,in}^+(a), \tag{203}$$

$$t_2^-(a) = t_{2,in}^-(a), \tag{204}$$

$$\left.\frac{dt_{2,in}}{dx}\right|_{x=0} = 0 \text{ and} \tag{205}$$

$$t_1^-(a) = 0. \tag{206}$$

Results

The exact expression of t_m (see appendix 8.2.2) is obtained through lengthy but standard calculations. To simplify this expression, we consider the small density limit $a/b \to 0$ and finally obtain the following very good approximation of t_m (see figure 68) :

$$t_m = \frac{(\tau_1+\tau_2)b}{\alpha^{3/2}}\left(\left(\frac{b}{3}+L_1\right)\sqrt{\alpha} + \Gamma L_2(\sqrt{\alpha}+L_2)\right), \tag{207}$$

where :

$$\Gamma = \frac{(\sqrt{\alpha}-L_1)(L_1+L_2) + \sqrt{\alpha}(L_2-L_1+\sqrt{\alpha})X + X^2L_2(L_2-L_1)}{((L_1+\sqrt{\alpha})X^2 + (L_1-\sqrt{\alpha}))(\sqrt{\alpha}+L_2-L_1)}, \tag{208}$$

$$X = e^{\frac{2a}{L_2}}, \tag{209}$$

$$\alpha = L_1^2 + L_2^2, \tag{210}$$

$$L_2 = V\tau_2 \text{ and} \tag{211}$$

$$L_1 = v_l\tau_1. \tag{212}$$

A numerical analysis indicates (see figure 68 and table 2) that, depending on the parameters, there are two possible optimal strategies :

- $\tau_1 \to \infty$. Intermittence is not favorable.
- $\tau_1 \to 0, \tau_2 = \tau_2^{opt}$. Intermittence is favorable.

We now study analytically these regimes.

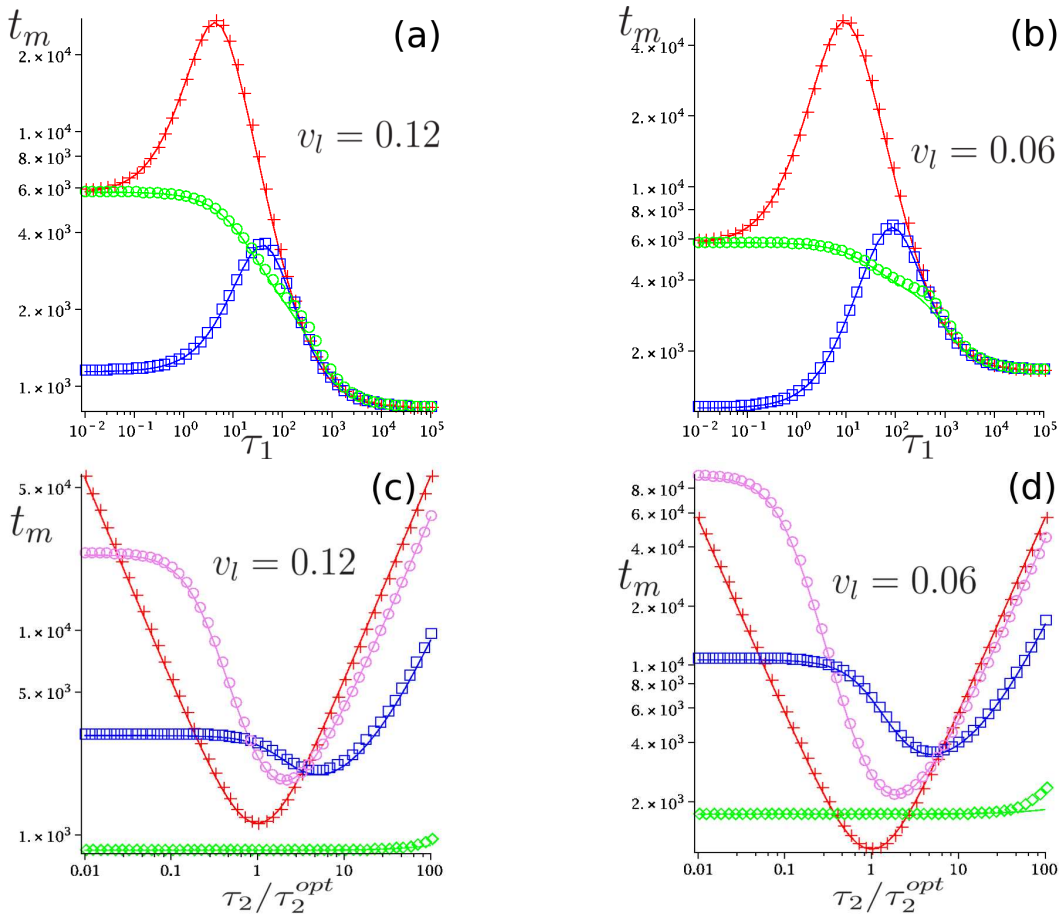


Figure 68: Ballistic mode in one dimension. Comparison between low density approximation (207) (symbols) and the exact expression of t_m (403) (line). (a), (b) : t_m as a function of τ_1, with $\tau_2 = 0.1\tau_2^{opt}$ (red, +), $\tau_2 = \tau_2^{opt}$ (blue, □), $\tau_2 = 10\tau_2^{opt}$ (green, ∘). (c), (d) : t_m as a function of τ_2/τ_2^{opt}, with $\tau_1 = 0$ (red, +), $\tau_1 = 10$ (violet, ∘), $\tau_1 = 100$ (blue, □), $\tau_1 = 10000$ (green, ⋄). (a), (c) : $v_l = 0.12 > v_l^c$: intermittence is not favorable. (b), (d) : $v_l = 0.06 < v_l^c$: intermittence is favorable. τ_2^{opt} is from the analytical prediction (217). $a = 1$, $V = 1$, $b = 100$.

	$v_l = 1$	$v_l = 0.1$	$v_l = 0.01$	$v_l = 0.001$	$\tau_2^{opt,th}(217)$
$b = 5$	$\tau_1 \to \infty$	$\tau_1 \to 0$, $\tau_2^{opt} = 0.86$			$\tau_2^{opt,th} = 0.91$
$b = 50$	$\tau_1 \to \infty$		$\tau_1 \to 0$, $\tau_2^{opt} = 2.9$		$\tau_2^{opt,th} = 2.9$
$b = 500$	$\tau_1 \to \infty$		$\tau_1 \to 0$, $\tau_2^{opt} = 9.1$		$\tau_2^{opt,th} = 9.1$
$b = 5000$	$\tau_1 \to \infty$			$\tau_1 \to 0$, $\tau_2^{opt} = 29$	$\tau_2^{opt,th} = 29$

Table 2: Ballistic mode in one dimension. Numerical minimization of the exact t_m (403). Values of τ_1 and τ_2 at the minimum. Comparison with theoretical τ_2. $a = 0.5$, $V = 1$.

Regime without intermittence : $\tau_1 \to \infty$

In this regime, there is no intermittence. The searcher starts either inside the target (x in $[-a, a]$) and immediately finds the target, or it starts at a position x outside the target. We can therefore take $x \in [a, b]$. If the searcher goes in the $-x$

direction, it find its target after $T = (x-a)/v_l$. If the searcher goes in the $+x$ direction, it finds its target after $T = ((b-x)+(b-a))/v_l$. This leads to :

$$t_{bal} = \frac{1}{b}\int_a^b \frac{b-a}{v_l}dx = \frac{(b-a)^2}{bv_l}. \tag{213}$$

Intermittent regime

We take the limit $\tau_1 \to 0$ in the expression of t_m (207) and obtain :

$$\lim_{\tau_1\to 0} t_m = \frac{b}{V}\left(\frac{b}{3L_2} + \frac{e^{\frac{2a}{L_2}}+1}{e^{\frac{2a}{L_2}}-1}\right). \tag{214}$$

Taking the derivative with respect to L_2 yields :

$$\frac{d}{dL_2}\left(\lim_{\tau_1\to 0} t_m\right) \propto 12ae^{\frac{2a}{L_2}} + 2be^{\frac{2a}{L_2}} - b - be^{\frac{4a}{L_2}}, \tag{215}$$

which has only one positive root :

$$L_2^{opt} = \frac{2a}{\ln\left(1+6a/b+2\sqrt{3a/b+9a^2/b^2}\right)}. \tag{216}$$

In the limit $b \gg a$ it leads to :

$$\tau_2^{opt} = \frac{a}{3V}\sqrt{\frac{b}{a}}, \tag{217}$$

which is in agreement with the numerical minimization of the exact mean detection time shown in the table 2.

The mean first passage time to the target is minimized in the intermittent regime for $\tau_1 \to 0$ and $\tau_2 = \tau_2^{opt}$. We replace τ_2 by (217) in the expression (214), and take $b \gg a$ to finally obtain :

$$t_m^{opt} = \frac{2}{\sqrt{3}}\frac{b}{V}\sqrt{\frac{b}{a}}, \tag{218}$$

$$gain = \frac{\sqrt{3}}{2}\frac{V}{v_l}\sqrt{\frac{a}{b}}. \tag{219}$$

This shows that the gain is larger than 1 for $v_l < v_l^c = V\frac{\sqrt{3}}{2}\sqrt{\frac{a}{b}}$, which defines the regime where intermittence is favorable.

Summary

In the case where the phase 1 is modeled by the ballistic mode in one dimension, we have calculated the exact mean first passage time t_m at the target. t_m can be minimized as a function of τ_1 and τ_2, yielding two possible optimal strategies :

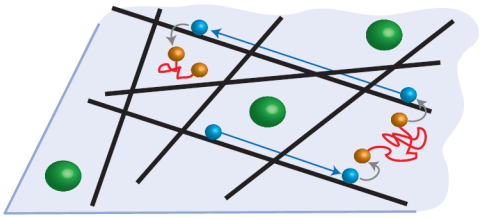
- for $v_l > v_l^c = V\frac{\sqrt{3}}{2}\sqrt{\frac{a}{b}}$, intermittence is not favorable : $\tau_1^{opt} \to \infty$, $\tau_2^{opt} \to 0$,
- for $v_l < v_l^c = V\frac{\sqrt{3}}{2}\sqrt{\frac{a}{b}}$, intermittence is favorable, with $\tau_1^{opt} \to 0$ and $\tau_2^{opt} = \frac{a}{3V}\sqrt{\frac{b}{a}}$.

Note that the model studied by Viswanathan et al. [1999] shows that when targets are not revisitable, the optimal strategy for a single state searcher is to perform a straight ballistic motion. This strategy corresponds to $\tau_1 \to \infty$ in our model. Our results show that if a faster phase without detection is allowed, this straight line strategy can be outperformed.

5.3.4 Conclusion in one dimension

Intermittent search strategies in one dimension share similar features for the static, diffusive and ballistic detection modes. In particular, all modes show regimes where intermittence is favorable and lead to a minimization of the search time. Strikingly, the optimal duration of the non-reactive relocation phase 2 is quite independent of the modeling of the reactive phase : $\tau_2^{opt} = \frac{a}{3V}\sqrt{\frac{b}{a}}$ for the static mode, for the ballistic mode (in the regime $v_l < v_l^c \simeq \frac{V}{2}\sqrt{\frac{3a}{b}}$), and for the diffusive mode (in the regime $b > \frac{D}{V}$ and $a \gg \frac{D}{V}\sqrt{\frac{b}{a}}$). This shows the robustness of the optimal value τ_2^{opt}.

5.4 Two dimensions



Microscopic scale : vesicles traffic on a membrane.

Macroscopic scale : gatherers searching for cockles hidden below the surface. Photo from Michel Sokolowski [Lepolard, 2007].

Figure 69: Biological examples of search in two dimensions.

The 2-dimensional problem is particularly well suited to model animal behaviors; it is also relevant to the microscopic scale, since it mimics for example the case of cellular traffic on membranes[Alberts, 2002] (see figure 69). While in one dimension the mean search time can be calculated analytically, we introduce in two dimensions (and later in three dimensions) approximation schemes, which we check by numerical simulations.

5.4.1 Static mode

We study here the case where the detection phase is modeled by the static mode : the searcher does not move during the detection phase and has a finite reaction rate with the target if it is within its detection radius a (see figure 70).

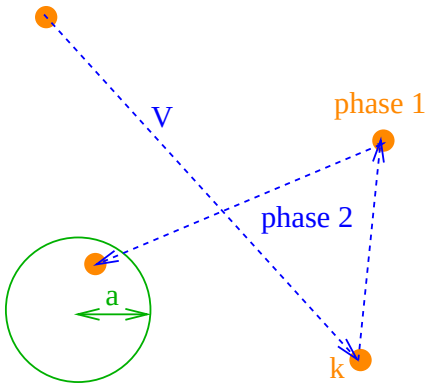


Figure 70: Static mode in two dimensions.

Equations

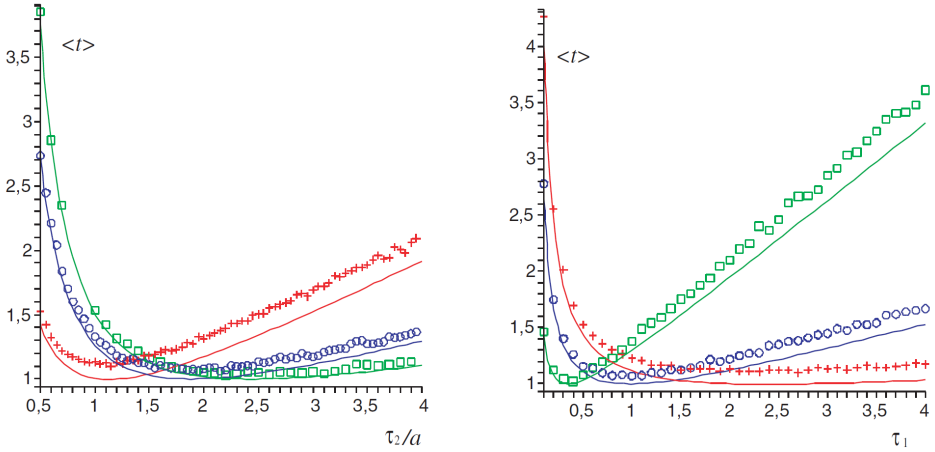


Figure 71: Static mode in two dimensions. Simulations (symbols) and analytical approximate (223) (lines). $k = 1$, $V = 1$, $b = 56$; $a = 10$ (red, +) ($\tau_1^{opt} = 2.41$, $\tau_2^{opt} = 11.2$), $a = 1$ (blue, $\circ$) ($\tau_1^{opt} = 0.969$, $\tau_2^{opt} = 1.88$), $a = 0.1$ (green, $\square$) ($\tau_1^{opt} = 0.348$, $\tau_2^{opt} = 0.242$). *Left* : mean search time t_m as a function of τ_2/a, with $\tau_1 = \tau_1^{opt}$. *Right* : mean search time t_m as a function of τ_1, with $\tau_2 = \tau_2^{opt}$.

The mean first passage time at a target satisfies the following backward equations [Redner, 2001]:

$$\frac{1}{2\pi\tau_1}\int_0^{2\pi}(t_2(\overrightarrow{r})-t_1(\overrightarrow{r}))d\theta_{\overrightarrow{V}} - k\mathrm{I}_a(\overrightarrow{r})t_1(\overrightarrow{r}) = -1. \tag{220}$$

$$\overrightarrow{V}\cdot\nabla_{\mathbf{r}}t_2(\overrightarrow{r}) - \frac{1}{\tau_2}(t_2(\overrightarrow{r})-t_1(\overrightarrow{r})) = -1 \tag{221}$$

The function I_a is defined by $\mathrm{I}_a(\overrightarrow{r}) = 1$ inside the target (if $|\overrightarrow{r}| \leq a$) and $\mathrm{I}_a(\overrightarrow{r}) = 0$ outside the target (if $|\overrightarrow{r}| > a$). In the present form, these integro-differential equations (completed with boundary conditions) do not seem to allow for an exact resolution with standard methods. t_2 is the mean first passage time to the target,

starting from $\overrightarrow{r}$ in phase 2, with speed $\overrightarrow{V}$, of angle $\theta_{\mathbf{V}}$, and with projections on the axes V_x, V_y. i and j can take either x or y as a value. We use the following decoupling assumption :

$$\langle V_i V_j t_2 \rangle_{\theta_{\mathbf{V}}} \simeq \langle V_i V_j \rangle_{\theta_{\mathbf{V}}} \langle t_2 \rangle_{\theta_{\mathbf{V}}} \tag{222}$$

and finally obtain the following approximation of the mean search time, which can be checked by numerical simulations (see figure 71) :

$$t_m = \frac{\tau_1+\tau_2}{2k\tau_1 y^2}\left\{\frac{1}{x}(1+k\tau_1)(y^2-x^2)^2\frac{\mathrm{I}_0(x)}{\mathrm{I}_1(x)} + \frac{1}{4}\left[8y^2+(1+k\tau_1)\left(4y^4\ln(y/x)+(y^2-x^2)(x^2-3y^2+8)\right)\right]\right\}, \tag{223}$$

$$\text{where } x = \sqrt{\frac{2k\tau_1}{1+k\tau_1}}\frac{a}{V\tau_2} \text{ and } y = \sqrt{\frac{2k\tau_1}{1+k\tau_1}}\frac{b}{V\tau_2}. \tag{224}$$

In that case, intermittence is trivially necessary to find the target. In the regime $b \gg a$, the optimization of the search time (223) leads to :

$$\tau_1^{opt} = \left(\frac{a}{Vk}\right)^{1/2}\left(\frac{2\ln(b/a)-1}{8}\right)^{1/4}, \tag{225}$$

$$\tau_2^{opt} = \frac{a}{V}\left(\ln(b/a)-1/2\right)^{1/2}, \tag{226}$$

and the minimum search time is given in the large volume limit by :

$$t_m^{opt} = \frac{b^2}{a^2 k} - \frac{2^{1/4}}{\sqrt{Vka^3}}\frac{(a^2-4b^2)\ln(b/a)+2b^2-a^2}{(2\ln(b/a)-1)^{3/4}} - \frac{\sqrt{2}}{48ab^2V}\frac{(96a^2b^2-192b^4)\ln^2(b/a)+(192b^4-144a^2b^2)\ln(b/a)+46a^2b^2-47b^4+a^4}{(2\ln(b/a)-1)^{3/2}}. \tag{227}$$

Summary

In the case of a static detection mode in two dimensions, we obtained a simple approximate expression of the mean first passage time t_m at the target. With the static detection mode, intermittence is always favorable and leads to a single optimal intermittent strategy. The minimal search time is realized for $\tau_1^{opt} = \left(\frac{a}{Vk}\right)^{1/2}\left(\frac{2\ln(b/a)-1}{8}\right)^{1/4}$ and $\tau_2^{opt} = \frac{a}{V}\left(\ln(b/a)-1/2\right)^{1/2}$. Importantly, the optimal duration of the relocation phase does not depend on k, *i.e.* on the description of the detection phase, like in one dimension.

5.4.2 Diffusive mode

We now assume that the searcher diffuses during the detection phase (see figure 72). For this process, the mean first passage time to the target satisfies the following backward equation [Redner, 2001]:

$$D\nabla^2_{\mathbf{r}} t_1(\overrightarrow{r}) + \frac{1}{2\pi\tau_1}\int_0^{2\pi}(t_2(\overrightarrow{r})-t_1(\overrightarrow{r}))d\theta_{\mathbf{V}} = -1, \tag{228}$$

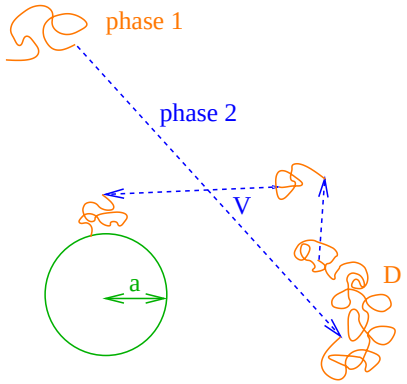


Figure 72: Diffusive mode in two dimension.

$$\overrightarrow{V} \cdot \nabla_{\mathbf{r}} t_2(\overrightarrow{r}) - \frac{1}{\tau_2}(t_2(\overrightarrow{r}) - t_1(\overrightarrow{r})) = -1, \tag{229}$$

with $t_1(\overrightarrow{r}) = 0$ inside the target ($r \leq a$). We use the same decoupling assumption than for the static case (222). It eventually leads to the following approximation of the mean search time, which is checked by numerical simulations (see figure 73) :

$$t_m = (\tau_1+\tau_2)\frac{1-a^2/b^2}{(\alpha^2 D\tau_1)^2}\left\{a\alpha(b^2/a^2-1)\frac{M}{2L_+} - \frac{L_-}{L_+} - \frac{\alpha^2 D\tau_1}{8\widetilde{D}\tau_2}\frac{(3-4\ln(b/a))b^4 - 4a^2b^2 + a^4}{b^2-a^2}\right\}, \tag{230}$$

with :

$$\begin{aligned} \text{with } L_\pm =& \mathrm{I}_0\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_1(a\alpha) - \mathrm{I}_1(a\alpha)\mathrm{K}_1(b\alpha)) \\ & \pm \alpha\sqrt{\widetilde{D}\tau_2}\,\mathrm{I}_1\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_0(a\alpha) + \mathrm{I}_0(a\alpha)\mathrm{K}_1(b\alpha))\,, \end{aligned} \tag{231}$$

and :

$$\begin{aligned} M =& \mathrm{I}_0\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_0(a\alpha) + \mathrm{I}_0(a\alpha)\mathrm{K}_1(b\alpha)) \\ & - 4\frac{a^2\sqrt{\widetilde{D}\tau_2}}{\alpha(b^2-a^2)^2}\mathrm{I}_1\left(\frac{a}{\sqrt{\widetilde{D}\tau_2}}\right)(\mathrm{I}_1(b\alpha)\mathrm{K}_1(a\alpha) - \mathrm{I}_1(a\alpha)\mathrm{K}_1(b\alpha))\,, \end{aligned} \tag{232}$$

where $\alpha = (1/(D\tau_1) + 1/(\widetilde{D}\tau_2))^{1/2}$ and $\widetilde{D} = v^2\tau_2$. We then minimize this time as a function of τ_1 and τ_2.

$a < b \ll D/V$: intermittence is not favorable

In that regime, intermittence is not favorable. Indeed, the typical time required to explore the whole domain of radius b is of order b^2/D with diffusive motion, which

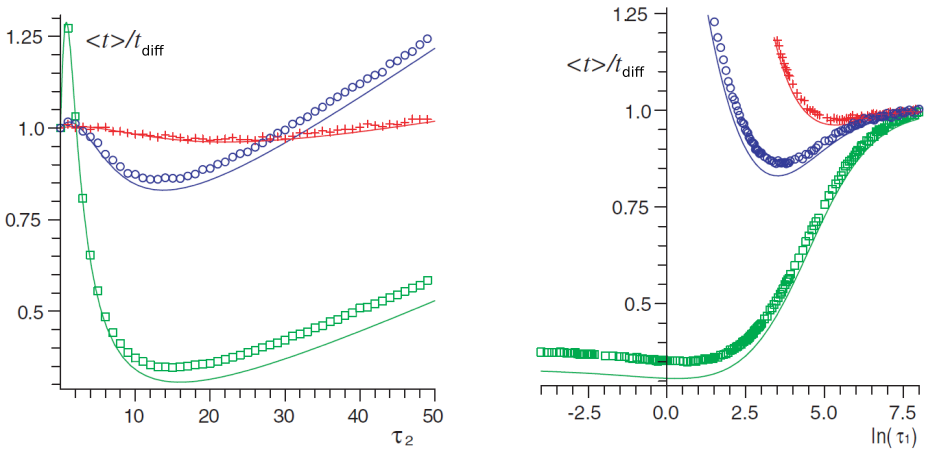


Figure 73: Diffusive mode in two dimensions. Simulations (symbols) versus analytical approximate (230) (line) of the search time, rescaled by the value in the absence of intermittence $t_{\rm diff}$ as a function of τ_2 (left) and $\ln(\tau_1) =$ (right), for $D = 1$, $V = 1$, $b = 226$. *Left* : $a = 10$, $\tau_1 = 1.37$ (green $\square$); $a = 1$, $\tau_1 = 33.6$ (blue $\circ$); $a = 0.1$, $\tau_1 = 213$ (red $+$). *Right*: $a = 10$, $\tau_2 = 15.9$ (green, $\square$); $a = 1$, $\tau_2 = 13.7$ (blue, $\circ$); $a = 0.1$, $\tau_2 = 22$ (red, $+$).

is shorter than the corresponding time b/V with ballistic motion. As a consequence, it is never useful to interrupt the diffusive phases by mere relocating ballistic phases. We use standard methods to calculate the mean first passage time to the target in this optimal regime of diffusion only :

$$\frac{D}{r}\frac{d}{dr}\left(r\frac{dt}{dr}\right) = -1. \tag{233}$$

The boundary conditions $t(a) = 0$ and $\frac{dt}{dr}(r = b) = 0$ lead to :

$$t_{\rm diff} = \frac{1}{8b^2 D_{\rm eff}}\left(4a^2b^2 - a^4 - 3b^4 + 4b^4 \ln\frac{b}{a}\right), \tag{234}$$

and we find in the limit $b \gg a$:

$$t_{\rm diff} = \frac{b^2}{8D_{\rm eff}}\left(-3 + 4\ln\frac{b}{a}\right). \tag{235}$$

$a \ll D/V \ll b$ **: first regime of intermittence**

In this second regime, one can use the following approximate formula for the search time:

$$t_m = \frac{b^2}{4DV^2\alpha^2}\frac{\tau_1+\tau_2}{\tau_1\tau_2^2}\left\{4\ln(b/a) - 3 - 2\frac{(V\tau_2)^2}{D\tau_1}(\ln(\alpha a) + \gamma - \ln 2)\right\}, \tag{236}$$

γ being the Euler constant. An approximate criterion to determine if intermittence is useful can be obtained by expanding t_m in powers of $1/\tau_1$ when $\tau_1 \to \infty$ ($\tau_1 \to \infty$ corresponds to the absence of intermittence), and requiring that the coefficient of

the term $1/\tau_1$ is negative for all values of τ_2. Using this criterion, we find that intermittence is useful if

$$\sqrt{2}\exp(-7/4+\gamma)Vb/D - 4\ln(b/a) + 3 > 0, \tag{237}$$

In this regime, using Eq(236), the optimization of the search time leads to :

$$\tau_1^{opt} = \frac{b^2}{D}\frac{4\ln w - 5 + c}{w^2(4\ln w - 7 + c)},\ \tau_2^{opt} = \frac{b}{V}\frac{\sqrt{4\ln w - 5 + c}}{w}, \tag{238}$$

where w is the solution of the implicit equation $w = 2Vbf(w)/D$ with :

$$\frac{\sqrt{4\ln w - 5 + c}}{f(w)} = -8(\ln w)^2 + (6+8\ln(b/a))\ln w - 10\ln(b/a) + 11 - c(c/2 + 2\ln(a/b) - 3/2), \tag{239}$$

and $c = 4(\gamma - \ln(2))$, γ being the Euler constant. An useful approximation for w is given by :

$$w \simeq \frac{2Vb}{D} f\left(\frac{Vb}{2D\ln(b/a)}\right). \tag{240}$$

The gain for this optimal strategy reads :

$$gain = \frac{t_{\rm diff}}{t_m^{opt}} \simeq \frac{1}{2}\frac{4\ln b/a - 3 + 4a^2/b^2 - a^4/b^4}{4\ln b/a - 3 + 2(4\ln w)\ln(b/aw)}\left(\frac{1}{4\ln w - 5} + \frac{wD}{bV}\frac{4\ln w - 7}{(4\ln w - 5)^{3/2}}\right)^{-1}. \tag{241}$$

If intermittence significantly speeds up the search in this regime (typically by a factor 2), it does not change the order of magnitude of the search time.

$D/V \ll a \ll b$: "universal" regime of intermittence

In the last regime $D/V \ll a \ll b$, the optimal strategy is obtained for :

$$\tau_1^{opt} \simeq \frac{D}{2V^2}\frac{\ln^2(b/a)}{2\ln(b/a) - 1},\ \tau_2^{opt} \simeq \frac{a}{V}(\ln(b/a) - 1/2)^{1/2}. \tag{242}$$

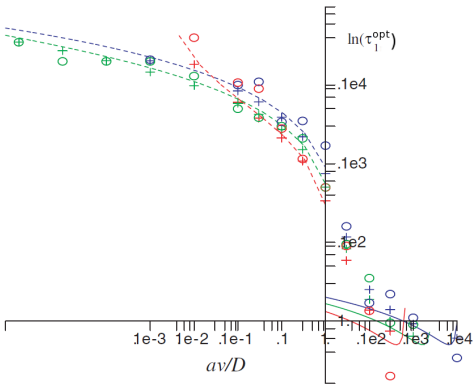
and the gain reads :

$$gain = \frac{t_{\rm diff}}{t_m^{opt}} \simeq \frac{\sqrt{2}aV}{8D}\left(\frac{1}{4\ln(b/a) - 3}\frac{\mathrm{I}_0\left(2/\sqrt{2\ln(b/a) - 1}\right)}{\mathrm{I}_1\left(2/\sqrt{2\ln(b/a) - 1}\right)} + \frac{1}{2\sqrt{2\ln(b/a) - 1}}\right)^{-1}. \tag{243}$$

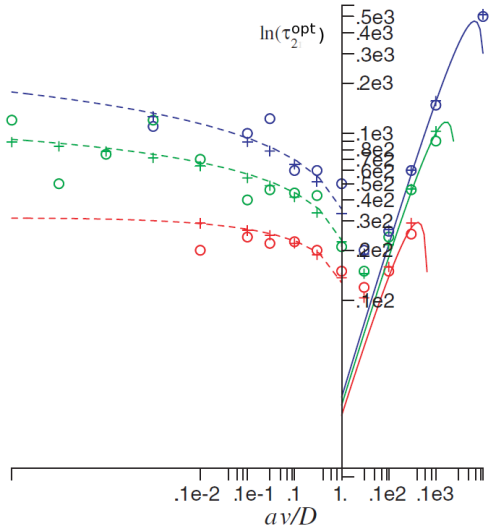
Here, the optimal strategy leads to a significant decrease of the search time which can be rendered arbitrarily smaller than the search time in absence of intermittence.

Summary

We have studied the case where the detection phase 1 is modeled by the diffusive mode, and obtained an approximation of the mean first passage time to the target. We have found that intermittence is favorable (*i.e.* better than diffusion alone), in the regime of large system size $b \gg D/V$. The optimal intermittent strategy then follows two subregimes :



$\ln(\tau_1^{opt})$ as a function of $\ln(aV/D)$



$\ln(\tau_2^{opt})$ as a function of $\ln(aV/D)$

Figure 74: Values of τ_1 and τ_2 at the minimum of t_m, obtained via the approximation in the case $a \ll D/V$ (238) (discontinuous line), via the approximation in the case $a \gg D/V$ (242) (continuous line), via the minimization of analytical t_m (230) (+), and via the simulations ($\circ$) (which are not very precise due to the flatness of the minimum), for $V = 1$, $D = 1$, $b = 113$ (red), $b = 451$ (light green), and $b = 1800$ (dark blue).

- if $a \ll D/V$, the best strategy is given by (238). The search is significantly reduced by intermittence but keeps the same order of magnitude as in the case of a single state diffusive search.
- if $a \gg D/V$, the best strategy is given by (242), and weakly depends on b. In this regime, intermittence is very efficient as shown by the large gain obtained for V large.

5.4.3 Ballistic mode

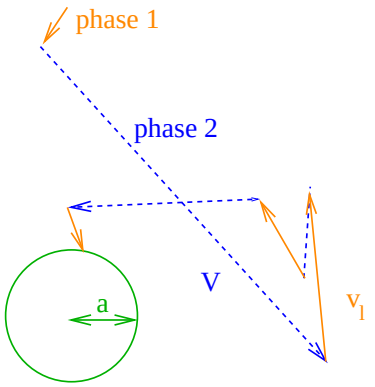


Figure 75: Ballistic mode in two dimensions.

In this case, the searcher has access to two different speeds: one (V) is fast but

prevents the searcher from finding its target, and the other one (v_l) is slower but enables the searcher to detect it target (see figure 75).

Simulations

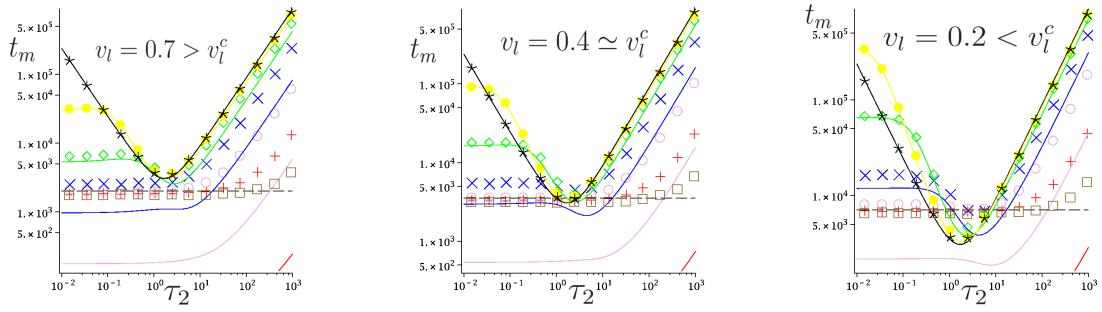


Figure 76: Ballistic mode in two dimensions. $\ln(t_m)$ as a function of $\ln(\tau_2)$. Simulations (symbols), diffusive/diffusive approximation (230) with (247) (colored lines), $\tau_1 \to 0$ limit (248) (black line), $\tau_1 \to \infty$ (no intermittence) (246) (dotted black line). $b = 30$, $a = 1$, $V = 1$. $\tau_1 = 0$ (black, $\star$), $\tau_1 = 0.17$ (yellow, $\bullet$), $\tau_1 = 0.92$ (green, $\diamond$), $\tau_1 = 5.0$ (blue, $\times$), $\tau_1 = 28$ (purple, $\circ$), $\tau_1 = 150$ (red, $+$), $\tau_1 = 820$ (brown, $\square$).

Since an explicit expression of the mean search time is not available, a numerical study is performed. Exploring the parameter space numerically enables to identify the regimes where the mean search time is minimized. Then, for each regime, approximation schemes are developed to provide analytical expression of the mean search time. The numerical results presented in figure 76 suggest two regimes defined according to a threshold value v_l^c of v_l to be determined later on :

- for $v_l > v_l^c$, t_m is minimized for $\tau_2 \to 0$
- for $v_l < v_l^c$, t_m is minimized for $\tau_1 \to 0$.

Regime without intermittence ($\tau_2 \to 0$, $\tau_1 \to \infty$)

Qualitatively, it is rather intuitive that for v_l large enough (the precise threshold value v_l^c will be determined next), phase 2 is inefficient since it does not allow for target detection. The optimal strategy is therefore $\tau_2 \to 0$ in this case. In this regime, the searcher performs a ballistic motion, which is randomly reoriented with frequency $1/\tau_1$. Along the same line as in Viswanathan et al. [1999] (where however the times between successive reorientations are Lévy distributed), it can be shown that the optimal strategy to find a target (which is assumed to disappear after the first encounter) is to minimize oversampling and therefore to perform a purely ballistic motion. In our case this means that in the regime $\tau_2 \to 0$, the optimal τ_1 is given by $\tau_1^{opt} \to \infty$.

In this regime, we can propose an estimate of the optimal search time t_{bal}. The surface scanned during δt is $2av_l\delta t$. $p(t)$ is the proportion of the total area which has not yet been scanned at t. If we neglect correlations in the trajectory, $p(t)$ is solution of :

$$\frac{dp}{dt} = -\frac{2av_l p(t)}{\pi b^2}. \tag{244}$$

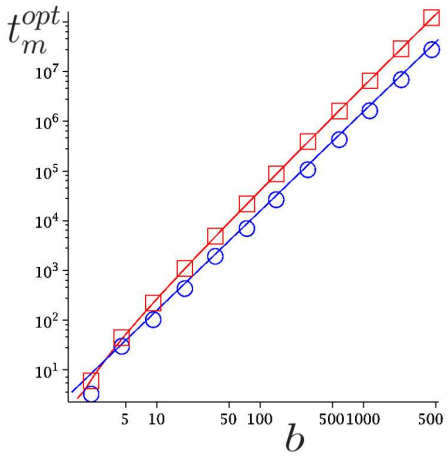


Figure 77: Ballistic mode in two dimension. t_m^{opt} as a function of b, logarithmic scale. Regime without intermittence ($\tau_2 = 0$ and $\tau_1 \to \infty$, $v_l = 1$), analytical approximation (246) (blue line), numerical simulations (○). Regime with intermittence (with $\tau_1 = 0$, $\tau_2 = \tau_2^{opt}$, $V = 1$), analytical approximation (250) (red line), numerical simulations (□). $a = 1$.

Then, given that $p(t = 0) = 1$, we obtain :

$$p(t) = \exp\left(-\frac{2av_l}{\pi b^2}\right), \tag{245}$$

and the mean first passage time to the target in these conditions is :

$$t_{bal} = -\int_0^\infty t\frac{dp}{dt}dt = \frac{\pi b^2}{2av_l}. \tag{246}$$

This expression yields (see figure 77) a good agreement with numerical simulations. Note in particular that $t_{bal} \propto \frac{1}{v_l}$.

Regime with intermittence $\tau_1 \to 0$

In this regime where $v_l < v_l^c$, the numerical study shows that the search time is minimized for $\tau_1 \to 0$ ((see figure 76)). We here determine the optimal value of τ_2 in this regime. To proceed we approximate the problem by the case of a diffusive mode previously studied (230), with an effective diffusion coefficient :

$$D = \frac{v_l^2\tau_1}{2}. \tag{247}$$

This approximation is very satisfactory in the regime $\tau_1 \to 0$ as shown in (see figure 76).

We can then use the results of the previous section for the diffusive mode in the $\tau_1 \to 0$ regime and obtain:

$$t_m = \tau_2\left(1 - \frac{a^2}{b^2}\right)\left(1 - \frac{1}{4}\frac{\left(3 + 4\ln\left(\frac{a}{b}\right)\right)b^4 - 4a^2b^2 + a^4}{\tau_2^2V^2\left(b^2 - a^2\right)} + \frac{a}{V\tau_2\sqrt{2}}\left(\frac{b^2}{a^2} - 1\right)\frac{I_0\left(\frac{a\sqrt{2}}{\tau_2 V}\right)}{I_1\left(\frac{a\sqrt{2}}{\tau_2 V}\right)}\right). \tag{248}$$

The calculation of τ_2^{opt} minimizing t_m then gives:

$$\tau_2^{opt} = \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}. \tag{249}$$

In turn, replacing τ_2 by τ_2^{opt} (249) in (248), we obtain the minimal mean time of target detection :

$$t_m^{opt} = \frac{a}{u\sqrt{2}V}\left(1 - \frac{a^2}{b^2}\right)\left(1 - u^2\frac{\left(3 + 4\ln\left(\frac{a}{b}\right)\right)b^4 - 4a^2b^2 + a^4}{a^2 2\left(b^2 - a^2\right)} + \frac{u\left(\frac{b^2}{a^2} - 1\right)I_0\left(2u\right)}{I_1\left(2u\right)}\right), \tag{250}$$

with $u = \left(\ln\left(\frac{b}{a}\right) - 1\right)^{-\frac{1}{2}}$. It can be noticed that $t_m^{opt} \propto \frac{1}{V}$. Note that if $b \gg a$ this last expression can be greatly simplified:

$$t_m^{opt} \simeq \frac{2b^2}{aV}\sqrt{\ln\left(\frac{b}{a}\right)}. \tag{251}$$

Finally the gain reads (using (246) and (251)):

$$gain = \frac{t_{bal}}{t_m^{opt}} \simeq \frac{\pi V}{4v_l}\left(\ln\left(\frac{b}{a}\right)\right)^{-0.5}. \tag{252}$$

Numerical simulations of figure 77 show the validity of these approximations.

Determination of v_l^c

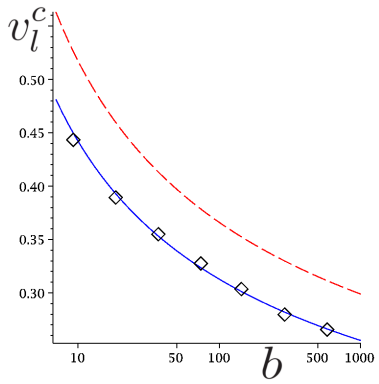


Figure 78: Ballistic mode in two dimension. v_l^c as a function of $\ln(b)$ by simulations (symbols), predicted expression (253) (red dotted line), predicted expression multiplied by a fitted numerical constant (blue line). $V = 1$, $a = 1$.

It is straightforward than $v_l^c < V$. Indeed, if $v_l = V$, phase 2 is useless, since the target cannot be detected. Actually, an estimate of v_l^c can be obtained from (252) as the value of v_l for which $gain = 1$:

$$v_l^c \simeq \frac{\pi V}{4}\left(\ln\left(\frac{b}{a}\right)\right)^{-0.5} \propto \frac{V}{\sqrt{\ln(b/a)}}. \tag{253}$$

We note that this expression ((see figure 78)) gives the correct dependence on b, but it however departs from the value obtained by numerical simulations. This is due to fact that the expression of t_m^{opt} with intermittence (251) is underestimated, while t_{bal} given in (246) is overestimated. It is noteworthy that intermittence is less favorable with increasing b. This effect is similar to the 1 dimensional case, even though it is less important here. It can be understood as follows: at very large scales the intermittent trajectory is reoriented many times and therefore scales as diffusion, which is less favorable than the non intermittent ballistic motion.

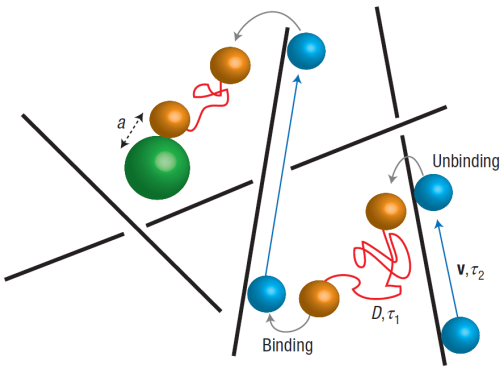
Summary

We have studied the case of the ballistic mode for the detection phase in two dimensions. When $v_l > v_l^c$, the optimal strategy is to remain in phase 1 and to explore the domain in a purely ballistic way. Therefore, $\tau_2^{opt} \to 0$, $\tau_1^{opt} \to \infty$. When $v_l < v_l^c$, we find on the contrary $\tau_1^{opt} \to 0$ and $\tau_2^{opt} = \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$. The threshold value is given by $v_l^c \propto \frac{V}{\sqrt{\ln(b/a)}}$ and shows that when the target density decreases, intermittence is less favorable.

5.4.4 Conclusion in two dimensions

Remarkably, for the three different modes of detection (static, diffusive and ballistic), we find a regime where intermittence minimizes the search time for one and the same τ_2^{opt}, given by $\tau_2^{opt} = \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$. As in one dimension, this indicates that optimal intermittent strategies are robust and widely independent of the details of the description of the detection mechanism.

5.5 Three dimensions



microscopic scale : vesicles in the bulk cytoplasm of a cell

Macroscopic scale : fish lives in a 3-dimensional space

Figure 79: Biological examples in 3 dimensions.

The 3 dimensional case is also relevant to biology (see figure 79). At the microscopic scale, it corresponds for example to intracellular traffic in the bulk cytoplasm

of cells, or at larger scales to animals living in 3 dimensions, such as plankton [Bartumeus et al., 2003], or *C.elegans* in its natural habitat (soil) [Kiontke and Sudhaus, 2005]. As it was the case in two dimensions, different assumptions have to be made to obtain analytical expressions of the search time. We checked the validity of our assumptions with numerical simulations using the same algorithms as in two dimensions.

5.5.1 Static mode

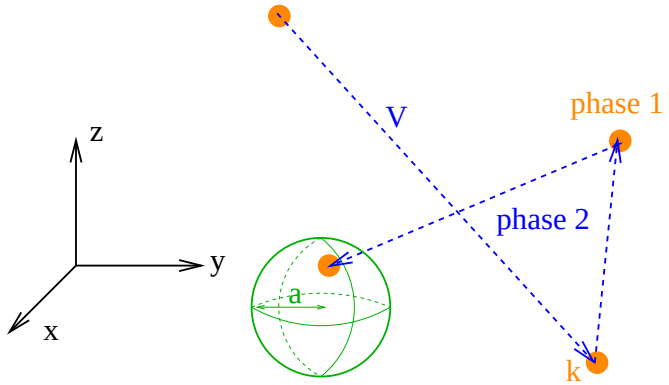


Figure 80: Static mode in three dimensions.

We study in this section the case where the detection phase is modeled by the static mode, for which the searcher does not move during the detection phase and has a finite reaction rate with the target if it is within a detection radius a (see figure 80).

Equations

Denoting $t_1(r)$ the mean first passage time to the target starting from a distance r from the target in phase 1 (detection phase), and $t_{2,\theta,\phi}(r)$ the mean first passage time to the target starting from a distance r from the target in phase 2 (relocation phase) with a ballistic motion in a direction characterized by θ and ϕ, we obtain :

$$\overrightarrow{V}.\overrightarrow{\bigtriangledown} t_{2,\theta,\phi} + \frac{1}{\tau_2}\left(t_1 - t_{2,\theta,\phi}\right) = -1. \tag{254}$$

Then outside the target $(r > a)$:

$$\frac{1}{\tau_1}\left(\frac{1}{4\pi}\int_0^{\pi} d\theta sin\theta \int_0^{2\pi} d\phi t_{2,\theta,\phi} - t_1\right) = -1, \tag{255}$$

and inside the target $(r \leq a)$:

$$\frac{1}{\tau_1}\frac{1}{4\pi}\int_0^{\pi} d\theta sin\theta \int_0^{2\pi} d\phi t_{2,\theta,\phi} - \left(\frac{1}{\tau_1} + k\right) t_1 = -1. \tag{256}$$

With $t_2 = \frac{1}{4\pi}\int_0^{\pi} d\theta sin\theta \int_0^{2\pi} d\phi t_{2,\theta,\phi}$, we obtain outside the target ($r > a$) :

$$\frac{1}{\tau_1}\left(t_2 - t_1\right) = -1, \tag{257}$$

and inside the target ($r < a$) :

$$\frac{1}{\tau_1} t_2 - \left(\frac{1}{\tau_1} + k\right) t_1 = -1. \tag{258}$$

Making a similar decoupling approximation as in two dimensions, we finally obtain :

$$\frac{V^2\tau_2}{3} \triangle\, t_2 - \frac{1}{\tau_2}(t_1 - t_2) = -1. \tag{259}$$

We solve these equations inside and outside the target, using the following boundary conditions :

$$\left.\frac{dt_2^{out}}{dr}\right|_{r=b} = 0, \tag{260}$$

$$t_2^{out}(a) = t_2^{in}(a), \tag{261}$$

$$\left.\frac{dt_2^{out}}{dr}\right|_{r=a} = \left.\frac{dt_2^{in}}{dr}\right|_{r=a}, \tag{262}$$

and the condition that $t_2^{in}(0)$ should be finite.

Results

We find an explicit expression of the mean search time :

$$t_m = (\tau_1 + \tau_2)\left(\frac{1}{k\tau_1} + \frac{1}{b^3V^2\tau_2^2}\left(-2\,b^3\left(b^2 - a^2\right) + \left(b^3 - a^3\right)\left(3\,\frac{a^2}{\alpha^2} + \beta\right) + \frac{1}{5}(b^5 - a^5)\right)\right), \tag{263}$$

with :

$$\beta = \frac{-\sinh(\alpha)\,a^3 + \alpha\,\cosh(\alpha)\,b^3}{a\left(-\sinh(\alpha) + \alpha\,\cosh(\alpha)\right)}, \tag{264}$$

and $\alpha = \sqrt{3\frac{k\tau_1}{1+k\tau_1}}\frac{a}{V\tau_2}$.

In the limit $b \gg a$, this can be simplified to :

$$t_m = (\tau_1 + \tau_2)\left(\frac{1}{k\tau_1} + \frac{1}{\tau_2^2V^2}\left(\frac{-\sinh(\alpha)\,a^3 + \alpha\,\cosh(\alpha)\,b^3}{a\left(-\sinh(\alpha) + \alpha\,\cosh(\alpha)\right)} - \frac{9}{5}b^2 + \frac{3a^2}{\alpha^2}\right)\right). \tag{265}$$

Assuming further that α is small, we use the expansion $\beta \simeq \frac{b^3}{a}\left(1 - \tanh(\alpha)/\alpha\right)^{-1} \simeq \frac{b^3}{a}\left(\frac{3}{\alpha^2} + \frac{6}{5}\right)$ and rewrite mean search time as :

$$t_m = \frac{b^3(\tau_2 + \tau_1)}{a}\left(\frac{(1 + k\tau_1)}{\tau_1 k a^2} + \frac{6}{5\tau_2^2V^2}\right). \tag{266}$$

This expression of t_m can be minimized for :

$$\tau_1^{opt} = \left(\frac{3}{10}\right)^{\frac{1}{4}}\sqrt{\frac{a}{Vk}}, \tag{267}$$

$$\tau_2^{opt} = \sqrt{1.2}\frac{a}{V}, \tag{268}$$

and the minimum mean search time finally reads :

$$t_m^{opt} = \frac{1}{\sqrt{5}}\frac{1}{k}\frac{b^3}{a^3}\left(\sqrt{\frac{ak}{V}}24^{1/4} + 5^{1/4}\right)^2. \tag{269}$$

Comparisons with simulations

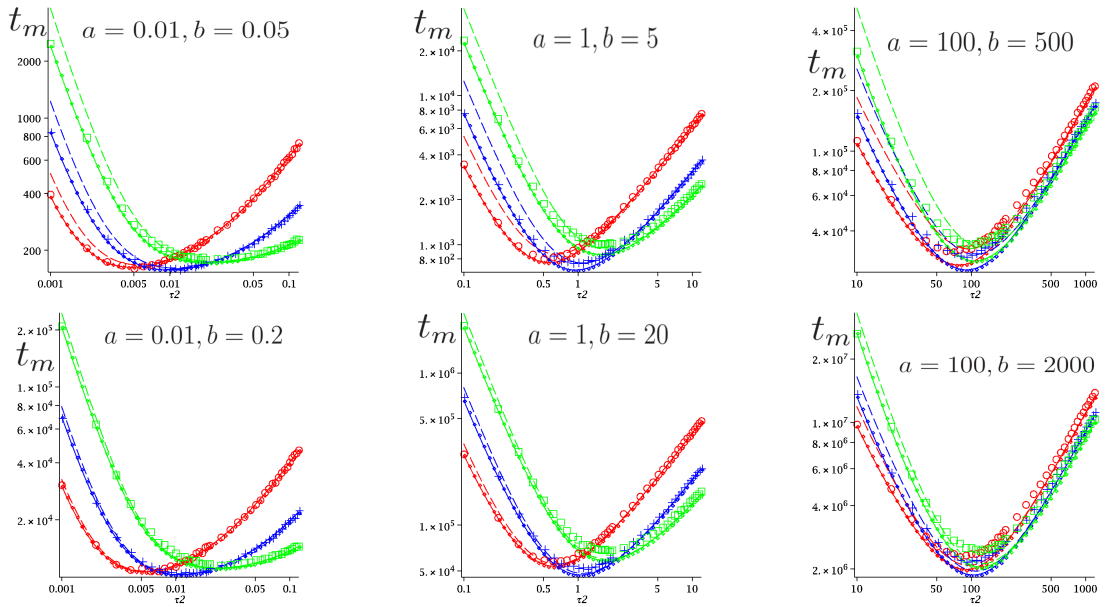


Figure 81: Static mode in 3 dimensions. $\ln(t_m)$ as a function of $\ln(\tau_2)$ for different values of τ_1, a and b/a. Comparison between simulations (symbols), analytical expression (263) (line), expression for $b \gg a$ (265) (small dots), and simple expression for $b \gg a$ and α small (266) (dashed line). $\tau_1 \simeq \tau_1^{opt} \simeq 0.74\sqrt{\frac{aV}{k}}$ (267) (blue, +), $\tau_1 = 0.25\sqrt{\frac{aV}{k}}$ (red, $\circ$), $\tau_1 = 2.5\sqrt{\frac{aV}{k}}$ (green, $\square$). $V = 1$, $k = 1$.

Data obtained by numerical simulations (figure 81, and additionally in appendix 8.2.3) are in good agreement with the analytical expression (263). In particular, the position of the minimum is very well approximated, and the error on the value of the mean search time at the minimum is close to 10%. Note that the very simple expression (266) fits also rather well the numerical data, except for small τ_2 or small b.

Summary

In the case of a static detection mode in three dimensions, we obtained a simple approximation of the mean first passage time to the target $t_m = \frac{b^3(\tau_2+\tau_1)}{a}\left(\frac{(1+k\tau_1)}{\tau_1 k a^2} + \frac{6}{5\tau_2^2 V^2}\right)$. t_m has a single minimum for $\tau_1^{opt} = \left(\frac{3}{10}\right)^{\frac{1}{4}}\sqrt{\frac{a}{Vk}}$ and $\tau_2^{opt} = \sqrt{1.2}\frac{a}{V}$, and the minimal mean search time is $\frac{1}{\sqrt{5}}\frac{1}{k}\frac{b^3}{a^3}\left(\sqrt{\frac{ak}{V}}24^{1/4} + 5^{1/4}\right)^2$. With the static detection mode, intermittence is always favorable and leads to a single optimal intermittent strategy.

As in one and two dimensions, the optimal duration of the relocation phase does not depend on k, *i.e.* on the description of the detection phase. In addition, this optimal strategy does not depend on the typical distance between targets b.

One can notice than for the static mode in the three cases studied (1, 2, and 3 dimensions), we have the relation : $\tau_1^{opt} = \sqrt{\tau_2^{opt}/(2k)}$. This relation between the optimal durations of the two phases is independent from the dimension.

5.5.2 Diffusive mode

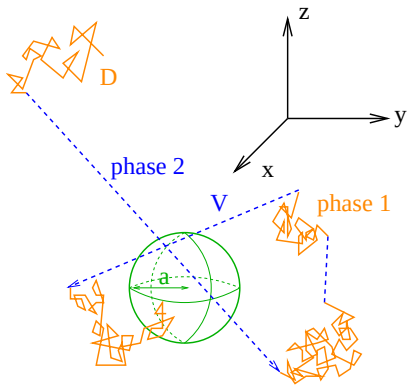


Figure 82: Diffusive mode in three dimensions.

We now study the case where the detection phase is modeled by a diffusive mode. During the detection phase, the searcher diffuses and detects the target as soon as their respective distance is less than a (see figure 82).

Equations

Outside the target ($r > a$), we write :

$$\overrightarrow{V}.\overrightarrow{\bigtriangledown} t_{2,\theta,\phi} + \frac{1}{\tau_2}\left(t_1 - t_{2,\theta,\phi}\right) = -1, \tag{270}$$

$$D \bigtriangleup t_1 + \frac{1}{\tau_1}\left(\frac{1}{4\pi}\int_0^{\pi} d\theta sin\theta \int_0^{2\pi} d\phi t_{2,\theta,\phi} - t_1\right) = -1, \tag{271}$$

and inside the target ($r \leq a$) :

$$\overrightarrow{V}.\overrightarrow{\bigtriangledown} t_{2,\theta,\phi} - \frac{1}{\tau_2} t_{2,\theta,\phi} = -1, \tag{272}$$

$$t_1 = 0. \tag{273}$$

With $t_2 = \frac{1}{4\pi}\int_0^{\pi} d\theta sin\theta \int_0^{2\pi} d\phi t_{2,\theta,\phi}$, we get outside the target ($r > a$) :

$$D \bigtriangleup t_1^{out} + \frac{1}{\tau_1}\left(t_2^{out} - t_1^{out}\right) = -1. \tag{274}$$

The decoupling approximation described in previous sections then yields outside the target :

$$\frac{V^2\tau_2}{3} \triangle t_2^{out} + \frac{1}{\tau_2}(t_1^{out} - t_2^{out}) = -1, \tag{275}$$

and inside the target ($r \leq a$) :

$$\frac{V^2\tau_2}{3} \triangle t_2^{int} - \frac{1}{\tau_2} t_2^{int} = -1. \tag{276}$$

These equations are completed by the following boundary conditions :

$$\left.\frac{dt_2^{out}}{dr}\right|_{r=b} = 0, \tag{277}$$

$$t_2^{out}(a) = t_2^{int}(a), \tag{278}$$

$$\left.\frac{dt_2^{out}}{dr}\right|_{r=a} = \left.\frac{dt_2^{int}}{dr}\right|_{r=a}. \tag{279}$$

Results in the general case

Through standard but lengthy calculations we solve the above system and obtain an analytical approximation of t_m (see appendix 8.2.4). In the regime $b \gg a$, we use the assumption $b\sqrt{(\tau_1 D)^{-1} + 3(\tau_2 v)^{-2}} \gg 1$ and obtain :

$$t_m = \frac{b^3\kappa_2^4(\tau_1+\tau_2)}{\kappa_1} \frac{\tanh(\kappa_2 a) + \frac{\kappa_1}{\kappa_2}}{\kappa_1\kappa_2^2\tau_1 Da\left(\tanh(\kappa_2 a) + \frac{\kappa_1}{\kappa_2}\right) - \tanh(\kappa_2 a)} \tag{280}$$

with $\kappa_1 = \frac{\sqrt{\tau_2^2 V^2 + 3\tau_1 D}}{\tau_2 V\sqrt{D\tau_1}}$ and $\kappa_2 = \frac{\sqrt{3}}{V\tau_2}$. As shown in Fig.83 left or in the additional figure 124 in appendix 8.2.4, t_m only weakly depends on τ_1, which indicates that this variable will be less important than τ_2 in the minimization of the search time. The relevant order of magnitude for τ_1^{opt} can be evaluated by comparing the typical diffusion length $L_{\text{diff}} = \sqrt{6Dt}$ and the typical ballistic length $L_{bal} = Vt$. An estimate of the optimal time τ_1^{opt} can be given by the time scale for which those lengths are of same order, which gives :

$$\tau_1^{opt} \sim \frac{6D}{V^2}. \tag{281}$$

Note that taking $\tau_1 = 0$ does not change significantly t_m^{opt} (see figure 83 left), and enables us to significantly simplify t_m :

$$t_m = \frac{b^3\sqrt{3}}{V^3\tau_2^2}\left(\frac{\sqrt{3}a}{V\tau_2} - \tanh\left(\frac{\sqrt{3}a}{V\tau_2}\right)\right)^{-1}. \tag{282}$$

In turn, the minimization of this expression leads to :

$$\tau_2^{opt} = \frac{\sqrt{3}a}{Vx}, \tag{283}$$

with x solution of :

$$2\tanh(x) - 2x + x\tanh(x)^2 = 0. \tag{284}$$

This finally yields :

$$\tau_2^{opt} \simeq 1.078\frac{a}{V}. \tag{285}$$

Importantly this approximate expression is very close to the expression obtained for the static mode ($\tau_2^{opt} = \sqrt{\frac{6}{5}}\frac{a}{V} \simeq 1.095\frac{a}{V}$) (268), and there is no dependence with the typical distance between targets b. The simplified expression of the minimal t_m (282) can then be obtained as:

$$t_m^{opt} = \frac{b^3x^2}{\sqrt{3}a^2V}\left(x - \tanh(x)\right)^{-1} \simeq 2.18\frac{b^3}{a^2V}, \tag{286}$$

and the gain reads:

$$gain = \frac{t_{\text{diff}}}{t_m^{opt}} \simeq 0.15\frac{aV}{D}. \tag{287}$$

Intermittence is then favorable for $a > a_c \simeq 6D/V$.

Comparison between analytical approximations and numerical simulations

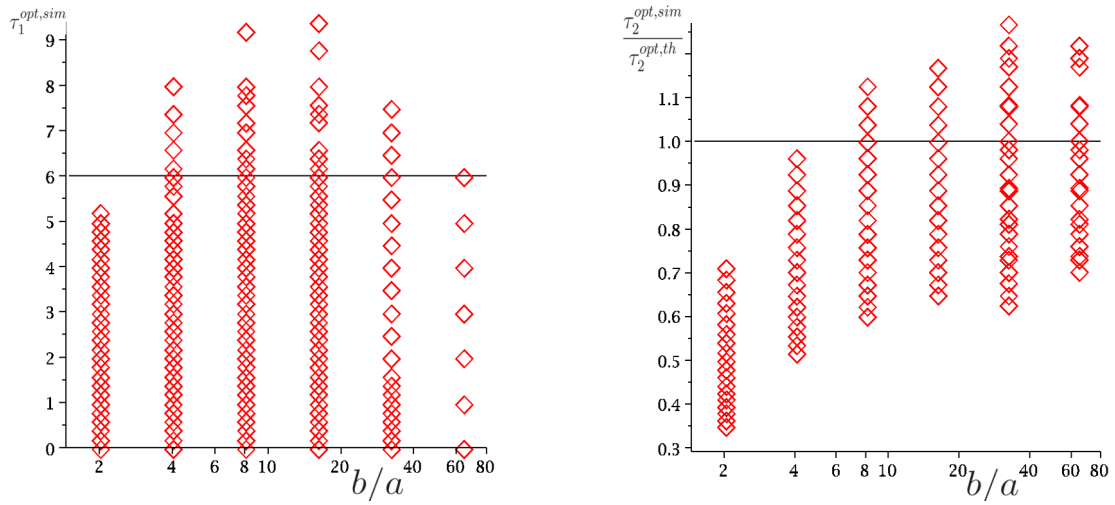


Figure 83: Diffusive mode in 3 dimensions. Comparison between analytical approximations (281) (283) (black lines) and numerical simulations : the symbols are the values of τ_1 and τ_2 for which $t_m^{simulation} < 1.05t_m^{opt,simulation}$. $a = 100$, $V = 1$, $D = 1$.

Numerical simulations reveal that the minimum of t_m with respect to τ_1 is shallow as it was expected (cf Fig.83 left). It approximately ranges from 0 to the theoretical estimate (281). The value $\tau_2^{opt,sim}$ at the minimum is close to the expected values (283) (cf. Fig.83 right), except for very small b, which is consistent with our assumption $b \gg a$. We can then conclude than the position of the optimum in τ_1 and τ_2 is very well described by the analytical approximations, even if the value of t_m at the minimum is underestimated by our analytical approximation by about 10-20% ((see figure 84)).

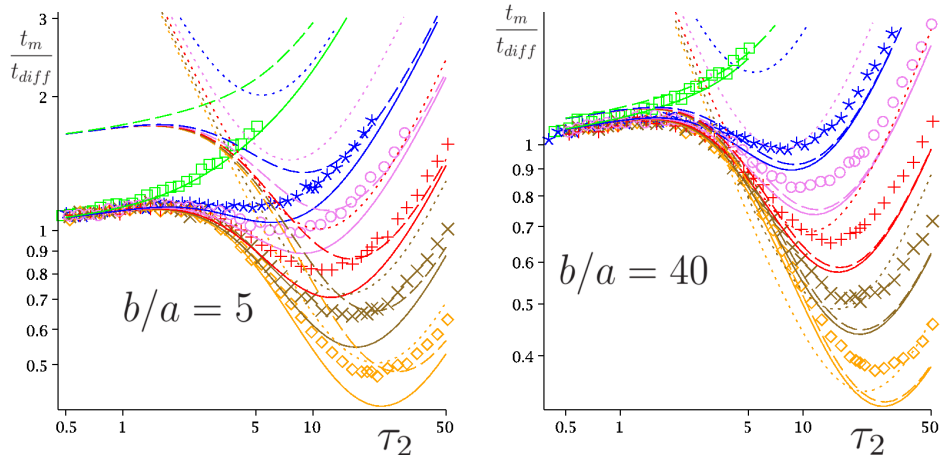


Figure 84: Diffusive mode in 3 dimensions. t_m/t_{diff} (t_{diff} given by (288)) as a function of τ_2 for different values of the ratio b/a (logarithmic scale). The full analytical form (441) (plain lines) is plotted against the simplified expression (280) (dashed lines), the simplified expression with $\tau_1 = 0$ (282) (dotted line), and numerical simulations (symbols) for the following values of the parameters (arbitrary units): $a = 1$ (green, $\square$), $a = 5$ (blue, $\star$), $a = 7$ (purple, $\circ$), $a = 10$ (red, $+$), $a = 14$ (brown, $\times$), $a = 20$ (orange, $\diamond$). $\tau_1 = 6$ everywhere except for the small dots, $V = 1$, $D = 1$. t_m/t_{diff} presents a minimum only for $a > a_c \simeq 4$.

Case without intermittence : single state diffusive searcher

If the searcher always remains in the diffusive mode, it is straightforward to obtain (see appendix 8.2.4):

$$t_{\text{diff}} = \frac{1}{15Dab^3}\left(5b^3a^3 + 5b^6 - 9b^5a - a^6\right), \tag{288}$$

which gives in the limit $b/a \gg 1$:

$$t_{\text{diff}} = \frac{b^3}{3Da}. \tag{289}$$

Criterion for intermittence to be favorable

There is a range of parameters for which intermittence is favorable, as indicated by (see figure 84). Both the analytical expression for t_m^{opt} in the regime without intermittence (288) and with intermittence (286) scale as b^3. However, the dependence on a is different (see appendix 8.2.4). In the diffusive regime, $t_m \propto a^{-1}$, whereas in the intermittent regime $t_m \propto a^{-2}$. This enables us to define a critical a_c, such that when $a > a_c$, intermittence is favorable: $a_c \simeq 6.5\frac{D}{V}$ is the value for which the gain (287) is 1.

Summary

We have studied the case where the detection phase 1 is modeled by a diffusive mode, and calculated explicitly an approximation of the mean first passage time

to the target. We found that intermittence is favorable (*i.e.* better than diffusion alone) when $a > a_c \simeq 6.5\frac{D}{V}$:

- if $a < a_c$, the best strategy is a single state diffusion, without intermittence, and the mean first passage time to the target is $t_m \simeq \frac{b^3}{3Da}$.
- if $a > a_c$, intermittence is favorable. The dependence on τ_1 is not crucial, as long as it is smaller than $\frac{6D}{V^2}$. The value of τ_2 at the optimum is $\tau_2^{opt} \simeq 1.08\frac{a}{V}$. The minimum search time is then $t_m^{opt} \simeq 2.18\frac{b^3}{a^2V}$.

5.5.3 Ballistic mode

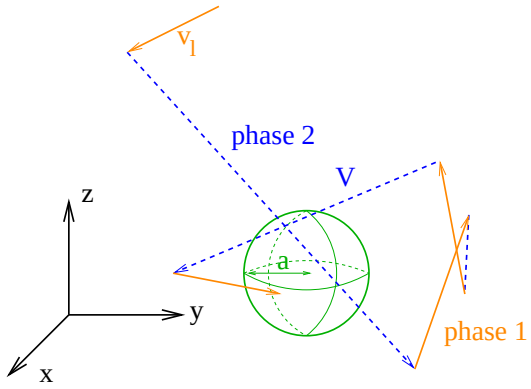


Figure 85: Ballistic mode in three dimensions.

We now discuss the last case, where the detection phase 1 is modeled by a ballistic mode (see figure 85). Since an explicit analytical determination of the search time seems out of reach, we proceed as in two dimensions and first explore numerically the parameter space to identify the regimes where the search time can be minimized. We then develop approximation schemes in each regime to obtain analytical expressions (more details are given in appendix 8.2.5).

Numerical study

The numerical analysis puts forward two strategies minimizing the search time, depending on a critical value v_l^c to be determined ((see figure 86)) :

- when $v_l > v_l^c$, $\tau_1^{opt} \to \infty$ and $\tau_2^{opt} \to 0$. In this regime intermittence is not favorable.
- when $v_l < v_l^c$, $\tau_1^{opt} \to 0$, and τ_2^{opt} is finite. In this regime the optimal strategy is intermittent.

Regime without intermittence (single state ballistic searcher) : $\tau_2 \to 0$

Following the same argument as in two dimensions, without intermittence the best strategy is obtained in the limit $\tau_1 \to \infty$ (see appendix 8.2.5) in order to

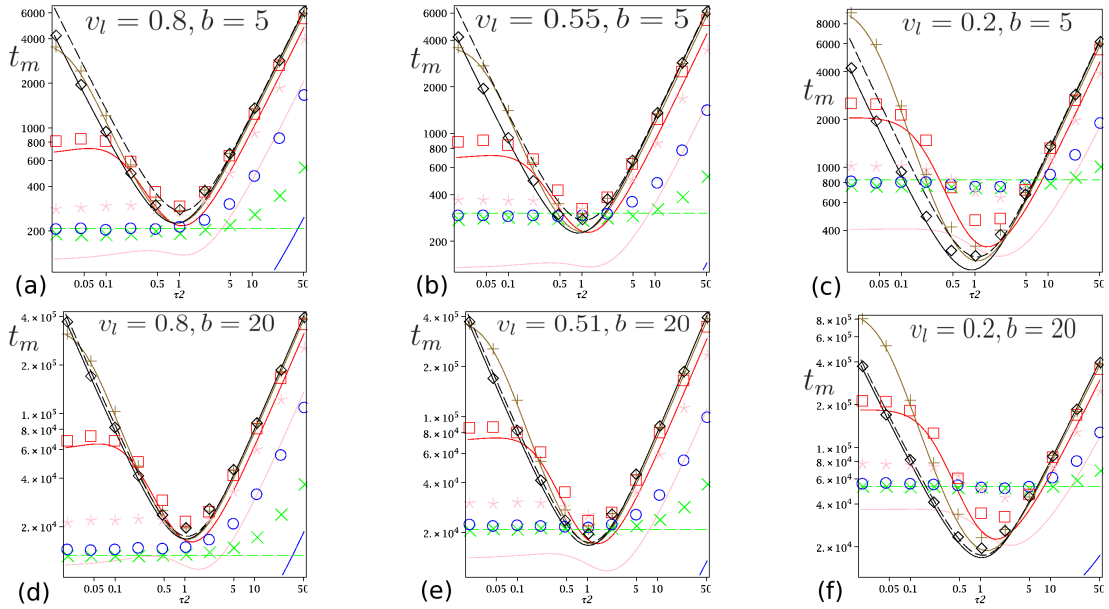


Figure 86: Ballistic mode in 3 dimensions. t_m as a function of τ_2 in loglogscale. Simulations (symbols). Approximation $v_l\tau_1 \leq a$ (441 with 456) (colored lines), approximation $\tau_1 = 0$ (291) (black line), approximation $\tau_1 = 0$ and $b \gg a$ (282) (dotted black line). Ballistic limit ($\tau_2 \to 0$ and $\tau_1 \to \infty$) (no intermittence) (290) (green dotted line). (a),(d) : $v_l > v_l^c$, $\tau_{1,1} = 0.04$, $\tau_{1,2} = 0.2$, $\tau_{1,3} = 1$, $\tau_{1,4} = 5$, $\tau_{1,5} = 25$. (b),(e) : $v_l \simeq v_l^c$, $\tau_{1,1} = 0.08$, $\tau_{1,2} = 0.4$, $\tau_{1,3} = 2$, $\tau_{1,4} = 10$, $\tau_{1,5} = 50$. (c),(f) : $v_l < v_l^c$, $\tau_{1,1} = 0.2$, $\tau_{1,2} = 1$, $\tau_{1,3} = 5$, $\tau_{1,4} = 25$, $\tau_{1,5} = 125$. $V = 1$, $a = 1$. $\tau_1 = 0$ (black, $\diamond$), $\tau_1 = \tau_{1,1}$ (brown, $+$), $\tau_1 = \tau_{1,2}$ (red, $\square$), $\tau_1 = \tau_{1,3}$ (pink, $\star$), $\tau_1 = \tau_{1,4}$ (blue, $\circ$), $\tau_1 = \tau_{1,5}$ (green, $\times$).

minimize oversampling of the search space. Following the derivation of (246) (see appendix for details), it is found that the search time reads :

$$t_{bal} = \frac{4b^3}{3a^2 v_l}. \tag{290}$$

Regime with intermittence

In the regime of favorable intermittence, the numerical study suggests that the best strategy is realized for $\tau_1 \to 0$ (see figure 86). In this regime $\tau_1 \to 0$, the phase 1 can be well approximated by a diffusion with effective diffusion coefficient D_{eff} (see (456)). We can then use the analytical expression t_m derived in (441). We therefore take $\tau_1 = 0$ in the expression of t_m (441), which yields :

$$t_m(\tau_1 = 0) \simeq \frac{u}{b^3 a V}\left(\frac{\sqrt{3}}{5}\left(5b^3a^2 - 3b^5 - 2a^5\right) + \frac{\left(b^3 - a^3\right)^2 u}{\sqrt{3}a(u - \tanh(u))}\right), \tag{291}$$

where $u = \frac{\sqrt{3}a}{\tau_2 V}$. In the limit $b \gg a$, this expression can be further simplified (see (282)) to :

$$t_m = \frac{b^3\sqrt{3}}{V^2\tau_2^2}\left(\frac{\sqrt{3}a}{V\tau_2} - \tanh\left(\frac{\sqrt{3}a}{V\tau_2}\right)\right)^{-1}, \tag{292}$$

and one finds straightforwardly that $\tau_2^{opt} = \frac{\sqrt{3}a}{Vx}$, where x is solution of $x\tanh(x)^2 + 2\tanh(x) - 2x = 0$, that is $x \simeq 1.606$. Using this optimal value of τ_2 in the expression of t_m (282), we finally obtain :

$$t_m^{opt} = \frac{2}{\sqrt{3}} \frac{x}{\tanh(x)^2} \frac{b^3}{a^2 V} \simeq 2.18 \frac{b^3}{a^2 V}. \tag{293}$$

These expressions show a good agreement with numerical simulations (see figures 86 and 87).

Discussion of the critical value v_l^c

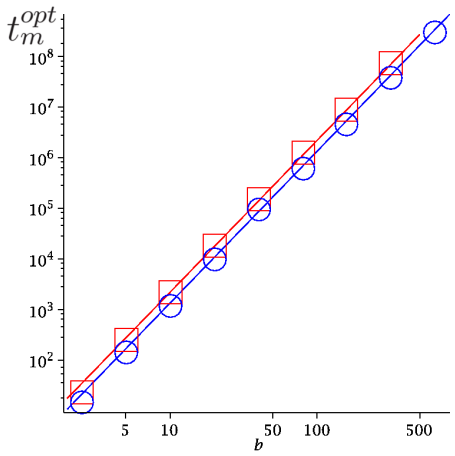


Figure 87: Ballistic mode in 3 dimension. t_m^{opt} as a function of b, logarithmic scale. Regime without intermittence ($\tau_2 = 0$ and $\tau_1 \to \infty$, $v_l = 1$), analytical approximation (290) (blue line), numerical simulations (○). Regime with intermittence (with $\tau_1 = 0$, $V = 1$), analytical approximation (293) (red line), numerical simulations (□). $a = 1$.

The gain is given by :

$$gain = \frac{t_{bal}}{t_m^{opt}} \simeq 0.61 \frac{V}{v_l}. \tag{294}$$

As in two dimensions, it is trivial that $v_l^c < V$, and the critical value v_l^c can be defined as the value of v_l such that $gain = 1$. This yields :

$$v_l^c \simeq 0.6V. \tag{295}$$

Importantly, v_l^c neither depends on b nor a. Simulations are in good agreement with this result (see appendix 8.2.5), except for a small numerical shift.

Summary

We have studied the case where the detection phase 1 is modeled by a ballistic mode in three dimensions. We have shown by numerical simulations that there are two possible optimal regimes, that we have then studied analytically.

- In the first regime $v_l > v_l^c$, the optimal strategy is a single state ballistic search ($\tau_1 \to \infty$, $\tau_2 = 0$) and $t_m \simeq \frac{4b^3}{3a^2 v_l}$.

- In the second regime $v_l < v_l^c$, the optimal strategy is intermittent ($\tau_1 = 0$, $\tau_2 \simeq 1.1\frac{a}{V}$), and $t_m \simeq 2.18\frac{b^3}{a^2 V}$ (in the limit $b \gg a$).

The critical speed is obtained numerically as $v_l^c \sim 0.5V$ (analytical prediction : $v_l^c \sim 0.6V$). It is noteworthy that when $b \gg a$, the values of τ_1 and τ_2 at the optimum, and the value of v_l^c does not depend on the typical distance between targets b.

5.5.4 Conclusion in three dimensions

We have found that for the three possible modelings of the detection mode (static, diffusive and ballistic) in three dimensions, there is a regime where the optimal strategy is intermittent. Remarkably, and as was the case in one and two dimensions, the optimal time to spend in the fast non-reactive phase 2 is independent of the modeling of the detection mode and reads $\tau_2^{opt} \simeq 1.1\frac{a}{V}$. Additionally, while the mean first passage time to the target scales as b^3, the optimal values of the durations of the two phases do not depend on the target density a/b.

5.6 Discussion and conclusion

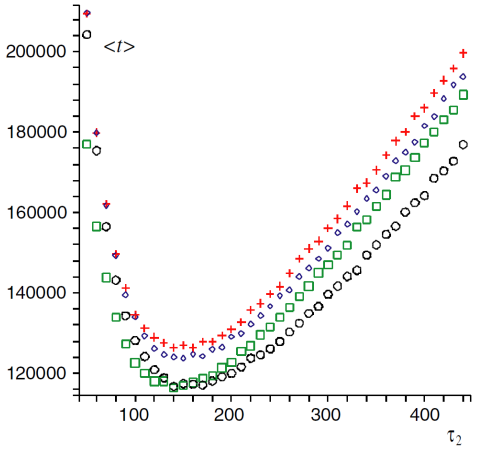


Figure 88: Robustness of the minimum of t_m in τ_2, here for static, diffusive and composite modes in two dimensions. The mean search time t_m as a function of τ_2, with $\tau \simeq \tau_1^{opt}$, for different descriptions of the slow reactive phase : diffusive mode ($D = 1$, $k = \infty$, $\tau_1 = 9.19$) ($\circ$), static mode ($D = 0$, $k = 1$, $\tau_1 = 8.8$) ($+$), composite mode ($D = 1$, $k = 100$, $\tau_1 = 0.165$) ($\square$), and another composite mode ($D = 1$, $k = 1$, $\tau_1 = 10$) ($\diamond$). For all the simulations, $a = 100$, $b = 1800$, $V = 1$.

The starting point of this model is the observation that intermittent trajectories are observed in various biological examples of search behaviors, ranging from the microscopic scale, where searchers can be molecules looking for reactants, to the macroscopic scale of foraging animals. In these different problems, we have developed particular models, which share common characteristics. More precisely, this model is a generalization of the models of sections 3.4 and 4.2. We addressed the

general question of determining whether such kind of intermittent trajectories could be favorable from a purely kinetic point of view, that is whether they could allow to minimize the search time for a target. On very general grounds, we have proposed a generic model of search strategy based on intermittence, where the searcher switches between two phases, one slow where detection is possible, the other one faster but preventing target detection. This model is minimal in the sense that the search is assumed to have very limited memory skills. We have studied this generic model in one, two and three dimensions, and under several modeling hypotheses. We believe that this systematic analysis can be used as a basis to study quantitatively various real search problems involving intermittent behaviors. In addition, it provides the framework for analyzing the mean first passage time of intermittent random walks.

More precisely, we have calculated the mean first passage time to the target for an intermittent searcher, and minimized this search time as a function of the mean duration of each of the two phases. The table 3 summarizes the results. In particular, this study shows that for certain ranges of the parameters which we have determined, the optimal search strategy is intermittent. In other words, there is an optimal way for the intermittent searcher to tune the mean time it spends in each of the two phases. We have found that the optimal durations of the two phases and the gain of intermittent search (as compared to a single state search) do depend on the target density in one dimension. In particular, the gain can be very high at low target concentration. Interestingly, this dependence is smaller in two dimensions, and vanishes in three dimensions. The fact that intermittent search is more advantageous in low dimensions (1 and 2) can be understood as follows. At large scale, the intermittent searcher of our model effectively performs a random walk, and therefore scans a space of dimension 2. In an environment of one dimension (and critically of two dimensions), the searcher therefore oversamples the space, and it is favorable to perform large jumps to go to previously unexplored areas. On the contrary, in three dimensions, the random walk is transient, and the searcher on average always scans previously unexplored areas, which makes large jumps less beneficial.

Additionally, our results show that, for various modeling choices of the slow reactive phase, there is one and the same optimal duration of the fast non reactive phase, which depends only on the space dimension (see figure 88). This further supports the robustness of optimal intermittent search strategies. Such robustness and efficiency could explain why intermittent trajectories are observed so often, and in various forms.

	Static mode	Diffusive mode			Ballistic mode	
1D	always intermittence	$b < \frac{D}{V}$	$b > \frac{D}{V}$, $a \ll \sqrt{\frac{b}{a}}\frac{D}{V}$	$b > \frac{D}{V}$, $a \gg \sqrt{\frac{b}{a}}\frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq \frac{V}{2}\sqrt{\frac{3a}{b}}$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt{\frac{a}{2Vk}}\sqrt[4]{\frac{b}{3a}}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \simeq \left(\frac{b^2 D}{36V^4}\right)^{\frac{1}{3}}$	$\tau_1^{opt} \simeq \frac{Db}{48V^2 a}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \left(\frac{2b^2 D}{9V^4}\right)^{\frac{1}{3}}$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\frac{b}{3a}}$
	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}\left(1+\sqrt{\frac{V}{2ka}}\sqrt[4]{\frac{3a}{b}}\right)^2$	$t_m^{opt} \simeq \frac{b^2}{3D}$	$t_m^{opt} \simeq \left(\frac{3^5 b^4}{2^4 DV^2}\right)^{\frac{1}{3}}$	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}$	$t_m^{opt} \simeq \frac{b}{v_l}$	$t_m^{opt} \simeq \frac{2b}{V}\sqrt{\frac{b}{3a}}$
2D	always intermittence	$b < \frac{D}{V}$	$b \gg \frac{D}{V} \gg a$	$b \gg a \gg \frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq \pi V / \left(4\sqrt{\ln\left(\frac{b}{a}\right)}\right)$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt{\frac{a}{2Vk}}\sqrt[4]{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \simeq \frac{b^2}{D}\frac{4\ln w - 5 + c}{w^2(4\ln w - 7 + c)}$	$\tau_1^{opt} \simeq \frac{D}{2V^2}\frac{\left(\ln\left(\frac{b}{a}\right)\right)^2}{2\ln\left(\frac{b}{a}\right)-1}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{b}{V}\frac{\sqrt{4\ln w - 5 + c}}{w}$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq \frac{a}{V}\sqrt{\ln\left(\frac{b}{a}\right) - \frac{1}{2}}$
	$t_m^{opt} \simeq \frac{2b^2}{aV}\left(\sqrt[4]{\ln\left(\frac{b}{a}\right)} + \sqrt{\frac{V}{2ak}}\right)^2$	$t_m^{opt} \simeq \frac{b^2}{2D}\ln\left(\frac{b}{a}\right)$	t_m^{opt}, c, w : see section5.4.2	$t_m^{opt} \simeq \frac{2b^2}{aV}\sqrt{\ln\left(\frac{b}{a}\right)}$	$t_m^{opt} \simeq \frac{\pi b^2}{2av_l}$	$t_m^{opt} \simeq \frac{2b^2}{aV}\sqrt{\ln\left(\frac{b}{a}\right)}$
3D	always intermittence	$a \lesssim 6\frac{D}{V}$		$b \gg a \gtrsim 6\frac{D}{V}$	$v_l > v_l^c$	$v_l < v_l^c \simeq 0.6V$
	$\tau_1^{opt} = \sqrt{\frac{\tau_2^{opt}}{2k}} \simeq \sqrt[4]{\frac{3}{10}}\sqrt{\frac{a}{Vk}}$	$\tau_1^{opt} \to \infty$		$\tau_1^{opt} \simeq \frac{6D}{V^2}$	$\tau_1^{opt} \to \infty$	$\tau_1^{opt} \to 0$
	$\tau_2^{opt} \simeq 1.1\frac{a}{V}$	$\tau_2^{opt} \to 0$		$\tau_2^{opt} \simeq 1.1\frac{a}{V}$	$\tau_2^{opt} \to 0$	$\tau_2^{opt} \simeq 1.1\frac{a}{V}$
	$t_m^{opt} \simeq \frac{b^3}{Va^2}\left(\sqrt[4]{\frac{24}{5}} + \sqrt{\frac{V}{ak}}\right)^2$	$t_m^{opt} \simeq \frac{b^3}{3Da}$		$t_m^{opt} \simeq 2.2\frac{b^3}{Va^2}$	$t_m^{opt} \simeq \frac{4b^3}{3a^2 v_l}$	$t_m^{opt} \simeq 2.2\frac{b^3}{Va^2}$

Table 3: Recapitulation of the main results of the generic model : strategies minimizing the mean first passage time to the target. In each cell, validity of the regime, optimal τ_1, optimal τ_2, minimal t_m (t_m with $\tau_i = \tau_i^{opt}$). Yellow background highlight the value of τ_2^{opt} independent from the description of the slow detection phase 1. Results are given in the limit $b \gg a$.

6 Extensions and perspectives

Contents

6.1 Beyond the first moment 144

6.1.1 Regularly spaced targets in one dimension with infinite correlation 144

6.1.2 Randomly located targets in one dimension with infinite correlation 148

6.1.3 Regularly spaced targets in two dimensions without correlation 148

6.1.4 Regularly spaced targets in three dimensions without correlation 148

6.1.5 Conclusion on the distribution of the first passage time to the target 148

6.2 Taking into account partial correlations in ballistic phases 150

6.2.1 Motivation 150

6.2.2 Model 150

6.2.3 Infinite range correlations 150

6.2.4 Intermediate correlation 152

6.2.5 An argument on the form of the minimum 153

6.2.6 Optimization with low correlations 154

6.2.7 Conclusion 155

6.3 Other distributions of phases duration 155

6.3.1 Fixed durations of the phases 156

6.3.2 Lévy distribution of the fast phase durations 156

6.4 Other models of intermittent search 159

6.5 Designing efficient searches 163

6.5.1 Microscopic scale 163

6.5.2 Macroscopic scale 164

6.6 Comparison with experimental data 165

6.6.1 Motivation 165

6.6.2 Experiment 165

6.6.3 Data analysis 167

6.6.4 Preliminary results 167

Far from closing the problem, the generic model presented in the previous section on the contrary opens interesting perspectives. We highlight a few promising directions in this section, some of them being currently under investigation. First, in the previous section the quantity optimized is the mean first passage time : in section 6.1 we study the complete distribution of the first passage time. Second,

in the generic model of section 5, the searcher has minimal memory skills. Indeed, on the one hand, the phase duration distribution is exponential, which means that there is no temporal memory : the effect of other duration distributions is studied in section 6.3. On the other hand, the direction of each new ballistic "blind" phase is taken at random, independently of the previous phases, meaning that there is no spatial memory : we study the effect of correlations in section 6.2. Last, we stress that our works on intermittent search strategies have raised interest, and subsequent models of intermittent search have been developed, as detailed in section 6.4. Further models could also be applied to design efficient searches instead of interpreting biological systems (see section 6.5). Last but not least, we show preliminary results of a comparison of this model with experimental data on humans, which is a return back to the first example of a lost key (see section 6.6).

6.1 Beyond the first moment

Until now, the quantity calculated has been the mean first passage time to the target. Nonetheless, the mean time might not be representative of the typical time. Thus, it is important to know what the complete distribution looks like.

In this section, we study the distribution of the first passage time for the intermittent search with a diffusive detection phase. We expect that the other modes of detection give similar results.

6.1.1 Regularly spaced targets in one dimension with infinite correlation

We start with the case of regularly spaced targets in one dimension with infinite correlation. More precisely, the searcher alternates between phases of diffusion of diffusion coefficient D, where the target can be immediately detected (here the detection radius $a \to 0$), and ballistic blind phases of velocity $+V$ (like in the model of section 3.2). Correlation is infinite in the sense that ballistic phases are all in the same direction. The distance between two successive targets is $2b = L$. They are regularly spaced (or equivalently, there is one centered target on a segment with reflective boundaries).

Analytic distribution

In this case, it is possible to calculate analytically the distribution, at least in some limits.

If $\sqrt{D\tau_1}, V\tau_2 \ll L$, the probability p that the searcher detects the next target is the probability that the searcher is in phase 1 (the detection phase) instead of phase 2 (the fast non-reactive phase) when it passes on the target. We define L_i as the mean span of one phase i. If $L_2 \gg L_1$, the probability p writes :

$$p = \frac{L_1}{L_2}. \tag{296}$$

As phase 2 is ballistic of mean duration τ_2, $L_2 = V\tau_2$. The mean span of a diffusive phase of duration t is $S_1 = 4\sqrt{\frac{Dt}{\pi}}$ (see appendix 8.3.1), then $L_1 = \int_0^\infty dt S_1(t) \exp(-t/\tau_1)/\tau_1 =$

$2\sqrt{D\tau_1}$. Thus p writes :

$$p = 2\frac{\sqrt{D\tau_1}}{V\tau_2}. \tag{297}$$

If the searcher finds the first target, the traveled distance is in $[0, L]$, with uniform probability. Indeed, the starting position is random, taken uniformly in $[0, L]$. The exception is when the starting point if very close to the target (as compared to the typical scales L_1 and L_2) : as the searcher starts in phase 1, if the target is within L_1, the probability of finding it is higher than p.

If the first target is missed, but the second found, there is an additional distance L traveled, leading to a distance traveled in $[L, 2L]$. If the first n targets are missed but the $(n+1)$-th is found, the traveled distance is in $[nL, (n+1)L]$, with uniform probability.

The probability of missing the first n targets but of finding the $(n+1)^{th}$ target is $p(1-p)^n$.

The last step is to convert distance into time. Diffusive phases imply a zero mean displacement, as they are symmetric. Under the assumption that $x \gg V\tau_2, \sqrt{D\tau_1}$ the mean time t to travel over a distance x writes :

$$t = \frac{x}{V\tau_2}\left(\tau_1 + \tau_2\right). \tag{298}$$

In particular, we define τ^* the mean travel time between 2 targets :

$$\tau^* = \frac{L}{V\tau_2}\left(\tau_1 + \tau_2\right). \tag{299}$$

Finally, in the regime $L \gg V\tau_2, \sqrt{D\tau_1}$, the density of first passage time to the target $f(t)$ writes for $t \in [n\tau^*; (n+1)\tau^*]$ $(n \in \mathbb{N})$:

$$f(t) = \frac{p(1-p)^n}{\tau^*} = p(1-p)^n \frac{V\tau_2}{L\left(\tau_1 + \tau_2\right)}. \tag{300}$$

Numerical simulations

If L is much larger than $\sqrt{D\tau_1}, V\tau_2$, then the analytical expression (300) is in agreement with simulations. Else, the distribution converges to an exponential distribution (see figure 89).

The analytical value of p (297) is in agreement with simulations, as soon as $\sqrt{D\tau_1} \ll V\tau_2$ (see figure 90).

Moments of the distribution

Now that we have obtained the complete distribution of the first passage time to the target, all the moments of the distribution can be calculated :

$$\langle t^m \rangle = \sum_{n=0}^{\infty} p(1-p)^n \int_{n\tau^*}^{(n+1)\tau^*} t^m \frac{dt}{\tau^*} = (\tau^*)^m \sum_{n=0}^{\infty} p(1-p)^n \int_n^{n+1} u^m du. \tag{301}$$

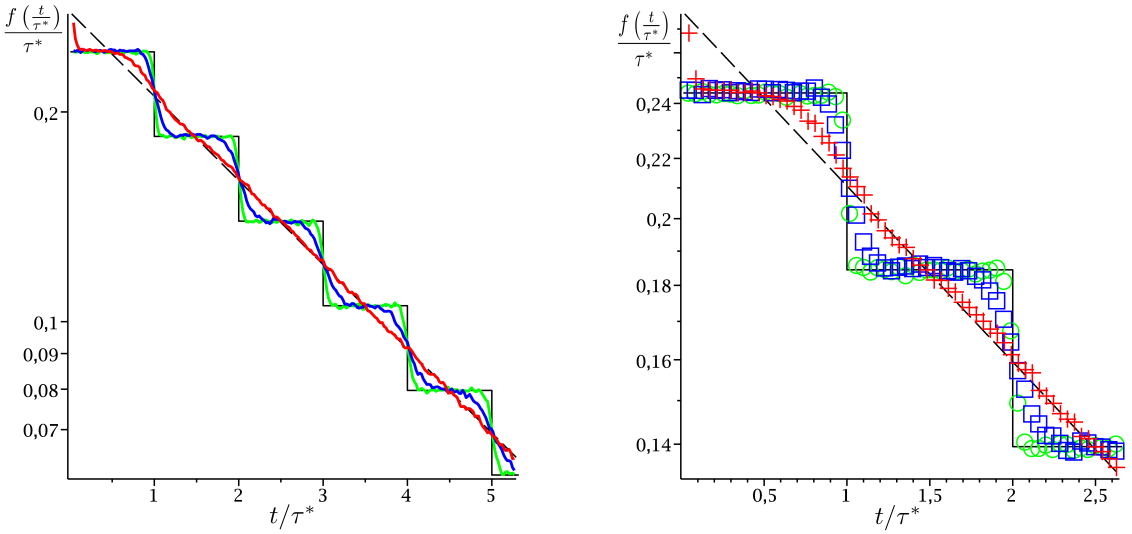


Figure 89: Distribution of the first passage time to the target in the correlated one dimensional case with regularly spaced targets. $f(t/\tau^*)/\tau^*$ as a function of t/τ^*. τ^* defined in (299). Theoretical prediction (300) with fitted $p = 0.244$ (black line). Simulations (colors, symbols) with $L = 2.10^5$ (red, +), $L = 2.10^6$ (blue, □), $L = 2.10^7$ (green, ∘). Exponential distribution (dashed black line). $\tau_1 = 10^6$, $\tau_2 = 7100$, $D = 1$, $V = 1$.

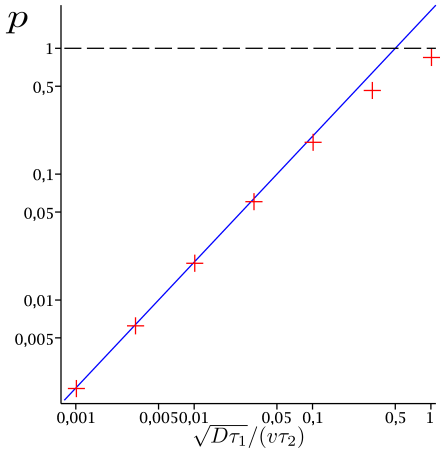


Figure 90: Probability p of finding the first target in the correlated one dimensional case with regularly spaced targets. p as a function of $\sqrt{D\tau_1}/(V\tau_2)$. Simulations (+), theoretical prediction $p = 2\sqrt{D\tau_1}/(V\tau_2)$ (297) (blue line), upper limit $p = 1$ (horizontal dotted black line). $L = 10^5$, $\tau_2 = 10^3$, $D = 1, V = 1$.

$$\begin{aligned}
\frac{\langle t^m \rangle}{(\tau^*)^m} &= \frac{p}{m+1}\left(\sum_{n=0}^{\infty}(1-p)^n(n+1)^{m+1} - \sum_{n=0}^{\infty}(1-p)^n n^{m+1}\right) \\
&= \frac{p}{m+1}\left(\sum_{n=0}^{\infty}(1-p)^n(n+1)^{m+1} - (1-p)\sum_{n=0}^{\infty}(1-p)^n(n+1)^{m+1}\right) \qquad (302) \\
&= \frac{p^2}{m+1}\sum_{n=0}^{\infty}(1-p)^n(n+1)^{m+1}.
\end{aligned}$$

In particular, the first two moments write :

$$\begin{aligned}\frac{\langle t\rangle}{\tau^*} &= \frac{p^2}{2}\sum_{n=0}^{\infty}(1-p)^n(n+1)^2 = -\frac{p^2}{2}\frac{\partial}{\partial p}\sum_{n=0}^{\infty}(n+1)(1-p)^{n+1} \\ &= -\frac{p^2}{2}\frac{\partial}{\partial p}(1-p)\sum_{n=0}^{\infty}(n+1)(1-p)^n = \frac{p^2}{2}\frac{\partial}{\partial p}(1-p)\frac{\partial}{\partial p}\sum_{n=0}^{\infty}(1-p)^{n+1} \\ &= \frac{p^2}{2}\frac{\partial}{\partial p}(1-p)\frac{\partial}{\partial p}(1-p)\sum_{n=0}^{\infty}(1-p)^n = \frac{p^2}{2}\frac{\partial}{\partial p}(1-p)\frac{\partial}{\partial p}(1-p)\frac{1}{p} \\ &= -\frac{p^2}{2}\frac{\partial}{\partial p}\frac{1-p}{p^2} = \frac{p^2}{2}\frac{2-p}{p^3} \\ &= \frac{2-p}{2p},\end{aligned} \tag{303}$$

$$\frac{\langle t^2\rangle}{(\tau^*)^2} = \frac{p^2}{3}(-1)\frac{\partial}{\partial p}(1-p)\frac{2-p}{p^3} = \frac{p^2}{3}\frac{6-6p+p^2}{p^4} = \frac{6-6p+p^2}{3p^2}. \tag{304}$$

More generally :

$$\frac{\langle t^m\rangle}{(\tau^*)^m} = \frac{p^2}{m+1}\left[\frac{\partial}{\partial p}(p-1)\right]^{m+1}\frac{1}{p} = \frac{p^2}{m+1}\frac{\partial}{\partial p}(p-1)\frac{m}{p^2}\frac{\langle t^{m-1}\rangle}{(\tau^*)^{m-1}}. \tag{305}$$

Comparison with previous result

$\langle t\rangle$ has already be obtained with a different method (equation 32). In the limit $L \gg V\tau_2, \sqrt{D\tau_1}, D\tau_1/(V\tau_2)$, this expression has been simplified (equation 36) :

$$\langle t\rangle = \frac{L}{V}\frac{V^2}{D}(\tau_1+\tau_2)\frac{\frac{D^2}{V^4\tau_2^2}+2\frac{D}{V^2\tau_1}}{\sqrt{\frac{D^2}{V^4\tau_2^2}+4\frac{D}{V^2\tau_1}}}. \tag{306}$$

In the limit $\sqrt{D\tau_1} \gg V\tau_2$ (little or no overlap between successive diffusive phases), it can be further simplified :

$$\langle t\rangle = L\frac{V}{D}(\tau_1+\tau_2)\frac{2\frac{D}{V^2\tau_1}}{\sqrt{4\frac{D}{V^2\tau_1}}} = \frac{L}{\sqrt{D\tau_1}}(\tau_1+\tau_2). \tag{307}$$

Now we compare with the first moment of the distribution (equation 303) :

$$\langle t\rangle = \tau^*\frac{2-p}{2p} = \frac{2L}{V\tau_2}(\tau_1+\tau_2)\frac{2-p}{2p}. \tag{308}$$

If $\sqrt{D\tau_1} \gg V\tau_2$, $1 \gg p \simeq 2\sqrt{D\tau_1}/(V\tau_2)$ (equation 297), hence :

$$\langle t\rangle \simeq \frac{2L}{pV\tau_2}(\tau_1+\tau_2) \simeq \frac{L}{\sqrt{D\tau_1}}(\tau_1+\tau_2). \tag{309}$$

The different methods lead to the same result.

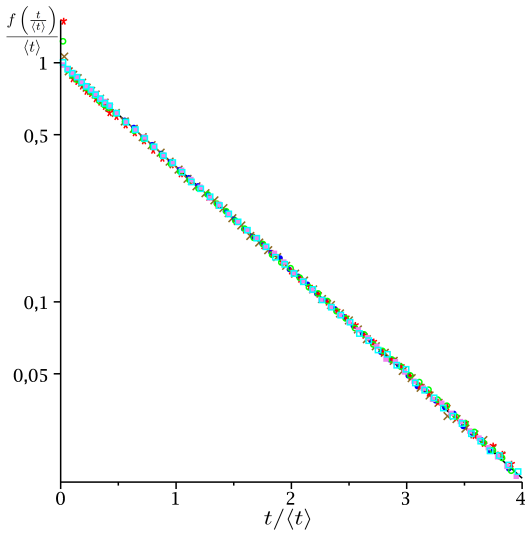


Figure 91: $f(t/\langle t\rangle)/\langle t\rangle)$ as a function of $t/\langle t\rangle$ in the correlated one dimensional case, with Poissonian targets. Expected exponential (black line), simulations (symbols). $L = 100$: $\tau_1 = 1$, $\tau_2 = 1$ ($\star$); $\tau_1 = 10^{-2}$, $\tau_2 = 0.05$ ($\times$). $L = 10^4$: $\tau_1 = 10^4$, $\tau_2 = 10^3$ ($\circ$); $\tau_1 = 10^2$, $\tau_2 = 32$ ($\square$); $\tau_1 = 1$, $\tau_2 = 1$ ($\bullet$); $\tau_1 = 10^{-2}$, $\tau_2 = 0.05$ ($\blacksquare$). $D = 1$, $V = 1$.

6.1.2 Randomly located targets in one dimension with infinite correlation

In the case of regularly spaced targets, we have obtained a "stair" distribution for the first passage time to the target. It is very specific of regularly spaced targets in one dimension with infinite correlation in the directions of the ballistic phases. As expected, taking the same model but with a Poissonian distribution of targets averages stairs, resulting in an exponential distribution of the first passage time to the target (see figure 91).

6.1.3 Regularly spaced targets in two dimensions without correlation

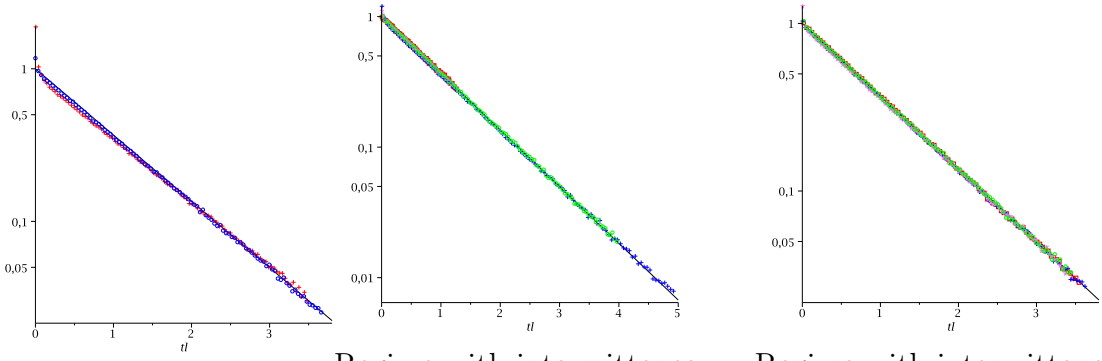
In the case of regularly spaced targets for the diffusive mode in two dimensions without correlation, the distribution is exponential, except for very small t (see figure 92).

6.1.4 Regularly spaced targets in three dimensions without correlation

For regularly spaced targets for the diffusive mode in three dimensions without correlation, the distribution of the detection time is exponential, except for very small t ($t \ll \langle t\rangle$) (see figure 93).

6.1.5 Conclusion on the distribution of the first passage time to the target

In the very specific case of regularly spaced targets in one dimension with infinite correlations of the direction of the ballistic phase, the first passage time to the tar-

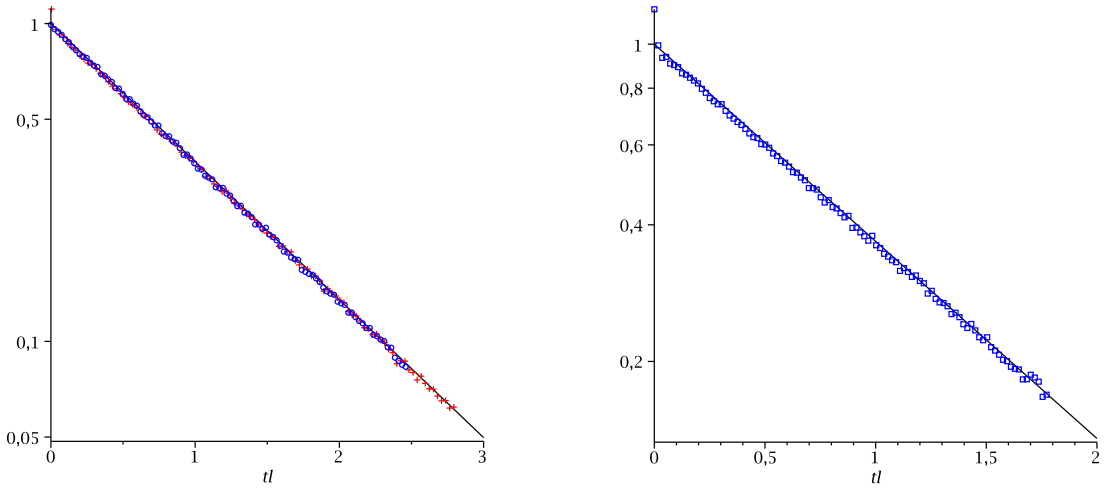


Diffusive regime.
$b = 0.1$.
$\tau_1 \to \infty$, $\tau_2 = 0$.
$a = 0.01$ (+),
$a = 0.001$ (∘).

Regime with intermittence
$b \gg D/v \gg a$.
$a = 0.1$, $b = 10$ (+);
$a = 0.01$, $b = 10$ (⋆);
$a = 0.1$, $b = 100$ (∘);
$a = 0.01$, $b = 100$ (□);

Regime with intermittence
$b \gg a \gg D/v$.
$a = 10$, $b = 1000$ (+);
$a = 100$, $b = 1000$ (×);
$a = 100$, $b = 10000$ (∘);
$a = 100$, $b = 100000$ (□).

Figure 92: Regularly spaced targets for the 2-dimensional case. $f(t/\langle t\rangle)/\langle t\rangle)$ as a function of $t/\langle t\rangle$. Simulations (symbols), exponential distribution (black line). $D = 1$, $V = 1$. $\tau_1 \simeq \tau_1^{opt}$, $\tau_2 \simeq \tau_2^{opt}$.



Diffusive regime. $a = 0.1$
$b = 1$ (+), $b = 10$ (∘).

Regime with intermittence.
$a = 10$, $b = 100$.

Figure 93: Regularly spaced targets for the 3-dimensional case. $f(t/\langle t\rangle)/\langle t\rangle)$ as a function of $t/\langle t\rangle$. simulations (symbols), exponential distribution (black line). $D = 1$, $V = 1$. $\tau_1 \simeq \tau_1^{opt}$, $\tau_2 \simeq \tau_2^{opt}$.

get is stair-distributed. But as soon as one constraint is relaxed (for example, with targets randomly located, or with no correlations between ballistic phases), the distribution tends to an exponential (except for the very beginning of the distribution, for $t \ll \langle t\rangle$). It means that the mean search time is the only time scale of the distribution. Our choice of studying the mean search time is consequently justified.

6.2 Taking into account partial correlations in ballistic phases

6.2.1 Motivation

In section 3.2 and section 5, models are very similar. The searcher alternates between a slow reactive phase, and a fast ballistic blind phase. The difference is that in section 3.2, ballistic phases are always in the same direction, whereas at the contrary, in section 5, for each new ballistic phase, direction is taken at random, without any correlation with the previous ballistic phase.

In the case of animals, there is usually correlation in the angles between two successive ballistic phases [O'Brien et al., 1990]. We have studied the two extremes : no correlation or perfect correlation. In both cases, there are regimes of favorable intermittence. However, in the case of infinite correlation, there is a relation between the mean duration of the two phases minimizing the search time, but the shorter the durations are, the smaller the search time. In contrast, in the case without any correlation, the minimal search time is obtained for finite τ_1 and τ_2. In the intermediate case of finite range correlations, determining the nature of the minimum is an interesting theoretical question. Besides, as real biological systems often present correlation, it is an important issue to add correlations to our generic model.

We present in the following preliminary results on correlations, obtained in the simplest case of the static mode in one dimension.

6.2.2 Model

The model studied is the static mode in one dimension. The searcher is either in the reactive phase 1 (no motion, reaction rate k if the target is at a distance smaller than a), or in the ballistic phase 2, of velocity V, in the same direction as in the previous ballistic phase with probability p, and in the opposite direction with probability $1-p$. Distribution of the duration of the phases is exponential, of mean τ_i. The distance between two targets is $2b$.

As calculated in the section 5.3.1, when $p = 0.5$, the mean search time is (150) :

$$t_m = \frac{\tau_1+\tau_2}{b}\left(\frac{b}{k\tau_1}+\frac{(b-a)^3}{3V^2\tau_2^2}+\frac{\beta(b-a)^2}{V\tau_2}\coth\left(\frac{a}{V\tau_2\beta}\right)+(b-a)\beta^2\right), \qquad (310)$$

where $\beta = \sqrt{(k\tau_1)^{-1}+1}$. At the optimum, $\tau_1^{opt} = \sqrt{\frac{a}{Vk}}\left(\frac{b}{12a}\right)^{1/4}$ (154) and $\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{3a}}$ (155).

We first calculate and optimize the mean search time when $p = 1$, then we calculate the mean search time for arbitrary p. As it is not easy to minimize the resulting expression, we give an argument for the form of the minimum, and finally we conclude.

6.2.3 Infinite range correlations

Mean search time

t_i is the mean search time starting from x at $t = 0$ in state i. The target is in $[-a, a]$, and we study the segment $[-a, 2b-a]$. Inside the target ($x \in [-a, a]$), the

equations for t_i are :

$$\lambda_1 \left(t_2^{int} - t_1^{int}\right) - k t_1^{int} = -1, \tag{311}$$

$$V \frac{dt_2^{int}}{dx} + \lambda_2 \left(t_1^{int} - t_2^{int}\right) = -1. \tag{312}$$

Outside the target ($x \in [a, 2b-a]$) :

$$\lambda_1 \left(t_2^{ext} - t_1^{ext}\right) = -1, \tag{313}$$

$$V \frac{dt_2^{ext}}{dx} + \lambda_2 \left(t_1^{ext} - t_2^{ext}\right) = -1. \tag{314}$$

t_2 is continuous, and the boundaries are circular. It leads to :

$$t_2^{int}(a) = t_2^{ext}(a), \tag{315}$$

$$t_2^{int}(-a) = t_2^{ext}(2b-a). \tag{316}$$

With this set of equations, we solve the system. The mean search time starting from a random position in state 1 is :

$$t_m = \frac{2b-a}{bk}\left(\frac{\tau_2}{\tau_1}+1\right) + \frac{b-a}{b}\left(\tau_1+\tau_2+\frac{b-a}{V}\left(\frac{\tau_1}{\tau_2}+1\right)\coth\left(\frac{ak\tau_1}{V\tau_2(1+k\tau_1)}\right)\right). \tag{317}$$

The next step is the minimization of the mean search time.

Minimization

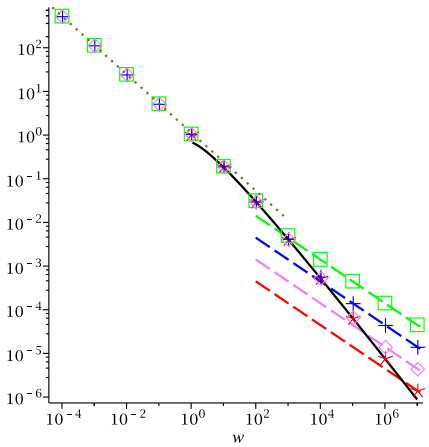


Figure 94: Minimization of the mean search time for the static mode with infinite correlation ($p = 1$). α^{opt} as a function of w. Theoretical expression for w small (492) (brown dots), for w intermediate (495) (black solid line), for w large (497) (colored dashed lines). Optimization of the exact expression (with $\tau_2 \to 0$ and $\tau_1 = \alpha\tau_2$) (symbols). $b = 100$ (green, $\square$), $b = 10^3$ (blue, $+$), $b = 10^4$ (violet, $\diamond$), $b = 10^5$ (red, $\star$). $a = 1$, $V = 1$.

As detailed in appendix 8.3.3, the mean search time (317) is minimized for τ_1 and τ_2 tending to 0, with $\tau_1 = \alpha\tau_2$. We define $w = ak/v$, and depending on this parameter :

- $w < 1 : \alpha^{opt} \simeq \left(\frac{3}{2w^2}\right)^{\frac{1}{3}}$,
- $1 < w < w^* : \alpha^{opt} \simeq \frac{\ln(4w)}{2w}$,
- $w > w^* : \alpha^{opt} \simeq \sqrt{\frac{2a}{wb}}$.

w^* is defined as the solution of $\frac{\ln(4w^*)}{2w^*} = \sqrt{\frac{2a}{w^*b}}$. These expressions are in good agreement with the numerical minimization of the entire exact expression of the mean search time (see figure 94).

6.2.4 Intermediate correlation

We have obtained the mean search time and optimized it for the case without correlation ($p = 0.5$), and for the case of infinite correlation ($p = 1$). Now, let us study the case with an arbitrary correlation p.

Equations

In what follows, + denotes a phase 2 in the direction $+x$, or a phase 1 coming after a phase 2 in the direction $+x$, and equivalently for $-$. i means "inside the target".

Inside the target :

$$V\frac{dt_2^{+,i}}{dx} + \lambda_2\left(t_1^{+,i} - t_2^{+,i}\right) = -1, \tag{318}$$

$$-V\frac{dt_2^{-,i}}{dx} + \lambda_2\left(t_1^{-,i} - t_2^{-,i}\right) = -1, \tag{319}$$

$$\lambda_1\left(pt_2^{+,i} + (1-p)t_2^{-,i} - t_1^{+,i}\right) - kt_1^{+,i} = -1, \tag{320}$$

$$\lambda_1\left(pt_2^{-,i} + (1-p)t_2^{+,i} - t_1^{-,i}\right) - kt_1^{-,i} = -1. \tag{321}$$

Outside the target :

$$V\frac{dt_2^{+}}{dx} + \lambda_2\left(t_1^{+} - t_2^{+}\right) = -1, \tag{322}$$

$$-V\frac{dt_2^{-}}{dx} + \lambda_2\left(t_1^{-} - t_2^{-}\right) = -1, \tag{323}$$

$$\lambda_1\left(pt_2^{+} + (1-p)t_2^{-} - t_1^{+}\right) = -1, \tag{324}$$

$$\lambda_1\left(pt_2^{-} + (1-p)t_2^{+} - t_1^{-}\right) = -1. \tag{325}$$

Resolution

These equations can be written differently, with $t_j = (t_j^+ + t_j^-)/2$ and $t_j^d = (t_j^+ - t_j^-)/2$. Inside the target :

$$V\frac{dt_2^{d,i}}{dx} + \lambda_2\left(t_1^{i} - t_2^{i}\right) = -1, \tag{326}$$

$$V\frac{dt_2^{i}}{dx} + \lambda_2\left(t_1^{d,i} - t_2^{d,i}\right) = 0, \tag{327}$$

$$\lambda_1 \left(t_2^i - t_1^i\right) - k t_1^i = -1, \tag{328}$$

$$\lambda_1 \left((2p-1) t_2^{d,i} - t_1^{d,i}\right) - k t_1^{d,i} = 0. \tag{329}$$

Outside the target :

$$V \frac{dt_2^d}{dx} + \lambda_2 \left(t_1 - t_2\right) = -1, \tag{330}$$

$$V \frac{dt_2}{dx} + \lambda_2 \left(t_1^d - t_2^d\right) = 0, \tag{331}$$

$$\lambda_1 \left(t_2 - t_1\right) = -1, \tag{332}$$

$$\lambda_1 \left((2p-1) t_2^d - t_1^d\right) = 0. \tag{333}$$

We have symmetry in $x = 0$ (the center of the target) and in $x = b$ (where the distances from the target on the left and on the right are equal), consequently in these points $t_2^+ = t_2^-$. It leads to $t_2^{d,i}(x = 0) = 0$ and $t_2^d(x = b) = 0$. t_2^+ and t_2^- are continuous in a (at the target boundary). Thus $t_2^i(x = a) = t_2(x = a)$ and $t_2^{d,i}(x = a) = t_2^d(x = a)$.

Mean search time

We solve using these boundary conditions. We are interested in the mean first passage time to the target, starting from phase 1 in a random direction, *i.e.* $t_m = \frac{1}{b}\left(\int_0^a t_1^i(x)dx + \int_a^b t_1(x)dx\right)$. We obtain :

$$t_m = (\tau_2 + \tau_1)\left(\frac{1}{k\tau_1} + \frac{b-a}{b}\left(\frac{2}{3}\frac{(1-p)\,(b-a)^2}{{\tau_2}^2 V^2} + 1 + \frac{1}{k\tau_1} + \frac{u\,(b-a)}{\sqrt{k\tau_1}\tau_2 V}\coth\left(\frac{\sqrt{k\tau_1}ua}{\tau_2\,V\,(1+k\,\tau_1)}\right)\right)\right), \tag{334}$$

with $u = \sqrt{2(1-p) + k\tau_1}$.

6.2.5 An argument on the form of the minimum

The mean search time obtained (334) is hard to optimize. A question raised is if it is minimized for finite τ_1 and τ_2 (like the case $p = 0.5$), or for τ_1 and τ_2 tending to 0 (like the case $p = 1$).

A lower bound can be set for the mean search time :

$$t_m \geq (\tau_1 + \tau_2)\left(\frac{1}{k\tau_1} + \frac{(b-a)^3}{b}\frac{2(1-p)}{3\tau_2^2 V^2}\right). \tag{335}$$

Supposing the minimum is for at least one of the $\tau_i \to 0$, 3 cases arise.

- $\tau_1 \to 0$ with $\tau_1 \ll \tau_2$. In this case $t_m \geq \tau_2/(k\tau_1) \to \infty$.
- $\tau_2 \to 0$ with $\tau_2 \ll \tau_1$. In this case $t_m \geq \tau_1 \frac{(b-a)^3}{b}\frac{2(1-p)}{3\tau_2^2 V^2} \to \infty$.
- $\tau_1 \sim \tau_2$ and both $\to 0$. In this case $t_m \geq \tau_2 \frac{(b-a)^3}{b}\frac{2(1-p)}{3\tau_2^2 V^2} \sim 1/\tau_2 \to \infty$.

The conclusion is that the minimum is for finite values of τ_1 and τ_2 as soon as $p < 1$. However, these values could depend on p, and be small when p is close to 1.

6.2.6 Optimization with low correlations

We first study the case with low correlations, with $p > 0.5$, but not too close to $p = 1$.

We assume $b \gg a$. It can be noticed that as $p \geq 0.5$, $\sqrt{k\tau_1}\sqrt{2(1-p)+k\tau_1}/(1+k\tau_1) \leq \sqrt{k\tau_1}\sqrt{1+k\tau_1}/(1+k\tau_1) \leq 1$. We assume that the optimal τ_2 value will not be too different from the value for $p = 0.5$, leading to $\tau_2 V \gg a$. With this supposition, $\coth\left(\sqrt{k\tau_1}ua/\left(\tau_2 V\left(1+k\tau_1\right)\right)\right)$ can be developed. With again $b \gg a$, and with the additional hypothesis $2(1-p)b^2/(3\tau_2^2V^2) \gg 1$, the expression of the mean search time can be greatly simplified :

$$t_m \simeq (\tau_2+\tau_1)\left(\frac{2}{3}\frac{(1-p)\,b^2}{{\tau_2}^2V^2}+\frac{b\,(1+k\tau_1)}{k\tau_1\,a}\right). \tag{336}$$

The minimum of the search time is obtained for :

$$\tau_1^{opt} = \sqrt{\frac{a}{Vk}}\left(\frac{b}{a}\frac{(1-p)}{6}\right)^{1/4} \text{ and} \tag{337}$$

$$\tau_2^{opt} = \frac{a}{V}\sqrt{\frac{b}{a}\frac{2(1-p)}{3}}. \tag{338}$$

Interestingly, note that the relation $\tau_2^{opt} = 2k\left(\tau_1^{opt}\right)^2$ obtained initially in the case of the absence of correlations Table 3 still holds in this case.

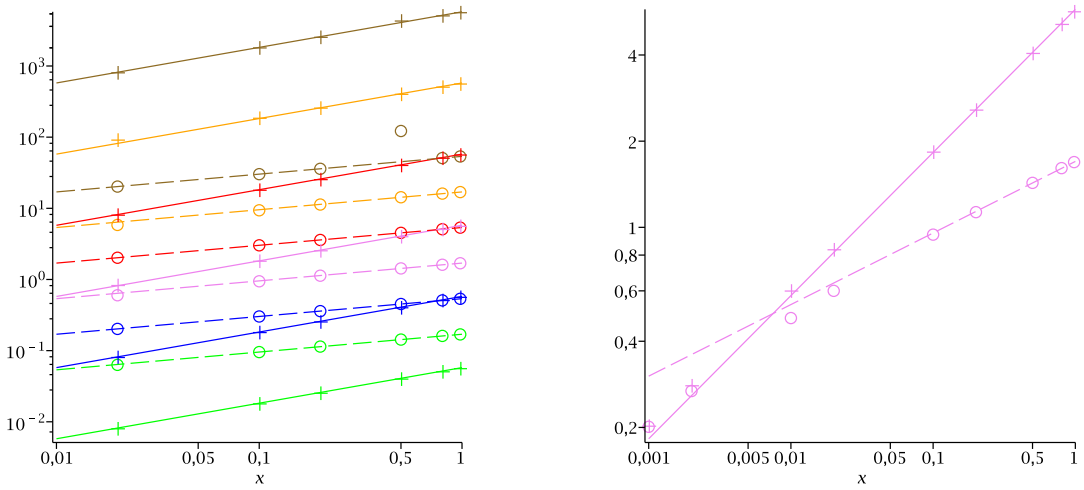


Figure 95: τ^{opt} as a function of $2(1-p)$. Theoretical value of τ_1^{opt} (337) (dashed line) and theoretical value of τ_2^{opt} (338) (solid lines), compared to the numerical minimization of the full exact mean search time, leading to τ_1^{opt} ($\circ$) and τ_2^{opt} ($\square$). $a = 0.01$, $b = 1$ (green), $a = 0.01$, $b = 100$ (blue), $a = 1$, $b = 100$ (violet), $a = 1$, $b = 10^4$ (red), $a = 100$, $b = 10^4$ (orange), $a = 100$, $b = 10^6$ (brown). $k = 1$, $V = 1$.

These expressions are in agreement with the numerical minimization of the exact expression of the mean search time (see figure 95), except when $1-p$ is very small. We assumed $V\tau_2 \gg a$. When correlations increase, τ_2 decreases, and when $(1-p)$ is of the same order of magnitude or smaller than a/b, this approximation breaks down.

6.2.7 Conclusion

In the simple case of the static mode in one dimension, the influence of correlations on the mean search time and its minimization can be studied. We have obtained the exact expression of the mean search time. Its minimum is for finite values of τ_1 and τ_2 as soon as $p < 1$. When $(1-p) \gg a/b$, the optimal durations τ_1^{opt} and τ_2^{opt} can be explicitly given, and they are in continuity with the case without correlation $p = 0.5$.

In the future, a first step would be to give τ_1^{opt} and τ_2^{opt} for $p \to 1$, and then to study other cases : other modes, and other dimensions. However, it is probable that solving analytically in other cases is very hard, and maybe answers can only be found numerically.

6.3 Other distributions of phases duration

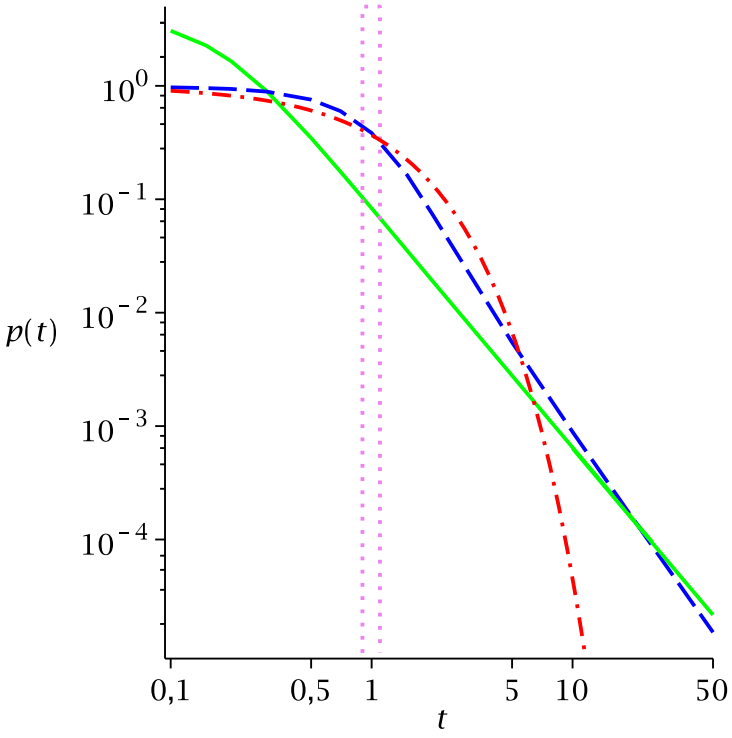


Figure 96: Different distributions for the duration of the ballistic non-reactive phase. Red dashed-dotted line : exponential distribution, used in our generic model (see section 5); violet dots : probability peaked around the mean, a form of temporal memory (see section 6.3.1); blue dashed line ($\alpha = 1.5$) and green line ($\alpha = 1.1$) : Lévy distribution (see section 6.3.2); All these distributions have a mean of 1.

The model presented in section 5 is minimal in the sense that the searcher has no memory. As seen in the previous section, a possibility is to add some spatial memory. Another possibility is to add temporal memory. In the generic model, we have assumed a "Markovian" searcher, in the sense that the rate of switching from one

phase to the other is constant. It leads to exponential distribution of the durations of the phases. In this section, we study the influence of the distribution of the duration of the phases. On the one hand, we have studied the effect of distributions peaked around the the mean duration [Bénichou et al., 2007]. On the other hand, Lomholt et al. [2008] have studied the case of Lévy-distributed "blind" phases. We present their main results and show that this strategy can be outperformed.

6.3.1 Fixed durations of the phases

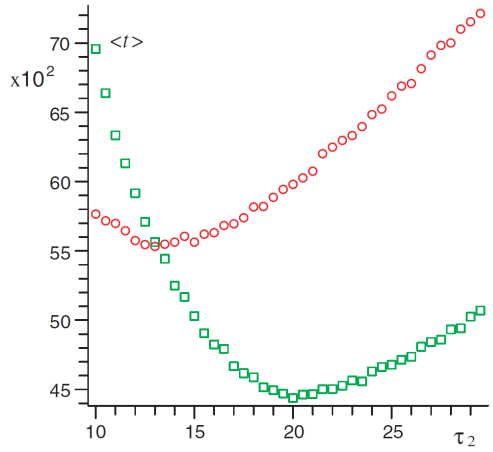


Figure 97: Comparison between the search without temporal memory (○) and the search with temporal memory (□). Static mode in two dimensions. t_m as a function of τ_2. $k = 1$, $V = 1$, $b = 113$, $a = 10$, $\tau_1 = 2.6$.

In the generic model previously described, we have considered exponential durations of phases. A minimal form of memory could be to switch from one phase to another at deterministic times instead of exponentially-distributed times. As calculations are more complex than for an exponential, we only propose a hint in one case. We study numerically [Bénichou et al., 2007] the effect of such temporal memory for the static mode in 2 dimensions. The optimal τ_1^{opt} and τ_2^{opt} are superior to their values without memory, but are of the same order of magnitude (see figure 97). Such temporal memory decreases the mean search time. Indeed, there is neither time lost due to very short relocation which are inefficient, nor time lost due to too large excursions that overshoot the target (see figure 96). However, the gain from this temporal memory is quite low (less than 40% in an extended range of parameters, and decreasing with b/a increasing).

6.3.2 Lévy distribution of the fast phase durations

Reynolds [2006] proposes to combine intermittence with Lévy walks. In one dimension, steps are taken from a Lévy law ($p(l) \propto l^{-\alpha-1}$), but steps smaller than a cut-off l_0 are assumed to allow the searcher to detect targets, whereas steps larger than l_0 are assumed to be "blind". Reynolds' numerical study is very artificial in the sense that there is an unjustified upper cut-off l^*_{max}. Viswanathan et al. [1999]

also introduce an upper cut-off, but it is a natural outcome of the details of their model. Since in their study the searcher can find its target all along the trajectory, if the typical distance between targets is L, then the maximum flight length will be of order L. In Reynolds [2006], there is no natural justification for an upper cut-off. It dramatically changes the validity of his results. In particular, for $0 < \alpha < 1$, the mean is not defined for a real Lévy flight, and it would result in an efficiency tending to 0, while it is in this regime that Reynolds [2006] finds the optimum for non-revisitable targets. The effect of the cut-off for $1 < \alpha < 2$ is less dramatic, but probably still important.

Lomholt et al. [2008] study analytically and through simulations a combination of intermittence with real Lévy walks. Their model is similar to our generic model, more precisely to the diffusive mode in one dimension (with $a \to 0$), except that the ballistic phase duration is taken from a Lévy law ($p(l) \propto l^{-\alpha-1}$, with $1 < \alpha < 2$, see figure 96 : more exactly, the characteristic function of the distribution is $\phi(l) = \exp(-(\sigma l)^\alpha)$, σ being defined as the scaling factor of the distribution). L is the distance between two successive targets.

The relation between σ, α, the velocity V and τ_2 (the mean duration of phase 2, which is defined since $\alpha > 1$) is (equation 10 of Lomholt et al. [2008]) :

$$\sigma = \frac{\pi V \tau_2}{2\Gamma\left(1 - 1/\alpha\right)} \tag{339}$$

The mean search time is evaluated with the exact formula (equation 9 of Lomholt et al. [2008]) :

$$\langle t \rangle = \sum_{n=1}^{\infty} \frac{2(\tau_1 + \tau_2)}{D\tau_1 k_n^2 + 1 - \lambda(k_n)} \tag{340}$$

where $k_n = 2\pi/L$, and λ is the characteristic function of the distribution : $\lambda(k) = \exp(-\sigma^\alpha |k|^\alpha)$ ($p(k) = \int_{-\infty}^{\infty} e^{ikx} p(x) dx$). This infinite sum can be approximated with the more tractable expression (equation 14 of Lomholt et al. [2008]) :

$$\langle t \rangle = 2(\tau_1 + \tau_2) \left(\frac{L}{4\sqrt{D\tau_1}} + \left(\frac{L}{2\pi\sigma} \right)^\alpha \zeta(\alpha) \right) \tag{341}$$

with $\zeta(\alpha) = \sum_{n=1}^{\infty} n^{-\alpha}$ is the Riemann ζ function.

Lomholt et al. argue that as diffusion is recurrent in one dimension, large relocations are favorable, in particular for large distance L between targets, and that Lévy walks, thanks to their infinite variance, are good candidates for optimizing the search. They show that when L is large, Lévy-distributed relocations decrease the mean search time more than exponentially distributed relocations. They find that when $L \to \infty$, the search time is minimized for $\alpha \to 1^+$.

However, relocations larger than L cannot be profitable, as L is the distance between two targets. And the closer to 1 is α, the more often these time-wasting relocations occur. A simple idea is that a Lévy distribution with a cut-off at L should be better. We performed simulations (see appendix 8.3.2 for technical details). For $L = 10^4$ (see figure 98), the optimum without cut-off is for approximately $\alpha \simeq 1.4$, with $t_m \simeq 195\ 000$; and the optimum with a cut-off of L is for $\alpha \simeq 1.3$, with $t_m \simeq 188\ 000$, that is $\simeq 3.7\%$ lower. For $L = 10^5$ (see figure 99), the optimum

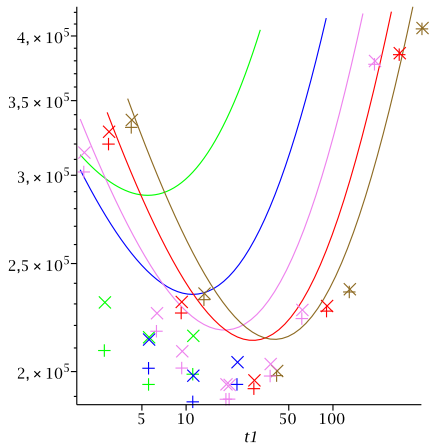


Figure 98: t_m as a function τ_1, σ being at the theoretical minimum (numerical minimization of equation 341). lines : approached analytical formula 341. $\times$: simulations without cut-off. $+$: simulations with cut-off at L. sienna : $\alpha = 1.6$; red : $\alpha = 1.5$; violet : $\alpha = 1.4$; blue : $\alpha = 1.3$; green : $\alpha = 1.2$. $L = 10^4$. $D = 1$, $V = 1$.

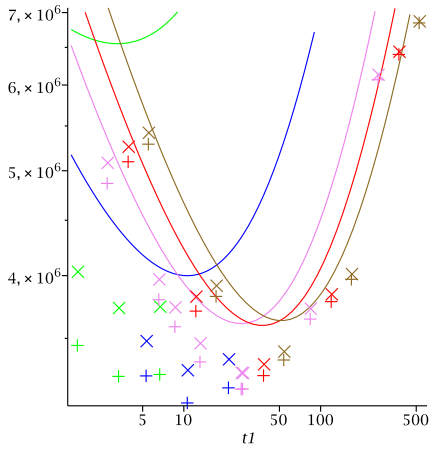


Figure 99: t_m as a function τ_1, σ being at the theoretical minimum (numerical minimization of equation 341). lines : approached analytical formula 341. $\times$: simulations without cut-off. $+$: simulations with cut-off at L. sienna : $\alpha = 1.4$; red : $\alpha = 1.35$; violet : $\alpha = 1.3$; blue : $\alpha = 1.2$; green : $\alpha = 1.1$. $L = 10^5$. $D = 1$, $V = 1$.

without cut-off is for $\alpha \simeq 1.3$, $t_m \simeq 3\ 260\ 000$; the optimum with a cut-off of L is for $\alpha \simeq 1.2$, $t_m \simeq 3\ 060\ 000$, that is $\simeq 6.5\%$ lower. As explained in appendix 8.3.2, even if there can be some imprecisions, they are not large enough to change the main conclusions : as expected, using a cut-off decreases the mean search time; and the gain is larger when L is increasing.

The Lévy-distributed relocations decrease the mean search time more efficiently than exponentially distributed relocations, but it is not because of their infinite variance. Indeed, Lévy-distributed relocations with a cut-off (no relocation larger than the distance between targets) have a finite variance and outperform pure Lévy-distributed relocations.

6.4 Other models of intermittent search

In the previous section, we have seen that the Lévy walks combined to intermittence have been studied by Lomholt et al. [2008]. Our work on intermittence has raised interest in the physics community. We review in the following other models on intermittence, departing more from the models we have proposed, but still relying on the idea of intermittence.

Oshanin et al. [2007]

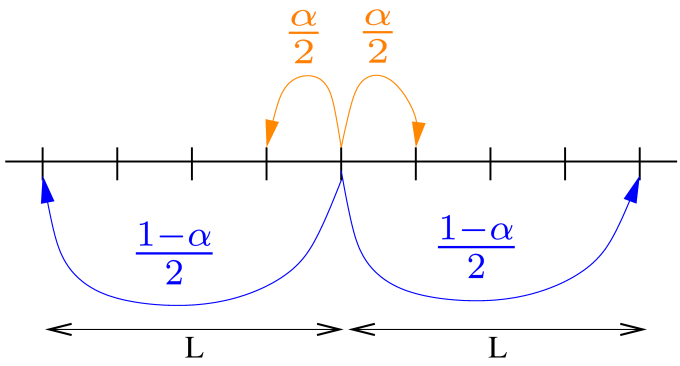


Figure 100: Model used by Oshanin et al. [2007].

Oshanin et al. [2007] propose a model very similar to the diffusive mode in one dimension of our generic model, but in discrete space, on an infinite lattice. At each time step, with probability α, the searcher jumps to the neighboring node of the line (with equal probabilities for each side, which corresponds to diffusion). With probability $1-\alpha$, it stays off-lattice during a time T and after this time, it lands at a distance L from its initial position (once again, with equal probabilities for each side) (see figure 100). This phase is equivalent to a ballistic non-reactive phase. Its duration is exactly T, whereas the duration of the diffusive phase with target detection is exponentially distributed, with mean duration $1/(1-\alpha)$. There is one target, but an infinite set of searchers, initially randomly distributed. The quantity maximized is the probability that at a given time t, the target has already been found by any of the searchers. Oshanin et al. do find an optimal α, but dependent on t.

Rojo et al. [2009]

Rojo et al. [2009] propose a model which displays some similarities with the previous model (see figure 101). The search domain is also a one-dimensional discrete infinite lattice with one target, there are also an infinite set of searchers, and the quantity optimized is also the probability that the target is found by any of the searchers at a given t. The detection phase consists of jumps to the nearest-neighbors, with a given frequency. Such a rule is equivalent to diffusion. The non-reactive phase consists of jumps to the next nearest neighbors. It is again diffusion, but if the jump frequency is the same as in the other phase, it is a faster

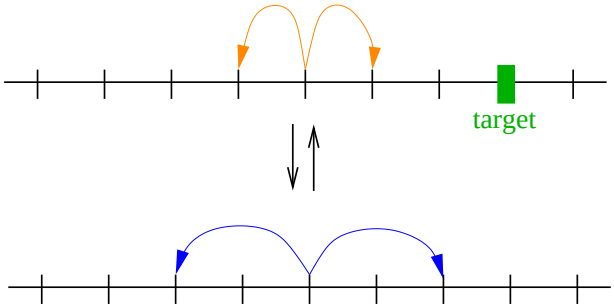


Figure 101: Model used by Rojo et al. [2009].

diffusion. In both phases, there is a fixed rate of switch to the other phase, leading to exponentially distributed durations of the phases. If one of the mean durations is fixed, the probability that the target is already found at t is minimized for a finite duration of the other phase. But the optimum is for infinitely short phases, enabling the searcher to combine the faster diffusion of one phase and the detection capacities of the other phase.

Reingruber and Holcman [2009]

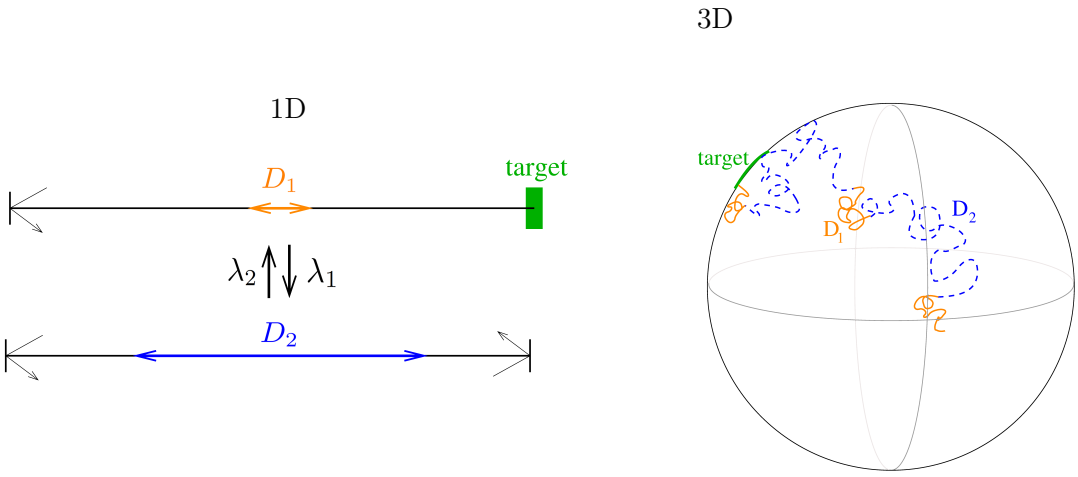


Figure 102: Model used by Reingruber and Holcman [2009].

Reingruber and Holcman [2009] propose a model which is also diffusive/diffusive (see figure 102). They study this model first in one dimension : the searcher's starting point is at one extremity of a segment, a reflecting boundary. The target is at the other end of the segment. However, in phase 1 (diffusion of coefficient D_1), the target can be found, whereas in phase 2 (diffusion of coefficient D_2), both extremities are reflecting. There are fixed rates of switching from one phase to another. The results show that there are two regimes : if $D_1 > D_2$, straightforwardly, the optimum for the searcher is to be in phase 1 only; if $D_2 > D_1$, the optimum is to switch very rapidly between the two phases, such as to spend almost all the time in the faster phase 2, but not to miss the target. This model is extended to a 3-dimensional

ball (the initial position is almost without importance in this geometry), but with a target of radius a on the border (which is reflecting everywhere else). The two phases are defined like in one dimension. Reingruber and Holcman [2009] give two limits in this case. It can be noticed that the expression we have obtained in the generic model for the diffusive mode in 3 dimensions (441) could be used, with $3V^2\tau_2^2 = D_2\tau_2$. In fact, our calculations use a "diffusive/diffusive" approximation, with an effective $D_2^{\text{eff}} = 3V^2\tau_2$. The optimization will be quite different, because the dependence in τ_2 is dramatically changed if instead of a fixed D_2, D_2 is a function of τ_2. Indeed, the optimum for our generic model is for finite τ_1 and τ_2, whereas, even if not explicitly calculated, it is probable that the optimum for diffusion/diffusion in three dimensions is similar to the one-dimensional case, *i.e.* for phases durations as small as possible. Their goal is to model cellular signaling, with a ligand binding to a target which will transmit a signal.

Bressloff and Newby [2009], Newby and Bressloff [2009]

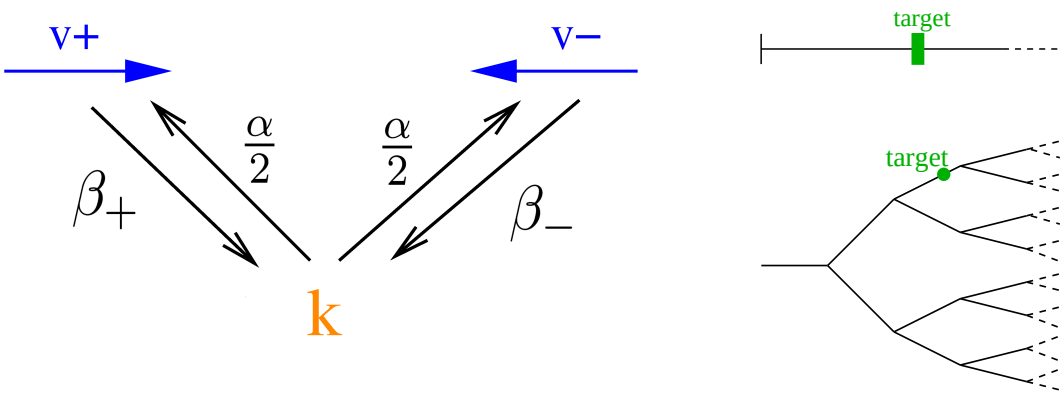


Figure 103: Model used by Bressloff and Newby [2009], Newby and Bressloff [2009].

Bressloff and Newby [2009] present another model applied to intracellular transport, more precisely here to the transport of mRNA granules inside neurons. They present a model in one dimension, standing for example for an axon with little branching. The starting point is at one extremity, which is reflecting : it models granules produced in the soma of the neuron and that have to be exported to the axon. The target, a synapse, is somewhere in the segment. The other end of the segment is an absorbing boundary, representing that the vesicles containing the mRNA can be degraded, or that there can be other targets further away in the axon that can absorb the searcher. To complete this idea that there are several targets that are not equivalent, and that these targets are in competition, they also calculate explicitly the probability that the searcher finds a target more often than the others. In this model, there are 3 states (see figure 103) : an immobile detection phase, similar to the static mode, switching to ballistic modes with probability α per unit time; a ballistic phase in direction $+$, with speed v_+, and with a transition rate to the detection mode β_+; a ballistic phase in direction $-$, with speed v_-, and with a transition rate to the detection mode β_-. During the two ballistic phases, the searcher cannot detect the target. Movement is biased to the direction $+$ if $v_+/\beta_+ > v_-/\beta_-$.

This modeling implies that for each direction, there are different motors, with different velocities and binding properties. It can be for example the case in an axon, where microtubules are known to be aligned in the same direction, and with vesicles binding with different affinities to the motors walking in the two opposite directions. On the contrary, in dendrites for example, microtubules are not all oriented in the same direction. In this case, if the vesicles bind to one type of motors, $\beta_+ = \beta_-$, and $v_+ = v_-$, and if the concentration of microtubules pointing to one or the other direction are not equal, α, the switching rate from detection phase to ballistic phase will be different for the two directions of the ballistic phases. This last modeling may be more realistic in certain cases. However, it is probable that the main results would not change by taking a bias due to the α instead of the β.

The results are based on the fact that on the segment, there are two contradicting constraints : maximizing the hitting probability (as the searcher can be degraded before finding the target), and minimizing the time to find the target when the target is found. Indeed, if there is more bias, the time to find the target will be smaller, but the target will be missed more often. With a fixed hitting probability, the mean first passage time to the target (on the condition that the target is found) is minimized when there is more bias. In other words, unidirectional motion is better than bi-directional motion in this case.

Newby and Bressloff [2009] extend this problem to the case of a directed tree. In this case, unidirectional motion has a drawback : a wrong branch can be taken, annihilating any possibility to find the target. Biased bidirectional motion can be seen as an effective combination of a ballistic and a diffusive motion. It exists a critical hitting probability p^*. If the mean first passage time to the target is minimized given that the probability of finding the target is a given $p < p^*$, unidirectional motion is better; but if the given probability is $p > p^*$, there is an optimal finite bias which minimizes the mean search time in case of success.

Ramezanpour [2007]

Intermittence in networks such as the tree we have just seen is an interesting extension. Ramezanpour [2007] proposes (see figure 104) to explore a network in which the degree (= number of neighbors) distribution is $p(k) \propto k^{-3}$, constructed as proposed by Barabasi and Albert [1999], or with some modifications. On this finite network, at each time step, the searcher chooses randomly one of the edges connected to the node where it is, and goes to the node connected by this edge. Every t_w, the searcher jumps to a completely random node. The question is whether the mean time to cover the nodes and the edges of the network can be optimized as a function of t_w. For the nodes, the random jumping is a way to visit all the nodes with equal probability, thus t_w should be as small as possible. For the edges, there is an optimal finite t_w. Indeed, if t_w is small, most edges visited will emanate from low-connected nodes (as the low connected nodes are the more numerous nodes, such edges are more likely to be visited after a random jump), but if t_w is large, the searcher would spend most of its time on the edges connecting high degree nodes, and will take time to explore the whole network, especially for remote edges connecting nodes of low degree.

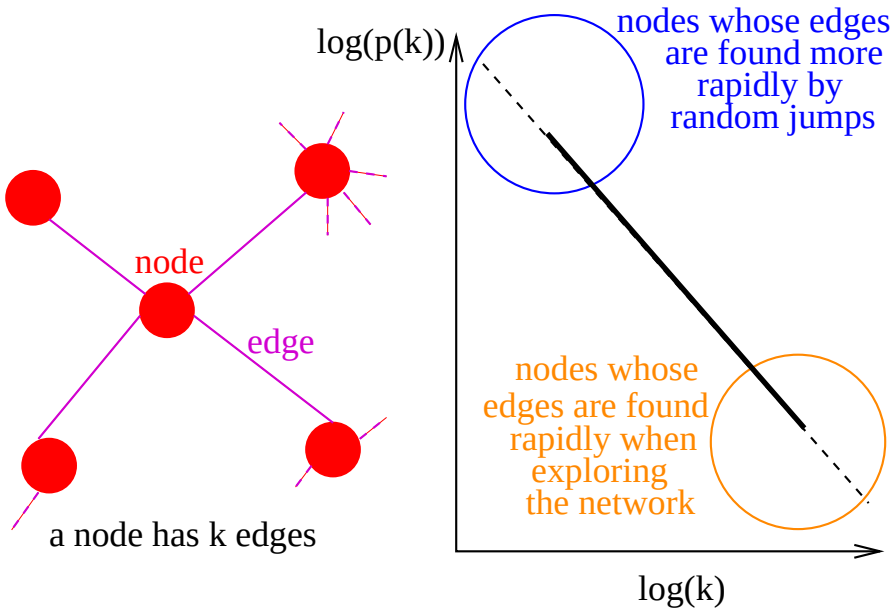


Figure 104: Model presented by Ramezanpour [2007].

Summary of these other models of intermittent search

Finally, the simple idea of intermittence has been developed in various different directions. These models present rich behaviors, and show that, depending on the situation, intermittence can be an optimal search strategy.

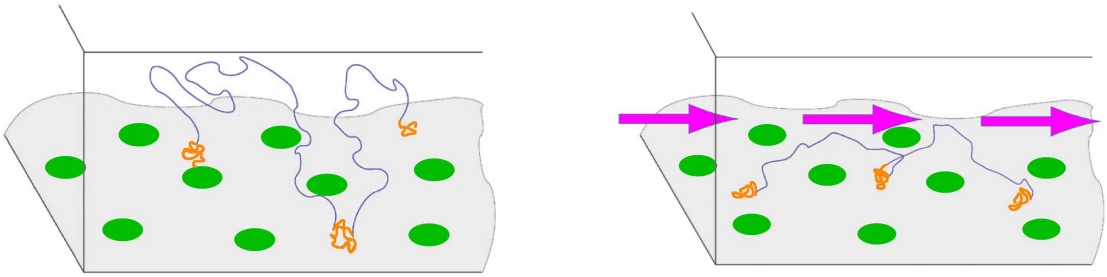
6.5 Designing efficient searches

As seen previously, intermittent reaction paths are involved in various search problems involving biomolecules at the microscopic scale, as well as biological organisms at the macroscopic scale. Simple analytical models show that intermittent transport can actually minimize search time. A reason why such intermittent trajectories are widely observed could be simply that they constitute very generic optimal search strategies, and consequently they could have been selected by evolution.

Beyond *modeling* what is observed in real-life biological examples, such intermittent strategies could also be used to *design* searches, at the microscopic and macroscopic scales.

6.5.1 Microscopic scale

Heterogeneous chemical reactions, with targets fixed at an interface (1D (polymer) or 2D (surface)) are intermittent [Bénichou et al., 2008a]. Indeed, the reactant can either diffuse in the bulk volume, where it cannot find a target, or bind to the interface and diffuse more slowly (see figure 105 left). Beyond evident optimizations (increasing the target and the reactant concentrations, increasing the diffusion coefficients of the reactant in the bulk or at the interface, etc.), the mean duration of the phases (free or bound to the surface) are the main adjustable parameters. Indeed, the mean search time can be optimized tuning these parameters, in particular using



Without flow

With a flow

Figure 105: Design of heterogeneous chemical reactions, with targets (green disks) here fixed on a 2-dimensional surface. The reactant either diffuses in the volume (blue), or diffuses on the surface (orange thick line). The flow is represented by magenta arrows.

the "teleportation" approximation. In this approximation, after a bulk excursion, the distance between the reactant landing point on the surface and its starting point is larger than the typical distance between targets. With such a condition, each new bound phase is independent from the previous one, limiting overlap. The mean durations of the phases can actually be tuned in real systems. The mean time a reactant remains bound to the interface depends on its affinity with the interface, which could be tuned. The mean time spent in the bulk is mainly controlled by the confinement volume [Kac, 1959, Blanco and Fournier, 2003, Bénichou et al., 2005b, Condamin et al., 2007]. The mean time spent in the bulk should be large enough to make the approximation of "teleportation" valid, but otherwise as small as possible to save time. Another possibility is to apply a hydrodynamic flow parallel to the surface, which makes the teleportation approximation valid even for a very small time spent is the bulk, provided the velocity of the flow is high enough (see figure 105 right). There are regimes in which intermittence is favorable and can be optimized [Bénichou et al., 2008a].

6.5.2 Macroscopic scale

Systematic search

Stochastic search

Figure 106: Design of macroscopic searches. Blue line : searcher trajectory. Red line (left) : target trajectory.

At the microscopic scale, stochastic strategies are often the only available strategies. But at the macroscopic scale, for example in robotics, what is the justification for a stochastic search? Indeed, systematic strategies often find targets faster than stochastic strategies. There could be material considerations (limitation in robot memory, accumulations of mistakes, ...), coupled with the fact that the difference between the mean search time using systematic strategies compared to stochastic

strategies can be small. Beyond these contingent aspects, if the target does not want to be found, and if it can watch the strategy adopted by the robot, the target can easily escape a systematic strategy. For example, if police search apartments in a building one by one, respecting the number, only preventing people from leaving the building but not from circulating from one apartment to another, it is enough that the wanted person switch apartment before the police arrive to the apartment where he initially was. In contrast, if police search at random, some time is lost as already searched apartments are searched again, but there is no strategy ensuring escape of the wanted man (see figure 106). In the case of targets that do not want to be found and that have memory, stochastic search, by its unpredictability, will be more efficient (except if the target guesses which aleatory number generator is used, which is more difficult than understanding a systematic strategy). Such stochastic strategies could be optimized depending on the context.

6.6 Comparison with experimental data

6.6.1 Motivation

Our models show that intermittent search is efficient and robust. Consequently, we expect intermittent search to be a widespread strategy. Indeed, as seen in sections 3 and 4, intermittent search is observed in real systems. However, to go further, a more quantitative comparison with experiments is needed.

We have a project with Ana-Maria Lennon-Duménil ("Protéase et immunité" team, Institut Curie, Paris, France) on dendritic cells. These cells are of particular interest because they travel in the organism searching for pathogens. We have another project with Vincent Fourcassié (CRCA, Université Paul Sabatier, Toulouse, France) on ants. However these projects are not advanced enough to be presented here.

Another interesting animal is the human being. Indeed, the idea itself of intermittent strategies stems from the every-day life situation of searching for a key, or any small object. We have shown that when a searcher alternates between a slow phase with detection and a fast phase without detection, there are regimes of favorable intermittence. But how does a human really search for a small object?

To answer this question, we observed trajectories of humans searching for a small object, with the pressure of finding it rapidly.

The work presented in the following is done in collaboration with Michel Sokolowski, searcher in psychology at the University of Picardie - Jules Verne (Amiens, France). We present here preliminary results.

6.6.2 Experiment

The set-up is a circle of 20 meters of diameter in a lawn (see figure 107 right). A camera records the subjects searching for targets in the arena (see figure 107 left) (telling how the camera is controlled, how data are acquired, how to power all this, the technical problems that aroused and how they were eventually solved, would be a long story, not uninteresting, but probably irrelevant here).

Image taken from the camera, where you can see a subject in an arena.

The camera is on top of a vertical tube, which is installed on a bank close to the arenas.

Figure 107: Experimental set-up.

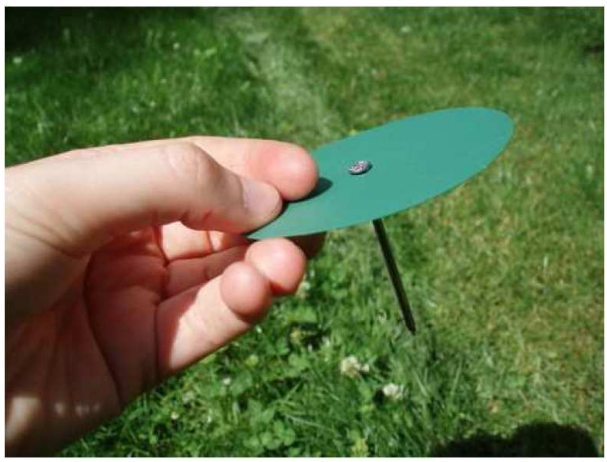


Figure 108: One of the large targets. The nail fixes the target to the ground. The smaller targets are of the same color.

For practical reasons, subjects are mainly students in psychology. There is always only one subject at the same time, for not influencing the subject. For the same reason, we ask them not to tell about the experiment to their fellow students. First, we explain to them the procedure of the experiment, and tell them that the subjects finding the more rapidly the targets will get a prize (a sum of money depending of their rank; they are not paid otherwise, except that participating to such experiments increases their grades).

There are 3 conditions in which the experiment is performed.

- The "normal" condition : subjects are not restricted in their vision, and the

only perturbation is that they are clothed in white, to facilitate subsequent analysis of the movies.

- The "hat" condition : subjects wear a hat preventing them from seeing what is around them. They can see the ground up to 1 meter in front of them.
- The "blind" condition : they are blindfolded.

The "normal" condition is the most natural condition. The "hat" condition is aimed at preventing the subjects from using external benchmarks to orientate themselves. The "blind" condition is aimed at enforcing an intermittent behavior : it is really hard to move while scanning the ground for the target. In the "normal" and the "hat" conditions, there are 3 targets in the arena. A target is a green plastic disk of 3 mm of diameter, mounted on a small nail to remain fixed on the ground (of the same color than the larger target of figure 108). In the "blind" condition, there are 12 targets (as moves are much slower in this condition, a higher density is needed to obtain a similar probability of finding a target), the targets are larger (10 cm of diameter), and can be sensed when scanned by hand, but not sensed when walking on them. The targets are placed at random. To ensure that the randomness is really "random", random patterns are made before the experiment with a computer program, and targets are placed according to these patterns (with more or less accuracy, but the randomness is satisfying).

The subjects are first trained to find the target in a 1 meter wide corridor. They also stay for one minute in the empty arena to get used to it. Then they are placed near the center of the arena and have 7 minutes to search for targets. This search is repeated three times (with the old targets removed and new targets added).

Some of the subjects do two of the conditions, with the same procedure.

6.6.3 Data analysis

The acquired movies are first decomposed into images. Then images are analyzed using ImageJ, an open source software [Rasband, 1997-2009]. Figure 109 shows the process of analysis from an image of the movie to the detection of the subject's position : the last image (Figure 109 (f)) shows the position of the subject. The horizontal position is taken as the horizontal of the barycenter, and the vertical position is taken as the lowest point of the detected shape. Positions are then renormalized for the vision angle. The output are trajectories. These trajectories are smoothed (in the vertical direction in particular, there are moves of one pixel leading to abrupt changes, and unsmoothed trajectories present regular "wave" patterns corresponding to the walk itself) (see figure 110). After all this process, the result is exploitable trajectories of the subjects.

6.6.4 Preliminary results

The "normal" and "hat" conditions give very similar results. Full vision does not seem to really help the search in these conditions, at least in the light of the first data gathered. Most trajectories look like spirals, or like Pearson walks (straight lines reflected by the boundaries) [Pearson, 1905], or a combination of both. However,

(a) starting point

(b) blue canal of image (a)

(d) result of the substraction (b)-(c)

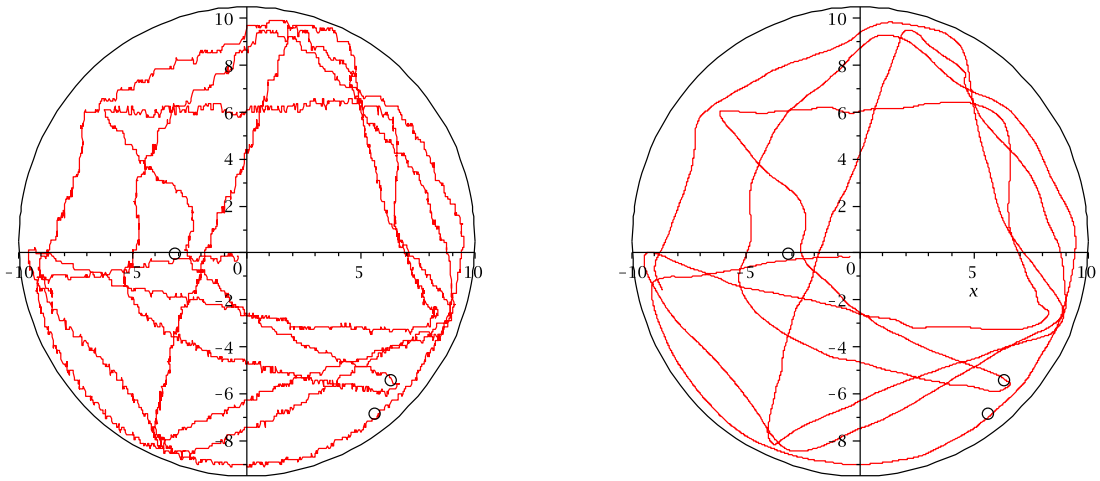
(c) image of the empty arena

(e) thresholded image (d)

(f) automatically detected shape

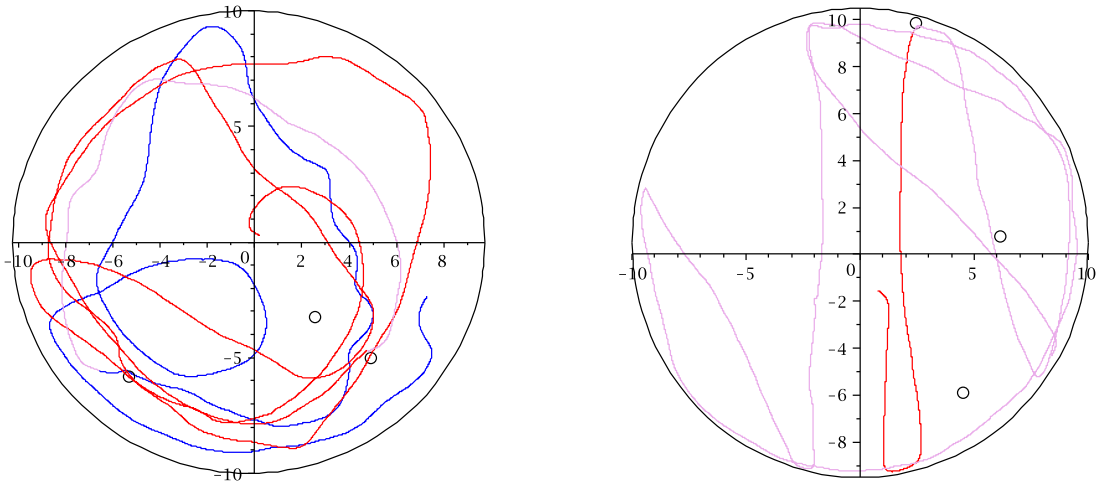
Figure 109: Image analysis : The initial image (a) is decomposed into the different color canals. The canal used is the blue one (b), as it has the best contrast. An image of the empty arena (c) is subtracted, leading to the image (d). This image is then thresholded, resulting in the image (e). Then a routine of ImageJ identifies the shape (image (f)) (with here the additional rule that shapes are made of a number of pixels larger than a cut-off, in order to remove noise).

spirals (see for example figure 111 left) are far from being perfect spirals. The subjects performing such trajectories declared afterwards (they were asked at the end if they applied a strategy, and if so, which one) that they wanted to make a spiral to cover the space without going back to previously scanned areas : when



raw trajectory smoothed trajectory

Figure 110: Trajectory analysis : trajectories are first smoothed.

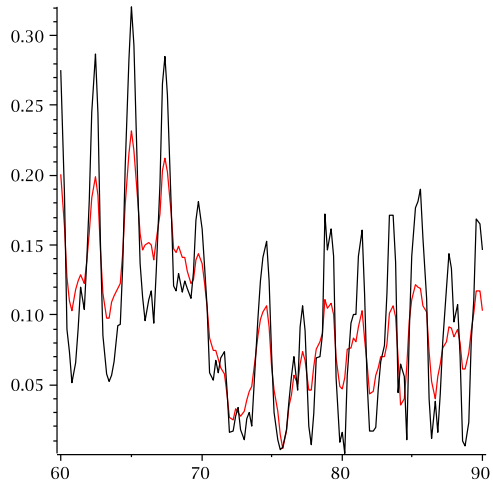


Spiral trajectory Pearson trajectory

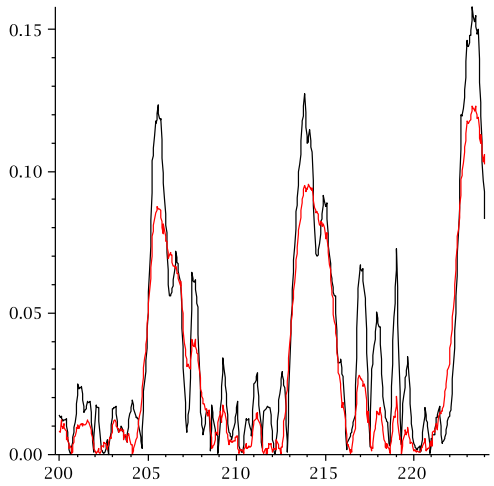
Figure 111: Representative examples of 7 minute trajectories in the “normal“ and “hat“ conditions (here, the ”hat” condition). The large black circle represents the arena (with the scale in meters), the colored lines are the trajectory, with a change of color when a target is detected. The targets positions are represented by small black circles.

we see the real trajectories, it seems that they did not succeed! However, as they can miss the target even passing on it, it did not really seem a handicap, at least qualitatively. In these conditions, search is not intermittent. The subjects look at the ground all along the trajectory. These trajectories are however interesting and further analysis is on the way.

The “blind” condition makes the search intermittent (see figure 112). Indeed, the ground has to be scanned by hand. The subjects most often kneel down on the ground, and use their hand to ease their displacements. So either they scan, or they move, but not both at the same time. There are very strong correlations between “ballistic” phases : subjects declared trying to follow a straight line until arriving to



Fast scan. One knee is moved, then one hand scans, then the symmetric, and so on



Slow scan. Both knees are moved, then the area is scanned with both hands, and so on

Figure 112: Velocity on "blind" trajectories. Velocity (m/s) as a function of time (s). Black and red lines represent different degrees of smoothing of the trajectory.

the boundary, and they succeeded quite well in this.

These data are quite qualitative, because only a small part of the movies have been analyzed. The analysis of the other movies is under way. The total number of subjects is however relatively small, but interesting information can be extracted, even if not fully quantitative.

7 Conclusion

Intermittent search strategies rely on a simple assumption : the searcher alternates between two phases, one during which the target can be detected, but with slow motion, and another of faster motion but without target detection. The latter phase could be detrimental, as it means spending time in a state with no target detection, but it also could be beneficial, as it allows the searcher to better explore the search domain.

This simple idea of intermittence comes from the observation of real-life biological searches. At the macroscopic scale, an example is given by animals searching for hidden food, who alternate between fast ballistic relocation phases with no target detection, and phases of slower motion aimed at detecting the target. After reviewing a simple model based on this observation, we have extended it in two directions. On the one hand, we have studied the influence of the target distribution. On the other hand, as the initial model assumes infinite correlations in the direction of the searcher, we have extended it to the case of a random reorientations in two dimensions. This model is particularly relevant for animals living on the ground, which is a 2-dimensional interface. We have shown analytically that the mean search time can be minimized as a function of the phases mean duration. There is one single way to share time between the two phases to find the target as fast as possible. This intermittent search is then an optimal search strategy. In this aspect, this model is an alternative to the famous Lévy walks model which is optimal only in restrictive conditions.

Intermittence is also observed at the microscopic scale. A first example involves specific DNA sequence localization by proteins in cells : as proposed by Berg et al. [1981], a protein either diffuses in the bulk of the cell, moving rapidly, or is bound non-specifically to DNA, where the target is located, but moving much slower. When a protein leaves the DNA and diffuses, as the DNA is coiled, it can either bind back to DNA to a position of the sequence correlated with the starting point (hop), or far away on the linear sequence, even if it is close in the 3-dimensional space (jump). An accurate description of the distribution of the hop size and of the proportion of hops among 3-dimensional excursions was missing. First, we have obtained the hop distribution in a simple model (normal diffusion, infinite observation time and idealized geometry). Then, we have adapted this distribution to the interpretation of a single-molecule experiment of enzyme interaction with DNA, with a finite observation time and a specific geometry. The observations have been compared quantitatively with the model, showing that the enzyme does combine sliding and hopping. Finally, we have extended our calculations to the hop distribution in a crowded medium. Another example of intermittence at the microscopic scale is given by active transport of vesicles reacting with a specific target within cells. Vesicles can either freely diffuse or bind to motors performing ballistic motion. We have studied this transport in the framework of a simple analytical model. As expected, vesicles, diffusing slowly, react faster with intermittent transport. The gain is small in the cytoplasm (3D), but for a membrane (2D) or a tubular structure (1D), the gain can be very large if the target concentration is low.

Since these intermittent search strategies are observed at various scales, we have

proposed the hypothesis that they could constitute a generic search mechanism. In a more technical section, we have studied systematically a generic model in the framework of intermittent random walks, in 1, 2 and 3 dimensions, and for three different descriptions of the slow reactive phase, in order to account for the variety of real-life situations on the one hand, and to assess the robustness of this model on the other hand. This study shows that the optimality of these search strategies is a widely robust result. Finally, if intermittent random walks are observed in real biological systems, it is probably because they do constitute an efficient search strategy.

8 Appendices

Contents

8.1 Complements to hops and jumps 173

8.1.1 Numerical inversion 174

8.1.2 Effective geometry 175

8.1.3 Simulations on percolation clusters 179

8.2 Complements to the generic intermittent search model . 182

8.2.1 Diffusive mode in one dimension 183

8.2.2 Ballistic mode in one dimension : exact result 187

8.2.3 Static mode in 3 dimensions : more comparisons between the analytical expressions and the simulations 189

8.2.4 Diffusive mode in 3 dimensions 190

8.2.5 Ballistic mode in 3 dimensions 191

8.3 Complements to the extensions and perspectives 193

8.3.1 Appendix : span of a phase of 1D diffusion 193

8.3.2 Appendix : details of the simulation of intermittent model with Lévy-distributed relocations 196

8.3.3 Appendix : static mode in one dimension with perfect correlation 197

8.4 An example of the importance of transport in biology . 200

8.4.1 Introduction 200

8.4.2 HEK cells 200

8.4.3 Neurons 207

8.4.4 Conclusion 210

The appendices are in fact of two types. Appendices 8.1, 8.2 and 8.3 give details on calculations or simulations too cumbersome for the main text (even if important from the technical and practical point of view). Appendix 8.4 is presenting another work which is not directly related to intermittent search strategies. However, similarly to section 4.2, this appendix is about vesicle transport within cells : we have worked on a model interpreting experimental data showing the relation between a membrane receptor activity and its distribution, which has important functional implications.

8.1 Complements to hops and jumps

Here we give complements to the section on hops and jumps of a protein interacting with DNA (see section 4.1).

8.1.1 Numerical inversion

We use a numerical inversion of a Laplace transform (see section 4.1.3). Here its practical implementation is detailed.

Method for numerical inversion

$P_{t_{obs}}(z)$ is the probability of observing a hop larger than z during t_{obs}, $P_{\parallel}(z|t)$ is the probability of a displacement z along the cylinder axis after a time t by Brownian diffusion, $F_{\perp}(t)$ is the probability of going back to the cylinder at t. These quantities are related :

$$P_{t_{obs}}(z) = \int_0^{t_{obs}} P_{\parallel}(z|t) F_{\perp}(t) \mathrm{d}t. \tag{342}$$

Movement is independent along the different directions (not that z is here the absolute value of the distance along the axis of the cylinder) :

$$P_{\parallel}(z|t) = \frac{1}{\sqrt{\pi D t}} e^{-\frac{z^2}{4Dt}}. \tag{343}$$

$F_{\perp}(t)$ is the probability of coming back at t in 2 dimensions on a disk of radius a and with the boundary condition characterized by κ, starting at $t = 0$ on $r_0 \geq a$. The Laplace transform $F_{\perp}(s)$ of $F_{\perp}(t)$ writes [Redner, 2001] (K_0 is the second kind of modified Bessel function of order 0) :

$$F_{\perp}(s) = \frac{K_0\left(r_0\sqrt{\frac{s}{D}}\right)}{K_0\left(a\sqrt{\frac{s}{D}}\right) + \frac{\sqrt{s}}{\kappa\sqrt{D}} K_1\left(a\sqrt{\frac{s}{D}}\right)}. \tag{344}$$

Using the Stehfest algorithm in Maple [Vogt, 2006], we inverse the Laplace transform $F_{\perp}(s)$. Then we take the integral as a sum :

$$P_{t_{obs}}(z) = \int_0^{t_{obs}} P_{\parallel}(z|t) F_{\perp}(t) dt \simeq \sum_{i=1}^{N} P_{\parallel}(z|t_i) F_{\perp}(t_i) \delta^i t. \tag{345}$$

The problem is to choose the right interval $\delta^i t$: short enough to have an accurate estimate, and not too long to keep calculation time reasonable. When $\kappa \to \infty$ and $r_0 > a$, $F_{\perp}(t)$ has a maximum, at approximately $t_r = \frac{(r_0-a)^2}{6D}$ (minimal time to go back to the cylinder, that can also be found using the small time approximation in Levitz et al. [2008]). After some tests, we have chosen $\alpha = 0.06 t_r$. When κ is finite and $r_0 = a$, $F_{\perp}(t) \to \infty$ for $t \to 0$. However, $tF_{\perp}(t)$ decreases when $t \to 0$. After some tests, we have chosen α such that $\alpha F_{\perp}(\alpha) \simeq 10^{-5}$. The integral on t from 0 to t_{obs} is sliced as follows :

- the first interval $t \in [0, \alpha]$, with value $F_{\perp}(\alpha)/2$,
- n_{max} intervals ($i \in [1, n_{max}]$) $t \in [\alpha\beta^{i-1}, \alpha\beta^i]$, with value $F_{\perp}(\alpha\beta^{i-1}(\beta-1)/2)$, and n_{max} defined such that $\alpha\beta^{n_{max}} < t_{obs}$ and $\alpha\beta^{n_{max}+1} > t_{obs}$,
- the last interval $t \in [\alpha\beta^i, t_{obs}]$, of value $F_{\perp}(\alpha\beta^{n_{max}} - t_{obs})/2)$.

We have chosen $\beta = 1.05$ for a good precision.

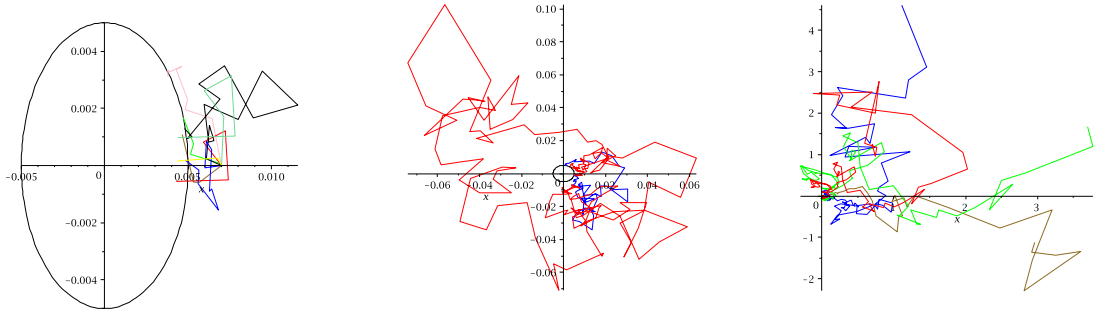
To obtain the cumulative instead of the direct probability :

$$C_{t_{obs}}(z) = \int_{z}^{z_{max}} P_{t_{obs}}(z)dz + \int_{z_{max}}^{\infty} P_{t_{obs}}(z)dz. \tag{346}$$

As the numerical calculation from a finite z to infinity is long, we calculate first the integral from z_{max} (taken here $= 2.2\mu$m) to infinity, then add it to integrals between finite z.

Simulations methods

To check the results of the numerical inversion, we have performed simulations.



Some trajectories quickly coming back to the disk.

Two trajectories coming back to the disk after a large excursion.

Four trajectories not returning to the disk before t_{obs}.

Figure 113: some trajectories from simulations, starting from $(x = a + r_0, y = 0)$, with $a = 5$ nm, $r_0 = 2$ nm. Disk of radius a in black, x and y in μm. $D = 54$ μm^2.s^{-1}, and $t_{obs} = 0.04$ s.

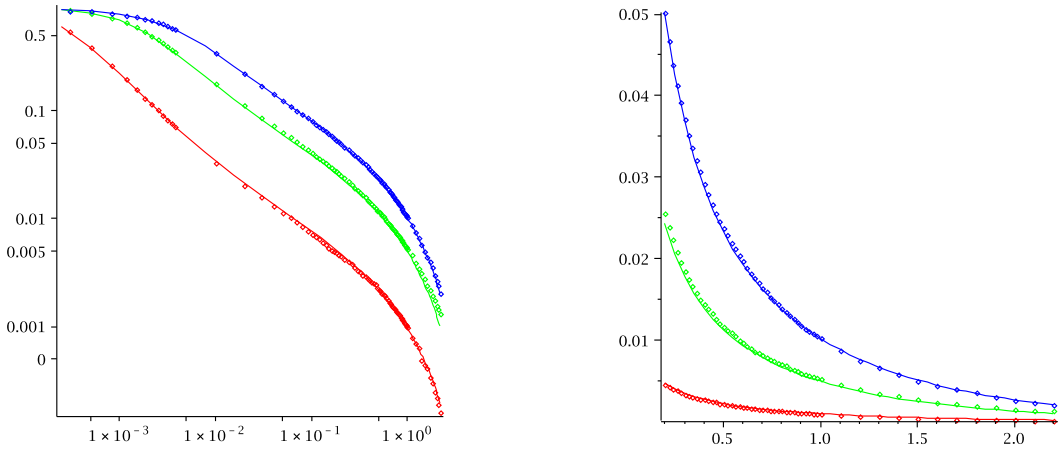
For the simulations we take $\kappa \to \infty$ and $r_0 > a$, and the simple case of an infinite cylinder in a infinite space. In this case, the problem reduces to a disk of radius a in an infinite plane. At each time step the distance from the disk is calculated, and time steps are taken smaller close to the disk, as proposed by Berezhkovskii et al. [1998] : steps need to be small for accuracy near the disk, but it saves time to make large steps far away from the disk (see figure 113).

Comparison between numerical inversion and simulations

Simulations and numerical inversion are in good agreement (see figure 114).

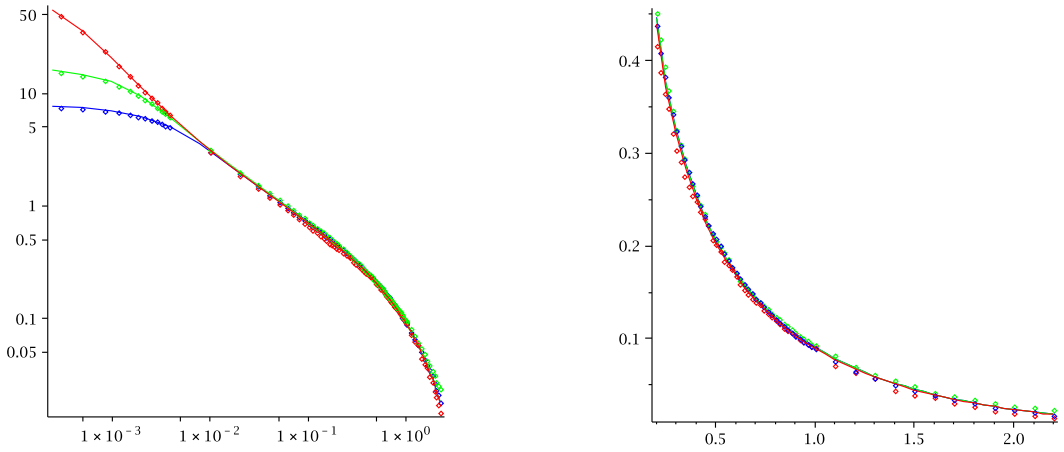
8.1.2 Effective geometry

To interpret experimental data, we have to choose an effective geometry where analytical calculations are tractable (see section 4.1.3). Here, we present two other model geometries that could have been thought of, but that are not as good as the chosen geometry. The chosen geometry is a cylinder between reflective planes : we also present the practical implementation of the result, and additionally some other calculations in this model.



loglog scale linear scale

Probability that for a 3D excursion, the resulting hop is larger than z, as a function of z (µm).



loglog scale linear scale

Probability of observing a hop larger than z during an effective interaction as a function of z (µm).

Figure 114: Comparison between numerical inversion (lines) and simulations (symbols). $D = 54$ µm^2.s^{-1}, $a = 5$ nm, $t_{obs} = 0.04$ s. $r_0 - a = 0.34$ nm (red), $r_0 - a = 2$ nm (green), $r_0 - a = 5$ nm (blue).

Cylinder partly absorbing

A first possibility to take into account the fact that the DNA molecule is finite is to represent the cylinder as infinite, but to define hops returning at $z < 0$ or $z > L$ as losses, along with returns taking longer than t_{obs}. If $P(z)$ is the probability of a return of size z for an infinite cylinder, then for such a partly absorbing cylinder, the corresponding probability is $P^*(z) = \frac{L-z}{L} P(z)$.

However, such a representation does not take into account at all that there is also a reflecting surface, that will allow enzymes to return more often. In particular, the distribution for large returns will be biased. Yet, the experimental observation is

the number of hops by effective interaction. The theoretical distribution of hops by 3D excursion is thus renormalized by the probability that an enzyme is lost during a 3D excursion. As a consequence, a bias on the large hops would be a problem for a quantitative analysis.

Cylindrical confinement

A way to represent the reflecting surface could be to confine the cylinder in a reflecting cylinder of larger radius R.

Following the method of Redner [2001], in 2 dimensions, a partially reflecting disk stands at $a \leq r_0$ (parameter κ), a reflecting circle at $R > r_0$, and the walker starts at r_0. $p(r,t)$ is the probability of being at a radius r at t :

$$p(r, t = 0) = \frac{\delta(r - r_0)}{2\pi r}. \tag{347}$$

$c(x,s)$ is the Laplace transform of p, with the change in coordinates $x = r\sqrt{\frac{s}{D}}$ ($x_0 = r_0\sqrt{\frac{s}{D}}$, $x_a = r_a\sqrt{\frac{s}{D}}$, $X = R\sqrt{\frac{s}{D}}$, $x_k = \frac{1}{\kappa}\sqrt{\frac{s}{D}}$). c is solution of :

$$\frac{\partial^2 c}{\partial x^2} + \frac{1}{x}\frac{\partial c}{\partial x} - c = -\frac{\delta(x - x_0)}{2\pi x_0} \tag{348}$$

The solutions are the modified Bessel functions $I_0(x)$ and $K_0(x)$. We have to study separately the domains $x < x_0$ and $x > x_0$, then use the boundary conditions. $c_> = AI_0(x) + BK_0(x)$ and $c_< = CI_0(x) + EK_0(x)$. At the outer circle :

$$\left.\frac{\partial c_>}{\partial x}\right|_{x=X} = 0. \tag{349}$$

At the inner circle :

$$\left.\frac{\partial c_<}{\partial x}\right|_{x=X} = \frac{1}{x_\kappa} c_<(x = x_a). \tag{350}$$

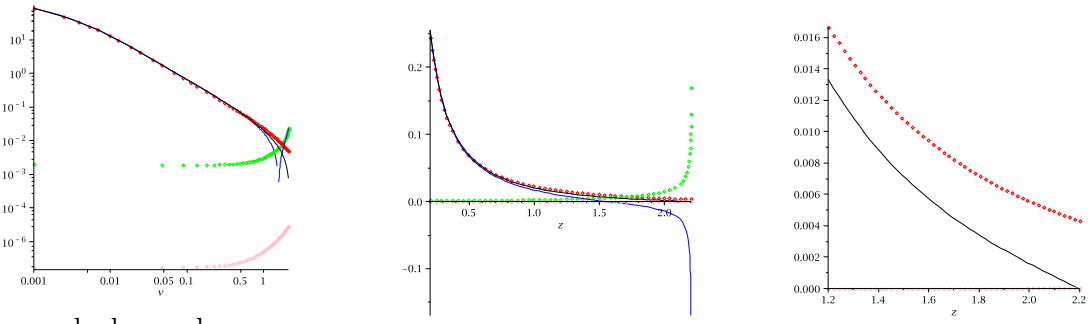
c is continuous in $x = x_0$ ($c_>(x_0) = c_<(x_0)$). Last, the initial conditions are translated by :

$$\left.\frac{\partial c_>}{\partial x}\right|_{x_0^+} - \left.\frac{\partial c_<}{\partial x}\right|_{x_0^-} = -\frac{1}{2\pi x_0 D}. \tag{351}$$

We obtain $c_>$ and $c_<$, using the Wronskian relation ($K_0(x)I_1(x) + K_1(x)I_0(x) = \frac{1}{x}$).

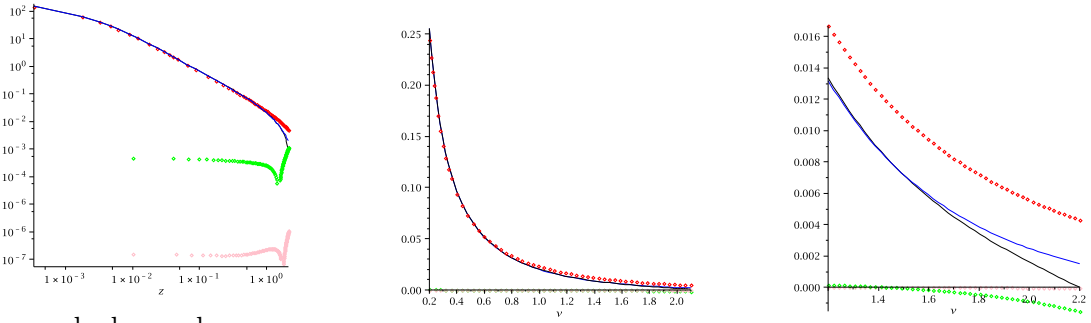
$$F(s) = D2\pi x_a \left.\frac{dc_<}{dx}\right|_{x=x_a} = \frac{K_1(X)I_0(x_0) + I_1(X)K_0(x_0)}{K_1(X)\left(I_0(x_a) - x_\kappa I_1(x_a)\right) + I_1(X)\left(K_0(x_a) - x_\kappa K_1(x_a)\right)}. \tag{352}$$

This representation of the reflective plane can be sophisticated by adding a finite-size effect as proposed previously (cylinder absorbing outside $z \in [0, L]$). However, in this geometry, either we take R large, and reflection would not be significant within t_{obs}, whereas the real surface is relatively close to DNA in experiments; or we take R small, and there would be much more returns than with a planar surface : indeed, in this model geometry, there is a confining surface all around. It is not a satisfying model geometry to model the experiment.



loglog scale (absolute value) | linear scale | linear scale, zoom

Figure 115: Probability that a 3D excursion leads to a hop of length z before t_{obs} as a function of z (µm). Calculation of $P^*(z)$ using equations 105, 108 and 110. $P(z)$ for an infinite cylinder (red dots), W_0 (108) (blue line), W_1 (110) (green dots), W_2 (110) (pink dots), and $P^*(z) = W_0 + W_1$ (black line). With $r_0 = a$, $t_{obs} = 40$ ms, $D_{3D} = 54$ µm^2.s^{-1}, $a = 5$ nm, $L = 2.2$ µm, and $\kappa^{-1} = 10$ nm.



loglog scale (absolute value) | linear scale | linear scale, zoom

Figure 116: Probability that a 3D excursion leads to a hop of length z before t_{obs} as a function of z (µm). Calculation of $P^*(z)$ using equations 113, 114 and 115. $p(z)$ for an infinite cylinder (red dots), Y_0 (114) (blue line), Y_1 (115) (green dots), Y_2 (115) (pink dots), and $P^*(z) = Y_0 + Y_1$ (black line). With $r_0 = a$, $t_{obs} = 40$ ms, $D_{3D} = 54$ µm^2.s^{-1}, $a = 5$ nm, $L = 2.2$ μm, and $\kappa^{-1} = 10$ nm.

Practical numerical implementation of the reflective boundaries

We have chosen a cylinder between reflective planes as the model geometry. For the numerical calculation of $P^*(z)$, we use the equation 113, but keeping only $n = 0$ (equation 114) and $n = 1$ (equation 115); for small values of z, $n = 0$ is sufficient (see figure 116). However, the formula 105 is useful too, because integrating from v finite to $v \to \infty$ is long (and is a required step for the calculation of Y_0 (114)). As it is the case for Y_i, the first two terms of the W_i sum are enough (even the first term for z small) (see figure 115).

Additional calculations in the case of reflective boundaries

When we calculate the influence of the small hops on the effective diffusion coefficient, we need N the number of 3D excursions by interaction, the interaction being ended if there is a hop larger than $z_m = 0.2$ µm, or if the return time to the cylinder is longer than t_{obs}. We also need $\langle l^2 \rangle = \int_0^{z_m} z^2 P^*(z) dz$. Indeed, the

correction to the 1D diffusion coefficient due to small hops is $\frac{1}{2}(N-1)\langle l^2\rangle$. As z_m is small, we can make the approximation $P^*(z) \simeq W_0$ or $P^*(z) \simeq Y_0$. q is the probability that a 3D excursion results in losing the enzyme (loss if $t_{return} > t_{obs}$ and if $z_{return} > z_m$), $N = 1/q$. We define :

$$R_W = \frac{1}{L}\int_0^{z_m} W_0 dv, \tag{353}$$

$$R_Y = \frac{1}{L}\int_0^{z_m} Y_0 dv. \tag{354}$$

We will use $q = 1 - R_w$ or $q = 1 - R_Y$. After some calculations :

$$R_W = \int_0^{z_m} p(v)dv + \frac{1}{L}\left(z_m \int_{z_m}^{L-z_m} p(v)dv + \frac{L+z_m}{2}\int_{L-z_m}^{L+z_m} p(v)dv - \frac{1}{2}\int_{L-z_m}^{L+z_m} vp(v)dv\right), \tag{355}$$

$$R_Y = \int_0^{z_m} p(v)dv + \frac{z_m}{L}\int_{z_m}^{\infty} p(v)dv. \tag{356}$$

We define :

$$\langle l^2\rangle_W = \frac{1}{L}\int_0^{z_m} v^2 W_0 dv, \tag{357}$$

$$\langle l^2\rangle_Y = \frac{1}{L}\int_0^{z_m} v^2 Y_0 dv. \tag{358}$$

After some calculations :

$$\begin{aligned}\langle l^2\rangle_W = \; & \int_0^{z_m} v^2 p(v)dv \\ & + \tfrac{1}{6L}\left(-4\int_0^{z_m} v^3 p(v)dv + z_m^3\left(2\int_{z_m}^{L-z_m} p(v)dv + \int_{L-z_m}^{L+z_m} p(v)dv\right) + \int_{L-z_m}^{L+z_m}(L-v)^3 p(v)dv\right)\end{aligned} \tag{359}$$

$$\langle l^2\rangle_Y = \int_0^{z_m} v^2 p(v)dv + \frac{1}{3L}\left(z_m^3\int_{z_m}^{\infty} p(v)dv - 2\int_0^{z_m} v^3 p(v)dv\right). \tag{360}$$

The simplified expressions (equation 355 for instance) are faster to compute in numerical calculations than the direct expression (equation 353 for instance). The difference between R_W and R_Y, and between $\langle l^2\rangle_W$ and $\langle l^2\rangle_Y$, is less than 0.4%, at least in the tested cases.

8.1.3 Simulations on percolation clusters

Simulations are made on percolation clusters to check the analytical formula of hops in a fractal medium (see section 4.1.4).

Introduction to percolation clusters

To check the hop distribution when the environment around the DNA is fractal (equation 124), we perform simulations on percolation clusters. More precisely, we use percolation of bonds in a square lattice. DNA is represented by a cylinder. In the simulations, it is represented by a line of the lattice. We define :

- p_c the critical probability (simulations are done with $p_{bound} = p_c$)

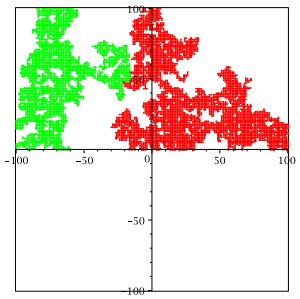
- d_f the fractal dimension of the critical cluster
- $d_f^{\perp}$ the dimension of the projection of the critical cluster on the plane perpendicular to the cylinder
- $d_f^{\parallel}$ the dimension of the projection of the critical cluster on the plane parallel to the cylinder
- d_w the dimension of a walk in the critical cluster (same dimension if projected on a line or on a plane)
- d'_w the dimension of a walk in a lattice with $p_{bound} = p_c$, but on any cluster.

The relevant parameters for us are (for more information about percolation, see Ben-Avraham and Havlin [2000] (chapters 2 and 6), and Hughes [1996] (in particular page 176)) :

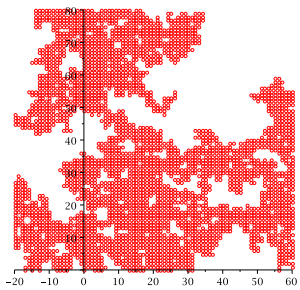
	p_c	d_f	$d_f^{\perp}$	$d_f^{\parallel}$	d_w	d'_w
2D	0.5	91/48	1	1	2.878	3.04
3D	0.248812	2.53	2	1	3.88	5.07

We checked that on simulated lattices, the obtained values of the dimensions are coherent with these expected values.

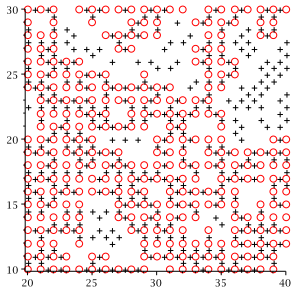
2 dimensions



The largest cluster (red), and another cluster (green)



Zoom on the largest cluster

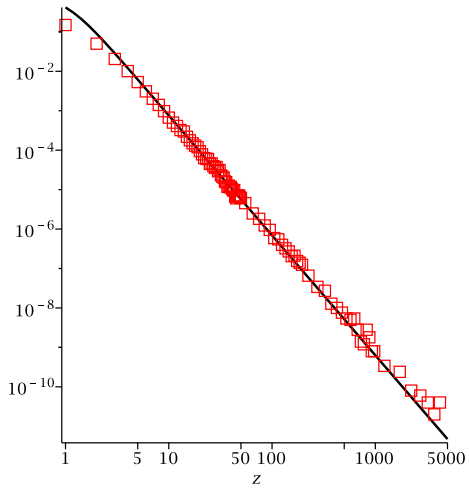


Zoom : bonds (+) and nodes (○) belonging to the largest cluster

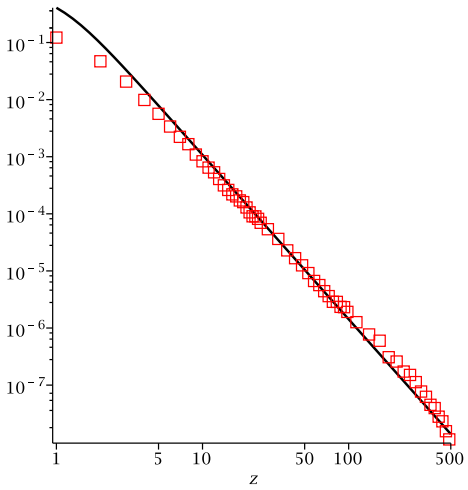
Figure 117: Percolation clusters in 2 dimensions. Simulation on a finite lattice (201×100 nodes). The cylinder representing DNA in $y = 0$. The largest cluster (in red) does not extend from side to side in the horizontal direction : indeed, the probability to percolate from one side to the other tends to 1 when the size of the system tends to infinity.

Figure 117 shows how the largest cluster looks like in a simulation in two dimensions.

The analytical formula (Eq. 124) works well (see figure 118) in two dimensions, with a walk on any cluster as well as on the largest cluster.

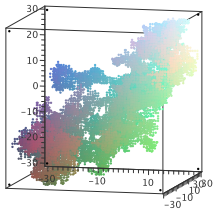


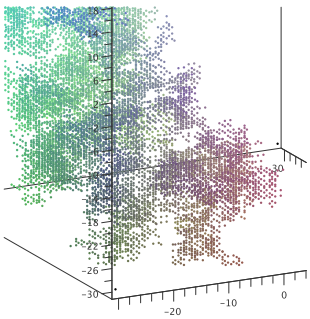
Start from any cluster. Simulation : 10 000 × 20 001 nodes



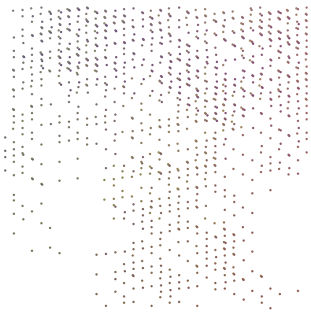
Start on the largest cluster. Simulations : start close to the center of the largest cluster, in a lattice of 1001 × 1001 nodes

Figure 118: $log(p(z))$ the probability of returning to the cylinder at a distance z as a function of $log(z)$ in 2 dimensions. Black line : analytical formula (Eq. 124). Red squares : simulations. Simulations are averaged over 10 000 clusters, and over 10 000 walks by cluster.





Zoom



Zoom

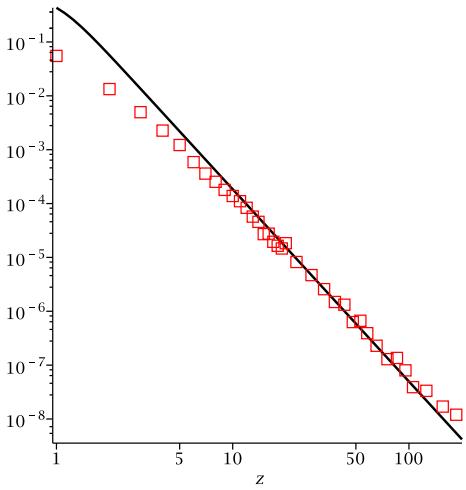
Figure 119: Percolation cluster in 3 dimensions : simulations in a finite lattice ($61 \times 61 \times 61$ nodes). Only the nodes of the largest cluster are represented. The cylinder representing DNA is the line $y = 0$, $x = 0$.

3 dimensions

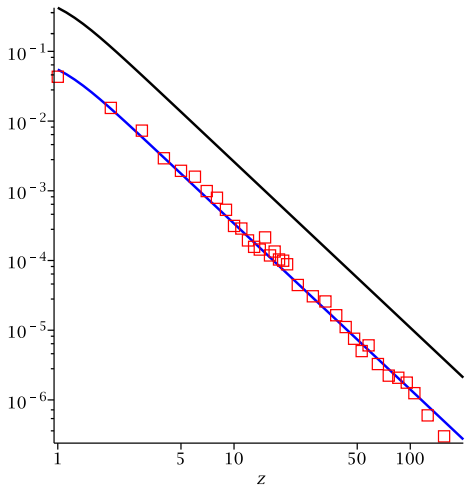
Figure 119 shows how the largest cluster looks like in a simulation in three dimensions.

The agreement between simulations and the analytical formula (124) is not as good in three dimensions as in two dimensions (see figure 120). Problems are concentrated on z small. Indeed, the calculation leading to the formula (124) has been made in continuous space, whereas in simulations space is discrete. For a walk on any clusters, simulation points are well fitted by the theoretical curve for z larger than a few nodes. For a walk starting on the largest cluster, there is a considerable bias for the point $z = 0$, that shifts the rest of the curve : when divided by the adequate coefficient, the simulation points and the theoretical curve are in good agreement.

For the largest cluster in 3 dimensions, we renormalize the results by the number



Start from any cluster.
Simulation : average on 10 000 clusters



Start on the largest cluster.
Simulations : average on 1000 clusters.

Figure 120: $log(p(z))$ the probability of returning to the cylinder at a distance z as a function of $log(z)$ in 3 dimensions. Black line : analytical formula (Eq. 124). Blue line : analytical formula multiplied by a fitted coefficient (0.13). □ : simulations. Simulations are in lattices of $401 \times 401 \times 401$ nodes, and for each cluster 10 000 walks are performed.

of returns equal or larger than 2 nodes, to remove the influence of the first points, where the discretization does not represent well a continuous space (see figure 121). It is what we use in the figure 50 in the main text.

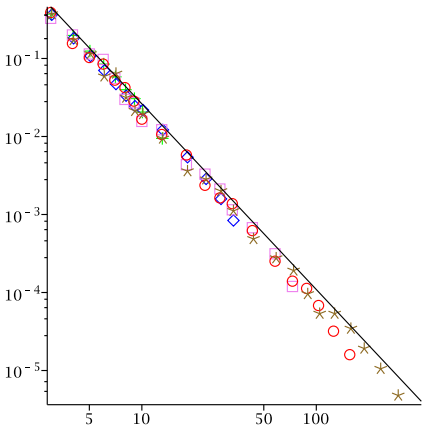


Figure 121: $log(C(z))$ the probability of returning to the cylinder at a distance larger than z as a function of $log(z)$ in 3 dimensions, renormalized by the number of returns larger than 2 nodes. Black line : analytical formula. Symbols : simulation, with a box of side length 41 (+), 101 (◇), 201 (□), 401 (○), 801 (⋆).

8.2 Complements to the generic intermittent search model

Here we present additional material about the generic search model (see section 5).

8.2.1 Diffusive mode in one dimension

Exact results (see section 5.3.2)

The mean detection time for the diffusive mode in one dimension exactly writes :

$$t_m = \frac{1}{3}\frac{(\tau_1+\tau_2)}{\beta^2 b}\frac{Num}{Den}, \tag{361}$$

with :

$$Num = \alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7 \tag{362}$$

$$Den = \gamma_1+\gamma_2+\gamma_3+\gamma_4 \tag{363}$$

$$\alpha_1 = {L_2}^3\left(\left(3{L_2}^2\left(L_1^2-L_2^2\right)+2h^2\beta\right)h\sqrt{\beta}S+3L_1L_2\left({L_2}^4-2h^2\beta\right)C\right) \tag{364}$$

$$\alpha_2 = -L_1hL_2^5\left(2\beta+3L_2^2\right)RC \tag{365}$$

$$\alpha_3 = L_1\left(2h^4\beta^2-3L_2^8\right)BC \tag{366}$$

$$\alpha_4 = h^2\sqrt{\beta}\left(6L_2^6+h^2\beta\left(\beta+L_1^2\right)\right)RS \tag{367}$$

$$\alpha_5 = \sqrt{\beta}hL_2^3\left(4h^2\beta+3L_2^2\left(L_2^2-L_1^2\right)\right)BS \tag{368}$$

$$\alpha_6 = L_1L_2^3\left(3\left(2h^2L_2\beta+L_2^5\right)B+h\left(3L_2^2\left(\beta+L_2^2\right)+2h^2\beta\right)R\right) \tag{369}$$

$$\alpha_7 = -L_1\left(3L_2^8+2h^4\beta^2\right) \tag{370}$$

$$\gamma_1 = L_2^3L_1R\left(C-1\right) \tag{371}$$

$$\gamma_2 = \sqrt{\beta}h\left(2L_1^2+L_2^2\right)RS \tag{372}$$

$$\gamma_3 = \sqrt{\beta}L_2^3\left(B-1\right)S \tag{373}$$

$$\gamma_4 = 2h\beta L_1\left(BC-1\right) \tag{374}$$

$$B = \cosh\left(\frac{2a}{L_2}\right) \tag{375}$$

$$C = \cosh\left(2h\sqrt{L_1^{-2}+L_2^{-2}}\right) \tag{376}$$

$$R = \sinh\left(\frac{2a}{L_2}\right) \tag{377}$$

$$S = \sinh\left(2h\sqrt{L_1^{-2}+L_2^{-2}}\right) \tag{378}$$

$$\beta = L_1^2+L_2^2 \tag{379}$$

$$L_1 = \sqrt{D\tau_1} \tag{380}$$

$$L_2 = V\tau_2 \tag{381}$$

$$h = b-a. \tag{382}$$

b/a		100	10^3	10^4	10^5	10^6	10^7
a= 0.005	$gain^{th,1}$	0.085	0.39	1.8	8.5	39	180
	$gain$	1	1	2.1	8.7	40	180
	$gain^{th,2}$	0.014	0.046	0.14	0.46	1.4	4.6
	$\tau_1^{th,1}$	0.19	0.89	4.1	19	89	410
	τ_1^{opt}	∞	∞	6.1	21	90	410
	$\tau_1^{th,2}$	2.1	21	210	2100	21000	$2.1.10^5$
	$\tau_2^{th,1}$	0.38	1.8	8.2	38	180	820
	τ_2^{opt}	0	0	8.4	38	180	820
	$\tau_2^{th,2}$	0.029	0.091	0.29	0.91	2.9	9.1
a= 0.5	$gain^{th,1}$	1.8	8.5	39	180	850	4000
	$gain$	2.4	9.4	41	190	850	4000
	$gain^{th,2}$	1.4	4.6	14	46	140	460
	$\tau_1^{th,1}$	4.1	19	89	410	1900	8900
	τ_1^{opt}	3.5	15	78	390	1900	8800
	$\tau_1^{th,2}$	2.1	21	210	2100	21000	$2.1.10^5$
	$\tau_2^{th,1}$	8.2	38	180	820	3800	18000
	τ_2^{opt}	7.6	36	170	810	3800	18000
	$\tau_2^{th,2}$	2.9	9.1	29	91	290	910
a= 50	$gain^{th,1}$	39	180	850	3900	18000	85000
	$gain$	150	470	1500	5500	21000	91000
	$gain^{th,2}$	140	460	1400	4600	14000	46000
	$\tau_1^{th,1}$	89	410	1900	8900	41000	$1.9.10^5$
	τ_1^{opt}	2.2	22	230	2500	21000	$1.4.10^5$
	$\tau_1^{th,2}$	2.1	21	210	2100	21000	$2.1.10^5$
	$\tau_2^{th,1}$	180	820	3800	18000	82000	$3.8.10^5$
	τ_2^{opt}	290	980	3500	15000	72000	$3.6.10^5$
	$\tau_2^{th,2}$	290	910	2900	9100	29000	91000
a= 5000	$gain^{th,1}$	850	3900	18000	85000	$3.9.10^5$	$1.8.10^6$
	$gain$	15000	46000	$1.4.10^5$	$4.6.10^5$	$1.5.10^5$	$4.7.10^6$
	$gain^{th,2}$	14000	46000	$1.4.10^5$	$4.6.10^5$	$1.4.10^6$	$4.6.10^6$
	$\tau_1^{th,1}$	1900	8900	41000	$1.9.10^5$	$8.9.10^5$	$4.1.10^6$
	τ_1^{opt}	2.2	21	210	2100	21000	$2.2.10^5$
	$\tau_1^{th,2}$	2.1	21	210	2100	21000	$2.1.10^5$
	$\tau_2^{th,1}$	3800	1800	82000	$3.8.10^5$	$1.8.10^6$	$8.2.10^6$
	τ_2^{opt}	29000	91000	$2.9.10^5$	$9.2.10^5$	$2.9.10^6$	$9.8.10^6$
	$\tau_2^{th,2}$	29000	91000	$2.9.10^5$	$9.1.10^5$	$2.9.10^6$	$9.1.10^6$

Table 4: Diffusive mode in 1 dimension. Optimization of t_m as a function of τ_1 and τ_2 for different sets of parameters ($D = 1$, $V = 1$). For each (a, b), numerical values for the exact analytical function (361) are given with the theoretical values in the regimes where intermittence is favorable, either with $\frac{bD^2}{a^3V^2} \gg 1$ $(th, 1)$, or with $\frac{bD^2}{a^3V^2} \ll 1$ $(th, 2)$. $gain^{th,1}$ (183), $gain = t_m^{opt}/t_{\text{diff}}$, $gain^{2,th}$ (188). $\tau_1^{th,1}$ (180), τ_1^{opt}, $\tau_1^{th,2}$ (185). $\tau_2^{th,1}$ (181), τ_2^{opt}, $\tau_2^{th,2}$ (186). Colors indicate the regime : red when intermittence is not favorable, green in the $\frac{bD^2}{a^3V^2} \gg 1$ regime, blue in the $\frac{bD^2}{a^3V^2} \ll 1$ regime.

Numerical study (see section 5.3.2)

We have numerically studied the optimum of the exact t_m expression (361) (Table 4). 3 regimes can be distinguished : one with no intermittence, and two with favorable intermittence, but with different scalings. Intermittence is favorable when

$b > \frac{D}{V}$. The demarcation line between the two intermittent regimes is $\frac{bD^2}{a^3V^2} = 1$.

Details of the optimization of the regime where intermittence is favorable, with $\frac{bD^2}{a^3V^2} \gg 1$ (see section 5.3.2)

We suppose that the target density is low, *i.e.* $\frac{a}{b} \ll 1$.

We are interested in the regime where intermittence is favorable. It is clear that $2(b-a)\sqrt{L_1^{-2}+L_2^{-2}} > 2\frac{b-a}{L_1}$ and $2(b-a)\sqrt{L_1^{-2}+L_2^{-2}} > 2\frac{b-a}{L_2}$. In a regime of intermittence, one diffusion phase does not explore a significant part of the system, leading to $b/L_1 \gg 1$. Alternatively, having a ballistic phase of the size of the system is a waste of time, thus close to the optimum $b/L_2 \gg 1$. Consequently $2(b-a)\sqrt{L_1^{-2}+L_2^{-2}} \gg 1$.

We use the numerical results (Table 4) to make assumptions on the dependence of τ_1^{opt} and τ_2^{opt} with the parameters. We define k_1 and k_2 :

$$\tau_1 = (k_1)^{-1}\left(\frac{b^2D}{V^4}\right)^{\frac{1}{3}}, \tag{383}$$

$$\tau_2 = (k_2)^{-1}\left(\frac{b^2D}{V^4}\right)^{\frac{1}{3}}. \tag{384}$$

We make a development of t_m for $b \gg a$. We suppose that k_1 and k_2 do not depend on b/a :

$$t_m = \frac{1}{3}\frac{D}{V^2}\left(\frac{bV}{D}\right)^{\frac{4}{3}}\frac{k_1+k_2}{k_1k_2}\left(k_2^2+3\sqrt{k_1}\right). \tag{385}$$

We checked that this expression gives a good approximation of t_m in this regime, in particular around the optimum (see figure 65 in section 5.3.2).

Derivatives of (179) as a function of k_1 and k_2 must be equal to 0 at the optimum, leading to :

$$-3k_1^{\frac{3}{2}} + 3k_2^3 + 3\sqrt{k_1}k_2 = 0, \tag{386}$$

$$3k_1^{\frac{3}{2}} - 2k_2^3 - k_1k_2^2 = 0. \tag{387}$$

On four pairs of solutions, only one is strictly positive :

$$\tau_1^{opt} = \frac{1}{2}\sqrt[3]{\frac{2b^2D}{9V^4}}, \tag{388}$$

$$\tau_2^{opt} = \sqrt[3]{\frac{2b^2D}{9V^4}}. \tag{389}$$

Details of the optimization of the universal intermittent regime $\frac{bD^2}{a^3V^2} \ll 1$ (see section 5.3.2)

We start from the exact expression of t_m (361). We have to make assumption on the dependence of τ_2^{opt} with b and a. We define f by $\tau_2 = \frac{1}{f}\frac{a}{V}\sqrt{\frac{b}{3a}}$, and we suppose

that f is independent from a/b. We make a development of $a/b \to 0$. The first two terms give :

$$t_m \simeq \frac{b}{a}\left(\sqrt{\frac{ab}{3}}\frac{1}{Vf}+\tau_1\right)\frac{a+af^2+\sqrt{D\tau_1}f^2}{a+\sqrt{D\tau_1}}. \tag{390}$$

This expression gives a very good approximation of t_m in the $bD^2/(a^3v^2) \ll 1$ regime, especially close to the optimum (see figure 66, in section 5.3.2).

We then minimize t_m (390) as a function of f and τ_1. We introduce w defined as :

$$w = \frac{aV}{D}\sqrt{\frac{a}{b}}. \tag{391}$$

We make an assumption on the dependence of τ_1^{opt} with a/b, inferred via the numerical results :

$$s = \frac{1}{\tau_1}\frac{D}{V^2}\frac{b}{a}. \tag{392}$$

We write the equation (390) with these quantities. Its derivatives with f and s should be equal to zero at the optimum, leading to :

$$-\sqrt{3}s^{3/2}w^2+\sqrt{3}s^{3/2}w^2f^2+\sqrt{3}s\,wf^2+6\sqrt{s}wf^3+6\,f^3=0, \tag{393}$$

$$6\,s^{3/2}w^2f^3+6\,s^{3/2}f\,w^2-w^2s^2\sqrt{3}+3\,ws\,f+12\,ws\,f^3+6\sqrt{s}f^3=0. \tag{394}$$

Starting from the equation (393), we make the assumption that $a \ll b \ll \frac{a^3V^2}{D^2}$, leading to :

$$\sqrt{3}s^{3/2}w^2(f^2-1)=0. \tag{395}$$

Consequently $f = 1$. We incorporate this result to the equation (394) :

$$12\,s^{3/2}w^2-w^2s^2\sqrt{3}+15ws+6\sqrt{s}=0. \tag{396}$$

The relevant solution is :

$$s_{sol} = \left(\frac{1}{3}\frac{\sqrt[3]{u}}{w}+\frac{5\sqrt{3}+16\,w}{\sqrt[3]{u}}+\frac{4}{\sqrt{3}}\right)^2, \tag{397}$$

with :

$$u = \left(270\,w+27\sqrt{3}+192\,w^2\sqrt{3}+9\sqrt{55\sqrt{3}w+84\,w^2+27}\right)w. \tag{398}$$

When $w \to \infty$, $s_{sol} = 48$. As we have made the assumption $\frac{bD^2}{a^3V^2} = w^{-2} \ll 1$, the difference from the asymptote will be small (see figure 122).

It leads to :

$$\tau_2^{opt} = \frac{1}{f^{opt}}\frac{a}{V}\sqrt{\frac{b}{3a}} = \frac{a}{V}\sqrt{\frac{b}{3a}}, \tag{399}$$

$$\tau_1^{opt} = \frac{Db}{s^{opt}V^2a} = \frac{Db}{48V^2a}. \tag{400}$$

These equations are in agreement with numerical results (Table 4).

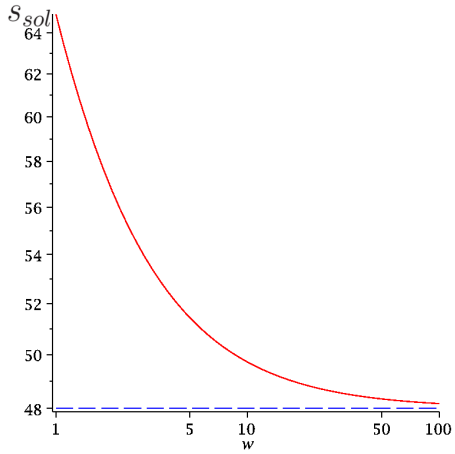


Figure 122: Diffusive mode in 1 dimension. s_{sol} (397) (red line) as a function of $\ln(w)$, with its asymptote (blue dotted line).

We use equations (177), (184) to calculate the gain :

$$t_m^{opt} \simeq \frac{2a}{V\sqrt{3}} \left(\frac{b}{a}\right)^{3/2}, \text{ and} \tag{401}$$

$$gain \simeq \frac{1}{2\sqrt{3}} \frac{aV}{D} \sqrt{\frac{b}{a}}. \tag{402}$$

The latter equation is in very good agreement with numerical data of table 4. the gain can be very large if the target density a/b is low.

8.2.2 Ballistic mode in one dimension : exact result

For the ballistic mode in one dimension, the exact mean search time is calculated analytically (see section 5.3.3). It writes :

$$t_m = \frac{\tau_1 + \tau_2}{b}(\gamma_1 + \gamma_2 + \gamma_3) \tag{403}$$

$$\gamma_1 = \frac{h^2\,(h + 3L_1)}{3\alpha} \tag{404}$$

$$\gamma_2 = \frac{L_2\,(h + L_1)}{\alpha^{3/2} den} \left(g_4(-1)e^{-\frac{2a}{L_2}} - g_4(1)\right)\left(g_3(-1)e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} + g_3(1)e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}}\right) \tag{405}$$

$$den = g_1(1)e^{-2\frac{a\,(L_1 - \sqrt{\alpha})}{L_1 L_2}} + g_1(-1)e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}} + g_2(1)e^{-2\frac{aL_1 - \sqrt{\alpha}b}{L_1 L_2}} + g_2(-1)e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} \tag{406}$$

$$\gamma_3 = -\frac{L_2}{4\alpha^{3/2}} \frac{num_1 num_2}{den_1 den_2} \tag{407}$$

$$num_1 = f_1(1) + f_1(-1)e^{-\frac{2a}{L_2}} + \sigma_1 + \sigma_2 + \sigma_3 + \sigma_4 \tag{408}$$

$$f_1(\epsilon) = 2\left(\alpha g_4(\epsilon)\,(h + L_1) + L_2^4\,(\epsilon L_2 - L_1)\right) \tag{409}$$

$$\sigma_1 = \left(f_2(-1) + f_4(1,1) + f_3(1)\right) e^{\frac{\sqrt{\alpha}2h}{L_1 L_2}} \tag{410}$$

$$\sigma_2 = \left(f_2(1) + f_4(1,-1) + f_3(-1)\right) e^{-\frac{\sqrt{\alpha}2h}{L_1 L_2}} \tag{411}$$

$$\sigma_3 = \left(f_4(-1,1) + f_5(1) + f_6(1)\right) e^{2\frac{-aL_1+\sqrt{\alpha}h}{L_1 L_2}} \tag{412}$$

$$\sigma_4 = \left(f_4(-1,-1) + f_5(-1) + f_6(-1)\right) e^{-2\frac{aL_1+\sqrt{\alpha}h}{L_1 L_2}} \tag{413}$$

$$f_2(\epsilon) = \left(\sqrt{\alpha} + \epsilon L_2\right) L_2 \left(L_1 - L_2\right) g_3(\epsilon) \tag{414}$$

$$f_3(\epsilon) = -L_2^2 L_1 \sqrt{\alpha} \left(h + L_1\right) \left(\sqrt{\alpha} + \epsilon L_2 + \epsilon L_1\right) \tag{415}$$

$$f_4(\epsilon_1, \epsilon_2) = h\alpha \left(h + L_1\right) \left(\left(2L_2 + \epsilon_1 L_1\right) \left(\epsilon_2 \sqrt{\alpha} + L_2\right) + L_1^2\right) \tag{416}$$

$$f_5(\epsilon) = -\epsilon h \sqrt{\alpha} L_2 \left(L_1 + L_2\right) \left(2 \left(\epsilon\sqrt{\alpha} + L_2\right) L_2 + L_1^2\right) \tag{417}$$

$$f_6(\epsilon) = L_2^2 L_1 \left(L_2 \left(\epsilon\sqrt{\alpha} + L_2\right) \left(L_1 + L_2\right) - \sqrt{\alpha} \left(h + L_1\right) \left(\sqrt{\alpha} + \epsilon L_2 - \epsilon L_1\right)\right) \tag{418}$$

$$num_2 = \varsigma_1 + \varsigma_2 - g_4(1) e^{\frac{2a}{L_2}} \left(f_7(1) + f_8(1)\right) - g_4(-1) e^{-\frac{2a}{L_2}} \left(f_7(-1) + f_8(-1)\right) \tag{419}$$

$$\varsigma_1 = 2\sqrt{\alpha} \left(\left(L_2^2 - h^2\right) \alpha - L_1^3 h\right) \left(e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} + e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}}\right) \tag{420}$$

$$\varsigma_2 = 2L_2 \left(h \left(h + L_1\right) \alpha - L_2^4\right) \left(e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} - e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}}\right) \tag{421}$$

$$f_7(\epsilon) = \left(\alpha + \left(L_1 + \epsilon L_2\right) h\right) \sqrt{\alpha} \left(e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} + e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}}\right) \tag{422}$$

$$f_8(\epsilon) = \left(\epsilon h \alpha + L_2^3 + \epsilon L_1^3\right) \left(-e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} + e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}}\right) \tag{423}$$

$$den_1 = \sqrt{\alpha} \left(\xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 + \xi_6 + \xi_7\right) \tag{424}$$

$$\xi_1 = 2L_1 \left(h + L_1\right) \alpha \tag{425}$$

$$\xi_2 = L_2 \sqrt{\alpha} \left(\left(\alpha + L_2^2\right) \sinh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right) + 2L_2 \sqrt{\alpha} \cosh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right)\right) \tag{426}$$

$$\xi_3 = L_1 L_2 \left(\alpha \sinh\left(\frac{2a}{L_2}\right) - 2L_1 L_2 \cosh\left(\frac{2a}{L_2}\right)\right) \tag{427}$$

$$\xi_4 = -L_2 \sqrt{\alpha} \left(\alpha + 2L_1 h + L_2^2\right) \cosh\left(\frac{2a}{L_2}\right) \sinh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right) \tag{428}$$

$$\xi_5 = -2 \left(L_1 \left(h + L_1\right) \alpha + L_2^4\right) \cosh\left(\frac{2a}{L_2}\right) \cosh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right) \tag{429}$$

$$\xi_6 = -L_2 \alpha \left(2h + L_1\right) \sinh\left(\frac{2a}{L_2}\right) \cosh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right) \tag{430}$$

$$\xi_7 = -\sqrt{\alpha} \left(\left(2h + L_1\right) \alpha + L_1^3\right) \sinh\left(\frac{2a}{L_2}\right) \sinh\left(\frac{2h\sqrt{\alpha}}{L_1 L_2}\right) \tag{431}$$

$$den_2 = g_1(1) e^{-2\frac{a\left(L_1 - \sqrt{\alpha}\right)}{L_1 L_2}} + g_1(-1) e^{2\frac{\sqrt{\alpha}b}{L_1 L_2}} + g_2(1) e^{-2\frac{aL_1 - \sqrt{\alpha}b}{L_1 L_2}} + g_2(-1) e^{2\frac{\sqrt{\alpha}a}{L_1 L_2}} \tag{432}$$

$$g_1(\epsilon) = L_2 \left(L_1 + L_2\right) \left(\sqrt{\alpha} - \epsilon L_2\right) \tag{433}$$

$$g_2(\epsilon) = \sqrt{\alpha}\,(2h + L_1)\left(-\epsilon\sqrt{\alpha} - L_2 + \epsilon L_1\right) + \epsilon\alpha^{3/2} + L_2^3 - \epsilon L_1^3 \tag{434}$$

$$g_3(\epsilon) = h\sqrt{\alpha}\left(-\epsilon\sqrt{\alpha} - L_2\right) + \epsilon 2\, L_2^2 L_1 \tag{435}$$

$$g_4(\epsilon) = \left((\epsilon L_2 - L_1)\, h + L_2^2\right) \tag{436}$$

$$h = b - a \tag{437}$$

$$\alpha = L_1^2 + L_2^2 \tag{438}$$

$$L_1 = v_l \tau_1 \tag{439}$$

$$L_2 = V\tau_2. \tag{440}$$

This result has been checked by numerical simulations and by comparison with known limits.

8.2.3 Static mode in 3 dimensions : more comparisons between the analytical expressions and the simulations

Here we further compare the analytical expressions and the simulations for the static mode in 3 dimensions (see section 5.5.1).

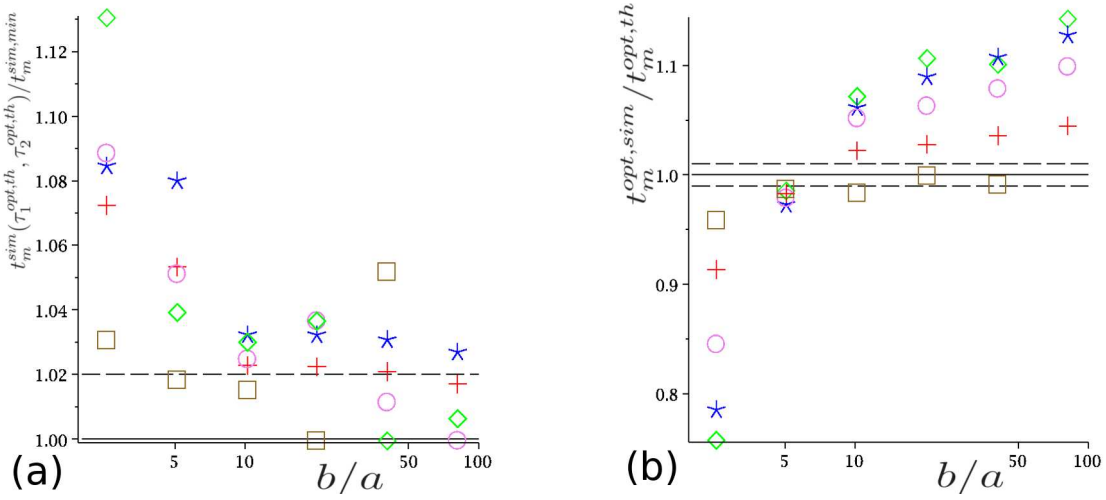


Figure 123: Static mode in 3 dimensions. Study of the minimum : its location in the τ_1, τ_2 space (a), and its value (b). *sim* means values obtained through numerical simulations, *th* means analytical values. Value expected if there was a prefect agreement between theory and simulations (black line), and values taking into account the simulations noise (dotted black lines) (we performed 10 000 walks for each point). $a = 0.01$ (□), $a = 0.1$ (+), $a = 1$ (○), $a = 10$ (⋆), $a = 100$ (◇). $V = 1$, $k = 1$.

The numerical study of the minimum mean search time (see figure 123) shows that the analytical values give the good position of the minimum in τ_1 and τ_2 as soon as b/a is not too small. However, the value of the minimum is underestimated by about 10%.

8.2.4 Diffusive mode in 3 dimensions

Full analytical approached expression of t_m (see section 5.5.2)

For the diffusive mode in three dimensions, the approached analytical expression of the mean detection time is :

$$t_m = \frac{1}{b^3\alpha^4 dpD}\left(X + Y + Z\right), \tag{441}$$

with :

$$X = \frac{\left(\tau_1^{-1} + \alpha^2 dp\right)\left(\frac{\alpha^2\left(b^3-a^3\right)}{a} - 3\,S\right)\left(\frac{1}{3}\frac{\left(b^3-a^3\right)\left(\alpha^2 dp - \tau_2^{-1}\right)}{a} + \frac{\tau_2^{-1}\left(\alpha\, aR+1\right)}{\alpha^2} + \frac{\alpha\, dp\left(-1+TT\right)}{{\alpha_2}^2}\right)}{\tau_1\left(\left(\tau_1^{-1} + \alpha^2 dp\right)\tau_2^{-1}\, R\alpha + \frac{\left(-\alpha^2 dp + \tau_2^{-1}\right)\tau_1^{-1}}{a} + \frac{TT\,\alpha^2 dp\left(\tau_1^{-1}+\tau_2^{-1}\right)}{a}\right)} \tag{442}$$

$$Y = 3\frac{\tau_1^{-1}\, aS}{\alpha^2} \tag{443}$$

$$Z = -\frac{1}{15}, \frac{\left(-b+a\right)^3\alpha^2\left(a^3 + 3\,ba^2 + 6\,b^2a + 5\,b^3\right)\left(\tau_1^{-1} + \alpha^2 dp\right)}{a} \tag{444}$$

$$\alpha = \sqrt{(\tau_1 D)^{-1} + (\tau_2 D_2)^{-1}} \tag{445}$$

$$D_2 = \frac{1}{3}V^2\tau_2 \tag{446}$$

$$dp = \frac{DD_2}{D - D_2} \tag{447}$$

$$\alpha_2 = (\tau_2 D_2)^{-1} \tag{448}$$

$$R = \frac{\alpha\, b\tanh\left(\alpha\ (b-a)\right) - 1}{\alpha\, b - \tanh\left(\alpha\ (b-a)\right)} \tag{449}$$

$$S = \frac{\left(\alpha^2 ba - 1\right)\tanh\left(\alpha\ (b-a)\right) + \alpha\ (b-a)}{\alpha\, b - \tanh\left(\alpha\ (b-a)\right)} \tag{450}$$

$$TT = \frac{\alpha_2\, a}{\tanh\left(\alpha_2\, a\right)}. \tag{451}$$

Dependence of t_m with τ_1

The mean detection time is very weakly dependent on τ_1 as long as $\tau_1 < 6D/V^2$ (see figure 124).

t_m in the regime of diffusion alone (see section 5.5.2)

For a diffusive random walk starting from $r = r_0$ in a sphere with reflective boundaries at $r = b$ and absorbing boundaries at $r = a$, the mean time of absorption $t(r_0)$ is solution of the following equation :

$$D_{\text{eff}}\frac{1}{r_0^2}\left(\frac{d}{dr_0}\left(r_0^2\frac{dt(r_0)}{dr_0}\right)\right) = -1. \tag{452}$$

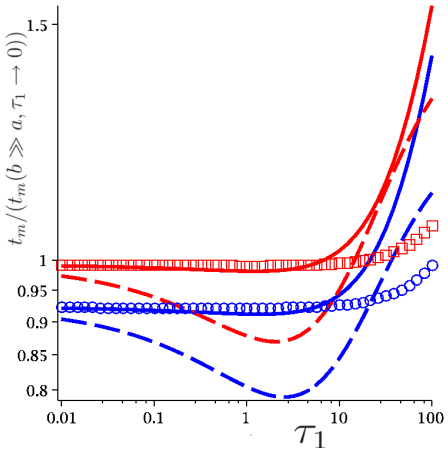


Figure 124: Diffusive mode in 3 dimension. t_m from (441), $t_m(b \gg a, \tau_1 \to 0)$ from (282). $\tau_2 = \tau_2^{opt,th}$ (283), $D = 1$, $V = 1$, $a = 10$ (dotted lines), $a = 100$ (lines), $a = 1000$ (symbols), $b/a = 10$ (blue,circles), $b/a = 100$ (red, squares).

With the boundary conditions, the solution is :

$$t(r_0) = \frac{1}{6D_{\text{eff}}} \left(\frac{2b^3}{a} + a^2 - r_0^2 - \frac{2b^3}{r_0} \right). \tag{453}$$

As the searcher starts from a random point of the sphere, we average on r_0 :

$$t_{\text{diff}} = \frac{1}{15Dab^3} \left(5b^3a^3 + 5b^6 - 9b^5a - a^6 \right), \tag{454}$$

which in the limit $b/a \gg 1$ simplifies :

$$t_{\text{diff}} = \frac{b^3}{3Da}. \tag{455}$$

Criterion for intermittence : additional figure (see section 5.5.2)

The figure 125 shows the dependence of t_m^{opt} with a. In the regime $a \lesssim D/V$, $t_m^{opt} \propto a^{-1}$, whereas in the regime $a \gg D/V$, $t_m^{opt} \propto a^{-2}$.

8.2.5 Ballistic mode in 3 dimensions

Without intermittence (see section 5.5.3)

In the regime without intermittence, τ_1 is not necessarily equal to 0. We calculate t_m in two limits : τ_1 small or τ_1 large.

Limit $\tau_2 \to 0$, $v_l\tau_1 \leq a$

In the limit $v_l\tau_1 \leq a$, we can consider the phase 1 as diffusive, with :

$$D = \frac{1}{3} v_l^2 \tau_1. \tag{456}$$

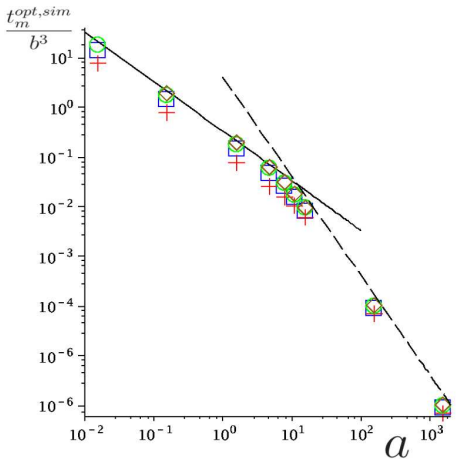


Figure 125: Diffusive mode in 3 dimensions. Simulations : $b/a = 2.5$ (+), $b/a = 5$ (□), $b/a = 10$ (○), $b/a = 20$ (◇). Analytical expressions in the low target density approximation ($b/a \gg 1$) : $\tau_1 = 0$ (282) (with $\tau_2 = \tau_2^{opt,th}$ (283)) (dotted line), diffusion alone (288) (continuous line). $V = 1$, $D = 1$.

We use the approached expression of t_m obtained in the diffusive mode (288) with this effective diffusive coefficient, leading to :

$$t_m = \frac{1}{5v_l^2\tau_1 ab^3}\left(5b^3a^3 + 5b^6 - 9b^5a - a^6\right), \tag{457}$$

and in the limit $b \gg a$:

$$t_m = \frac{b^3}{v_l^2\tau_1 a}. \tag{458}$$

limit $\tau_2 \to 0$, $\tau_1 \to \infty$

We name V_{ol} the volume of the sphere. $g(t)$ is the volume already explored by the searcher at a time t. The volume explored during dt is $\pi v_l a^2 dt$. If we assume that the probability of encountering unexplored space is uniform, which is wrong at short times but close to reality at long times, the volume explored for the first time after t and before $t+dt$ is $\frac{V_{ol}-g(t)}{V_{ol}}\pi v_l a^2 dt$. Then with this hypothesis, $g(t)$ is solution of :

$$g(t) = \int_0^t \frac{V_{ol} - g(u)}{V_{ol}} \pi v_l a^2 du. \tag{459}$$

This equation can be simplified taking a renormalized time $r = \frac{\pi v_l a^2}{V_{ol}} t$, and $f = g/V_{ol}$:

$$f(r) = \int_0^r (1 - f(w))dw. \tag{460}$$

Then, as $f(0) = 0$ (nothing has been explored at time 0), $f(r) = 1 - e^{-r}$. The probability of encountering the target at time t during dt (and not before) is the volume explored for the first time at t divided by the whole volume V_{ol} in the mean-field approximation. Then the probability $p(r)$ that the target is not yet found at time r is solution of :

$$\frac{dp}{dr} = -(1 - f(u)). \tag{461}$$

Since $p(0) = 1$, the result is $p(r) = e^{-r}$. Then the mean detection time of the target is 1 in renormalized time, which in real time means :

$$t_{bal} = \frac{4b^3}{3a^2 v_l}. \tag{462}$$

Numerical study

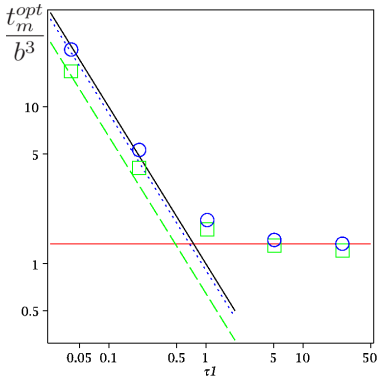


Figure 126: Ballistic mode in 3 dimensions. Regime without intermittence ($\tau_2 = 0$). $\ln(t_m/b^3)$ as a function of $\ln(\tau_1)$, simulations for $b = 5$ (□) and $b = 20$ (○). Ballistic limit ($\tau_1 \to \infty$) (no intermittence) (290) (red horizontal line), diffusive limit ($v\tau_1 < a$) (457) with $b = 5$ (green dashed line), $b = 20$ (blue dotted line), $b \gg a$ limit (458) (black line). $a = 1$, $v_l = 1$.

These expressions give a very good approximation of the values obtained through simulations (see figures 87 and 126). In the regime without intermittence, t_m is minimized for $\tau_1 \to \infty$.

Numerical v_l^c (see section 5.5.3)

In simulations (see figure 127), when b is small, v_l^c decreases, but stabilizes for larger b, which is coherent with the fact that this value is obtained through a development for b/a large. The value of v_l for large b is different (even if close) to the expected value. The main explanation of this discrepancy is that in the regime of intermittence, the approached value of t_m is about 20% away from the value obtained through simulations.

8.3 Complements to the extensions and perspectives

8.3.1 Appendix : span of a phase of 1D diffusion

Here, we complete section 6.1.1. We calculate the span of a phase of 1D diffusion.

Starting from x_0 at $t = 0$ with a coefficient of diffusion D in one dimension, the probability of being in x at t is :

$$f(x_0, x, t) = \frac{1}{2\sqrt{\pi D t}} \exp\left(\frac{-(x - x_0)^2}{4Dt}\right). \tag{463}$$

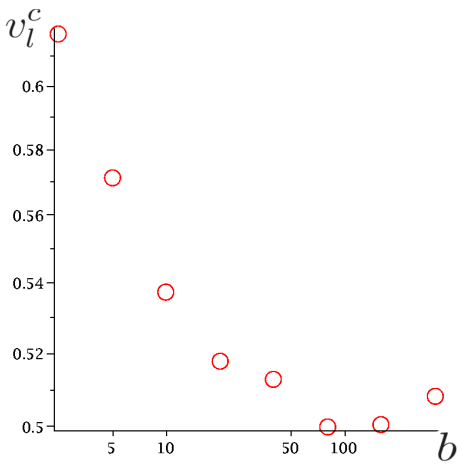


Figure 127: Ballistic mode in 3 dimensions. v_l^c as a function of $\ln(b)$, obtained by simulations. $a = 1$, $V = 1$.

Now, adding absorbing boundaries at $-a$ and b, the probability of being at x at t starting from $x = 0$ at $t = 0$ is :

$$g(x,t) = \sum_{n=-\infty}^{+\infty} g_n(x,t), \text{ with} \tag{464}$$

$$g_n(x,t) = f(2n(a+b), x, t) - f(2n(a+b) + 2b, x, t). \tag{465}$$

The survival probability at t is :

$$S(t) = \int_{-a}^{b} g(x,t)dx. \tag{466}$$

If a walk is absorbed for $(-a, b+db)$ but not absorbed for $(-a, b)$, it means that the maximum of the walk is at b. Consequently, the probability that the minimum of the walk is at $-a$ and the maximum at b is :

$$p_m(t,a,b) = \frac{\partial^2 S}{\partial a \partial b}. \tag{467}$$

If the span of the walk is L, $b = L - a$ and $a \in [0, L]$. The probability that the span of a walk is L is :

$$P(L,t) = \int_0^L p_m(t, a, L-a)da. \tag{468}$$

We decompose $P(L,t)$ in several components according to (464) :

$$P(L,t) = \sum_{n=-\infty}^{\infty} P_n(L,t), \text{ with} \tag{469}$$

$$P_0(L,t) = \left(-1 + \mathrm{e}^{\frac{3}{4}\frac{L^2}{Dt}}\right) \mathrm{e}^{-\frac{L^2}{Dt}} \frac{1}{\sqrt{\pi D t}}, \tag{470}$$

$$P_{-1}(L,t) = \left(-5 + 3\,\mathrm{e}^{-\frac{5}{4}\frac{L^2}{D\,t}} + 2\,\mathrm{e}^{\frac{3}{4}\frac{L^2}{D\,t}}\right)\mathrm{e}^{-\frac{L^2}{D\,t}}\frac{1}{\sqrt{\pi D t}} \text{ and} \tag{471}$$

$$P_n(L,t) = \frac{\mathrm{e}^{-\frac{n^2 L^2}{D\,t}}}{\sqrt{\pi\,D\,t}}\left(\left(5\,n + 6\,n^2 + 1\right)\mathrm{e}^{\frac{1}{4}\frac{-(1+4\,n)L^2}{D\,t}} - 6\,n^2 - n + \left(2\,n^2 - n\right)\mathrm{e}^{\frac{1}{4}\frac{L^2(4\,n-1)}{D\,t}} - \left(3\,n + 2\,n^2 + 1\right)\mathrm{e}^{-\frac{L^2(1+2\,n)}{D\,t}}\right). \tag{472}$$

We check that the distribution is normalized, i.e :

$$\int_0^\infty P(L,t)dL = 1. \tag{473}$$

The terms $n = 0$ and $n = -1$ give :

$$\int_0^\infty P_0(L,t)dL = 0.5, \tag{474}$$

$$\int_0^\infty P_{-1}(L,t)dL = 0.5. \tag{475}$$

$\forall n$ different from 0 and -1 :

$$\int_0^\infty P_n(L,t)dL = 0. \tag{476}$$

Therefore, the distribution is indeed normalized. After this verification, we calculate the mean span :

$$\langle L\rangle = \int_0^\infty LP(L,t)dL. \tag{477}$$

The first terms have to be calculated separately :

$$\int_0^\infty LP_0(L,t)dL = \frac{3}{2}\sqrt{\frac{Dt}{\pi}}, \tag{478}$$

$$\int_0^\infty LP_1(L,t)dL = \frac{5}{12}\sqrt{\frac{Dt}{\pi}}, \tag{479}$$

$$\int_0^\infty LP_{-1}(L,t)dL = \frac{13}{6}\sqrt{\frac{Dt}{\pi}}, \tag{480}$$

$$\int_0^\infty LP_2(L,t)dL = \frac{1}{20}\sqrt{\frac{Dt}{\pi}}, \tag{481}$$

$$\int_0^\infty LP_{-2}(L,t)dL = -\frac{7}{60}\sqrt{\frac{Dt}{\pi}}, \tag{482}$$

$$\int_0^\infty LP_3(L,t)dL = \frac{13}{840}\sqrt{\frac{Dt}{\pi}} \text{ and} \tag{483}$$

$$\int_0^\infty LP_3(L,t)dL = -\frac{11}{420}\sqrt{\frac{Dt}{\pi}}. \tag{484}$$

For the other n :

$$\int_0^\infty LP_n(L,t)dL = \sqrt{\frac{Dt}{\pi}}\frac{4n+1}{2n(1+n)(2n+1)(2n-1)}. \tag{485}$$

Finally :

$$\begin{aligned}\langle L\rangle =&\sqrt{\frac{Dt}{\pi}}\left(\frac{3}{2}+\frac{13}{6}+\frac{5}{12}+\frac{1}{20}-\frac{7}{60}+\frac{13}{840}-\frac{11}{420}-\sum_{n=4}^{\infty}\frac{3}{(4n^2-1)(n^2-1)}\right)\\ &\sqrt{\frac{Dt}{\pi}}\left(\frac{3}{2}+\frac{13}{6}+\frac{5}{12}+\frac{1}{20}-\frac{7}{60}+\frac{13}{840}-\frac{11}{420}-\frac{1}{168}\right)=4\sqrt{\frac{Dt}{\pi}}.\end{aligned} \tag{486}$$

It is different from the result page 202 of Weiss [1994] $\left(\langle L\rangle=\sqrt{\frac{8Dt}{\pi}}\right)$, but it is what we expect.

The mean span of a diffusive phase which duration is exponentially distributed (of mean τ_1) is :

$$L_1=\int_0^{\infty}\frac{1}{\tau_1}\exp\left(\frac{-t}{\tau_1}\right)4\sqrt{\frac{Dt}{\pi}}dt=2\sqrt{D\tau_1}. \tag{487}$$

8.3.2 Appendix : details of the simulation of intermittent model with Lévy-distributed relocations

We have performed simulations (see section 6.3.2) using variable diffusion steps for simulating the phase 1 [Berezhkovskii et al., 1998], and the Chambers-Mallows-Stuck algorithm for the Lévy distribution of phase 2 [Chambers et al., 1976]. We have rejected flights larger than L when simulating a cut-off. There are two main issues when trying to determine the minimum : there could be too much noise, and the minimum could be missed.

On the one hand, the mean search time is obtained by averaging over a finite number of simulations. Consequently, there is noise in the values obtained, especially when $\alpha\to 1$, since the distribution of the ballistic phase duration becomes very large. The mean search times obtained cannot be compared if the noise is larger than the differences seen between these values. To rule out this possibility, we have averaged the search time over a large number of iterations (the results shown are obtained for 10^7 iterations, and are within 0.5% of the results with 10^6 iterations).

On the other hand, we estimate the function for a finite number of α, τ_1 and σ. The real minimum could be missed, and comparing a value close to the real minimum, and another value, thought to be minimal but in fact non-representative of the minimum, could lead to wrong conclusions. The values of α that we have used are represented in figures 98 and 99 : it can be seen that the minimum is flat enough in α so that there is no risk to have missed the minimum to the point to make the conclusions uncertain. We have chosen τ_1 and σ by minimizing numerically the approached analytical expression for the mean search time without cut-off (341). We have checked that taking these theoretical optimal values divided or multiplied by 2 leads to a larger mean search time. Moreover, a shift in the minimum is more likely for the simulations with a cut-off, because the function (341) used to choose where to calculate t_m is an approximation of the function without cut-off. It means that if imprecisions for not taking the real minimum exist, the mean search time with a cut-off is more likely to be overestimated than the mean search time without cut-off.

8.3.3 Appendix : static mode in one dimension with perfect correlation

We give here the details of the optimization of the mean search time for the static mode in one dimension with perfect correlation (see section 6.2.3).

Another way to write t_m

t_m is minimized for $\tau_2 \to 0$ and $\tau_1 \to 0$. We can introduce α and β such that : $\tau_1 = \alpha\tau_2^\beta$. It can be seen easily that if $\beta \neq 1$, $t_m \to \infty$ when $\tau_2 \to 0$. Thus $\tau_1 = \alpha\tau_2$. What is the optimal α? Taking $b \gg a$, $\tau_1 = \alpha\tau_2$ and $\tau_2 \to 0$ in (317), the mean search time can be greatly simplified :

$$kt_m = \frac{w}{d}(1+\alpha)\coth(w\alpha) \tag{488}$$

with $w = \frac{ak}{V}$, $d = a/b$. The value of α minimizing this expression is z/w, with z solution of $4ze^{2z} - e^{4z} + 1 + 4we^{2z} = 0$. This approximate works, except for very large w (of the same order or larger than b/a) (see figure 128).

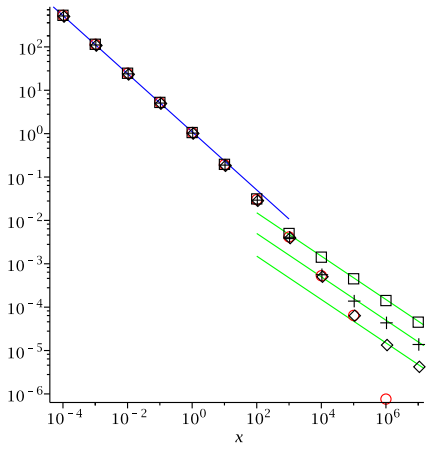


Figure 128: Minimization of t_m when $p = 1$: α^{opt} as a function of w, optimization of the exact expression (black points), for $b = 100$ (□), $b = 10^3$ (+), $b = 10^4$ (◇) (with $\tau_2 \to 0$), z/w (○). For w small, slope seems to be $-2/3$ (blue line), for w large, slope seems to be $-1/2$ (green lines). ($a = 1$, $V = 1$).

w small

When w is small, t_m can be simplified, because the argument of the coth can be considered small :

$$kt_m = \frac{w}{d}\left(\frac{1}{\alpha w} + \frac{1}{3}\alpha w\right). \tag{489}$$

For $dt_m/d\alpha = 0$, α should be :

$$\alpha = \frac{1}{6}\left(u + \frac{1}{u} - 1\right), \tag{490}$$

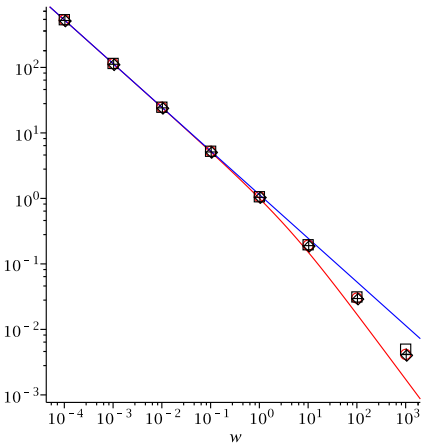


Figure 129: Minimization of t_m when $p = 1$, pour w small : α^{opt} as a function of w, optimization of the exact expression (black symbols), for $b = 100$ ($\square$), $b = 10^3$ ($+$), $b = 10^4$ ($\diamond$) (with $\tau_2 \to 0$), z/w ($\circ$). Theoretical α for small w (490) (red line), simplified form (492) ($a = 1$, $V = 1$).

with :

$$u = \left(\frac{9}{w} + \sqrt{\left(\frac{9}{w}\right)^2 - 1} \right). \tag{491}$$

w being small, this expression can be simplified :

$$\alpha = \left(\frac{3}{2w^2} \right)^{\frac{1}{3}}. \tag{492}$$

This expression works well (see figure 129).

w intermediate

In the regime $b/a >> w >> 1$, we assume that the argument of the coth is large, thus coth can be developed :

$$\coth(x) = \frac{1 + e^{-2x}}{1 - e^{-2x}} \simeq 1 + 2e^{-2x}. \tag{493}$$

The expression to minimize is proportional to :

$$t_m \propto (1 + \alpha)\left(1 + 2e^{-2w\alpha}\right) \simeq 1 + \alpha + 2e^{-2w\alpha}. \tag{494}$$

$dt_m/d\alpha = 0$ requires :

$$\alpha = \frac{\ln(4w)}{2w}. \tag{495}$$

This expression gives a good approximation of the intermediate regime (see figure 130).

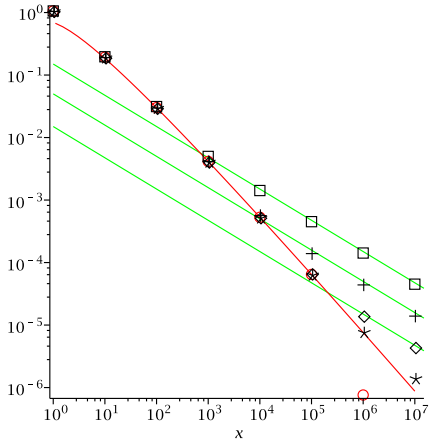


Figure 130: Minimization of t_m when $p = 1$, for intermediate w : α^{opt} as a function of w, optimization of the exact expression (black symbols), for $b = 100$ (□), $b = 10^3$ (+), $b = 10^4$ (◇), $b = 10^5$ (⋆) (with $\tau_2 \to 0$), z/w (○). Theoretical α for intermediate w (495) (red line) ($a = 1$, $V = 1$).

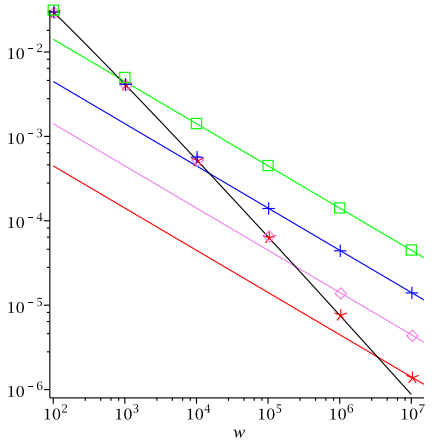


Figure 131: Minimization of t_m when $p = 1$, for large w : α^{opt} as a function of w, optimization of the exact expression (symbols), for $b = 100$ (□), $b = 10^3$ (+), $b = 10^4$ (◇), $b = 10^5$ (⋆) (with $\tau_2 \to 0$). theoretical α^{opt} for intermediate w (495) (black line), theoretical α^{opt} for large w (497) (colored lines). ($a = 1$, $V = 1$).

w large

We start from the whole expression of t_m, with $\tau_1 = \alpha\tau_2$ and $\tau_2 \to 0$. We assume that the argument of the coth is very large, thus $coth \simeq 1$. It leads to :

$$kt_m \simeq (\alpha + 1)\left(\frac{2}{\alpha} - \frac{a}{b\alpha} + \frac{(b-a)^2 k}{bv}\right). \tag{496}$$

$dt_m/d\alpha = 0$ for :

$$\alpha = \frac{\sqrt{kv(2b-a)}}{k(b-a)} \simeq \sqrt{\frac{2v}{kb}} = \sqrt{\frac{2a}{wb}}. \tag{497}$$

This expression gives a good approximation of α^{opt} when w is large (see figure 131).

8.4 An example of the importance of transport in biology

8.4.1 Introduction

The main subject of this thesis is search strategies. A search strategy consists in finding how to move to find a target as fast as possible. Transport is important to biology in a more general way.

We present here an example of the importance of transport in biology : the dynamics of the distribution of a membrane receptor in HEK cells and in neurons. We have worked on the interpretation of data of Zsolt Lenkei and Anne Simon, team "Dynamique des Récepteurs Neuronaux" of the laboratory "Neurobiologie et Diversité Cellulaire" of the ESPCI, Paris. An article is submitted [Simon et al., 2009].

They work on a membrane protein (CB1R cannabinoid type-1 receptors : a target for marijuana ; see Marzo et al. [2004] for a review on the endocannabinoid system). At the membrane, this receptor is either active or inactive. It has a constitutive activity, meaning that a wild type receptor (denoted WT afterwards), even without any ligand, switches between these two states. Ligands shift the equilibrium to more activity (WIN for example), or to less activity (AM for example). There are also mutant receptors that have more (these mutants are denoted I in the following) or less (these mutants are denoted A in the following) intrinsic activity. But before being on the membrane, receptors have to be transported from inside the cell, where they are produced, to their destination. After some time at the membrane, they are endocytosed to be recycled or degraded [Koenig and Edwardson, 1997]. Consequently, these receptors are located either at the membrane, or in vesicles inside the cell. They can be visualized with fluorescence. However, the techniques used are not the same for labeling all the receptors or only the receptors at the surface of the cell. This explains why most measurements are values relative to the WT without ligand, and not absolute values. Different phenotypes are observed. The distribution of the receptors between the inside and the surface of the cell is variable (see section 8.4.2), and in polarized cells the geography of receptors is also variable (see section 8.4.3). The main idea is that differences in the activity of the receptor could be responsible for these changes. Endocytosis could happen only to receptors in the active state, and not in the inactive state. We shall see in the following if data and model support this hypothesis.

8.4.2 HEK cells

Model

CB1R receptors are first studied in HEK cells (HEK means "human embryonic kidney cells", this cell lineage is often used because these cells are easy to modify). These cells can be more or less considered as spherical. Receptors are either inside the cell (state C in figure 132), or at its membrane (states M_i and M_a in figure 132). At the membrane, the receptor is either active (M_a) or inactive (M_i). We consider that the kinetic constants between these states are fast enough to always assume

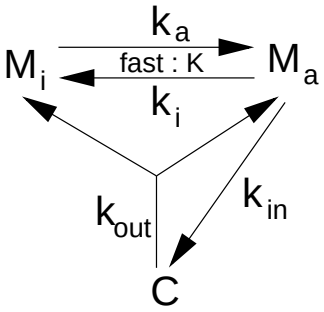


Figure 132: Model of receptor dynamics for HEK cells.

that equilibrium between M_i and M_a is immediate, with an equilibrium constant K ($K = M_a/M_i$, X representing implicitly the quantity of receptors in state X). The switching rate from state C to state $M_{a,i}$ is k_{out}. k_{out} is thought to be linked mainly to the timescale of receptor recycling. We assume that k_{out} and k_{in} are constant. Consequently, the only parameter of the equilibrium that will change in the experiments is K, characterizing the activity of receptors. We assume that the total number of receptors in the cell is fixed : $T = M_a + M_i + C$ is a constant

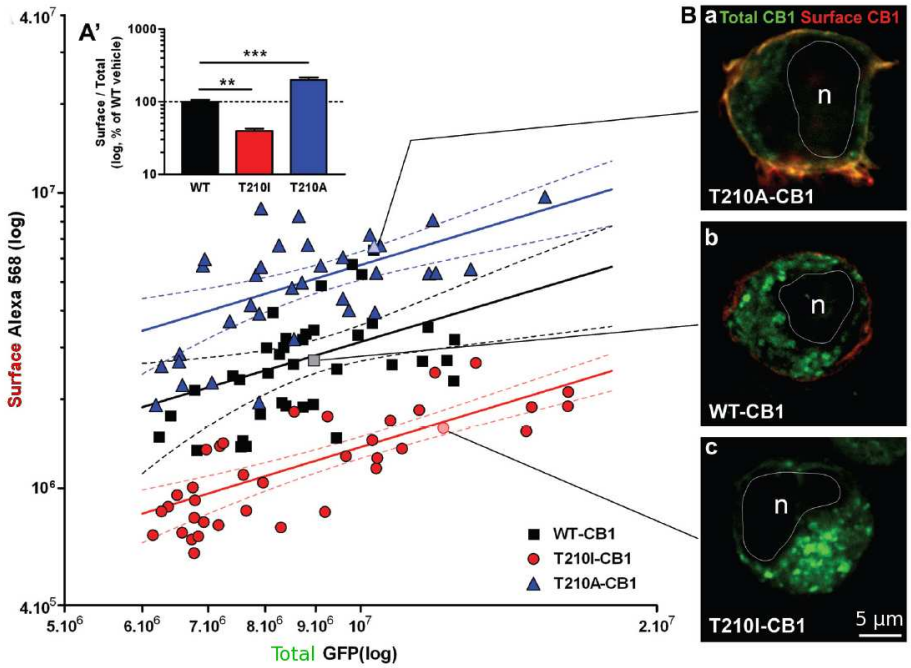


Figure 133: Fluorescence measurements for HEK cells. Here, receptors are fluorescent (green), and the green fluorescence is measured to estimate the quantity of receptors in the whole cell. At the surface, receptors are targeted by antibodies that are eventually labeled with fluorescence (red), and the red fluorescence is measured to estimate the quantity of receptors at the membrane. Here, results are shown without ligand, for the WT, A and I receptors. Inset A' shows the average ratio of fluorescence, compared to the WT. Inset B shows how a cell typically looks like.

As can be seen figure 133, the distribution depends on whether the receptor is mutant or not. A similar figure could have been obtained with the WT receptor in a solution with ligands that change activity. And reversely, if the ligand with the opposite effect on activity is added to the environment of a cell with mutant proteins (for example if AM is added to the I mutant), a normal phenotype can be recovered for adequate concentrations, supporting the hypothesis that activity (characterized by the parameter K) does control the distribution of the receptors.

Endocytosis and activity

	Endocytosis	Activity
I	135 ± 7	133 ± 8
A	59 ± 3	66 ± 20

Table 5: Antibody feeding and activity in mutants compared to the WT (the value for WT receptors is set at 100).

Activity is indirectly measured, and is supposed to be proportional to the quantity of active receptors on the membrane, *i.e.* M_a. Endocytosis is measured using antibody feeding. Antibody feeding consists in adding antibodies in the environment that bind on the receptors at the membrane. If the receptors are endocytosed, antibodies are endocytosed with them. After washing the environment, fluorescence is measured, and is proportional to the quantity of receptors endocytosed. It should then be proportional to $k_{in}M_a$. Variations of the antibody feeding and of the activity for the mutants are coherent, supporting the hypothesis of k_{in} constant (see table 5).

Mean fluorescence Ratio (MFR)

Leterrier et al. [2004] (figure 4Cb of this article) measured ratios of fluorescence between the inside of the cell and the membrane. These fluorescence measurements inside the cell and at the membrane are supposed to be comparable. Ratios are absolute, not relative.

However, a first correction, geometric, is to be made on original measurements. We make the assumption that a HEK cell is spherical. We define R as its radius, m the membrane thickness, e the thickness of the optical slice, d_m the density of receptors in the membrane, d_c the density of receptors inside the cell. The MFR is defined as the total quantity of receptors at the membrane divided by the total quantity of receptors inside the cell. Thus :

$$MFR^{real} = \frac{M}{C} = \frac{4\pi R^2 m d_m}{\frac{4}{3}\pi R^3 d_c} = \frac{3md_m}{Rd_c}. \tag{498}$$

Fluorescence is measured at the equatorial plane of the cell, leading to the measurement :

$$MFR^{measured} = \frac{2\pi Remd_m}{\pi R^2 e d_c} = \frac{2md_m}{Rd_c}. \tag{499}$$

The measurements have to be renormalized as follows :

$$MFR^{real} = 1.5 MFR^{measured}. \tag{500}$$

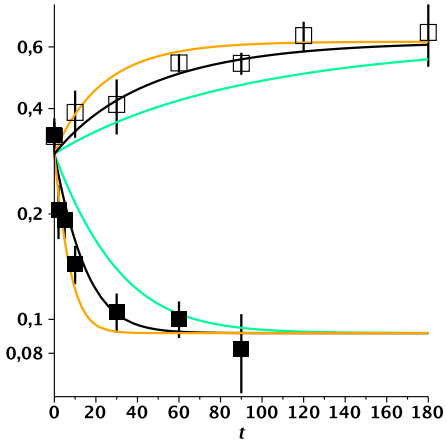


Figure 134: MFR (the fluorescence on the membrane divided by the fluorescence inside the cell) as a function of time (minutes). Ligand is added a $t = 0$. □ : ligand added is AM (at a concentration of 7 µM), which decreases activity. ■ : ligand added is WIN (at a concentration of 330 nM), which increases activity. Black lines are the fit (504) for $k_{out}^{-1} = 135$ min. Orange lines are the same curves for $k_{out}^{-1} = 77.5$ min, and green lines for $k_{out}^{-1} = 270$ min. Other parameters of the fit are initial and final values, that are taken as measured.

The change in MFR is measured in WT cells that are initially in a medium without ligand, in which a ligand is added at $t = 0$ (see figure 134). Temporal variations of the quantities of receptors follow these equations : :

$$\frac{dC}{dt} = -k_{out} C + k_{in} M_a, \tag{501}$$

$$\frac{dM}{dt} = k_{out} C - k_{in} M_a, \tag{502}$$

$$M = \left(\frac{1}{K} + 1\right) M_a. \tag{503}$$

Defining MFR_0 as the value of the MFR at $t = 0$ and MFR_∞ as the value of the MFR at $t = \infty$, the solution is :

$$MFR = \frac{MFR_\infty(1 + MFR_0) + (MFR_0 - MFR_\infty) \exp\left(-\left(1 + \frac{1}{MFR_\infty}\right) t k_{out}\right)}{1 + MFR_0 + (MFR_\infty - MFR_0) \exp\left(-\left(1 + \frac{1}{MFR_\infty}\right) t k_{out}\right)}. \tag{504}$$

Experimental data with different ligands can be fitted with a single k_{out}, supporting the hypothesis that this parameter is indeed constant (see figure 134). It gives a direct estimate of $k_{out}^{-1} \simeq 135$ min.

We can also express MFR_∞ with the parameters of our model (since when the distribution does not change anymore, $k_{out}M_a = k_{in}C$) :

$$MFR = \frac{M_i + M_a}{C} = \frac{M_a/K + M_a}{k_{out}M_a/k_{in}} = \frac{k_{out}}{k_{in}}\left(1 + \frac{1}{K}\right). \tag{505}$$

In the case of the measurements with highly concentrated WIN, the MFR does not change anymore with higher concentrations. We can assume that in this case, K is very large, consequently k_{out}/k_{in} can be estimated :

$$\frac{k_{out}}{k_{in}} \lesssim MFR_{WIN} \simeq 0.124 \tag{506}$$

As we have already an estimation of k_{out}, we can estimate $k_{in}^{-1} \simeq 16.7$ min. Values obtained for k_{in} and k_{out} seemed of the expected order of magnitude according to the biologists.

Endocytosis and relative fluorescence

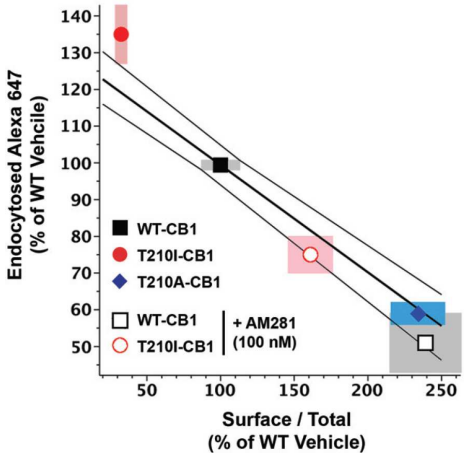


Figure 135: Endocytosis as a function of SCOP, *i.e.* the ratio between surface and total fluorescence (both are relative to the value for the WT), for different receptors and with or without ligand (symbols, with light color areas representing error bars). Black thick line is the theoretical fit (513), and the other lines are theoretical fit with error bars on the parameters used.

$SCOP$ is the fluorescence at the surface divided by the total fluorescence. It has only a meaning in comparison with the WT, because the two fluorescences measured cannot be directly compared. For different receptors and with or without ligand, SCOP is measured, along with the endocytosis. We check what relation we expect with the model :

$$SCOP = \frac{M}{T} = \frac{M_a + M_i}{M_a + M_i + C} = \frac{1 + 1/K}{1 + 1/K + k_{in}/k_{out}}, \tag{507}$$

$$1 - SCOP = \frac{k_{in}/k_{out}}{1 + 1/K + k_{in}/k_{out}}. \tag{508}$$

Endocytosis is proportional to $k_{in}M_a$. As k_{in} and T are assumed to be constant, it is equivalent to say that endocytosis is proportional to $M_a/T = 1/(1+1/K+k_{in}/k_{out})$. Thus :

$$1 - SCOP = \frac{k_{in}}{k_{out}} \frac{M_a}{T}. \tag{509}$$

It is also true for the WT :

$$1 - SCOP_{WT} = \frac{k_{in}}{k_{out}} \left(\frac{M_a}{T}\right)_{WT}. \tag{510}$$

We combine these equations :

$$\frac{M_a}{T} \Big/ \left(\frac{M_a}{T}\right)_{WT} = \frac{1 - SCOP}{1 - SCOP_{WT}} = 1 + \frac{SCOP_{WT} - SCOP}{1 - SCOP_{WT}} = 1 + \frac{1 - \frac{SCOP}{SCOP_{WT}}}{\frac{1}{SCOP_{WT}} - 1}. \tag{511}$$

As $MFR = M/C$, $SCOP = M/T$ and $T = M+C$, we have $MFR = (SCOP^{-1} - 1)^{-1}$. This leads to :

$$\frac{M_a}{T} \Big/ \left(\frac{M_a}{T}\right)_{WT} = 1 + MFR_{WT} \left(1 - \frac{SCOP}{SCOP_{WT}}\right). \tag{512}$$

Using now quantities in percentage of WT, the equation for figure 135 writes :

$$\left(\frac{M_a}{T} / \left(\frac{M_a}{T}\right)_{WT}\right)_{\%} = 100 + MFR_{WT} \left(100 - \left(\frac{SCOP}{SCOP_{WT}}\right)_{\%}\right). \tag{513}$$

We are able to fit the data with this expression (see figure 135). Note that there is no adjustable parameter.

Predictions

Now that we have gained some confidence in our model and that we have evaluated k_{out}/k_{in}, we study what the model predict for the activity (K value), and for the proportion of active receptors (M_a). As $K = M_a/M_i$, $M_a/M = M_a/(M_a+M_i) = M_a/(M_a + M_a/K) = 1/(1 + 1/K)$.

$M/T = 1/(1 + 1/MFR)$: as the MFR is measured experimentally, M/T is obtained straightforwardly. We assume that $k_{out}/k_{in} \simeq MFR_{WIN}$. This leads to $K = 1/((MFR/MFR_{WIN}) - 1)$. We can then estimate K for the different receptors (see figure 136).

Now we have obtained K, we can estimate the proportion of the receptors of the cell that are in the active state : $M_a/T = 1/(1 + 1/K + k_{in}/k_{out})$ (see figure 137). An issue is that the relative variations of M_a/T between the mutants and the WT obtained with the model are smaller than the relative variations measured with the endocytosis (supposed to be proportional to M_a/T). Error bars may be underestimated. Different experiments with different neuron batches may be hard to compare : MFR and $SCOP$ measurements are also not fully compatible. However, the relative M_a/T variations, obtained by the model or by measurements, even if out of the errors bars, are of the same order of magnitude.

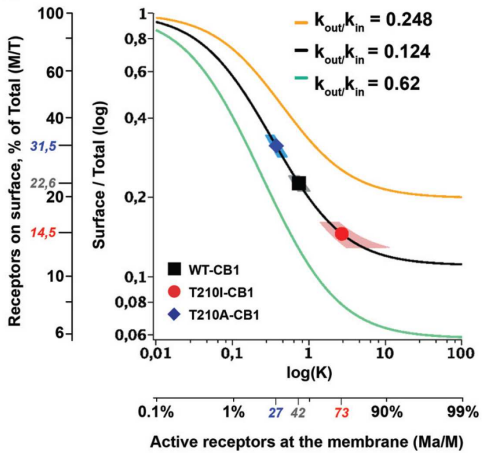


Figure 136: M/T as a function of K. Symbols have for coordinates experimental values of M/T, and K estimated with the model, with the estimation of k_{out}/k_{in} stemming from MFR_{WIN}. Lines represent the prediction from our model.

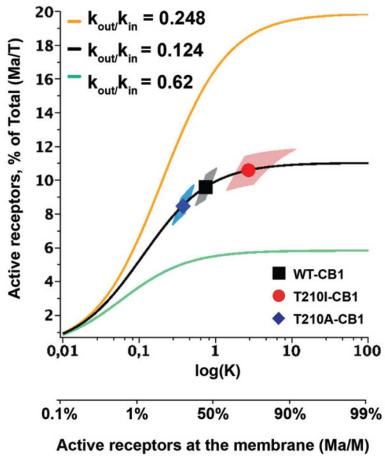


Figure 137: M_a/T as a function of K. Both are inferred via the model and the MFR measurements. Lines represent the prediction from the model.

Implications

The idea that the activity of the receptor (equilibrium between inactive and active state at the membrane) controls its distribution (inside the cell or at the membrane) is supported by the results of this simple kinetic model applied to the experimental data. These results tend to show that the proportion of receptors that are actually in the active state at the membrane only represent about 10% of all the receptors in the cell. From a design perspective, the proportion of receptors inside the cell may regulate signaling function in two opposing ways. In cells like here, with a low proportion of receptors in the active state, the initial signaling response to a ligand may not have a large amplitude, but the stock of receptors can maintain the response. Reversely, in cells with most receptors in the active state, a change in the environment can lead to a highly efficient activation of intracellular signal-

ing, but the following desensitization period limits cellular response to subsequent activation. Consequently, by regulating the availability of the receptors, biological characteristics of different cell types may have an important influence to establish the optimal trade-off level between efficiency and robustness of receptor signaling.

8.4.3 Neurons

CB1 in neurons

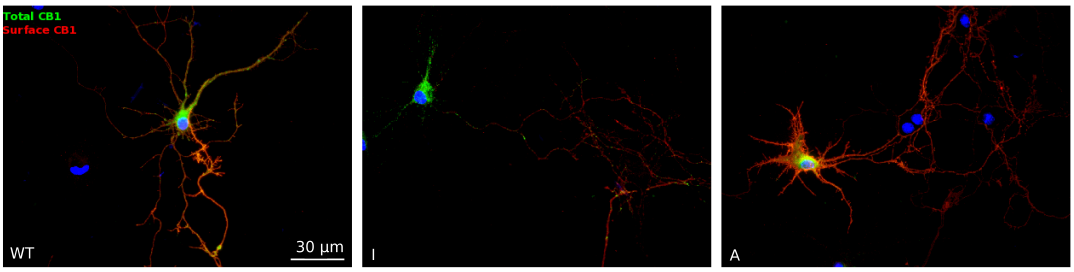


Figure 138: Distribution of receptors in neurons. Blue is a marker for the soma, receptors are green, receptors on the membrane are labeled in red.

The same receptors are studied in neurons. The distribution of receptors between the soma and the axon is very different depending on the receptor type (WT, I, A) (see figure 138). If the receptors are the A mutants, they are mainly located in the soma, whereas if they are the I mutants, they are mainly located in the axon. The WT has an intermediate distribution.

We first explain here how we imagine the "life cycle" of these receptors in a neuron. Receptors are produced inside the soma. Then they are exported to the membrane in vesicles by a first transport pathway which does not send the vesicles too far away. They arrive at the soma membrane, or at the axon membrane close to the soma. Activity will change the effective internalization rate k_{soma}, as it was the case in the HEK cells. But when the receptors being at the soma membrane are internalized, the endocytosis vesicles are exported via a second transport pathway, sending them to the axon. Once on the axon membrane, the receptors will diffuse, and the internalization rate k_{axon} will again be dependent on the activity. Once internalized, receptors are sent back to the soma to be degraded or recycled. In what follows, we shall see if the distribution of receptors along axons is compatible with this model with two transport pathways.

Parameters

Receptors are transported inside the axon starting from the soma (see figure 139). The model relies on the idea that there are two transport pathways in the axon. For each transport the probability that the receptors are released to the membrane per unit time is constant. The two transports have not the same release rate.

- Assuming that the direct transport is diffusive (with coefficient D_d), and have a release rate k_d, the mean length over which receptors are transported before

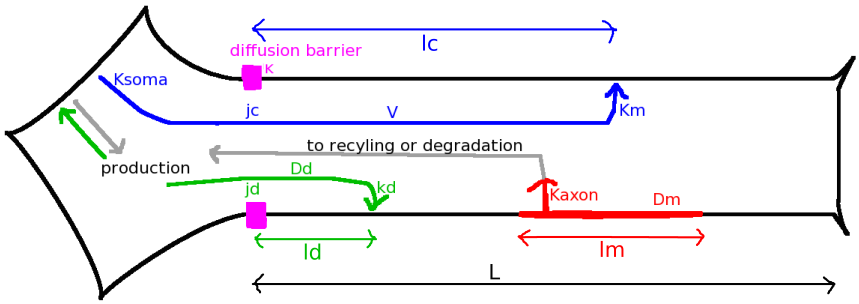


Figure 139: Model for the distribution of receptors in neurons.

being released to the membrane is $l_d = \sqrt{D_d/k_d}$. We assume that the flux j_d of receptors using this transport pathway is fixed.

- Assuming that the transport after the endocytosis at the soma membrane is ballistic (velocity v) and that the release rate is k_m, the mean length over which receptors are transported before being released to the membrane is $l_c = v/k_m$. The flux of receptors using this pathways is assumed to be proportional to the the endocytosis rate at the soma k_{soma}.

On the axon membrane, receptor diffuse with a constant D_m, and are endocytosed back with a rate of k_{axon}. It means that $l_m = \sqrt{D_m/k_{axon}}$ is the typical length on which they diffuse before "dying" (in the point of view of the model, since afterwards they are sent back to the soma, where other processes recycle or degrade them). In fact, only l_d, l_m and l_c matters. The details of the transport are not important, only the typical length over which receptors are released to the membrane matters.

The axon is assumed to be a simple segment of length L. It is a simplification, as real axons are branched. Receptors transported inside the axon are assumed to be released when they are at the end of the axon opposite from the soma ($x = L$). This boundary is reflective for the receptors diffusing at the membrane. The other boundary (separating the axon from the soma) ($x = 0$) is considered to be semi-reflective for the receptors diffusing at the membrane, with a parameter κ translating this condition ($\partial_x c = \kappa c$ at $x = 0$). Indeed, a diffusion barrier is observed experimentally at the beginning of axons [Nakada et al., 2003]. κ can be either positive or negative depending on the net flux of receptors at the membrane (to the soma or to the axon).

Equations

$P_c(x,t)$ is the concentration of receptors transported inside the axon by the ballistic pathway. $P_d(x,t)$ is the concentration of receptors transported inside the axon by the diffusive pathway. $P_m(x,t)$ is the concentration of receptors at the membrane.

$$\frac{\partial P_c}{\partial t} = -v\frac{\partial P_c}{\partial x} - k_m P_c, \tag{514}$$

$$\frac{\partial P_d}{\partial t} = D_d\frac{\partial^2 P_d}{\partial x^2} - k_d P_d, \tag{515}$$

$$\frac{\partial P_m}{\partial t} = D_m \frac{\partial^2 P_m}{\partial x^2} + k_m P_c + k_d P_d - k_e P_m. \tag{516}$$

In the stationary case, all the derivatives relative to t are equal to 0. This leads to :

$$P_c = C_1 e^{-x/l_c}. \tag{517}$$

Since $j_c = v P_c(x=0)$:

$$P_c = \frac{j_c}{v} e^{-\frac{x}{l_c}} = \frac{j_c}{l_c k_m} e^{-\frac{x}{l_c}}, \tag{518}$$

and with a similar reasoning :

$$P_d = \frac{j_d}{\sqrt{D_d k_d}} e^{-x\sqrt{\frac{k_d}{D_d}}} = \frac{j_d}{l_d k_d} e^{-\frac{x}{l_d}}. \tag{519}$$

These expressions replace P_d and P_c in the equation on P_m :

$$D_m \frac{\partial^2 P_m}{\partial x^2} + \frac{j_c}{l_c} e^{-\frac{x}{l_c}} + \frac{j_d}{l_d} e^{-\frac{x}{l_d}} - k_e P_m = 0. \tag{520}$$

The boundary conditions are :

$$\left.\frac{dP_m}{dx}\right|_{x=0} = \kappa P_m(x=0) \text{ and} \tag{521}$$

$$D_m \left.\frac{dP_m}{dx}\right|_{x=L} = v P_c(x=L) - D_d \left.\frac{dP_d}{dx}\right|_{x=L}. \tag{522}$$

Finally, P_m divided by the mean concentration writes :

$$p(x) = \frac{L}{l_m} \frac{ch\left(\frac{-x+L}{l_m}\right)\alpha - \left(e^{-\frac{x}{l_d}} l_d \hat{j}_d + e^{-\frac{x}{l_c}} l_c \hat{j}_c\right) l_m \left(sh\left(\frac{L}{l_m}\right) + ch\left(\frac{L}{l_m}\right) l_m \kappa\right) - \left(e^{-\frac{L}{l_d}} {l_d}^2 \hat{j}_d + e^{-\frac{L}{l_c}} {l_c}^2 \hat{j}_c\right)\left(ch\left(\frac{x}{l_m}\right) + l_m \kappa\, sh\left(\frac{x}{l_m}\right)\right)}{-\left(sh\left(\frac{L}{l_m}\right) + ch\left(\frac{L}{l_m}\right) l_m \kappa\right)\left({l_c}^2 \hat{j}_c + {l_d}^2 \hat{j}_d\right) + \alpha sh\left(\frac{L}{l_m}\right) + \left(e^{-\frac{L}{l_c}} {l_c}^2 \hat{j}_c + e^{-\frac{L}{l_d}} {l_d}^2 \hat{j}_d\right) l_m \kappa}, \tag{523}$$

with $\alpha = {l_m}^2\left(\hat{j}_c + \hat{j}_d + \kappa\left(l_c \hat{j}_c + l_d \hat{j}_d\right)\right)$, $\hat{j}_i = j_i/(l_m^2 - l_i^2)$.

Results

We use this expression to fit profiles of concentration in axons. These profiles are averaged over several axons, grouped into "long" and "short" axons. These profiles are for the CB1 receptor wild type (WT), and for the mutants A and I. There are additional profiles of the SST receptor, which is thought to be only transported with a pathway similar to the diffusive pathway.

Fit parameters appear at first numerous. But with a closer examination, it is not completely true.

- In fact, if you look at equation 523, $r_j = j_c/j_d$ matters, but not each flux. the WT value of r_j is a fit parameter. However, as j_d is assumed to be fixed, and as j_c is assumed to be proportional to k_{soma}, measurements of k_{soma} for the mutants compared to the WT ($k_{soma}^{WT} = 32$ (in arbitrary units), $k_{soma}^{I} = 65 \simeq 2k_{soma}^{WT}$, $k_{soma}^{A} = 17.5 \simeq 0.5 k_{soma}^{WT}$ (see figure 140)) lead to : $r_j^I \simeq 2 r_j^{WT}$, $r_j^A \simeq 0.5 r_j^{WT}$.

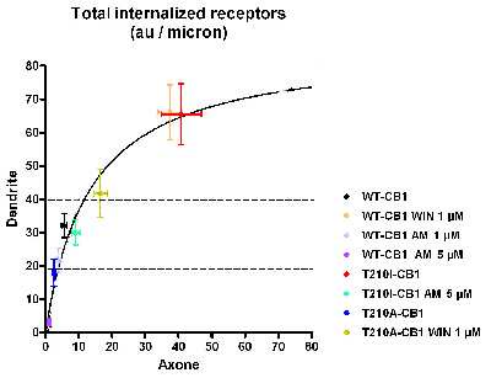


Figure 140: Internalization rates

- It is the same for l_m : l_m^{WT} is a fit parameter, however, as k_{axon} is measured for mutants relative to the WT ($k_{axon}^{WT} = 5.0$ (arbitrary units), $k_{axon}^{I} = 40.5 \simeq 8.1 k_{axon}^{WT}$, $k_{axon}^{A} = 2.3 \simeq 2.2 k_{axon}^{WT}$ (see figure 140)), we obtain $l_m^A \simeq 1.5 l_m^{WT}$ and $l_m^I \simeq 0.35 l_m^{WT}$.
- l_d does not matter as long as it is short compared to the other length scales.
- l_c is supposed to be the same for all the CB1 data, be it WT or mutant.

We can qualitatively fit the gradient of concentration of the receptors (see figure 141). However, κ allows us a lot of freedom in the adjustment. The results should not be coined as quantitative.

The distribution of the SST receptor, thought to be only transported via the "diffusive" pathway, can be roughly fitted (however here there is only one adjustable parameter) (see figure 142).

The concentration close to the soma ($x \lesssim 50\mu$m : gray area in figures 141 and 142) is not well captured by the fit. Indeed, it is the typical scale below which the diffusion barrier has a significant influence, and our model has only included it in the parameter κ at $x = 0$ [Nakada et al., 2003].

Our model shows qualitatively that the observed concentration gradients of receptors are compatible with two transport pathways, linked to endocytosis rates. However, too many fit parameters (in particular κ) remain, thus our fits are not very discriminating, and besides that, the data are very noisy.

8.4.4 Conclusion

The activity of the receptors changes the distribution of the receptors in cells. For the spherical HEK cells, more activity means that the receptors will be more endocytosed, and thus there will be more receptors inside the cells than at their membrane. The presence (or the absence) of such a "reservoir" inside the cell could change the robustness of the response to a signal. For the polarized neurons, the concentration of receptors varies on the axon membrane with the distance from the soma. Depending on the activity of these receptors, this distribution is radically changed. It matters, because it would change the connections with other neurons. Besides,

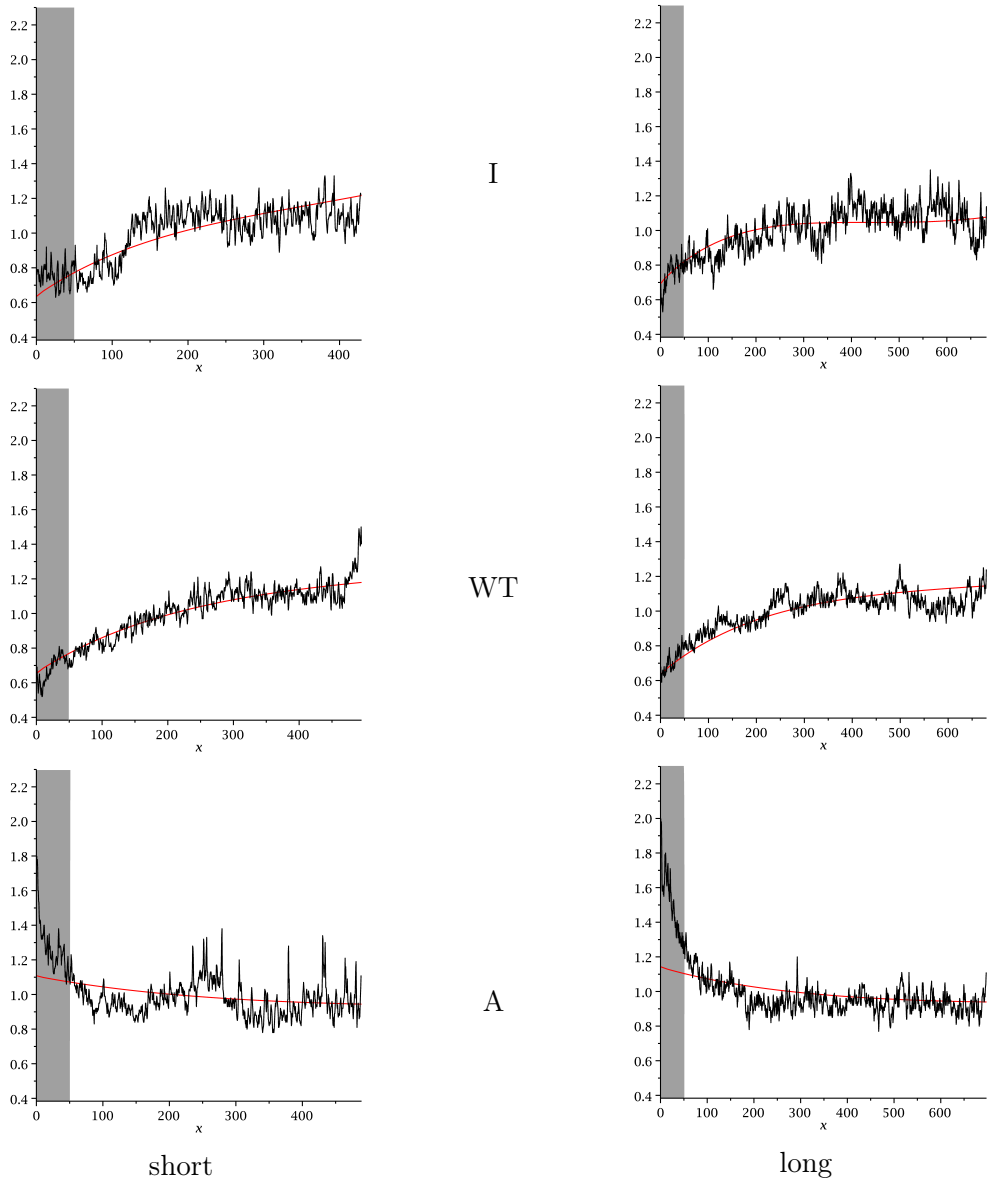
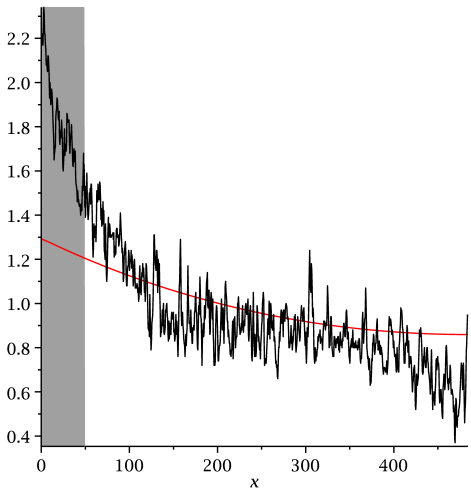
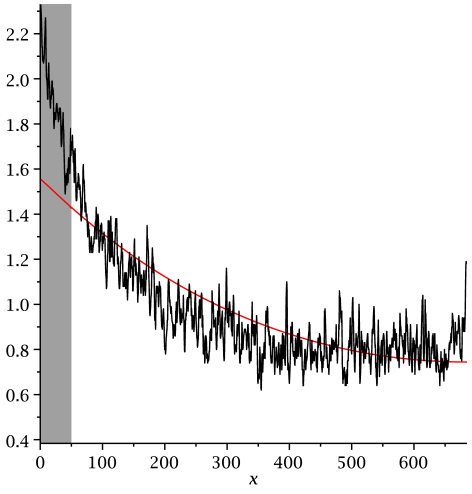


Figure 141: Concentration $p(x)$ as a function of x (µm), renormalized by the mean concentration. Average on several axons. Fit parameters : $l_d = 1$ µm (not important as long as it is short compared to other length scales : it would only change the local shape close to $x = 0$). $l_m^{WT} = 1500$ µm (If l_m^{WT} is larger, there would be little change. Another value $l_m \geq 1000$ µm would also work). $l_c = 250$ µm (another value in 200 µm – 300 µm would also work). $(j_c/j_d)^{WT} = 0.5$ (a value in 0.3-1 would also work). $\kappa^I = -0.003$ µm^{-1}, $\kappa^{WT} = 0.012$ µm^{-1}, $\kappa^A = 0.012$ µm^{-1}.

cultured neurons with mutant receptors have axons more or less branched than the WT, depending on the mutant type : it could be that with fewer receptors, an axon does not receive signals triggering its growth. It could radically change the network properties. The activity of these receptors being modified by pharmacological drugs as well as by cannabis, their use could affect overall brain networks properties. Zsolt Lenkei and Anne Simon are now on a project studying links between chronic exposure to cannabinoids, structural abnormalities and onset of psychosis : microscopic



short



long

Figure 142: SST receptor : $p(x)$ as a function of x (µm), renormalized by the mean concentration. Average on several neurons. Fit parameters : $l_d = 1$ µm (not important as long as small, except for the local shape in $x = 0$). $l_m = 500$ µm. Here $\kappa = 0$ (reflecting boundary), but $\kappa \neq 0$ would only make a local difference : indeed, κ practically adds a source (or a well) in $x = 0$, and here, as l_d is short, and as there is no other transport pathway, inside transport is also equivalent to a source in $x = 0$: the shape only depends on l_m (except very locally, close to the soma).

properties leading to macroscopic changes.

This example of the distribution of receptors shows the importance of transport in biology.

References

A Ajdari. Transport by active filaments. *Europhysics Letters*, 31:69–74, 1995.

B Alberts. *Molecular Biology of the Cell*. Garland, New York, 2002.

A Bancaud, S Huet, N Daigle, J Mozziconacci, J Beaudouin, and J Ellenberg. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. 2009.

A L Barabasi and R Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

M Barbi, C Place, V Popkov, and M Salerno. Base-sequence-dependent sliding of proteins on DNA. *Physical Review E*, 70:041901, 2004.

F Bartumeus. Behavioral intermittence, lévy patterns, and randomness in animal movement. *Oikos*, 118:488–494, 2009.

F Bartumeus and S A Levin. Fractal reorientation clocks: Linking animal behavior to statistical patterns of search. *Proceedings of the National Academy of Sciences*, 105:19072–19077, 2008.

F Bartumeus, J Catalan, U L Fulco, M L Lyra, and G M Viswanathan. Optimizing the encounter rate in biological interactions: Levy versus Brownian strategies. *Physical Review Letters*, 88:097901, 2002.

F Bartumeus, F Peters, S Pueyo, C Marrase, and J Catalan. Helical Levy walks: Adjusting searching statistics to resource availability in microzooplankton. *Proceedings of the National Academy of Sciences*, 100(22):12771–12775, 2003.

W J Bell. *Searching behaviour: the behavioural ecology of finding resources*. Chapman and Hall, 1991.

D Ben-Avraham and S Havlin. *Diffusion and reaction in fractals and disorders systems*. Cambridge University Press, 2000.

S Benhamou. How many animals really do the Levy walk? *Ecology*, 88:1962–1969, 2007.

S Benhamou. Efficiency of area-concentrated searching behaviour in a continuous patchy environment. *Journal of Theoritical Biology*, 159:67–81, 1992.

O Bénichou, M Coppey, M Moreau, P-H Suet, and R Voituriez. Optimal search strategies for hidden targets. *Physical Review Letters*, 94:198101, 2005a.

O Bénichou, M Coppey, M Moreau, P-H Suet, and R Voituriez. Averaged residence times of stochastic motions in bounded domains. *Europhysics Letters*, 70:42–48, 2005b.

O Bénichou, M Coppey, M Moreau, P-H Suet, and R Voituriez. A stochastic model for intermittent search strategies. *Journal of Physics: Condensed Matter*, 17:1–12, 2005c.

O Bénichou, C Loverdo, M Moreau, and R Voituriez. Two-dimensional intermittent search processes: An alternative to lévy flight strategies. *Physical Review E*, 74: 020102, 2006.

O Bénichou, C Loverdo, M Moreau, and R Voituriez. A minimal model of intermittent search in dimension two. *Journal of Physics: Condensed Matter*, 19:065141, 2007.

O Bénichou, C Loverdo, M Moreau, and R Voituriez. Optimizing intermittent reaction paths. *Physical Chemistry Chemical Physics*, 10:7059–7072, 2008a.

O Bénichou, C Loverdo, and R Voituriez. How gene colocalization can be optimized by tuning the diffusion constant of transcription factors. *Europhysics Letters*, 84: 38003, 2008b.

O Bénichou, Y Kafri, M Sheinman, and R Voituriez. Searching fast for a target on a DNA without falling to traps. *Physical Review Letters*, 103:138102, 2009.

A M Berezhkovskii, V Zaloj, and N Agmon. Residence time of a brownian particle. *Physical Review E*, 57(4):3937–3947, 1998.

H C Berg. *E.Coli in motion*. Springer, New York, 2004.

O G Berg and C Blomberg. Association kinetics with coupled diffusional flows - special application to lac repressor-operator system. *Biophysical Chemistry*, 4: 367–381, 1976.

O G Berg, R B Winter, and P H Von Hippel. Diffusion-driven mechanisms of protein translocation on nucleid acids. 1. models an theory. *Biochemistry*, 20:6929–6948, 1981.

S Blanco and R Fournier. An invariance property of diffusive random walks. *Europhysics Letters*, 61:168–173, 2003.

S M Block, C L Asbury, J W Shaevitz, and M J Lang. Probing the kinesin reaction cycle with a 2d optical force clamp. *PNAS*, 100:2351–2356, 2003.

I Bonnet, A Biebricher, P-L Porté, C Loverdo, O Bénichou, R Voituriez, C Escudé, W Wende, A Pingoud, and P Desbiolles. Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleid Acids Research*, 36:4118–27, 2008.

D Boyer, G Ramos-Fernandez, O Miramontes, J L Mateos, G Cocho, H Larralde, H Ramos, and F Rojas. Scale-free foraging by primates emerges from their interaction with a complex environment. *Proceedings of the Royal Society B*, 273: 1743–1750, 2006.

P Bressloff and J Newby. Directed intermittent search for hidden targets. *New Journal of Physics*, 11:023033, 2009.

J M Chambers, C L Mallows, and B W Stuck. A method for simulating stable random variables. *Journal of the american statistical association*, 71:340–344, 1976.

L Champagne, R G Carl, and R Hill. Search theory, agent-based simulation, and u-boats in the bay of biscay. *Proceedings of The 2003 Winter Simulation Conference*, 1-2:991–998, 2003.

E L Charnov. Optimal foraging, marginal value theorem. *Theoretical Population Biology*, 9:129–136, 1976.

A V Chechkin, I M Zaid, M A Lomholt, I M Sokolov, and R Metzler. Bulk-mediated surface diffusion along a cylinder: Propagators and crossovers. *Physical Review E*, 79:040105, 2009.

S Condamin, O Bénichou, V Tejedor, R Voituriez, and J Klafter. First-passage times in complex scale-invariant media. *Nature*, 450:77–80, 2007.

S Condamin, V Tejedor, R Voituriez, O Bénichou, and J Klafter. Probing microscopic origins of confined subdiffusion by first-passage observables. *Proceedings of the National Academy of Sciences*, 105:5675–5680, 2008.

M Coppey, O Bénichou, R Voituriez, and M Moreau. Kinetics of target site localization of a protein on DNA. *Biophysical Journal*, 87:1640–1649, 2004.

V Dahirel, F Paillusson, M Jardat, M Barbi, and J-M Victor. Nonspecific DNA-protein interaction: Why proteins can diffuse along DNA. *Physical Review Letters*, 102:228101, 2009.

P G de Gennes. La percolation: un concept unificateur. *La Recherche*, 7:919, 1976.

A Dussutour, J L Deneubourg, and V Fourcassie. Amplification of individual preferences in a social context: the case of wall-following in ants. *Proceedings of the Royal Society B*, 272:705–714, 2005.

A M Edwards, R A Phillips, N W Watkins, M P Freeman, E J Murphy, V Afanasyev, S V Buldyrev, M G E Da Luz, E P Raposo, H E Stanley, and G M Viswanathan. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449:1044, 2007.

J Elf, G W Li, and X S Xie. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, 316:1191–1194, 2007.

I Eliazar, T Koren, and J Klafter. Searching circular DNA strands. *Journal of Physics-Condensed Matter*, 19:065140, 2007.

I Eliazar, T Koren, and J Klafter. Parallel search of long circular strands: Modeling, analysis, and optimization. *Journal of Physical Chemistry B*, 112:5905–5909, 2008.

A-M Florescu and M Joyeux. Description of nonspecific DNA-protein interaction and facilitated diffusion with a dynamical model. *Journal of Chemical Physics*, 130:015103, 2009.

J R Frost and L D Stone. Review of search theory: Advances and applications to search and rescue decision support. *http.//www.rdc.uscg.gov/reports/2001/cgd1501dpexsum.pdf*, 2001.

M Fujiwara, P Sengupta, and S L McIntire. Regulation of body size and behavioural state of *C.elgans* by sensory perception and the EGL-4 cGMP-dependant protein kinase. *Neuron*, 36:1091–1102, 2002.

C W Gardiner. *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*. Springer, 1996.

I Golding and E C Cox. Physical nature of bacterial cytoplasm. *Physical Review Letters*, 96:098102, 2006.

J Gorman and E C Greene. Visualizing one-dimensional diffusion of proteins along DNA. *Nature structural and molecular biology*, 15:768–774, 2008.

D M Gowers, G G Wilson, and S E Halford. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proceedings of the National Academy of Sciences*, 102(44):15883–15888, 2005.

S E Halford. An end to 40 years of mistakes in DNA-protein association kinetics? *Biochemical Society Transactions*, 37:343–348, 2009.

S E Halford and J F Marko. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Research*, 32:3040–3052, 2004.

P Hanggi, P Talkner, and M Borkovec. Reaction-rate theory - 50 years after Kramers. *Reviews of Modern Physics*, 62:251–341, 1990.

S Hill, M T Burrows, and R N Hughes. Increased turning per unit distance as an area-restricted search mechanism in a pause-travel predator, juvenile plaice, foraging for buried bivalves. *Journal of Fish Biology*, 56:1497–1508, 2000.

D Holcman. Modeling DNA and virus trafficking in the cell cytoplasm. *Journal of Statistical Physics*, 127:471–494, 2007.

J Howard, A J Hudspeth, and R D Vale. Movement of microtubules by single kinesin molecules. *Nature*, 342:154–158, 1989.

M Hsieh and M Brenowitz. Comparison of the DNA association kinetics of the lac repressor tetramer, its dimeric mutant LacI(adi) and the native dimeric Gal repressor. *Journal of Biological Chemistry*, 272:22092–22096, 1997.

L H Hu, A Y Grosberg, and R Bruinsma. Are DNA transcription factor proteins Maxwellian demons? *Biophysical journal*, 95:1151–1156, 2008.

T Hu, A Y Grosberg, and B I Shklovskii. How proteins search for their specific sites on DNA : the role of DNA conformation. *Biophysical journal*, 90:2731–2744, 2006.

S Huet, E Karatekin, V S Tran, I Fanget, S Cribier, and J P Henry. Analysis of transient behavior in complex trajectories: Application to secretory vesicle dynamics. *Biophysical Journal*, 91:3542–3559, 2006.

R B Huey. *The psychology and pedagogy of reading*. MIT Press, 1968.

B D Hughes. *Random Walks and Random environments. Volume 2 : random environments*. Oxford science publications, 1996.

A James, M J Plank, and R Brown. Optimizing the encounter rate in biological interactions: Ballistic versus Levy versus Brownian strategies. *Physical Review E*, 78:051128, 2008.

B Jiang, J Yin, and S Zhao. Characterizing the human mobility pattern in a large street network. *Physical Review E*, 80:021136, 2009.

H Kabata, O Kurosawa, I Arai, M Washizu, S A Margarson, R E Glass, and N Shimamoto. Visualization of single molecules of RNA-polymerase sliding along DNA. *Science*, 262:1561–1563, 1993.

M Kac. *Probability and related topics in physical sciences*. Interscience publisher, 1959.

Y Kafri and R A Da Silveira. Steady-state chemotaxis in Escherichia coli. *Physical Review Letters*, 100:238101, 2008.

M Kampmann. Facilitated diffusion in chromatin lattices: mechanistic diversity and regulatory potential. *Molecular Microbiology*, 57:889–899, 2005.

K Kiontke and W Sudhaus. Ecology of *Caenorhabditis* species. In The C. elegans Research Community, editor, *WormBook*, page http://www.wormbook.org. WormBook, 2005.

J A Koenig and J M Edwardson. Endocytosis and recycling of G protein-coupled receptors. *18*, pages 276–287, 1997.

G Kolesov, Z Wunderlich, O N Laikova, M S Gelfand, and L A Mirny. How gene order is influenced by the biophysics of transcription regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 104:13948–13953, 2007.

G Komazin-Meredith, R Mirchev, D E Golan, A M van Oijen, and D M Coen. Hopping of a processivity factor on DNA revealed by single-molecule assays of diffusion. *Proceedings of the National Academy of Sciences of the United States of America*, 105:10721–10726, 2008.

D L Kramer and R L McLaughlin. The behavioral ecology of intermittent locomotion. *American Zoologist*, 41:137–153, 2001.

J Lepolard. Les stratégies d'approvisionnement social chez l'Humain : Étude d'une population de pêcheurs à pied en Baie de Somme. Master's thesis, Université de Picardie, France, 2007.

C Leterrier, D Bonnard, D Carrel, J Rossier, and Z Lenkei. Constitutive endocytic cycle of the CB1 cannabinoid receptor. *The Journal of Biological Chemistry*, 279 (34):36013–36021, 2004.

P Levitz, M Zinsmeister, P Davidson, D Constantin, and O Poncelet. Intermittent brownian dynamics over a rigid strand : Heavily tailed relocation statistics in simple geometry. *Physical Review E*, 78:030102, 2008.

G W Li, O G Berg, and J Elf. Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nature Physics*, 5:294–297, 2009.

L Li, S F Nørrelykke, and E C Cox. Persistent cell motion in the absence of external signals : a search strategy for eukaryotic cells. *Plos One*, 3:e2093, 2008.

T M Lohman. Kinetics of protein-nucleic acid interactions - use of salt effects to probe mechanisms of interaction. *CRC Critical Reviews In Biochemistry*, 19:191–245, 1986.

M A Lomholt, T Ambjornsson, and R Metzler. Optimal target search on a fast-folding polymer chain with volume exchange. *Physical Review Letters*, 95:260603, 2005.

M A Lomholt, I M Zaid, and R Metzler. Subdiffusion and weak ergodicity breaking in the presence of a reactive boundary. *Physical Review Letters*, 98:200603, 2007.

M A Lomholt, T Koren, R Metzler, and J Klafter. Levy strategies in intermittent search processes are advantageous. *Proceedings of the National Academy of Sciences*, 105:11055–11059, 2008.

M A Lomholt, B van den Broek, S M J Kalisch, and G L Wuiteand R Metzler. Facilitated diffusion with DNA coiling. *PNAS*, 106:8204–8208, 2009.

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Enhanced reaction kinetics in biological cells. *Nature physics*, 4:134–137, 2008.

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Robustness of optimal intermittent search strategies in one, two, and three dimensions. *Physical review E*, 80:031146, 2009a.

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Reaction kinetics in active media. *Journal of statistical mechanics - theory and experiments*, page P02045, 2009b.

C Loverdo, O Bénichou, R Voituriez, A Biebricher, I Bonnet, and P Desbiolles. Quantifying hoping and jumping in facilated diffusion of DNA-binding proteins. *Physical Review Letters*, 102:188101, 2009c.

K Maeda, Y Imae, J I Shioi, and F Oosawa. Effect of temperature on motility and chemotaxis of Escheria coli. *Journal of Bacteriology*, 127(3):1039–1046, 1976.

V Di Marzo, M Bifulco, and L De Petrocellis. The endocannabinoid system and its therapeutic exploitation. *Nature Reviews Drug Discovery*, 3:771–784, 2004.

Y Meroz, I Eliazar, and J Klafter. Facilitated diffusion in a crowded environment : from kinetics to stochastics. *Journal of Physics A - Mathematical and Theoretical*, 42:434012, 2009.

R Metzler and J Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports – Review Section of Physics Letters*, 339:1–77, 2000.

L Mirny. News and views : Cell commuters avoid delays. *Nature Physics*, 4:93–95, 2008.

M Moreau, O Bénichou, C Loverdo, P-H Suet, and R Voituriez. Intermittent search processes: Chance against strategy. *AIP Conference Proceedings*, 913:065141, 2007a.

M Moreau, O Bénichou, C Loverdo, and R Voituriez. Intermittent search process in disordered medium. *Europhysics Letters*, 77:20006, 2007b.

M Moreau, O Bénichou, C Loverdo, and R Voituriez. Dynamical and spatial disorder in an intermittent search process. *Journal of physics A : mathematical and theoretical*, 42:434007, 2009.

C Nakada, K Ritchie, Y Oba, M Nakamura, Y Hotta, R Iino, R S Kasai, K Yamaguchi, T Fujiwara, and A Kusumi. Accumulation of anchored proteins forms membrane diffusion barriers during neuronal polarization. *Nature Cell Biology*, 5: 626–U3, 2003.

G Nardone, J George, and J G Chirikjian. Differences in the kinetic-properties of BamHI endonuclease and methylase with linear DNA substrates. *Journal of Biological Chemistry*, 261:2128–2133, 1986.

F Nedelec, T Surrey, and A C Maggs. Dynamic concentration of motors in microtubule arrays. *Physical Review Letters*, 86:3192–3195, 2001.

J M Newby and P C Bressloff. Directed intermittent search for a hidden target on a dendritic tree. *Physical Review E*, 80:021913, 2009.

W J O'Brien, H I Browman, and B I Evans. Search strategies of foraging animals. *American Scientist*, 78:152–160, 1990.

G Oshanin, H S Wio, K Lindenberg, and S F Burlatsky. Intermittent random walks for an optimal search strategy: one-dimensional case. *Journal of Physics: Condensed Matter*, 19:065142, 2007.

B O'Shaughnessy and I Procaccia. Analytical solutions for diffusion on fractal objects. *Physical Review Letters*, 54:455–458, 1985.

K Pearson. The problem of the random walk. *Nature*, 72:294–294, 1905.

J T Pierce-Shimonura, T M Morse, and S R Lockery. The fundamental role of pirouettes in *Caenorhabditis elegans* chemotaxis. *The journal of neuroscience*, 19 (21):9557–9569, 1999.

A Ramezanpour. Intermittent exploration on a scale-free network. *Europhysics Letters*, 77:60004, 2007.

E P Raposo, S V Buldyrev, M G E Da Luz, M C Santos, H E Stanley, and G M Viswanathan. Dynamical robustness of Levy search strategies. *Physical Review Letters*, 91(24), 2003.

W S Rasband. ImageJ. *U. S. National Institutes of Health, Bethesda, Maryland, USA*, page http://rsb.info.nih.gov/ij/, 1997-2009.

S Redner. *A guide to first passage time processes*. Cambridge University Press, 2001.

J Reingruber and D Holcman. The gated narrow escape time for molecular signaling. *Physical Review Letters*, 103:148102, 2009.

A M Reynolds. On the intermittent behaviour of foraging animals. *Europhysics Letters*, 75:517–520, 2006.

A M Reynolds and F Bartumeus. Optimising the success of random destructive searches: Levy walks can outperform ballistic motions. *J Theor Biol*, 260:98–103, 2009.

S A Rice. Diffusion-limited reactions. In C H Bamford, C F H Tipper, and R G Compton, editors, *Compr. Chem. Kinetics*, page 25. Elsevier, New York, 1985.

H R Richardson and L D Stone. Operations analysis during the underwater search for Scorpion. *Naval Research Logistics Quarterly*, 18:141 – 157, 1971.

A D Riggs, S Bourgeoi, and M Cohn. Lac repressor-operator interaction .3. kinetic studies. *Journal of Molecular Biology*, 53:401, 1970.

F Rojo, C E Budde, and H S Wio. Optimal intermittent search strategies. *Journal of Physics A-Mathematical and Theoretical*, 42:125002, 2009.

H Salman and A Libchaber. A concentration-dependent switch in the bacterial response to temperature. *Nature Cell Biology*, 9:1098–U78, 2007.

H Salman, A Abu-Arish, S Oliel, A Loyter, J Klafter, R Granek, and M Elbaum. Nuclear localization signal peptides induce molecular delivery along microtubules. *Biophysical Journal*, 89:2134–2145, 2005.

H Salman, A Zilman, C Loverdo, M Jeffroy, and A Libchaber. Solitary modes of bacterial culture in a temperature gradient. *Physical Review Letters*, 97:118101, 2006.

M C Santos, E P Raposo, G M Viswanathan, and M G E Da Luz. Optimal random searches of revisitable targets: Crossover from superdiffusive to ballistic random walks. *Europhysics Letters*, 67(5):734–740, 2004.

M C Santos, D Boyer, O Miramontes, G M Viswanathan, E P Raposo, J L Mateos, and M G E Da Luz. Origin of power-law distributions in deterministic walks: The influence of landscape geometry. *Physical Review E*, 75:061114, 2007.

J M Schurr. The one-dimensional diffusion coefficient of proteins absorbed on DNA. hydrodynamic considerations. *Biophysical Chemistry*, 9:413–414, 1979.

M P Sheetz and J A Spudich. Movement of myosin-coated fluorescent beads on actin cables in vitro. *Nature*, 303:31–35, 1983.

M F Shlesinger. Mathematical physics: Search research. *Nature*, 443:281, 2006.

M F Shlesinger and J Klafter. Random-walks in liquids. *Journal of Physical Chemistry*, 93:7023–7026, 1989.

M F Shlesinger and J Klafter. Lévy walk vs. Lévy flights. In H E Stanley and N Ostrowski, editors, *On growth and forms*, pages 279–283. Martinus Nijhof Publishers, Amsterdam, 1986.

A Simon, C Loverdo, A L Gaffuri, A Roland, C Leterrier, O Bénichou, M Piel, R Voituriez, and Z Lenkei. Model for activation-state dependent regulation of GPCR availability. submitted, 2009.

M Slutsky and L Mirny. Kinetics of protein-DNA interaction: Facilitated target location in sequence-dependent potential. *Biophysical Journal*, 87:4021–4035, 2004.

I M Sokolov, R Metzler, K Pant, and M C Williams. First passage time of n excluded-volume particles on a line. *Physical Review E*, 72:041102, 2005.

W W Sprenger, W D Hoff, J P Armitage, and K J Hellingwerf. the eubacterium Ectothiorhodospira-halophila is negatively phototactic, with a wavelength dependence that fits the absorption-spectrum of the photoactive yellow protein. *Journal of Bacteriology*, 175:3096–3104, 1993.

J D Taylor and S E Halford. Discrimination between DNA-sequences by the EcoRV restriction endonuclease. *Biochemistry*, 28:6198–6207, 1989.

I M Tolic-Norrelykke, E L Munteanu, G Thon, L Oddershede, and K Berg-Sorensen. Anomalous diffusion in living yeast cells. *Physical Review Letters*, 93:078102, 2004.

J Travis. Ecology - do wandering albatrosses care about math? *Science*, 318: 742–743, 2007.

B van den Broek, M A Lomholt, S M J Kalisch, R Metzler, and G J L Wuite. How DNA coiling enhances target localization by proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 105:15738–15742, 2008.

M Vergassola, E Villermaux, and B I Shraiman. 'infotaxis' as a strategy for searching without gradients. *Nature*, 445:406–409, 2007.

G M Viswanathan, S V Buldyrev, S Havlin, M G E Da Luz, E P Raposo, and H E Stanley. Optimizing the success of random searches. *Nature*, 401(6756):911–914, 1999.

G M Viswanathan, E P Raposo, and M G E Da Luz. Levy flights and superdiffusion in the context of biological encounters and random searches. *Physics of Life Reviews*, 5:133–150, 2008.

A Vogt. Stehfest algorithm. *http://www.mapleprimes.com/blog/alec/numerical-inverse-laplace-transform-0*, 2006.

P H Von Hippel. From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annual Review of Biophysics and Biomolecular Structure*, 36:79–105, 2007.

M von Smoluchowski. Experiments on a mathematical theory of kinetic coagulation of coloid solutions. *Zeitschrift fur physikalische chemie–stochiometrie und verwandtschaftslehre*, 92:129–168, 1917.

G H Weiss. *Aspects and applications of the random walk*. North-Holland, New York, 1994.

G G Wilson and N E Murray. Restriction and modification systems. *Annual Review of Genetics*, 25:585–627, 1991.

R B Winter and P H Von Hippel. Diffusion-driven mechanisms of protein translocation on nucleid acids. 2. the Escherichia coli repressor-operator interaction : equilibrium measurements. *Biochemistry*, 20:6948–6960, 1981.

R B Winter, O G Berg, and P H Von Hippel. Diffusion-driven mechanisms of protein translocation on nucleid acids. 2. the Escherichia coli repressor-operator interaction : kinetic measurements and conclusion. *Biochemistry*, 20:6961–6977, 1981.

Z Wunderlich and L A Mirny. Spatial effects on the speed and reliability of protein-DNA search. *Nucleic Acids Research*, 36:3570–3578, 2008.

E A Yuzbasyan, H Lin, N C Darnton, J B Stock, P Silberzan, S Park, P M Wolanin, and R H Austin. Influence of topology on bacterial social interaction. *Proceedings of the National Academy of Sciences*, 100:13910–13915, 2003.

Publications linked to this thesis

The contributions of this thesis are included in the following articles :

A Simon, C Loverdo, A L Gaffuri, A Roland, C Leterrier, O Bénichou, M Piel, R Voituriez, and Z Lenkei. Model for activation-state dependent regulation of GPCR availability. submitted, 2009

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Robustness of optimal intermittent search strategies in one, two, and three dimensions. *Physical review E*, 80:031146, 2009a

M Moreau, O Bénichou, C Loverdo, and R Voituriez. Dynamical and spatial disorder in an intermittent search process. *Journal of physics A : mathematical and theoretical*, 42:434007, 2009

C Loverdo, O Bénichou, R Voituriez, A Biebricher, I Bonnet, and P Desbiolles. Quantifying hoping and jumping in facilated diffusion of DNA-binding proteins. *Physical Review Letters*, 102:188101, 2009c

O Bénichou, C Loverdo, M Moreau, and R Voituriez. Optimizing intermittent reaction paths. *Physical Chemistry Chemical Physics*, 10:7059–7072, 2008a

O Bénichou, C Loverdo, and R Voituriez. How gene colocalization can be optimized by tuning the diffusion constant of transcription factors. *Europhysics Letters*, 84:38003, 2008b

I Bonnet, A Biebricher, P-L Porté, C Loverdo, O Bénichou, R Voituriez, C Escudé, W Wende, A Pingoud, and P Desbiolles. Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleid Acids Research*, 36:4118–27, 2008

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Enhanced reaction kinetics in biological cells. *Nature physics*, 4:134–137, 2008

O Bénichou, C Loverdo, M Moreau, and R Voituriez. A minimal model of intermittent search in dimension two. *Journal of Physics: Condensed Matter*, 19:065141, 2007

M Moreau, O Bénichou, C Loverdo, and R Voituriez. Intermittent search process in disordered medium. *Europhysics Letters*, 77:20006, 2007b

O Bénichou, C Loverdo, M Moreau, and R Voituriez. Two-dimensional intermittent search processes: An alternative to lévy flight strategies. *Physical Review E*, 74: 020102, 2006

They are also included in the following conference proceedings :

C Loverdo, O Bénichou, M Moreau, and R Voituriez. Reaction kinetics in active media. *Journal of statistical mechanics - theory and experiments*, page P02045, 2009b

M Moreau, O Bénichou, C Loverdo, P-H Suet, and R Voituriez. Intermittent search processes: Chance against strategy. *AIP Conference Proceedings*, 913:065141, 2007a

Remerciements

Tout d'abord, je remercie Pierre Levitz et Ralf Metzler pour avoir accepté de rapporter cette thèse, ainsi que Didier Chatenay, Jean-François Joanny et Emmanuel Trizac pour avoir accepté de faire partie du jury.

Un immense merci à Olivier Bénichou, mon directeur de thèse, qui m'a laissé beaucoup de liberté, tout en étant disponible quand j'en ai eu besoin. Il est à la fois brillant, et très à l'écoute de ceux qui travaillent avec lui. Dans l'équipe du LPTMC, merci à Raphaël Voituriez qui a aussi beaucoup contribué à l'encadrement de cette thèse, notamment (mais pas seulement) pour le lien avec la biologie et les biologistes. Merci aussi à Michel Moreau qui a contribué par son expérience. Pour la partie ADN/protéines, Pierre Desbiolles a eu aussi un grand rôle dans l'encadrement de cette thèse. Je suis admirative de sa connaissance approfondie du sujet, et de sa grande honnêteté intellectuelle.

Un des aspects qui me plait beaucoup dans la recherche, c'est de discuter avec d'autres personnes. Voici donc une liste de collaborations, qui ont été plus ou moins loin, mais qui m'ont toutes appris beaucoup de choses :

- Sur EcoRV : merci à toute l'équipe optique et biologie, notamment Pierre Desbiolles comme je l'ai déjà dit, Isabelle Bonnet qui m'a expliqué beaucoup de choses, Natacha Porté avec qui j'ai eu de longues discussions sur les simulations, et Andreas Biebriecher qui a bien voulu me montrer l'expérience en détails. Au Muséum d'histoire naturelle, Christophe Escudé m'a montré avec pédagogie des bases de biologie moléculaire (en particulier la méthode de construction de la molécule d'ADN utilisée dans les expériences).

- Merci à Zsolt Lenkei et Anne Simon sur les récepteurs CB1R : ils ont expliqué avec beaucoup de patience un univers initialement peu compréhensible, habité de multiples sigles. Ils ont aussi toujours exhumé leurs données rapidement et efficacement à la moindre demande.

- Merci à Michel Sokolowski pour la recherche chez les humains, à la fois chercheur en psychologie et bricoleur hors-pair.

- Merci à tous les membres de l'ANR DIOPTRY!

- Merci à Juliette Ben Arous de m'avoir montré beaucoup de choses sur C.elegans, même si au final le déplacement de ce ver n'est sans doute pas intermittent au sens où on l'entend.

- Merci à Anna-Maria Lennon-Dumenil et Matthieu Piel pour leurs explications sur les cellules dendritiques.

- Merci à François Darchen, Sébastien Huet et Erdem Karatekin pour nous avoir indiqué un cas où les vésicules réagissent peu quand elles sont transportées activement, et pour en avoir discuté en détails.

- Merci à Tal Koren pour des discussions approfondies sur les méthodes de simulations et sur les marches de Lévy.

Et au LPTMC même j'ai beaucoup de gens à remercier.

J'ai aussi souvent discuté avec les autres thésards/post-docs de mon équipe : Pierre-Henri Suet, Sylvain Condamin, Bob Meyer, Vincent Tejedor, Claire Chevalier. Merci à tous!

J'ai eu aussi des discussions particulièrement intéressantes avec l'autre équipe du LPTMC qui s'intéresse à l'interaction enzyme ADN : Maria Barbi, Jean-Marc Victor, Julien Mozziconacci, Fabien Paillusson. Merci!

D'une manière plus générale, la liste des personnes intéressantes avec qui j'ai discuté de choses et d'autres au laboratoire est presque égale à la liste des membres du LPTMC, je n'ai donc pas besoin de répéter ici une liste facile à trouver sur le web.

Merci à Laura Messio pour l'organisation du séminaire jeune, qui m'a permis de parler de mon travail, mais aussi de voir les travaux des autres.

Merci aux co-bureaux! J'ai passé les deux premières années de thèse dans le bureau du fond, froid et sombre, mais éclairé par la présence en première année de Xavier Rozanska et en deuxième année de Laura Messio. Et depuis la troisième année de thèse, je suis passée au très sympathique "bureau des filles" avec Laura Messio encore, mais aussi Juliana Restrepo et Rhoda Hawkins, toutes les deux toujours de bonne humeur (et sans oublier le discret Mathieu Bauchy, mis (par manque de place probablement) dans le "bureau des filles").

Merci aussi aux directeurs successifs du LPTMC, Bertrand Guillot et Pascal Viot, efficaces, ouverts, attentifs. Pascal Viot est aussi un rouage essentiel de l'infrastructure informatique du laboratoire, ainsi que Michel Quaggetto. Merci!

Merci à Martine Postic et à Dalla Foglia Sylvie, des secrétaires très efficaces et chaleureuses, en un seul mot parfaites.

Et je n'oublie pas non plus Nacera Lebled puis Kader Zeggaï pour le ménage, sans qui la poussière envahirait la tour 24 et ferait tousser les ordinateurs!

Pour revenir à cette thèse, je dois remercier tous les relecteurs, certains qui ont tout relu (ou presque), et d'autres seulement une sous-partie, mais qui ont tous contribué à améliorer sensiblement le manuscrit, et qui ont répondu présent au moment clé : Olivier Bénichou, Alexeï Chepelianskii, Francesca Chiodi, Pierre Desbiolles, Myrian Fenina, Rhoda Hawkins, Marie Jardat, François Levrat, Blaise Li, Claude Loverdo (senior), Maï Nguyen, Émilia Robin, Michel Sokolowski, Pierre-Henri Suet, Laurent Tournier, Nicolas Verzelen, Hugo Viciana et Raphaël Voituriez.

Sur la forme, j'ai eu des conseils précieux sur LaTeXpar Émilia Robin, Florent Hivert et Laurent Tournier : merci! Et aussi via le forum informatique de l'ENS, où à n'importe quelle heure des gens compétents répondent dans un délai tellement court qu'il en devient presque inquiétant.

Il faudrait aussi que je remercie tous ceux qui ont contribué à ma formation scientifique, mais ce serait fort long. Je dois juste souligner que je me sens très redevable de l'environnement favorable et des conditions avantageuses que j'ai trouvés à l'ENS.

J'ai eu la chance de faire un monitorat pendant ma thèse. J'ai beaucoup aimé enseigner à l'université, et il faut dire que j'ai eu des collègues très compétents, et des étudiants agréables.

Et j'ai aussi une une vie en dehors de l'université...

Ces années de thèse je me suis investie dans deux associations, l'association d'alphabétisation du foyer Pinel, et le club cirque de l'ENS. J'y ai rencontré beaucoup de gens intéressants, ouverts, passionnés, et je me suis énormément amusée.

"L'amitié double les joies et réduit de moitié les peines" : ce n'est pas de moi, vous vous en doutez bien, mais je partage ce sentiment. Merci donc à mes amis, sans faire de liste sinon je commettrai probablement des injustices.

Merci aussi à la famille (au sens étendu bien sûr, tous n'étant pas directement liés par la génétique), ceux qui sont là, et ceux qui sont déjà partis (et c'est toujours trop tôt!).

Et merci à Hugo bien sûr, compagnon de tous les jours, des tristesses et des bonheurs.

Et enfin, comme il y a sans doute des trous dans mon énumération, je conclurai par ces quelques mots, merci à tous...

Résumé

Cette thèse concerne les stratégies de recherches de cible dites intermittentes, qui alternent des phases lentes permettant la détection de la cible, et des phases rapides sans détection.

Un exemple à l'échelle macroscopique est celui d'animaux en quête de nourriture. Nous en proposons un modèle, alternatif aux célèbres stratégies de Lévy, et montrons analytiquement que le temps moyen de recherche peut être minimisé en fonction des durées moyennes de chaque phase.

Un premier exemple à l'échelle microscopique est celui de la recherche par des protéines de cibles sur l'ADN. Nous calculons analytiquement la distribution de la distance parcourue le long de l'ADN lors d'une excursion 3D, l'adaptons à une expérience de molécule unique et montrons que les trajectoires observées combinent des diffusions 1D et 3D. Un autre exemple cellulaire concerne le transport actif de vésicules, qui diffusent ou se lient à des moteurs assurant un déplacement balistique. Nous optimisons la constante cinétique dans un modèle général de réaction limitée par ce type de transport.

Finalement, ces stratégies intermittentes pourraient constituer un mécanisme de recherche générique. Nous étudions de manière systématique l'influence de la modélisation de la phase de détection et de la dimension de l'espace, et montrons que l'optimalité des stratégies intermittentes est un résultat robuste.

Mot-clés : physique statistique, marches aléatoires, biophysique, stratégies de recherche, processus stochastiques, temps de premier passage

Abstract

This thesis deals with intermittent target search strategies, which combine slow phases, allowing the searcher to detect the target, and fast phases without detection.

Foraging animals are an example at the macroscopic scale. We propose a model, alternative to the famous Lévy strategies, and show analytically that the mean search time can be minimized as a function of the mean duration of both phases.

Our first example at the microscopic scale is given by proteins searching for targets on DNA. We analytically calculate the distribution of the distance travelled along DNA during a 3D excursion, adapt it to a single-molecule experiment and show that the observed trajectories combine 1D and 3D diffusion. Another cellular example is provided by active transport of vesicles, which diffuse or bind to motors performing ballistic motion. We optimize the global kinetic constant within a general framework of reactions limited by this kind of transport.

Finally, these intermittent strategies could constitute a generic search mechanism. We systematically study the influence both of the modeling of the detection phase and of the space dimension, and show that the optimality of intermittent strategies is a robust result.

Keywords : statistical physics, random walks, biophysics, search strategies, stochastic processes, first passage time