Split of Territories in Concurrent Optimization

Jean-Antoine Désidéri

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Split of Territories in Concurrent Optimization

Jean-Antoine Désidéri

N° 6108

8th October 2007

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6108--FR+ENG

Split of Territories in Concurrent Optimization

Jean-Antoine Désidéri*

Thème NUM — Systèmes numériques
Projet Opale

Rapport de recherche n° 6108 — 8th October 2007 —34 pages

Abstract: We propose a methodology for the numerical treatment of a concurrent optimization problem in which two criteria are considered, one, J_A, being more critical than the second, J_B. After completion of the parametric, possibly–constrained minimization of the single, primary functional J_A alone, approximations of the functional gradient and Hessian matrix, as well as K constraint gradients, are assumed to be available or calculated using *meta-models*. Then, the entire parametric space (a subset of $\mathbb{R}^{n+1}$) is split into two supplementary subspaces on the basis of a criterion related to the second variation. The construction is such that from the initial convergence point of the primary–functional minimization in full dimension, infinitesimal perturbations of the parameters lying in the second subspace, of specified dimension $p \leq n + 1 - K$, potentially cause the least degradation to the value of the primary functional. The latter subspace is elected as the support of the parameterization strategy of the secondary functional, J_B, in a concurrent optimization realized by an algorithm simulating a Nash game between players associated with the two functionals respectively. We prove a second result indicating that the original global optimum point of the primary problem in full dimension is Pareto-optimal for a trivial concurrent problem. This latter result permits us to define a continuum of Nash equilibrium points originating from the initial single-criterion optimum, in which the designer could potentially make a rational election of operating point. Thirdly, the initial single-criterion optimum is found to be robust. A simple minimization problem involving quadratic criteria is treated explicitly to demonstrate these properties in both cases of a linear and a nonlinear constraint. Lastly we note that the hierarchy introduced between the criteria applies to the split of parameters in preparation of a Nash game. The bias is therefore different in nature from the one that a Stackelberg-type game would introduce.

Key-words: Optimum–shape design, concurrent engineering, multi-criterion optimization, split of territories, Nash and Stackelberg game strategies, Pareto optimality

* Project Team Opale: `htpp://www-sop.inria.fr/opale`; Jean-Antoine.Desideri@sophia.inria.fr

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Partage de Territoires en Optimisation Concourante

Résumé : On propose une méthodologie pour le traitement numérique d'un problème d'optimisation concourante dans lequel deux critères sont à considérer, l'un, J_A, étant plus critique que le second, J_B. A convergence de la minimisation paramétrique, éventuellement sous contraintes, de la seule fonctionnelle principale, J_A, des approximations du gradient et de la matrice hessienne, ainsi que des K gradients de contraintes sont par hypothèse disponibles ou calculées en utilisant des *métamodèles*. Alors, on partage l'espace paramétrique entier (un sous-ensemble de $\mathbb{R}^{n+1}$) en deux sous-espaces supplémentaires sur la base d'un critère lié à la seconde variation. La construction est telle qu'à partir du point initial de convergence de la minimisation de la fonctionnelle principale en dimension complète, des perturbations infinitésimales des paramètres dans le second sous-espace, dont la dimension $p \leq n+1-K$ est spécifiée, causent potentiellement la moindre dégradation de la valeur de la fonctionnelle principale. Ce sous-espace est choisi comme support de la stratégie de paramétrisation de la seconde fonctionnelle, J_B, dans une optimisation concourante mise en oeuvre par un algorithme simulant un jeu de Nash entre des joueurs associés aux deux fonctionnelles respectivement. On prouve un deuxième résultat selon lequel le point optimum original du problème principal en dimension complète est Pareto-optimal vis-à-vis d'un problème trivial d'optimisation concourante. Ce dernier résultat nous permet de définir un continuum de points d'équilibre de Nash ayant pour origine le point optimum initial, donnant ainsi au concepteur la possibilité de choisir un point opérationnel. Tertio, on démontre que l'optimum initial mono-critère est robuste. Un problème simple de minimisation de critères quadratiques est traité explicitement pour illustrer ces propriétés dans les deux cas d'une contrainte linéaire ou non linéaire. Enfin on note que la hiérarchie introduite entre les critères s'applique au partage des paramètres en préparation d'un jeu de Nash. Le biais est donc par nature différent de celui qu'introduirait un jeu de type Stackelberg.

Mots-clés : Conception optimale de forme, ingénierie concourante, optimisation multi-critère, partage de territoires, stratégies de jeux de Nash ou de Stackelberg, Pareto-optimalité

1 Introduction, framework and general notations

When devising a numerical shape–optimization method in the context of a complex engineering situation, the practitioner is faced to an acute difficulty related to the participation, in a realistic formulation, of several relevant physical criteria originating from several disciplines. Perhaps the most commonly–used treatment of multi–criterion problems is the penalization approach in which one minimizes a single functional agglomerating all the criteria weighted by penalty constants. The method is computationally economical, but evidently, the resulting solution depends on the penalty constants whose adjustment is usually made with a fair amount of arbitrariness. Alternately, at the other extreme in computational cost involved, when feasible, identifying Pareto fronts made of non–dominated design points, has the great merit of providing the designer with a very rich information about the system supporting the decision making. However the corresponding implementation requires a very large number of simultaneous evaluations of several functionals.

A treatment of multi–criterion problems that eludes the question of adjusting the penalty constants, and that is computationally more economical than identifying the Pareto equilibrium front, is to seek a pseudo–optimal solution as the equilibrium point of a simulated dynamic game in which the set of design parameters is split into subsets, each subset being considered as the strategy or territory of a given functional. Nash or Stackelberg games are usually considered [1], [2]. Of course, the adopted definition of the splitting also introduces a bias, but we momentarily put aside this question. Examples of successful concurrent optimizations realized numerically by such dynamic games can be found for example in [3] and [4].

Another difficulty is very acute when optimum–shape design is sought w.r.t. an aerodynamic criterion as well as other criteria also related to complex distributed systems governed by partial–differential equations, mainly for two reasons. The first is that aerodynamics alone is costly to analyze in terms of functional evaluation. The second is that, for aircraft design, generally only a small degradation of the aerodynamic performance absolute optimum can be tolerated (sub–optimality) when introducing the other criteria without compromising the whole engineering concept. Hence, we would like to introduce the concurrent process in a scheme as harmless as possible to a *primary functional* of our choice.

To formulate this notion more precisely, we consider first an optimum–shape design problem in which a *primary functional* is minimized:

$$\begin{cases} \min\limits_{\Gamma} \; J_A(\Gamma) \\ \text{Subject to: } g(\Gamma) = 0 \end{cases} \tag{1}$$

We have in mind a functional related to a distributed physical system such as compressible aerodynamics governed by the Euler or Navier–Stokes equations in a region external to the shape Γ. This functional problem is reformulated as a parametric optimization problem by the choice of a shape parameterization (here assumed in two dimensions for notational simplicity). For example, following [5]:

$$\begin{cases} \min\limits_{X,Y} \; J_A(\Gamma(X,Y)) \\ \text{Subject to: } g(X,Y) = 0 \end{cases} \tag{2}$$

in which

$$X = \{\, x_k \,\} \quad (k = 0, 1, ..., n) \tag{3}$$

is the *support* of a Bézier representation, and

$$Y = \{\, y_k \,\} \quad (k = 0, 1, ..., n) \tag{4}$$

is the unknown *design vector*. In this case, the unknown shape is parameterized as follows:

$$x(t) = B_n(t)^T X \,, \quad y(t) = B_n(t)^T Y \tag{5}$$

in which t is the parameter ($0 \le t \le 1$), and $B_n(t)^T$ is an $(n+1)$-dimensional row–vector

$$B_n(t)^T = \left(B_n^0(t),\, B_n^1(t),\, ...,\, B_n^n(t) \right) \tag{6}$$

whose components are the Bernstein polynomials $B_n^k(t) = C_n^k \, t^k \, (1-t)^{n-k}$, and $C_n^k = n!/[k!(n-k)!]$. Many other types of parametric representations such as B–splines or NURBS, or supported by the Hicks–Henne functions [6] (as in [4]) can serve the same purpose of reduction to the finite dimension.

In our algorithms, the minimization is conducted iteratively with the support vector X held fixed over the design vector Y, as the sole unknown of the parametric optimization. This strategy has revealed to be effective provided the support vector X is statically or dynamically adapted once or several times in the course of the optimization; this important question is addressed elsewhere [5], [8]. At complete or nearly–complete convergence, $Y = Y_A^*$, we further assume that we dispose of reliable approximations of the gradient ∇J_A^* and the Hessian matrix H_A^*, as well as the possibly multi–dimensional constraint gradient, ∇g^*. If the numerical optimization is conducted by a gradient–based method such as the well–known *BFGS* algorithm [7], such approximations are provided by the iteration itself. However more rustic, but robust methods such as the simplex method [5] or evolutionary algorithms (e.g. genetic, as in [3], [4], or particle–swarm, as in [8], algorithms) have merits of their own. In the latter case, we additionally assume that *surrogate* or *meta* models can be elaborated to compute approximations of the desired quantities. This can be achieved in particular by techniques of *Design of Experiments* [9] and/or by training an *Artificial Neural Network* [10], or simply by *curve fitting*.

After completion of the above single–criterion optimization in full dimension $n + 1$, we would like to conduct a concurrent optimization initiated from Y_A^* and involving a secondary functional denoted J_B, perhaps originating from another physical discipline such as a structural mechanics, electro-magnetics, or other. We decide to perform this concurrent optimization by letting

$$\boxed{Y = Y(U, V) = Y_A^* + S \begin{pmatrix} U \\ V \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_{n+1-p} \end{pmatrix}, \quad V = \begin{pmatrix} v_p \\ \vdots \\ v_1 \end{pmatrix}} \tag{7}$$

in which S is an invertible $(n + 1) \times (n + 1)$ adjustable matrix, thereafter referred to as the *splitting matrix*, and by implementing a Nash game involving two players utilizing the sub-vectors $U \in \mathbb{R}^{n+1-p}$ and $V \in \mathbb{R}^p$ as their strategies to minimize the primary and secondary functionals respectively. We denote by $\bar{Y}$ the corresponding Nash equilibrium point if it exists. Recall that by definition, the subvectors $\bar{U}$ and $\bar{V}$ associated with the vector $\bar{Y} = Y(\bar{U}, \bar{V})$ are solutions, for fixed support X, to the partial problems:

$$\boxed{\left\{\begin{array}{l} \min\limits_{U \in \mathbb{R}^{n+1-p}} J_A\Big[\, Y(U, \bar{V}) \,\Big] \\ \text{Subject to: } g\Big[\, Y(U, \bar{V}) \,\Big] = 0 \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l} \min\limits_{V \in \mathbb{R}^{p}} J_B\Big[\, Y(\bar{U}, V) \,\Big] \\ \text{Subject to: no constraints} \end{array}\right.} \tag{8}$$

Note that the dimension of the vector U is $n + 1 - p$; it should be at least equal to the number K of scalar constraints (or dimension of the vector g). Thus, the adjustable dimension p is limited by the condition:

$$\boxed{p \le n + 1 - K} \tag{9}$$

The limiting case corresponding to $K + p = n + 1$ is examined in a special section. Otherwise mentioned, it is assumed throughout that a strict inequality holds above.

In the cited examples, [3] and [4], the splittings are partitions of the *primitive variables*, that is, the components of the vector Y. In a parametric shape optimization, these primitive variables are shape–control variables, as in the above Bézier parameterization, and are associated with the spatial locations where these controls are made active, or preponderant. The splitting matrix S is then a permutation matrix and it reflects an actual choice made of territories in the physical spatial domain where governing partial–differential equations, such as the equations of gas-dynamics for aerodynamics, are solved. This choice is then made on the basis of an *a priori* knowledge of the portions of the shape that are the most influential on the physical criteria, such as the shock region of the upper surface of a wing geometry when minimizing the pressure drag in transonic flow. In particular, physically–irrelevant choices are well–known to yield dynamic games that do not converge [3]. Thus, our objective is to gain generality in our numerical approaches by making a more abstract and systematic definition of the territories, in particular independent of *a priori* knowledge, or intuition, of the geometrical characteristics of the underlying possibly–complex physical situation.

In the more general splitting (7), the span of the first $n+1-p$ column–vectors of the splitting matrix S is viewed as the territory of the primary functional, and the span of the remaining p column–vectors as the secondary functional territory. Thus, the open question is how should the splitting matrix S be defined to realize an adequate split of territories. The point of view adopted throughout is to chose the splitting matrix S *a priori*, that is, after solving the primary problem in full dimension alone, but before the actual numerical implementation of the concurrent optimization; hence the choice is made on the sole basis of information related to the primary minimization problem, and in such a way that infinitesimal perturbations of the parameters lying in the secondary functional territory should potentially cause the least possible degradation to the value of the primary functional achieved after completion of the initial single–criterion minimization phase.

To support our discussion, we consider the expansion of the primary functional about the convergence point Y_A^* in the direction of a unit vector $\omega \in \mathbb{R}^{n+1}$:

$$J_A\left(Y_A^* + \varepsilon\omega\right) = J_A\left(Y_A^*\right) + \varepsilon\,\nabla J_A^* \,.\, \omega + \frac{\varepsilon^2}{2}\,\omega\,.\,H_A^*\,\omega + O(\varepsilon^3) \tag{10}$$

Our objective is to propose a splitting associated with the rational choice of a basis of vectors $\{\,\omega^j\,\}$ $(j=1,...,n+1)$, in a manner such that the tail p vectors $(n+2-p \le j \le n+1)$ are associated with variations $|\,J_A\left(Y_A^*+\varepsilon\omega^j\right) - J_A\left(Y_A^*\right)\,|$, that are, for fixed small–enough ε, as small as possible. The choice is made first in the case of an unconstrained problem for which $\nabla J_A^* = 0$ at convergence of the initial phase of minimization of the primary functional alone, and then extended to the more general case of a constrained problem.

It is assumed throughout that both criteria J_A and J_B are positive, and that:

$$J_A^* = J_A\left(Y_A^*\right) > 0\,, \quad J_B^* = J_B\left(Y_A^*\right) > 0 \tag{11}$$

These assumptions are made although not always essential; however if necessary, these conditions can easily be enforced by substituting the following criteria to the original ones:

$$\mathcal{J}_{\mathcal{A}} = \exp\left[\alpha\left(J_A - J_A^*\right)\right]\,, \quad \mathcal{J}_{\mathcal{B}} = \exp\left[\beta\left(J_B - J_B^*\right)\right] \tag{12}$$

in which α and β are positive constants chosen such that

$$\alpha\,\|\nabla J_A^*\| \sim \beta\,\|\nabla J_B^*\| \tag{13}$$

for a better numerical conditioning. Note that these new criteria vary in the same way as the original ones; additionally, the exponential transform has the merit that it tends to strengthen convexity; for example, $\mathcal{J}_{\mathcal{A}}$ is convex under weaker conditions that J_A itself, since:

$$\mathcal{H}_A^* = E\,H_A^* + E^2\,\nabla J_A^*\,\left(\nabla J_A^*\right)^T \tag{14}$$

where $\mathcal{H}_A^*$ and H_A^* are the Hessian matrices of $\mathcal{J}_A$ and J_A respectively, and $E = \alpha\,\exp\left[\alpha\left(J_A - J_A^*\right)\right] > 0$.

2 Splitting strategy for unconstrained problems

The Hessian matrix is real and symmetric and as such it admits a complete set of orthogonal eigenvectors $\{\,\omega^j\,\}$ $(j=1,...,n+1)$. Thus:

$$\begin{cases} H_A^* = \Omega_H\,\Lambda_H\,\Omega_H^T \\ \Lambda_H = \mathrm{Diag}(\,h_j\,) \quad \text{(real eigenvalues of the Hessian matrix)} \\ \Omega_H = \begin{pmatrix} & \vdots & \\ \dots & \omega^j & \dots \\ & \vdots & \end{pmatrix} \quad (\Omega_H\text{: orthogonal matrix of eigenvectors; } \Omega_H^T\,\Omega_H = I) \end{cases} \tag{15}$$

Consistently with the assumption that Y_A^* realizes a local or global unconstrained minimum of the functional J_A, we have:

$$\nabla J_A^* = 0 \tag{16}$$

and

$$h_j \geq 0 \quad \text{(for all } j) \tag{17}$$

Thus, for small enough ε, omitting third–order terms, we have:

$$J_A\left(Y_A^* + \varepsilon\omega\right) - J_A\left(Y_A^*\right) = \frac{\varepsilon^2}{2}\,\omega\,.\,H_A^*\,\omega + \cdots \geq 0 \tag{18}$$

and the problem stated in the introduction is equivalent to minimizing the non–negative quadratic term $\omega\,.\,H_A^*\,\omega$. This problem has a clear solution given by the classical Rayleigh–Ritz characterization of the eigenvalues of the positive semidefinite Hessian matrix H_A^* (see e.g. [11]). The territory of the secondary functional should be taken to be the span of p eigenvectors of the Hessian matrix associated with the least eigenvalues $\{\,h_j\,\}$.

In other words, after ordering of the eigenvectors in a such a way that the sequence of eigenvalues $\{\,h_j\,\}$ is monotone non–increasing, the split indicated in (7) is optimal with:

$$S = \Omega_H \tag{19}$$

3 Splitting strategy for constrained problems

We now turn our attention to constrained problems. From a computational viewpoint, we assume reliably accurate approximations of the primary–functional gradient and Hessian, as well as constraint gradients are available. This may require the auxiliary construction of meta-models. Then we propose to identify p linearly independent orthogonal vectors $\{\,\omega^j\,\}$ $(n+2-p \leq j \leq n+1$ lying in the hyperplane orthogonal to the gradient, $\left(\nabla J_A^*\right)^{\perp}$, to define the strategy of minimization of the secondary functional.

Precisely, since Y_A^* is a point of (local) minimum of the constrained functional J_A, a vector $\lambda_A^* \in \mathbb{R}^K$ of Lagrange multipliers exists such that:

$$\nabla J_A^* + {\lambda_A^*}^T\,\nabla g^* = 0 \tag{20}$$

where the superscript * indicates an evaluation at $Y = Y_A^*$. This gives explicitly:

$$\nabla J_A^* + \sum_{k=1}^{K} \lambda_k^*\,L_k = 0\,, \quad L_k = \nabla g_k^* \tag{21}$$

Then, we assume that the constraint gradient vectors $\{\,L_k\,\}$ $(1 \leq k \leq K)$ are linearly–independent. If this is not the case, we reduce the family to a maximal subfamily of linearly–independent vectors and redefine K accordingly.

Then, we apply the classical Gram–Schmidt orthogonalization process to the vectors $\{\,L_k\,\}$ $(1 \leq k \leq K)$ to define an orthonormal basis $\{\,\omega^1, \omega^2, ..., \omega^K\,\}$ of the subspace generated by them, $Sp\,(\,L_1,\,L_2,\,...,\,L_K\,)$ that is, the subspace of dimension K tangent, at $Y = Y_A^*$, to the constraint hyper-surfaces $g_k = 0$ $(1 \leq k \leq K)$:

$$\begin{cases} Set:\ \tilde{\omega}^1 = L_1\,;\ set:\ \omega^1 = \tilde{\omega}^1/\|\tilde{\omega}^1\|\,; \\ Set:\ \tilde{\omega}^2 = L_2 - c_1^2\,\omega^1\,;\ compute\ c_1^2\ so\ that\ \tilde{\omega}^2 \perp \omega^1\,;\ set:\ \omega^2 = \tilde{\omega}^2/\|\tilde{\omega}^2\|\,; \\ Set:\ \tilde{\omega}^3 = L_3 - c_1^3\,\omega^1 - c_2^3\,\omega^2\,;\ compute\ c_1^3\ and\ c_2^3\ so\ that\ \tilde{\omega}^3 \perp \omega^1\ and\ \omega^2\,;\ set:\ \omega^3 = \tilde{\omega}^3/\|\tilde{\omega}^3\|\,; \\ etc. \end{cases} \tag{22}$$

In what follows, we identify vectors of $\mathbb{R}^{n+1}$ with the $(n+1) \times 1$ column–vectors made of their components in the canonical basis. Then, for fixed k, the matrix $\omega^k\,{\omega^k}^T$ represents the axial projection along the unit vector ω^k, and the matrix

$$\boxed{P = I - \omega^1\,{\omega^1}^T - \omega^2\,{\omega^2}^T - \cdots - \omega^K\,{\omega^K}^T} \tag{23}$$

the orthogonal projection onto the subspace tangent to the constraint hyper-surfaces. Consequently, the vectors $\{\omega^k\}$ $(1 \le k \le K)$ generate the null space of P, which can easily be verified:

$$\begin{aligned} P\,\omega^1 &= \left(I - \omega^1\,{\omega^1}^T - \omega^2\,{\omega^2}^T - \cdots - \omega^K\,{\omega^K}^T\right)\omega^1 \\ &= \left(I - \omega^1\,{\omega^1}^T\right)\omega^1 \quad \text{since } \omega^1 \perp \omega^2, \omega^3, \ldots, \omega^K \\ &= \omega^1 - \omega^1\,\|\omega^1\|^2 \\ &= 0 \quad \text{since } \|\omega^1\| = 1 \end{aligned} \tag{24}$$

and similarly by permutation of symbols:

$$P\,\omega^2 = \cdots = P\,\omega^K = 0 \tag{25}$$

Now define the following matrix:

$$\boxed{H'_A = P\,H^*_A\,P} \tag{26}$$

This matrix is real–symmetric and thus admits a complete set of orthonormal eigenvectors:

$$\boxed{H'_A = \Omega\,\Lambda\,\Omega^T} \tag{27}$$

where the matrix Ω is orthogonal and the matrix Λ real and diagonal:

$$\Omega^T\,\Omega = I\,, \quad \Lambda = \mathrm{Diag}(\,h'_j\,) \tag{28}$$

The eigenvectors of the matrix H'_A are the column–vectors of the matrix Ω. These include the vectors $\omega^1, \omega^2, \ldots, \omega^K$ that belong to the null space of P and thus of H'_A. The remaining eigenvectors, $\{\omega^k\}$ $(K+1 \le k \le n+1)$ are orthogonal to them. Thus, by virtue of (21):

$$\forall k > K\,: \quad \nabla J^*_A\,.\,\omega^k = 0 \tag{29}$$

Additionally for such a k, since $\omega^k \perp \omega^1,\ \omega^2,\ \ldots,\ \omega^K$, we have trivially:

$$\omega^k = P\,\omega^k \tag{30}$$

which implies that:

$$\omega^k\,.\,H^*_A\,\omega^k = {\omega^k}^T\,H^*_A\,\omega^k = {\omega^k}^T\,\underbrace{P^T\,H^*_A\,P}_{H'_A}\,\omega^k = h'_k \tag{31}$$

and therefore:

$$\Delta J^k_A = J_A\left(Y^*_A + \varepsilon\omega^k\right) - J_A\left(Y^*_A\right) = \varepsilon\,\nabla J^*_A\,.\,\omega^k + \frac{\varepsilon^2}{2}\,\omega^k\,.\,H^*_A\,\omega^k + O(\varepsilon^3) = \frac{\varepsilon^2\,h'_k}{2} + O(\varepsilon^3) \tag{32}$$

Hence if, in the dynamic game, the secondary functional is minimized in a subspace of $Sp(\omega^{K+1}, \ldots, \omega^{n+1})$, the minimization results in a degradation of the primary functional that is only second–order in terms of the distance from the original optimal point Y^*_A. This splitting strategy is favored in the present approach.

Then several situations are possible. First, if the constraints are linear, the points $Y^*_A + \varepsilon\omega^k$ $(k > K)$ satisfy them. Then, a necessary condition for the optimality of Y^*_A is that:

$$\forall k\,, \quad h'_k \ge 0 \tag{33}$$

which is equivalent to stating that the matrix H'_A is positive semidefinite. More generally, whether the constraints are linear or not, if the matrix H'_A is positive semidefinite, the *a priori best* definition of a p–dimensional subspace as a strategy to minimize the secondary functional is clearly the subspace spanned by the p eigenvectors associated with the smallest strictly–positive eigenvalues h'_k. Again, by *a priori best* we mean *optimal blind choice* of parameterization splitting that can be made without information about the secondary functional.

Otherwise, if the constraints are nonlinear and the matrix H'_A turns out to be indefinite, we propose to apply the same strategy but after reformulating the primary problem not only by linearization of the constraints, but also, as suggested by J.P. Zolésio, by regularization of the criterion:

$$\begin{cases} \min\limits_{Y} \; J_A(Y) + \dfrac{\alpha}{2} \left\| Y - Y_A^* \right\|^2 \\ \text{Subject to: } L_k \,.\, (Y - Y_A^*) = 0 \quad (1 \le k \le K) \end{cases} \tag{34}$$

for some small but large enough positive α.

In summary, once the real–symmetric matrix H'_A is constructed and diagonalized, first retain the null–space eigenvectors $\{\,\omega^k\,\}$ $(1 \le k \le K)$ in the strategy of the primary functional. Arrange the remaining eigenvectors so that α_k decreases, and retain the tail p eigenvectors in the strategy of the secondary functional:

$$Y = Y_A^* + \sum_{k=1}^{n+1-p} u_k\, \omega^k + \sum_{j=1}^{p} v_j\, \omega^{n+2-j} \tag{35}$$

In this way, the set of parameters associated with the two functionals are respectively:

$$\begin{cases} J_A : U = \{\, u_k \,\} \quad (k = 1, \dots, n+1-p) \\ J_B : V = \{\, v_j \,\} \quad (j = 1, \dots, p) \end{cases} \tag{36}$$

Thus, the method of the previous section has been extended to constrained problems by replacing the Hessian matrix H_A^* by the matrix H'_A calculated assuming constraint gradients can be computed sufficiently accurately at completion of the primary minimization process.

4 Working out a simple case of concurrent optimization

In this section, in order to illustrate the effect of the territory–splitting strategy on the result of a concurrent optimization, we work out explicitly a simple case of minimization of two quadratic forms of four variables, subject to a linear constraint.

Let $A \ge 1$ be a positive constant and consider the following (constrained) *primary minimization* problem:

$$A : \begin{cases} \min\limits_{Y \in \mathbb{R}^4} J_A(Y) := \displaystyle\sum_{k=0}^{3} \frac{y_k^2}{A^k} \\ \text{Subject to: } \displaystyle\sum_{k=0}^{3} \left(y_k - A^k \right) = 0 \end{cases} \tag{37}$$

as well as the following (unconstrained) *secondary minimization* problem:

$$B : \min_{Y \in \mathbb{R}^4} J_B(Y) := \sum_{k=0}^{3} y_k^2 \tag{38}$$

The parameter A may be viewed as a measure of the antagonism between the two criteria.

The solution to the primary optimization problem is straightforward. Enforcing the stationarity of the *augmented Lagrangian*,

$$\mathcal{L}_A = \sum_{k=0}^{3} \left[\frac{y_k^2}{A^k} + \lambda \left(y_k - A^k \right) \right] \tag{39}$$

yields the following

$$y_k^* = -\frac{\lambda}{2} A^k \tag{40}$$

which, injected in the constraint equation, gives:

$$(1 + \frac{\lambda}{2}) \sum_{k=0}^{3} A^k = 0 \Longrightarrow \lambda^* = -2 \tag{41}$$

Thus, the optimal solution to the primary minimization problem is given by:

$$\begin{cases} y_k^* = A^k \quad (k = 0, ..., 3) \\ J_A^* = J_A(Y_A^*) = \sum_{k=0}^{3} A^k = \dfrac{A^4 - 1}{A - 1} \\ J_B^* = J_B(Y_A^*) = \sum_{k=0}^{3} A^{2k} = \dfrac{A^8 - 1}{A^2 - 1} \end{cases} \tag{42}$$

where we have also indicated the value J_B^* of the secondary criterion J_B for $Y = Y_A^*$.

Now, restricting ourselves to splitting strategies into two packets of two primitive variables $\{ y_k \}$ $(k = 0, ..., 3)$, that is, $p = 2$, a total of six different combinations or games can be identified according to the variables that are respectively allocated to the two minimization problems:

$$G_{j,k}\ (j < k) : \quad A : \{ y_j, y_k \}, \quad \text{and } B : \{ y_\ell, y_m \} (\ell \neq j, k\, ;\ m \neq j, k, \ell) \tag{43}$$

Let us calculate the Nash equilibrium point corresponding to the splitting of $G_{j,k}$. Clearly, at equilibrium:

$$\bar{y}_\ell = \bar{y}_m = 0 \tag{44}$$

Thus, conditioned to this constraint, the primary minimization problem reduces to:

$$\begin{cases} \min\limits_{y_j, y_k} \left(\dfrac{y_j^2}{A^j} + \dfrac{y_k^2}{A^k} \right) \\ \text{Subject to: } y_j + y_k = \dfrac{A^4 - 1}{A - 1} \end{cases} \tag{45}$$

Again, we have formally:

$$\bar{y}_j = -\frac{\lambda}{2} A^j, \quad \bar{y}_k = -\frac{\lambda}{2} A^k \tag{46}$$

but, injecting this in the constraint equation, now gives:

$$-\frac{\lambda}{2} \left(A^j + A^k \right) = \frac{A^4 - 1}{A - 1} \tag{47}$$

and:

$$\bar{y}_j = \frac{A^j}{A^j + A^k} \frac{A^4 - 1}{A - 1}, \quad \bar{y}_k = \frac{A^k}{A^j + A^k} \frac{A^4 - 1}{A - 1} \tag{48}$$

and correspondingly:

$$J_A^{G_{j,k}} = \frac{1}{A^j + A^k} \left(\frac{A^4 - 1}{A - 1} \right)^2, \quad J_B^{G_{j,k}} = \frac{A^{2j} + A^{2k}}{(A^j + A^k)^2} \left(\frac{A^4 - 1}{A - 1} \right)^2 \tag{49}$$

Expressing this result in terms of relative variations, we get:

$$\frac{J_A^{G_{j,k}}}{J_A^*} = \frac{1}{A^j + A^k} \frac{A^4 - 1}{A - 1}, \quad \frac{J_B^{G_{j,k}}}{J_B^*} = \frac{A^{2j} + A^{2k}}{(A^j + A^k)^2} \frac{A^4 - 1}{A - 1} \frac{A + 1}{A^4 + 1} \tag{50}$$

The case $A = 1$ corresponds to the absence of antagonism between the criteria (except from the standpoint of the constraint) and equal sensitivities of the primary criterion J_A w.r.t. all four primitive variables. For this case, all four games are equivalent and their performance is measured by the ratios:

$$\frac{J_A^{G_{j,k}}}{J_A^*} = \frac{J_B^{G_{j,k}}}{J_B^*} = 2 \tag{51}$$

This degradation in both criteria is due to the fact that two variables have been set to 0 in the *unconstrained* minimization of J_B. Hence for $A = 1$, and by extension locally also, both criteria are degraded by all these games!

Inversely, consider now the case of a large A. Recall that $j < k$ and observe that:

$$\frac{J_A^{G_{j,k}}}{J_A^*} \sim A^{3-k}\,, \quad \frac{J_B^{G_{j,k}}}{J_B^*} \sim 1 \quad (A \gg 1) \tag{52}$$

For the games $G_{0,3}$, $G_{1,3}$ and $G_{2,3}$ ($k = 3$), the relative increase in criterion J_A is measured by the ratio

$$\frac{J_A^{G_{j,3}}}{J_A^*} = \frac{1 + A + A^2 + A^3}{A^j + A^3} = 1 + \frac{1 + A + A^2 - A^j}{A^3 + A^j} \quad (j = 0, 1, 2) \tag{53}$$

which is greater than 1 for all A but tends to 1 as $A \to \infty$, and is thus bounded (as A varies). For the games $G_{0,2}$ and $G_{1,2}$ ($k = 2$), the relative increase exhibits a presumably–acceptable asymptotic linear growth with A. Finally, for the game $G_{0,1}$ ($k = 1$), the relative increase is asymptotically quadratic. These trends are depicted on Fig. 1 (*top*).

Now, concerning J_B, observe that it is proportional to the following monotone–increasing function of the variable $\eta = A^{k-j}$:

$$\frac{1 + \eta^2}{(1 + \eta)^2} \tag{54}$$

Thus, w.r.t. J_B, the performance of the games $G_{0,1}$, $G_{1,2}$ and $G_{2,3}$ ($k - j = 1$) is the same and optimal (least value of η); the games $G_{0,2}$ and $G_{1,3}$ ($k - j = 2$) are equivalent and second–best, and the game $G_{0,3}$ corresponds to the least effective strategy. The corresponding trends are depicted on Fig. 1 (*bottom*).

Hence, the best strategy w.r.t. both criteria is the game $G_{2,3}$. But

$$\frac{J_B^{G_{2,3}}}{J_B^*} > 1\,, \quad \forall A \tag{55}$$

and all the games considered so far fail to reduce the secondary criterion.

Thus, we conclude to the failure of all the games based on splitting strategies of the primitive variables $\{\, y_k \,\}$ into two packets of two. All operational points perform more poorly w.r.t. both criteria than the absolute optimum of the primary optimization problem. For one game, $G_{0,1}$, particularly inappropriate, the equilibrium point corresponds to an important relative increase in the primary criterion that grows quadratically with the antagonism parameter A. Even more deceivingly, none of the games succeeds to reduce the secondary criterion. These dramatic conclusions should however be tampered by the observation that, with such a small number of variables, too little latitude is given to improve the performance of the initial point.

Then we develop the proposed optimal split. We calculate the gradient ∇J_A, the Hessian H_A^*, the constraint gradient ∇g and the unit vector ω^1:

$$\nabla J_A = \begin{pmatrix} \frac{2y_0}{1} \\ \frac{2y_1}{A} \\ \frac{2y_2}{A^2} \\ \frac{2y_3}{A^3} \end{pmatrix}, \quad H_A^* = \begin{pmatrix} 2 & & & \\ & \frac{2}{A} & & \\ & & \frac{2}{A^2} & \\ & & & \frac{2}{A^3} \end{pmatrix}, \quad \nabla g = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \omega^1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \tag{56}$$

Consequently, the symmetric matrix H'_A is easily calculated:

$$\begin{aligned} H'_A &= \left(I - \omega^1\, {\omega^1}^T\right) H_A^* \left(I - \omega^1\, {\omega^1}^T\right) \\ &= \frac{1}{8A^3} \begin{pmatrix} 9A^3 + A^2 + A + 1 & . & . & . \\ -3A^3 - 3A^2 + A + 1 & A^3 + 9A^2 + A + 1 & . & . \\ -3A^3 + A^2 - 3A + 1 & A^3 - 3A^2 - 3A + 1 & A^3 + A^2 + 9A + 1 & . \\ -3A^3 + A^2 + A - 3 & A^3 - 3A^2 + A - 3 & A^3 + A^2 - 3A - 3 & A^3 + A^2 + A + 9 \end{pmatrix} \end{aligned} \tag{57}$$

where dots stand for omitted entries known by symmetry. Note that the row sums (or column sums) of this matrix are all equal to 0 since the vector ω^1 belongs to (and generates) the kernel. Also note that since the matrix H_A^* is positive definite, the matrix H'_A is positive semidefinite, as it must be when the constraint is linear.

The matrix H'_A is then diagonalized:

$$H'_A = \Omega\,\Lambda\,\Omega^T\,, \quad \Omega = \{\,\omega_{j,k}\,\}\,, \quad \Lambda = \mathrm{Diag}(\,h'_j\,) \quad (h'_1 = 0\,;\, h'_2 \geq h'_3 \geq h'_4 > 0) \tag{58}$$

in which the eigenvectors, or column–vectors of matrix Ω, have been arranged by placing ω^1 first, and the remaining three in decreasing order of the associated eigenvalue h'_j. Thus, in particular, the first column–vector is the vector ω^1:

$$\omega_{1,1} = \omega_{2,1} = \omega_{3,1} = \omega_{4,1} = \tfrac{1}{2} \tag{59}$$

Consistently with the notation introduced previously, for a *"2+2 split of territory"*, we define two unknown vectors of dimension 2:

$$U = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad V = \begin{pmatrix} v_2 \\ v_1 \end{pmatrix} \tag{60}$$

and parameterize the unknown vector Y according to (7):

$$\begin{cases} y_0 = 1 + \omega_{1,1}\,u_1 + \omega_{1,2}\,u_2 + \omega_{1,3}\,v_2 + \omega_{1,4}\,v_1 \\ y_1 = A + \omega_{2,1}\,u_1 + \omega_{2,2}\,u_2 + \omega_{2,3}\,v_2 + \omega_{2,4}\,v_1 \\ y_2 = A^2 + \omega_{3,1}\,u_1 + \omega_{3,2}\,u_2 + \omega_{3,3}\,v_2 + \omega_{3,4}\,v_1 \\ y_3 = A^3 + \omega_{4,1}\,u_1 + \omega_{4,2}\,u_2 + \omega_{4,3}\,v_2 + \omega_{4,4}\,v_1 \end{cases} \tag{61}$$

In this *perturbation formulation*, the equality constraint imposed on the minimization of J_A is expressed by the following homogeneous condition:

$$S_1\,u_1 + S_2\,u_2 + S_3\,v_2 + S_4\,v_1 = 0 \tag{62}$$

in which:

$$S_j = \sum_{i=1}^{4} \omega_{i,j} \tag{63}$$

But since $\omega_{i,1} = \frac{1}{2}$ (for all i),

$$S_1 = 2 \tag{64}$$

whereas, for $j > 1$:

$$S_j = 2\,\omega^1 \,.\, \omega^j = 0 \tag{65}$$

by orthogonality of the eigenvectors. Hence:

$$u_1 = 0 \tag{66}$$

and (61) is simplified as follows:

$$\begin{cases} y_0 = 1 + \omega_{1,2}\,u_2 + \omega_{1,3}\,v_2 + \omega_{1,4}\,v_1 = 1 + y'_0 \\ y_1 = A + \omega_{2,2}\,u_2 + \omega_{2,3}\,v_2 + \omega_{2,4}\,v_1 = A + y'_1 \\ y_2 = A^2 + \omega_{3,2}\,u_2 + \omega_{3,3}\,v_2 + \omega_{3,4}\,v_1 = A^2 + y'_2 \\ y_3 = A^3 + \omega_{4,2}\,u_2 + \omega_{4,3}\,v_2 + \omega_{4,4}\,v_1 = A^3 + y'_3 \end{cases} \tag{67}$$

Then, we obtain a linear equation for u_2 by expressing that the criterion J_A should be stationary w.r.t. u_2 for fixed v_1, v_2:

$$\frac{\partial J_A}{\partial u_2} = \sum_{k=0}^{3} \frac{2y_k}{A^k}\,\frac{\partial y_k}{\partial u_2} = 2\sum_{k=0}^{3} \frac{y_k}{A^k}\,\omega_{k+1,2} = 0 \tag{68}$$

Now, since

$$y_k = A^k + \omega_{k+1,2}\,u_2 + \omega_{k+1,3}\,v_2 + \omega_{k+1,4}\,v_1 \tag{69}$$

we get:

$$c\,u_2 + d = 0 \tag{70}$$

where:

$$c = \sum_{k=0}^{3} \frac{\omega_{k+1,2}^2}{A^k} \tag{71}$$

and

$$d = \sum_{k=0}^{3} \frac{\omega_{k+1,2}}{A^k} \left(A^k + \omega_{k+1,3}\, v_2 + \omega_{k+1,4}\, v_1 \right) = S_2 + \sigma_{2,3}\, v_2 + \sigma_{2,4}\, v_1 \tag{72}$$

where $S_2 = 0$, and:

$$\sigma_{i,j} = \sum_{k=0}^{3} \frac{\omega_{k+1,i}\, \omega_{k+1,j}}{A^k} \tag{73}$$

This gives:

$$u_2 = -\frac{d}{c} = -\frac{\sigma_{2,3}\, v_2 + \sigma_{2,4}\, v_1}{\sigma_{2,2}} \tag{74}$$

Finally we calculate v_1 and v_2 by expressing the stationarity of the second criterion, J_B, for fixed $u_1 = 0$, and u_2. Since

$$J_B = \sum_{k=0}^{3} \left[A^k + \sum_{j=0}^{3} \omega_{k+1,j+1}\, y'_j \right]^2 \tag{75}$$

the condition is:

$$\frac{\partial J_B}{\partial y'_\ell} = 2 \sum_{k=0}^{3} \left[A^k + \sum_{j=0}^{3} \omega_{k+1,j+1}\, y'_j \right] \omega_{k+1,\ell+1} = 0 \quad (\ell = 2, 3) \tag{76}$$

Hence, the linear equation for $\ell = 3$ is:

$$\alpha\, v_1 + \beta\, v_2 + \gamma = 0 \tag{77}$$

where:

$$\alpha = \sum_{k=0}^{3} \omega_{k+1,4}^2 = 1 \tag{78}$$

by normalization of the eigenvector,

$$\beta = \sum_{k=0}^{3} \omega_{k+1,3}\, \omega_{k+1,4} = 0 \tag{79}$$

by orthogonality of the eigenvectors, and:

$$\gamma = \sum_{k=0}^{3} \left[A^k + \sum_{j=0}^{1} \omega_{k+1,j+1} y'_j \right] \omega_{k+1,4} = \sum_{k=0}^{3} A^k\, \omega_{k+1,4} \tag{80}$$

where the orthogonality of the eigenvectors has again been used. By proceeding in the same way for the second equation ($\ell = 2$), we get:

$$\begin{cases} v_1 = -\displaystyle\sum_{k=0}^{3} A^k\, \omega_{k+1,4} \\ v_2 = -\displaystyle\sum_{k=0}^{3} A^k\, \omega_{k+1,3} \end{cases} \tag{81}$$

and again:

$$\begin{cases} u_1 = 0 \\ u_2 = -\dfrac{\sigma_{2,3}\, v_2 + \sigma_{2,4}\, v_1}{\sigma_{2,2}} \end{cases} \tag{82}$$

where:

$$\sigma_{i,j} = \sum_{k=0}^{3} \frac{\omega_{k+1,i}\, \omega_{k+1,j}}{A^k} \tag{83}$$

From this quasi–explicit solution, two observations can readily be made.

First, the optimal solution is such that $u_1 = 0$, that is, the component of the state vector along the constraint gradient is preserved.

Second, in the case of a vanishing antagonism, $A = 1$:

$$v_1 = -S_4 = v_2 = -S_3 = 0 \tag{84}$$

and hence also:

$$u_1 = u_2 = 0 \tag{85}$$

that is:

$$\bar{Y} = Y_A^* \tag{86}$$

Thus, contrasting with the situation of the previous splittings, here, the optimal solution of the primary minimization problem is preserved when the criteria are equal, and this constitutes a definite advantage of this new formulation. This property holds in a more general context for which it is established separately in the next section.

For values of the parameter A greater than 1, we have calculated numerically the matrices H'_A, Ω and Λ, the above solution U, V, and its performance in terms of the ratios J_A/J_A^* and J_B/J_B^*. These new results have also been collected on FIG. 1, on top for J_A, and on the bottom for J_B.

We observe that, contrasting with the previous strategies based on splits of the primitive variables ($G_{j,k}$), the strategy based on the split of the eigenvectors does succeed to diminish the secondary criterion J_B, despite the reduced number of parameters permitting to organize the concurrent game. This is realized, of course at the expense of some increase of the primary criterion J_A, whose growth with the antagonism parameter seems to be asymptotically linear.

Additionally, the figure provides evidence that for the proposed strategy based on the eigensplit, and for this strategy only, the derivatives of both criteria J_A at the equilibrium point w.r.t. the parameter A are equal to zero at $A = 1$ (case of vanishing antagonism). To explain this, let

$$A = 1 + \varepsilon \tag{87}$$

Suppose ε sufficiently small and let $\bar{Y}_\varepsilon$ be an equilibrium point in the neighborhood of the optimum. Since $u_1 = 0$, the perturbation in design vector, $\bar{Y}_\varepsilon - Y_A^*$ is orthogonal to ω^1 which is the unit vector along ∇g, which is itself proportional to ∇J_A^* at the point $Y = Y_A^*$ of optimality of the constrained primary criterion J_A. Hence, $J_A(\bar{Y}_\varepsilon) - J_A^* = O(\varepsilon^2)$, which explains the result for J_A. Besides, since J_B and J_A functionally identify at $A = 1$, their gradients also do, and the result is also true for J_B. This demonstrates the *robustness* of the equilibrium point w.r.t. the parameter A at the initial equilibrium point. This behavior is due to a general result established in the next section.

Besides, we have examined how does this Nash–equilibrium point compare in the reduction of the secondary criterion J_B with the absolute minimum of this criterion attainable under the constraint of equal value of the primary criterion J_A, that is, at the corresponding point on the Pareto–equilibrium front defined by:

$$\begin{cases} \min\limits_{Y \in \mathbb{R}^{n+1}} J_B(Y) \\ \text{Subject to: } g = 0\,,\ J_A = \bar{J}_A \end{cases} \tag{88}$$

To solve this problem we introduce the augmented Lagrangian

$$\mathcal{L}(Y, \lambda, \mu) = J_B + \lambda\, g + \mu\, (J_A - \bar{J}_A) \tag{89}$$

Expressing the stationarity of $\mathcal{L}(Y, \lambda, \mu)$ w.r.t. its three arguments yields the following representation of the state vector that realizes the Pareto–optimal point:

$$y_k = -\frac{\dfrac{\lambda}{2}}{1 + \dfrac{\mu}{A^k}} \tag{90}$$

Injecting this result into the constraint equation gives:

$$\frac{\lambda}{2} = -\frac{\dfrac{A^4 - 1}{A - 1}}{\sum_{k=0}^{3} \dfrac{1}{1 + \dfrac{\mu}{A^k}}} \tag{91}$$

so that:

$$y_k = \frac{\dfrac{1}{1+\dfrac{\mu}{A^k}}}{\sum_{j=0}^{3} \dfrac{1}{1+\dfrac{\mu}{A^j}}} \, \frac{A^4-1}{A-1} \tag{92}$$

Lastly, the above expression is injected into the additional constraint

$$J_A = \sum_{k=0}^{3} \frac{y_k^2}{A^k} = \bar{J}_A \tag{93}$$

where $\bar{J}_A$ is specified. This permits us to solve for the unknown Lagrange multiplier μ and identify the Pareto–optimal point. This results in a polynomial equation that has been solved numerically by an *ad hoc* Newton's method. The corresponding value of the secondary criterion J_B is indicated on FIG. 1 (bottom) for a range of values of the antagonism parameter A. From the figure it appears that if the performance of the Pareto–optimal point is slightly superior, the curves drawn by the Nash–equilibrium point and the Pareto–optimal point as A varies are close to each other and follow similar trends.

Lastly, unsymmetrical splits of variables have also been analyzed; see Appendix A for 3+1 splittings, and Appendix B for 1+3 splittings. In all cases, qualitatively very similar results have been obtained.

5 Continuum of Nash equilibrium points

In this section, we generalize a number of results observed in the particular case of the four-dimensional problem of the previous section, to a more general setting.

We first consider the particular situation of identical criteria ($J_B = J_A = J$) for which the only difference between the primary and the secondary problems lies in the application of K scalar equality constraints,

$$g = 0 \tag{94}$$

in the primary problem only. Again, we propose the split of territories according to the vectors U and V of (7) with $S = \Omega$ as in (35):

$$Y = Y_A^* + \Omega \begin{pmatrix} U \\ 0 \end{pmatrix} + \Omega \begin{pmatrix} 0 \\ V \end{pmatrix} \tag{95}$$

in which "0" stands for a column–vector of zeroes of adequate dimension.

By assumption, Y_A^* is one, or the unique optimal solution to the full–dimension primary optimization problem, and thus solves the equations:

$$\frac{\partial}{\partial Y}\left(J + \lambda^T g\right) = 0\,, \quad g = 0 \tag{96}$$

for some $\lambda \in \mathbb{R}^K$, say λ_A^*. Here, $\frac{\partial}{\partial Y}\left(J + \lambda^T g\right)$ is understood as a row vector of dimension $1 \times (n+1)$, transpose of the gradient.

In the concurrent game, considering the form given to the split of parameters, the primary optimization problem of reduced dimension is:

$$A : \begin{cases} \min\limits_{U \in \mathbb{R}^{n+1-p}} J(Y) \\ \text{Subject to: } g = 0 \end{cases} \tag{97}$$

whereas the secondary problem is the following unconstrained minimization on V:

$$B : \min_{V \in \mathbb{R}^p} J(Y) \tag{98}$$

Thus, assuming now that a strict inequality holds in (9), the equations defining the Nash equilibrium are:

$$\left(\frac{\partial}{\partial U}\right)_V \left(J + \mu^T g\right) = 0 \quad \text{for some } \mu \in \mathbb{R}^K\,, \quad g = 0 \tag{99}$$

$$\left(\frac{\partial}{\partial V}\right)_U J = 0 \tag{100}$$

where the subscript U or V over a closing parenthesis indicates the variable held fixed in the partial differentiation. Thus, $\left(\frac{\partial}{\partial U}\right)_V \left(J + \mu^T g\right)$ and $\left(\frac{\partial}{\partial V}\right)_U$ are row vectors of dimension $1 \times (n+1-p)$ and $1 \times p$ respectively.

In this setting, does the optimum solution Y_A^*, equivalently defined by $U^* = V^* = 0$, correspond to an equilibrium point for the game? In other words: is the solution Y_A^* (associated with λ_A^*) of (96) also a solution of (99)–(100)?

First, observe that variations in V alone about the optimal point $U = U^* = 0$ and $V = V^* = 0$, cause variations in Y in the span of the *tail p eigenvectors*. But these, by construction, are orthogonal to the span of the K constraint gradients which contains the gradient of the criterion at $Y = Y_A^*$, $\nabla J^* = \nabla J_A^*$. Therefore:

$$\left(\frac{\partial J}{\partial V}\right)_U \delta V = 0\,, \quad \forall \delta V \in \mathbb{R}^p \tag{101}$$

at $U = U^* = 0$ and $V = V^* = 0$, which proves that (100) holds.

Second, by virtue of (96):

$$\left(\frac{\partial}{\partial U}\right)_V J = \frac{\partial J}{\partial Y}\left(\frac{\partial Y}{\partial U}\right)_V = -\lambda_A^{*\,T}\frac{\partial g}{\partial Y}\left(\frac{\partial Y}{\partial U}\right)_V = -\lambda_A^{*\,T}\left(\frac{\partial g}{\partial U}\right)_V = -\left(\frac{\partial}{\partial U}\right)_V \left(\lambda_A^{*\,T} g\right) \tag{102}$$

where the Jacobian matrix $\left(\frac{\partial Y}{\partial U}\right)_V$ is $(n+1)\times(n+1-p)$. Thus (99) is also satisfied with $Y = Y_A^*$ and $\mu = \lambda_A^*$. Hence, the optimum solution Y_A^* (associated with λ_A^*) also satisfies the equilibrium equations, and we conclude to the existence of a Nash–equilibrium point:

$$\bar{Y} = Y_A^* \tag{103}$$

Remark 1

In his doctoral thesis, Wang [3] has implemented certain virtual games *in aerodynamics in the context of parallel computing where a split of geometrical variables is introduced naturally, although all the variables cooperate to minimize the same criterion. He considered in particular the optimization of an aircraft high–lift devices in shape and position concurrently to maximize lift. Thus, practical situations in which $J_B = J_A$ do exist. However, we consider this situation here as a theoretical step to initiate a continuum of Nash equilibrium points.*

Although quite simple, the above result is not trivial, since, for example, games based on the split of the primitive variables usually do not satisfy this *consistency* property, as illustrated by the particular example of the previous section.

We now examine certain algorithmic questions. In a concurrent optimization formulated as several minimization subproblems, each one of which being subject to its own subset of constraints, and iterated on using only a subset of variables, it would be ideal to operate in a working space accounting for the whole set of constraints, and to let the concurrent optimization be organized in a dynamic game solely in terms of the criteria being minimized. However, this is not realistic, because the constraints as well as the criteria often arise, in practical problems, from different physics usually governed by independent PDE's possibly defined over three–dimensional domains and thus costly to evaluate. However, a practical formulation could be to replace in the treatment of one particular subproblem, all the functionals and constraints related to the other subproblems, by adequate approximate *meta-models*.

Although retaining the above possibility of using meta-models, note that (103) offers an alternative possibility since it indicates that it is not strictly necessary to apply the constraint related to the primary problem to the secondary problem. Indeed, by this result, we anticipate a continuity of behavior for standard cases in which $J_B \neq J_A$, that is: if the antagonism between the criteria is small or introduced smoothly, then games based on such compatible splits are expected to define equilibrium points that deviate from the originally–optimal point by small perturbations only, and this, even if the constraint $g = 0$ is not explicitly enforced in the formulation of the secondary problem. In particular, for strongly–antagonistic criteria, we propose to organize the concurrent optimization by considering simultaneously the following two problems:

$$\begin{cases} \min\ J_A \\ \text{Subject to: } g = 0 \end{cases} \quad \text{and} \quad \begin{cases} \min\ J_{AB} := \dfrac{J_A}{J_A^*} + \varepsilon\left(\theta\,\dfrac{J_B}{J_B^*} - \dfrac{J_A}{J_A^*}\right) \\ \text{Subject to: no constraints} \end{cases} \tag{104}$$

where we now assume that both criteria are positive and that:

$$J_A^* = J_A\left(Y_A^*\right) > 0\,, \quad J_B^* = J_B\left(Y_A^*\right) > 0 \tag{105}$$

Here, ε is a small parameter permitting, as it is increased progressively, to let the formulation smoothly evolve from a trivial concurrent problem in which the two criteria are equivalent ($\varepsilon = 0$), to the original setting ($\varepsilon = 1$), and we expect to be able to follow the Nash equilibrium point from the initial optimal point Y_A^* on by continuity.

The parameter θ is strictly positive. If not set equal to 1, it is meant to introduce under-relaxation if $\theta < 1$, or over-relaxation if $\theta > 1$ in the iterative realization of the Nash equilibrium. This algorithmic aspect is not treated in this report. It can easily be shown that the Nash equilibrium point depends only upon the self-similarity parameter:

$$\varepsilon_1 = \frac{\varepsilon\theta}{1 + \varepsilon(\theta - 1)} \tag{106}$$

Technically, one would consider a positive sequence $\{\varepsilon_n\}$, with $\varepsilon_0 = 0$, and devise an iteration in which the Nash equilibrium point associated with the above concurrent formulation, say $\bar{Y}_{\varepsilon_n}$, would be determined successively for $\varepsilon = \varepsilon_n$ ($n = 0, 1, 2, ...$). Letting $\bar{Y}_0 = Y_A^*$ (which is an equilibrium point for $\varepsilon = \varepsilon_0 = 0$ by virtue of (103), at step n, the new equilibrium point $\bar{Y}_{n+1}$ would be found by the simulation of a dynamic game initiated from $Y = \bar{Y}_{\varepsilon_n}$. In this way, one would expect to get discrete values $\bar{Y}_{\varepsilon_n}$ of a parameterized path $\bar{Y}_\varepsilon$ to which one could associate the functions

$$j_A(\varepsilon) := J_A\left(\bar{Y}_\varepsilon\right), \quad j_{AB}(\varepsilon) := J_{AB}\left(\bar{Y}_\varepsilon\right) \tag{107}$$

expected to be respectively monotone increasing and monotone decreasing functions of ε, according to trends similar to those in Fig. 1, or Fig. 2, or Fig. 3. From these functions, the designer could decide which value of ε is adequate, that is, which *trade–off* between the concurrent criteria is acceptable.

Remark 2
If the calculation of the criterion J_A is computationally costly, assume we dispose of a meta-model $\tilde{J}_A$ *for J_A that can be evaluated economically. Then, we would first apply the above technique to the meta-model itself, substituting $\tilde{J}_A$ to J_B, up to $\varepsilon = 1$. From this equilibrium point, the technique would then be applied to the following modified formulation:*

$$\begin{cases} \min J_A \\ \textit{Subject to: } g = 0 \end{cases} \quad \textit{and} \quad \begin{cases} \min J_{AB} := \dfrac{\tilde{J}_A}{\tilde{J}_A^*} + \varepsilon\left(\theta\,\dfrac{J_B}{J_B^*} - \dfrac{\tilde{J}_A}{\tilde{J}_A^*}\right) \\ \textit{Subject to: no constraints} \end{cases} \tag{108}$$

We now return to theoretical considerations making first the additional assumption of linear equality constraints:

$$g_k = L_k \,.\, Y - b_k = L_k \,.\, \left(Y - Y_A^*\right) = 0 \quad (1 \le k \le K) \tag{109}$$

where the constraint gradient vectors $\{L_k\}$ $(1 \le k \le K)$ are again assumed to be linearly–independent. If this is not the case, we reduce the family $\{L_1, L_2, ... L_K\}$ to a maximal subfamily of linearly–independent vectors and redefine K accordingly.

We consider again the concurrent problem defined, as ε varies by (104), and we assume that a continuum of Nash–equilibrium points $\bar{Y}_\varepsilon$ exists originating from $\bar{Y}_0 = Y_A^*$ at $\varepsilon = 0$:

$$\bar{Y}_\varepsilon = Y_A^* + \sum_{j=1}^{n+1-p} u_j(\varepsilon)\,\omega^j + \sum_{j=1}^{p} v_j(\varepsilon)\,\omega^{n+2-j} \tag{110}$$

The constraint equations thus write:

$$< L_k, \sum_{j=1}^{n+1-p} u_j(\varepsilon)\,\omega^j + \sum_{j=1}^{p} v_j(\varepsilon)\,\omega^{n+2-j} > = 0 \quad (1 \le k \le K) \tag{111}$$

But, by construction, the basis $\{\omega^j\}$ $(1 \le j \le n+1)$ is orthogonal, and the gradient L_k belongs to the span of the first K basis vectors ($n + 1 - p \ge K$); hence, the above reduces to:

$$< L_k, \sum_{j=1}^{K} u_j(\varepsilon)\,\omega^j > = 0 \quad (1 \le k \le K) \tag{112}$$

which constitutes a homogeneous linear system of K equations for the K unknowns $\{u_j(\varepsilon)\}$ ($1 \le j \le K$). This system is invertible since the constraint gradients $\{L_k\}$ ($1 \le k \le K$) are linearly–independent. Therefore, along the assumed–continuous path of Nash–equilibrium points, we have identically:

$$u_1(\varepsilon) = u_2(\varepsilon) = \ldots = u_K(\varepsilon) = 0 \tag{113}$$

as observed in the example of Section 4 for which a single linear constraint was imposed ($K = 1$). Consequently:

$$\nabla J_A^* \,.\, \left(\bar{Y}_\varepsilon - Y_A^*\right) = 0 \tag{114}$$

and we conclude that:

$$j_A'(0) = 0\,, \quad J_A\left(\bar{Y}_\varepsilon\right) - J_A^* = O(\varepsilon^2) \tag{115}$$

Concerning the function $j_{AB}(\varepsilon)$, note that J_{AB} is functionally proportional to J_A at $\varepsilon = 0$, which implies that ∇J_{AB} and ∇J_A^* are also proportional; hence, we have also:

$$\nabla J_{AB}^* \,.\, \left(\bar{Y}_\varepsilon - Y_A^*\right) = 0 \tag{116}$$

However, J_{AB} contains a term that depends explicitly on ε; thus:

$$j_{AB}'(0) = \theta\, \frac{J_B\left(\bar{Y}_0\right)}{J_B^*} - \frac{J_A\left(\bar{Y}_0\right)}{J_A^*} = \theta - 1 \le 0 \tag{117}$$

We now extend this result to the case of nonlinear constraints. Suppose we modify (104) by replacing the condition $g = 0$ by a linear expansion of g about Y_A^* as in (34). Suppose that this results in the definition of a new continuum of Nash–equilibrium points, $\bar{Y}_\varepsilon^L$. Then the above result applies to $\bar{Y}_\varepsilon^L$:

$$J_A\left(\bar{Y}_\varepsilon^L\right) = J_A^* + O(\varepsilon^2) \tag{118}$$

Let

$$\bar{Y}_\varepsilon - \bar{Y}_\varepsilon^L = v + w \tag{119}$$

where $v \in Sp\,(L_1, L_2, \ldots, L_K)$ and $w \in Sp\,(L_1, L_2, \ldots, L_K)^\perp$. Assuming local regularity and smoothness of the hyper-surfaces $g_k = 0$, we have:

$$v = O(\varepsilon)\,, \quad w = O(\varepsilon^2) \tag{120}$$

Then:

$$\begin{aligned} J_A\left(\bar{Y}_\varepsilon\right) &= J_A\left(\bar{Y}_\varepsilon^L + v + w\right) \\ &= J_A\left(\bar{Y}_\varepsilon^L\right) + \nabla J_A\left(\bar{Y}_\varepsilon^L\right) . (v + w) + O(\varepsilon^2) \\ &= J_A\left(\bar{Y}_\varepsilon^L\right) + \nabla J_A^* \,.\, (v + w) + O(\varepsilon^2) \quad \text{provided } \nabla J_A \text{ is smooth} \\ &= J_A\left(\bar{Y}_\varepsilon^L\right) + O(\varepsilon^2) \quad \text{since } \nabla J_A^* \,.\, v = 0 \text{ and } \nabla J_A^* \,.\, w = O(\varepsilon^2) \\ &= J_A^* + O(\varepsilon^2) \end{aligned} \tag{121}$$

We conclude that (135) still holds, and we summarize our theoretical results by the following

Theorem 1
Let n, p and K be positive integers such that:

$$1 \le p \le n\,, \quad 0 \le K < n + 1 - p \tag{122}$$

Let J_A, J_B and, if $K \ge 1$, $\{g_k\}$ ($1 \le k \le K$) be $K + 2$ smooth real–valued functions of the vector $Y \in \mathbb{R}^{n+1}$. Assume that J_A and J_B are positive, and consider the following primary optimization problem,

$$\min_{Y \in \mathbb{R}^{n+1}} J_A(Y) \tag{123}$$

that is either unconstrained ($K = 0$), or subject to the following K equality constraints:

$$g(Y) = \left(g_1, g_2, \ldots, g_K\right)^T = 0 \tag{124}$$

Assume that the above minimization problem admits a local or global solution at a point $Y_A^* \in \mathbb{R}^{n+1}$ *at which* $J_A^* = J_A\left(Y_A^*\right) > 0$ *and* $J_B^* = J_B\left(Y_A^*\right) > 0$*, and let* H_A^* *denote the Hessian matrix of the criterion* J_A *at* $Y = Y_A^*$.

If $K = 0$*, let* $P = I$ *and* $H_A' = H_A^*$*; otherwise, assume that the constraint gradients,* $\{\nabla g_k^*\}$ $(1 \le k \le K)$*, are linearly independent and apply the Gram–Schmidt orthogonalization process to the constraint gradients, and let* $\{\omega^k\}$ $(1 \le k \le K)$ *be the resulting orthonormal vectors. Let* P *be the matrix associated with the projection operator onto the* K*–dimensional subspace tangent to the hyper–surfaces* $g_k = 0$ $(1 \le k \le K)$ *at* $Y = Y_A^*$,

$$P = I - \omega^1 {\omega^1}^T - \omega^2 {\omega^2}^T - \cdots - \omega^K {\omega^K}^T \tag{125}$$

and consider the following real–symmetric matrix:

$$H_A' = P\,H_A^*\,P \tag{126}$$

Let Ω *be an orthogonal matrix whose column-vectors are normalized eigenvectors of the matrix* H_A' *organized in such a way that the first* K *are precisely* $\{\omega^k\}$ $(1 \le k \le K)$*, and the subsequent* $n + 1 - K$ *are arranged by decreasing order of the eigenvalue*

$$h_k' = \omega^k \,.\, H_A'\,\omega^k = \omega^k \,.\, H_A^*\,\omega^k \quad (K+1 \le k \le n+1) \tag{127}$$

Consider the splitting of parameters defined by:

$$Y = Y_A^* + \Omega \begin{pmatrix} U \\ V \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_{n+1-p} \end{pmatrix}, \quad V = \begin{pmatrix} v_p \\ \vdots \\ v_1 \end{pmatrix} \tag{128}$$

Let ε *be a small positive parameter* $(0 \le \varepsilon \le 1)$*, and let* $\bar{Y}_\varepsilon$ *denote the Nash equilibrium point associated with the concurrent optimization problem:*

$$\begin{cases} \min\limits_{U \in \mathbb{R}^{n+1-p}} J_A \\ \textit{Subject to: } g = 0 \end{cases} \quad \textit{and} \quad \begin{cases} \min\limits_{V \in \mathbb{R}^p} J_{AB} \\ \textit{Subject to: no constraints} \end{cases} \tag{129}$$

in which again the constraint $g = 0$ *is not considered when* $K = 0$*, and*

$$J_{AB} := \frac{J_A}{J_A^*} + \varepsilon \left(\theta \frac{J_B}{J_B^*} - \frac{J_A}{J_A^*} \right) \tag{130}$$

where θ *is a strictly-positive relaxation parameter* ($\theta < 1$*: under-relaxation;* $\theta > 1$*: over-relaxation).*

Then:

- *[Optimality of orthogonal decomposition] If the matrix* H_A' *is positive semi-definite, which is the case in particular if the primary problem is unconstrained* ($K = 0$)*, or if it is subject to linear equality constraints, its eigenvalues have the following structure:*

$$h_1' = h_2' = \ldots = h_K' = 0 \quad h_{K+1}' \ge h_{K+2}' \ge \ldots \ge h_{n+1}' \ge 0 \tag{131}$$

  and the tail associated eigenvectors $\{\omega^k\}$ $(K+1 \le k \le n+1)$ *have the following variational characterization:*

$$\begin{aligned} \omega^{n+1} &= \text{Argmin}_\omega \left|\omega \,.\, H_A^*\,\omega\right| \quad \textit{s.t. } \|\omega\| = 1 \textit{ and } \omega \perp \left\{\omega^1, \omega^2, \ldots, \omega^K\right\} \\ \omega^n &= \text{Argmin}_\omega \left|\omega \,.\, H_A^*\,\omega\right| \quad \textit{s.t. } \|\omega\| = 1 \textit{ and } \omega \perp \left\{\omega^1, \omega^2, \ldots, \omega^K, \omega^{n+1}\right\} \\ \omega^{n-1} &= \text{Argmin}_\omega \left|\omega \,.\, H_A^*\,\omega\right| \quad \textit{s.t. } \|\omega\| = 1 \textit{ and } \omega \perp \left\{\omega^1, \omega^2, \ldots, \omega^K, \omega^{n+1}, \omega^n\right\} \\ &\vdots \end{aligned} \tag{132}$$

- *[Preservation of optimum point as a Nash equilibrium] For* $\varepsilon = 0$*, a Nash equilibrium point exists and it is:*

$$\bar{Y}_0 = Y_A^* \tag{133}$$

- *[Robustness of original design] If the Nash equilibrium point exists for $\varepsilon > 0$ and sufficiently small, and if it depends smoothly on this parameter, the functions:*

$$j_A(\varepsilon) = J_A\left(\bar{Y}_\varepsilon\right), \quad j_{AB}(\varepsilon) = J_{AB}\left(\bar{Y}_\varepsilon\right) \tag{134}$$

are such that:

$$j'_A(0) = 0 \tag{135}$$
$$j'_{AB}(0) = \theta - 1 \tag{136}$$

and

$$j_A(\varepsilon) = J_A^* + O(\varepsilon^2) \tag{137}$$
$$j_{AB}(\varepsilon) = 1 + (\theta - 1)\,\varepsilon + O(\varepsilon^2) \tag{138}$$

- *In case of linear equality constraints, the Nash equilibrium point satisfies identically:*

$$u_k(\varepsilon) = 0 \quad (1 \le k \le K) \tag{139}$$
$$\overline{Y}_\varepsilon = Y_A^* + \sum_{k=K+1}^{n+1-p} u_k(\varepsilon)\,\omega^k + \sum_{j=1}^{p} v_j(\varepsilon)\,\omega^{n+2-j} \tag{140}$$

- *For $K = 1$ and $p = n$, the Nash equilibrium point* $\overline{Y}_\varepsilon$ is Pareto optimal.

Thus for under-relaxation ($\theta < 1$), it is guaranteed that along the continuum of Nash equilibrium points, the criterion J_{AB} initially decreases as ε increases. However, this conclusion is not sufficient to guarantee the same behavior for the criterion J_B.

The very last conclusion of the above theorem concerning the Pareto-optimality in a special case will be established in the next section. Prior to this, in order to illustrate the case of a nonlinear constraint, we have calculated, in a formulation only slightly different from the above, a variant of the simple problem treated in Section 4, obtained by modifying the constraint equation to be the following nonlinear condition:

$$g = \sqrt[10]{y_0^4\, y_1^3\, y_2^2\, y_3} - r = 0 \tag{141}$$

in which r is the constant $\sqrt[10]{96\sqrt{3}}$. The $\sqrt[10]{(\,)}$ is not essential: it permits, in the numerical experiment, to manipulate a scaled g function that involves only a homogeneous (of degree 1) term and a constant.

Now the primary minimization problem in dimension 4 admits the following solution:

$$y_k^* = \sqrt{(4-k)\,A^{k-1}} \quad (0 \le k \le 3) \tag{142}$$

The gradient vector, ∇J_A, and the Hessian matrix, H_A, admit the same formal expressions in terms of the variables, and in particular H_A^* is the same matrix. However, the constraint gradient vector ∇g is now given by:

$$\nabla\left(y_0^4\, y_1^3\, y_2^2\, y_3\right) = \nabla\,(g+r)^{10} = 10\,(g+r)^9\,\nabla g \tag{143}$$

Thus, ∇g^*, and the corresponding vector ω^1 are now given by:

$$10\,(g^*+r)^9\,\nabla g^* = \begin{pmatrix} 4/y_0^* = 2\sqrt{A} \\ 3/y_1^* = \sqrt{3} \\ 2/y_2^* = \sqrt{2/A} \\ 1/y_3^* = 1/A \end{pmatrix}, \quad \omega^1 = \frac{1}{\sqrt{4A^3+3A^2+2A+1}} \begin{pmatrix} 2A\sqrt{A} \\ A\sqrt{3} \\ \sqrt{2A} \\ 1 \end{pmatrix} \tag{144}$$

Consequently, the matrix H'_A is now:

$$H'_A = \frac{2}{A^3\,(4A^3+3A^2+2A+1)^2}\,\mathcal{H} \tag{145}$$

where the matrix $\mathcal{H}$ admits the following four column–vectors:

$$\mathcal{H}_1 = \begin{pmatrix} {B_3}^2A^3 + 12\,A^7 + 8\,A^5 + 4\,A^3 \\ -2\,\sqrt{3}A^{11/2}B_3 - 2\,\sqrt{3}A^{9/2}B_2 + 4\,A^{9/2}\sqrt{3} + 2\,A^{5/2}\sqrt{3} \\ -2\,\sqrt{2}A^5B_3 + 6\,A^6\sqrt{2} - 2\,\sqrt{2}A^3B_1 + 2\,A^2\sqrt{2} \\ -2\,A^{9/2}B_3 + 6\,A^{11/2} + 4\,A^{7/2} - 2\,A^{3/2}B_0 \end{pmatrix} \quad \mathcal{H}_2 = \begin{pmatrix} . \\ 12\,A^8 + A^2{B_2}^2 + 6\,A^4 + 3\,A^2 \\ 4\,\sqrt{6}A^{15/2} - \sqrt{6}A^{7/2}B_2 - \sqrt{6}A^{5/2}B_1 + \sqrt{6}A^{3/2} \\ 4\,A^7\sqrt{3} - \sqrt{3}A^3B_2 + 2\,A^3\sqrt{3} - \sqrt{3}AB_0 \end{pmatrix} \tag{146}$$

$$\mathcal{H}_3 = \begin{pmatrix} . \\ . \\ 8\,A^7 + 6\,A^5 + A{B_1}^2 + 2\,A \\ 4\,A^{13/2}\sqrt{2} + 3\,A^{9/2}\sqrt{2} - \sqrt{2}A^{3/2}B_1 - \sqrt{2}\sqrt{A}B_0 \end{pmatrix} \quad \mathcal{H}_4 = \begin{pmatrix} . \\ . \\ . \\ 4\,A^6 + 3\,A^4 + 2\,A^2 + {B_0}^2 \end{pmatrix} \tag{147}$$

where

$$B = 1 + 2\,A + 3\,A^2 + 4\,A^3 \tag{148}$$

and

$$B_0 = B - 1\,, \quad B_1 = B - 2\,A\,, \quad B_2 = B - 3\,A^2\,, \quad B_3 = B - 4\,A^3 \tag{149}$$

and . stands for terms known by symmetry and omitted for brevity. The diagonalization of the matrix $\mathcal{H}$ permits us to identify a new eigenvector matrix Ω. To maintain the continuity of the optimization variables U and V as the antagonism parameter A varies, the sign of the eigenvectors, after normalization and appropriate ordering, has been defined uniquely by forcing the diagonal elements of the matrix Ω to be positive.

From this point, for the 2+2 splitting, the formal solution of Section 4 is amended. In particular, since the change of variables now reflects a different expression for y_k^* and a nonzero u_1,

$$y_k = y_k^* + \omega_{k+1,1}\,u_1 + \omega_{k+1,2}\,u_2 + \omega_{k+1,3}\,v_2 + \omega_{k+1,4}\,v_1 \quad (0 \le k \le 3) \tag{150}$$

The derivation of v_1 and v_2 is similar; we get $\alpha = 1$ and $\beta = 0$, but:

$$\gamma = \sum_{k=0}^{3} \left(y_k^* + \sum_{j=0}^{1} \omega_{k+1,j+1} \right) \omega_{k+1,4} = \sum_{k=0}^{3} y_k^*\,\omega_{k+1,4} \tag{151}$$

Consequently, independently of u_1 and u_2:

$$\begin{cases} v_1 = -\displaystyle\sum_{k=0}^{3} y_k^*\,\omega_{k+1,4} \\ v_2 = -\displaystyle\sum_{k=0}^{3} y_k^*\,\omega_{k+1,3} \end{cases} \tag{152}$$

The remaining unknowns u_1 and u_2 are obtained by expressing that the primary criterion J_A, subject to the constraint $g = 0$, is stationary w.r.t. the vector U for fixed V, that is:

$$\left(\frac{\partial}{\partial U}\right)_V (J_A + \lambda g) = 0\,, \quad g = 0 \tag{153}$$

By eliminating the Lagrange multiplier λ, we get the following system of two nonlinear equations for u_1 and u_2:

$$\begin{cases} g = 0 \\ h = \dfrac{\partial J_A}{\partial u_1}\dfrac{\partial g}{\partial u_2} - \dfrac{\partial J_A}{\partial u_2}\dfrac{\partial g}{\partial u_1} = 0 \end{cases} \tag{154}$$

The Nash equilibrium point has been evaluated numerically for different values of the parameter A ($1 \ge A \ge 3$). For each value, the eigenvector matrix Ω is identified by diagonalization of the matrix $\mathcal{H}$ and appropriate ordering of the eigenvalues and scaling of the eigenvectors. Then $\{y_k^*\}$ $(0 \le k \le 3)$, v_1 and v_2 are computed;

lastly, u_1 and u_2 are calculated by solving the nonlinear system $g = h = 0$ for u_1 and u_2 by Newton's method, by which, at the n-th iteration, after solving the system:

$$\begin{pmatrix} \dfrac{\partial g}{\partial u_1} & \dfrac{\partial g}{\partial u_2} \\ \dfrac{\partial h}{\partial u_1} & \dfrac{\partial h}{\partial u_2} \end{pmatrix}^{(n)} \begin{pmatrix} \delta u_1 \\ \delta u_2 \end{pmatrix} = -\begin{pmatrix} g \\ h \end{pmatrix}^{(n)} \tag{155}$$

the unknowns $u_1^{(n)}$ and $u_2^{(n)}$ are incremented of δu_1 and δu_2. To implement the iteration, we have used the following expressions for derivatives, for $i = 1, 2$:

$$\frac{\partial g}{\partial u_i} = \frac{g+r}{10} \sum_{k=0}^{3} \frac{4-k}{y_k}\, \omega_{k+1,i} \tag{156}$$

$$\frac{\partial^2 g}{\partial u_i\, \partial u_j} = \frac{1}{g+r} \frac{\partial g}{\partial u_i} \frac{\partial g}{\partial u_j} - \frac{g+r}{10} \sum_{k=0}^{3} \frac{4-k}{y_k^2}\, \omega_{k+1,i}\, \omega_{k+1,j} \tag{157}$$

$$\frac{\partial J_A}{\partial u_i} = \sum_{k=0}^{3} \frac{2 y_k}{A^k}\, \omega_{k+1,i} \tag{158}$$

$$\frac{\partial^2 J_A}{\partial u_i\, \partial u_j} = \sum_{k=0}^{3} \frac{2\, \omega_{k+1,i}\, \omega_{k+1,j}}{A^k} \tag{159}$$

$$\frac{\partial h}{\partial u_1} = \frac{\partial^2 J_A}{\partial u_1^2} \frac{\partial g}{\partial u_2} + \frac{\partial J_A}{\partial u_1} \frac{\partial^2 g}{\partial u_2 \partial u_1} - \frac{\partial^2 J_A}{\partial u_2 \partial u_1} \frac{\partial g}{\partial u_1} - \frac{\partial J_A}{\partial u_2} \frac{\partial^2 g}{\partial u_1^2} \tag{160}$$

$$\frac{\partial h}{\partial u_2} = \frac{\partial^2 J_A}{\partial u_1 \partial u_2} \frac{\partial g}{\partial u_2} + \frac{\partial J_A}{\partial u_1} \frac{\partial^2 g}{\partial u_2^2} - \frac{\partial^2 J_A}{\partial u_2^2} \frac{\partial g}{\partial u_1} - \frac{\partial J_A}{\partial u_2} \frac{\partial^2 g}{\partial u_1 \partial u_2} \tag{161}$$

This solution has been calculated numerically for values of A varying from 1 to 3. The corresponding ratios $\bar{J}_A / J_A^*$ and $\bar{J}_B / J_B^*$ are plotted on Fig. 4. The derivatives w.r.t. A of both criteria are again visibly equal to zero at $A = 1$. On Fig. 5 is plotted the variation of the optimization variables u_1 and u_2 with the parameter A. The variable u_1 would be identically equal to zero if the constraint were linear; here it is found, as expected, equal to zero at $A = 1$ with a zero derivative. Thus, these results provide some evidence to confirm in this case of nonlinear constraint, that the initial optimum is robust. To be complete, we also provide on Fig. 6 the variation of the optimization variables v_1 and v_2 that control the secondary criterion J_B; these are also equal to zero at $A = 1$ but have visibly nonzero derivatives.

6 The special case $K + p = n + 1$ and the connection with Pareto fronts

6.1 Generalities

When $K + p = n + 1$, the dimension of the vector U is just equal to the number of scalar constraints applied to the primary problem. In this case, in the Nash game, the primary functional J_A does not play any role, and the primary subproblem only acts as an algebraic constraint of dimension K put on the secondary subproblem. Thus, the formulation of the Nash game reduces to:

$$\begin{cases} \min\limits_{V} J_B(Y), \quad Y = Y_A^* + S \begin{pmatrix} U \\ V \end{pmatrix} \\ \text{Subject to: } g = 0 \end{cases} \tag{162}$$

and the equations defining the Nash equilibrium point are:

$$\begin{cases} \left(\dfrac{\partial}{\partial V} \right)_U \left(J_B + \nu^T g \right) = 0 \\ g = 0 \end{cases} \tag{163}$$

where ν is a K–dimensional vector of Lagrange multipliers. Let $\bar{Y}$ be a solution of this problem, and $\bar{J}_A$, $\bar{J}_B$ the corresponding values of the functionals, and consider the Pareto–optimal point $\bar{Y}^*$ for which $J_A = \bar{J}_A$. This point is the solution of the problem:

$$\begin{cases} \min\limits_{Y \in \mathbb{R}^{n+1}} J_B(Y) \\ \text{Subject to: } g = 0\,,\ J_A(Y) \leq \bar{J}_A \end{cases} \tag{164}$$

At $Y = \bar{Y}^*$, the inequality constraint is saturated, and we have:

$$\begin{cases} \dfrac{\partial}{\partial Y}\left(J_B + \alpha^T g + \beta J_A\right) = 0 \\ g = 0\,,\ J_A(Y) = \bar{J}_A \end{cases} \tag{165}$$

for some Lagrange multipliers α and β ($\alpha \in \mathbb{R}^K$, $\beta \in \mathbb{R}$). Hence at this Pareto–optimal point, $g = 0$ and:

$$\begin{aligned} \left(\frac{\partial}{\partial V}\right)_U \left(J_B + \nu^T g\right) &= \frac{\partial J_B}{\partial Y}\left(\frac{\partial Y}{\partial V}\right)_U + \nu^T \left(\frac{\partial g}{\partial V}\right)_U \\ &= -\left(\alpha^T \frac{\partial g}{\partial Y} + \beta \frac{\partial J_A}{\partial Y}\right)\left(\frac{\partial Y}{\partial V}\right)_U + \nu^T \left(\frac{\partial g}{\partial V}\right)_U \\ &= \left[\left(\frac{\partial g}{\partial V}\right)_U^T (\nu - \alpha) - \beta \left(\frac{\partial J_A}{\partial V}\right)_U^T\right]^T \end{aligned} \tag{166}$$

The transposed Jacobian matrix $\left(\dfrac{\partial g}{\partial V}\right)_U^T$ is of dimension $p \times K$, and the column–vector $\left(\dfrac{\partial J_A}{\partial V}\right)_U^T$ of dimension $p \times 1$.

6.2 Sub-case $K \geq p$

Given α and β, if $K \geq p$ and if the rank of the matrix $\left(\dfrac{\partial g}{\partial V}\right)_U^T$ is equal to p, it is possible to find ν so that the above expression, in (166), is equal to zero. In such a case, the Pareto–optimal point is also a Nash equilibrium point.

6.3 Sub-case $K = 1$ and $p = n$

In this more practical special case:

$$U = u_1\,, \quad V = \begin{pmatrix} v_n \\ \vdots \\ v_1 \end{pmatrix} \tag{167}$$

Let ψ be a smooth real–valued function of the vector Y:

$$d\psi = \frac{\partial \psi}{\partial Y}\, dY = \frac{\partial \psi}{\partial u_1}\, du_1 + \frac{\partial \psi}{\partial V}\, dV \tag{168}$$

But:

$$\begin{pmatrix} du_1 \\ dV \end{pmatrix} = \Omega^T\, dY \tag{169}$$

Let Π be the $(n+1) \times (n+1)$ matrix whose elements are all equal to 0 except for the first element in the first row that is equal to 1. Then:

$$\begin{pmatrix} du_1 \\ 0 \end{pmatrix} = \Pi \begin{pmatrix} du_1 \\ dV \end{pmatrix} = \Pi\, \Omega^T\, dY\,, \quad \begin{pmatrix} 0 \\ dV \end{pmatrix} = (I - \Pi)\, \Omega^T\, dY \tag{170}$$

where the first "0" stands for a column–vector of 0. Consequently:

$$\frac{\partial \psi}{\partial Y} = \left[(\psi_{u_1})_V\,, 0\right] \Pi\, \Omega^T + \left[0, (\psi_V)_{u_1}\right] (I - \Pi)\, \Omega^T \tag{171}$$

where here "0" stands for an n–dimensional row–vector.

Now assume that a Nash equilibrium point exists satisfying (163) for some $\bar{Y}, \nu$. Let $\alpha = \nu$ and $\psi = J_B + \alpha^T g$, yielding:

$$v^T = \frac{\partial}{\partial Y}\left(J_B + \alpha^T g\right) = \left[(\psi_{u_1})_V\,, 0\right] \Pi\, \Omega^T = \left[(\psi_{u_1})_V\,, 0\right] \Omega^T \tag{172}$$

Let w be the projection of the column–vector v orthogonal to ω^1:

$$w = \left(I - \omega^1\, {\omega^1}^T\right) v = \left(I - \omega^1\, {\omega^1}^T\right) \Omega \begin{pmatrix} (\psi_{u_1})_V \\ 0 \end{pmatrix} = \left(0\,, \omega^1\,, \ldots\,, \omega^n\right) \begin{pmatrix} (\psi_{u_1})_V \\ 0 \end{pmatrix} = 0 \tag{173}$$

Hence

$$v = \rho_1\, \omega^1 = \rho\, \nabla J_A^* \tag{174}$$

for some real number ρ, and (165) is satisfied for

$$\beta = -\frac{\rho}{\|\nabla J_A\|} \tag{175}$$

which proves that the Nash–equilibrium point belongs to the Pareto–equilibrium front.

The result applies to the particular case of the *"1+3 splitting"* of Appendix B, where $K = 1$ and $p = n = 3$.

7 Successive kernels, constraints, robust design

The evaluation of the numbers α_j should be used to guide the choice of the dimension p of the subspace allocated to the secondary functional.

In the limit, in a concurrent optimization of two strongly–antagonistic functionals, variations of the parameters about the original optimum point Y_A^* may cause excessive modifications to the original–design functional value J_A^*, to the extent of making nearly all practical games unstable. In such a difficult situation, one could, in a first step, restrict p to 1, giving best chance of success to a subsequent first concurrent optimization. From this iteration, one would create of new database supporting a new or updated surrogate model yielding updated values for the gradient and Hessian. Hence, a fresh matrix H_{AS_1} could be calculated to replace H'_A, permitting us to define a new null space, and consider again solely the smallest nonzero eigenvalue α_n, and so on.

Besides, our approach opens the way to the treatment of other difficulties.

A game strategy can be used to satisfy an additional constraint that was not included in the original design loop, by treating it as a secondary functional to be minimized in a subspace of least performance degradation.

In fact, our procedure is a form of *locally–robust design* approach, since we are looking for designs close to the original one that best preserve it.

8 Conclusions

We have considered a general setting in which a constrained *primary minimization* problem is treated concurrently with an unconstrained *secondary minimization* problem. Several theoretical results have been established. First, a *split of territories* has been proposed based on the eigensystem of a real symmetric matrix representing the restriction of the primary–criterion Hessian matrix to the subspace tangent to the constraint hyper-surfaces. The split is made in the perspective of identifying a Nash equilibrium point between the two minimization problems and is optimal in the sense that it results in a least degradation of the primary functional w.r.t. the original optimal point of the full–dimension primary problem. Second, the eigensplit has been proved to be such that in case of identical criteria, the Nash equilibrium point is identical to the original optimal point (Theorem 1). Thus, the formulation yields, as the difference between the criteria is introduced progressively, a continuum of Nash equilibrium points initiated at the original optimal point. Thirdly, the original optimal point is found to be *robust*, in the sense that variations in the primary functional are second–order in terms of the deviation of the Nash–equilibrium point from the original single–criterion optimal point.

These theoretical results have been illustrated in the particular case of the minimization of quadratic forms of four variables subject to a linear constraint. In this simple problem, all fourteen possible splittings of the primitive variables result in a Nash equilibrium point in which both criteria are degraded. Thus all these trivial splits of territories fail. Inversely, the eigensplitting succeeds in all cases to reduce the secondary criterion, at the expense of a degradation, that is an increase, of the primary criterion. Subsequently, a variant of this problem involving a nonlinear constraint has also been treated numerically, demonstrating similar behaviors.

A practical way to implement a continuation technique to introduce smoothly the antagonism between two general concurrent criteria has been proposed and will be tested in future work.

Lastly, the connection between the Nash equilibrium points and Pareto–optimal points has been partially examined.

References

[1] J.F. Nash, *Non-cooperative games*, Annals of Mathematics, **54** (2), pp. 286-295, 1951.

[2] T. Basar and G.J. Olsder, *Dynamic Noncooperative Game Theory*, Academic Press, Bodmin, Cornwall, Great Britain, 1995.

[3] J.F. Wang, *Optimisation Distribuée Multicritère par Algorithmes Génétiques et Théorie des Jeux & Application à la Simulation Numérique de Problèmes d'Hypersustentation en Aérodynamique*, Doctoral Thesis *(Spécialité: Mathématiques Appliquées)*, University of Paris 6, 2001.

[4] Z. Tang and J.-A. Désidéri and J. Périaux, *Multi-criterion Aerodynamic Shape-Design Optimization and Inverse Problems Using Control Theory and Nash Games*, Journal of Optimization Theory and Applications, to appear.

[5] J.-A. Désidéri and B. Abou El Majd and A. Janka, *Nested and Self-Adaptive Bézier Parameterizations for Shape Optimization*, J. Comput. Phys., to appear.

[6] R.M. Hicks and P.A. Henne, *Wing design by numerical optimization*, J. of Aircraft, **15** (7), pp. 407-412, 1978.

[7] J. L. Zhou, A. L. Tits and C. T. Lawrence, *User's Guide for FFSQP Version 3.7 : A Fortran Code for Solving Optimization Programs, Possibly Minimax, with General Inequality Constraints and Linear Equality Constraints, Generating Feasible Iterates*, Institute for Systems Research, University of Maryland,Technical Report SRC-TR-92-107r5, College Park, MD 20742, 1997. http://www.aemdesign.com/

[8] R. Duvigneau, *Adaptive Parameterization using Free-Form Deformation for Aerodynamic Shape Optimization*, INRIA Research Report No. 5949, 2006. (http://hal.inria.fr/inria-00085058)

[9] J. Sacks, W. J. Welch, T. J. Mitchell and H. P. Wynn, *Design and Analysis of Computer Experiments. Statistical Science*, Vol. 4, No. 4, 409-435, 1989.

[10] M. K. Karakasis and K. C. Giannakoglou, *On the use of surrogate evaluation models in multi-objective evolutionary algorithms*. In P. Neittaanmäki *et al* Eds., *Eccomas 2004, Proc. of the European Congress on Computational Methods in Applied Sciences and Engineering*, Jÿvaskÿla, Finland, July 2004.

[11] *Introduction to Matrix Analysis, Second Edition*, R. Bellman, Classics in Applied Mathematics, SIAM, Philadelphia, 1997. *(In particular: Chapter 7, Variational Description of Characteristic Roots, Theorem 1.)*

Acknowledgements

The author gratefully acknowledges the fruitful scientific discussions with his colleagues P. Bernhard, R. Duvigneau, A. Habbal, J. Périaux, O. Pourtallier and J.-P. Zolésio.

This work was partially supported by the ANR/RNTL Program "OMD" on Multidisciplinary Optimization under contract ANR05RNTL00908.

A "3+1" splittings

In this appendix, we consider again the simple problem of Section 4 and derive the formulas corresponding to "3+1" splittings in which three variables are assigned to the constrained minimization of the primary criterion J_A and only one to the unconstrained minimization of the secondary criterion J_B.

We first examine the games in which the splitting applies to the primitive variables $\{ y_k \}$. Four games are of this type:

$$G'_j \, (0 \le j \le 3) \; : \quad A \; : \; \{ y_k, y_\ell, y_m \} \, (k, \ell, m \neq j) \, , \quad \text{and } B \; : \; \{ y_j \} \tag{176}$$

Hence, at equilibrium:

$$\bar{y}_j = 0 \tag{177}$$

and for the other three variables ($i \neq j$):

$$\bar{y}_i = -\frac{\lambda}{2} A^i \tag{178}$$

Thus, the constraint equation writes:

$$-\frac{\lambda}{2} \sum_{\nu \neq j} A^\nu = \frac{A^4 - 1}{A - 1} \tag{179}$$

and:

$$\bar{y}_i = \frac{A^i}{\sum_{\nu \neq j} A^\nu} \frac{A^4 - 1}{A - 1} \tag{180}$$

and correspondingly:

$$\begin{cases} J_A^{G'_j} = \dfrac{\sum_{i \neq j} A^i}{\left(\sum_{\nu \neq j} A^\nu \right)^2} \left(\dfrac{A^4 - 1}{A - 1} \right)^2 = \dfrac{1}{\left(\sum_{i \neq j} A^i \right)} \left(\dfrac{A^4 - 1}{A - 1} \right)^2 \\ J_B^{G'_j} = \dfrac{\sum_{i \neq j} A^{2i}}{\left(\sum_{\nu \neq j} A^\nu \right)^2} \left(\dfrac{A^4 - 1}{A - 1} \right)^2 = \dfrac{\sum_{i \neq j} A^{2i}}{\left(\sum_{i \neq j} A^i \right)^2} \left(\dfrac{A^4 - 1}{A - 1} \right)^2 \end{cases} \tag{181}$$

or equivalently:

$$\begin{cases} \dfrac{J_A^{G'_j}}{J_A^*} = \dfrac{1}{\left(\sum_{i \neq j} A^i \right)} \left(\dfrac{A^4 - 1}{A - 1} \right) = \dfrac{1 + A + A^2 + A^3}{\left(\sum_{i \neq j} A^i \right)} \ge 1 \\ \dfrac{J_B^{G'_j}}{J_B^*} = \dfrac{\sum_{i \neq j} A^{2i}}{\left(\sum_{\nu \neq j} A^\nu \right)^2} \left(\dfrac{A^4 - 1}{A - 1} \right)^2 \dfrac{A^2 - 1}{A^8 - 1} = \dfrac{\sum_{i \neq j} A^{2i}}{\left(\sum_{i \neq j} A^i \right)^2} \dfrac{A + 1}{A - 1} \dfrac{A^4 - 1}{A^4 + 1} \end{cases} \tag{182}$$

Hence, in particular for $A = 1$:

$$\frac{J_A^{G'_j}}{J_A^*} = \frac{J_B^{G'_j}}{J_B^*} = \frac{4}{3} > 1 \tag{183}$$

In fact, for all $A \ge 1$:

$$\forall j \, , \quad \frac{J_A^{G'_j}}{J_A^*} \ge 1 \, , \quad \frac{J_B^{G'_j}}{J_B^*} \ge 1 \tag{184}$$

as illustrated by Fig. 2. Hence, all four splittings result in games that fail to reduce the secondary criterion, even at the expense of a degradation of the primary criterion. Interestingly, w.r.t. the secondary criterion the splittings associated with G'_0 and G'_3 are equivalent.

We now examine the recommended split of territory, using the gradient vector, the Hessian matrix H_A^* and the matrix H'_A of Section 4. Here, the split of parameters is:

$$U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} , \quad V = \begin{pmatrix} v_1 \end{pmatrix} \tag{185}$$

and it is associated with the following change of variables:

$$\begin{cases} y_0 = 1 + \omega_{1,1}\, u_1 + \omega_{1,2}\, u_2 + \omega_{1,3}\, u_3 + \omega_{1,4}\, v_1 \\ y_1 = A + \omega_{2,1}\, u_1 + \omega_{2,2}\, u_2 + \omega_{2,3}\, u_3 + \omega_{2,4}\, v_1 \\ y_2 = A^2 + \omega_{3,1}\, u_1 + \omega_{3,2}\, u_2 + \omega_{3,3}\, u_3 + \omega_{3,4}\, v_1 \\ y_3 = A^3 + \omega_{4,1}\, u_1 + \omega_{4,2}\, u_2 + \omega_{4,3}\, u_3 + \omega_{4,4}\, v_1 \end{cases} \tag{186}$$

Then again, since $S_1 = 2$ and $S_2 = S_3 = S_4 = 0$, the application of the constraint to the primary minimization yields:

$$u_1 = 0 \tag{187}$$

The stationarity of the criterion J_A w.r.t. u_2 and u_3 writes:

$$\sum_{k=0}^{3} \frac{y_k}{A^k}\, \omega_{k+1,2} = \sum_{k=0}^{3} \frac{y_k}{A^k}\, \omega_{k+1,3} = 0 \tag{188}$$

that is, since $S_2 = S_3 = u_1 = 0$:

$$\begin{cases} \sigma_{2,2}\, u_2 + \sigma_{2,3}\, u_3 + \sigma_{2,4}\, v_1 = 0 \\ \sigma_{3,2}\, u_2 + \sigma_{3,3}\, u_3 + \sigma_{3,4}\, v_1 = 0 \end{cases} \tag{189}$$

where again:

$$\sigma_{i,j} = \sum_{k=0}^{3} \frac{\omega_{k+1,i}\, \omega_{k+1,j}}{A^k} \tag{190}$$

Hence:

$$u_2 = \frac{\sigma_{2,3}\, \sigma_{3,4} - \sigma_{2,4}\, \sigma_{3,3}}{\sigma_{2,2}\, \sigma_{3,3} - \sigma_{2,3}^2}\, v_1\,, \quad u_3 = \frac{\sigma_{2,3}\, \sigma_{2,4} - \sigma_{3,4}\, \sigma_{2,2}}{\sigma_{2,2}\, \sigma_{3,3} - \sigma_{2,3}^2}\, v_1 \tag{191}$$

Lastly, we express the stationarity of the secondary criterion J_B w.r.t. the unique parameter v_1. We find again:

$$v_1 = -\sum_{k=0}^{3} A^k\, \omega_{k+1,4} \tag{192}$$

From this quasi–explicit solution, we observe again that for $A = 1$:

$$u_1 = u_2 = u_3 = v_1 = 0 \tag{193}$$

by orthogonality of the eigenvectors, and this is equivalent to:

$$\bar{Y} = Y_A^* \tag{194}$$

which illustrates again Theorem 1.

The ratios $J_A^{G'_j}/J_A^*$ and $J_B^{G'_j}/J_B^*$ corresponding to the 3+1 splittings of the primitive variables, as well as the ratios $\bar{J}_A/J_A^*$ and $\bar{J}_B/J_B^*$ corresponding to the 3+1 eigensplitting have been calculated for A varying from 1 to 3 and plotted on Fig. 2.

Thus, the main conclusion is that 1+3 splittings of the primitive variables fail to reduce the secondary criterion from the initial optimal point Y_A^*. In contrast, with the proposed 3+1 eigensplitting, one finds a continuum of equilibrium points evolving from Y_A^* as A increases from 1 to ∞. For any value of the antagonism parameter $A > 1$, the strategy succeeds to reduce the secondary criterion J_B at the expense of an increase of the primary criterion J_A.

At the initial equilibrium point ($A = 1$), the derivatives of both criteria w.r.t. the antagonism parameter are equal to zero.

Again, we have indicated on Fig. 2 (bottom) the value of the secondary criterion J_B achieved by the Pareto–optimal point corresponding to the same value of the primary criterion ($J_A = \bar{J}_A$). The curves drawn by the Nash–equilibrium point and the Pareto–optimal point follow similar trends and remain close to each other.

B "1+3" splittings

In this appendix, we consider again the simple problem of Section 4 and derive the formulas corresponding to "1+3" splittings in which only one variable is assigned to the constrained minimization of the primary criterion J_A and three to the unconstrained minimization of the secondary J_B.

We first examine the games in which the splitting applies to the primitive variables $\{ y_k \}$. Four games are of this type:

$$G''_j \ (0 \le j \le 3) : \quad A : \{ y_j \}, \quad \text{and } B : \{ y_k, y_\ell, y_m \} \ (k, \ell, m \neq j) \tag{195}$$

Thus, at equilibrium:

$$\bar{y}_k = \bar{y}_\ell = \bar{y}_m = 0 \tag{196}$$

and y_j is forced by the constraint:

$$\bar{y}_j = \frac{A^4 - 1}{A - 1} \tag{197}$$

Consequently:

$$J_A^{G''_j} = \frac{\bar{y}_j^2}{A^j} = \frac{1}{A^j} \left(\frac{A^4 - 1}{A - 1} \right)^2 , \quad J_B^{G''_j} = \bar{y}_j^2 = \left(\frac{A^4 - 1}{A - 1} \right)^2 \tag{198}$$

which is equivalent in terms of relative variations to the following:

$$\frac{J_A^{G''_j}}{J_A^*} = \frac{1}{A^j} \left(\frac{A^4 - 1}{A - 1} \right) , \quad \frac{J_B^{G''_j}}{J_B^*} = \left(\frac{A^4 - 1}{A - 1} \right)^2 \frac{A^2 - 1}{A^8 - 1} = \frac{A^4 - 1}{A - 1} \, \frac{A + 1}{A^4 + 1} \tag{199}$$

Hence for $A = 1$:

$$\frac{J_A^{G''_j}}{J_A^*} = \frac{J_B^{G''_j}}{J_B^*} = 4 \tag{200}$$

and both ratios are greater than 1 for all $A \ge 1$. Therefore, none of these four strategies succeeds. The least degradation of the criterion J_A is realized when j is maximum, that is $j = 3$, corresponding to the splitting of G''_3.

We now examine the recommended split of territory, for which we use here a more compact notation:

$$Y = Y_A^* + u_1 \, \omega^1 + v \quad (v \perp \omega^1) \tag{201}$$

which reflects the splitting of the perturbation about Y_A^* into a component along the vector ω^1, that is along the constraint gradient, and one, v, orthogonal to it. Thus:

$$J_B = \| Y \|^2 =< Y_A^* + u_1 \, \omega^1 + v \, , \, Y_A^* + u_1 \, \omega^1 + v >= \| Y_A^* + v \|^2 + u_1^2 + 2 \, u_1 \, < \omega^1 , Y_A^* > \tag{202}$$

since $< \omega^1 , v >= 0$. Therefore the strategy of the player in control of J_B is evident. For fixed u_1, it minimizes $\| Y_A^* + v \|^2$, with $v \in \omega^{1\perp}$. The solution is the vector opposite to the orthogonal projection of Y_A^* onto $\omega^{1\perp}$:

$$\bar{v} = - \left(I - \omega^1 \, {\omega^1}^T \right) Y_A^* \tag{203}$$

Consequently:

$$\bar{Y}_A + \bar{v} = \omega^1 \, {\omega^1}^T \, Y_A^* = ({\omega^1}^T \, Y_A^*) \, \omega^1 \, , \quad < \omega^1 , Y_A^* >= {\omega^1}^T \, Y_A^* \tag{204}$$

and:

$$J_B = ({\omega^1}^T \, Y_A^*)^2 + u_1^2 + 2 \, u_1 \, ({\omega^1}^T \, Y_A^*) = (u_1 + {\omega^1}^T \, Y_A^*)^2 \tag{205}$$

Lastly, u_1 is determined by the strategy of the player in charge of the minimization of the criterion J_A. But again this strategy is here forced by the constraint alone, which writes in our particular case of linear constraint:

$$\sum_{k=0}^{3} y_k = \nabla g^* \, . \, Y = 2 \, {\omega^1}^T \, Y = \frac{A^4 - 1}{A - 1} \tag{206}$$

In view of (201), this gives:

$${\omega^1}^T \, Y_A^* + \bar{u}_1 = \tfrac{1}{2} \, \frac{A^4 - 1}{A - 1} \tag{207}$$

and the equilibrium solution is:

$$\bar{Y} = Y_A^* + \bar{v} + \bar{u}_1\,\omega^1 = \left(\omega^{1^T} Y_A^* + \bar{u}_1\right)\omega^1 = \tfrac{1}{2}\,\frac{A^4-1}{A-1}\,\omega^1 \tag{208}$$

or component-wise:

$$\bar{y}_k = \tfrac{1}{4}\,\frac{A^4-1}{A-1} \quad (0 \le k \le 3) \tag{209}$$

Hence:

$$\bar{J}_A = \frac{1}{16}\left(\frac{A^4-1}{A-1}\right)^2 \sum_{k=0}^{3} \frac{1}{A^k} = \frac{1}{16}\left(\frac{A^4-1}{A-1}\right)^2 \frac{\frac{1}{A^4}-1}{\frac{1}{A}-1} = \frac{1}{16A^3}\left(\frac{A^4-1}{A-1}\right)^3 \tag{210}$$

and

$$\bar{J}_B = \tfrac{1}{4}\left(\frac{A^4-1}{A-1}\right)^2 \tag{211}$$

Equivalently:

$$\frac{\bar{J}_A}{J_A^*} = \frac{1}{16A^3}\left(\frac{A^4-1}{A-1}\right)^2 \tag{212}$$

and:

$$\frac{\bar{J}_B}{J_B^*} = \tfrac{1}{4}\left(\frac{A^4-1}{A-1}\right)^2 \frac{A^2-1}{A^8-1} = \tfrac{1}{4}\,\frac{A^4-1}{A-1}\,\frac{A+1}{A^4+1} \tag{213}$$

Consequently:

$$\left(\frac{\bar{J}_A}{J_A^*}\right)_{A=1} = 1\,, \quad \frac{\bar{J}_A}{J_A^*} \sim \frac{A^3}{16} \quad (\text{as } A \to \infty) \tag{214}$$

indicating a rapid (cubic) degradation of the criterion J_A as A increases from 1, but also:

$$\left(\frac{\bar{J}_B}{J_B^*}\right)_{A=1} = 1\,, \quad \lim_{A\to\infty} \frac{\bar{J}_B}{J_B^*} = \tfrac{1}{4}\,, \quad \forall A \ge 1 : \tfrac{1}{4} \le \frac{\bar{J}_B}{J_B^*} \le 1 \tag{215}$$

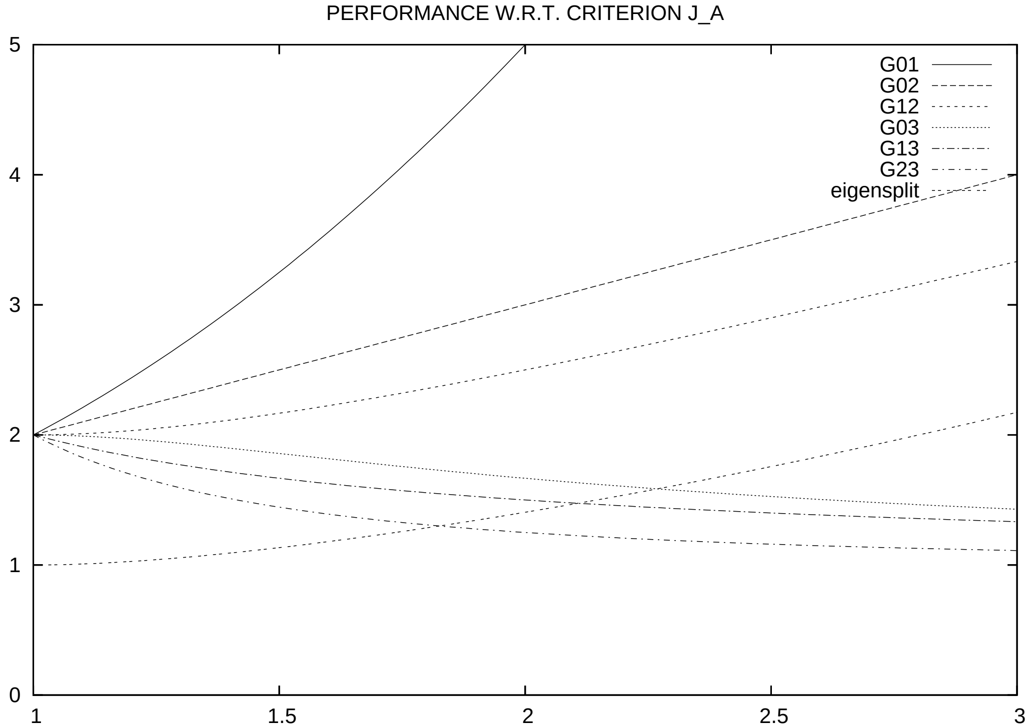
Therefore this strategy succeeds.

The ratios $J_A^{G''_j}/J_A^*$ and $J_B^{G''_j}/J_B^*$ corresponding to the 1+3 splittings of the primitive variables, as well as the ratios $\bar{J}_A/J_A^*$ and $\bar{J}_B/J_B^*$ corresponding to the 1+3 eigensplitting have been calculated for A varying from 1 to 3 and plotted on Fig. 3.

Thus, the main conclusion is that 1+3 splittings of the primitive variables fail to reduce the secondary criterion from the initial optimal point Y_A^*. In contrast, with the proposed 1+3 eigensplitting, one finds a continuum of equilibrium points evolving from Y_A^* as A increases from 1 to ∞. For any value of the antagonism parameter $A > 1$, the strategy succeeds to reduce the secondary criterion J_B at the expense of an increase of the primary criterion J_A.

At the initial equilibrium point ($A = 1$), the derivatives of both criteria w.r.t. the antagonism parameter are equal to zero.

Again, we have indicated on Fig. 3 (bottom) the value of the secondary criterion J_B achieved by the Pareto–optimal point corresponding to the same value of the primary criterion ($J_A = \bar{J}_A$). In this case, the curves drawn by the Nash–equilibrium point and the Pareto–optimal point are identical by virtue of a result established in Section 6.



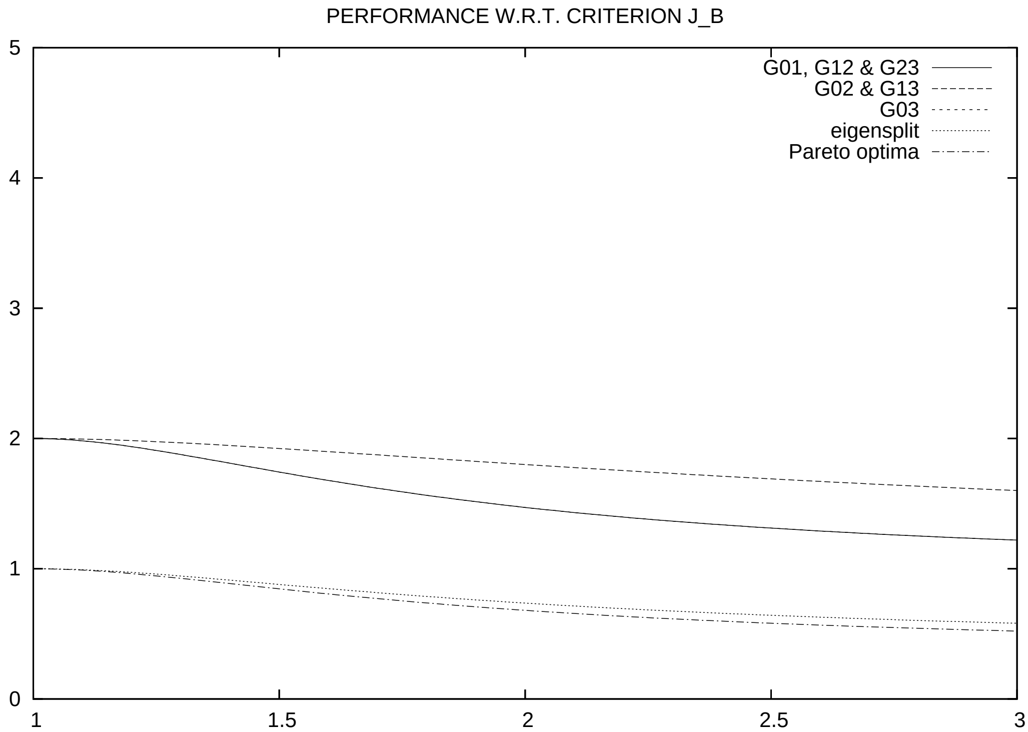
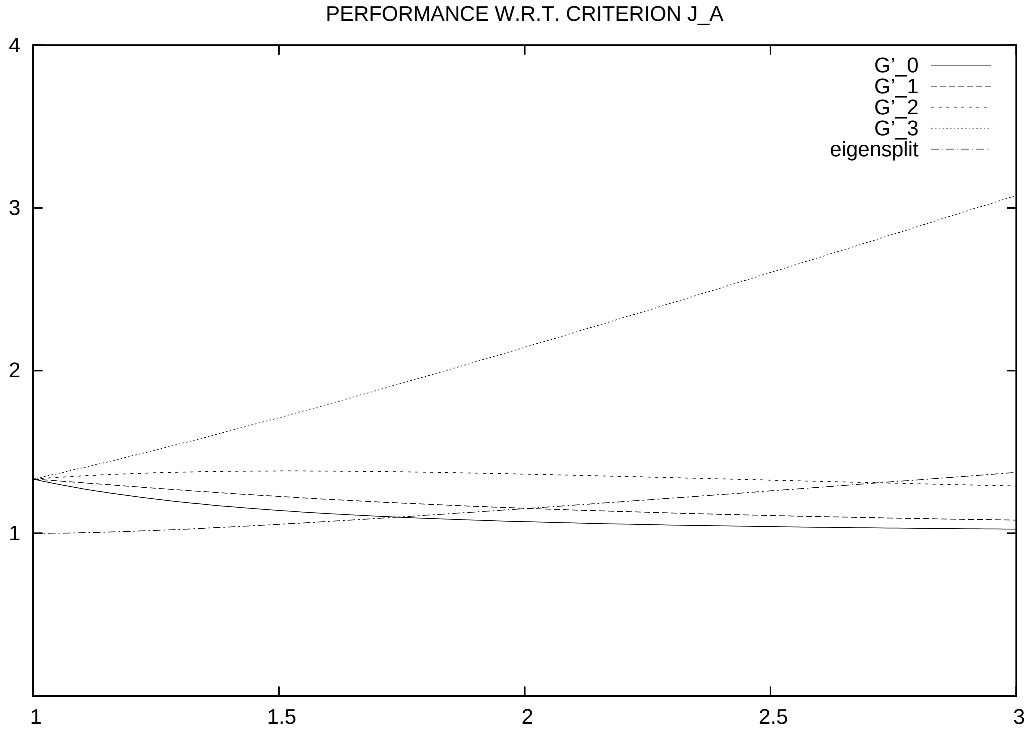


Figure 1: Performance of various strategies based on 2+2 splittings in terms of the criteria J_A (top) and J_B (bottom): ratios J_A/J_A^* (top) and J_B/J_B^* (bottom) corresponding to the equilibrium points of the games $G_{j,k}$ on the primitive variables and of the strategy based on the split of eigenvectors, as the antagonism parameter A varies.



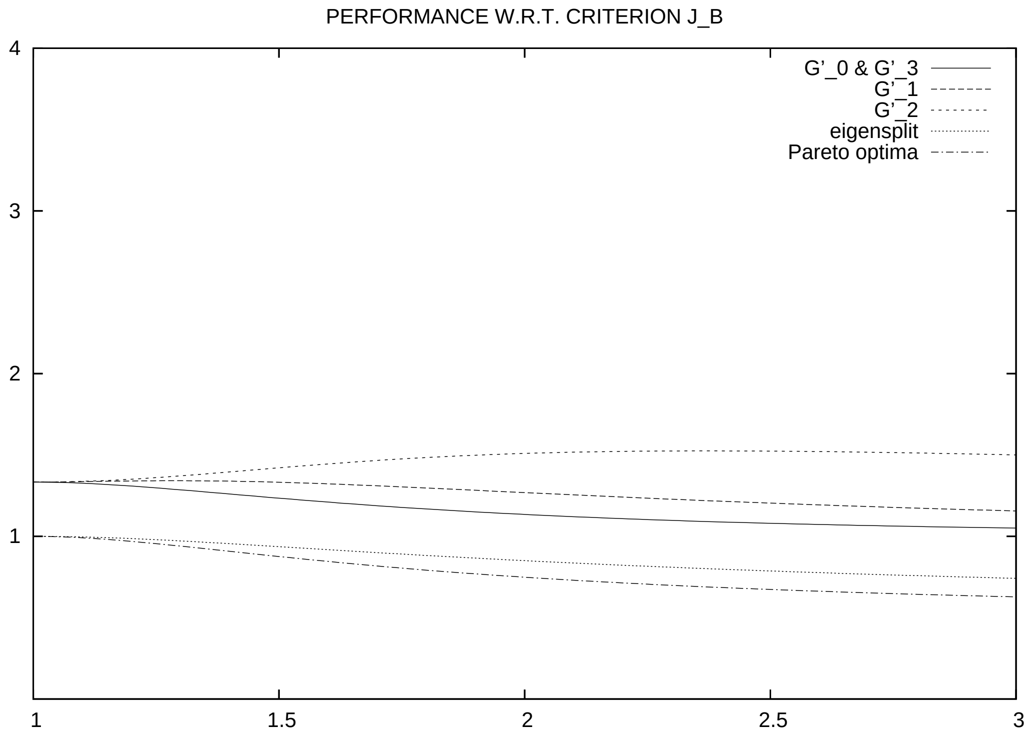


Figure 2: Performance of various strategies based on 3+1 splittings in terms of the criteria J_A (top) and J_B (bottom): ratios $\bar{J}_A/J_A^*$ (top) and $\bar{J}_B/J_B^*$ (bottom) corresponding to the equilibrium points of the games G'_j on the primitive variables and of the strategy based on the split of eigenvectors, as the antagonism parameter A varies.

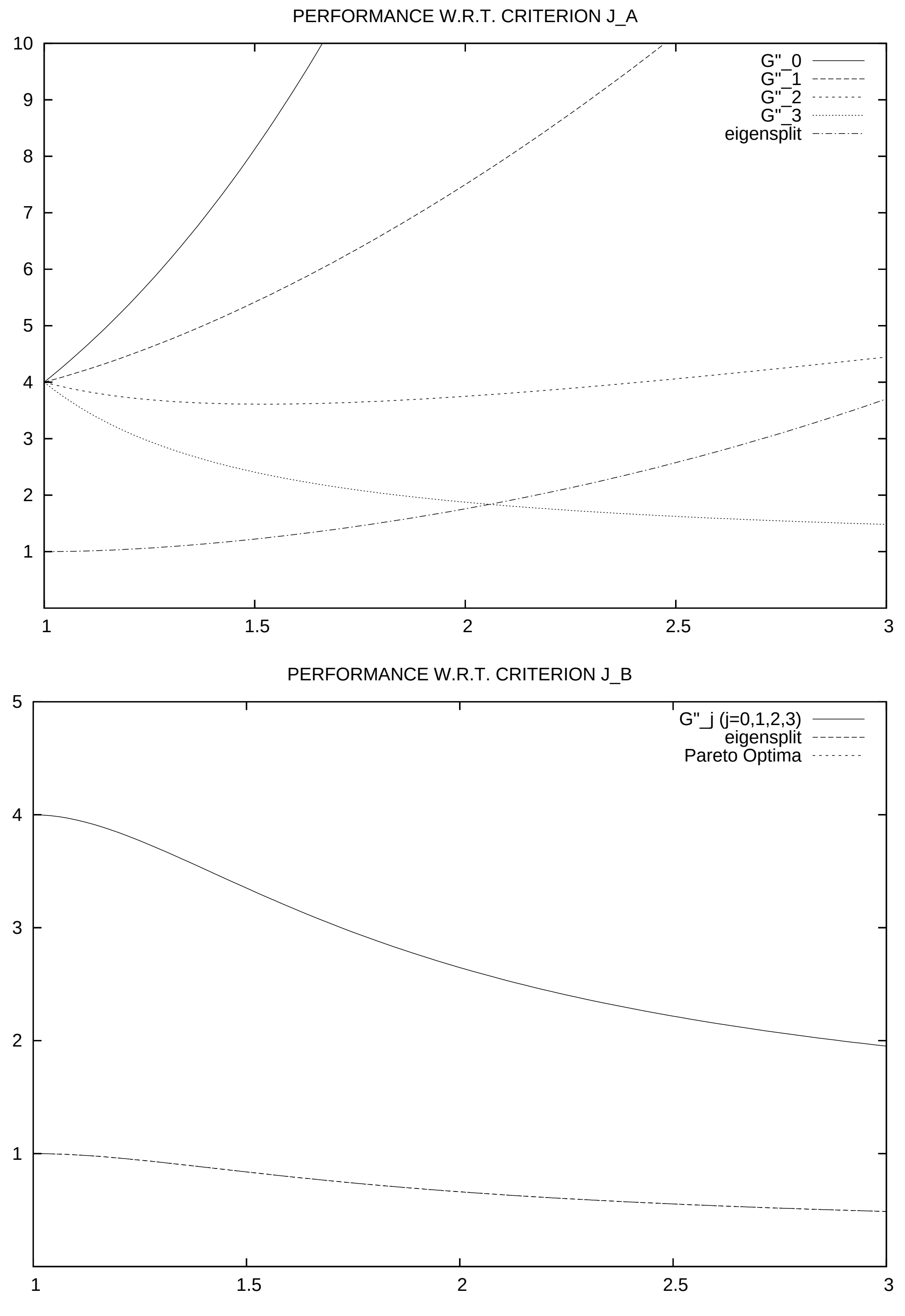


Figure 3: Performance of various strategies based on 1+3 splittings in terms of the criteria J_A (top) and J_B (bottom): ratios $\bar{J}_A/J_A^*$ (top) and $\bar{J}_B/J_B^*$ (bottom) corresponding to the equilibrium points of the games G_j'' on the primitive variables and of the strategy based on the split of eigenvectors, as the antagonism parameter A varies.

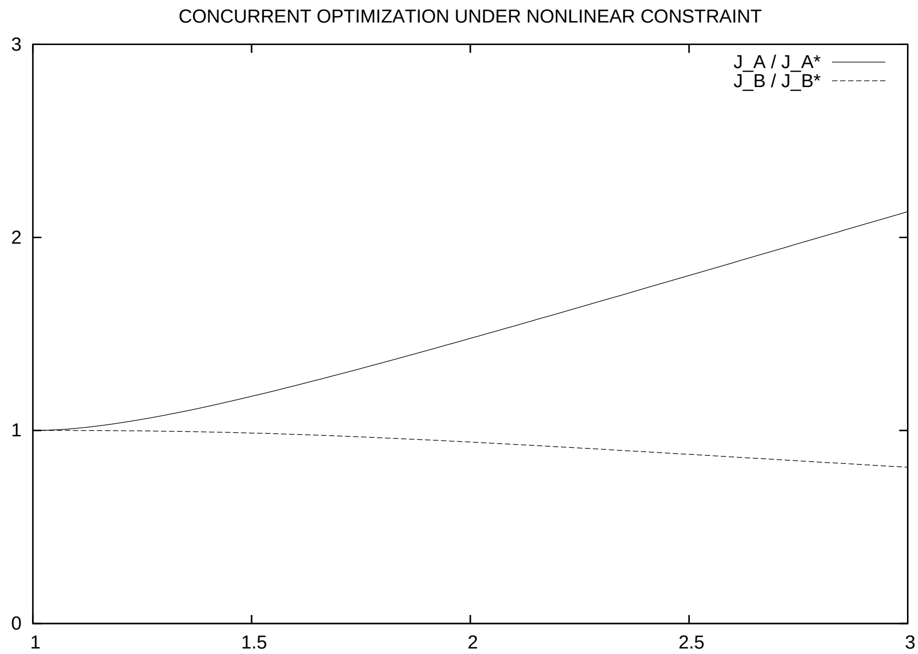


Figure 4: Concurrent minimization of the criteria J_A and J_B in the case of a nonlinear equality constraint; variations with A of the two criteria.

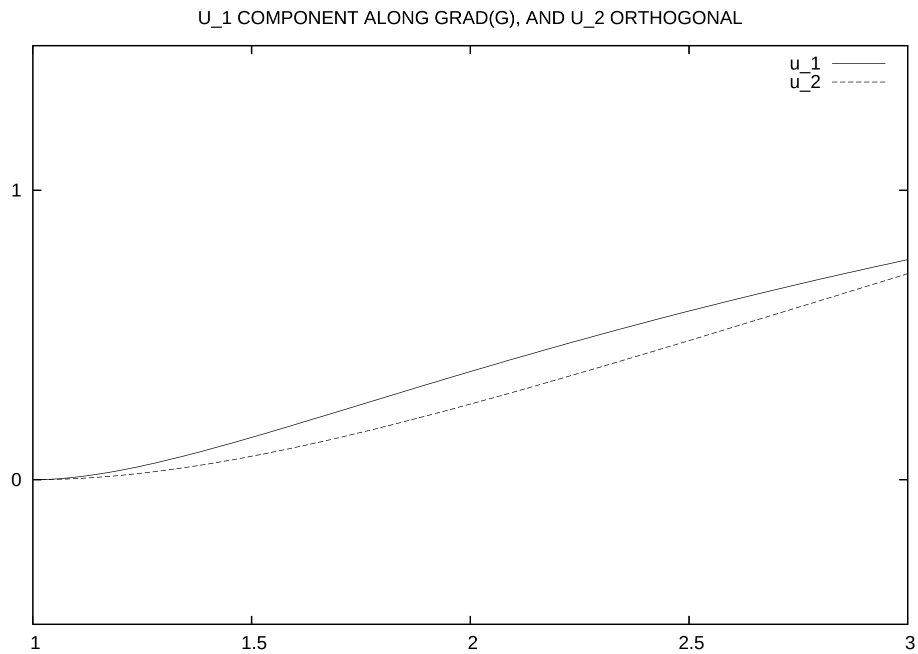


Figure 5: Concurrent minimization of the criteria J_A and J_B in the case of a nonlinear equality constraint; variation with A of the optimization variables u_1, controling the perturbation along the constraint gradient, and u_2 also participating in the minimization of J_A.

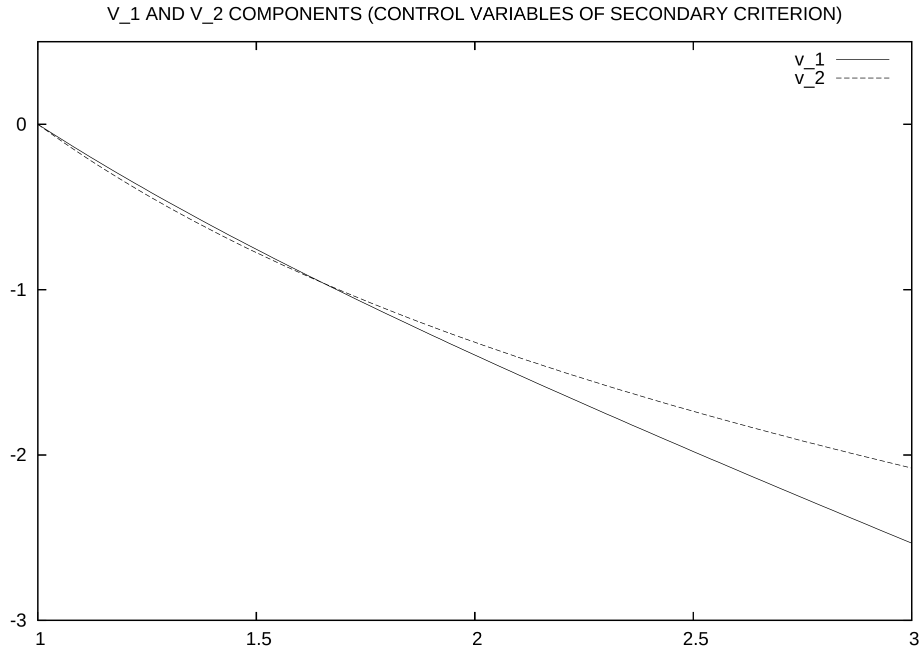


Figure 6: Concurrent minimization of the criteria J_A and J_B in the case of a nonlinear equality constraint; variation with A of the optimization variables v_1 and v_2 controling the minimization of J_B.

Contents

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399