Adaptation by Nash games in gradient-based multi-objective/multi-disciplinary optimization

Jean-Antoine Désidéri

Adaptation by Nash games in gradient-based multi-objective/multi-disciplinary optimization

J.-A. Désidéri [a]

[a] Université Côte d'Azur, Inria, CNRS, LJAD, 2004 Routes des Lucioles, 06902 Sophia Antipolis, France

ARTICLE HISTORY
Compiled June 6, 2021

ABSTRACT
A two-phase numerical process is proposed for gradient-based multi-disciplinary optimization. In a first phase, one or several Pareto-optimal solutions associated with a subset of the cost functions, the primary objective functions, subject to constraints, are determined by some effective multi-objective optimizer. In the second phase, a continuum of Nash equilibria is constructed tangent to the Primary Pareto Front along which secondary cost functions are to be reduced, while best preserving the Pareto-optimality of the primary cost functions. The focus of the article is on estimating the rate at which the secondary cost functions are diminished. The method is illustrated by the numerical treatment of the optimal sizing problem of a sandwich panel w.r.t. structural resistance under bending loads and blast mitigation.

KEYWORDS
gradient-based optimization, multi-objective optimization, multi-disciplinary optimization, descent direction, Nash game, adaptivity

AMS CLASSIFICATION
49M 49Q 90C 91A

1. Introduction

In engineering, the optimization of a complex system is almost invariably multi-objective, and this, for several reasons. Hence the problematics associated with multi-objective optimization has raised important developments in recent years [15].

Firstly, the problem is very often multi-disciplinary by the physics involved. To illustrate this, consider the prototypic example of the shape optimization of an aircraft wing, or configuration. Certainly the designer wishes to optimize the aerodynamic coefficients of the wing, six in total, since these impact directly the aircraft flight performance (mass, range, stability and maneuverability). However, during flight, the configuration is submitted to important loads and it must be assessed also structurally, statically and dynamically. Additionally, an aircraft is not a rigid body and a proper assessment of the effects requires the coupling of these disciplines (aero-elasticity). Other considerations are also accounted for, and somewhat differently for civil or military aircraft in relation to acoustics, propulsion, stealth, etc.

CONTACT J.-A. Désidéri. Email: jean-antoine.desideri@inria.fr

The multi-disciplinary aspect has fostered important developments in Industry, since it impacts the way different groups within the enterprise cooperate to assemble a coordinated concept. Multi-level strategies have been proposed to address this question, and in particular the BLISS method [27] is widely used.

In relation to this, speciliazed computing architectures have also been devised [22].

Secondly, each discipline may require a multi-point treatment, since the system can be operated in several ways. In the aeronautical design, the entire flight envelope must be analyzed. In particular, for a civil aircraft, the optimization of take-off requires to maximize lift in the subsonic-flow regime (small Mach number, large angle of attack), whereas cruise requires the minimization of drag under lift constraint in transonic flow conditions (high Mach number, small angle of attack) for minimum kerosene consumption and air pollution. Hence numerical simulation serves to analysis and optimization in turns.

An additional major difficulty resides in the fact that the critical disciplines are most often evaluated via the integration of a system of partial-differential equations (PDE). For the aircraft aerodynamics alone, the steady Euler equations in three dimensions may be considered sufficient for drag minimization in transonic conditions, since the most important effect is related to the main shock wave. However, in the take-off phase, during which high-lift devices are deployed, the complexity of the flow requires a simulation of the complete Reynolds-Averaged Navier-Stokes (RANS) equations, in steady or time-dependent form. For structural analysis, simulations involve either a simple structural model, such as a beam model, [19], or a more general elasticity model embedded in a complete finite-element discretization (e.g. [18]; for acoustics, a wave equation using the pressure field as input; for stealth, the Maxwell equations; etc. Hence the cost functions, and the constraints, the latter sometimes causing even greater difficulty, are usually functionals and in large number.

Furthermore, in a complex engineering optimization, even when the set of potential important cost functions may be known a priori exhaustively by experience, in a multi-point optimization, after the treatment of a reduced subset of preponderant criteria, those additional objective functions that require a subsequent correction may depend on the operational point that is selected and be revealed by the process itself. Hence a certain form of adaptivity is required.

Concerning the PDE-constrained optimization aspect, differential elements (gradients and Hessians, mostly) can be calculated by adjoint approaches, e.g. [12], automatic differentiation, e.g. [16], or well-calibrated finite-differences. These issues are not discussed here. In the second phase of our algorithm, the question is simplified by the construction of either global or local quadratic meta-models.

Numerous contributions have been made in the field of evolutionary strategies, and more specifically using genetic algorithms. Perhaps one of the most employed is the Multi-objective Genetic Algorithm NSGA-II [3], but variants demonstrate specific merits [3] [17] [24] [28] [25]. These methods only require procedures for function evaluations, and are usually the most robust in finding global optima. This, of course, at the expense of a high computational demand. Nevertheless, such methods can indeed be applied in complex engineering problems [26]. In particular, in the framework of evolutionary algorithms, Greiner *et al* [13] have documented a number of test-cases in structural engineering problems in which the multi-objective problematics is handled by a game formulation. They examined mostly Nash and Stackelberg games and their relationship with Pareto-optimality.

More recently, Bayesian Optimization (BO) methods have emerged as an alternative to evolutionary algorithms to solve global optimization problems using only

function evaluations [20]. These statistical approaches are especially well suited to the case of expensive evaluations and are designed to account for limited computational budgets. In the context of multi-objective optimization, approximations of the Pareto front can be obtained with a relatively moderate number of evaluations [10], typically some dozens, whereas genetic algorithms require some hundreds of evaluations, and sometimes more.

While evolutionary strategies act on populations that are iteratively enriched by application of pseudo-Darwinian operators, classical methods operate on individual design points. Certain algorithms have been devised to directly enrich the Pareto front from a set of initial points, in particular:

- the normal boundary intersection [2] aiming to produce evenly-distributed points on the Pareto set, and related weights;
- the normalized normal constraint method [23], which incorporates an additional filter for a more proper identification;
- the Pareto-front interpolation [14], in which the authors construct a sub-complex of a Delaunay triangulation of a finite set of Pareto optimal outcomes, and devise special rules for checking the inherent non-dominance of complexes; the method, was further developed in various publications, e.g. [15], and is supported by a surrogate model to alleviate the high computational cost of function(al) evaluations.

Alternately, several authors have investigated the possibility of extending the steepest-descent method to multi-objective optimization. Among them, note the method of Fliege and Svaiter [11].

In this article, the choice of the effective multi-objective optimizer to conduct the first phase of optimization is left open. Our focus is here on the second phase of our proposed adaptation strategy, meant to reduce secondary cost functions, along a continuum of Nash equilibria, originating from a design-point of the Primary Pareto Front. In this way, a certain form of adaptation is introduced in the entire multi-objective optimization process.

Our adaptation strategy is defined step by step in the next section that also includes fundamental results related to convex hulls and the definition of a descent direction common to a set of cost functions, excluding the case of a weakly-Pareto optimal point. In Section 3, the theoretical analysis is developed and the existence of the Nash equilibria and the efficiency of the method are established. In particular, the negative directional derivatives of the secondary cost functions at the initial point on the continuum of Nash equilibria are evaluated. In Section 4, the method is illustrated by the numerical treatment of the optimal sizing of a structural element according to analytical models. As expected, the plot of cost functions normalized to their initial values exhibits a characteristic funnel-shaped pattern.

2. Nash-game-based adaptation algorithm

2.1. Mathematical problem definition

One considers a multi-objective differentiable optimization problem involving a set of M cost functions $\{f_j(\mathbf{x})\}$ $(j = 1, \ldots, M;\ \mathbf{x} \in \Omega_a \subseteq \mathbb{R}^n)$, the first m $(1 \leq m < M)$ of which are thought of preponderant, or primary, and the remaining ones $(j = m+1, \ldots, M)$, secondary, and all are subject to K equality constraints, $c_k(\mathbf{x}) = 0$

$(k = 1, \ldots, K)$. Objective and cost functions are all smooth, say $C^2(\Omega_a)$.

A first phase of optimization (omitted here) has led to the knowledge of a point $\mathbf{x}_A^\star$ in the open interior of the admissible domain Ω_a that is Pareto-optimal for the minimization of the sole primary cost functions under the constraints.

The second phase of optimization is aimed to construct a continuum of Nash equilibria $\{\overline{\mathbf{x}}_\varepsilon\}$ parameterized by ε ($\varepsilon \geq 0$) in way such that: firstly, $\overline{\mathbf{x}}_0 = \mathbf{x}_A^\star$ (compatibility), and secondly, for ε sufficiently small, the Pareto-optimality condition of the primary cost functions is violated by a term $O(\varepsilon^2)$ only, whereas the secondary cost functions decrease at least linearly with ε.

2.2. Convex hull, common descent direction and directional derivatives

Problematics: Given m differentiable cost functions $\{f_j(\mathbf{x})\}$ $(j = 1, \ldots, m)$ $(\mathbf{x} \in \Omega_{ad} \subseteq \mathbb{R}^n)$, and their gradients at some admissible point $\mathbf{x}_0 \in \Omega_{ad}$,

$$\mathbf{g}_j = \nabla f_j(x_0), \tag{1}$$

one seeks a vector $\mathbf{d}^\star \in \mathbb{R}^n$ in the direction of which the derivatives of all cost functions are positive:

$$D_{\mathbf{d}^\star} f_j(x_0) = (\mathbf{g}_j)^t \mathbf{d}^\star \geq 0 \quad (\forall j) \tag{2}$$

(superscript ^{t} for transposition). When $\mathbf{x}_0$ is not (weakly) Pareto-optimal (that is, Pareto-stationary), one such direction exists, and if a strict inequality holds in (2), the vector $(-\mathbf{d}^\star)$ provides a descent direction common to all cost functions.

To construct one such direction, we consider an $n \times n$ positive-definite matrix $\mathbf{A}_n$, and define the scalar product in $\mathbb{R}^n$

$$(\mathbf{u} \in \mathbb{R}^n) \quad (\mathbf{v} \in \mathbb{R}^n) \quad \big(\mathbf{u}, \mathbf{v}\big) = \mathbf{u}^t \mathbf{A}_n \mathbf{v} \tag{3}$$

and associated Euclidean norm

$$(\mathbf{u} \in \mathbb{R}^n) \quad \|\mathbf{u}\| = \sqrt{\big(\mathbf{u}, \mathbf{u}\big)} = \sqrt{\mathbf{u}^t \mathbf{A}_n \mathbf{u}}, \tag{4}$$

and consider the convex hull of the family of gradients:

$$\overline{\mathbf{U}} = \left\{ \mathbf{u} \in \mathbb{R}^n \text{ such that: } \mathbf{u} = \sum_{j=1}^{m} \alpha_j \mathbf{g}_j,\ \alpha_j \geq 0\ (\forall j),\ \sum_{j=1}^{m} \alpha_j = 1 \right\} \tag{5}$$

which evidently, is a closed, bounded and convex set.

Proposition 2.1. *The convex hull $\overline{\mathbf{U}}$ admits a unique element $\boldsymbol{\omega}^\star$ of minimum Euclidean norm*

$$\boldsymbol{\omega}^\star = \arg\min_{\mathbf{u} \in \overline{\mathbf{U}}} \|\mathbf{u}\|. \tag{6}$$

Proof. Existence: the norm is continuous, and $\overline{\mathbf{U}}$ is closed and bounded. Uniqueness: $\overline{\mathbf{U}}$ is convex. □

Proposition 2.2. *Let* $\mathbf{d}^\star = \mathbf{A}_n\boldsymbol{\omega}^\star$. *The elements* $(\boldsymbol{\omega}^\star, \mathbf{d}^\star)$ *are such that:*

$$\forall \mathbf{u} \in \overline{\mathbf{U}}: \quad \big(\mathbf{u}, \boldsymbol{\omega}^\star\big) = \mathbf{u}^t\mathbf{d}^\star \geq \|\boldsymbol{\omega}^\star\|^2 \tag{7}$$

$$\forall j: \quad \big(\mathbf{g}_j, \boldsymbol{\omega}^\star\big) = (\mathbf{g}_j)^t\mathbf{d}^\star \geq \|\boldsymbol{\omega}^\star\|^2 \tag{8}$$

and if $\boldsymbol{\omega}^\star \neq 0$, *the vector* $(-\mathbf{d}^\star)$ *provides a common descent direction.*

Proof. Let $\mathbf{u} \in \overline{\mathbf{U}}$, arbitrary, and $\mathbf{v} = \mathbf{u} - \boldsymbol{\omega}^\star$. Since $\overline{\mathbf{U}}$ is convex, one has:

$$\forall \theta \in [0,1]: \ (1-\theta)\boldsymbol{\omega}^\star + \theta\mathbf{u} = \boldsymbol{\omega}^\star + \theta\mathbf{v} \in \overline{\mathbf{U}}, \tag{9}$$

and since $\boldsymbol{\omega}^\star$ is the minimum-norm element in $\overline{\mathbf{U}}$:

$$\|\boldsymbol{\omega}^\star + \theta\mathbf{v}\|^2 = \big(\boldsymbol{\omega}^\star + \theta\mathbf{v}, \boldsymbol{\omega}^\star + \theta\mathbf{v}\big) \geq \|\boldsymbol{\omega}^\star\|^2. \tag{10}$$

By developing the scalar product, one gets

$$\forall \theta \in [0,1]: \ 2\theta\big(\boldsymbol{\omega}^\star, \mathbf{v}\big) + \theta^2\|\mathbf{v}\|^2 \geq 0 \tag{11}$$

and this requires that the coefficient of θ, that is, $\big(\boldsymbol{\omega}^\star, \mathbf{v}\big) = \big(\mathbf{u} - \boldsymbol{\omega}^\star, \boldsymbol{\omega}^\star\big)$, be positive, which yields the conclusion. □

Without great loss of generality, we now assume that the gradients are distinct. Let $\mathcal{F} = \{\mathbf{g}_j\}$ $(j = 1, \ldots, m)$ denote their family. Let $\mathcal{F}_i$ be a subfamily of $\mathcal{F}$, $\overline{\mathbf{U}}_i$ its convex hull, and $\boldsymbol{\omega}_i^\star$ the minimum Euclidean norm element in $\overline{\mathbf{U}}_i$. Then, there is at least one subfamily $\mathcal{F}_i$ of minimal cardinal such that $\boldsymbol{\omega}_i^\star = \boldsymbol{\omega}^\star$. Such a subfamily may be the sole $\mathcal{F}$ itself. Inversely, it may not be unique. Hereafter, one such subfamily will be said to be a "minimal generating subfamily" in the sense that it contains the minimal number of distinct elements permitting to generate $\boldsymbol{\omega}^\star$ as a convex combination of its elements. Then, let $J^\star$ be the set of indices of the constitutive elements of the subfamily $\mathcal{F}_i$ so that:

$$\boldsymbol{\omega}^\star = \sum_{j \in J^\star} \alpha_j^\star \mathbf{g}_j \tag{12}$$

in which $J^\star$ is a subset of $\{1, 2, \ldots, m\}$, $\alpha_j^\star \geq 0$ $(\forall j \in J^\star)$ and $\sum_{j\in J^\star} \alpha_j^\star = 1$ for a convex combination. Furthermore, $\alpha_j^\star \neq 0$ $(\forall j \in J^\star)$ since the subfamily is of minimal cardinal; hence, $\alpha_j^\star > 0$ $(\forall j \in J^\star)$. Then:

Proposition 2.3. *Let* $\sigma = \|\boldsymbol{\omega}^\star\|^2$. *The directional derivatives satisfy the following:*

$$D_{\mathbf{d}^\star} f_j = (\mathbf{g}_j)^t\mathbf{d}^\star = \sigma \ (\forall j \in J^\star), \quad \textit{and} \ \geq \sigma \ \textit{otherwise.} \tag{13}$$

Proof. Let $\sigma_j = D_{\mathbf{d}^\star} f_j$ $(j \in J^\star)$, $\sigma_{\max} = \max_{(j\in J^\star)} \sigma_j$, and $\sigma_{\min} = \min_{(j\in J^\star)} \sigma_j$. Assume that $\sigma_{\max} \neq \sigma_{\min}$. Let $j_{\max} \in J^\star$ (resp. $j_{\min} \in J^\star$) be such that $\sigma_{j_{max}} = \sigma_{\max}$ (resp. $\sigma_{j_{min}} = \sigma_{\min}$). Recall that $\alpha_{j_{\max}}^\star > 0$, and consider strictly-positive numbers ϵ

such that $0 < \epsilon < \alpha^\star_{j_{\max}}$, and let

$$\omega = \omega^\star - \epsilon \mathbf{g}_{j_{\max}} + \epsilon \mathbf{g}_{j_{\min}} = \sum_{j \in J^\star} \alpha_j \mathbf{g}_j \tag{14}$$

where $\alpha_j = \alpha^\star_j$ except for $\alpha_{j_{\max}} = \alpha^\star_{j_{\max}} - \epsilon > 0$ and $\alpha_{j_{\min}} = \alpha^\star_{j_{\min}} + \epsilon > 0$. Clearly, for sufficiently small and positive ϵ, $\alpha_j > 0$ $(\forall j)$, and $\sum_{j \in J^\star} \alpha_j = 1$. Hence $\omega \in \overline{\mathbf{U}}$. Additionally:

$$\begin{aligned} \|\omega\|^2 = \big(\omega, \omega\big) &= \|\omega^\star\|^2 - 2\epsilon\big(\omega^\star, \mathbf{g}_{j_{\max}} - \mathbf{g}_{j_{\min}}\big) + \epsilon^2 \left\|\mathbf{g}_{j_{\max}} - \mathbf{g}_{j_{\min}}\right\|^2 \\ &= \|\omega^\star\|^2 - 2\epsilon\left(\sigma_{\max} - \sigma_{\min}\right) + \epsilon^2 \left\|\mathbf{g}_{j_{\max}} - \mathbf{g}_{j_{\min}}\right\|^2 < \|\omega^\star\|^2 \end{aligned} \tag{15}$$

for all sufficiently small ϵ, and this is a contradiction since $\boldsymbol{\omega}^\star$ is the minimum-norm element in $\overline{\mathbf{U}}$. The contradiction is removed by abandoning the hypothesis $\sigma_{\max} \neq \sigma_{\min}$. Hence $\sigma_j = \sigma = \|\omega^\star\|^2$ $(\forall j \in J^\star)$. Lastly, for $j \notin J^\star$, $\sigma_j \geq \sigma$ by virtue of Proposition 2.2. □

Proposition 2.4. *Let $\mathcal{F}_i$ be a minimal generating subfamily. Then, only two cases are possible:*

(1) either $\boldsymbol{\omega}^\star = 0$ in a situation of Pareto-stationarity;
(2) or the subfamily $\mathcal{F}_i$ is linearly independent.

Proof. By induction on m.

If $m = 2$, and if the two vectors are linearly dependent, either they are in opposite direction (Pareto stationarity), or the one of smaller norm constitutes alone the subfamily $\mathcal{F}_i$. Otherwise the two vectors form a linearly independent family.

Now, assume the statement is true when $m \leq \mu$, for some μ, and let us examine a case for which $m = \mu + 1$. In an affine-space representation, the vectors are given an origin O, and the convex hull associated with $\mathcal{F}_i$ is viewed as a polytope. If O belongs to the interior of this polytope, the gradients are in a situation of Pareto stationarity (1st case). Otherwise, $\boldsymbol{\omega}^\star$ is identified by projection of O onto a boundary element, and the subfamily of vectors defining this boundary element constitute a minimal generating subfamily and it contains at most $m - 1 \leq \mu$ elements: these, by inductive assumption, must be associated with linearly independent vectors. □

The configuration of the convex hull is illustrated in Appendix A in various cases of two and three gradients vectors.

Proposition 2.4 provides one of the possible ways to compute the element $\boldsymbol{\omega}^\star$. In situations other than Pareto stationary, one way to proceed is to explore recursively the boundary of the polytope to identify a minimal generating subfamily. Then, one can evaluate the matrix $\mathbf{A}_n$ and the related metrics associated with the basis change from the canonical basis to this subset of gradients. In the new basis, $\boldsymbol{\omega}^\star$ is associated with the bisector of the coordinate system and the back transformation yields $\boldsymbol{\omega}^\star$ explicitly in the canonical basis [5][4].

In all what follows, the standard Euclidean scalar product is used:

$$\mathbf{A}_n = \mathbf{I}_n, \quad \mathbf{d}^\star = \boldsymbol{\omega}^\star. \tag{16}$$

2.3. Preliminary calculations

The preliminary calculations are made of three main steps:

- Definition of a convex "Primary Steering Function" $f_A^+(\mathbf{x})$
- Definition of a "Territory Splitting"
- Definition of a "Secondary Steering Function" $f_B(\mathbf{x})$.

The cost functions are all assumed to be strictly positive. This may require the preliminary application of an adequate exponential transform, or the like. The exponential transform additionally permits us to consider gradients in the logarithmic form, and examine the variations of these functions in terms of proportions. Furthermore, this tends to improve convexity.

The Pareto-stationarity condition satisfied at $\mathbf{x} = \mathbf{x}_A^\star$ is expressed as follows:

$$\sum_{j=1}^{m} \alpha_j^\star \frac{\mathbf{P}\nabla f_j^\star}{f_j^\star} = 0 \tag{17}$$

where the supercript * on any symbol denotes throughout an evaluation at $\mathbf{x} = \mathbf{x}_A^\star$, ∇ is the gradient operator w.r.t. $\mathbf{x}$, $\{\alpha_j^\star\}$ $(j = 1, \ldots, m)$ are the known coefficients of a convex combination, $\mathbf{P}$ is the projection operator onto the linear subspace orthogonal to all constraint gradients, $\{\nabla c_k^\star\}$ $(k = 1, \ldots, K)$. Then the primary steering function is defined as follows:

$$f_A^+(\mathbf{x}) = \sum_{j=1}^{m} \alpha_j^\star \frac{f_j(\mathbf{x})}{f_j^\star} + \frac{c}{2} \left\| \mathbf{x} - \mathbf{x}_A^\star \right\|^2 \tag{18}$$

where c is a nonnegative constant adjusted to enforce the local convexity of $f_A^+(\mathbf{x})$ and the associated Lagrangian. Then, by virtue of (17), the primary steering function is stationary and convex at $\mathbf{x}_A^\star$, and therefore achieves a local minimum.

The territory splitting is based on the orthogonal decomposition of the following $n \times n$ reduced Hessian matrix [7, 8]:

$$\mathbf{H}'_A = \mathbf{P}\mathbf{H}_A^{+,\star}\mathbf{P} = \mathbf{\Omega}\mathcal{H}\mathbf{\Omega}^t \tag{19}$$

where $\mathbf{H}_A^{+,\star} = (\nabla\nabla^t f_A^+)^\star = \mathbf{H}_A^\star + c\mathbf{I}_n$ and $\mathbf{H}_A^\star = \nabla\nabla^t f_A^\star$, and $\nabla\nabla^t$ is the Hessian operator. Matrix $\mathbf{H}'_A$ is positive semi-definite, of rank $n - K$ exactly since matrix $\mathbf{H}_A^{+,\star}$ is strictly-positive definite by convexity fix. The eigenvectors, column-vectors of matrix $\mathbf{\Omega}$, are conventionally arranged in a special way: the first K are associated with the directions of the constraint gradients (null space of $\mathbf{P}$), and the remaining ones are ordered by decreasing eigenvalue. In this way, the tail eigenmodes are associated with the smaller sensitivities of the primary steering function. The splitting is associated with the following change of variables:

$$\mathbf{x} = \mathbf{x}_A^\star + \mathbf{\Omega} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} := \mathbf{X}(\mathbf{u}, \mathbf{v}) \tag{20}$$

where $\mathbf{u} \in \mathbb{R}^{n-p}$, $\mathbf{v} \in \mathbb{R}^p$, and p is a free integer parameter to be chosen such that and

$1 \leq p < n - K$. In this setting, the following notations are used:

$$\mathbf{\Omega} = (\mathbf{\Omega_u}\ \mathbf{\Omega_v})\,, \quad \mathbf{\Omega_u} = \begin{pmatrix} \mathbf{\Omega_{uu}} \\ \mathbf{\Omega_{vu}} \end{pmatrix}, \quad \mathbf{\Omega_v} = \begin{pmatrix} \mathbf{\Omega_{uv}} \\ \mathbf{\Omega_{vv}} \end{pmatrix} \tag{21}$$

and the following scaled logarithmic gradients are associated with the secondary cost functions:

$$\mathbf{g}_j = \mathbf{S}^{-\frac{1}{2}} \frac{\nabla_{\mathbf{v}} f_j^\star}{f_j^\star} = \mathbf{S}^{-\frac{1}{2}} \frac{\mathbf{\Omega}_{\mathbf{v}}^t \nabla f_j^\star}{f_j^\star} \qquad (j = m+1, \ldots, M) \tag{22}$$

where the scaling matrix $\mathbf{S} = (\nabla\nabla^t_{\mathbf{vv}} f_A^+)^\star$ is the lower $p \times p$ diagonal block of matrix $\mathcal{H}$ of (19), positive-definite by convexity fix. Hence these gradients reflect the sensitivities of the secondary cost functions when $\mathbf{v}$ varies, and a scaling by elements related to the primary steering function as well. Following the strategy employed in the Multiple Gradient Descent Algorithm (MGDA) [7], one identifies the minimum-norm element $\boldsymbol{\omega}_B^\star$ in the convex hull of the gradients $\mathbf{g}_j$:

$$\boldsymbol{\omega}_B^\star = \sum_{j=m+1}^{M} \alpha_j^\star \mathbf{g}_j \tag{23}$$

and the coefficients $\{\alpha_j^\star\}$ of this convex combination are used to define the secondary steering function:

$$f_B(\mathbf{x}) = \sum_{j=m+1}^{M} \alpha_j^\star \frac{f_j(\mathbf{x})}{f_j^\star} \tag{24}$$

and this completes the preliminaries.

2.4. Nash game formulation and statement of theoretical result

One defines a continuation parameter ε ($0 \leq \varepsilon \leq 1$) and the auxiliary cost function:

$$f_{AB}(\mathbf{x}) = (1-\varepsilon) f_A^+(\mathbf{x}) + \varepsilon f_B(\mathbf{x}) \tag{25}$$

For every fixed value of ε, two virtual players, *Player A* and *Player B*, respectively controling the sub-vectors $\mathbf{u}$ and $\mathbf{v}$ in (20), are involved in a Nash game defined by the following strategies:

– Strategy A : *Player A* attempts to minimize $f_A^+\big(\mathbf{X}(\mathbf{u},\mathbf{v})\big)$ by the strategy $\mathbf{u}$, subject to the equality constraint $\mathbf{c}\big(\mathbf{X}(\mathbf{u},\mathbf{v})\big) = 0$, and by accounting for *Player B*'s fixed strategy $\mathbf{v}$,

whereas:

– Strategy B : *Player B* attempts to minimize $f_{AB}\big(\mathbf{X}(\mathbf{u},\mathbf{v})\big)$ by the strategy $\mathbf{v}$, subject to no constraints, but by accounting for *Player A*'s fixed strategy $\mathbf{u}$.

The following results hold:

- The Nash equilibrium point $\overline{\mathbf{x}}_\varepsilon = \mathbf{X}\left(\overline{\mathbf{u}}_\varepsilon, \overline{\mathbf{v}}_\varepsilon\right)$ exists for all sufficiently small ε, and

$$\overline{\mathbf{x}}_0 = \mathbf{x}_A^\star \tag{26}$$

  that is, $\overline{\mathbf{u}}_0 = 0 \in \mathbb{R}^{n-p}$ and $\overline{\mathbf{v}}_0 = 0 \in \mathbb{R}^p$, an essential property referred to as compatibility or consistency.
- Let $\phi_A(\varepsilon) = f_A^+\left(\overline{\mathbf{x}}_\varepsilon\right)$. Then $\phi'_A(0) = 0$ and

$$f_A^+ = 1 + O(\varepsilon^2) \tag{27}$$

  which reflects the fact that, as ε varies, the Pareto-optimality of the primary cost functions is degraded by $O(\varepsilon^2)$ only.
- Let $\phi_B(\varepsilon) = f_B\left(\overline{\mathbf{x}}_\varepsilon\right)$. Then $\phi'_B(0) = -\sigma_B$, where $\sigma_B = \left\|\boldsymbol{\omega}_B^\star\right\|^2 \geq 0$. Hence, the secondary steering function $f_B(\mathbf{x})$ decreases linearly as ε increases, and by virtue of (13), all the secondary cost functions do as well, or faster.

Thus the objectives of our construction are achieved.

3. Convergence rate

3.1. Compatibility

Proposition 3.1. *For $\varepsilon = 0$, the Nash equilibrium exists and corresponds to $\overline{\mathbf{x}}_0 = \mathbf{x}_A^\star$.*

Proof. For $\varepsilon = 0$, $f_{AB} = f_A^+$ and both players attempt to reduce f_A^+, but by acting in different subdomains: *Player A* acts on the subvector $\mathbf{u}$, for fixed $\mathbf{v}$, and under the constraint $\mathbf{c}\big(\mathbf{X}(\mathbf{u},\mathbf{v})\big) = 0$, whereas *Player B* acts on subvector $\mathbf{v}$, for fixed $\mathbf{u}$, and under no constraints. We have to prove that this concurrency results in the equilibrium

$$\mathbf{u} = \overline{\mathbf{u}}_0 = 0 \in \mathbb{R}^{n-p}, \quad \mathbf{v} = \overline{\mathbf{v}}_0 = 0 \in \mathbb{R}^p, \tag{28}$$

for which $\overline{\mathbf{x}} = \mathbf{x}_A^\star$. Equivalently: is $(\overline{\mathbf{u}}_0, \overline{\mathbf{v}}_0) = (0,0) \in \mathbb{R}^{n-p} \times \mathbb{R}^p$ a solution to the following two uncoupled problems:

$$(A)\ \begin{cases} \min_{\mathbf{u}} f_A^+\big(\mathbf{X}(\mathbf{u},0)\big) \\ \text{s. t. } \mathbf{c}\big(\mathbf{X}(\mathbf{u},0)\big) = 0 \end{cases} \qquad (B)\ \begin{cases} \min_{\mathbf{v}} f_A^+\big(\mathbf{X}(0,\mathbf{v})\big) \\ \text{s. t. no constraints} \end{cases} \tag{29}$$

The affirmative answer is evident for *Subproblem (A)* which is a reformulation of the primary multi-objective problem in a subdomain where $\mathbf{u} = 0$ yields the solution $\mathbf{x}_A^\star$.

Concerning *Subproblem (B)*, note that it is the unconstrained minimization of a function that is convex by virtue of the convexity fix (18). Hence only stationarity remains to be verified for $\mathbf{v} = 0$, that is, $\mathbf{x} = \mathbf{x}_A^\star$. For this, consider an index $j \leq p$:

$$\left(\frac{\partial f_A^+}{\partial v_j}\right)^\star = \nabla f_A^{+,\star}.\frac{\partial \mathbf{x}}{\partial v_j} = \nabla f_A^{+,\star}.\left[\boldsymbol{\Omega}\right]_{n-j+1} \tag{30}$$

where the notation $[\boldsymbol{\Omega}]_\ell$ is used for the ℓth column vector of matrix $\boldsymbol{\Omega}$, that is, the ℓth eigenvector of the reduced Hessian matrix $\mathbf{H}'_A$ of (19). Now, by optimality of $\mathbf{x}_A^\star$:

$\nabla f_A^{+,\star} = -\sum_{k=1}^{K} \lambda_k \nabla c_k^\star \in Span\{\nabla c_k^\star\} = \mathrm{Ker}(\mathbf{P}) = Span\{[\mathbf{\Omega}]_1, \ldots, [\mathbf{\Omega}]_K\}$. Recall that the matrix $\mathbf{\Omega}$ is orthogonal, and its column vectors 2 by 2 orthogonal. Thus, provided that $K < n - p + 1$, $n - j + 1 > K$ and $[\mathbf{\Omega}]_{n-j+1}$ is orthogonal to this span and the above scalar product is indeed equal to 0. This confirms that *Subproblem (B)* does admit $\mathbf{v} = 0$ for trivial solution. □

3.2. Continuum of Nash equilibria

Proposition 3.2. *Suppose that the following bounds are imposed on the subvectors* $\mathbf{u}$ *and* $\mathbf{v}$*:*

$$\|\mathbf{u}\| \leq B_{\mathbf{u}}, \quad \|\mathbf{v}\| \leq B_{\mathbf{v}}, \tag{31}$$

Then, the Nash equilibrium point exists for all sufficiently small ε*.*

Proof. We begin by citing literally a theorem by Nash given in [1] (p. 268):

Theorem 3.3 (Nash)**.** *Suppose that the multistrategy set*
(i) $X(N)$ *is a convex compact subset*
and that, for each player i*, the loss function*
(ii) f_i *is continuous and* $f_i(., \mathbf{x}^i)$ *is convex for all* $\mathbf{x}^i \in X^i$*.*
Then there exists a non-cooperative equilibrium.

Here, the strategy subsets of the two players are:

$$X^A = \left\{ \mathbf{u} \in \mathbb{R}^{n-p} \ / \ \|\mathbf{u}\| \leq B_{\mathbf{u}} \right\}, \quad X^B = \left\{ \mathbf{v} \in \mathbb{R}^p \ / \ \|\mathbf{v}\| \leq B_{\mathbf{v}} \right\}. \tag{32}$$

These are finite-dimensional closed balls, thus indeed convex compact subsets.

Player A, in effect, minimizes the Lagrangian associated to f_A^+ w.r.t. its strategy $\mathbf{u}$. By the adjustment of the convexity-fix constant c, this function is strictly convex, and so is also the corresponding partial function of this player's strategy.

Player B minimizes the cost function f_{AB}^+ w.r.t. its strategy $\mathbf{v}$. This function is strictly-convex for all sufficiently small ε, and so is also the corresponding partial function of this player's strategy.

Hence Theorem 3.3 applies, and this establishes Proposition 3.2. □

Furthermore, in all what follows, we assume that the dependency of the Nash equilibrium solution $\mathbf{x}_\varepsilon$ on the continuation parameter ε is smooth, and refer to the continuum of Nash equilibria as "the continuum":

Hypothesis 1. For strictly-positive but small enough ε, the Nash equilibrium solution $\overline{\mathbf{x}}_\varepsilon$, as a function of ε, is assumed to have C^2-regularity.

Remark 1. At the origin of the continuum, $(\bar{\mathbf{u}}_0, \bar{\mathbf{v}}_0) = 0$ and the bounds in (31) are satisfied trivially. Since continuity is assumed, this remains true for small enough ε. In this sense, these bounds are not really stringent.

3.3. Variation of the primary cost functions along the continuum

Since the constraints

$$c_k\left(\overline{\mathbf{x}}_\varepsilon\right) = 0 \quad (\forall k) \tag{33}$$

are satisfied all along, by differentiating the above w.r.t. ε (denoted by $'$) and setting $\varepsilon = 0$ (denoted by the subscript $_0$), one gets the following initial directional derivatives:

$$(c_k)'_0 = \nabla c_k^\star.\overline{\mathbf{x}}'_0 = 0 \quad (\forall k). \tag{34}$$

But, due to the optimality condition,

$$\nabla f_A^\star = -\sum_{k=1}^{K} \lambda_k \nabla c_k^\star \quad (\lambda_k : \text{ Lagrange multiplier}), \tag{35}$$

one gets accordingly:

$$(f_A)'_0 = \nabla f_A^\star.\overline{\mathbf{x}}'_0 = 0, \tag{36}$$

that is:

$$f_A = 1 + O(\varepsilon^2) \tag{37}$$

(and the same holds for f_A^+). Consequently, the Pareto-stationarity condition is degraded by a term $O(\varepsilon^2)$ only.

3.4. Variation of the secondary cost functions along the continuum

3.4.1. Estimating $\overline{\mathbf{v}}$

By definition of *Subproblem (B)*:

$$\overline{\mathbf{v}} = \arg_{\mathbf{v}} \min f_{AB} \quad (\mathbf{u} \text{ fixed}) \tag{38}$$

in which f_{AB} is defined by (25) and expanded as follows in terms of $\mathbf{v}$ for fixed $\mathbf{u}$:

$$f_{AB}(\mathbf{v}) = f_{AB}(0) + \nabla_{\mathbf{v}} f_{AB}.\mathbf{v} + \tfrac{1}{2}\mathbf{v}.\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_{AB}\mathbf{v} + O(\varepsilon^3). \tag{39}$$

Hence, optimality requires that

$$\nabla_{\mathbf{v}} f_{AB} + \nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_{AB}\mathbf{v} + O(\varepsilon^2) = 0 \tag{40}$$

which gives the following estimate:

$$\overline{\mathbf{v}} = -(\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_{AB})^{-1}\nabla_{\mathbf{v}} f_{AB} + O(\varepsilon^2). \tag{41}$$

Now, as $\varepsilon \to 0$,

$$\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_{AB} = (1-\varepsilon)\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_A^+ + \varepsilon\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}^t f_B = \mathbf{S} + O(\varepsilon) \tag{42}$$

where $\mathbf{S} = \left(\nabla_{\mathbf{v}}\nabla_{\mathbf{v}} f_A^+\right)^\star$. On the other hand:

$$\nabla_{\mathbf{v}} f_{AB} = (1-\varepsilon)\nabla_{\mathbf{v}} f_A^+ + \varepsilon \nabla_{\mathbf{v}} f_B \tag{43}$$

and: firstly, $\nabla_{\mathbf{v}} f_A^+ = O(\varepsilon^2)$ since f_A^+, as a function of ε, differs from 1 by $O(\varepsilon^2)$ uniformly; and secondly, $\nabla_{\mathbf{v}} f_B = \nabla_{\mathbf{v}} f_B^\star + O(\varepsilon)$. This gives:

$$\nabla_{\mathbf{v}} f_{AB} = \varepsilon \nabla_{\mathbf{v}} f_B^\star + O(\varepsilon^2). \tag{44}$$

Finally:

$$\overline{\mathbf{v}} = -\varepsilon \mathbf{S}^{-1} \nabla_{\mathbf{v}} f_B^\star + O(\varepsilon^2). \tag{45}$$

We now turn to the evaluation of matrix $\mathbf{S}$. Observe that:

$$\nabla_{\mathbf{v}} = \mathbf{\Omega}_{\mathbf{v}}^t \nabla, \quad \nabla_{\mathbf{v}} \nabla_{\mathbf{v}}^t = \mathbf{\Omega}_{\mathbf{v}}^t \nabla \nabla^t \mathbf{\Omega}_{\mathbf{v}} \tag{46}$$

$$\mathbf{S} = \mathbf{\Omega}_{\mathbf{v}}^t \mathbf{H}_A^{+\star} \mathbf{\Omega}_{\mathbf{v}} = \mathbf{\Omega}_{\mathbf{v}}^t \Big[\mathbf{P} + (\mathbf{I}_n - \mathbf{P})\Big] \mathbf{H}_A^{+\star} \Big[\mathbf{P} + (\mathbf{I}_n - \mathbf{P})\Big] \mathbf{\Omega}_{\mathbf{v}}. \tag{47}$$

Now, since $(\mathbf{I}_n - \mathbf{P})$ is the matrix associated with the orthogonal projection onto the subspace spanned by the constraint gradients, the column-vectors of matrix $\mathbf{\Omega}_{\mathbf{v}}$ are orthogonal to them. Consequently:

$$(\mathbf{I}_n - \mathbf{P})\mathbf{\Omega}_{\mathbf{v}} = 0_{n\times(n-p)}, \quad \mathbf{\Omega}_{\mathbf{v}}^t (\mathbf{I}_n - \mathbf{P}) = 0_{(n-p)\times n}, \tag{48}$$

so that

$$\mathbf{S} = \mathbf{\Omega}_{\mathbf{v}}^t \mathbf{P} \mathbf{H}_A^{+\star} \mathbf{P} \mathbf{\Omega}_{\mathbf{v}} = \mathbf{\Omega}_{\mathbf{v}}^t \mathbf{H}_A^{+\prime} \mathbf{\Omega}_{\mathbf{v}} = \mathbf{\Omega}_{\mathbf{v}}^t \mathbf{\Omega} \mathcal{H} \mathbf{\Omega}^t \mathbf{\Omega}_{\mathbf{v}} = \mathcal{H}_{\mathbf{v}} \tag{49}$$

in which $\mathcal{H}_{\mathbf{v}}$ is the "$\mathbf{v}$ diagonal block", that is, the lower $p \times p$ diagonal block of the positive-definite diagonal matrix $\mathcal{H}$ of (19). Hence the matrix $\mathbf{S}$ is diagonal positive-definite since $f_A^{+,\star}$ is convex.

3.4.2. Estimating $\overline{\mathbf{u}}$

By construction, the column vectors of matrix $\mathbf{\Omega}$ are the orthogonal eigenvectors of matrix $\mathbf{H}'_A$ of (19). The vector $\mathbf{\Omega}_{\mathbf{u}}\mathbf{u}$ is in the span of the first $n-p$ ($n-p > K$) which contains the null space of the projection matrix $\mathbf{P}$ onto the subspace tangent to the constraint manifold. Thus, this vector is orthogonal to the constraint manifold, and the vector $\mathbf{\Omega}_{\mathbf{v}}\mathbf{v}$ is orthogonal to it. The sum of these two vectors is a "chord" for the manifold since both points $\mathbf{x}_A^\star$ and $\overline{\mathbf{x}}_\varepsilon$ satisfy the constraints. Consequently, assuming regularity of the manifold:

$$\|\mathbf{\Omega}_{\mathbf{u}}\overline{\mathbf{u}}\| = O\left(\|\mathbf{\Omega}_{\mathbf{v}}\overline{\mathbf{v}}\|^2\right). \tag{50}$$

Remark 2. This result describes how the Nash game does preserve the optimum of the primary cost function to second-order in ε: vector $\mathbf{u}$ alters f_A^+ and f_A by a term $O(\varepsilon^2)$, whereas vector $\mathbf{v}$ has no effect on it (to second-order) since $\mathbf{\Omega}_{\mathbf{v}}\mathbf{v} \perp \nabla f_A^{+\star} = \nabla f_A^\star$.

3.5. Secondary cost functions

The variation of the secondary cost functions is examined by letting:

$$\phi_j(\varepsilon) = \frac{f_j(\bar{\mathbf{x}}_\epsilon)}{f_j^\star} \quad (j = m+1, \ldots, M). \tag{51}$$

This function is expanded in terms of ε using the established asymptotics:

$$\phi_j(\varepsilon) - 1 = \frac{1}{f_j^\star} \nabla f_j^\star . \big(\mathbf{\Omega_u u} + \mathbf{\Omega_v v} \big) + O(\varepsilon^2) = O(\varepsilon^2) - \frac{\varepsilon}{f_j^\star} \nabla f_j^\star . \mathbf{\Omega_v} \mathbf{S}^{-1} \nabla_{\mathbf{v}} f_B^\star \tag{52}$$

and this provides the derivative:

$$\phi_j'(0) = -\frac{1}{f_j^\star} \nabla f_j^\star . \mathbf{\Omega_v} \mathbf{S}^{-1} \nabla_{\mathbf{v}} f_B^\star = -\frac{1}{f_j^\star} (\nabla_{\mathbf{v}} f_B^\star)^t \mathbf{S}^{-1} \nabla_{\mathbf{v}} f_j^\star \tag{53}$$

To symmetrize this formula, the following new vector variable is introduced

$$\mathbf{w} = \mathbf{S}^{\frac{1}{2}} \mathbf{v} \tag{54}$$

so that $\nabla_{\mathbf{w}} = \mathbf{S}^{-\frac{1}{2}} \nabla_{\mathbf{v}}$ (- recall that $\mathbf{S}$ is positive-definite -) and (53) is rewritten:

$$\phi_j'(0) = -(\nabla_{\mathbf{w}} f_B^\star)^t \nabla_{\mathbf{w}} (f_j / f_j^\star) == -\mathbf{g}_j^t \boldsymbol{\omega}_B^\star \tag{55}$$

where the definitions made in (22)-(23) have been used. Lastly, the directional derivative is given by (13):

$$\phi_j'(0) = -\sigma_B \quad (j \in J_B^\star), \quad \leq -\sigma_B \text{ otherwise} \tag{56}$$

where $\sigma_B = \|\boldsymbol{\omega}_B^\star\|^2 \geq 0$, and $J_B^\star$ is a set of indices of minimal cardinal permitting to define the convex hull of the secondary gradients (w.r.t. $\mathbf{w}$).

Equation (56) establishes that our strategy for reducing the secondary cost functions linearly with ε is effective for small enough ε in all cases except if the gradients of the secondary cost functions w.r.t. the variable $\mathbf{w}$ are in a configuration of Pareto stationarity for which $\boldsymbol{\omega}_B^\star = 0$.

In effect, the steering function f_B forces all secondary cost functions to diminish at the same rate, at least, and this completes the proof of the theoretical statement. □

Remark 3. The definition of the variable $\mathbf{w}$, (54), is informative of how the negotiation between the two virtual players does operate. Intuitively, one would have guessed that the secondary cost functions react according to their partial gradients w.r.t. $\mathbf{v}$ since this vector constitutes *Player B*'s strategy. However, it is not exactly so. In effect, these partial gradients are scaled by the inverse square root of the block $\mathbf{S}$, and this reflects how *Player B* accounts for *Player A*'s strategy.

4. Illustration of the prioritized multi-objective optimization approach

4.1. Numerical implementation

The present method can be applied formally to simple cases defined analytically. In the more general case in which the formal derivation is not at hand, we propose to resort to models of substitution, and formulate the Nash game in terms of these meta-models. Precisely, quadratic models can be constructed once for all for the two steering functions based on local differential elements at $\mathbf{x} = \mathbf{x}_A^\star$ (typically approximated by central differencing), and function values over a database to account on second derivatives more globally. For constraints, quadratic meta-models can also be constructed, but it is recommended to upgrade these models, as the determination of the continuum of Nash equilibria proceeds, because constraints participate critically to the definition of Pareto-optimality, and this notion cannot be altered legitimately. Such techniques have been implemented over the software platform `http://mgda.inria.fr` according to procedures fully defined in [6].

4.2. Application to the sizing of a sandwich element for optimal mechanical resistance

The present technique of prioritized optimization has been applied to four different test-cases with respect to 4 parameters that size an aluminum sandwich panel in order to maximize its mechanical performance in terms of strength under bending loads and blast mitigation [9]. Two of these test-cases are briefly summarized here. The reader is directed to the report for further methodological details and the precise data used in the numerical experiments.

4.2.1. Geometry.

The structural element under consideration is a three-layer sandwich aluminum panel depicted in Figure 1. The thicknesses of the upper and lower layers and core are respectively $\mathbf{t}_u$, $\mathbf{t}_l$ and $\mathbf{t}_c$. The following dimensions are used throughout $b = 5$ cm, $\ell = 50$ cm.

The core is made of honeycomb foam. Such structural elements are widely used by industrial manufacturers for their exceptional mechanical properties. Illustrations can be found in many engineering sites. In the present case, the honeycomb is symmetrical and its dimensions are defined Figure 1.

In summary, the thicknesses $\mathbf{t}_u$, $\mathbf{t}_l$ and $\mathbf{t}_c$, as well as the ratio $\mathbf{R} = t/h$ of thickness to length of the honeycomb cell walls (see Figure 1) constitute the design optimization variables.

4.2.2. Constraints

Expressing all sizes in meters, the design variables are subject to the following interval constraints

$$\mathbf{t}_u,\ \mathbf{t}_l \in [0.03, 0.14], \quad \mathbf{t}_c \in [0.05, 0.19], \quad \mathbf{R} \in [0.01, 0.20], \tag{57}$$

and the total thickness to the following bound:

$$e = \mathbf{t}_u + \mathbf{t}_l + \mathbf{t}_c \leq 0.25. \tag{58}$$

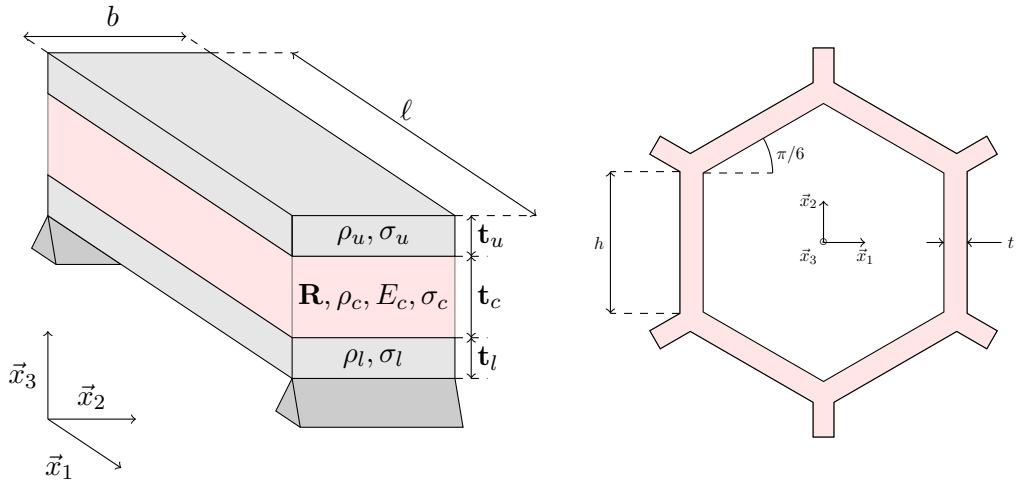


Figure 1. Geometry of the sandwich panel (left) and honeycomb core (right)

4.2.3. Parameterization

In order to satisfy the interval constraints automatically, the following change of variables is adopted:

$$\begin{aligned}\mathbf{t}_u &= 0.085 + 0.055 \sin x_1 \\ \mathbf{t}_l &= 0.085 + 0.055 \sin x_2 \\ \mathbf{t}_c &= 0.120 + 0.070 \sin x_3 \\ \mathbf{R} &= 0.105 + 0.095 \sin x_4.\end{aligned}$$

where the variables x_i's are initially set in the interval $[-\pi/2, \pi/2]$ but may not remain in it in the course of the optimization.

Besides, in order to satisfy the bound constraint on the total thickness, an additional variable is introduced, x_5, as a slack variable, and the constraint is transformed into the following equality constraint:

$$c_1(\mathbf{x}) = 0.29 + 0.055\,(\sin x_1 + \sin x_2) + 0.070 \sin x_3 - 0.25 + x_5^2 = 0. \qquad (59)$$

In this way, the problem involves 5 variables, and 1 constraint of equality type.

4.2.4. Cost functions.

The following five dimensionless cost functions have considered:

- φ_m representing mass (per unit transversal area);
- (φ_1, φ_2), inversely proportional to the critical failure forces $(F_{c,1}, F_{c,2})$ under bending loads, according to core indentation (1st mode) and exceeded lower-skin elastic resistance (second mode);
- φ_u, inversely proportional to the energy absorbed by the core under blast;
- φ_w, proportional to the panel central point deflection under blast.

All five cost functions are calculated from the geometrical definition, and thus in terms of $\mathbf{x}$, according to analytical models [21]-[9].

4.2.5. Equating the critical flexural forces.

The first numerical experiment relates to mass and critical flexural forces only. Suppose that a first phase of optimization has consisted of the concurrent minimization of $f_1 = \varphi_m$ and $f_2 = \varphi_1$, under constraint (59), and resulted in the "Primary Pareto Front" represented by magenta symbols on Figure 2. Then, the corresponding Pareto-optimal designs have been evaluated w.r.t. the secondary objective function $f_3 = \varphi_2$. It then appears that for most of these designs, $f_3 \gg f_2$ and that is $F_{c,2} \ll F_{c,1}$. In other words, the second mode of failure can be triggered much prior to the first, and this not acceptable.

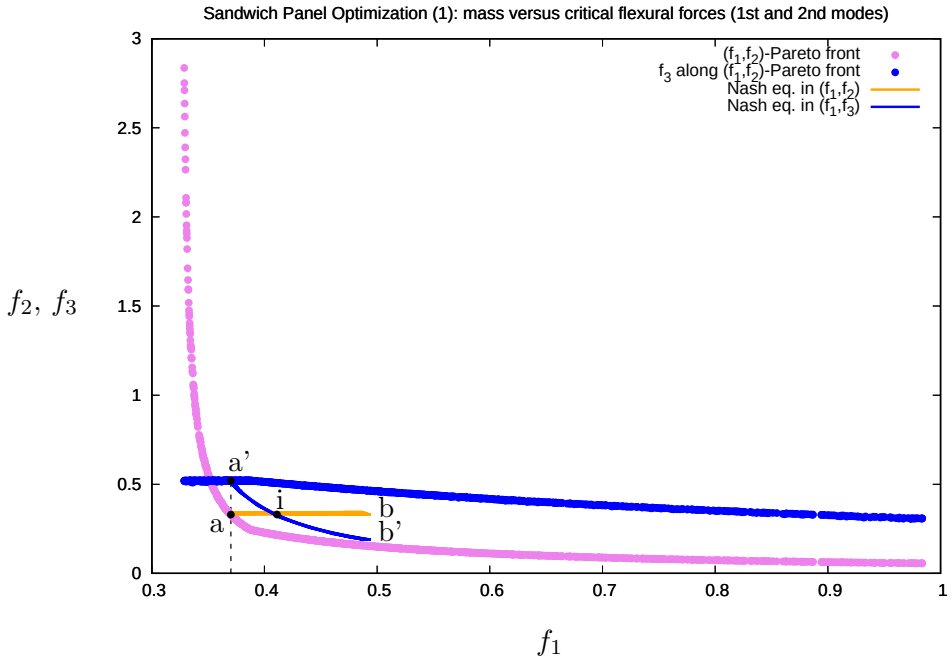


Figure 2. Sandwich Panel Optimization (1): Primary Pareto Front (byzantine) and continuum of Nash equilibria (yellow) ($f_1 = \varphi_m,\ f_2 = \varphi_1,\ f_3 = \varphi_2$)

However, starting from such a design, the prioritized optimization procedure can be used to correct this deficiency, and achieve a design for which both critical forces, $F_{c,1}$ and $F_{c,2}$, are equal to a value only slightly inferior to the original value of $F_{c,1}$.

As an example, the (f_1, f_2)-Pareto-optimal design point for which $f_1 \approx 0.3702$ was considered, and defined as the starting point $\mathbf{x}_A^\star$ to initiate the computation of the continuum of Nash equilibria conducted to reduce f_3 using the MGDA platform [4]. The continuum of Nash equilibria is represented by the yellow path ab in terms of (f_1, f_2) and the blue path a'b' in terms of (f_1, f_3). These paths intersect at a point i where $f_1 \approx 0.4111$. There, mass has increased of some 11%, the first critical force has diminished marginally, of about 0.2%, the second critical force, inversely proportional to f_3, importantly increased of more than 57%, and, at point i, the two forces equate (see Table 1).

Remark 4. In a way, the prioritized approach has offered an alternative to the non-

Sandwich Panel Optimization (1)

ε	f_1	f_2	f_3
0.	0.3702	0.3297	0.5194
0.6298	0.4111	0.3304	0.3304
	(+11.1%)	(+0.2%)	(-36.4%)

Table 1. Mass versus flexural forces - values of the objective functions at start and at the interpolated intersection point

differentiable optimization of mass versus $\max_{\mathbf{x}} \min(F_{c,1}, F_{c,2})$.

4.2.6. Blast mitigation.

The second numerical experiment relates to blast mitigation. The "Primary Pareto Front" between mass, $f_1 = \varphi_m$, and the following blend of inverse critical flexural forces, $f_2 = \frac{1}{2}(\varphi_1 + \varphi_2)$, under constraint (59), has been determined first, and is represented by magenta symbols on Figure 3. Then, the corresponding designs have been evaluated w.r.t. the following two secondary cost functions:

$$f_3 = \varphi_w, \quad f_4 = \tfrac{1}{2} e^{\varphi_u}. \tag{60}$$

(In f_4, the factor $\frac{1}{2}$ is for plotting convenience only; however the exponential transform is introduced to improve convexity [9], Appendix B.)

Remark 5. Although the Pareto front has been determined densely and smoothly, the close examination of the Pareto set reveals that designs of different structures produce identical cost functions. As a result, the secondary cost functions "along the Pareto front" are multi-valued and exhibit an evident dispersion on the figure.

The second phase of optimization is launched from a design at which $f_1 \approx 0.4260$, whose performance is indicated on Figure 3 by a, a' and a" in terms of f_2, f_3 and f_4 respectively. The continuum of Nash equilibria is represented by a yellow path ab in terms of (f_1, f_2), hardly distinguished from the Primary Pareto Front, a blue path a'b' in terms of (f_1, f_3), and a green path a"b" in terms of (f_1, f_4). At the endpoint, mass has increased of some 32%: the cost to pay. In return, the other three objective functions have been notably reduced: the average inverse critical flexural force of about 31% (point b), the deflection w of more than 84% (point b'), perhaps the most significant improvement, and f_4, based on the core blast energy absorption less, but significantly (7%) (point b"), corresponding to nearly 15% increase in energy absorption (see Table 2). Beyond this point, the variation of the cost functions in terms of the continuation parameter ε, indicated on Figure 4, is no longer monotonic and eventually becomes unstable near the limit of convexity $\varepsilon_{\max}$ of the numerical formulation. These functions normalized to their respective initial values are represented on Figure 5. This plot demonstrates the funnel-shaped pattern, a characteristic feature of the present approach:

- an increasing primary steering function with initial null slope, confirming that the Pareto-stationarity is degraded by a term $O(\varepsilon^2)$,
- a decreasing secondary steering function with finite negative initial slope, $-\sigma_B$, and the identical behavior of all secondary cost functions.

The variation with ε of the optimization variables is given by Figure 6. The Nash game affects all five variables. The core thickness increases and the ratio $\mathbf{R}$ here augments somewhat.

Figure 7 indicates the variation with ε of the constraint function $c_1(\mathbf{x})$. Apart from the very end of the continuum, when the instability is triggered, the constraint violation $\|c_1(\mathbf{x})\|$ remains less than 10^{-6}, two orders of magnitude inferior to the accuracy tolerance employed in this computation.

In conclusion, the prioritized optimization process was able to diminish both blast-mitigation-related objective functions. The reduction in deflection is impressive, the increase in energy absorption by the core more moderate, but significant. The numerical procedure demonstrates the general feature of the prioritized approach, in particular w.r.t. the funnel-shaped pattern of the cost functions.

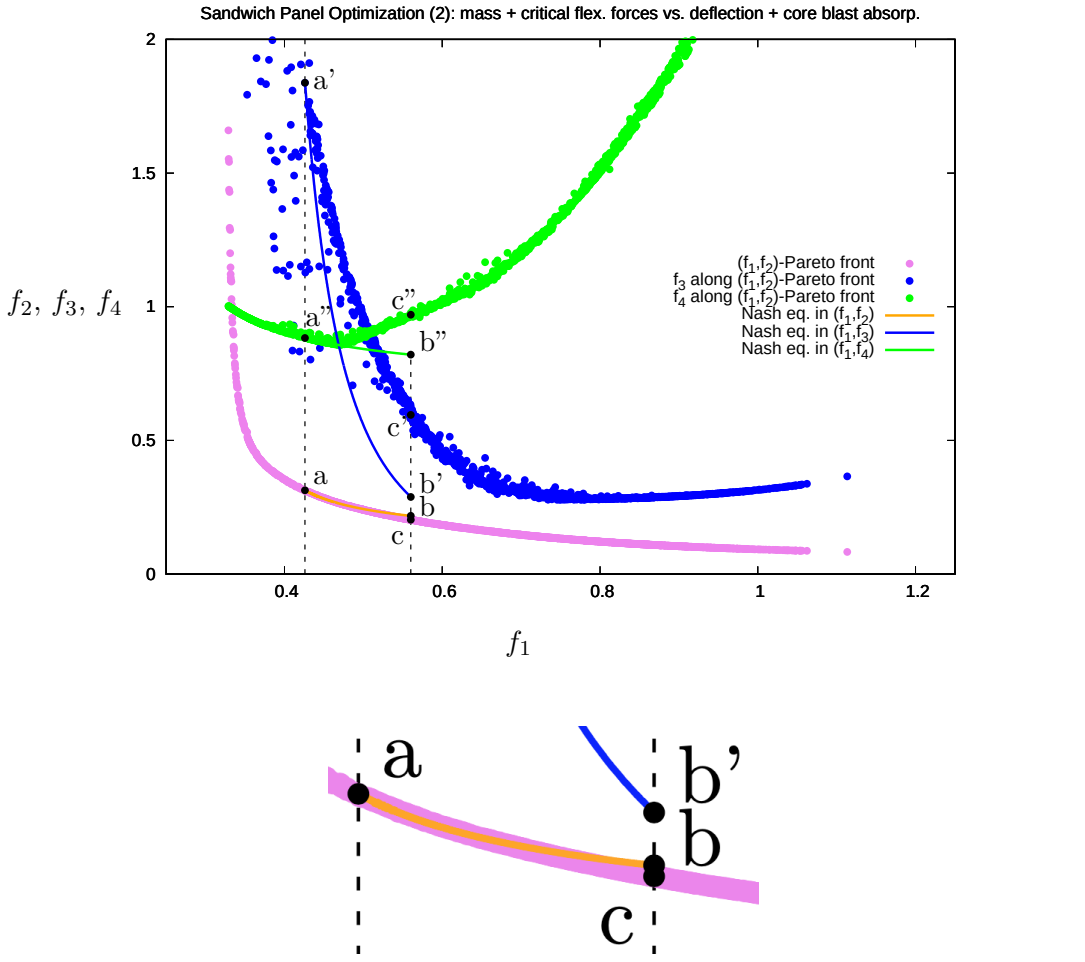


Figure 3. Sandwich Panel Optimization (2): Primary Pareto Front (byzantine) and continuum of Nash equilibria (yellow) involving four objective functions ($f_1 = \varphi_m,\ f_2 = \frac{1}{2}(\varphi_1 + \varphi_2),\ f_3 = \varphi_w,\ f_4 = \frac{1}{2}e^{\varphi_u}$ (top), and zoom of arc ab of continuum of Nash equilibria (below)

Sandwich Panel Optimization (2)

ε	f_1	f_2	f_3	f_4	φ_u^{-1}
0.	0.4260	0.3137	1.8380	0.8832	1.7576
0.9133	0.5602	0.2180	0.2885	0.8208	2.0175
	(+31.5%)	(-30.5%)	(-84.3%)	(-7.1%)	(+14.8%)

Table 2. Blast mitigation versus Pareto-optimality between mass and the average of inverse critical flexural forces - values of the objective functions at start and near the end of the continuum ($\varphi_u^{-1} = 1/\ln(2f_4) \propto U_{\text{lost}}$)

Sandwich Panel Optimization (2)
Different functions along the continuum in terms of the continuation parameter ε
($f_1 = \varphi_m,\ f_2 = \frac{1}{2}(\varphi_1 + \varphi_2),\ f_3 = \varphi_w,\ f_4 = \frac{1}{2}e^{\varphi_u}$)

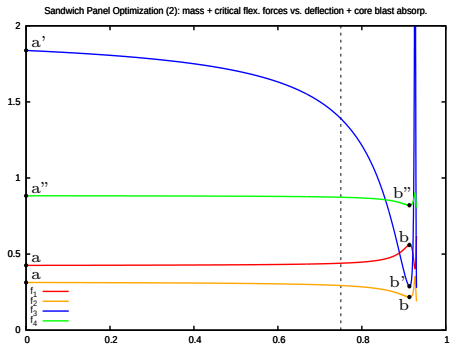


Figure 4. Objective functions $\{f_j\}$

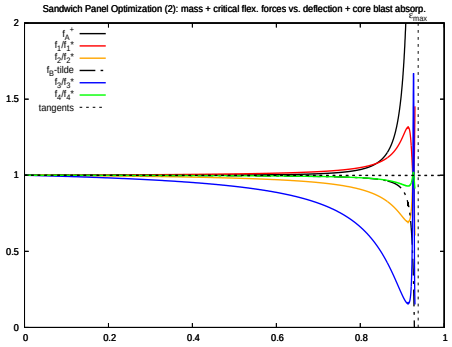


Figure 5. Funnel-shaped pattern of objective functions as ratios to their initial values $\{f_j/f_j(\mathbf{x}_A^\star)\}$

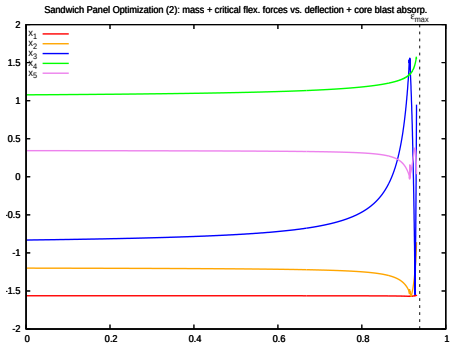


Figure 6. Optimization variables $\{x_i\}$

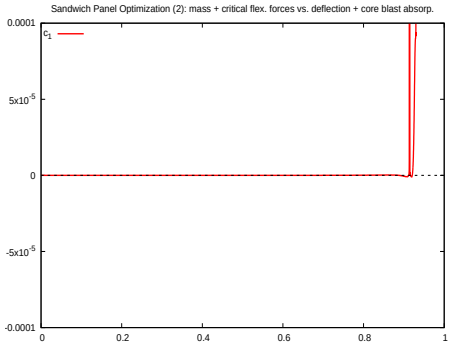


Figure 7. Constraint function c_1

5. Conclusion and perspective

An adaptation technique based on Nash games has been proposed for multi-objective optimization. The technique operates after a first phase of multi-objective optimization has produced one, or several design points that are Pareto-optimal w.r.t. preponderant or primary cost functions under constraints. This prioritized approach permits to reduce additional or secondary cost functions in a second phase of optimization. A continuum of Nash equilibria, parameterized by ε, is constructed and its image in function space is tangent to the "Primary Pareto Front" which demonstrates that the

Pareto-stationarity condition of the primary cost functions is degraded by $O(\varepsilon^2)$ only, while the secondary cost functions are reduced at least linearly with ε, with an initial negative directional derivative bounded from above by $-\sigma_B$. Thus the entire set of cost functions normalized to their values at start ($\varepsilon = 0$) exhibit a funnel-shaped pattern.

The theoretical development has permitted to relate the positive constant σ_B to gradients of the secondary cost functions w.r.t. a vector variable, $\mathbf{w}$, defined from an orthogonal decomposition of the domain referred to as "territory splitting", based on a reduced Hessian associated to the primary cost functions and the constraints. The definition of the variable $\mathbf{w}$ reflects in part the negotiation between the two virtual players defining the Nash game. The compatibility of the second phase of optimization to the result of the first phase has been justified, and the existence of the Nash equilibrium for small enough ε established.

To illustrate these theoretical results, a test-case in structural mechanics was solved numerically: a sandwich panel sizing problem meant to optimize the performance in terms of strength under bending loads, and blast mitigation, and involving four physical variables, one slack variable, one constraint, and up to four cost functions. The approach has permitted to identify Nash-equilibrium design points which, at the cost of a mass increase, were found, in terms of strength and blast mitigation, superior to the initial design point, while the trajectory of the continuum remains close to the Primary Pareto Front. The new design points were also found superior in efficiency to neighboring design points strictly on this Pareto front. This surprising result was alleged to the multi-valued nature of the secondary cost functions along the front.

Note that the same technique has also been successfully applied in [6] to a Supersonic Business Jet (SSBJ) sizing problem meant to optimize the aircraft flight performance in terms of take-off mass and distance, range, and approach speed involving 15 physical variables, one slack variable, and one constraint, yielding a similar behavior of the numerical method.

In the future, attempt will be made to apply the approach to a PDE-constrained optimization context in which the evaluation of the objective functions is necessarily more computationally demanding. In such a case, our plan is to develop additional local surrogate models based on statistical techniques and valid in the vicinity of the point chosen to initiate the continuum.

Acknowledgements

The author wishes to express his warmest thanks to his colleagues from Inria and the University Côte d'Azur R. Duvigneau, A. Habbal, L. Hascoët, L. Monasse and B. Mourrain for very fruitful scientific discussions on Nash games and polytope exploration in large dimension, as well as software implementation of algorithms.

Special thanks are also due to the Inria Service of Experimentation and Development (SED), T. Kloczko, N. Niclausse and J.Wintz particularly, for the development of the web interface of the software platform `http://mgda.inria.fr`.

References

[1] J.P. Aubin, *Mathematical methods of game and economic theory*, Courier Corporation, 2007.

[2] I. Das and J. Dennis, *Normal-boundary intersection: An alternate method for generating pareto optimal points in multicriteria optimization problems*, ICASE Report 96-62, ICASE, 1996.

[3] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, *A fast and elitist multiobjective genetic algorithm: NSGA-II*, IEEE Transactions on Evolutionary Computation 6 (2002), pp. 182–197.

[4] J.A. Désidéri, *MGDA Software Platform for Multi-Objective Differentiable Optimization*, `http://mgda.inria.fr`.

[5] J.A. Désidéri, *Révision de l'algorithme de descente à gradients multiples (MGDA) par orthogonalisation hiérarchique*, Research Report 8710, Inria, 2015. `https://hal.inria.fr/hal-01139994`.

[6] J.A. Désidéri, *Platform for prioritized multi-objective optimization by metamodel-assisted Nash games*, Research Report 9290, Inria, 2019. `https://hal.inria.fr/hal-02285197`.

[7] J.A. Désidéri and R. Duvigneau, *Direct and adpative approaches to multi-objective optimization*, Research Report 9291, Inria, 2019. `https://hal.inria.fr/hal-02285899`.

[8] J.A. Désidéri, R. Duvigneau, and A. Habbal, *Computational Intelligence in Aerospace Sciences, V. M. Becerra and M. Vassile Eds.*, Progress in Astronautics and Aeronautics, T. C. Lieuwen Ed.-in-Chief Vol. 244, chap. Multi-Objective Design Optimization Using Nash Games, American Institute for Aeronautics and Astronautics Inc., Reston, Virginia (2014).

[9] J.A. Désidéri, P. Leite, and Q. Mercier, *Prioritized multi-objective optimization of a sandwich panel*, Research Report 9362, Inria, 2020. `https://hal.inria.fr/hal-02931770`.

[10] M. Emmerich, K. Yang, A. Deutz, H. Wang, and C.M. Fonseca, *Advances in Stochastic and Deterministic Global Optimization, P. M. Pardalos and A. Zhigljavsky and J. Žilinskas eds.*, chap. A Multicriteria Generalization of Bayesian Global Optimization, Springer Optimization and Its Applications, Springer, Cham (2016).

[11] J. Fliege and B.F. Svaiter, *Steepest descent methods for multicriteria optimization*, Mathematical Methods of Operations Research (2000), pp. 479–494.

[12] M.B. Giles and N.A. Pierce, *An introduction to the adjoint approach to design*, Flow, Turbulence and Combustion (2000), pp. 393–415.

[13] D. Greiner, J. Périaux, J.M. Emperador, B. Galván, and G. Winter, *Game Theory Based Evolutionary Algorithms: A Review with Nash Applicationss in Structural Engineering Optimization Problems*, Arch Comput Methods Eng (Springer) (2017), pp. 703–750.

[14] M. Hartikainen and K. Miettinen, *Constructing a pareto front approximation for decision making*, Mathematical Methods of Operations Research (2011), pp. 209–234.

[15] M. Hartikainen, K. Miettinen, and M.M. Wiecek, *PAINT: Pareto front interpolation for nonlinear multiobjective optimization*, Computational Optimization and Applications (2012), pp. 845–867.

[16] L. Hascoët, *Tapenade, version 3*, `https://team.inria.fr/ecuador/fr/tapenade/`.

[17] J. Horn, N. Nafpliotis, and D.E. Goldberg, *A niched Pareto genetic algorithm for multiobjective optimization*, in *Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence*. 1991.

[18] A. Jameson, K. Leoviriyakit, and S. Shankaran, *Multi-point Aero-Structural Optimization of Wings Including Planform Variations*, in *45th Aerospace Sciences Meeting and Exhibit, Reno, Nevada*, January 811. 2007.

[19] A. Jameson, J.C. Vassberg, and S. Shankaran, *Aerodynamic-structural design studies of low-sweep transonic wings*, Journal of Aircraft 47 (2010).

[20] D.R. Jones, M. Schonlau, and W.J. Welch, *Efficient global optimization of expensive black-box functions*, Journal of Global Optimization 13 (1998), pp. 455–492.

[21] P. Leite, *Conception architecturale appliquée aux matériaux sandwichs pour propriétés multifonctionnelles*, Ph.D. diss., Université de Grenoble, École Doctorale I-MEP2, 2013. `http://www.theses.fr/2013GRENI053`.

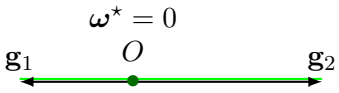
[22] J.R.R.A. Martins and A.B. Lambe, *Multidisciplinary design optimization: A survey of architectures*, AIAA Journal 51 (2013).

[23] A. Messac, A. Isamail-Yahaya, and C. Mattson, *The normalized normal constraint method for generating the pareto frontier*, Structural and Multidisciplinary Optimization (2003), pp. 86–98.

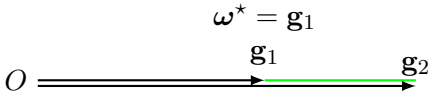
[24] T. Murata and H. Hishibuchi, *MOGA: Multi-Objective Genetic Algorithms*, in *Proc. of 1995 IEEE International Conference on Evolutionary Computation, Perth, Australia.* 1995, pp. 289–294.
[25] K. Parsopoulos and M. Vrahatis, *Recent approaches to global optimization problems through particle swarm optimization*, Natural Computing (2002), pp. 235–306.
[26] J. Périaux, F. Gonzalez, and D.S.C. Lee, *Evolutionary Optimization and Game Strategies for Advanced Multi-Disciplinary Design - Applications to Aeronautics and UAV Design*, Intelligent Systems, Control and Automation: Science and Engineering Vol. 75, Springer, 2015.
[27] J. Sobieszczanski-Sobieski, T.D. Altus, M. Phillips, and R. Sandusky, *Bilevel integrated system synthesis for concurrent and distributed processing*, AIAA Journal 41 (2003).
[28] E. Zitzler and L. Thiele, *Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach*, IEEE Transactions on Evolutionary Computation 3 (1999), pp. 257–271.

Appendix A. Configuration of convex hulls

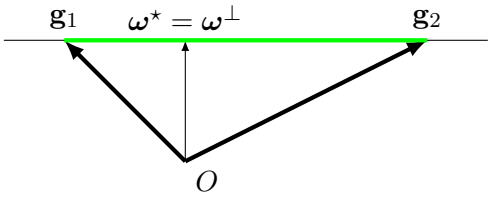
(a) Two vectors of opposite direction: Pareto stationarity
$J^\star = \{1, 2\}$



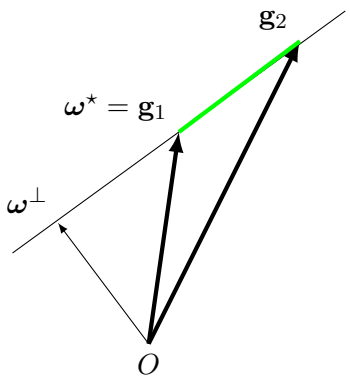
(b) Two vectors of same direction: $\boldsymbol{\omega}^\star$ is the vector of smaller norm
$J^\star = \{1\}$



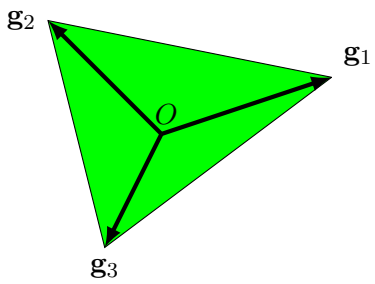
(c) Two vectors forming an obtuse angle: orthogonal projection
$\boldsymbol{\omega}^\star = \boldsymbol{\omega}^\perp,\ J^\star = \{1, 2\}$



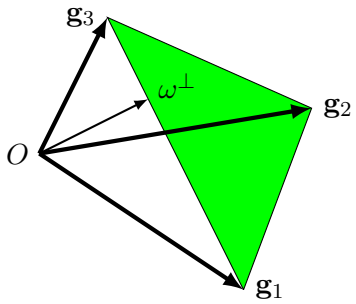
(d) Two vectors of very different norm and forming an acute angle:
$\boldsymbol{\omega}^\star = \mathbf{g}_1$ (of smaller norm),
$J^\star = \{1\}$



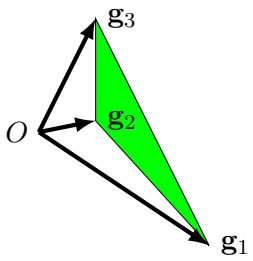
(e) Three coplanar vectors in Pareto-stationarity configuration: $\boldsymbol{\omega}^\star = 0,\ J^\star = \{1,2,3\}$



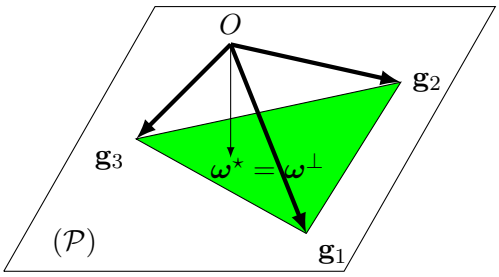
(f) Three coplanar vectors – orthogonal projection on edge: $\boldsymbol{\omega}^\star = \boldsymbol{\omega}^\perp,\ J^\star = \{1,3\}$



(g) Three coplanar vectors – non-orthogonal projection on vertex: $\boldsymbol{\omega}^\star = \mathbf{g}_2,\ J^\star = \{2\}$



(h) Three l.i. vectors – orthogonal projection onto convex hull: $\boldsymbol{\omega}^\star = \boldsymbol{\omega}^\perp,\ J^\star = \{1,2,3\}$



(i) Three l.i.vectors – orthogonal projection onto boundary edge: $J^\star = \{2,3\}$

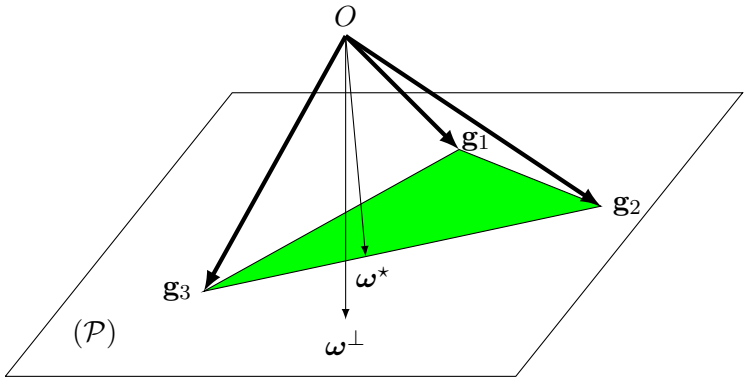


Figure A1. Various geometrical configurations of the convex hull for two and three gradients vectors – In this affine-space representation, the vectors of the convex hull are given the origin O and are pointing over the green line or triangle.