SINGULAR CONTINUATION: GENERATING PIECEWISE LINEAR APPROXIMATIONS TO PARETO SETS VIA GLOBAL ANALYSIS [*]

ALBERTO LOVISON[†]

Abstract. We propose a strategy for approximating Pareto optimal sets based on the global analysis framework proposed by Smale [*Global analysis and economics.* I. *Pareto optimum and a generalization of Morse theory*, in Dynamical Systems, Academic Press, New York, 1973, pp. 531–544]. The method highlights and exploits the underlying manifold structure of the Pareto sets, approximating Pareto optima by means of simplicial complexes. The method distinguishes the hierarchy between singular set, Pareto critical set, and stable Pareto critical set, and it can handle the problem of superposition of local Pareto fronts, occurring in the general nonconvex case. Furthermore, a quadratic convergence result in a suitable setwise sense is proven and tested in a number of numerical examples.

Key words. Multiobjective optimization, multicriteria optimization, Pareto critical set, Delaunay tessellations in general dimension, stability of mappings

AMS subject classifications. 90C29, 58K25

DOI. 10.1137/100784746

1. Introduction.

1.1. Multiobjective optimization and Pareto optimality. Multiobjective optimization is concerned with the problem of optimizing several functions (or objectives) simultaneously. A precise mathematical statement in an economics framework was first given by Pareto [38], [39] in the 1880s. In recent years a strong interest has grown, as a variety of problems in structural mechanics, automotive industry, aerospace, production planning, environmental policy, and many others involve more than one objective function, and different numerical strategies have been developed subsequently [34].

In the single objective case, an optimum is defined as a point $x \in W \subseteq \mathbb{R}^n$, where a given function $u \colon W \to \mathbb{R}$ assumes its maximum, if the maximum exists. In multiobjective optimization we consider two or more functions, $u_1, \ldots, u_m \colon W \to \mathbb{R}$, and in all the nontrivial cases the optima for one function are distinct from the optima of the remaining ones. A key point is that one not only has to consider the optima of the individual functions, but there usually arises an infinite number of so-called nondominated points.[1] They are defined precisely as follows.

DEFINITION 1 (Pareto optimality). *Let W be an open subset of $\mathbb{R}^n$, or an n-dimensional manifold, and let $u_1, \ldots, u_m \colon W \to \mathbb{R}$ be smooth functions.[2] A point $\bar{x} \in W$ is called a nondominated point, or a Pareto optimum, if there is no $x \in W$ such that $u_i(x) \geq u_i(\bar{x})$ for all $i = 1, \ldots, m$ and $u_j(x) > u_j(\bar{x})$ for some j. If there exists a neighborhood $V \subseteq W$ of $\bar{x}$ where $\bar{x}$ is Pareto optimum, then $\bar{x}$ is called a local Pareto optimum.*

1.2. The necessity for global representations of the Pareto sets. As pointed out, for instance, in [9], the set of Pareto optima is in many cases a large and complicated nonconvex set and most of the existing algorithms, being inspired by local search ideas

---

†Dipartimento di Matematica Pura ed Applicata, Via Trieste, 63 35121–Padova, Italy (lovison@math.unipd.it).

[1]On the other hand, because of Morse's lemma, the optima of the individual functions are usually isolated.

[2]We may equivalently refer to a unique smooth vector function or mapping $u \colon W \to \mathbb{R}^m$.
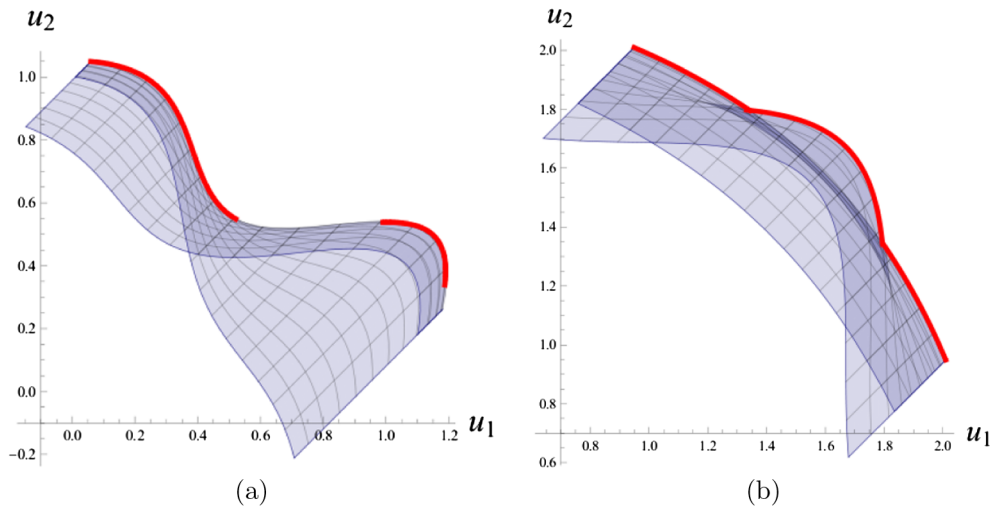
FIG. 1.1. *Possible problems arising in the objectives space when u_i are nonconvex functions.* (a) *The Pareto set is composed of separate branches.* (b) *A connected global Pareto set is composed of separate local branches crossing each other.*

from traditional linear and nonlinear programming, fail at giving a truly global representation of the set of Pareto optima. Also, [12] stresses that "a whole collection of Pareto optimal points, representative of the entire spectrum of efficient solutions" would be helpful in facilitating design in engineering applications.

Recent multiobjective optimization literature tackled this issue focusing on defining algorithms producing even distributions of Pareto points [12], [31], [32], [64], while an alternative philosophy [43], [44], [45], [51] dealt with producing local meshes approximating Pareto sets, relying on continuation (homotopy) strategies. In the recent paper [40], both topics are addressed. Alternative techniques aiming to approximate the entire optimal set are described in the recent papers [19], [29], in the survey [46], and in the references therein.

We want here to highlight a key feature of the Pareto set which makes it, in general nonconvex cases, a complicated set. Its complexity is even amplified when the Pareto set is viewed in the output space.

Indeed, the set of all global Pareto optima can be disconnected, i.e., composed of separate portions of seemingly smooth surfaces (see Figure 1.1(a)). Furthermore, even when the image of the set of global Pareto optima is a connected set, it could be composed by cutting and sewing together different locally optimal branches (see Figure 1.1 (b)), coming from separate zones of the domain. We will illustrate in what follows that this kind of behavior is not an artifact obtained with unrealistic functions but in a sense represents a typical situation that is not destroyed by slight deformations of the functions. Those situations are persistent, or more technically, structurally stable.[3]

We notice that the algorithms mentioned above are expected to work properly only in a local sense, although they are intended to capture some of the global features of the optimal set. Moreover, apart from the homotopy techniques, they are *pointwise* strategies, in the sense that as an approximation of the Pareto optimal set they produce a scatter of points; the evenness of the distribution of points is then estimated on the image

---

[3]See section 1.3 below for a formal definition of structural stability.

space. In some applications those points are joined together in a compound structure, e.g., a Delaunay triangulation, but only a posteriori and in the output space. It should be noted that because of the effects of mappings described above, defining such a structure from the output space, i.e., joining nearby values of optimal points, is subject to failure: the corresponding preimage points, indeed, are not necessarily nearby, or even connected in the way suggested by the positions of their images. Conversely, the images of nearby optimal points are nearby, because of continuity.

We propose instead that a faithful global representation of the Pareto set in general nonconvex settings is obtained according to the following three steps. First, by shifting the focus from the output space to the input space, second by approximating the full set of the local Pareto optima, and third by adopting a *setwise* standpoint, namely, using compound geometrical objects as simplicial complexes instead of scatters of points. The first step unfolds the singularities (branches crossing, cusp points, and so on) occurring as an effect of the mapping. Indeed, as illustrated in what follows, the preimage of the Pareto set is nonsingular, as it exhibits in general a regular manifold structure. The second step, because of possible superpositions of local branches, guarantees that every portion of the global Pareto optimal set is represented. Simplicial complexes, i.e., *meshes*, faithfully reflect the manifold structures and explicitly offer the desired parametrization for each portion of the Pareto set, allowing one to perform "tradeoff studies" among the conflicting objectives. Indeed, tradeoff studies may be the application of greatest practical importance of multiobjective optimization. Nevertheless, from the above discussion it is clear that trying to track the surface of the Pareto set by picking points from the output space, as pointwise strategies are aimed to do, is supposed to work correctly only throughout limited branches.

There are at least two reasons why our program has not yet been pursued in its entirety. First of all, in a number of situations, the numerical techniques available in literature are able to build sufficiently faithful representations of the Pareto set. For example, when the functions at hand are convex, or relatively simple, or when the singularities are situated far away from interesting zones, a global investigation of the problem is not required. Moreover, typically, tradeoff studies are performed in the neighborhood of a previously determined solution; therefore, they can be limited to a nonproblematic branch of the Pareto set giving back as well the important information. Second, it is clear that a global exploration of the domain is a demanding task which could be far out of the scope of a typical design problem.

Nevertheless, faithful global representations of the Pareto set are a worthy goal to pursue, because they complement existing local exploitation strategies in two senses: they resolve the above-mentioned problematic superpositions, and they facilitate the location of important zones, which could merit further investigation. It is clear that this kind of program has to be implemented in an efficient way in order to be useful in applications. On the other hand, even a roughly sketched global picture of the whole situation can give crucial information on the problem at hand, suggesting correctly the location of paramount zones.

1.3. Global analysis and multiobjective nonlinear programming. With this in mind, we have devised a novel numerical strategy for approximating Pareto sets, theoretically based on the global analysis [4] framework established by Smale and others in the early 1970s [13], [14], [16], [17], [57], [58], [65], [66], [67] and in more recent work

---

[4]See [56]. For brevity, we speak a bit loosely of global analysis also when referring to concepts of singularity theory or differential topology.

[35], [36]. Motivated by his discussions with Debreu,[5] Smale investigated the problem of optimizing several functions within the dynamical systems arena. In the series of works that followed there emerged interesting topological and geometrical features of the sets of the Pareto optima. The notion of *Pareto critical set θ*, generalizing the concept of critical point for scalar functions, was introduced and furthermore, local Pareto optima were characterized by means of first and second derivatives. Quoting Smale [57], "We study the local and global nature of θ, as one uses freshman calculus to study the maximum of a single function."

One of the basic facts highlighted in Smale's global analysis framework is that under the assumptions of second order differentiability and some generic transversality condition, Pareto optimal sets are portions of $(m-1)$-dimensional manifolds. It is fundamental that a slight deformation of the functions do not alter substantially the Pareto set. Global analysis is the proper setting to study resilience properties. A mapping $u\colon W \to \mathbb{R}^m$ is said to be structurally stable[6] if there exists a neighborhood $N(u)$ in the C^r topology such that every function $\tilde{u}$ in $N(u)$ is equivalent to u; i.e., there exist diffeomorphisms h, k, close to the identities of the respective spaces, such that the diagram

$$
\begin{array}{ccc}
W & \xrightarrow{\ u\ } & \mathbb{R}^m \\
h \downarrow & & \downarrow k \\
W & \xrightarrow{\ \tilde{u}\ } & \mathbb{R}^m
\end{array}
$$

commutes. Clearly, if two mappings are equivalent, their Pareto sets are diffeomorphic. One of the main results of global analysis is that there exists an open and dense set in $C^r(W, \mathbb{R}^m)$ of structurally stable mappings.[7] In other words, the Pareto set of *almost* every mapping u is as close as desired to the Pareto set of any other mapping in a sufficiently small neighborhood of u. This is clearly of fundamental importance for the applications: when functions are known only with a certain approximation, as is usual in engineering design problems, the set of optimal points is guaranteed to be approximated correctly by any convergent sequence of functions [7], [56], [62].[8] Moreover, a generalization of Morse theory for several functions can be defined [57], [65].

The strategy presently proposed highlights and exploits the manifold structure underlying the Pareto sets and precisely reproduces the hierarchy described in Smale's work among the singular set, the Pareto critical set, and the stable Pareto critical set.[9] These sets are approximated by means of simplicial complexes, and by exploiting Newton-type estimates it is possible to prove quadratic convergence in a setwise sense, adopting the Hausdorff measure. Because of this result the present method can be considered a setwise variant of multiobjective Newton methods, as in [18].

---

[5]Debreu won the Nobel Prize for Economics in 1983 "for having incorporated new analytical methods into economic theory and for his rigorous reformulation of the theory of general equilibrium." For an account of the cooperation between Smale and Debreu, see [16], [17].

[6]To be precise, we should speak of *stability of mappings*, while *structural* stability is more often used when speaking about differential equations. On the other hand, we must speak about stability of Pareto optima, which is instead a concept deriving from the study of stability of equilibra and refers to critical points which are maxima. Therefore, we will keep speaking of structural stability when dealing with typical singularities of mappings.

[7]It is necessary that $m < 7$ and $n \neq 8$, or $m < 6$ and $n = 8$ [30].

[8]The original idea of structural stability is a joint work from an engineer, A. Andronov, and a mathematician, L. Pontryagin; see [55], [56].

[9]See below section 2 for definitions.

The algorithms of this paper can also be considered as a globalization and generalization to more than two objectives of the homotopy techniques, while the use of tessellations can be thought of as the specialization of the techniques of simplicial pivoting [3], [4], [5] to the problem of optimizing several functions. A strong similarity can be found in the method proposed by [51], where the authors detect and progressively refine the hypercubes containing the Pareto sets, relying on the standard Karush–Kuhn–Tucker conditions instead of Pareto criticality.

2. The global analysis framework. We recall now Smale's definitions and results. Let $W \subseteq \mathbb{R}^d$ be an open set or more generally a smooth n-dimensional manifold, $u \colon W \to \mathbb{R}^m$ a smooth vector function with $m \leq n$.[10] The singular set $\Sigma \subseteq W$ is the collection of singular points, i.e., the points where the rank of the Jacobian $Du(x)$ is nonmaximal. If $m = 1$, the singular set coincides with the set of critical, or stationary, points, i.e., $Du(x) = 0$. It can be proved that under generic conditions the singular set is a smooth manifold.

Let Pos be the open positive cone in $\mathbb{R}^m$, $\mathrm{Pos} := \{y \in \mathbb{R}^m | y_j > 0 \, \forall j = 1, \ldots, m\}$, and let C_x be the corresponding open cone in the tangent space $T_x W$, $C_x := Du^{-1}(\mathrm{Pos})$.

DEFINITION 2 (Pareto critical set θ). *The set*

$$(2.1) \qquad \qquad \theta := \{x \in W | C_x = \varnothing\}$$

is called the Pareto critical set.

We characterize θ in terms of the Jacobian of u.

PROPOSITION 3 (first order proposition). *Let $x \in W$. Then, $x \in \theta$ if and only if*

(a) *$\{Du_j(x)\}_{j=1,\ldots,m}$ do not belong to a unique open half-space of the cotangent space $T_x^\star W$.*

(b) *$\exists \lambda_j \geq 0$, $j = 1, \ldots, m$, not all zero such that $\sum_j \lambda_j Du_j(x) = 0$.*

Remark 4. The meaning of Proposition 3 is that for x to be critical, first order results imply that there cannot exist paths along which all the objectives u_j can be incremented at the same time. If there were an open half-space containing all Du_j, as in condition (a), any direction in this half-space would be a direction of improvement for every u_j. Equivalently, condition (b) states that the gradients Du_j should be linear dependent and furthermore should "oppose" each other. In other words, moving in the direction of steepest ascent according to one of the u_j causes one or more of the remaining u_i to strictly decrease.

Remark 5. In the bi-objective case, $m = 2$, Proposition 3 states that in Pareto critical points the two gradients are collinear and in opposition to each other. Also critical points for one of the two objectives are Pareto critical.

In analogy with freshman calculus, (Pareto) criticality is a necessary condition for x to be optimal. In order to discriminate the nature of the Pareto critical points we introduce a notion of stability and point out its relation with the second derivatives of u. This will give *sufficient* conditions for x to be Pareto optimal.

DEFINITION 6. *A curve $(a, b) \ni t \mapsto \varphi(t) \in W$ is said to be admissible if*

$$(2.2) \qquad \qquad \frac{d}{dt} u_i(\varphi(t)) > 0, \quad t \in (a, b) \quad \forall i = 1, \ldots, m.$$

Clearly, if a point is Pareto critical, there could not exist admissible curves passing through it. In order to establish its optimality it is necessary to investigate the behavior

---

[10]The case $m > n$ is less frequent. We will consider some aspect of this case in what follows.

of the admissible curves in a neighborhood of a critical point. Admissible curves are smooth paths along which every objective is incremented. Therefore, they move towards local Pareto optima; conversely, if a critical point captures all neighboring admissible curves, that point is a local Pareto optimum.

DEFINITION 7. *A Pareto critical point x is said to be stable, $x \in \theta_S$, if, given a neighborhood V_x of x in W, there exists a neighborhood U_x of x in V_x such that every admissible curve $\varphi \colon [a, b) \to W$ with $\varphi(a) \in U_x$ satisfies Image $(\varphi) \subset V_x$.*

Pareto stability can be fully decidable by carefully examining the second derivatives of the objectives. In the single objective case, by virtue of Morse's lemma, it is possible to find a coordinate system where the objective u can be written as a quadratic polynomial $u = \pm x_1^2 \pm \cdots \pm x_d^2$, which number of minus signs defines the Morse index, and therefore decides the nature of the critical point (maximum, minimum, or saddle) [37]. With some effort, results can be extended to multiple objectives: second derivatives are not defined invariantly, but if we think about them as a symmetric bilinear form restricted to the kernel of the differential $Du(x)$ assuming values on the cokernel $\mathbb{R}^m / \text{Image}(Du(x))$, then this form is invariantly defined. It is called "second intrinsic derivative" (see [30], [42]). The restriction to the kernel of the tangent map $Du(x)$ has also the following meaning. By investigating the attractive/repulsive behavior of admissible curves in a neighborhood of a critical point, we will not be interested in what happens along the directions parallel to the critical set, while the orthogonal space will be the arena where the stability of the critical points will be decided. The case of greatest importance is where corank $Du(x)$ is 1 (i.e., rank $Du(x)$ is $m - 1$). In this case the second intrinsic derivative assumes values in a 1-dimensional vector space. If we consider $x \in \theta$, we have Image $(Du(x)) \cap \text{Pos} = \varnothing$; thus $\mathbb{R}^m / \text{Image}(Du(x))$ has a canonical positive ray. We call the second intrinsic derivative, in this case, the generalized Hessian H_x. It makes sense to say that H_x is negative definite or positive definite, as well as to define an index, as the index of the symmetric form H_x. We set

$$(2.3) \qquad \partial \theta = \{x \in \theta | \text{Image}(Du(x)) \cap \{\text{Cl}(\text{Pos}) \setminus \{0\}\} \neq \varnothing\},$$

where Cl(Pos) is the closure of Pos.

PROPOSITION 8 (second order proposition). *Let $u \colon W \to \mathbb{R}^m$ be a smooth map with $x \in \theta$, $x \notin \partial \theta$, and corank $Du(x) = 1$. Then,*

(a) *if the generalized Hessian H_x is negative definite, then $x \in \theta_S$;*

(b) *let $\lambda_j \geq 0$, $j = 1, \ldots, m$ be as in the first order proposition; then (up to a positive scalar)*

$$(2.4) \qquad H_x = \sum_{j=1}^{m} \lambda_j D^2 u_j(x) \qquad on \; \ker \, Du(x).$$

The proposition is proved in [58], while a discussion of the genericity of the hypotheses on the rank assumption (rank $Du(x) \geq m - 1$ a.e.) is given in [11], [27], [57].

Most importantly, Proposition 8 offers a useful and workable criterion for deciding the stability of critical points. We will translate numerically this proposition in Algorithm 2.

2.1. The structure of Pareto sets. We start by recalling the notion of Thom's stratification (see [60], [61], [62], [63]).

DEFINITION 9. *Let $A \subset W$ be a closed subset. A stratification $\mathcal{S}$ of A is a finite collection of connected submanifolds of W satisfying the following properties:*

(1) $\cup_{S \in \mathcal{S}} S = A$.

(2) If $S \in \mathcal{S}$, then $\partial S = \mathrm{Cl}(S) \setminus S$ is a union of elements of $\mathcal{S}$ of lower dimension.

(3) If $S \in \mathcal{S}$ and U is a submanifold of W transversal to S at $x \in S$, then U is transversal to all elements of $\mathcal{S}$ in a neighborhood of x.

The following theorem has been proved in [13]. Consider the space $C^\infty(W, \mathbb{R}^m)$ endowed with the C^∞ topology. W is a compact manifold with dimension $n \geq m$.

THEOREM 10 (θ is a stratified set of dimension $m-1$). *There is an open and dense set $\mathcal{G} \subset C^\infty(W, \mathbb{R}^m)$ such that if $u \in \mathcal{G}$, then θ is a stratified set of dimension $m-1$.*

Remark 11. If $m > n$, it is possible to prove that, for a generic mapping u, θ is a stratified set of dimension n.

From the point of view of the numerical applications, we state that in the generic case the strata of the Pareto critical set θ can be discretized by means of a collection of $(m-1)$-dimensional meshes. Obviously, we would like to refine this procedure to θ_S. Unfortunately, the following conjecture has been proved only for $m = 2, 3$ (see [14], [65]).

CONJECTURE 12. *There is an open and dense set $\mathcal{G} \subset C^\infty(W, \mathbb{R}^m)$ such that if $u \in \mathcal{G}$, then θ is a stratified set and θ_S is a union of strata.*

Remark 13. The stable Pareto critical set θ_S is formed by all the local Pareto optimal points. The global Pareto optimal points cannot be distinguished from local optima by means of differential features as in the statements presented above. Global Pareto optima can only be filtered out a posteriori.

3. Numerical translation of the global analysis approach. In the following sections we illustrate numerical methods for approximating Pareto sets on the basis of Propositions 3 and 8. The procedure is reminiscent of contour plot algorithms for plotting level sets of functions, and it is a special instance of general strategies for piecewise-linear approximation algorithms for implicitly defined manifolds [1], [3], [5], [6]. The method determines a simplicial complex approximating the singular set Σ and then refines it to the critical set θ and to the stable critical set θ_s. Because the strategy proposed consists of a continuation method focused on the manifold structure of Pareto optima inherited by the singular set, we coined the term singular continuation.

3.1. First order search algorithm. Algorithm 1 translates numerically Proposition 3. We start by considering a set of data points $\mathcal{D} = \{P_1, \ldots, P_N\}$ where we will evaluate the Jacobian Du, and then we build a Delaunay tessellation having $\mathcal{D}$ as nodes.[11] We assume that the nodes $P_1, \ldots, P_n$ are in general position; i.e., they give rise to a valid Delaunay tessellation. Better results are obtained if the simplexes are "round"; i.e., they do not possess very thin or very large angles. Special tessellations, e.g., Freuenthal–Kuhn, simplify the operation of "pivoting" from a simplex to the adjacent, speeding up the process of glueing together the polytopes composing the implicitly defined manifold [5]. Hereafter, we also assume that the dataset is sufficiently dense to resolve all the features of the singular manifold Σ. More precisely, we assume that every connected component of Σ intersects at least one of the $(n-m)$-faces Δ of the tessellation, and the intersection is unique and transversal; i.e., $\dim T_x\Sigma \bigoplus T_x\Delta = n = \max$. Doing so, Σ is guaranteed to be homeomorphic to its piecewise linear approximation. We denote by Σ, θ, and θ_S the portions of the singular set, critical set, and stable critical set, respectively, which possibly are contained in a simplex Δ of

---

[11]In the implementation considered in what follows we employed the qhull software [8], based on the computation of convex hulls, and, in the 2-dimensional examples, we employed the TRIANGLE software [53], [54]. For iterative schemes, an efficient alternative is offered by the Bowyer–Watson algorithm [10], [68], which is incremental.

the tessellation of the domain considered. Hatted symbols, $\hat{\Sigma}$, $\hat{\theta}$, and $\hat{\theta}_S$ denote the corresponding piecewise-linear approximations. The details of the algorithm are discussed in subsection 3.2.

ALGORITHM 1. FIRST ORDER ALGORITHM FOR APPROXIMATING THE PARETO CRITICAL SET $\boldsymbol{\theta}$.

    1: Consider a set of data points $\mathcal{D} = \{P_1, \ldots, P_N\}$;
    2: evaluate the gradients of the u_j on the data points;
    3: build a Delaunay tessellation on the nodes $\mathcal{D}$;
    4: **for all** Delaunay simplex $\Delta = \langle P_{i_0}, \ldots, P_{i_n} \rangle$ in the tessellation **do**
    5: compute the $(m-1)$-polytope $\hat{\Sigma}$ where the first order approximation of the Jacobian of u vanishes;
    6: extract the subpolytope $\hat{\theta}$ where the vanishing linear combination $\lambda_1 Du_1 + \cdots + \lambda_m Du_m = 0$ has nonnegative coefficients;
    7: **end for**
    8: compose a simplicial complex glueing together adjacent polytopes $\hat{\theta}$

Remark 14. The algorithm assumes $n \geq m$. When $m > n$ things extend quite easily, because the singular set is all of the input domain, and as recalled in section 2.1 the critical set is a stratified set. More precisely, the gradients are always linearly dependent; thus it is sufficient to skip step 5 of Algorithm 1.

3.2. Analysis of simplexes. We cycle through the tessellation simplexes $\Delta = \langle P_{i_1}, \ldots, P_{i_{n+1}} \rangle$ and approximate the portion of the Pareto critical set θ possibly contained in Δ. To determine the linear approximation $\hat{\theta}_s$ of the stable Pareto critical portion $\theta_s \cap \Delta$, we recall that θ is contained in the singular set Σ, i.e., the set where the rank of the differential $Du(x)$ is less than maximal:

$$(3.1) \qquad \theta_s \subseteq \theta \subseteq \Sigma \subseteq W, \qquad (\Rightarrow \hat{\theta}_s \subseteq \hat{\theta} \subseteq \hat{\Sigma} \subseteq \Delta).$$

Adjacent approximate portions $\hat{\theta}_s$ are eventually sewed together.

3.2.1. Singular set $\hat{\Sigma}$. We fix a cell $\Delta := \langle P_1, \ldots, P_{n+1} \rangle$. The Jacobian is an $n \times m$ matrix whose rank is nonmaximal on the singular set Σ. The rank of Du drops when the rows are linearly dependent, i.e., when all of the m-order minors are zero. In practice, it is sufficient to consider only a suitable selection of square $m \times m$ submatrices, in a way that each column of the Jacobian occurs at least once. We consider, for instance, the following submatrices:

$$
M_1 = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \cdots & \frac{\partial u_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_m}{\partial x_1} & \cdots & \frac{\partial u_m}{\partial x_m} \end{pmatrix}, \quad
M_2 = \begin{pmatrix} \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_{m+1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_m}{\partial x_2} & \cdots & \frac{\partial u_m}{\partial x_{m+1}} \end{pmatrix}, \ldots
$$

$$(3.2) \qquad \ldots, \quad M_{n-m+1} = \begin{pmatrix} \frac{\partial u_1}{\partial x_{n-m+1}} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_m}{\partial x_{n-m+1}} & \cdots & \frac{\partial u_m}{\partial x_n} \end{pmatrix}.$$

Remark 15. It could happen in degenerate cases that the chosen minors are all zero while some of the discarded minors are nonzero. However, when the rank assumption holds, for almost every point in Σ we have rank $Du(x) = m - 1$. If the selected minors are zero, there exists a unique vanishing linear combination (up to a multiplicative factor) for every column of the Jacobian matrix; therefore all the remaining minors are zero.

We denote the number of minors by $r := n - m + 1$ and set $\omega_j(x) := \det M_j(x)$ for $j = 1, \ldots, r$, and we consider all the $(r+1)$-faces of the cell Δ; i.e., for every $\{i_1, \ldots, i_{r+1}\} \subseteq \{1, \ldots, n+1\}$ with $i_1 < i_2 < \cdots < i_{r+1}$, we consider the simplex $\langle P_{i_1}, \ldots, P_{i_{r+1}} \rangle$. The solution $(\mu_1, \ldots, \mu_{r+1})$ of the system

(3.3)
$$
\begin{cases}
\mu_1 \omega_1(P_{i_1}) + \cdots + \mu_{r+1} \omega_1(P_{i_{r+1}}) &= 0, \\
\vdots \quad \vdots \quad \vdots & \vdots \\
\mu_1 \omega_{r+1}(P_{i_1}) + \cdots + \mu_{r+1} \omega_{r+1}(P_{i_{r+1}}) &= 0, \\
\mu_1 + \cdots + \mu_{r+1} &= 1
\end{cases}
$$

leads to a singular vertex $Q := \mu_1 P_{i_1} + \cdots + \mu_{r+1} P_{i_{r+1}}$ of $\hat{\Sigma}$ if all $\mu_j > 0$, i.e., if Q is contained in the $(r+1)$-face of Δ considered.

The (possibly empty) singular set $\hat{\Sigma}$ is an $(m-1)$-polytope defined as the convex hull of the singular vertices Q.

3.2.2. Critical set $\hat{\theta}$. In the previous subsection we detected the singular set Σ, on the basis of the fact that on the singular set the gradients are linearly dependent. On the other hand, on the critical set θ there exists a positive linear combination of the gradients giving zero. Thus we proceed by estimating the coefficients λ_j of the vanishing linear convex combination of the gradients and cutting out the critical set θ from Σ by intersection with the half-spaces where the linear interpolations of the λ's are positive.

More precisely, we solve the system

(3.4)
$$
\begin{cases}
\lambda_1 Du_1(P) + \cdots + \lambda_m Du_m(P) = 0, \\
\lambda_1 + \cdots + \lambda_m = 1
\end{cases}
$$

for $\lambda_1, \ldots, \lambda_m$. The Jacobian of u has rank $m - 1$ in almost all the points of the singular set (generic hypothesis); thus the system (3.4) has rank m, and by the implicit function theorem λ_j are smooth functions of P. As a result the level sets $\{\lambda_j(P) = 0\}$, which define the boundary of θ, are smooth manifolds. At the first order we are working with, the requests $\lambda_j(P) \geq 0$ cut out half-spaces in $\hat{\Sigma}$, defining possibly a critical subpolytope $\hat{\theta}$ in Δ.

We notice that we do not know the actual values of Du on the singular vertices, i.e., the nodes of $\hat{\Sigma}$. Nevertheless, we can estimate them by linearly interpolating the values of Du on the data nodes $P_{i_1}, \ldots, P_{i_{r+1}}$ defining the vertex Q in $\hat{\Sigma}$. By taking the coefficients $\mu_1, \ldots, \mu_{r+1}$ solving the system (3.3), we are guaranteed that the $\widehat{Du}_j(Q) := \mu_1 Du_j(P_{i_1}) + \cdots + \mu_{r+1} Du_j(P_{i_{r+1}})$ are linearly dependent, and we are justified in solving for the vanishing linear combination $\lambda_1 \widehat{Du}_1(Q) + \cdots + \lambda_m \widehat{Du}_m(Q) = 0$.

3.3. Convergence analysis for θ. Let us consider for this section a single simplex Δ. Intuitively, it is clear that the approximation $\hat{\Sigma}$ of Σ obtained by linear interpolation is quadratically good because of Taylor's theorem. We state more precisely this result in the setwise context we have adopted.[12] The distance between the sets A and B can be measured in terms of Hausdorff distance:

---

[12]General estimates on the accuracy of piecewise-linear approximations of implicitly defined manifolds are proved in [2].

$$(3.5) \qquad d_{\mathcal{H}}(A, B) := \max\left(\sup_{x \in A} \inf_{y \in B} d(x, y), \sup_{y \in B} \inf_{x \in A} d(x, y)\right).$$

THEOREM 16 (quadratic precision for Σ). *Let $P_0, \ldots, P_n$ be in general position and such that Du has maximum rank. We denote by $\Delta = \langle P_0, \ldots, P_n \rangle$ the n-simplex whose vertices are those points. Let $\omega_1(x), \ldots, \omega_r(x)$ be a selection of independent minors of Du, and let $\hat{\omega}_j(x)$ be the first order interpolation of the values of ω_j on the nodes P_i. Assume 0 is a regular value for $\omega_1, \ldots, \omega_r$, that the zero levels of the ω_j are transversal, and that $\omega_j(P_i) \neq 0$ for all i, j. Then*

$$(3.6) \qquad \Sigma = \{\omega_1(x) = 0\} \cap \cdots \cap \{\omega_r(x) = 0\},$$

$$(3.7) \qquad \hat{\Sigma} = \{\hat{\omega}_1(x) = 0\} \cap \cdots \cap \{\hat{\omega}_r(x) = 0\},$$

and there exists a constant C:

$$(3.8) \qquad d_{\mathcal{H}}(\Sigma, \hat{\Sigma}) \leq \mathsf{C}\delta^2,$$

where $\delta > 0$ is the diameter of the simplex Δ.

Proof. First of all, we notice that the $\omega_k(x)$ are polynomials of the first derivatives of u; thus they are smooth in our hypotheses. Inductively, consider $r = 1$ and denote $\omega = \omega_1$. By Taylor's theorem,

$$(3.9) \qquad \omega(x) = \hat{\omega}(x) + O(|x - P_0|^2); \qquad \text{i.e.,} \quad |\omega(x) - \hat{\omega}(x)| \leq C\delta^2$$

for a suitable $C > 0$. Assume, without loss of generality, $\omega > 0$ on $P_0, \ldots, P_k$ and $\omega < 0$ on $P_{k+1}, \ldots, P_N$. Let $\varepsilon := C\delta^2$. (See panel (a) of Figure 3.1.) Thus the zero levels of ω and $\hat{\omega}$ are comprised between the $\pm\varepsilon$ levels of $\hat{\omega}$; i.e.,

$$(3.10) \qquad \{x \in \Delta | \omega(x) = 0\} \subseteq \{x \in \Delta | -\varepsilon \leq \hat{\omega}(x) \leq \varepsilon\}.$$

By the compactness of Δ, there exist $x_0 \in \{\hat{\omega} = 0\}$, $x_\varepsilon \in \{\hat{\omega} = \varepsilon\}$ such that

$$(3.11) \qquad d_{\mathcal{H}}(\{\hat{\omega} = 0\}, \{\hat{\omega} = \varepsilon\}) = |x_0 - x_\varepsilon|,$$
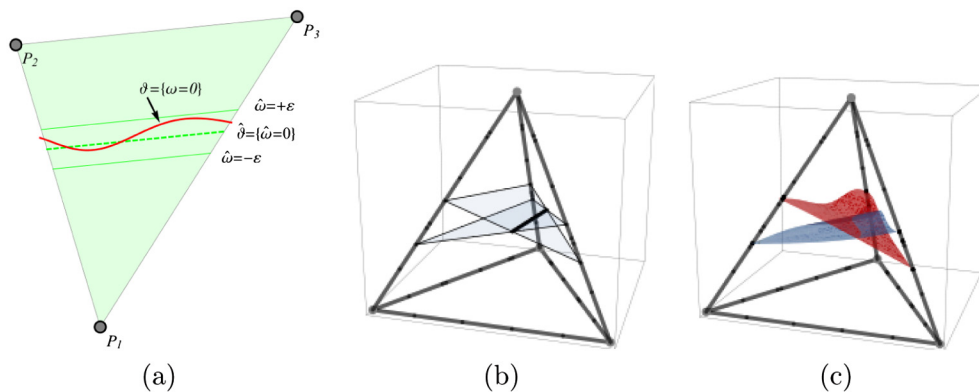
and it holds that

FIG. 3.1. *Critical simplexes with representations of the critical set θ and its first order approx $\hat{\theta}$. Panel (a): two functions in two dimensions. Panels (b) and (c): two functions in three dimensions. In panel (b) the the thick line is the first order approximation $\hat{\theta}$; in panel (c) the critical set θ is the curve of intersection of the two level surfaces $\omega_1(x) = 0$ and $\omega_2(x) = 0$.*

$$(3.12) \qquad \hat{\omega}(x_\varepsilon) - \hat{\omega}(x_0) = |\hat{\omega}'(x_0) \cdot (x_\varepsilon - x_0)| = \left| \frac{\partial \hat{\omega}}{\partial w}(x_0) \right| |x_\varepsilon - x_0|,$$

where $w = \frac{x_\varepsilon - x_0}{|x_\varepsilon - x_0|}$. By means of an elementary linear algebra argument we have also that

$$(3.13) \qquad \left| \frac{\partial \omega}{\partial w}(x_0) \right| \geq \min_{\substack{i=1,\dots,k, \\ i'=k+1,\dots,n}} \left| \frac{\hat{\omega}(P_i) - \hat{\omega}(P_{i'})}{P_i - P_{i'}} \right| =: B > 0,$$

so we can conclude

$$(3.14) \qquad |x_0 - x_\varepsilon| \leq \frac{\varepsilon}{B} = \frac{C}{B}\delta^2 = \mathsf{C}\delta^2$$

and eventually

$$(3.15) \qquad d_{\mathcal{H}}(\{\omega = 0\}, \{\hat{\omega} = 0\}) \leq \mathsf{C}\delta^2.$$

Consider now $r > 1$, and assume inductively that the Hausdorff distance between the intersection of the zero levels of $r - 1$ transversal functions and the intersection of the zero level of the respective linear interpolations on an n-simplex is quadratically smaller than the simplex diameter. Thus we have

$$(3.16) \qquad \Sigma_- = \{\omega_1(x) = 0\} \cap \cdots \cap \{\omega_{r-1}(x) = 0\},$$

$$(3.17) \qquad \hat{\Sigma}_- = \{\hat{\omega}_1(x) = 0\} \cap \cdots \cap \{\hat{\omega}_{r-1}(x) = 0\},$$

$$(3.18) \qquad d_{\mathcal{H}}(\Sigma_-, \hat{\Sigma}_-) \leq C\delta^2.$$

If we consider one more function $\omega_r(x)$ on the linear space $\hat{\Sigma}_-$, we are in the previous case, so there exists $A > 0$,

$$(3.19) \qquad d_{\mathcal{H}}(\hat{\Sigma}_- \cap \{\omega_r(x) = 0\}, \hat{\Sigma}_- \cap \{\hat{\omega}_r(x) = 0\}) \leq A\delta^2.$$

By transversality of the $\omega_1, \dots, \omega_r$, the fact holding for the linear space $\hat{\Sigma}_-$ holds also for the compact manifold with boundary Σ_- and the function ω_r (see Lemma 17 for the details). Thus there exists a $B > 0$ such that

$$(3.20) \qquad d_{\mathcal{H}}(\Sigma_- \cap \{\omega_r(x) = 0\}, \Sigma_- \cap \{\hat{\omega}_r(x) = 0\}) \leq B\delta^2.$$

On the other hand, for the intersection of the zero levels of the transversal functions $\omega_1, \dots, \omega_{r-1}$ on the linear space $\{\hat{\omega}_r(x) = 0\}$, by the inductive hypothesis there exists $C > 0$,

$$(3.21) \qquad d_{\mathcal{H}}(\{\omega_r(x) = 0\} \cap \Sigma_-, \{\omega_r(x) = 0\} \cap \hat{\Sigma}_-) \leq C\delta^2,$$

so the thesis is proved by the triangle inequality. □

LEMMA 17. *Let Σ be a manifold with boundary diffeomorphic to an n-simplex Δ, and let $\omega \colon \Sigma \to \mathbb{R}$ be differentiable and without critical points inside Σ. We have $\omega(x) = \hat{\omega}(x) + O(\delta^2)$, where $\hat{\omega}$ is an affine approximation and δ is the simplex diameter. Thus we have that*

$$(3.22) \qquad d_{\mathcal{H}}(\{\omega(x) = c\}, \hat{\omega}(x) = c\}) \leq C\delta^2 \quad for \ all \ c \in \mathbb{R}.$$

Proof. Let $\Delta \xrightarrow{\varphi} \Sigma$ be a diffeomorphism with $\xi > |\varphi'| > \eta > 0$. Thus we have, for all $y \in \Delta$,

$$\omega \circ \varphi(y) = \hat{\omega} \circ \varphi(y) + O(\delta^2).$$

For any $y^\star$ in the zero level of $\omega \circ \varphi$ we can find a line segment $[y_1, y_2]$, with y_1 being one of the nodes of Δ where $\omega \circ \varphi$ is negative and y_2 is a point on a face of Δ where on the forming nodes $\omega \circ \varphi$ is positive. By continuity there exists a point $\hat{y}$ on the line $[y_1, y_2]$ where $\hat{\omega} \circ \phi$ is zero.

Thus,

$$\omega \circ \varphi(y^\star) - \omega \circ \varphi(\hat{y}) = \hat{\omega} \circ \varphi(y^\star) - \hat{\omega} \circ \varphi(\hat{y}) + o(\delta^2) = \hat{\omega}' \circ \frac{\partial \varphi}{\partial w} |y^\star - \hat{y}| + o(\delta^2),$$

(3.23)

which gives

$$(3.24) \qquad\qquad |y^\star - \hat{y}| \leq \mathsf{C}\delta^2. \qquad \square$$

Note 18. The hypotheses of Theorem 16 are generic in the sense that they hold for an open and a dense set of functions. In particular, 0 is assumed to be a regular value for $\omega_1, \ldots, \omega_r$ because the set of the singular values has zero measure (Sard's theorem). See [7], [20], [28], [30], [33].

THEOREM 19. *In the simplex* $\Delta = \langle P_0, \ldots, P_n \rangle$, *if* θ *is the Pareto critical set and* $\hat{\theta}$ *is its linear approximation, there exists* $\mathsf{C} > 0$ *such that*

$$(3.25) \qquad\qquad d_{\mathcal{H}}(\theta, \hat{\theta}) \leq \mathsf{C}\delta^2.$$

Proof. The λ_j computed as described in Algorithm 1 are first order approximations to smooth functions, apart from a measure zero set of points. Thus the conclusions of Theorem 16 apply as well to the intersection of Σ with the half-spaces $\lambda_j(P) \geq 0$. $\square$

3.4. Second order algorithm. In Algorithm 2 we describe how to extract the stable critical set θ_s, i.e., the set of locally Pareto optimal points, from the critical set θ determined in the first order Algorithm 1.

ALGORITHM 2. SECOND ORDER ALGORITHM FOR THE STABLE PARETO CRITICAL SET θ_s.

1: Consider a set of data points $\mathcal{D} = \{P_1, \ldots, P_N\}$ and proceed as in Algorithm 1.
2: **for all** Delaunay simplex $\Delta = \langle P_{i_0}, \ldots, P_{i_n} \rangle$ in the tessellation **do**
3: compute the matrix of the second derivatives $D^2 u$ on the nodes $P_{i_0}, \ldots, P_{i_n}$;
4: compute a basis $w_1, \ldots, w_{n-m+1}$ for $\ker Du(Q)$, and set $\hat{H}(Q) := w^\top \cdot$
   $(\lambda_1(Q)\widehat{D^2 u_1}(Q) + \cdots + \lambda_m(Q)\widehat{D^2 u_m}(Q) \cdot w$;
6: compute the eigenvalues $\sigma_1, \ldots, \sigma_{n-m+1}$ of $\hat{H}(Q)$;
7: cut out from $\hat{\theta}$ the subpolytope $\hat{\theta}_s$, where $\sigma_k \leq 0$ for all $k = 1, \ldots, n - m + 1$;
8: **end for**
9: compose a simplicial complex glueing together adjacent polytopes $\hat{\theta}_s$.

The second derivatives could also be approximated computing the finite differences of the values of the gradients on the nodes of the n-simplex. Indeed, setting

$$v_i = P_i - P_0, \qquad i = 1, \ldots, n,$$

we have

$$(3.26) \quad D^2 u = \left(\frac{\partial^2 u}{\partial x_i \partial x_j} \right)_{i,j} = \sum_k \frac{\partial^2 u}{\partial v_k \partial x_j} \cdot \frac{\partial v_k}{\partial x_i} \simeq \sum_k (\nabla u(P_k) - \nabla u(P_0))_j \cdot (P_k - P_0)_i.$$

Using this formula, the quadratic precision cannot be guaranteed for locating boundary points of the stable critical set. Furthermore, because the boundary faces belong to different simplexes, the estimated boundary points for θ_s would jump from simplex to simplex.

On the other hand, the formula will be correct for discriminating the nature of inner stable critical points without extra computations. Boundary simplexes can thus be analyzed with second derivatives, allowing for the computation of the boundary of $\hat{\theta}_s$.

4. Applications.

4.1. Two functions in 2-dimensional examples. A series of examples in two dimensions is presented below. Via global analysis one sees that, for structurally stable mappings, the Pareto critical set is a 1-dimensional manifold with boundary contained in Σ. Critical points can only be of one of the following types:

1. *fold*; i.e., the mapping is locally equivalent to $u_1 = x_1$, $u_2 = x_2^2$;
2. *cusp*; i.e., the mapping is locally equivalent to $u_1 = x_1$, $u_2 = x_1 x_2 - \frac{1}{3} x_2^3$.

Therefore, the branches of Pareto critical points are composed by folds, which intersect only pairwise and at nonzero angles. Some local branches terminate in cusps, where the status of critical points can change from stable to unstable. Finally, images of folds and cusps do not intersect [7], [65].

Function gradients are evaluated on a grid of regular triangles, and the critical set θ is estimated according to the first order algorithm. Boundary points are marked with black diamonds. The generalized Hessian is estimated on the nodes of the critical set, computing second derivatives in the triangles where its index changes, allowing one to estimate the position of the points separating stable from unstable branches. Cusps are marked by a black star, stable branches are colored in red, unstable branches in orange, and finally noncritical branches are gray.

Example 1. Consider two negative definite quadratic polynomials. The critical stable set is a curve joining the two individual critical points. Other singular branches occur in outer regions of the domain.

$$u_1(x, y) = -1.05x^2 - 0.98y^2,$$
$$(4.1) \qquad u_2(x, y) = -0.99(x - 3)^2 - 1.03(y - 2.5)^2.$$

See Figure 4.1.

Example 2. This example is taken from [58].

$$u_1(x, y) = -y,$$
$$(4.2) \qquad u_2(x, y) = \frac{y - x^3}{x + 1}.$$

The critical set is a single curve split into a stable and an unstable branch, while the separating point is a cusp. See Figure 4.2.

Example 3. In the following mapping there are two second order polynomials, one negative definite and the other indefinite. The outcome is an (unbounded) global Pareto front and a local unbounded front terminating in a cusp.
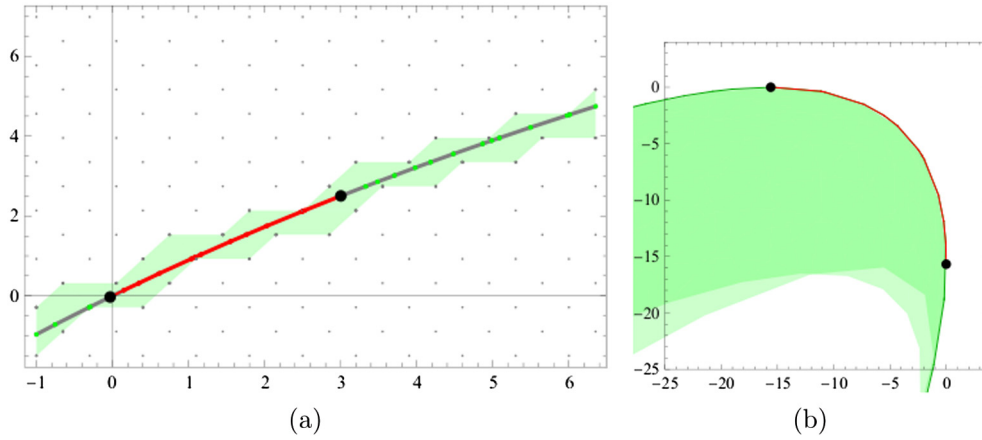
FIG. 4.1. *Pareto critical set* (a) *and the Pareto front* (b) *for Example* 1.

$$u_1(x, y) = -x^2 - y^2,$$
(4.3) $$u_2(x, y) = -(x - 6)^2 + (y + 0.3)^2.$$

See Figure 4.3.

Example 4. The following mapping is composed of a quadratic polynomial and a bimodal function. The resulting singular set is composed of an unbounded branch and two loops. One of the loops is critical and forms a local Pareto front delimited by two cusps, while the other loop is noncritical.
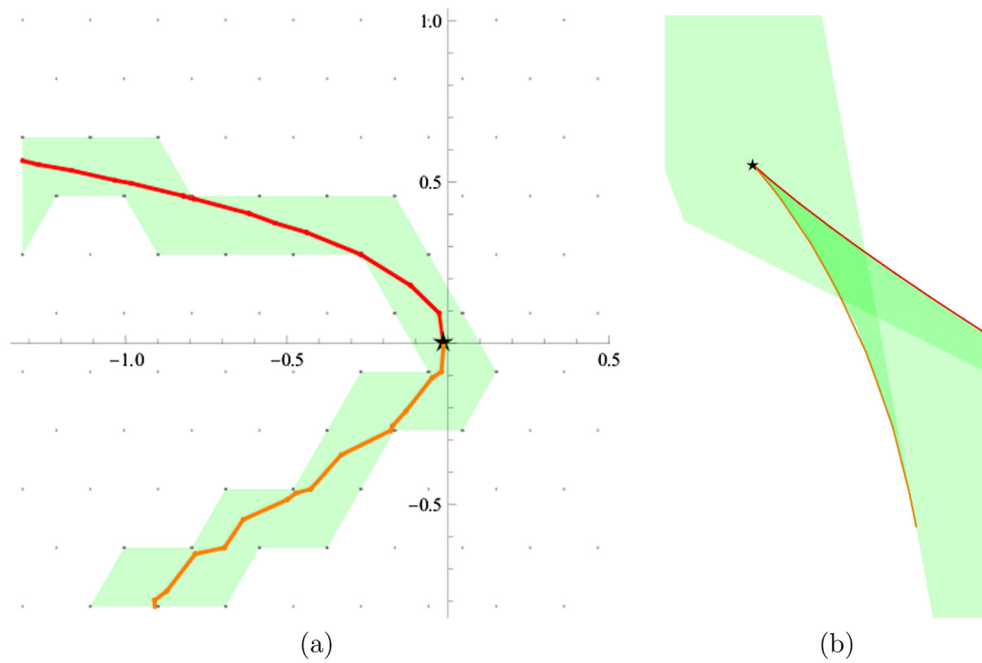

FIG. 4.2. *Example* 2. *Red line: stable critical set. Orange line: unstable critical set.*
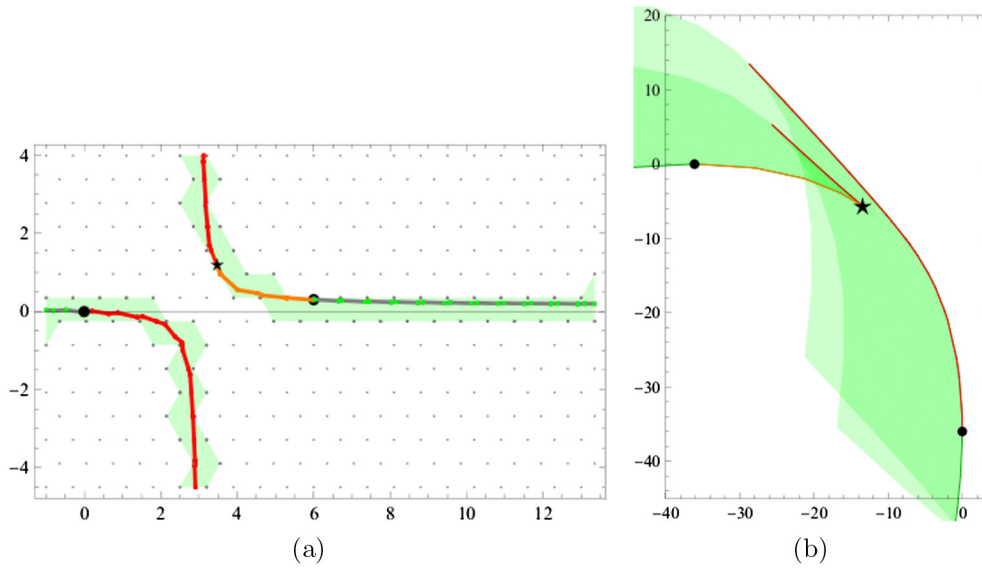
FIG. 4.3. *Example* 3.

$$u_1(x, y) = -x^2 - y^2 - 4(\exp(-(x+2)^2 - y^2) + \exp(-(x-2)^2 - y^2)),$$
$$(4.4) \quad u_2(x, y) = -(x-6)^2 - (y+0.5)^2.$$

See Figure 4.4.

4.2. Higher input dimension. *Example* 5. The following mapping demonstrates the capabilities of the method in distinguishing local and global features of the Pareto set. A widespread optimal branch is surpassed by a local branch. The sharper branch is composed of an unstable part (orange) and a stable part (red) which is interrupted by noncritical
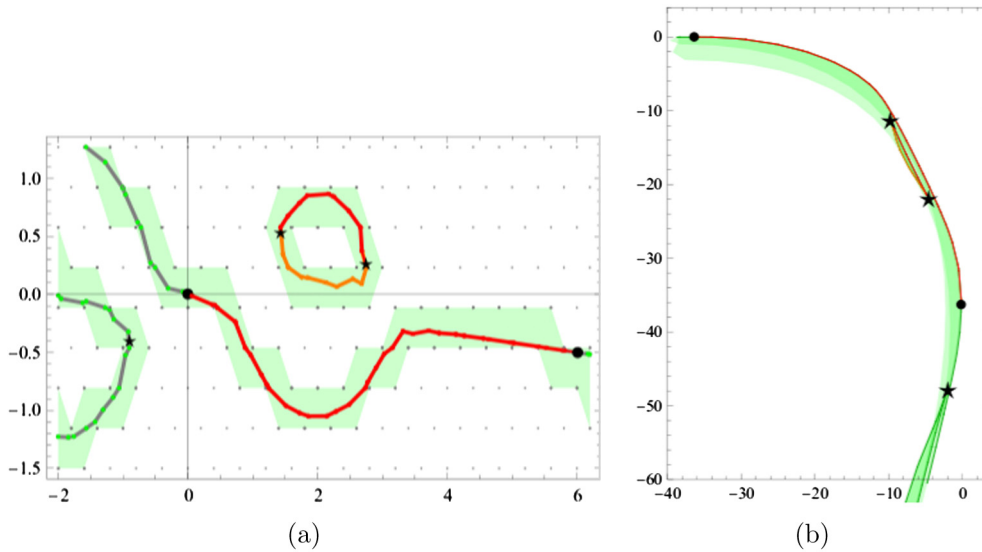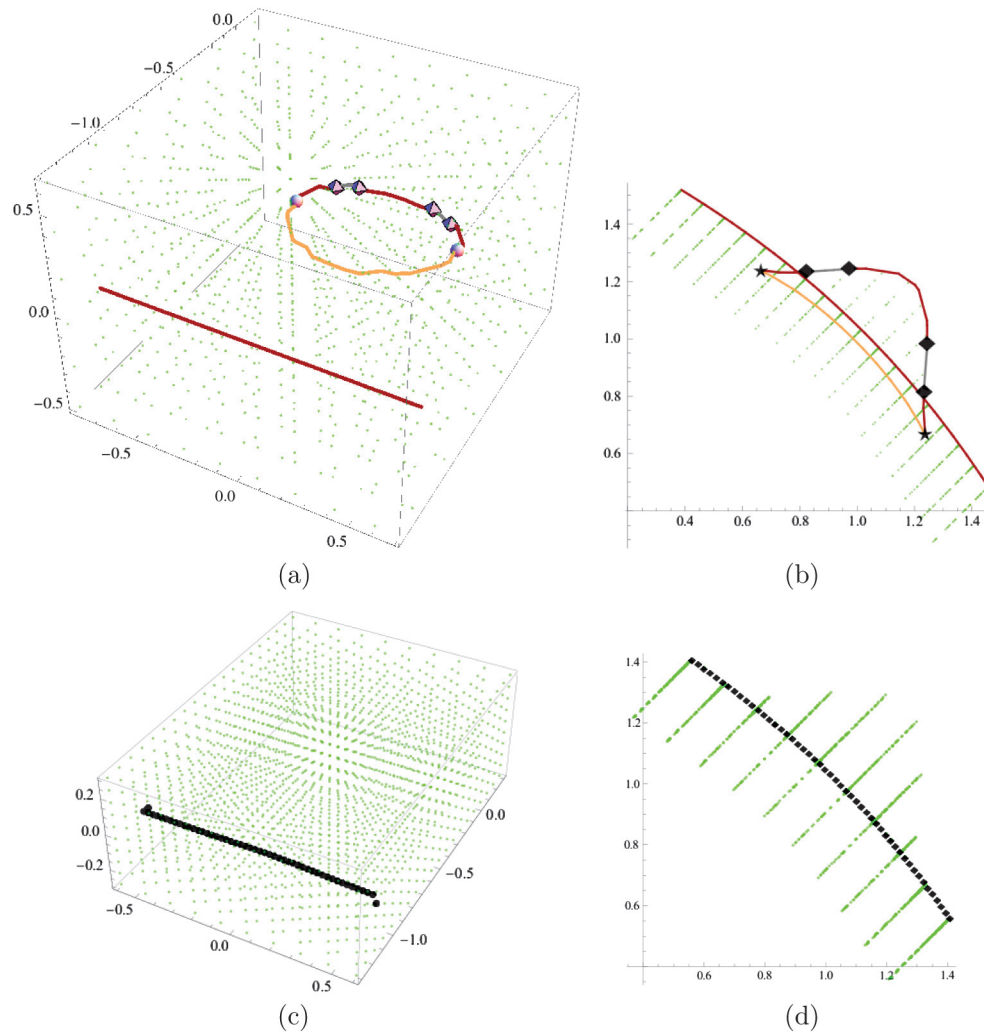

FIG. 4.4. *Example* 4.

FIG. 4.5. *Example* 5. *Panel* (a)*: singular (gray), Pareto critical (orange), and Pareto stable (red) sets in the problem domain. Green dots mark the nodes of the starting regular grid defining the tessellation. Octahedrons mark points separating critical and noncritical branches. Spheres separate stable from unstable branches; i.e., they mark cusps. Panel* (b)*: image of singular and Pareto sets. Diamonds separate critical from noncritical branches, while stars mark the cusps. Panels* (c)–(d)*: results obtained running the commercial implementation of NBI–AFSQP available in* **modeFRONTIERfi***, courtesy of E. Rigoni. Small green points are a starting regular grid, while marked points are the solutions of the* 50 *NBI subproblems.*

insertions (gray). Nevertheless, as illustrated in Figures 4.5(a)–(b), the two separate branches are properly recognized by the algorithm, and, moreover, the transitions among critical/noncritical and stable/unstable intervals are detected. For comparison, the outcome of the application of a commercial implementation of normal boundary intersection by Das and Dennis [12] is shown in Figures 4.5(c)–(d).[13] The starting grid (green dots) was $10 \times 20 \times 10$, and we considered 50 NBI subproblems. The sequence of NBI points is marked by black stars. For this particular problem, NBI tracks correctly the broad Pareto optimal branch, in the sense that it produces a parametrization of it. However, the smaller

---

[13]Applications of **modeFRONTIERfi** are courtesy of E. Rigoni at ESTECO.

branch is missed, although some of the points of the starting grid were close to this critical zone. It is clear that pointwise strategies suffer at tracking the Pareto optimal set and fail at performing widespread tradeoff studies, apart from small intervals where different fronts are far apart and do not change status (from critical to noncritical, stable to unstable, and so on).

$$p_0 = (0.0, 0.15, 0.0)^\top,$$
$$p_1 = (0.0, -1.1, 0.0)^\top,$$
$$M = \begin{pmatrix} -1.0 & -0.03 & 0.011 \\ -0.03 & -1. & 0.07 \\ 0.011 & 0.07 & -1.01 \end{pmatrix},$$
$$g(x, y, z, M, p, \sigma) = \sqrt{\frac{2\pi}{\sigma}} e^{\frac{((x,y,z)^\top - p)^\top M((x,y,z)^\top - p))}{\sigma^2}},$$
$$f(x, y, z) = g(x, y, z, M, p_0, 0.35) + g(x, y, 0.5z, M, p_1, 3.0),$$
$$u1(x, y, z) = \frac{\sqrt{2}}{2} x + \frac{\sqrt{2}}{2} f(x, y, z),$$
$$(4.5) \qquad u2(x, y, z) = -\frac{\sqrt{2}}{2} x + \frac{\sqrt{2}}{2} f(x, y, z).$$

Example 6. The following 6-dimensional example is a regularization of the third of the ZDT problems [15], which has degenerate second derivatives. The Pareto fronts of original and modified problems correspond to each other in output space. We used a Delaunay tessellation defined on 300 randomly generated points. The results are presented in Figure 4.6. Critical and merely singular branches are correctly represented. Note that the critical branches are correctly marked as unstable, being minima.

$$u1(x_1, \ldots, x_6) = x_1,$$
$$u2(x_1, \ldots, x_6) = 1 - \sqrt{x_1} - x_1 \sin(10\pi x_1) + x_2^2 + \ldots + x_6^2,$$
$$(4.6) \qquad x_1 \in [0.1, 0.425], \quad x_2, \ldots, x_6 \in [-0.16, 0.16].$$
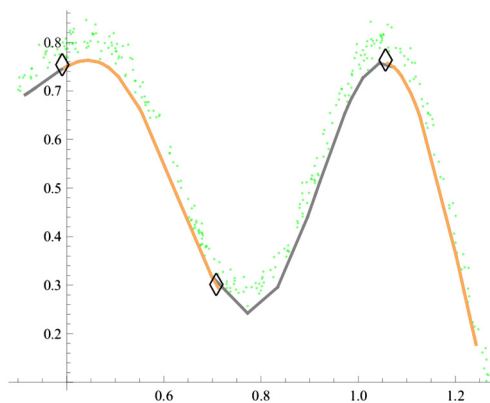

FIG. 4.6. *Example 6. Image of the singular (gray), Pareto critical (orange) sets for regularized ZDT3. Green dots mark the nodes of the starting random distribution of points defining the tessellation.*

4.3. Three functions examples. *Example* 7. The simplest nontrivial nondegenerate example we can build in the 3-dimensional case is composed of three negative definite second order polynomial functions $f_j(x), j = 1, 2, 3$. Additionally, we introduce a small nonpolynomial perturbation.

$$f_j(x) = (x - C_j)^\top \cdot \begin{pmatrix} -\alpha_{j,1} & 0 & 0 \\ 0 & -\alpha_{j,2} & 0 \\ 0 & 0 & -\alpha_{j,3} \end{pmatrix} \cdot (x - C_j), \qquad j = 1, 2, 3,$$

$$(4.7) \qquad \begin{pmatrix} u_1(x) \\ u_2(x) \\ u_3(x) \end{pmatrix} := \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} + \begin{pmatrix} 0 \\ \beta_2 \sin\left(\frac{\pi}{\gamma_2}(x+y)\right) \\ \beta_3 \cos\left(\frac{\pi}{\gamma_3}(x-y)\right) \end{pmatrix},$$

where $x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$, $\alpha_{j,i} > 0$, $i, j = 1, 2, 3$, $C_1, C_2, C_3 \in \mathbb{R}^3$ are distinct, noncollinear points, while β_j, γ_j are real numbers. In the generic case the singular set is a hypersurface of $\mathbb{R}^3$, while the critical set θ, which is stable, is diffeomorphic to a triangle; i.e., θ is a compact connected manifold with boundary and three corners, corresponding to the minima of the three functions u_1, u_2, u_3. See Figure 4.7(a).

Example 8. We break the convexity of the previous example by adding a secondary maximum to the first function. We define a further negative definite, second order polynomial $f_4(x)$ and set $u(x)$ as

$$(4.8) \qquad \begin{pmatrix} u_1(x) \\ u_2(x) \\ u_3(x) \end{pmatrix} := \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} + \begin{pmatrix} \beta_1 \exp\left(\frac{1}{\gamma_1} f_4(x)\right) \\ \beta_2 \sin\left(\frac{\pi}{\gamma_2}(x+y)\right) \\ \beta_3 \cos\left(\frac{\pi}{\gamma_3}(x-y)\right) \end{pmatrix}.$$

The main portion of the Pareto set is slightly deformed while a new branch appears. In Figure 4.7(b) is shown the resulting Pareto critical set θ obtained by iterative application of Algorithm 1 as described in section 5.
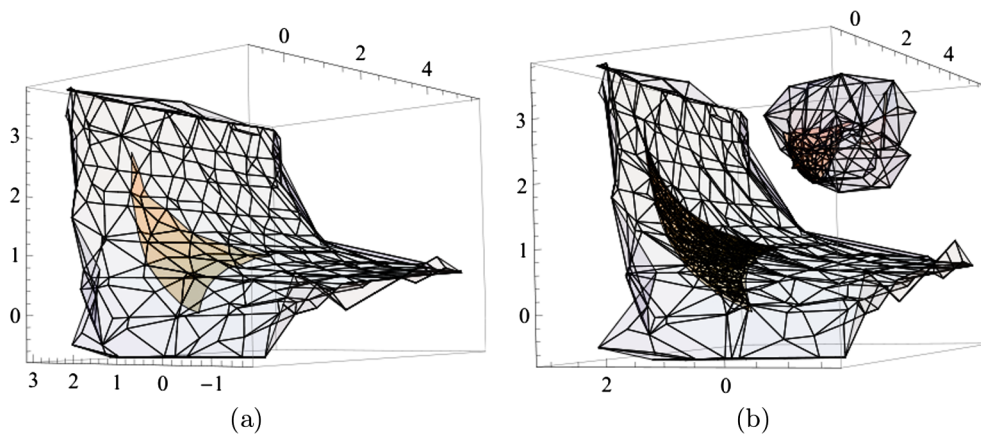

FIG. 4.7. (a) *Critical set θ of Example* 7. (b) *Critical set θ of Example* 8.

4.4. A constrained example. We briefly sketch here an adaptation of Algorithm 1 to the case of equality constraints. Next we illustrate a simple application. Let $W := \{x \in \mathbb{R}^n | g(x) = 0\}$, where $g\colon \mathbb{R}^n \to \mathbb{R}^{n-d}$ is a smooth function such that $\partial g|_W$ has maximum rank.

ALGORITHM 3. EQUALITY CONSTRAINT CASE FOR THE FIRST ORDER ALGORITHM.

1: Determine a piecewise-linear approximation $\hat{W}$ of W, with nodes $P_1, \ldots, P_n$;
2: **for all** simplex Δ in the tessellation $\hat{W}$, **do**
3: determine a piecewise-linear approximation of the singular, critical, and stable sets possibly crossing the simplex. This is done on the basis of the projections of the gradients of the u_j's on the tangent space to W. In principle, a basis for TW should be chosen respecting the orientation;
4: **for all** Node P of the simplex Δ **do**
5: compute $Dg(P)$ and project $\operatorname{grad} u_j(P)$ on $\ker Dg(P)$ via $\pi_{Dg(P)}$;
6: compute an independent set of minors for the matrix $(\pi \operatorname{grad} u_1(P), \ldots, \pi \operatorname{grad} u_m(P))$; or, equivalently,
7: compute an independent set of minors for the matrix $(\operatorname{grad} g_1, \ldots, \operatorname{grad} g_{n-d}, \operatorname{grad} u_1, \ldots, \operatorname{grad} u_m)$;
8: **end for**
9: determine if all minors vanish inside the simplex S, and in that case locate $\hat{\Sigma}$ via inverse linear interpolation;
10: estimate λ_j and determine the critical set $\hat{\theta}$ as in Algorithm 1;
11: **end for**
12: eventually glue together adjacent portions of $\hat{\Sigma}$ and $\hat{\theta}$.

Example 9. Maybe the simplest example of a constrained problem is when $W = \mathbb{S}^2$ and the objectives are the first two coordinates, $u_1(x_1, x_2, x_3) := x_1$, $u_2(x_1, x_2, x_3) := x_2$.[14] Explicit algebraic computation gives that the singular set Σ is the equator of the sphere, where the two curvilinear segments where $x_1 x_2 > 0$ are the critical set θ, as illustrated in Figure 4.8(a). By applying Algorithm 3, we start by approximating the sphere by an icosahedron. At every node $P = (x_1, x_2, x_3)$, $\frac{1}{2}(x_1^2 + x_2^2 + x_3^2 - 1) = 0$, we have $Dg(P) = (x_1, x_2, x_3)$; therefore, the projections of the gradients of u_j are $(1 - x_1^2, -x_1 x_2, -x_1 x_3)$ and $(-x_1 x_2, 1 - x_2^2, -x_1 x_3)$. The singular set Σ passes through the triangles where the pair of vectors $\pi \operatorname{grad} u_1$ and $\pi \operatorname{grad} u_2$ change orientation in the tangent plane to $\mathbb{S}^2$. It is equivalent then to compute the determinant of the matrix whose columns are $\operatorname{grad} g$, $\operatorname{grad} u_1$, and $\operatorname{grad} u_2$ and to determine the line along which it vanishes. This gives exactly the "equator" of the icosahedron. The signs of the λ_j depend on the sign of the scalar product among $\pi \operatorname{grad} u_j$, again giving as turning points the intersections with the axes. The results are summarized in Figure 4.8(a) and (b).

5. Iterative schemes. The previously presented approach defines an approximation of the Pareto optimal set given any distribution of points in the domain. Here we propose and discuss an iterative scheme. At every step a selection of points from the approximated Pareto optimal set is added to the dataset $\mathcal{D}$, the gradients in the new points are evaluated, the tessellation is updated, and a refined approximation of the Pareto set is built. The desired effect is obviously to get closer and closer to the actual optimal set, but an efficient strategy should produce an as uniform as possible discretization of the optimal set.

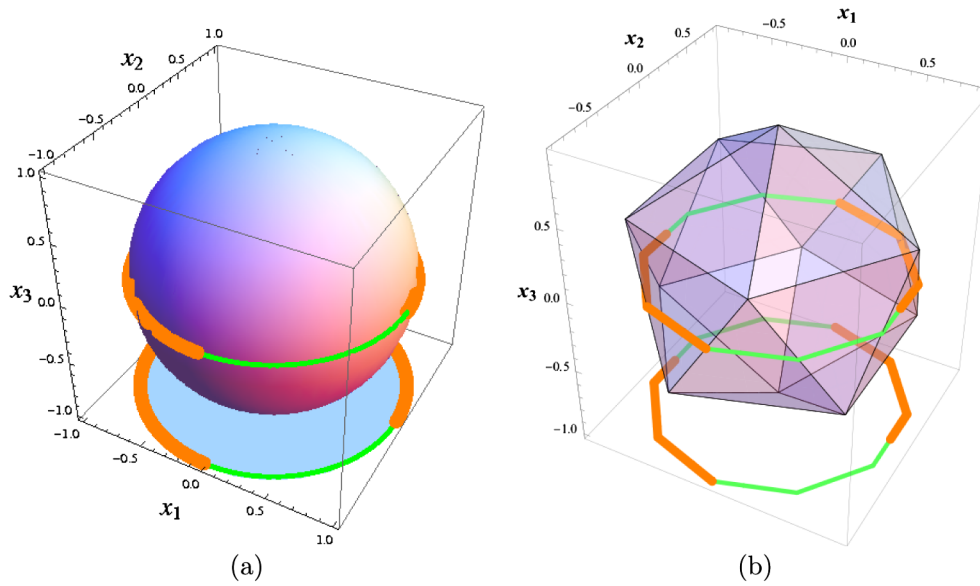[14] This example is also discussed in [13].

FIG. 4.8. (a) *Singular and critical sets determined analytically for Example* 9. (b) *Piecewise-linear approximation of the sphere, of the singular set (green solid curves), and of the critical set (orange and thicker curves).*

A naïve approach would suggest to insert in the set of the candidates for evaluation all of the nodes of the complexes, i.e., all the stable admissible vertices computed and all of the boundary points, both for criticality and stability. Nevertheless, a glance at the examples of the previous section reveals that the sizes of the optimal complexes cover a wide distribution; in particular, if Σ passes close to tessellation nodes, very small patches are generated. Moreover, little experience shows that large patches are reduced sensibly slowly if none of their internal points is introduced. With this criteria in mind we introduce an iterative scheme for the case of two functions.

5.1. Two-function iterative scheme. In the two-function case the Pareto optimal set is a 1-dimensional manifold with boundary, i.e., a collection of curved intervals. The discrete approximation is a collection of polygonal curves. For every interval a sequence of candidate points equally spaced along the polygonal curve is extracted. The number of points is chosen equal to the number of segments so that approximately every triangle containing optimal points is split as close as possible to the optimal set.

5.2. Higher number of functions. It seems reasonable to take into account the stratified structure of θ in the design of an iterative strategy. In fact, strata should be filled as uniformly as possible, where the uniformity is determined according to the k-dimensional measure if k is the dimension of the stratum. So, taking, for instance, the situation of Example 7, corners' approximations are reevaluated at each iteration, uniformly spaced points are taken along boundary lines, exactly as in the two-functions case, while internal points should be distributed proportionally to the area of the triangles and polygons composing $\hat{\theta}$. This is more difficult to be defined precisely. Indeed, the problem of uniformly filling a general n-dimensional region is a long-time crucial issue for statistical applications [48]. Furthermore, in our problem we have to fill uni-

formly a general n-dimensional *manifold*; thus we have to somehow take into account the effects of the curvature on the measure of the volumes.

Taking inspiration from a *Design of Experiments* strategy called *maximin distance design* [21], we proceed as described in Algorithm 4. This algorithm, because of point 6, can lead to long and thin simplexes and to numerical instabilities when iterated many times. This problem can be tackled by the application of mesh improvement strategies, as described below for the case of 2-dimensional domains. However, to the author's knowledge, general dimension mesh improvement strategies are still not available at present.

ALGORITHM 4. UNIFORMLY FILLING A SIMPLICIAL COMPLEX.
    1: Tessellate in simplexes the polytopes of the mesh;
    2: build the adjacency lists of the simplexes;
    3: evaluate the volume of each simplex;
    4: for every simplex define the accumulated volume as the sum of its volume and
       the volume of the adjacent simplexes;
    5: pick the simplex with the maximum accumulated volume;
    6: add to the candidates stack the center of mass of this maximal simplex;
    7: **repeat**
    8:    recompute the accumulated volumes excluding the already picked simplexes;
    9: **until** the desired number of candidate points is collected.

5.3. Stopping criteria. Analogously to gradient-based methods of single function optimization (nonlinear conjugate gradient, Newton, and Newton-like methods), a stopping criterion could be based on the magnitude of the minors $M_1, \ldots, M_r$ computed in the points of the last iteration. The magnitude of the minors is analogous to the magnitude of the gradients for single objective optimization.

In fact, we could define a different iterative strategy taking the rule of subdividing only stable critical triangles contained in simplexes where the minors are larger than a prescribed threshold.

5.4. Application. We show the behavior of the iterative scheme described above applied to the mapping in Example 4. At each iteration we generate a number of evenly spaced points along the approximate stable Pareto critical set. In order to exhibit the claimed quadratic convergence, it is necessary to sample the approximated optimal set by quadratically finer intervals, i.e., comparable to the precision gained. As a result the density of points will grow exponentially w.r.t. the number of iterations. Such a density of points rapidly deteriorates the mesh quality; i.e., skinny triangles suddenly appear leading to numerical instability. Thus, at each iteration, a number of extra nodes (namely, the circumcenters of the most skinny triangles) should be introduced in the mesh in order to produce a nice grading. At this extent we have coupled our method with Ruppert's algorithm, as implemented in the triangular mesh refinement software TRIANGLE by Shewchuk [53], [54].

Already at the fifth iteration the triangulation starts to suffer from numerical instability; thus we consider $\hat{\theta}_S^{(4)}$ generated at the fourth iteration as the optimum and evaluate the Hausdorff distances between $\hat{\theta}_S^{(i)}$ and $\hat{\theta}_S^{(4)}$ for $i = 1, \ldots, 3$. As can be seen in Figure 5.2(a), the Hausdorff distances between the approximated Pareto sets and the numerical optimum converge superlinearly. For reference also the convergence behavior of the maximum and the mean minors magnitude is reported.
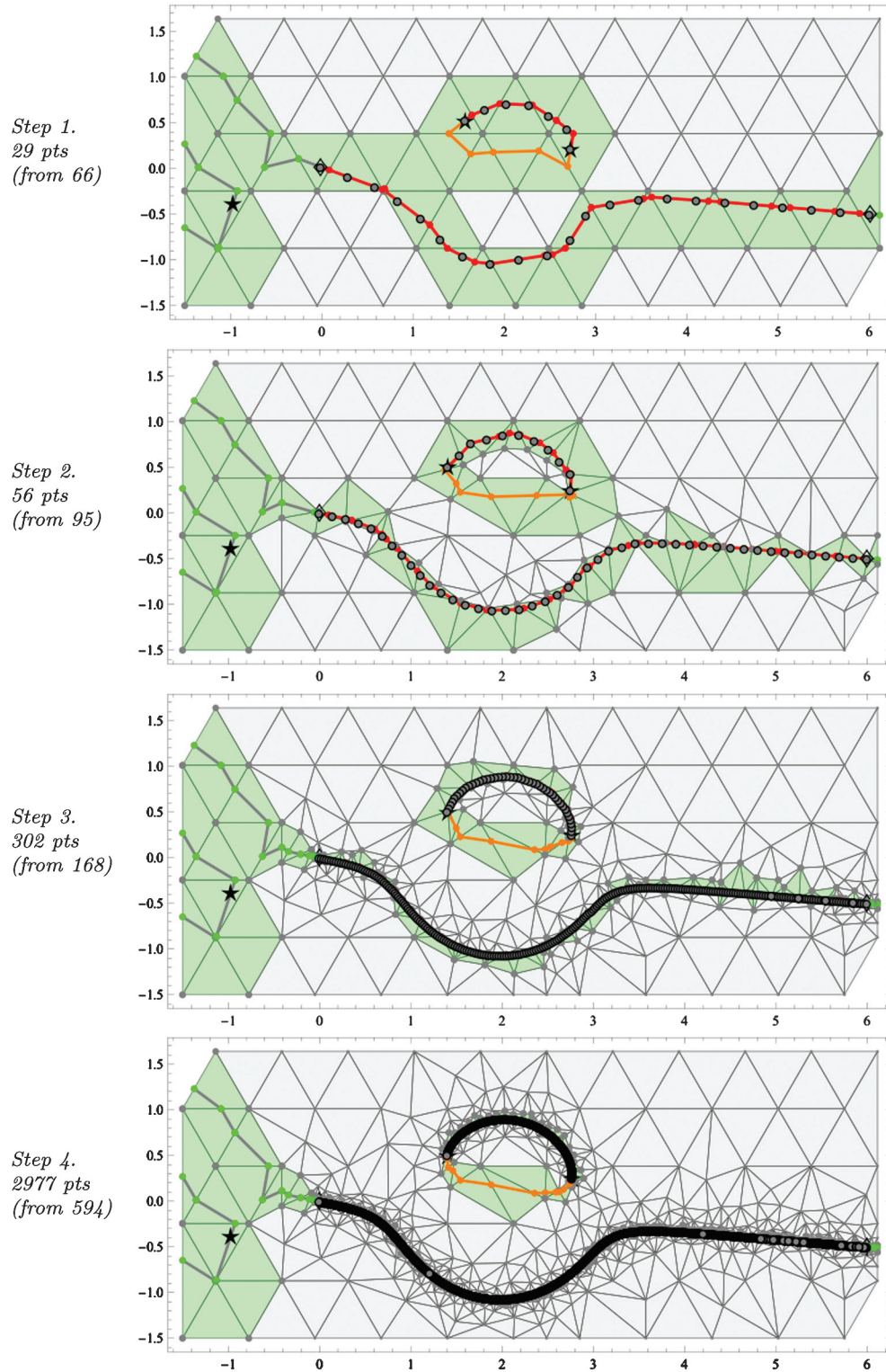
Step 1.
29 pts
(from 66)

Step 2.
56 pts
(from 95)

Step 3.
302 pts
(from 168)

Step 4.
2977 pts
(from 594)

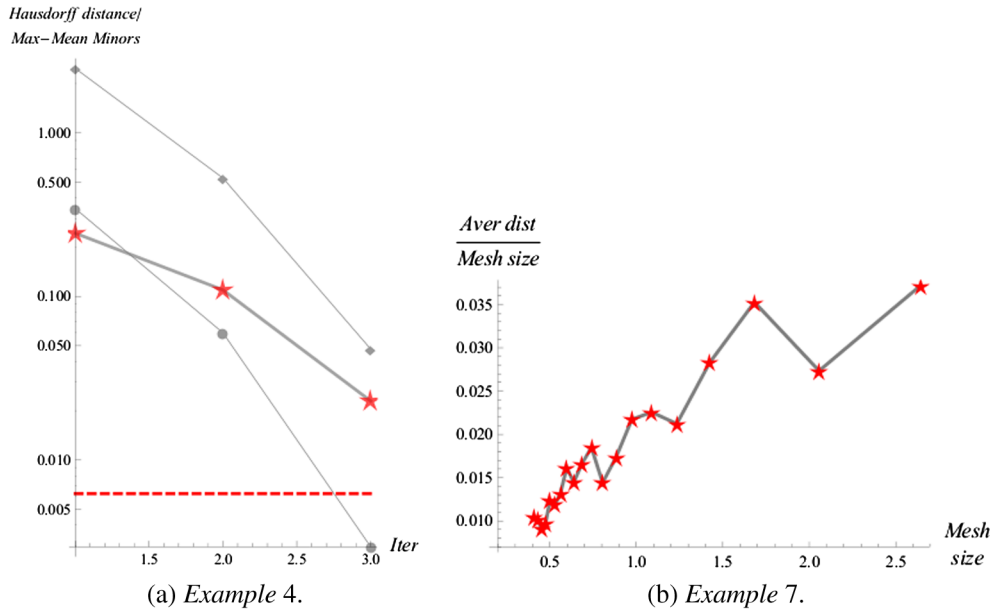Fig. 5.1. *Iterative scheme for the mapping in Example* 4.

FIG. 5.2. *Convergence behavior for iterative schemes applied to Examples 4 and 7. (a) Iterative scheme applied to Example 4. Stars represent the Hausdorff distance between the approximated Pareto set at each iteration and the Pareto set obtained at the fourth iteration, which is employed as an optimum. Diamonds and circles represent, respectively, the maximum and the mean absolute value of the minors of the Jacobian matrix computed on the points of the approximated Pareto set. Log scale reveals the superlinear convergence behavior. The horizontal dashed line represents the mesh size of the numerical optimum. (b) Algorithm 1 applied to Example 7 using progressively finer regular meshes. Stars represent the average distance between a node of the optimal set and the triangles of the approximation and vice versa. The ratio between the distance and the the mesh size decreases faster than linearly according to Theorem 16.*

In Figure 5.1 is illustrated how the triangulation and the representation of the Pareto set evolves from one iteration to the subsequent.

In Figure 5.3, the 3-dimensional problem of Example 7 is tackled by the procedure described in Algorithm 4. The algorithm has been applied by introducing only a small number ($\sim$10) of new points on the sites with the largest magnitude of the minors. In such a way it was possible to iterate 70 times the scheme reaching a very small magnitude for the minors.

Because of the mentioned exponentially growing number of samples necessary to exhibit quadratic convergence speed for the iterative scheme, the experiment described for the 2-dimensional case becomes prohibitive in three dimensions.

On the other hand, a superlinear precision can be verified as well by means of a sequence of approximations obtained from progressively finer regular meshes, corresponding to a sequence of mesh sizes $s = 2.8, \ldots, 0.4$. Because the Hausdorff distance among the first s-approximation and the optimal set is already comparable to the largest mesh size of the optimal set, we analyze the sequence of average distances between a point of one set and the triangles of the other set, instead of considering the maximum distances. These average distances decrease faster than linearly as it can be seen by plotting the ratio of distances and mesh sizes versus the mesh sizes, as reported in Figure 5.2(b).
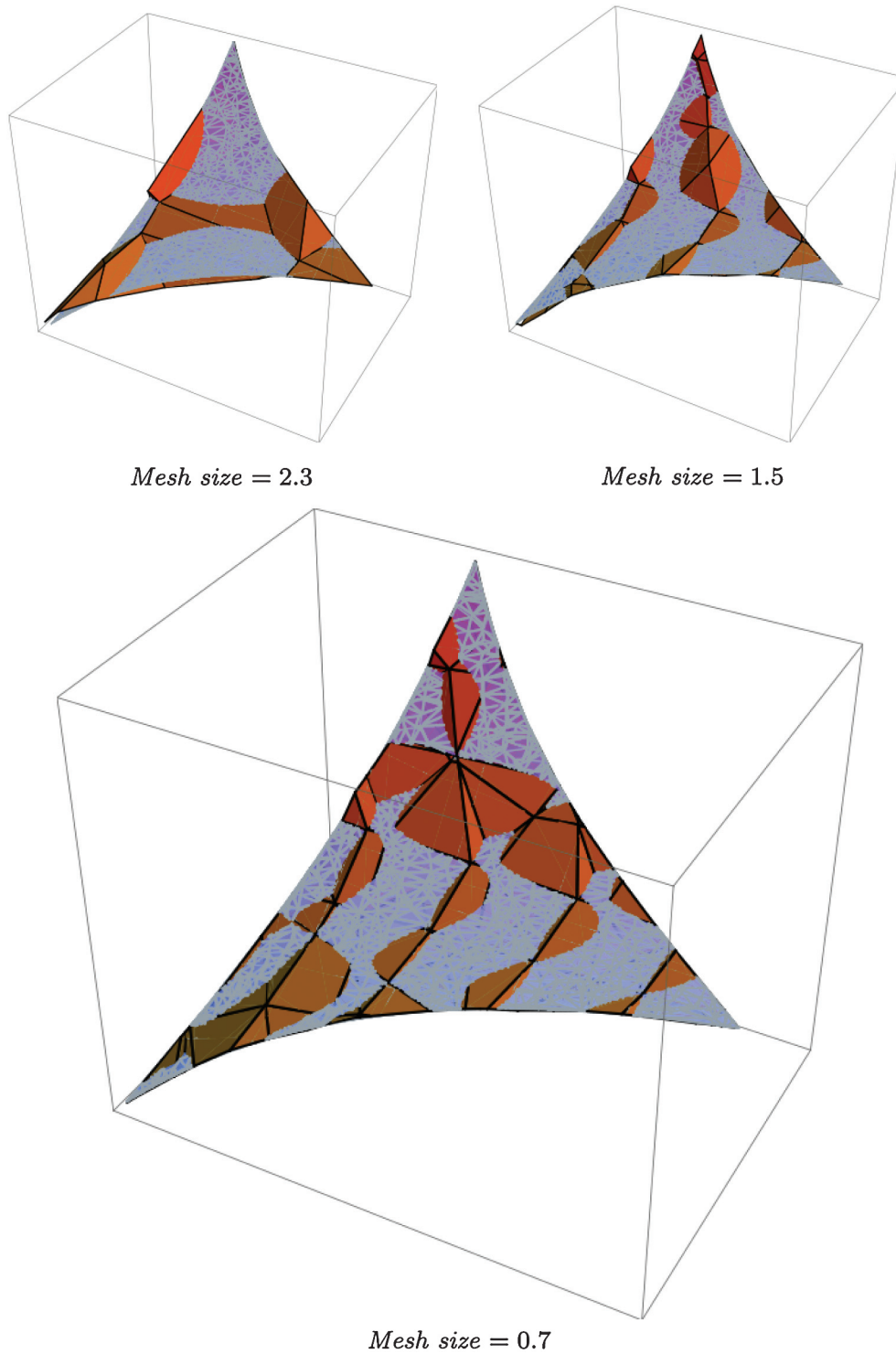
Mesh size = 2.3   *Mesh size = 1.5*

Mesh size = 0.7

FIG. 5.3. *Algorithm* 1 *applied to Example* 7 *for progressively finer meshes. The blue surface is the numerical optimum obtained with long application of the iterative scheme.*

6. Conclusions and perspectives. We have presented a novel multiobjective optimization method which exploits the manifold structure underlying the set of Pareto optimal points. Global analysis seems the proper setting where those structures arise and can be studied. We approximate Pareto sets via simplicial complexes, specializing simplicial pivoting techniques for detecting the singular manifold Σ and successively cutting out critical and stable subsets θ and θ_S. By contrast, most of the available strategies are aimed at producing a scatter of optimal points whose images should be evenly distributed. We have illustrated some generic situations where this program could not be successfully completed via such pointwise strategies because of nonconvexities of the functions. Adopting the Hausdorff measure, Newton-type estimates lead to quadratic convergence in a setwise sense.

Because of its global character, the method proposed here is demanding. Delaunay tessellations, in particular, are defined for every possible input dimension, but they are numerically workable only for cases of small dimension. The theory of singularities of mappings also highlights further limitations encountered when dealing with a large number of functions. Lastly, we have everywhere assumed the differentiability of the functions. Therefore, the method is not suitable for nonsmooth optimization; instead it is supposed to be applicable also via smooth surrogate functions when approximations are consistent with the functions at hand. Possible extensions of the algorithms described in this paper are conditioned by the issues enumerated below.

6.1. The curse of dimensionality. The first problem one encounters when trying to apply these algorithms to industrial strength problems is the limitations to the input dimension. The whole procedure is based on a Delaunay tessellation of the input domain, whose complexity grows exponentially with dimension. As pointed out, for instance, in the `qhull` documentation [8], building the convex hull of a 9-hypercube is computationally exhaustive. Analogous limitations are encountered in global optimization, where the search for optima in high-dimensional domains cannot realistically be performed on real case problems. Indeed, typically, global search algorithms are rarely tested and compared over dimensions larger than five (see [22], [23], [24], [41], [50], [52], [70]). This problem is structural and cannot be resolved by augmenting the computational resources. Therefore, the presented algorithms are best suited for low-dimensional problems. In fact, the *curse of dimensionality* is a strong motivation for reflecting carefully on the necessity of introducing extra input variables when tackling new problems and designing experiments. A possible exit strategy could be screening the input variables [49], [59]. This practice can be surprisingly successful, because usually sparsity of effects occurs, revealing a pronounced hierarchy among input variables, leading to sensible simplification of the problem formulation.[15]

Alternatively, as described in the recent paper [6], it is possible to redefine any problem in n-dimensional space in an equivalent problem in a linear subspace of dimension $2(m-1)+1$, if m is the number of objectives. This is because the singular set is an $(m-1)$-manifold, and by Whithney's embedding theorem, in the compact case, almost all projections on linear $(2(m-1)+1)$-dimensional subspaces are diffeomorphisms. This would mean that bi-objective problems could be equivalently discussed in a

---

[15]The sparsity of effects is an empirical law stating that in a generic physical experiment one usually observes that 80% of the effects are due to 20% of the factors. Related phenomena are that the first order contributions are the most important, while higher order contributions decay fast. Finally, one observes that the largest interactions (second order contributions) are a combination of the strongest factors. See, for instance, [69].

3-dimensional domain, 3 objectives would require only 5 input variables, and so on. This would dramatically reduce the computational burden of the tessellations involved.

6.2. Surrogate models. In industrial applications, when the objective functions at hand could be nondifferentiable, or may be computationally too expensive, preventing the computation of derivatives, we figure that the applicability of the algorithm proposed here will be significantly extended by using surrogate models. There exists an extensive literature developed in recent years on this subject (see [22], [23], [47], [48], and the references therein), also with specific applications to multiobjective optimization [25], [26].

The procedures of this paper can be adapted applying Algorithms 1 and 2 to a surrogate model $\tilde{u}$ fitted to the values of the true functions u computed on the given data points. On the outcoming candidate points, new evaluations of u are to be computed, and a new surrogate model is to be fitted to the increased dataset. This reduces the computational effort for computing derivatives and furthermore prevents premature stopping of the optimization process due to accidental failure of function evaluation at some data point. Again, the convergence to the Pareto sets of the true functions is guaranteed via global analysis.

REFERENCES

[1] E. L. Allgower and K. Georg, *Simplicial and continuation methods for approximating fixed points and solutions to systems of equations*, SIAM Rev., 22 (1980), pp. 28–85.

[2] E. L. Allgower and K. Georg, *Estimates for piecewise linear approximations of implicitly defined manifolds*, Appl. Math. Lett., 2 (1989), pp. 111–115.

[3] E. L. Allgower and K. Georg, *Piecewise linear methods for nonlinear equations and optimization*, J. Comput. Appl. Math., 124 (2000), pp. 245–261.

[4] E. L. Allgower and S. Gnutzmann, *Simplicial pivoting for mesh generation of implicitly defined surfaces*, Comput. Aided Geom. Design, 8 (1991), pp. 305–325.

[5] E. L. Allgower and P. H. Schmidt, *An algorithm for piecewise-linear approximation of an implicitly defined manifold*, SIAM J. Numer. Anal., 22 (1985), pp. 322–346.

[6] E. L. Allgower and A. J. Sommese, *Piecewise linear approximation of smooth compact fibers*, J. Complexity, 18 (2002), pp. 547–556.

[7] V. I. Arnol'd, *Singularities of smooth mappings*, Russian Math. Surveys, 23 (1968), pp. 1–43.

[8] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, *The quickhull algorithm for convex hulls*, ACM Trans. Math. Software, 22 (1996), pp. 469–483.

[9] H. Benson and S. Sayin, *Towards finding global representations of the efficient set in multiple objective mathematical programming*, Naval Res. Logist., 44 (1997), pp. 47–67.

[10] A. Bowyer, *Computing Dirichlet tessellations*, Comput. J., 24 (1981), pp. 162–166.

[11] E. Calabi, *Quasi-surjective mappings and a generalization of Morse theory*, in Proceedings of the U.S.-Japan Seminar in Differential Geometry, Nippon Hyoronsha, Tokyo, 1966, pp. 13–16.

[12] I. Das and J. E. Dennis, *Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems*, SIAM J. Optim., 8 (1998), pp. 631–657.

[13] W. de Melo, *On the structure of the Pareto set of generic mappings*, Bol. Soc. Brasil. Mat., 7 (1976), pp. 121–126.

[14] W. de Melo, *Stability and optimization of several functions*, Topology, 15 (1976), pp. 1–12.

[15] K. Deb, *Multi-objective genetic algorithms: Problem difficulties and construction of test problems*, Evolutionary Computation, 7 (1999), pp. 205–230.

[16] G. Debreu, *Regular differentiable economies*, The American Economic Review, 66 (1976), pp. 280–287.

[17] G. Debreu, *Stephen Smale and the economic theory of general equilibrium*, in From Topology to Computation: Proceedings of the Smalefest, Springer, New York, 1993, pp. 131–146.

[18] J. Fliege, L. M. G. Drummond, and B. F. Svaiter, *Newton's method for multiobjective optimization*, SIAM J. Optim., 20 (2009), pp. 602–626.

[19] D. Gourion and D. T. Luc, *Generating the weakly efficient set of nonconvex multiobjective problems*, J. Global Optim., 41 (2008), pp. 517–538.

[20] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[21] M. E. Johnson, L. M. Moore, and D. Ylvisaker, *Minimax and maximin distance designs*, J. Statist. Plann. Inference, 26 (1990), pp. 131–148.

[22] D. R. Jones, *A taxonomy of global optimization methods based on response surfaces*, J. Global Optim., 21 (2001), pp. 345–383.

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

[24] C. Khompatraporn, J. D. Pintér, and Z. B. Zabinsky, *Comparative assessment of algorithms and software for global optimization*, J. Global Optim., 31 (2005), pp. 613–633.

[25] J. Knowles, *ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems*, IEEE Trans. Evol. Comput., 10 (2006), pp. 50–66.

[26] J. Knowles and H. Nakayama, *Meta-modeling in multiobjective optimization*, in Multiobjective Optimization, Springer-Verlag, Berlin, 2008, pp. 245–284.

[27] H. Levine, *Singularities of differentiable mappings*, in Proceedings of Liverpool Singularities—Symposium I, Lecture Notes in Math. 192, C. Wall, ed., Springer, New York, 1971, pp. 1–21.

[28] H. Levine, *Stable maps: An introduction with low dimensional examples*, Bol. Soc. Brasil. Mat., 7 (1976), pp. 145–184.

[29] D. T. Luc, T. Q. Phong, and M. Volle, *Scalarizing functions for generating the weakly efficient solution set in convex multiobjective problems*, SIAM J. Optim., 15 (2005) pp. 987–1001.

[30] J. N. Mather, *Stability of C^∞ mappings. VI: The nice dimensions*, in Proceedings of Liverpool Singularities—Symposium, I, Lecture Notes in Math. 192, Springer, New York, 1971, pp. 207–253.

[31] A. Messac and C. A. Mattson, *Normal constraint method with guarantee of even representation of complete Pareto frontier*, AIAA J., 42 (2004), pp. 2101–2111.

[32] A. Messac and A. A. Mullur, *A computationally efficient metamodeling approach for expensive multiobjective optimization*, Optim. Eng., 9 (2008), pp. 37–67.

[33] P. W. Michor, *Elementary Catastrophe Theory*, Monografii Matematice (Timişoara) [Mathematical Monographs (Timişoara)] 24, Universitatea din Timişoara, Facultatea de Ştiinţe ale Naturii, Secţia Matematică, Timişoara, 1985.

[34] K. Miettinen, *Nonlinear Multiobjective Optimization*, International Series in Operations Research & Management Science 12, Kluwer Academic Publishers, Boston, MA, 1999.

[35] E. Miglierina and E. Molho, *Convergence of minimal sets in convex vector optimization*, SIAM J. Optim., 15 (2004), pp. 513–526.

[36] E. Miglierina, E. Molho, and M. Rocca, *Critical points index for vector functions and vector optimization*, J. Optim. Theory Appl., 138 (2008), pp. 479–496.

[37] J. Milnor, *Morse Theory*, Based on lecture notes by M. Spivak and R. Wells, Annals of Mathematics Studies 51, Princeton University Press, Princeton, N.J., 1963.

[38] V. Pareto, *Cours d'économie politique/Professé à l'université de Lausanne*, Rouge, Lausanne, 1896–1897.

[39] V. Pareto, *Manuale di economia politica con una introduzione alla scienza sociale*, Piccola biblioteca scientifica 13, Società editrice libraria, Milan, 1906.

[40] V. Pereyra, *Fast computation of equispaced Pareto manifolds and Pareto fronts for multiobjective optimization problems*, Math. Comput. Simulation, 79 (2009), pp. 1935–1947.

[41] J. D. Pintér, *Global optimization in practice: State of the art and perspectives*, in Advances in Applied Mathematics and Global Optimization, Adv. Mech. Math. 17, Springer, New York, 2009, pp. 377–404.

[42] I. R. Porteous, *Simple singularities of maps*, in Proceedings of Liverpool Singularities—Symposium I, Lecture Notes in Math. 192, Springer, New York, 1971, pp. 286–307.

[43] J. Rakowska, R. T. Haftka, and L. T. Watson, *Tracing the Efficient Curve for Multiobjective Control-structure Optimization*, Technical report, Virginia Polytechnic Institute & State University, Blacksburg, VA, 1991.

[44] J. Rakowska, R. T. Haftka, and L. T. Watson, *Multi-objective control-structure optimization via homotopy methods*, SIAM J. Optim., 3 (1993), pp. 654–667.

[45] J. R. Rao and P. Y. Papalambros, *A non-linear programming continuation strategy for one parameter design optimization problems*, in Proceedings of ASME Design Automation Conference, Montreal, Quebec, Canada, 1989, pp. 77–89.

[46] S. Ruzika and M. M. Wiecek, *Approximation methods in multiobjective programming*, J. Optim. Theory Appl., 126 (2005), pp. 473–501.

[47] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, *Design and analysis of computer experiments*, Statist. Sci., 4 (1989), pp. 409–423.

[48] T. J. Santner, B. J. Williams, and W. I. Notz, *The Design and Analysis of Computer Experiments*, Springer Series in Statistics, Springer-Verlag, New York, 2003.

[49] M. Schonlau and W. Welch, *Screening the input variables to a computer model via analysis of variance and visualization*, in Screening Methods for Experimentation in Industry, Drug Discovery and Genetics, Springer, New York, 2005, pp. 308–327.

[50] M. Schonlau, W. J. Welch, and D. R. Jones, *Global versus local search in constrained optimization of computer models*, in New Developments and Applications in Experimental Design, IMS Lecture Notes—Monograph Series 34, N. Flournoy, W. F. Rosenberger, and W. K. Wong, eds., Institute of Mathematical Sciences, Hayward, CA, 1998, pp. 11–25.

[51] O. Schütze, A. Dell'Aere, and M. Dellnitz, *On continuation methods for the numerical treatment of multi-objective optimization problems*, in Practical Approaches to Multi-Objective Optimization, Dagstuhl Seminar Proceedings 04461, J. Branke, K. Deb, K. Miettinen, and R. E. Steuer, eds., Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2005.

[52] Y. D. Sergeyev and D. E. Kvasov, *Global search based on efficient diagonal partitions and a set of Lipschitz constants*, SIAM J. Optim., 16 (2006), pp. 910–937.

[53] J. R. Shewchuk, *Triangle: Engineering a 2d quality mesh generator and Delaunay triangulator*, in Applied Computational Geometry: Towards Geometric Engineering, Lecture Notes in Comput. Sci. 1148, M. C. Lin and D. Manocha, eds., Springer, Berlin, 1996, pp. 203–222.

[54] J. R. Shewchuk, *Delaunay refinement algorithms for triangular mesh generation*, Comput. Geom., 22 (2002), pp. 21–74.

[55] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc., 73 (1967), pp. 747–817.

[56] S. Smale, *What is global analysis?*, Amer. Math. Monthly, 76 (1969), pp. 4–9.

[57] S. Smale, *Global analysis and economics.* I. *Pareto optimum and a generalization of Morse theory*, in Dynamical Systems (Proc. Sympos., Univ. Bahia, Salvador, 1971), Academic Press, New York, 1973, pp. 531–544.

[58] S. Smale, *Optimizing several functions*, in Manifolds–Tokyo 1973 (Proc. Internat. Conf., Tokyo, 1973), Univ. Tokyo Press, Tokyo, 1975, pp. 69–75.

[59] I. M. Sobol' and S. Kucherenko, *Derivative based global sensitivity measures and their link with global sensitivity indices*, Math. Comput. Simulation, 79 (2009), pp. 3009–3017.

[60] R. Thom, *Les singularités des applications différentiables*, Ann. Inst. Fourier (Grenoble), 6 (1956), pp. 43–87.

[61] R. Thom, *Généralisation de la théorie de Morse aux variétés feuilletées*, Ann. Inst. Fourier (Grenoble), 14 (1964), pp. 173–189.

[62] R. Thom, *Stabilité structurelle et morphogénèse: essai d'une théorie générale des modèles*, Benjamin, New York, 1971.

[63] R. Thom, *Structural stability, catastrophe theory, and applied mathematics: The John von Neumann lecture,* 1976, SIAM Rev., 19 (1977), pp. 189–201.

[64] S. V. Utyuzhnikov, P. Fantini, and M. D. Guenov, *A method for generating a well-distributed Pareto set in nonlinear multiobjective optimization*, J. Comput. Appl. Math., 223 (2009), pp. 820–841.

[65] Y. H. Wan, *Morse theory for two functions*, Topology, 14 (1975), pp. 217–228.

[66] Y. H. Wan, *On local Pareto optima*, J. Math. Econom., 2 (1975), pp. 35–42.

[67] Y. H. Wan, *On the structure and stability of local Pareto optima in a pure exchange economy*, J. Math. Econom., 5 (1978), pp. 255–274.

[68] D. F. Watson, *Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes*, Comput. J., 24 (1981), pp. 167–172.

[69] C. F. J. Wu and M. S. Hamada, *Experiments: Planning, Analysis, and Parameter Design Optimization*, John Wiley, New York, 2000.

[70] Y. Wu, L. Ozdamar, and A. Kumar, *Triopt: A triangulation-based partitioning algorithm for global optimization*, J. Comput. Appl. Math., 177 (2005), pp. 35–53.