Engineering Optimization
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713641621

Local Pareto approximation for multi-objective optimization

Sergei Utyuzhnikov [a]; Jeremy Maginot [b]; Marin Guenov [b]
[a] School of Mechanical, Aerospace & Civil Engineering, University of Manchester, UK [b] Department of PPAE, School of Engineering, Cranfield University, UK

First Published:September2008

To cite this Article Utyuzhnikov, Sergei, Maginot, Jeremy and Guenov, Marin(2008)'Local Pareto approximation for multi-objective optimization',Engineering Optimization,40:9,821 — 847
To link to this Article: DOI: 10.1080/03052150802086714
URL: http://dx.doi.org/10.1080/03052150802086714

Engineering Optimization
Vol. 40, No. 9, September 2008, 821–847

Local Pareto approximation for multi-objective optimization

Sergei Utyuzhnikov[a], Jeremy Maginot[b] and Marin Guenov[b]*

[a]*School of Mechanical, Aerospace & Civil Engineering, University of Manchester, UK;* [b]*Department of PPAE, School of Engineering, Cranfield University, UK*

(*Received 21 June 2007; final version received 7 January 2008*)

The design process of complex systems often resorts to solving an optimization problem, which involves different disciplines and where all design criteria have to be optimized simultaneously. Mathematically, this problem can be reduced to a vector optimization problem. The solution of this problem is not unique and is represented by a Pareto surface in the objective function space. Once a Pareto solution is obtained, it may be very useful for the decision-maker to be able to perform a quick local approximation in the vicinity of this Pareto solution for sensitivity analysis. In this article, new linear and quadratic local approximations of the Pareto surface are derived and compared to existing formulas. The case of non-differentiable Pareto points (solutions) in the objective space is also analysed. The concept of a local quick Pareto analyser based on local sensitivity analysis is proposed. This Pareto analysis provides a quantitative insight into the relation between variations of the different objective functions under constraints. A few examples are considered to illustrate the concept and its advantages.

Keywords: Pareto surface; approximation; multi-objective optimization; decision-making; sensitivity analysis

1. Introduction

In the process of designing complex systems, contributions and interactions of multiple disciplines are taken into account to achieve a consistent design. In a real industrial design procedure the problem is made more complicated due to the fact that the decision-maker (DM) has to consider many different and often conflicting criteria. In fact, during the optimization process, the DM often has to make compromises and look for trade-off solutions rather than a global optimum, which usually does not exist.

Multi-disciplinary design optimization (MDO) has become a field of comprehensive study for a few decades since the recent development of computer power has started to satisfy some minimal requirements to tackle this problem. MDO embodies a set of methodologies that provide a means of coordinating efforts and performing the optimization of a complex system. Two fundamental issues associated with the MDO concept are the complexity of the problem (large number of variables, constraints and objectives) and the difficulty of exploring the whole design space. Thus, in practice the DM would benefit from the opportunity to obtain additional information about the model without running it extensively.

*Corresponding author. Email:m.d.guenov@cranfield.ac.uk

ISSN 0305-215X print/ISSN 1029-0273 online

DOI: 10.1080/03052150802086714
http://www.informaworld.com

Finding a solution to an MDO problem implies solving a vector optimization problem under constraints. In general, the solution of such a problem is not unique. In this respect the existence of feasible solutions, *i.e.* solutions that satisfy all constraints but cannot be optimized further without compromising at least one of the other criteria, leads to the Pareto optimal concept (see *e.g.* Miettinen 1999). Each Pareto point is a solution of the multi-objective optimization problem. The DM often selects the final design solution among an available Pareto set based on additional requirements that are not taken into account in the mathematical formulation of the vector optimization problem.

In spite of the existence of many numerical methods for non-linear vector optimization, there are few methods suitable for real-design industrial applications. In many applications, each design cycle includes time-consuming and expensive computations for each discipline involved. In preliminary design it is important to get maximum information on a possible solution in a reasonably short time. Thus, it is very desirable for the DM to be able to approximate the Pareto surface in the vicinity of a current Pareto solution and to provide the sensitivity information (Hernandez 1995, Tappeta *et al.* 2000). It would also be very useful for the DM to be able to carry out a local approximation of other optimal solutions relatively quickly without additional full-run calculations. Such an approach is based on a local sensitivity analysis (SA) providing the relation between variations of different objective functions under constraints.

Currently, only a few articles are devoted to the SA of Pareto solutions in MDO. Some deal with the stability of the set of Pareto solutions (Tammer 1994, Fadel *et al.* 2002) while others concentrate on deriving methods to obtain local approximations of the Pareto surface (Hernandez 1995, Tappeta and Renaud 1999, Tappeta *et al.* 2000, Zhang 2003a, b). These methods are based on a local linearization of the solution set. The linear approximation geometrically results in finding the hyperplane tangent to the Pareto surface at some Pareto point using the gradient projection method (GPM) (Rosen 1958). Hernandez (1995) uses an approximate evaluation of the local Hessian for a quadratic approximation and relies on the assumption that other Pareto solutions are available in the vicinity of the Pareto solution under study. Methods are being developed to obtain evenly distributed Pareto solutions (Guenov *et al.* 2005, Utyuzhnikov *et al.* 2005). However, the assumption on the availability of several Pareto solutions may not be always valid. One of the most difficult problems in the SA of Pareto solutions is related to possible lack of smoothness of the Pareto surface in the objective space (Zhang 2003a, b).

In this context the aim of this work is to develop a method for local Pareto analysis and approximation of the Pareto surface. The objective is to derive linear and quadratic local approximations of the Pareto in the general case, including identification of non-differentiable Pareto points.

The rest of the article is organized as follows. Section 2 briefly outlines the formulation of the multi-objective optimization problem. The derivation of the Pareto approximation is presented in Section 3, where it is also shown that the existing formulas (Tappeta and Renaud 1999) are only valid under particular conditions. Some practical considerations on the use of these approximations for quick local Pareto solution analysis are given in Section 4. Section 5 gives an analysis of non-differentiable Pareto points. Several examples illustrating the concept and its advantages are presented in Section 6 and in Section 7 the method is tested on a typical engineering multi-objective optimization problem. Finally, conclusions and future work are outlined in Section 8.

2. The multi-objective optimization problem

It is assumed that an optimization problem is described in terms of a design variable vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)^T$ in the design space $\mathbf{X} \subset \mathrm{R}^N$. A function $\mathbf{f} \in R^M$ evaluates the quality of a solution by assigning it to an objective vector $\mathbf{y} = (y_1, y_2, \ldots, y_M)^T$ where $y_i = f_i(\mathbf{x})$,

f_i: $\mathrm{R}^N \rightarrow R^1$, $i = 1, 2, \ldots, M$ in the objective space $\mathbf{Y} \subset \mathrm{R}^M$. Thus, $\mathbf{X}$ is mapped onto $\mathbf{Y}$ by $\mathbf{f}$: $\mathbf{X}| \rightarrow \mathbf{Y}$. A multi-objective optimization problem may be formulated in the following form:

$$\text{Minimize } [\mathbf{y}(\mathbf{x})]. \tag{1}$$

Subject to L inequality constraints:

$$g_i(\mathbf{x}) \leq 0 \quad i = 1, \ldots, L \tag{2}$$

which may also include equality constraints.

The feasible design space $\mathbf{X}^*$ is defined as the set $\{\mathbf{x}|g_i(\mathbf{x}) \leq 0,\ j = 1, 2, \ldots, L\}$. A feasible design point is a point that does not violate any constraint. The feasible criterion (objective) space $\mathbf{Y}^*$ is defined as the set $\{\mathbf{Y}(\mathbf{x})|\mathbf{x} \in \mathbf{X}^*\}$. A design vector $\mathbf{a}$ ($\mathbf{a} \in \mathbf{X}^*$) is called a Pareto optimum if, and only if, it does not exist any $\mathbf{b} \in \mathbf{X}^*$ such that $y_i(\mathbf{b}) \leq y_i(\mathbf{a})$, $i = 1, \ldots, M$ and there exist $1 \leq j \leq M$ such that $y_j(\mathbf{b}) < y_j(\mathbf{a})$.

Henceforth, it is assumed that all vectors are considered in the appropriate Euclidean spaces.

3. Pareto approximation

In this section, it is assumed that the Pareto surface is smooth in the vicinity of the Pareto solution under study. A local approximation of the Pareto surface would allow the DM to quickly obtain both qualitative and quantitative information on the trade-off between different local Pareto optimal solutions.

A constraint is said to be active at a Pareto point $\mathbf{x}^*$ of the design space $\mathbf{X}$ if a strict equality holds at this point (Tappeta and Renaud 1999, Tappeta *et al.* 2000). In this section, it is assumed that constraints that are active at a particular Pareto point remain active in its vicinity. Thus, the sensitivity predicted at the given Pareto point is valid until the set of active constraints remains unchanged (Hernandez 1995, Tappeta and Renaud 1999). Without loss of generality, let us assume that the first I constraints are active and the first Q of those correspond to inequalities ($Q \leq I \leq L$).

Note the set of active constraints (Equation (2)) as $\mathbf{G} \in \mathrm{R}^I$. At the given point $\mathbf{x}^*$ of the design feasible space $\mathbf{X}^*$, this means that:

$$\mathbf{G}(\mathbf{x}^*) = \mathbf{0}. \tag{3}$$

Assume that $\mathbf{G} \in C^1(\mathrm{R}^I)$, then locally the constraints can be written in the linear form:

$$\mathbf{J}(\mathbf{x} - \mathbf{x}^*) = \mathbf{0}, \tag{4}$$

where $\mathbf{J}$ is the Jacobian of the active constraints set at $\mathbf{x}^*$: $\mathbf{J} = \nabla\mathbf{G}$. If all gradients of the active constraints are linearly independent at a point, then this point is called a regular point (Vincent and Grantham 1981). Thus, we say that a point $\mathbf{x}^* \in \mathbf{X}^*$ is regular if rank $(\mathbf{J}) = I$.

Further assume that in the objective space $\mathbf{Y}$ the Pareto surface is given by:

$$S(\mathbf{y}) = 0 \tag{5}$$

and at the Pareto point $\mathbf{y}^* = \mathbf{f}(\mathbf{x}^*)$, function $S \in C^2(\mathrm{R}^1)$.

The values of the gradient of any differentiable function F at point $\mathbf{x}^*$ under constraints are defined by the reduced gradient formula (Fletcher 1989):

$$\nabla F_{|S_l} = \mathbf{P}\nabla F, \tag{6}$$

where S_l is the hyperplane tangent to $\mathbf{X}^*$ the design space:

$$S_l = \{\mathbf{x}|\mathbf{J}(\mathbf{x} - \mathbf{x}^*) = 0\} \tag{7}$$

and $\mathbf{P}$ is projection matrix onto this hyperplane:

$$\mathbf{P} = \mathbf{I} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}. \tag{8}$$

Directional derivatives corresponding to Equation (6) in the objective space are represented by:

$$\frac{dF}{df_i} = \frac{dF}{d\mathbf{x}}\Big|_{S_l}\frac{d\mathbf{x}}{df_i}. \tag{9}$$

The first element of the product corresponds to the reduced gradient and the second element $d\mathbf{x}$ represents the infinitesimal change in the design vector $\mathbf{x}$ to accommodate the infinitesimal shift $d\mathbf{f}$ of the objective vector tangent to the Pareto surface.

The last derivative in Equation (9) can be represented via the gradients obtained in the design space $\mathbf{X}$. Assume that matrix $\mathbf{P}\nabla\mathbf{f}$, $(\mathbf{f} = (f_1, f_2, \ldots, f_M)^T$ has $n_f < M$ linearly independent columns. It is to be noted that $n_f \neq M$. Indeed, in view of Equation (5):

$$\sum_{i=1}^{M}\frac{\partial S}{\partial f_{i|S}}\,df_i = 0 \tag{10}$$

and it is easy to see that to move locally on the Pareto surface, $df_i (i = 1, \ldots, M)$ cannot be chosen independently. Given that:

$$d\mathbf{f} = (\mathbf{P}\nabla\mathbf{f})^T d\mathbf{x}, \tag{11}$$

one can prove that only $n_f < M$ objective functions f_i are locally linearly independent.

Without loss of generality, assume further that the first n_f components of $\mathbf{P}\nabla\mathbf{f}$ are linearly independent and represented by $\mathbf{P}\nabla\tilde{\mathbf{f}}$. Thus,

$$d\tilde{\mathbf{f}} = (\mathbf{P}\nabla\tilde{\mathbf{f}})^T d\mathbf{x}, \tag{12}$$

where matrix has all $\mathbf{P}\nabla\tilde{\mathbf{f}} \equiv (\mathbf{P}\nabla f_1, \ldots, \mathbf{P}\nabla f_{n_f})$ the columns linearly independent. Now, write $d\mathbf{x}$ in the following form:

$$d\mathbf{x} = \mathbf{A}d\tilde{\mathbf{f}}. \tag{13}$$

Then, having multiplied both sides of Equation (13) by $(\mathbf{P}\nabla\tilde{\mathbf{f}})^T$ and taking into account Equation (12) obtain:

$$(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{A}d\tilde{\mathbf{f}} = d\tilde{\mathbf{f}}. \tag{14}$$

Hence,

$$\mathbf{A} = \mathbf{P}\nabla\tilde{\mathbf{f}}[(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{P}\nabla\tilde{\mathbf{f}}]^{-1}. \tag{15}$$

Thus, matrix $\mathbf{A}$ is the right-hand generalized inverse matrix to $(\mathbf{P}\nabla\tilde{\mathbf{f}})^T$. It is possible to prove that the inverse matrix $[(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{P}\nabla\tilde{\mathbf{f}}]^{-1}$ is always non-singular because all the vectors $\mathbf{P}\nabla f_i$, $(i = 1, \ldots, n_f)$ are linearly independent. From the definition of matrix $\mathbf{A}$ it follows that $(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{A} = \mathbf{I}$ and $\mathbf{A}_i^T\mathbf{P}\nabla\tilde{f}_j = \delta_{ij}$, where $\mathbf{I}$ is the unit matrix and δ_{ij} is the Kronecker delta.

Hence, the system of vectors $\{\mathbf{A}_i\}(i = 1, \ldots, n_f)$ creates the basis reciprocal to the basis of vectors $\{\mathbf{P}\nabla\tilde{\mathbf{f}}_j\}(j = 1, \ldots, n_f)$. From Equation (15) it follows that $\mathbf{PA} = \mathbf{A}$ and $d\mathbf{x}$ in Equation (13) belongs to the tangent plane S_l at the Pareto point. Thus,

$$\frac{d\mathbf{x}}{d\tilde{\mathbf{f}}} = \mathbf{A} \equiv \mathbf{P}\nabla\tilde{\mathbf{f}}[(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{P}\nabla\tilde{\mathbf{f}}]^{-1} \tag{16}$$

and for any $i \leq n_f$

$$\frac{d\mathbf{x}}{df_i} = \mathbf{A}_i, \tag{17}$$

where $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{n_f})$ Then, from Equations (6), (9) and (16), it follows that for any $i \leq n_f$

$$\frac{dF}{df_i} = (\mathbf{P}\nabla F)^T\mathbf{A}_i = \mathbf{A}_i^T\mathbf{P}\nabla F = \mathbf{A}_i^T\nabla F. \tag{18}$$

If $F = f_j$ $(n_f < j < M)$, then we obtain the first-order derivative of objective f_j with respect to objective f_i on the Pareto surface. Thus,

$$\frac{df_j}{df_i} = \mathbf{A}_i^T\nabla f_j \quad (0 \leq i \leq n_f, n_f < j \leq M). \tag{19}$$

It is important to note that this formula coincides with the formula of sensitivity of objective f_j with respect to objective f_i along the greatest feasible descent direction for objective f_i,

$$\frac{df_j}{df_i} = \frac{(\mathbf{P}\nabla f_j, \mathbf{P}\nabla f_i)}{(\mathbf{P}\nabla f_i, \mathbf{P}\nabla f_i)} \equiv \frac{(\nabla f_j, \mathbf{P}\nabla f_i)}{(\nabla f_i, \mathbf{P}\nabla f_i)}, \tag{20}$$

(see Tappeta and Renaud 1999, Tappeta *et al.* 2000) if and only if either $n_f = 1$ or the vectors $\mathbf{P}\nabla\tilde{\mathbf{f}}$ create an orthogonal basis. In this case, the matrix $(\mathbf{P}\nabla\tilde{\mathbf{f}})^T\mathbf{P}\nabla\tilde{\mathbf{f}}$ is diagonal. In particular, in the case of a two-objective optimization these formulas always coincide since $n_f = 1$.

On the Pareto surface in the objective space, the operator of the first derivative can be defined by:

$$\frac{d}{df_i} = \mathbf{A}_i^T\nabla. \tag{21}$$

By applying this operator to the first-order derivative found previously, one can obtain the reduced Hessian as follows:

$$\frac{d^2F}{df_i df_j} = \mathbf{A}_i^T\nabla(\mathbf{A}_j^T\nabla F) \approx \mathbf{A}_i^T\nabla^2 F\mathbf{A}_j \quad (0 \leq i, j \leq n_f). \tag{22}$$

Thus, the Pareto surface can be locally represented as a linear hyperplane:

$$\sum_{i=1}^{n_f}\frac{dS}{df_i}\Delta f_i = 0 \tag{23}$$

or a quadratic surface:

$$\sum_{i=1}^{n_f}\frac{dS}{df_i}\Delta f_i + \frac{1}{2}\sum_{j,k=1}^{n_f}\frac{d^2S}{df_j df_k}\Delta f_j\Delta f_k = 0, \tag{24}$$

where $\Delta\mathbf{f} = \mathbf{f} - \mathbf{f}(\mathbf{x}^*)$. Approximations (23) and (24) can be rewritten with respect to the trade-off relations between the objective functions as follows:

$$f_p = f_p^* + \sum_{i=1}^{n_f} \frac{df_p}{df_i}\Delta f_i \quad (p = n_f + 1, \ldots, M) \tag{25}$$

and

$$f_p = f_p^* + \sum_{i=1}^{n_f} \frac{df_p}{df_i}\Delta f_i + \frac{1}{2}\sum_{j,k=1}^{n_f} H_{jk}^{(p)}\Delta f_j \Delta f_k \quad (p = n_f + 1, \ldots, M), \tag{26}$$

where

$$H_{jk}^{(p)} = \frac{d^2 f_p}{df_j df_k}.$$

A quadratic approximation where the reduced Hessian matrix H_{ij} is evaluated with a least-squared minimization using the Pareto set generated around the original Pareto point has been derived previously (Tappeta and Renaud 1999). However, in preliminary design such an evaluation can be unsuitable because it would require generating more Pareto points in the vicinity of the point under study. Instead, the local determination of the reduced Hessian using Equation (22) is more accurate and is entirely based on the value of the objective and constraint gradients with respect to the independent design variables. It is important to note that, in contrast to the approximations developed earlier (Tappeta and Renaud 1999), the ones derived in this article precisely correspond to the first three terms of the Taylor expansion in the general case.

4. Local quick Pareto analysis

The first-order derivatives df_p/df_i provide us with first-order sensitivity of an objective f_p along the feasible descent direction of an objective f_i when all other objectives are kept constant. It is to be noted here that all the derivatives df_p/df_i are non-positive for all $i = 1, \ldots, n_f$. Otherwise, two objectives could be locally improved, which would contradict the Pareto-solution assumption. The local approximations of the Pareto surface can be used to study the local adaptability of a Pareto solution. As in a real-life problem it can be time and computationally expensive to obtain a single Pareto solution; local approximate solution; around a Pareto point can be obtained using either Equation (25) or (26).

As discussed above, in the preliminary design it can be very beneficial to the DM if s/he is able to perform quick SA of the solution obtained. Using the local approximation of the Pareto surface the DM has the opportunity to perform the SA without additional full-run computations. It is also easy to obtain information on trade-offs between different objectives. It has to be emphasized that in the multi-objective case, the change of one objective does not fully determine changes in the others. If the DM freezes all objectives except two or three, it is then possible to obtain information that is useful for understanding the trade-off between the selected objectives. The analysis of solutions around a Pareto point allows the DM to correct locally the solution with respect to additional preferences. Furthermore, the DM is able to analyse possible violations of the constraints as part of the trade-off analysis. In design practice, the opportunity for further improvement of some objectives at the expense of local degradation of some other objectives can also be important. Representations given by Equations (25) and (26) are only local approximations and there is a question on the range of $\Delta\mathbf{x}$ where the approximation is valid. In the framework of a local analysis, giving a strict answer to this question is not possible. However, the reliable range of the variation of $\Delta\mathbf{x}$ can be evaluated qualitatively by comparing the solutions obtained

by linear and quadratic approximations. It is reasonable to expect that the approximations are suitable as long as the difference between the two approximations is small.

In the SA, due to a perturbation $\delta\mathbf{f}$ and appropriate displacement $\delta\mathbf{x}$, some constraints, which are inactive at point $\mathbf{x}^*$, can become either violated or active. The exact verification of the constraints validation may be time consuming. To investigate this problem, it was suggested that a local linear approximation of the inactive constraints along a feasible direction should be obtained to study the degree of constraint violation (Zhang 2003b). The approach developed in Section 3 can be used to obtain the appropriate linear and quadratic approximations for non-active inequality constraints, including their quadratic approximation, in any direction:

$$g_k(\mathbf{x}) = g_k(\mathbf{x}^*) + \sum_{i=1}^{n_f} \frac{dg_k}{df_i}\Delta f_i, \quad (I < k \leq L), \tag{27}$$

$$g_k(\mathbf{x}) = g_k(\mathbf{x}^*) + \sum_{i=1}^{n_f} \frac{dg_k}{df_i}\Delta f_i + \frac{1}{2}\sum_{i,j=1}^{n_f} \frac{d^2 g_k}{df_i df_j}\Delta f_i \Delta f_j. \tag{28}$$

These approximations can be used to verify that inactive constraints remain inactive at a new approximate Pareto point.

In general, there is no guarantee that the Pareto surface is smooth. If the designed Pareto solution appears at a point of lack of smoothness, the approximations derived above are not formally valid. In such a case, a substantial discrepancy can appear between the first- and second-order approximations in the nearest vicinity of the point. Another opportunity to reveal such non-smooth points is based on an analysis of the Lagrange multipliers described in the next section.

5. Pareto front analysis at non-differentiable Pareto points

Assume that an aggregate function Φ such that:

$$\Phi = \Phi(\mathbf{f}), \quad \frac{\partial \Phi}{\partial f_i}(\mathbf{x}_p) \geq 0, \quad i = 1, 2, \ldots, M \tag{29}$$

reaches the minimum at a Pareto point $\mathbf{x}_p$. In particular, such an aggregate function satisfying (29) is used in the physical programming method (Messac 1996, Utyuzhnikov *et al.* 2005). Then, from the Karush–Kuhn–Tucker (KKT) optimality condition (Fletcher 1989) and Equation (8) it follows that:

$$\mathbf{S}_\Phi \equiv -\nabla\Phi - \mathbf{J}^T\boldsymbol{\lambda} = \mathbf{P}(-\nabla\Phi) = 0, \tag{30}$$

where the vector $\boldsymbol{\lambda}$ corresponds to the Lagrange multipliers:

$$\begin{aligned} \boldsymbol{\lambda} &= -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Phi, \\ \lambda_j &\geq 0 (j = 1, \ldots, I). \end{aligned} \tag{31}$$

It is to be noted here that all analysis in this section is limited by the assumption of linear local constraints.

In the objective space $\mathbf{Y}$, the normal to the Pareto surface is determined by the vector:

$$\mathbf{n}_f = \frac{\partial \Phi}{\partial \mathbf{f}} \equiv \nabla_f \Phi. \tag{32}$$

Introduce a test function Ψ satisfying conditions:

$$\frac{\partial \Psi}{\partial f_i}(\mathbf{x}_p) \geq 0 \quad (i = 1, 2, \ldots, M) \tag{33}$$

in the vicinity of the Pareto point $\mathbf{x}_p$. This condition means that $\nabla\Psi \in C$, where C is a polyhedral cone defined by:

$$C = \{\mathbf{y} \in \mathrm{R}^N | \mathbf{y} = \sum_{i=1}^{M} a_i \nabla f_i,\ a_i \geq 0\}. \tag{34}$$

It is assumed that the Pareto point under study corresponds to a regular point in the feasible space. This condition implies that at this point, the gradients of all active constraints are linearly independent (Vincent and Grantham 1981).

From Equation (6) it follows that the descent direction $\mathbf{S}_\Psi$ of the function Ψ is determined by:

$$\mathbf{S}_\Psi = -\nabla\Psi - \mathbf{J}^T\boldsymbol{\mu} = \mathbf{P}(-\nabla\Psi), \tag{35}$$

where $\boldsymbol{\mu} \in \mathrm{R}^I$ is defined by:

$$\boldsymbol{\mu} = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Psi. \tag{36}$$

The vector $\mathbf{S}_\Psi$ is the orthogonal projection of $-\nabla\Psi$ onto the tangent plane to the Pareto surface in the design space. The vector $\mathbf{S}_\Psi$ is a null-vector if and only if $\nabla\Psi$ belongs to the linear manifold V_I defined by:

$$V_I = \left\{ \mathbf{y} \in R^N | \, \mathbf{y} = \sum_{i=1}^{I} v_i \nabla g_i \right\}. \tag{37}$$

If $\mathbf{S}_\Psi$ is not a null-vector, then the mapping of $\mathbf{S}_\Psi$ in the objective space gives a tangent direction to the Pareto surface. Indeed, multiplying $\mathbf{S}_\Phi$ by $\mathbf{S}_\Psi$ and taking into account the orthogonality condition $\mathbf{J}\mathbf{S}_\Psi = 0$, yields:

$$\sum_{i=1}^{M} \frac{\partial \Phi}{\partial f_i} \delta f_i = 0, \tag{38}$$

where $\delta f_i = \nabla f_i^T \mathbf{S}_\Psi \alpha$ is an infinitesimal variation of f_i along the direction $\mathbf{S}_\Psi$ and $\alpha \to 0$ is a scaling parameter. Therefore, if $\mathbf{S}_\Psi$ is not a null-vector, then it can be used to determine a tangent direction in the objective space and provide the SA along $\mathbf{S}_\Psi$:

$$\left.\frac{df_i}{df_j}\right|_{\mathbf{S}_\Psi} = \frac{df_i}{d\mathbf{S}_\Psi}\frac{d\mathbf{S}_\Psi}{df_j} = \frac{(\nabla f_i, \mathbf{S}_\Psi)}{(\nabla f_j, \mathbf{S}_\Psi)} < 0 \quad (i, j = 1, \ldots, M; i \neq j). \tag{39}$$

Assume now that $\mathbf{S}_\Psi$ is a null-vector. Then, $\nabla\Psi \in V_I$ and $\nabla\Psi$ is represented by a linear combination of the vectors normal to the hyperplanes of the active constraints:

$$\nabla\Psi = -\sum_{i=1}^{I} \mu_i \nabla g_i. \tag{40}$$

If, in Equation (40), $\mu_i \geq 0$ for $i = 1, \ldots, Q$, then there is no descent direction for the test function Ψ. This immediately follows from Farkas' lemma (Fletcher 1989). Indeed, the feasible directions

are defined by the set of vectors **e** that satisfy:

$$(\mathbf{e}, \nabla g_i) \leq 0 \text{ for } 1 \leq i \leq Q \quad \text{and} \quad (\mathbf{e}, \nabla g_i) = 0 \text{ for } Q+1 \leq i \leq I, \tag{41}$$

while a vector **e** which corresponds to a descent direction of the function Ψ, is such that:

$$(\mathbf{e}, \nabla \Psi) < 0. \tag{42}$$

From Farkas' lemma, the set of vectors **e** satisfying conditions (41) and (42) is empty if and only if in Equation (40) $\mu_i \geq 0\,(i = 1, \ldots, Q)$.

Now determine the set of Lagrange multipliers $\Omega_{I,Q}$ in such a way that the first Q Lagrange multipliers μ_i corresponding to the active constraints are positive:

$$\Omega_{I,Q} = \{\boldsymbol{\mu} \in \mathrm{R}^I | \mu_i \geq 0 \textit{ for } 1 \leq i \leq Q\}. \tag{43}$$

If some of the first Q multipliers are negative ($\boldsymbol{\mu} \notin \Omega_{I,Q}$), then the appropriate inequality constraints become inactive in the direction $\mathbf{S}_\Psi$ and the Pareto solution might correspond to a non-differentiable point of the Pareto frontier. In order to identify if the Pareto point is non-differentiable, it has been proposed to remove the constraints corresponding to negative components of $\boldsymbol{\mu}$ and repeat the analysis reconsidering the active constraints (Zhang 2003a, b). All the analysis is performed by representing Ψ by an objective function. The principal difference with respect to previous attempts (Zhang 2003a, b), is that the SA described below is done in the case of a general test function Ψ.

Now consider the requirements for the test function Ψ. It is to be noted that this question is not addressed by Zhang (2003a, b). The analysis is based on the consideration of the constraint and polar cones at a Pareto point.

The constraint cone K is defined as follows (Vincent and Grantham 1981):

$$K = \{\mathbf{y} \in \mathrm{R}^N | \mathbf{y} = -\sum_{i=1}^{l} \mu_i \nabla g_i, \boldsymbol{\mu} \in \Omega_{I,Q}\} \tag{44}$$

Along with the constraint cone K, the positive polar cone K^* is defined as:

$$K^* = \left\{\mathbf{z} \in \mathrm{R}^N | \mathbf{z}^T \mathbf{y} \geq 0 \ \ \forall\, \mathrm{y} \in K\right\}. \tag{45}$$

For any regular point of the feasible space, the cone T tangent to the feasible design space coincides with the cone $K^* : K^* = T$ (Vincent and Grantham 1981). It is to be noted that if the function Ψ reaches a minimum at point $\mathbf{x}_p$ then for any vector **e** satisfying $\mathbf{e} \in T$ the following condition must be valid:

$$\mathbf{e}^T \nabla \Psi \geq 0, \tag{46}$$

otherwise there is a feasible direction of minimization for Ψ. Inequality (46) also means that $\nabla \Psi \in T^* = K$, where T^* is the positive polar cone to the cone T. The identity $T^* = K$ follows from the Polar theorem: $(K^*)^* = K$ (Vincent and Grantham 1981). Note that since the function Φ reaches a minimum at point $\mathbf{x}_p$, then $\nabla \Phi \in K$.

If $\nabla \Psi \notin K$ then $\boldsymbol{\mu} \notin \Omega_{I,Q}$ and there is a direction along which the function Ψ can be further diminished. Assume that there is only one Lagrange multiplier μ_q in Equation (36): $\mu_q < 0$, $(q \leq Q)$. By removing the qth active constraint from $\mathbf{J}$, we obtain the reduced matrix $\mathbf{J}_r$. The new

descent vector $\tilde{\mathbf{S}}_\Psi$ is defined by the set of remaining active constraints as follows:

$$\tilde{\mathbf{S}}_\Psi = -\nabla\Psi - \mathbf{J}_r^T \boldsymbol{\mu}_r, \tag{47}$$

where

$$\boldsymbol{\mu}_r = -(\mathbf{J}_r \mathbf{J}_r^T)^{-1} \mathbf{J}_r \nabla\Psi. \tag{48}$$

Vector $\tilde{\mathbf{S}}_\Psi$ defines a descent direction along which the qth constraint becomes inactive. Then, make sure the new direction is feasible. Subtracting Equation (35) from (47) yields:

$$\tilde{\mathbf{S}}_\Psi = \mathbf{J}^T \boldsymbol{\mu} - \mathbf{J}_r^T \boldsymbol{\mu}_r = \mu_q \nabla g_q + \mathbf{J}_r^T (\tilde{\boldsymbol{\mu}}_r - \boldsymbol{\mu}_r), \tag{49}$$

where vector $\tilde{\boldsymbol{\mu}}_r \in R^{I-1}$ is obtained from $\boldsymbol{\mu}$ by dropping μ_q. Note that $\tilde{\mathbf{S}}_\Psi$ is not a null-vector, otherwise (from Equation (49)) there would exist a non-null-vector $\boldsymbol{\beta}$ such that $\mathbf{J}^T\boldsymbol{\beta} = \mathbf{0}$, which would contradict the assumption of local linear independence of the active constraints. Multiplying Equation (49) by itself and taking into account the orthogonality condition $\mathbf{J}_r \tilde{\mathbf{S}}_\Psi = 0$:

$$\left|\tilde{\mathbf{S}}_\Psi\right|^2 = \mu_q \tilde{\mathbf{S}}_\Psi^T \nabla g_q > 0. \tag{50}$$

Hence, $\tilde{\mathbf{S}}_\Psi^T \nabla g_q < 0$ and the direction $\tilde{\mathbf{S}}_\Psi$ is feasible.

If several Lagrange multipliers in the vector $\boldsymbol{\mu}$ are negative, the set Q^- is defined such that $\mu_q < 0$ if $q \in Q^-$. Only one ∇g_q with $q \in Q^-$ needs to be removed from $\mathbf{J}$ so that the above analysis remains valid. A strategy has been proposed (Zhang 2003b) to remove the gradient of the constraint that satisfies:

$$\frac{\lambda_k}{\mu_k} = \max_{q \in Q^-} \frac{\lambda_q}{\mu_q} < 0 \tag{51}$$

to obtain the reduced matrix $\mathbf{J}_r$ for the descent vector $\tilde{\mathbf{S}}_\Psi$. Along the descent direction $\tilde{\mathbf{S}}|_\Psi$:

$$\delta\Psi = \sum_{i=1}^{M} \frac{\partial\Psi}{\partial f_i} \delta f_i < 0 \tag{52}$$

and another Pareto solution is achievable in the framework of the linear approximation because all the coefficients $\partial\Psi/\partial f_i$ in Equation (52) are non-negative.

To find the new normal to the Pareto surface in the $\mathbf{Y}$ space, multiply the null-vector $\mathbf{S}_\Psi$ from Equation (35) by the multiplier α_Ψ defined as

$$\alpha_\Psi = -\max_{q \in Q^-} \frac{\lambda_q}{\mu_q} = -\frac{\lambda_k}{\mu_k} > 0 \tag{53}$$

and add it to the null-vector $\mathbf{S}_\Phi$ from Equation (30). Then:

$$\nabla(\Phi + \alpha_\Psi \Psi) = -\mathbf{J}^T(\boldsymbol{\lambda} + \alpha_\Psi \boldsymbol{\mu}). \tag{54}$$

This equality can be seen as the KKT optimality condition for the new aggregate function $\Phi^* = \Phi + \alpha_\Psi \Psi$ as it can be shown that:

$$\nabla\Phi^* + \mathbf{J}^T \boldsymbol{\lambda}^* = 0 \tag{55}$$

with the Lagrange multipliers: $\boldsymbol{\lambda}^* = \boldsymbol{\lambda} + \alpha_\Psi \boldsymbol{\mu}$. Additionally, it is easy to verify the following inequalities:

$$\frac{\partial\Phi^*}{\partial f_j} \geq 0, \quad (j = 1, \ldots, M), \tag{56}$$

$$\lambda_i^* \geq 0, \quad (i = 1, \ldots, I) \tag{57}$$

Note that in the KKT condition (55), there exists $k \in Q^-$ such that $\lambda_k^* = 0$. It corresponds to the active constraint that becomes inactive. As a result, the normal to the Pareto surface might shift to its new value:

$$\mathbf{n}_f = \nabla_f (\Phi + \alpha_\Psi \Psi). \tag{58}$$

It is to be noted here that the change of the number of active constraints, Q, does not necessarily mean the Pareto point is not differentiable. More precisely, the non-differentiability corresponds to a shift of the normal to the Pareto surface (Zhang 2003b), not a change in the set of active constraints.

It is possible to show that the vector $\tilde{\mathbf{S}}_\Psi$ defines a direction tangent to the Pareto surface. Indeed, multiplying the null-vector $\mathbf{S}_\Phi$ from Equation (30) by $\tilde{\mathbf{S}}_\Psi$ gives:

$$\nabla \Phi^T \tilde{\mathbf{S}}_\Psi = -\lambda_k \nabla g_k^T \tilde{\mathbf{S}}_\Psi. \tag{59}$$

Then, after multiplication of $\mathbf{S}_\Psi$ from Equation (35) by $\tilde{\mathbf{S}}_\Psi$ we have:

$$\nabla g_k^T \tilde{\mathbf{S}}_\Psi = -\frac{1}{\mu_k} \nabla \Psi^T \tilde{\mathbf{S}}_\Psi, \tag{60}$$

$$\nabla \Phi^T \tilde{\mathbf{S}}_\Psi = \frac{\lambda_k}{\mu_k} \nabla \Psi^T \tilde{\mathbf{S}}_\Psi = -\alpha_\Psi \nabla \Psi^T \tilde{\mathbf{S}}_\Psi, \tag{61}$$

hence

$$\nabla (\Phi + \alpha_\Psi \Psi)^T \tilde{\mathbf{S}}_\Psi = \nabla (\Phi^*)^T \tilde{\mathbf{S}}_\Psi = \mathbf{0}. \tag{62}$$

In turn, if $\nabla \Psi \in K$, then $\nabla \Psi$ can be expressed as a linear combination of the gradients of active constraints (cf. Equation (40)) with $\boldsymbol{\mu} \in \Omega_{I,Q}$. It follows that for such a test function, the feasible descent direction $\mathbf{S}_\Psi$ is a null-vector. From the previous analysis, it also follows that the test function Ψ does not add any additional information to the SA.

Thus, to verify the differentiability of a Pareto point, the function Ψ must be chosen in such a way that (see Equations (34), (37), (44))

$$\nabla \Psi \in C \cap (V_I \backslash K). \tag{63}$$

It is worth noting that the equality constraints have an effect on the matrix $\mathbf{J}$ but do not influence the SA with regard to the detection of non-differentiable Pareto points. This follows from the fact that equality constraints are to be strictly satisfied, *i.e.* they cannot be inactive, and the sign of the appropriate Lagrange multipliers is undetermined. Note that, in contrast to the work of Zhang (2003a, b) where only a weighted-sum function is considered as the test (preference) function Ψ, a general test function is considered here. Therefore, the approach becomes applicable to many other algorithms and is not limited to the weighted-sum method, which can only be used for convex problems. In particular, the developed SA method can be implemented in the physical programming based method (Guenov *et al.* 2005, Utyuzhnikov *et al.* 2005), which guarantees a well-distributed representation of the entire Pareto frontier. The entire approach for the local Pareto study is summarized in Figure 1.

6. Analytical examples

The first two examples are related to the smooth local approximation of the Pareto surface. They illustrate the linear and quadratic approximation of the Pareto surface. The requirements to the test function Ψ derived in the previous section are considered in the next two examples.

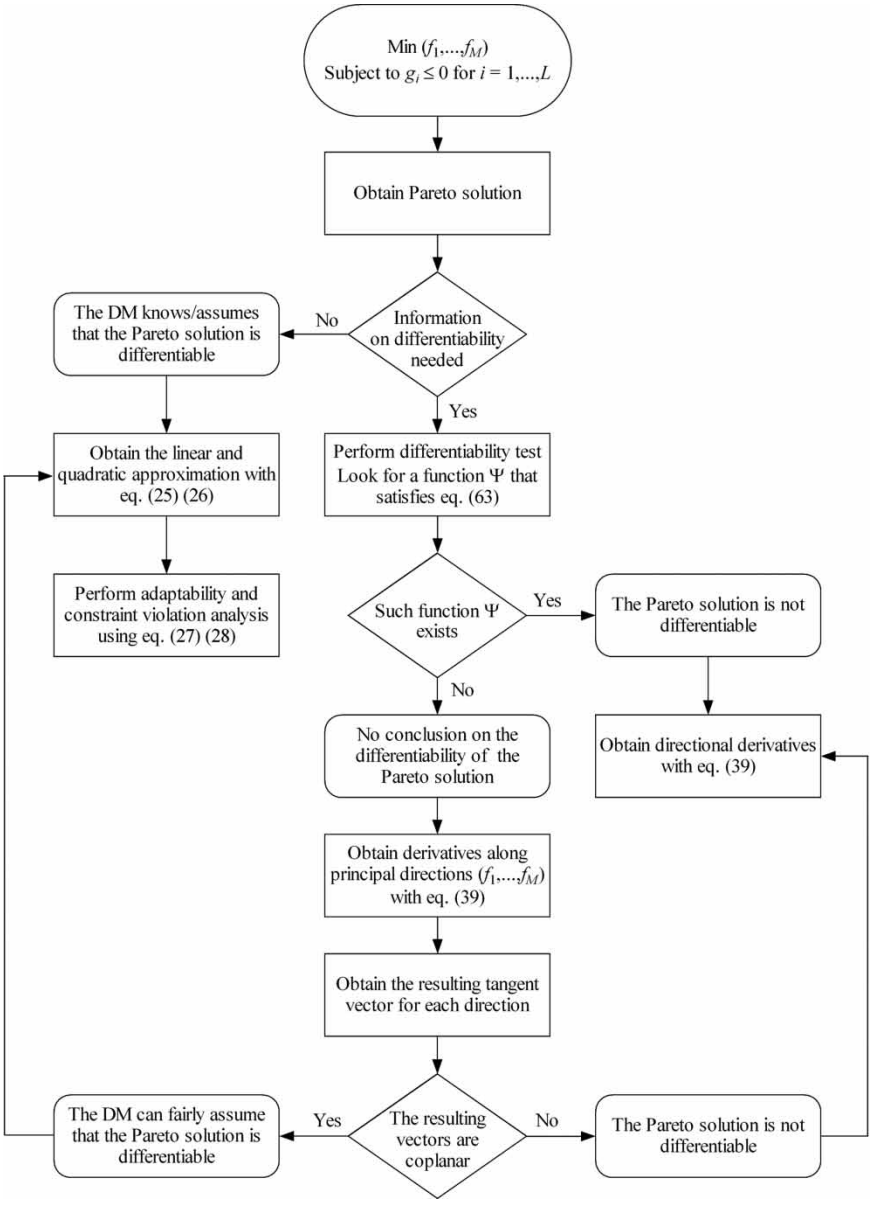

Figure 1. Local Pareto analysis flowchart.

6.1. *Local approximations—Comparison with existing formulas: Example 1*

To compare the approximations derived in this article with formulas already available (Tappeta and Renaud 1999, Tappeta *et al.* 2000), consider the following multi-objective problem:

$$\text{Minimize:} \quad \mathbf{f} = (x, y, z)^T \tag{64}$$

$$\text{Subject to:} \quad \begin{array}{c} g(\mathbf{x}) = 1 - x^2 - y^2 - z^2 \le 0, \\ x > 0, \\ y > 0, \\ z > 0. \end{array} \tag{65}$$

The design space and objective space coincide in this example. It is easy to see that the Pareto surface corresponds to the part of the unit sphere in the first quadrant and is represented by:

$$z = \sqrt{1 - x^2 - y^2}. \tag{66}$$

The first-order tangent derivatives are given by:

$$\begin{aligned} \frac{df_3}{df_1} &= \frac{\mathrm{d}z}{\mathrm{d}x} = \frac{-x}{\sqrt{1 - x^2 - y^2}} \\ \frac{df_3}{df_2} &= \frac{\mathrm{d}z}{\mathrm{d}y} = \frac{-y}{\sqrt{1 - x^2 - y^2}}. \end{aligned} \tag{67}$$

Derive the first-order approximation using the approach found in the literature and the method described in the present article. For any Pareto point, the matrices **J** and **P** are given by:

$$\mathbf{J} = \begin{bmatrix} -2x, & -2y, & -2z \end{bmatrix}, \tag{68}$$

$$\mathbf{P} = \begin{bmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{bmatrix}. \tag{69}$$

Using Equations (20) and (69) one can obtain the first-order derivatives as given previously (Tappeta and Renaud 1999, Tappeta *et al.* 2000)

$$\begin{aligned} \frac{df_3}{df_{1_{[3]}}} &= \frac{-xz}{y^2 + z^2} = \frac{(1 - x^2 - y^2)(-x)}{(1 - x^2)\sqrt{1 - x^2 - y^2}}, \\ \frac{df_3}{df_{2_{[3]}}} &= \frac{-yz}{x^2 + z^2} = \frac{(1 - x^2 - y^2)(-y)}{(1 - y^2)\sqrt{1 - x^2 - y^2}}. \end{aligned} \tag{70}$$

Note that Equations (70) are different from the exact analytical first-order derivatives given by Equations (67). They result in the approximation given in Figure 2. According to the method developed in this article:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\dfrac{x}{z} & -\dfrac{y}{z} \end{bmatrix}. \tag{71}$$

One can easily ensure that by using Equation (19) exact first-order derivatives (Equations (67)) are obtained. It is also possible to verify that the exact second-order derivatives can be obtained using Equation (22). The linear and quadratic approximations given by Equations (25) and (26) are shown in Figures 3 and 4, respectively. The relative error in predicting objective f_3 is shown in Figure 5. It is to be noted that only in the case of two design variables ($N = 2$) the previous linear approximation (Tappeta and Renaud 1999, Tappeta *et al.* 2000) provides the tangent plane (line) to the Pareto surface.

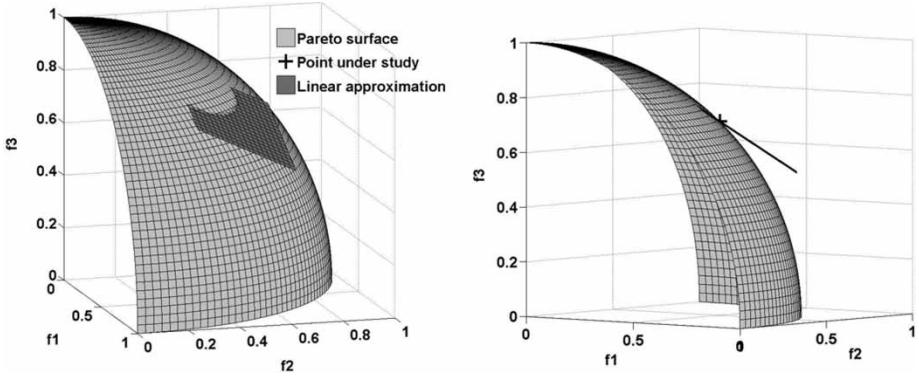

Figure 2. Example 1: linear approximation from literature (Equation (20)).

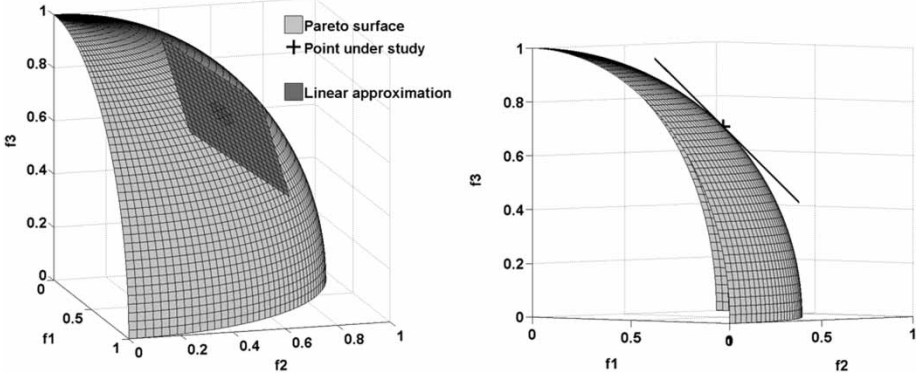

Figure 3. Example 1: new linear approximation (Equation (25)).

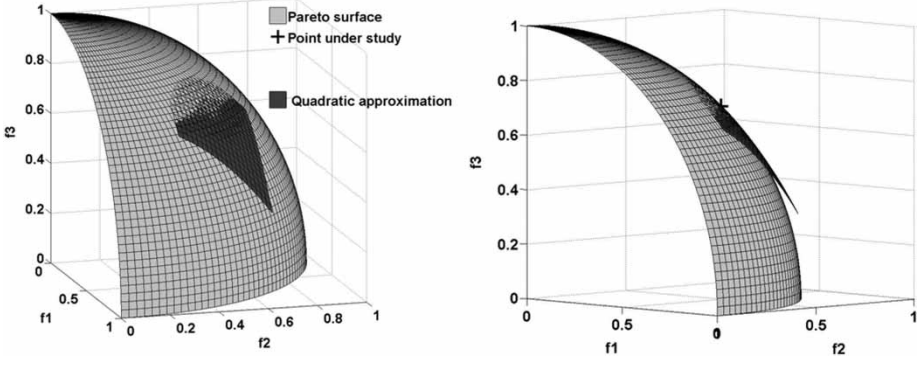

Figure 4. Example 1: new quadratic approximation (Equation (26)).

6.2. *Linear and quadratic approximations: Example 2*

Consider the following multi-objective problem studied by Tappeta *et al.* (2000):

$$\text{Minimize:} \quad \mathbf{f}(\mathbf{x}) = \{f_1(\mathbf{x}),\, f_2(\mathbf{x}),\, f_3(\mathbf{x})\} \tag{72}$$

$$\text{Subject to:} \quad \begin{array}{c} \mathrm{g}(\mathbf{x}) = 12 - x_1^2 - x_2^2 - x_3^3 \geq 0, \\ \mathbf{x} \geq 0 \end{array} \tag{73}$$

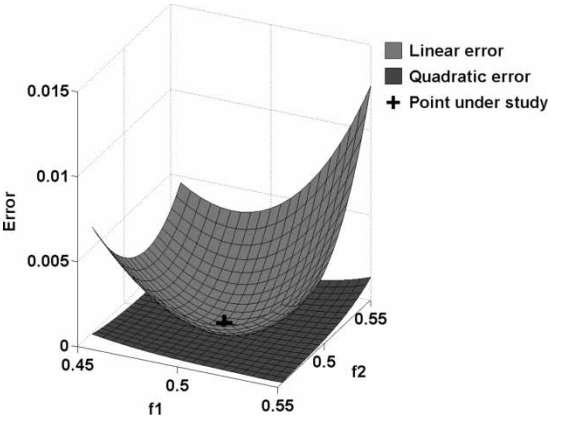

Figure 5. Example 1: relative error in predicting f_3.

where the objective functions are given by:

$$
\begin{aligned}
f_1 &= 25 - (x_1^3 + x_1^2(1 + x_2 + x_3) + x_2^3 + x_3^3)/10, \\
f_2 &= 35 - (x_1^3 + 2x_2^3 + x_2^2(2 + x_1 + x_3) + x_3^3)/10, \\
f_3 &= 50 - (x_1^3 + x_2^3 + 3x_3^3 + x_3^2(3 + x_1 + x_2))/10.
\end{aligned} \tag{74}
$$

The linear and quadratic approximations obtained by the analysis derived in Section 3 are given in Figures 6 and 7, respectively. The comparison between the two approximations is given in Figure 8. As expected, a much better approximation of the Pareto surface is provided by the quadratic approximation.

In the next two examples, non-smooth Pareto surfaces are considered. The Pareto analysis is used to detect non-differentiable Pareto points and limits of the approach are discussed.

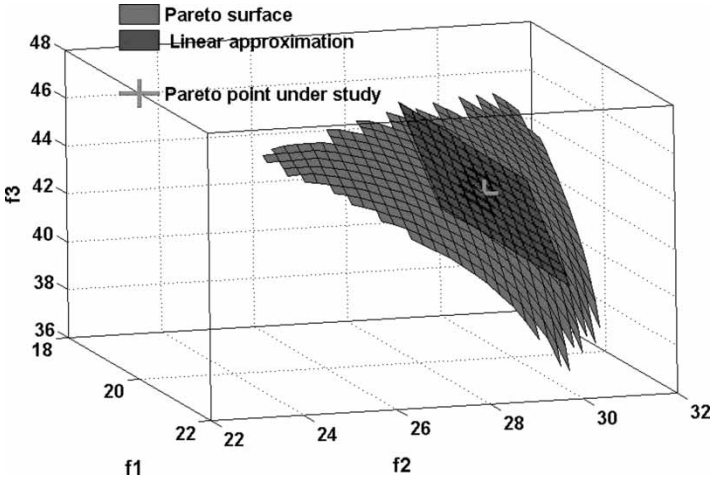

Figure 6. Example 2: linear approximation.

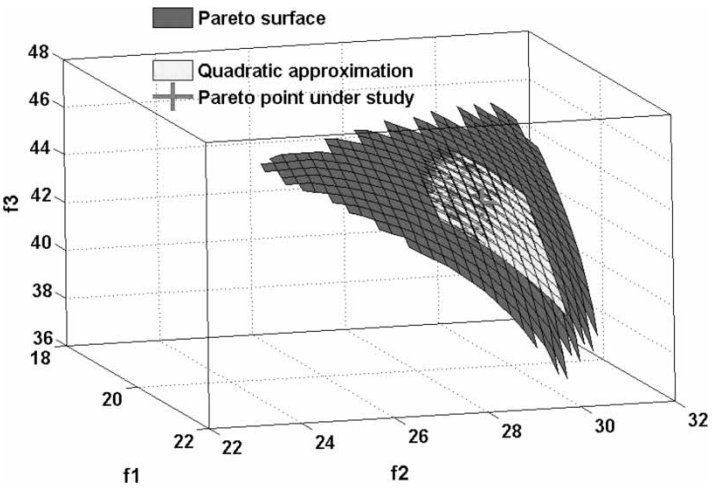

Figure 7. Example 2: quadratic approximation.

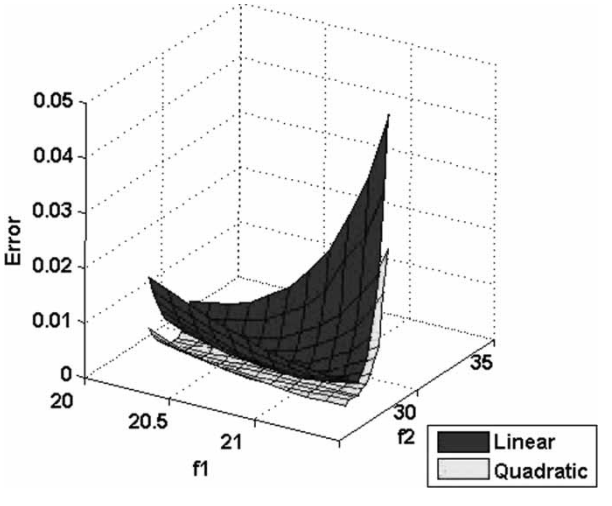

Figure 8. Example 2: relative error in predicting f_3.

6.3. *Detection of non-differentiable Pareto points—$K \subset K^* = T$: Example 3*

The linear bi-criteria test case taken from Zhang (2003b) is considered:

$$\text{Minimize:} \quad (x, y) \tag{75}$$

$$\text{Subject to:} \quad \begin{aligned} g_1(\mathbf{x}) &= -\frac{x}{4} - \frac{y}{4} + 1 \le 0, \\ g_2(\mathbf{x}) &= -\frac{x}{3} - \frac{y}{6} + 1 \le 0, \\ g_3(\mathbf{x}) &= -x \le 0, \\ g_4(\mathbf{x}) &= -y \le 0. \end{aligned} \tag{76}$$

In this test case, the design space coincides with the objective space. Hence, each point of the Pareto set is a point of the design space for which at least one constraint is active. The Pareto set

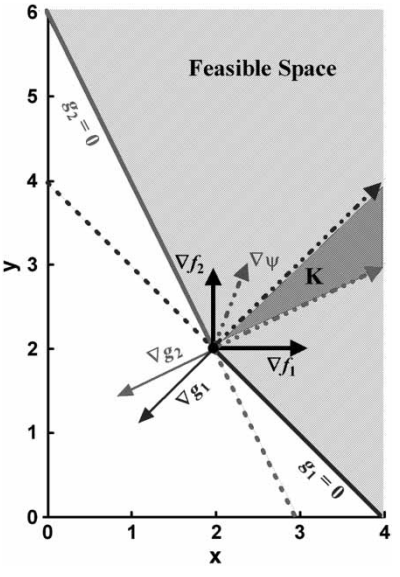

Figure 9. Example 3: relative position of the tangent cone and its polar at a non-differentiable Pareto point.

is represented by the lines $g_1 = 0$ and $g_2 = 0$. The kink in the Pareto surface corresponds to the only non-differentiable Pareto solution.

At the non-differentiable Pareto solution, the tangent cone $T = K^*$ and its polar cone K are constructed and shown in Figure 9. In this example $K \subset K^*$. A function Ψ that satisfies condition (63) can be obtained. This means that any small shift from the Pareto point leads to the loss of an active constraint. Therefore, according to the analysis given in the previous section, the Pareto point under study is non-differentiable. As the point is non-differentiable, the approximation derived in Section 3 cannot be used. Instead, tangent derivatives along the Pareto surface can be obtained using Equation (39).

At the non-differentiable point, both constraints are active and the corresponding matrix $\mathbf{J}$ is defined by:

$$\mathbf{J} = \begin{bmatrix} -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{3} & -\frac{1}{6} \end{bmatrix}. \tag{77}$$

Hence

$$\mathbf{P} = \mathbf{0}. \tag{78}$$

Let us study the sensitivity for the reduction of f_1 choosing $\Psi_1 = f_1$. The projection of $-\nabla\Psi_1$ gives $\mathbf{S}_1 = \mathbf{P}(-\nabla\Psi_1) = \mathbf{0}$ with $\boldsymbol{\mu}_1 = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Psi_1 = [-4\ 6]^T$. The negative value for the first component of $\boldsymbol{\mu}_1$ indicates that in the direction of reduction of f_1, constraint g_1 becomes inactive. Therefore, its gradient should be removed from $\mathbf{J}$ to obtain the reduced projection matrix:

$$\mathbf{P}_{\mathrm{r1}} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix}. \tag{79}$$

Thus, $\tilde{\mathbf{S}}_{\Psi 1}$ can be calculated as $\tilde{\mathbf{S}}_{\Psi 1} = -\mathbf{P}_{\mathrm{r}1}(\nabla\Psi_1)$ and we obtain $\tilde{\mathbf{S}}_{\Psi 1} = [-1/5\ 2/5]^T$. The sensitivity along $\tilde{\mathbf{S}}_{\Psi 1}$ for f_1 is then:

$$\left.\frac{df_2}{df_1}\right|_{\tilde{\mathbf{S}}_{\Psi 1}} = \frac{(\nabla f_2, \tilde{\mathbf{S}}_{\Psi 1})}{(\nabla f_1, \tilde{\mathbf{S}}_{\Psi 1})} = -2, \tag{80}$$

which corresponds to the slope of the line $g_2 = 0$.

The same analysis can be performed along the greatest feasible direction for objective f_2 by choosing $\Psi_2 = f_2$. It follows that $\mathbf{S}_2 = \mathbf{P}(-\nabla\Psi_2) = \mathbf{0}$ with the vector of Lagrange multipliers $\boldsymbol{\mu}_2 = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Psi_2 = [4\ -6]^T$. In this case, constraint g_2 becomes inactive and the new reduced projection matrix is:

$$\mathbf{P}_{\mathrm{r}2} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{81}$$

and $\tilde{\mathbf{S}}_{\Psi 2} = -\mathbf{P}_{\mathrm{r}2}(\nabla\Psi_2) = [1/2\ -1/2]^T$. The sensitivity along $\tilde{\mathbf{S}}_{\Psi 1}$ for f_2 is then:

$$\left.\frac{df_2}{df_1}\right|_{\tilde{\mathbf{S}}_{\Psi 2}} = \frac{(\nabla f_2, \tilde{\mathbf{S}}_{\Psi 2})}{(\nabla f_1, \tilde{\mathbf{S}}_{\Psi 2})} = -1, \tag{82}$$

which is the slope of the line $g_1 = 0$.

6.4. *Limit of the method—$K^* = \mathbf{T} \subset K$: Example 4*

In this example, the quadratic bi-criteria test case under linear constraints is considered:

$$\text{Minimize:} \quad (f_1, f_2) \tag{83}$$

$$\text{Subject to :} \quad \begin{aligned} g_1(\mathbf{x}) &= y - 2x + 1 \le 0, \\ g_2(\mathbf{x}) &= -y + \frac{x}{2} + \frac{1}{2} \le 0, \end{aligned} \tag{84}$$

with

$$\begin{aligned} f_1(\mathbf{x}) &= \left(y - \frac{3}{2}x - \frac{1}{2}\right)^2 \\ f_2(\mathbf{x}) &= \left(y - \frac{2}{3}x + \frac{1}{2}\right)^2 \end{aligned} \tag{85}$$

The feasible space is given in Figure 10. Only the two contour lines corresponding to the minimum value of the objectives are represented in the design space. Isovalue contour lines for each objective are straight lines parallel to the line corresponding to their respective minimum value. The set of Pareto solutions for the problem is given in Figure 11. The Pareto point, corresponding to the design point where both constraints are active, is a non-differentiable point in the objective space. Moving away from this point, the set of active constraints is changed and this clearly implies a shift in the direction tangent to the Pareto surface. In this example, the tangent cone $T = K^*$ is included in its polar cone K as shown in Figure 12. It is clear that there exists a vector $\nabla\Psi$ satisfying condition (63), therefore, the Pareto point is not differentiable. In this case, it is also possible to obtain the derivative along the greatest descent direction for each objective (Zhang 2003b) using Equation (39).

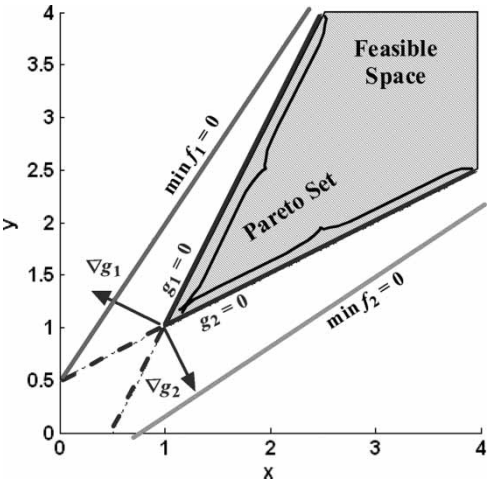

Figure 10. Example 4: design space/Pareto set.

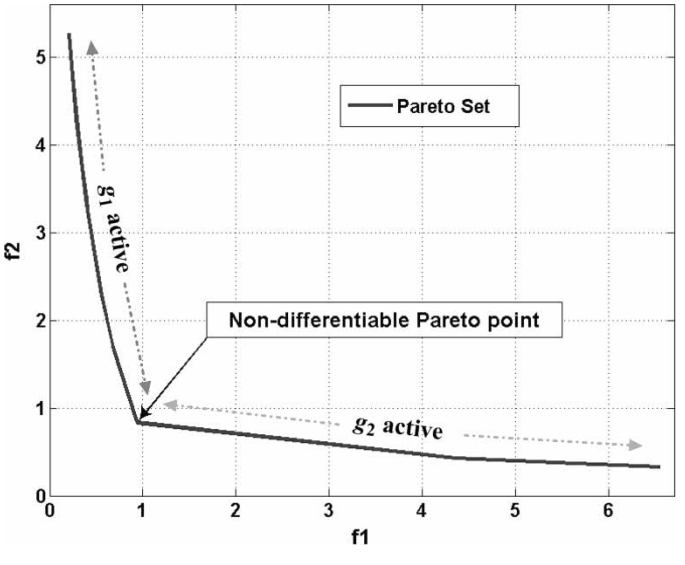

Figure 11. Example 4: objective space/Pareto set.

At the Pareto point under study, both constraints are active and the corresponding matrix $\mathbf{J}$ is defined by:

$$\mathbf{J} = \begin{bmatrix} -2 & 1 \\ \frac{1}{2} & -1 \end{bmatrix}. \tag{86}$$

Hence

$$\mathbf{P} = \mathbf{0}. \tag{87}$$

Evaluate the sensitivity for the reduction of objective f_1 and choose $\Psi_1 = f_1$. Proceeding as in the previous example, obtain $\mathbf{S}_1 = \mathbf{P}(-\nabla\Psi_1) = \mathbf{0}$ and the corresponding vector of Lagrange multipliers $\boldsymbol{\mu}_1 = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Psi_1 = [4/3 - 2/3]^T$. The negative value for the second component

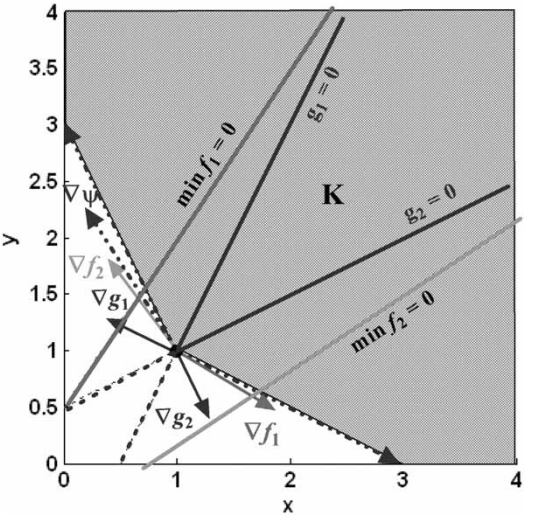

Figure 12. Example 4: relative position of the tangent cone and its polar at a non-differentiable Pareto point.

of $\boldsymbol{\mu}_1$ indicates that in the direction of reduction of f_1, constraint g_2 becomes inactive. Therefore, its gradient should be removed from $\mathbf{J}$ to obtain the reduced projection matrix:

$$\mathbf{P}_{\mathrm{r}1} = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}. \tag{88}$$

Thus, $\tilde{\mathbf{S}}_{\Psi 1}$ can be calculated as $\tilde{\mathbf{S}}_{\Psi 1} = -\mathbf{P}_{\mathrm{r}1}(\nabla\Psi_1)$ and $\tilde{\mathbf{S}}_{\Psi 1} = [1/5\ 2/5]^T$ is obtained. The sensitivity along the greatest feasible direction for f_1 is then given by:

$$\left.\frac{df_2}{df_1}\right|_{\tilde{\mathbf{S}}_{\Psi 1}} = \frac{(\nabla f_2, \tilde{\mathbf{S}}_{\Psi 1})}{(\nabla f_1, \tilde{\mathbf{S}}_{\Psi 1})} = -\frac{2}{9}. \tag{89}$$

The same analysis can be performed along the greatest feasible direction for objective f_2 by choosing $\Psi_2 = f_2$. It follows that $\mathbf{S}_2 = \mathbf{P}(-\nabla\Psi_2) = \mathbf{0}$ with the corresponding vector of Lagrange multipliers $\boldsymbol{\mu}_2 = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla\Psi_2 = [-0.1852\ 1.4815]^T$. In this case, constraint g_1 becomes inactive and the new reduced projection matrix is:

$$\mathbf{P}_{\mathrm{r}2} = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \tag{90}$$

and $\tilde{\mathbf{S}}_{\Psi 2} = -\mathbf{P}_{\mathrm{r}2}(\nabla\Psi_2) = [2/9\ 1/9]^T$. The sensitivity along the greatest feasible direction for f_2 is then:

$$\left.\frac{df_2}{df_1}\right|_{\tilde{\mathbf{S}}_{\Psi 2}} = \frac{(\nabla f_2, \tilde{\mathbf{S}}_{\Psi 2})}{(\nabla f_1, \tilde{\mathbf{S}}_{\Psi 2})} = -0.1389. \tag{91}$$

7. I-beam optimization problem

A typical engineering design optimization problem is considered: optimization of an I-beam, namely minimization of its weight and displacement subject to a particular load. The test case can be entirely described with analytical equations and is taken from Yang *et al.* (2002). The problem is to find the dimensions of the cross-section of a beam of fixed length, which satisfies geometric and strength constraints, and minimize its weight and static deflection under loads P and Q applied along the z and y axes, respectively. A description of the I-beam problem is given in Figure 13. Since the length of the beam is fixed, minimizing its weight is equivalent to minimizing its cross-sectional area. The optimization problem can be formulated as follows:

$$\underset{\mathbf{x}}{\text{Minimize}} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} area = 2x_2x_4 + x_3(x_1 - 2x_4) \\ \\ displacement = \dfrac{\mathrm{PL}^3}{48\mathrm{EI}} \end{pmatrix} \tag{92}$$

Subject to a strength constraint:

$$g(\mathbf{x}) = \frac{\mathrm{M_y}}{\mathrm{Z}_y} + \frac{\mathrm{M_z}}{\mathrm{Z_z}} - \sigma_{\mathrm{a}} \leq 0 \ \ (\mathrm{kN.cm^{-2}}) \tag{93}$$

and geometric constraints on the inputs range of variation:

$$\begin{aligned} 10 \leq x_1 \leq 80 \, (\mathrm{cm}) \\ 10 \leq x_2 \leq 50 \, (\mathrm{cm}) \\ 0.9 \leq x_3 \leq 5 \, (\mathrm{cm}) \\ 0.9 \leq x_4 \leq 5 \, (\mathrm{cm}) \end{aligned} \tag{94}$$

where the moments are given by:

$$\mathrm{M_y} = \frac{\mathrm{P}}{2} \times \frac{\mathrm{L}}{2}, \tag{95}$$

$$\mathrm{M_z} = \frac{\mathrm{Q}}{2} \times \frac{\mathrm{L}}{2} \tag{96}$$

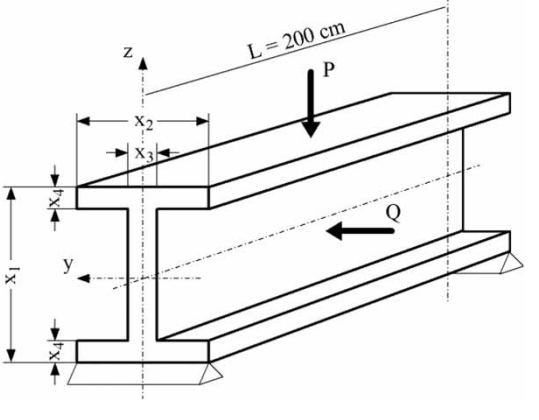

Figure 13. The I-beam design problem.

and the moments of inertia are calculated as follows:

$$Z_y = \frac{1}{6x_1}\{x_3(x_1 - x_4)^3 + 2x_2x_4[4x_4^2 + 3x_1(x_1 - 2x_4)]\}, \tag{97}$$

$$Z_y = \frac{1}{6x_2}\{(x_1 - x_4)x_3^3 + 2x_4x_2^3\}, \tag{98}$$

with $E = 2 \times 10^4\,\mathrm{kNcm^{-2}}$, $\sigma_a = 16\,\mathrm{kNcm^{-2}}$, $P = 600\,\mathrm{kN}$, $Q = 50\,\mathrm{kN}$ and $L = 200\,\mathrm{cm}$.

The set of solutions for this problem is obtained using the in-house developed gradient-based optimization tool for multi-objective optimization mentioned previously (Guenov *et al.* 2005, Utyuzhnikov *et al.* 2005). This tool searches for solutions that are uniformly distributed over the Pareto front. The set of Pareto solutions obtained for the test case is plotted in Figure 14. The high density of points gives a good representation of the Pareto front, particularly in the region where the curvature is important and will allow for comparison of the results of the approximation with the exact Pareto front.

The analysis focuses on the use of local Pareto approximation to understand the local trade-offs at different Pareto solutions. With respect to the Pareto analysis described in Section 3, the SA of the solution can be formulated as a single-objective optimization problem, which reduces to a linear problem in the case of two objectives:

Minimize and maximize δf_1

Subject to

$$\mathbf{x}_{min} < \mathbf{x} + \delta\mathbf{x} < \mathbf{x}_{max}, \tag{99}$$

$$g + \frac{dg}{df_1}\delta f_1 + \frac{1}{2}\frac{d^2g}{df_1^2}(\delta f_1)^2 < 0, \tag{100}$$

$$\left|\frac{(1/2)(d^2f_2/df_1^2)\delta f_1}{(df_2/df_1)}\right| < \alpha, \tag{101}$$

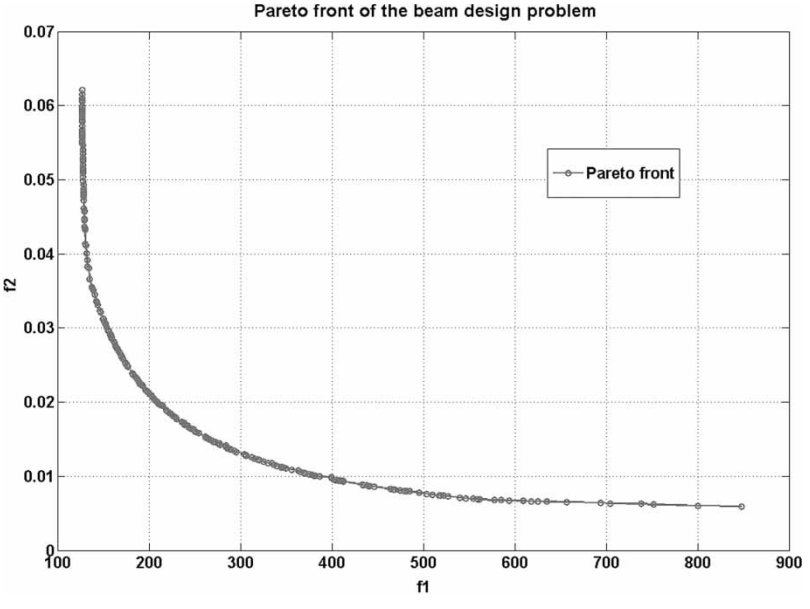

Figure 14. Pareto front for the I-beam design problem.

Table 1. Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$).

x_1 (cm)	80
x_2 (cm)	50
x_3 (cm)	0.9
x_4 (cm)	3.48
f_1 (cm^2)	413.83
f_2 (cm)	0.0093
g (kN.cm^{-2})	−12.930

where $\delta\mathbf{x}$ is given by (13) and alpha is fixed to 0.2 for this test case. Equation (101) is a constraint on the discrepancy between the first- and second-order approximation.

The approximation analysis is carried out on two Pareto solutions and approximations are derived to understand their local sensitivity. The first solution considered is given in Table 1. For this Pareto solution, three constraints on the input range are active: two constraints on the maximum bound for x_1 and x_2, and one constraint on the minimum bound for x_3. The relevant first- and second-order derivatives on the Pareto surface for objective f_2 and constraint g are given in Tables 2 and 3, respectively.

The values for input variables, objectives and constraints of the approximate solutions defining the bounds of the Pareto approximation are given in Table 4. Here, the values for objective f_1 are the results of the linear optimization problem and approximate values for f_2 and g are obtained with the quadratic approximation. Figure 15 plots the approximate solutions in the criterion space and Figure 16 shows the geometry of the corresponding beams.

Table 2. First- and second-order derivatives on the Pareto surface at Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$).

df_2/df_1	-2.29×10^{-5}
$d^2 f_2/df_1^2$	1.26×10^{-7}

Table 3. First- and second-order derivatives on the Pareto surface at Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$).

dg/df_1	-7.98×10^{-3}
$d^2 g/df_1^2$	4.48×10^{-5}

Table 4. Objectives and constraints values of Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$) and approximate Pareto solutions associated with the approximation bounds.

Variables	Approx solution 1		Pareto solution 1	Approx solution 2	
x_1 (cm)	80		80	80	
x_2 (cm)	50		50	50	
x_3 (cm)	0.9		0.9	0.9	
x_4 (cm)	2.74		3.48	4.22	
Outputs	Approx	Model	Model	Approx	Model
f_1 (cm^2)	341.175	341.175	413.83	486.493	486.493
f_2 (cm)	0.0113	0.0113	0.0093	0.0079	0.0078
g (kN.cm^{-2})	−12.232	−12.202	−12.930	−13.392	−13.412

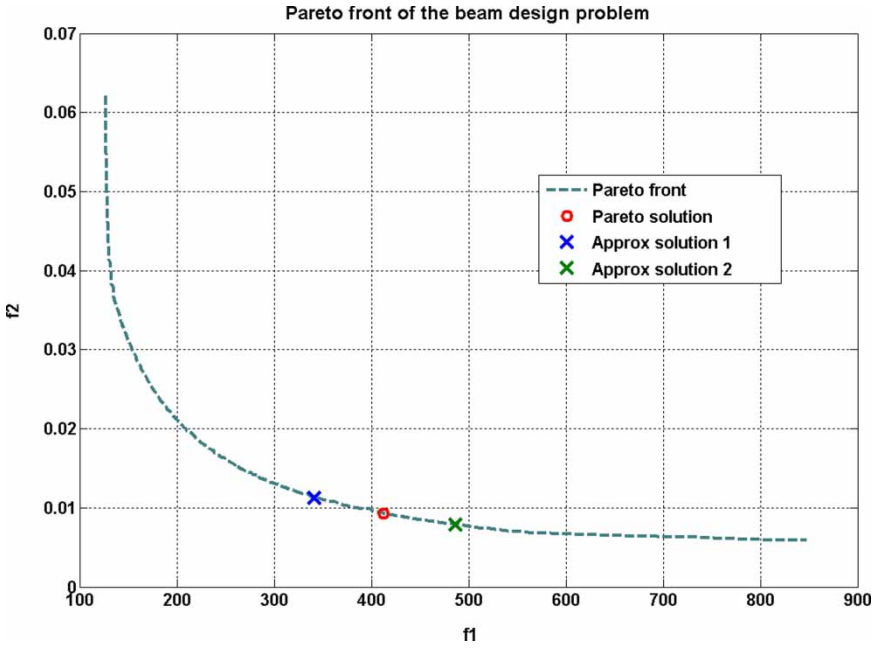

Figure 15. Extent of the approximation bounds at the Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$).

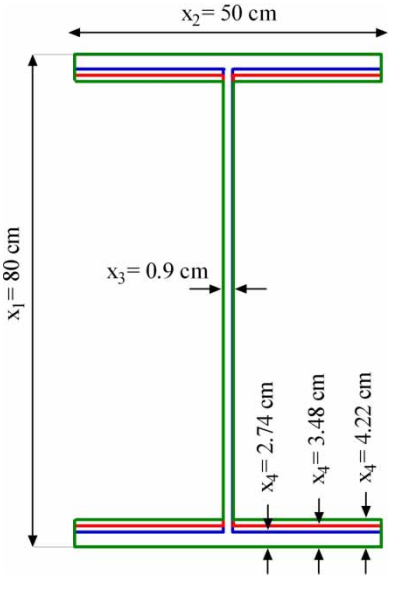

Figure 16. Extent of the approximation bounds at the Pareto solution ($f_1 = 413.83\,\text{cm}^2$, $f_2 = 0.0093\,\text{cm}$).

Back-mapping analysis (see Equation (13)) indicates what variables are changed as a result of the Pareto approximation. For this point, three constraints on the bounds of the inputs variables are active (x_1, x_2 and x_3) and remain active for the Pareto approximation. The Pareto approximation analysis indicates that, locally, this part of the Pareto front corresponds to changes with respect to x_4 only. This can be clearly seen in Figure 16 where only the thickness of the foot and head of the beam, *i.e.* x_4, is changed. Such a consideration provides a designer with valuable

information about the Pareto solution in the vicinity of the point under study. It also indicates what parameters can be adjusted locally if the compromise solution does not entirely satisfy some preferences on the type of solutions that were not articulated in the formulation of the optimization problem.

The second solution considered is given in Table 5. Two constraints on the minimum bounds for x_3 and x_4 are active at this point. The constraint on the strength is also activated here. This means that on this portion on the Pareto surface, input variables x_3 and x_4 are fixed to their minimum value. Only x_1 and x_2 are left free to move along the Pareto front. The first- and second-order derivatives on the Pareto surface in the criterion space are given in Table 6.

The extent of the approximation bounds is estimated with the same approach. The values for the values for input variables, objectives and constraints of the approximate solutions corresponding to the bounds of the Pareto approximation are given in Table 7 and plotted in Figures 17 and 18.

It is worth noting that when a constraint is active at the point under study, the approximation guarantees that it will remain active and therefore will not be violated within the bounds of the approximation. However, when moving away from the point under study, the model will (most of the time) depart from the approximation. Therefore, it is difficult to ensure that the active constraints will strictly remain active. Indeed, Table 7 shows that at the bounds of the approximation, the strength constraint does not remain active exactly. The assumption behind the approximation theory is that, locally, the move in the design space to accommodate for the change

Table 5. Pareto solution ($f_1 = 128.718\,\text{cm}^2$, $f_2 = 0.0472\,\text{cm}$).

x_1 (cm)	70.72
x_2 (cm)	37.05
x_3 (cm)	0.90
x_4 (cm)	0.90
f_1 (cm^2)	128.718
f_2 (cm)	0.0472
g (kN.cm^{-2})	0.000

Table 6. First- and second-order derivatives on the Pareto surface at Pareto solution ($f_1 = 128.718\,\text{cm}^2$, $f_2 = 0.0472\,\text{cm}$).

df_2/df_1	-4.01×10^{-3}
$d^2 f_2/df_1^2$	1.42×10^{-3}

Table 7. Objectives and constraints values of Pareto solution ($f_1 = 128.718\,\text{cm}^2$, $f_2 = 0.0472\,\text{cm}$) and approximate Pareto solutions associated with the approximation bounds.

Variables	Approx solution 1		Pareto solution 1	Approx solution 2	
x_1 (cm)	66.91		70.72	74.53	
x_2 (cm)	38.33		37.05	35.77	
x_3 (cm)	0.90		0.90	0.90	
x_4 (cm)	0.90		0.90	0.90	
Outputs	Approx	Model	Model	Approx	Model
f_1 (cm^2)	127.588	127.588	128.718	129.845	129.848
f_2 (cm)	0.0527	0.0521	0.0472	0.0436	0.0430
g (kN.cm^{-2})	0.000	0.046	0.000	0.000	0.045

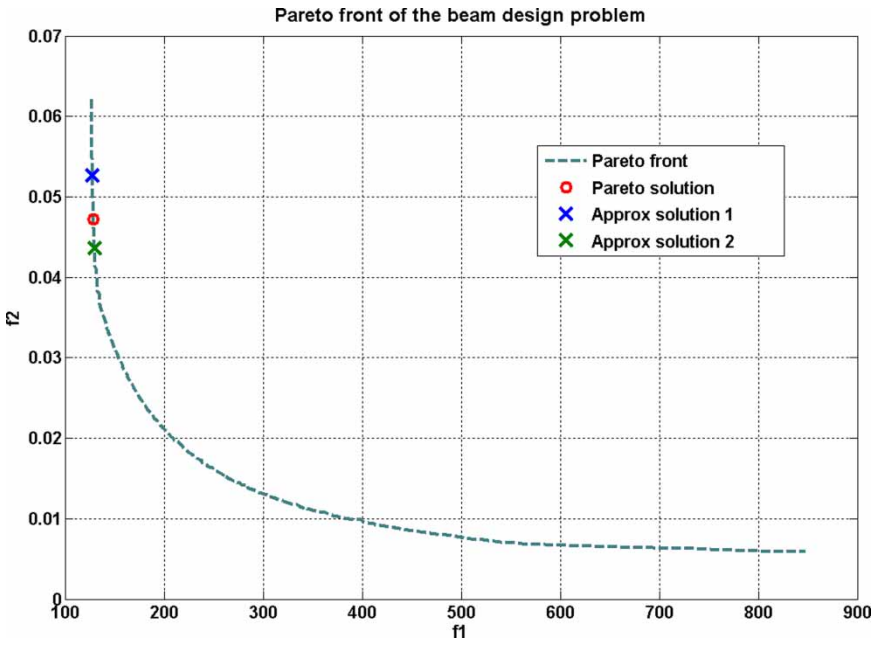

Figure 17. Extent of the approximation bounds at the Pareto solution ($f_1 = 128.718\,\text{cm}^2, f_2 = 0.0472\,\text{cm}$).

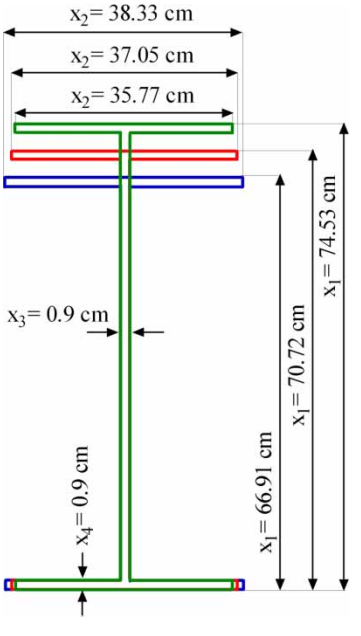

Figure 18. Beams associated with the approximation at the Pareto solution ($f_1 = 128.718\,\text{cm}^2,\ f_2 = 0.0472\,\text{cm}$).

in objective functions on the Pareto surface is locally normal to the set of gradients of the active constraints. For active bound constraints, this ensures that the constraint will remain active in this direction; however, for any other active constraint, this is only valid locally.

8. Conclusions

This article has shown that existing formulas for linear approximation (Tappeta and Renaud 1999) are only valid under particular conditions. The exact formulas for the first- and second-order approximations are derived in the general case. Test cases show that the new approximations outperform the existing ones. The concept of a local quick Pareto analysis based on these approximations is proposed to understand the trade-offs between objectives. It gives the DM the opportunity to articulate local preferences on the Pareto solution under consideration by improving some criteria and compromising on others. Along with understanding the trade-offs between objectives, the DM can ensure that inactive constraints remain inactive within the extent of the approximation. A method for detecting non-differentiable Pareto point has been proposed and its limitations pointed out. The approach was tested on a typical engineering multi-objective optimization problem. Future work will concentrate on testing and application of the method to complex industrial test cases.

Acknowledgements

The research reported in this article was carried out within the VIVACE Integrated Project (AIP3 CT-2003-502917), which is partly sponsored by the Sixth Framework Programme of the European Community under priority 4 'Aeronautics and Space'.

References

Fadel, G., Konda S., Blouin, V.Y. and Wiecek, M.M., 2002. Epsilon-optimality in bi-criteria optimization. *Proceedings of the 43rd Structures, Structural, Dynamics and Materials Conference*, 22–25 April, Denver, Co, Alexandria, VA: AIAA.

Fletcher, R., 1989. *Practical methods of optimization*. New York: Wiley.

Guenov, M.D., Utyuzhnikov, S.V. and Fantini, P., 2005. Application of the modified physical programming method to generating the entire Pareto frontier in multi-objective optimization. *Proceedings of the 6th Eurogen Conference*, 12–14 September, Munich, Germany.

Hernandez, S., 1995. A general sensitivity analysis for unconstrained and constrained Pareto optima in multi-objective optimization. *Proceedings of the 36th AIAA/ASME Structures Dynamics and Materials Conference*, AIAA-1288-CP, 18–21 April, New Orleans, LA, Alexandria, VA: AIAA.

Messac, A., 1996. Physical programming effective optimization for computational design. *AIAA Journal*, 34 (1), 149–158.

Miettinen, K.M., 1999. *Nonlinear multi-objective optimization*. Boston: Kluwer.

Rosen, J.B., 1958. The gradient projection method for nonlinear programming. Part I. Linear constraints. *Journal of the Society for Industrial and Applied Mathematics*, 8 (1), 181–217.

Tammer, C., 1994. Stability results for approximately efficient solutions. *OR Spektrum*, 16 (1), 47–52.

Tappeta, R.V. and Renaud, J.E., 1999. Interactive multi-objective optimization procedure. *AIAA* 99–1207. *Proceedings of the 40th Structures, Structural Dynamics and Materials Conference and Exhibit*, 12–15 April, St Louis, MO, Alexandria, VA: AIAA, Vol. 1, 27–41.

Tappeta, R.V., Renaud, J.E., Messac, A. and Sundararaj, G.J., 2000. Interactive physical programming: trade-off analysis and decision making in multidisciplinary optimization. *AIAA Journal*, 38 (5), 917–926.

Utyuzhnikov, S.V., Guenov, M.D. and Fantini, P., 2005. Numerical method for generating the entire Pareto frontier in multi-objective optimization. *Proceedings of the 6th Eurogen Conference*, 12–14 September, Munich, Germany.

Vincent, T.L. and Grantham, W.J., 1981. *Optimality in parametric systems*. New York: Wiley.

Yang B.S., Yeun, Y.S. and Ruy, W.S., 2002. Managing approximation models in multiobjective optimization. *Structural and Multidisciplinary Optimization*, 24 (2), 141–156.

Zhang, W.H., 2003a. Pareto sensitivity analysis in multi-criteria optimization. *Finite Elements in Analysis and Design*, 39 (5), 505–520.

Zhang, W.H., 2003b. On the Pareto optimum sensitivity analysis in multi-criteria optimization. *International Journal for Numerical Methods in Engineering*, 58 (6), 955–977.