

Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational
Mathematics (IMACM)

Preprint BUW-IMACM 20/04

Johanna Schultes, Michael Stiglmayr, Kathrin Klamroth,
and Camilla Hahn

Hypervolume Scalarization for Shape Optimization to Improve Reliability and Cost of Ceramic Components

February 21, 2020

<http://www.math.uni-wuppertal.de>

Hypervolume Scalarization for Shape Optimization to Improve Reliability and Cost of Ceramic Components

Johanna Schultes · Michael Stiglmayr ·
Kathrin Klamroth · Camilla Hahn

Received: date / Accepted: date

Abstract In engineering applications one often has to trade-off among several objectives as, for example, the mechanical stability of a component, its efficiency, its weight and its cost. We consider a biobjective shape optimization problem maximizing the mechanical stability of a ceramic component under tensile load while minimizing its volume. Stability is thereby modeled using a Weibull-type formulation of the probability of failure under external loads. The PDE formulation of the mechanical state equation is discretized by a finite element method on a regular grid.

To solve the discretized biobjective shape optimization problem we suggest a hypervolume scalarization, with which also unsupported efficient solutions can be determined without adding constraints to the problem formulation. We investigate the relation of the hypervolume scalarization to the weighted sum scalarization and to direct multiobjective descent methods. Since gradient information can be efficiently obtained by solving the adjoint equation, the scalarized problem can be solved by a gradient ascent algorithm. We evaluate our approach on a 2D test case representing a straight joint under tensile load.

Key words: biobjective shape optimization, hypervolume scalarization, shape gradients, probability of failure

This work was supported by the federal ministry of research and education (BMBF, Grant-No. 05M18PXA) as a part of the GIVEN consortium.

J. Schultes · M. Stiglmayr · K. Klamroth · C. Hahn
University of Wuppertal
School of Mathematics and Natural Sciences
Tel.: +49-202-439-2802
Fax.: +49-202-439-3061
{schultes, stiglmayr, klamroth, hahn}@math.uni-wuppertal.de

1 Introduction

In engineering design one often follows multiple incompatible goals, for example, when maximizing reliability and efficiency while simultaneously minimizing cost. In practical applications, conflicting goals are often combined into one single objective function using a weighted sum of the criteria. The main advantage of this approach is its simplicity and its numerical stability. On the downside, the selection of meaningful weights is generally difficult. Moreover, relevant compromise solutions may be missed when the problem is non-convex. In this paper, we consider a biobjective shape optimization problem with the two objectives reliability and cost. To carry over the main advantages of the weighted sum and simultaneously overcome its main drawbacks, we suggest to use parametric hypervolume scalarizations to efficiently generate compromise solutions. The scalarized subproblems are solved using a gradient ascent method that takes advantage of derivative information obtained from the numerical solution of the adjoint equation.

Shape optimization problems are typical examples of engineering problems where multiple conflicting objective functions have to be considered. In this paper, we focus on the maximization of the mechanical stability of a ceramic component under tensile load while simultaneously minimizing its volume (as a measure of cost and weight). We follow the modelling approach proposed in Doganay et al. (2019), where the mechanical stability is modelled using a Weibull formulation of the probability of failure. For a general introduction to shape optimization we refer to Haslinger and Mäkinen (2003).

The main advantage of using a probabilistic formulation for the mechanical stability is that in this way the probability of failure can be assessed with a differentiable objective function. This paves the way for the application of efficient optimization methods using gradient information. We adopt the formulation suggested in Bolten et al. (2015, 2019) that is based on the probabilistic model of Weibull (1939). In our numerical implementation we follow a first discretize then optimize paradigm as suggested in Doganay et al. (2019). In this approach, a finite dimensional optimization problem is solved on a finite element grid subject to a discretized PDE-constraint that is induced by the mechanical state equation.

The minimization of the probability of failure of a component usually leads to large and hence costly components. This unwanted effect can be avoided by setting an upper bound on the admissible volume, see, for example, Bolten et al. (2019). However, such a constrained formulation of the shape optimization problem does not provide information on the trade-off between reliability and cost. Indeed, interpreting the volume of the component as a second and independent objective function allows for the computation of alternative shapes, thus supporting the decision maker with the selection of a most preferred solution with respect to both criteria. Following the approach of Doganay et al. (2019), we complement the minimization of the probability of failure of a component by the minimization of its volume in a biobjective shape optimization model.

For a general introduction into the field of multiobjective optimization we refer to Miettinen (1998) and Ehrgott (2005). Problems of multiobjective shape optimization have mainly been approached with evolutionary multiobjective optimization (EMO) algorithms, see, for example, Deb and Goel (2002); Chirkov et al. (2018). However, recently multiobjective PDE-constrained optimization problems have also been considered from an optimal control point of view in Iapichino et al. (2017).

In this paper, we employ a scalarization-based approach to approximate the Pareto front of the biobjective shape optimization problem. Due to its simplicity and its numerical stability, the weighted sum method is a widely used scalarization method in this context. It combines all objective functions into one weighted sum objective, hence yielding a single objective counterpart problem without adding any additional constraints. However, the weighted sum method is quite limited in the context of non-convex problems. Indeed, by varying the weights only supported efficient solutions can be computed, i.e., Pareto optimal shapes for which the corresponding outcome vectors lie on the convex hull of the set of all feasible outcomes. Since relevant compromise solutions may thus be missed, we suggest to use nonlinear scalarizing functions as, for example, the hypervolume scalarization. While sharing the property that it leads to single-objective counterpart problems without adding constraints to the problem formulation, it is not restricted to supported efficient solutions. In the biobjective case, the hypervolume scalarization yields a quadratic objective function that has a nice geometrical interpretation. We note that quadratic scalarizations are also considered in Fliege (2004); Dandurand and Wiecek (2016).

The hypervolume scalarization is based on the hypervolume indicator that was introduced in Zitzler and Thiele (1999) as a measure to compare the performance of EMO algorithms. The hypervolume indicator has since become an important measure in a large variety of applications, of which the fitness evaluation of solutions sets (particularly in the context of EMO algorithms) and of individual solutions are only two examples. Given a reference point in the objective space, the hypervolume indicator of a solution set measures the volume that is (a) dominated by the solution set, and (b) dominates the reference point, c.f. Figure 1. A formal definition is given in Section 3 below. We exemplarily refer to Auger et al. (2009); Yang et al. (2019) as two examples from this active research field. The hypervolume indicator is also used as an objective function when searching for good representations of the Pareto front. Heuristic and exact solution approaches for this subset selection problem were discussed in Guerreiro et al. (2016) and in Bringmann et al. (2014); Kuhn et al. (2016), respectively. When restricting the search to individual solutions rather than solution sets, then the hypervolume indicator can be interpreted as a nonlinear scalarizing function, see, for example, Hernandez et al. (2018); Touré et al. (2019); Schulze et al. (2020).

This paper is organized as follows. In Section 2 a biobjective shape optimization model is introduced together with some fundamental properties of multiobjective optimization in the context of shape optimization. In Section 3 the

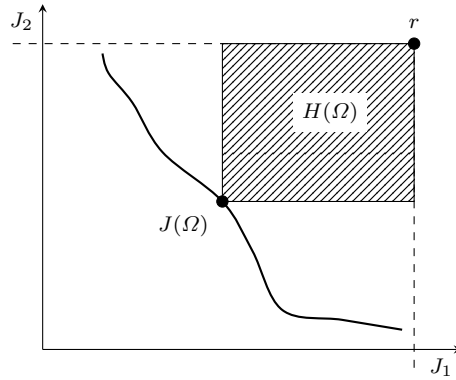


Fig. 1: Hypervolume contribution $H(\Omega)$ of a solution Ω with respect to the reference point r

hypervolume indicator is formally defined and its properties as a scalarizing functions are discussed and compared to the weighted sum method and the multiobjective gradient descent approach. Section 4 is devoted to the details of the numerical implementation. In particular, a finite element discretization and an iterative accent algorithm using smoothed gradients are presented. Furthermore, the choice of the reference point for the hypervolume scalarization is discussed. In Section 5 we consider in a 2D case study the optimal shape of a rod under tensile load. The article is concluded in Section 6.

2 Biobjective Shape Optimization

In this section, we first provide a general formulation of multiobjective shape optimization problems (Section 2.1) and then focus on a particular engineering application, namely the biobjective shape optimization of a ceramic component under tensile load w.r.t. the two objectives mechanical stability and volume (Section 2.2).

2.1 Multiobjective Optimization

We consider multiobjective optimization problems occurring in engineering applications and particularly focus on multiobjective shape optimization problems. To keep the exposition general, we denote the set of all admissible shapes by \mathcal{O}^{ad} and refer to individual shapes as $\Omega \in \mathcal{O}^{\text{ad}}$. Let $J : \mathcal{O}^{\text{ad}} \rightarrow \mathbb{R}^Q$ be a vector valued objective function comprising Q individual objective functions $J_i : \mathcal{O}^{\text{ad}} \rightarrow \mathbb{R}$, $i = 1, \dots, Q$ and let $J(\Omega) = (J_1(\Omega), \dots, J_Q(\Omega))^{\top}$. Then the *multiobjective shape optimization problem* is given by

$$\begin{aligned} \min J(\Omega) &= (J_1(\Omega), \dots, J_Q(\Omega))^{\top} \\ \text{s.t. } \Omega &\in \mathcal{O}^{\text{ad}}. \end{aligned} \quad (2.1)$$

Note that in general there does not exist a solution $\Omega^* \in \mathcal{O}^{\text{ad}}$ that minimizes all Q objective functions simultaneously (such a solution would be called *ideal solution*). Throughout the article we use the Pareto concept of optimality: A solution $\bar{\Omega} \in \mathcal{O}^{\text{ad}}$ is called *efficient* or *Pareto optimal*, if its objective function vector $J(\bar{\Omega})$ can not be improved in one component without deterioration in at least one other component. More precisely, $\bar{\Omega}$ is efficient if and only if there does not exist $\Omega' \in \mathcal{O}^{\text{ad}}$ such that $J(\Omega') \leq J(\bar{\Omega})$, i.e., $J_i(\Omega') \leq J_i(\bar{\Omega})$ for all $i = 1, \dots, Q$ and $J(\Omega') \neq J(\bar{\Omega})$. The objective vector $J(\bar{\Omega}) = (J_1(\bar{\Omega}), \dots, J_Q(\bar{\Omega}))^\top$ of an efficient solution $\bar{\Omega}$ is called *nondominated point* or *nondominated outcome vector*. An efficient solution $\bar{\Omega}$ (respectively a nondominated point $J(\bar{\Omega})$) is called *supported efficient* (or *supported nondominated*, respectively) if there is no point $y \in \text{conv}\{J(\Omega) : \Omega \in \mathcal{O}^{\text{ad}}\}$ dominating $J(\bar{\Omega})$. For an extensive introduction to multiobjective optimization see, e.g., Steuer (1986); Miettinen (1998); Ehrgott (2005).

2.2 Biobjective Optimization of Stability and Cost

In the biobjective case study described in this paper we consider a ceramic component under tensile load. The component is represented by a two dimensional compact body with Lipschitz boundary, denoted by $\Omega \subset \mathbb{R}^2$ (in general, $\Omega \subset \mathbb{R}^d$, $d = 2, 3$). We assume that the boundary of Ω is composed of three parts, i.e.,

$$\partial\Omega = \text{cl}(\partial\Omega_D) \cup \text{cl}(\partial\Omega_{N_{\text{fixed}}}) \cup \text{cl}(\partial\Omega_{N_{\text{free}}}).$$

Here, $\partial\Omega_D$ is the part where the boundary is fixed, i.e., where *Dirichlet* conditions are satisfied. The tensile load acts at $\partial\Omega_{N_{\text{fixed}}}$, which is also fixed and where *Neumann* conditions apply. The remaining parts of $\partial\Omega$ can be modified during the optimization and are denoted by $\Omega_{N_{\text{free}}}$. Furthermore, we assume that all reasonable shapes are contained in a (sufficiently large) bounded open set $\hat{\Omega} \subset \mathbb{R}^2$ that satisfies the cone property (see, e.g., Bolten et al. (2015); Doganay et al. (2019)). Using this notation, the set of admissible shapes can be defined as

$$\mathcal{O}^{\text{ad}} := \{\Omega \subset \hat{\Omega} : \partial\Omega_D \subset \partial\Omega, \partial\Omega_{N_{\text{free}}} \subset \partial\Omega, \Omega \text{ satisfies the cone property}\},$$

see Figure 2 for an exemplary setting.

A ceramic component under tensile load behaves according to the linear elasticity theory, see, e.g., Braess (2007). Its deformation under tensile loads is described by the state equation which is an elliptic PDE given by

$$\begin{aligned} -\text{div}(\sigma(u(x))) &= f(x) & \text{for } x \in \Omega \\ u(x) &= 0 & \text{for } x \in \partial\Omega_D \\ \sigma(u(x))n(x) &= g(x) & \text{for } x \in \partial\Omega_{N_{\text{fixed}}} \\ \sigma(u(x))n(x) &= 0 & \text{for } x \in \partial\Omega_{N_{\text{free}}}. \end{aligned} \tag{2.2}$$

Here, $f \in L^2(\Omega, \mathbb{R}^2)$ denotes the volume force density and $g \in L^2(\partial\Omega_{N_{\text{fixed}}}, \mathbb{R}^2)$ denotes the surface loads. Moreover, $u \in H^1(\Omega, \mathbb{R}^2)$ describes the displacement

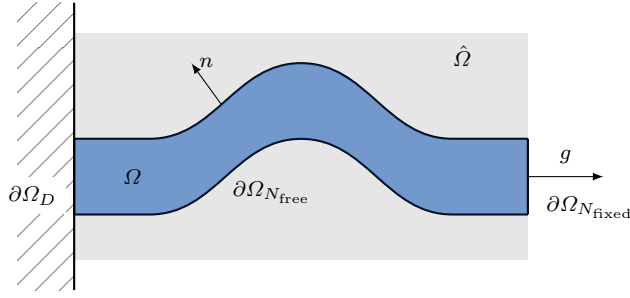


Fig. 2: Possible shape of a rod Ω under tensile load g

field as the reaction on the given loads, the stress tensor is denoted by $\sigma \in L^2(\Omega, \mathbb{R}^{2 \times 2})$, and $n(x)$ is the outward pointing normal vector at $x \in \partial\Omega$ (which is defined almost everywhere on $\partial\Omega$). We refer to Doganay et al. (2019) for further details.

Based on this formulation we consider the biobjective optimization of the mechanical stability and the cost of the component. Note that these two objective functions are usually conflicting since larger components tend to have a higher stability. We use the same formulation as Doganay et al. (2019) to assess the mechanical stability of the component via its probability of failure, that is to be minimized. Thus, our first objective function, J_1 , is based on a probabilistic Weibull-type model as suggested in Bolten et al. (2015):

$$J_1(\Omega) := J_1(\Omega, u(\Omega)) := \frac{1}{2\pi} \int_{\Omega} \int_{S^1} \left(\frac{n^\top \sigma(Du(x))n}{\sigma_0} \right)^m dn dx \quad (2.3)$$

Here, m denotes the Weibull module that typically assumes values between 5 and 25, Du denotes the Jacobian matrix of u and σ_0 is a positive constant.

The second objective function, J_2 , models the material cost which is approximated by the *volume* of the component $\Omega \in \mathcal{O}^{\text{ad}}$, and which is also to be minimized:

$$J_2(\Omega) := \int_{\Omega} dx \quad (2.4)$$

All in all, we obtain a *biobjective PDE constrained shape optimization problem* w.r.t. reliability and cost which is given by

$$\begin{aligned} & \min J_1(\Omega) \\ & \min J_2(\Omega) \\ & \text{s.t. } u(\Omega) \text{ solves the state equation (2.2)} \\ & \quad \Omega \in \mathcal{O}^{\text{ad}}. \end{aligned} \quad (2.5)$$

3 Hypervolume Scalarization

Given a biobjective optimization problem (2.5), the hypervolume indicator measures the hypervolume spanned by an outcome vector $J(\Omega) \in \mathbb{R}^2$ (or a set

of outcome vectors $J(S)$, respectively) and a predetermined reference point $r \in \mathbb{R}^2$ (that should be selected such that it is dominated by all considered outcome vectors), see Figure 1. The value of the hypervolume indicator for a set of solutions $S \subset \mathcal{O}^{\text{ad}}$ and an appropriate reference point $r \in \mathbb{R}^2$ can be written as

$$H(S) = \text{Vol} \left(\bigcup_{\Omega \in S} (J(\Omega) + \mathbb{R}_+^2) \cap (r - \mathbb{R}_+^2) \right),$$

where $\mathbb{R}_+^2 = \{y \in \mathbb{R}^2: y_i \geq 0, i = 1, 2\}$ denotes the positive orthant. Note that the hypervolume indicator is only meaningful when the reference point r is chosen such that $r_i \geq J_i(\Omega)$ for $i = 1, 2$ and for all $\Omega \in S$.

In the case of a single solution $\Omega \in \mathcal{O}^{\text{ad}}$ the *hypervolume indicator* function simplifies to

$$H(\Omega) = (r_1 - J_1(\Omega)) \cdot (r_2 - J_2(\Omega)),$$

which corresponds to the area of the *hypervolume rectangle* defined by the corner points $J(\Omega)$ and r .

While in EMO algorithms the hypervolume indicator is mainly used to evaluate the quality of already computed solutions (or solution sets, respectively), it can also be used as an objective function in a single-objective scalarization of (2.5) that aims at finding solutions that maximize the dominated hypervolume. For a given reference point $r := (r_1, r_2)^\top \in \mathbb{R}^2$, the *hypervolume scalarization* for (2.5) is formulated as

$$\begin{aligned} \max \quad & H(\Omega) = (r_1 - J_1(\Omega)) \cdot (r_2 - J_2(\Omega)) \\ \text{s.t.} \quad & u(\Omega) \text{ solves the state equation (2.2)} \\ & J_1(\Omega) \leq r_1 \\ & J_2(\Omega) \leq r_2 \\ & \Omega \in \mathcal{O}^{\text{ad}}. \end{aligned} \tag{3.1}$$

The constraints $J_1(\Omega) \leq r_1$ and $J_2(\Omega) \leq r_2$ ensure that $H(\Omega)$ is non-negative. They are redundant when all feasible outcome vectors lie below the reference point, i.e., when r is chosen such that $J(\Omega) \leq r$ for all $\Omega \in \mathcal{O}^{\text{ad}}$. If this is not the case, however, these two constraints can not be omitted in general, since also points that are strictly dominated by the reference point have positive hypervolume values and are thus potential candidates for maximizing $H(\Omega)$. Note that even in this case it is sufficient to include one of these constraints, since then the other will be automatically satisfied at optimality due to the maximization of the product $H(\Omega) = (r_1 - J_1(\Omega)) \cdot (r_2 - J_2(\Omega))$. Moreover, when problem (3.1) is solved with a steepest ascent algorithm (c.f. Section 4.2 below), starting with a shape $\Omega^{(0)}$ such that $J(\Omega^{(0)}) < r$ (i.e., $J_1(\Omega^{(0)}) < r_1$ and $J_2(\Omega^{(0)}) < r_2$), then the constraints $J_1(\Omega) \leq r_1$ and $J_2(\Omega) \leq r_2$ will always be satisfied during the course of the algorithm and can hence be omitted. We conclude that while some caution is necessary with regard to the choice of the reference point r , the constraints $J_1(\Omega) \leq r_1$ and $J_2(\Omega) \leq r_2$ can in most cases be omitted. If not stated otherwise, we will hence assume in

the following that the scalarized problem shares the same feasible set as the original problem.

The following result shows that under the above assumptions the hypervolume scalarization (3.1) always generates an efficient solution. This result is not new, however, we include a proof for the sake of completeness.

Theorem 1 *An optimal solution of the hypervolume scalarization (3.1) is an efficient solution of the corresponding biobjective optimization problem (2.5).*

Proof Suppose that $\Omega^* \in \arg \max_{\mathcal{O}^{\text{ad}}} H(\Omega)$ is not efficient for (2.5). Then there exists a shape $\Omega' \in \mathcal{O}^{\text{ad}}$ such that $J_1(\Omega') \leq J_1(\Omega^*)$ and $J_2(\Omega') \leq J_2(\Omega^*)$, with at least one strict inequality. It follows that

$$\begin{aligned} H(\Omega') &= (r_1 - J_1(\Omega')) \cdot (r_2 - J_2(\Omega')) \\ &> (r_1 - J_1(\Omega^*)) \cdot (r_2 - J_2(\Omega^*)) = H(\Omega^*), \end{aligned}$$

contradicting the assumption that $\Omega^* \in \arg \max_{\mathcal{O}^{\text{ad}}} H(\Omega)$. \square

Note that Theorem 1 also follows directly from the theory of *achievement scalarizing functions* of which the hypervolume indicator is a special case (see, e.g., Wierzbicki (1986a,b); Miettinen (1998)). Indeed, since $-H(\Omega)$ is a *strongly increasing* function (i.e., $J(\Omega) \leq J(\Omega')$ implies that $-H(\Omega) < -H(\Omega')$), an optimal solution of $\max_{\Omega \in \mathcal{O}^{\text{ad}}} H(\Omega) = \min_{\Omega \in \mathcal{O}^{\text{ad}}} -H(\Omega)$ is efficient for (2.5).

Efficient numerical optimization procedures require gradient information to obtain fast convergence. In this situation, we can take advantage of the fact that the gradient of the hypervolume indicator can be determined from the gradients of the individual objective functions using the chain-rule of differentiation:

$$\nabla H(\Omega) = (J_2(\Omega) - r_2) \cdot \nabla J_1(\Omega) + (J_1(\Omega) - r_1) \cdot \nabla J_2(\Omega) \quad (3.2)$$

Observe that the gradient vector of the hypervolume indicator is a negative weighted sum of the gradients of the individual objective functions.

3.1 Hypervolume Scalarization Versus Weighted Sum Scalarization

Let $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}_+^2$ with $(\lambda_1, \lambda_2) \neq (0, 0)^\top$ be a given weight vector. Then the *weighted sum scalarization* of problem (2.5) is given by

$$\begin{aligned} \min W(\Omega) &= \lambda_1 J_1(\Omega) + \lambda_2 J_2(\Omega) \\ \text{s.t. } u(\Omega) &\text{ solves the state equation (2.2)} \\ \Omega &\in \mathcal{O}^{\text{ad}}. \end{aligned} \quad (3.3)$$

Note that the weights are usually normalized such that $\lambda_1 + \lambda_2 = 1$. We omit this normalization to simplify the notation in the following. Note that this does not affect the properties of the weighted sum scalarization. Note also that under the assumption that the reference point in the hypervolume

scalarization (3.1) is chosen sufficiently large the feasible sets of the hypervolume scalarization (3.1) and the weighted sum scalarization (3.3) coincide. Thus, both scalarizations share the advantage that the feasible set of the original multiobjective optimization problem is not modified, i.e., no additional constraints have to be added to the problem formulation. However, since the level sets of the hypervolume indicator in the objective space are non-convex, non-supported efficient solutions can be determined using the hypervolume scalarization, which is not possible using the weighted sum scalarization. Formally, the *level set of the hypervolume indicator* in the objective space at level $H(\bar{\Omega})$ is given by

$$\mathcal{L}(H(\bar{\Omega})) := \{(J_1, J_2)^\top \in \mathbb{R}^2 : (r_1 - J_1) \cdot (r_2 - J_2) \leq H(\bar{\Omega})\}.$$

The non-convexity of the hypervolume indicator as a function of the outcome vector $J(\Omega)$ is easy to see when rewriting $H(\Omega)$ as a quadratic form

$$H(\Omega) = J(\Omega)^\top \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} J(\Omega) - \begin{pmatrix} r_2 \\ r_1 \end{pmatrix}^\top J(\Omega) + r_1 r_2.$$

Since the quadratic form is negative semi-definite, $H(\Omega)$ is a concave function. Thus, the hypervolume scalarization is not limited to supported efficient solutions. Even more so, by interpreting the reference point as a scalarization parameter, every efficient solution can be obtained as an optimal solution of a hypervolume scalarization with an appropriately chosen reference point.

Corollary 2 *If $\Omega^* \in \mathcal{O}^{\text{ad}}$ is an efficient solution of (2.5), then it is optimal for (3.1) with reference point $r := J(\Omega^*)$.*

Note that this is of rather theoretical interest since this choice of the reference point assumes that the corresponding efficient solution is already known. Practical choices for reference points in the context of biobjective shape optimization are discussed in Section 4.3.

In the following, let $r = (r_1, r_2)^\top$ be a fixed reference point and let $\bar{\Omega} \in \mathcal{O}^{\text{ad}}$ be a strictly feasible solution of (3.1), i.e., $J(\bar{\Omega}) < r$. To interrelate hypervolume scalarizations with weighted sum scalarizations, we analyze the level sets and the gradient of the hypervolume indicator. To simplify the notation in the biobjective case, we denote the two sides of the hypervolume rectangle spanned by $J(\bar{\Omega}) = (J_1(\bar{\Omega}), J_2(\bar{\Omega}))^\top$ and r by $a(\bar{\Omega}) := r_1 - J_1(\bar{\Omega})$ and $b(\bar{\Omega}) := r_2 - J_2(\bar{\Omega})$, respectively, see Figure 3. Now the hypervolume induced by $J(\bar{\Omega})$ can be rewritten as

$$H(\bar{\Omega}) = a(\bar{\Omega}) \cdot b(\bar{\Omega}). \quad (3.4)$$

Using this notation, the contour line of the hypervolume indicator at the level $H(\bar{\Omega})$ can be described by the parameterization $h_{\bar{\Omega}} : \mathbb{R} \rightarrow \mathbb{R}$ given by

$$h_{\bar{\Omega}}(J_1) = \frac{(J_1(\bar{\Omega}) - J_1) b(\bar{\Omega})}{a(\bar{\Omega}) + J_1(\bar{\Omega}) - J_1} + J_2(\bar{\Omega}), \quad (3.5)$$

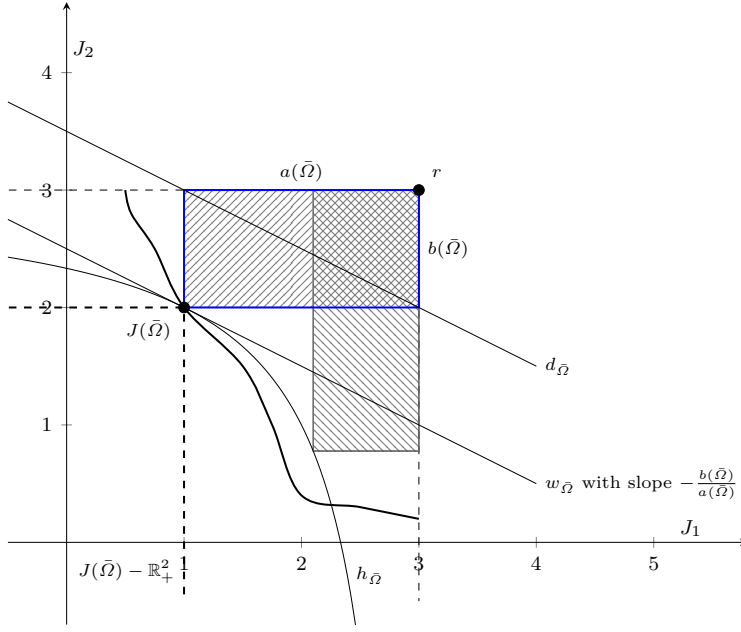


Fig. 3: A nondominated point which is suboptimal w.r.t. the hypervolume indicator

see Figure 3 for an illustration. Denoting the contour line by

$$h_{\bar{\Omega}} := \{(J_1, h_{\bar{\Omega}}(J_1))^\top : J_1 \in (-\infty, r_1)\},$$

we can conclude that all points $(J_1, h_{\bar{\Omega}}(J_1))^\top \in h_{\bar{\Omega}}$ attain the same hypervolume indicator value as $J(\bar{\Omega})$. For example, the two highlighted rectangles in Figure 3 have the same area, and hence the same hypervolume indicator value.

The derivative of $h_{\bar{\Omega}}(J_1)$ w.r.t. J_1 can be computed as

$$h'_{\bar{\Omega}}(J_1) = \frac{-b(\bar{\Omega}) a(\bar{\Omega})}{(a(\bar{\Omega}) + J_1(\bar{\Omega}) - J_1)^2}.$$

Evaluating this derivative at $J_1 = J_1(\bar{\Omega})$ yields

$$h'_{\bar{\Omega}}(J_1(\bar{\Omega})) = -\frac{b(\bar{\Omega})}{a(\bar{\Omega})},$$

which is the slope of the tangent to the level set of the hypervolume indicator at the point $J(\bar{\Omega})$, see again Figure 3. Note that this slope corresponds to the ratio of the two sides of the hypervolume rectangle, and thus also corresponds

to the slope of the diagonal passing through two opposite corners of the current hypervolume rectangle as illustrated in Figure 3, given by the equation

$$d_{\bar{\Omega}}(J_1) = -\frac{b(\bar{\Omega})}{a(\bar{\Omega})} (J_1 - J_1(\bar{\Omega})) + r_2. \quad (3.6)$$

As can be seen in Figure 3, the tangent to the level set of the hypervolume indicator at the point $J(\bar{\Omega})$ also corresponds to the contour line of the weighted sum scalarization $W(\Omega)$ with weights $(\lambda_1, \lambda_2) = (b(\bar{\Omega}), a(\bar{\Omega}))$ at this point. This contour line can be described by a function of J_1 , similar to (3.6), as

$$w_{\bar{\Omega}}(J_1) = -\frac{b(\bar{\Omega})}{a(\bar{\Omega})} \cdot J_1 + \frac{c}{a(\bar{\Omega})},$$

where $c := b(\bar{\Omega}) J_1(\bar{\Omega}) + a(\bar{\Omega}) J_2(\bar{\Omega})$ is defined by the point $J(\bar{\Omega})$. The corresponding contour line is denoted by

$$w_{\bar{\Omega}} := \{(J_1, w_{\bar{\Omega}}(J_1))^\top : J_1 \in \mathbb{R}\}.$$

As a consequence, the contour lines of both scalarizations, the hypervolume indicator and the weighted sum function, have the same slope at the point $J(\bar{\Omega})$, and hence both scalarizations share the same direction of largest improvement. Indeed, if we consider the gradient $\nabla H(\bar{\Omega})$ of the hypervolume indicator at the point $J(\bar{\Omega})$ (which is the steepest ascent direction, see (3.2)) we obtain

$$\begin{aligned} \nabla H(\bar{\Omega}) &= (J_2(\bar{\Omega}) - r_2) \cdot \nabla J_1(\bar{\Omega}) + (J_1(\bar{\Omega}) - r_1) \cdot \nabla J_2(\bar{\Omega}) \\ &= -b(\bar{\Omega}) \nabla J_1(\bar{\Omega}) - a(\bar{\Omega}) \nabla J_2(\bar{\Omega}). \end{aligned} \quad (3.7)$$

Similarly, the steepest descent direction for the weighted sum scalarization (3.3) with weights $(\lambda_1, \lambda_2) = (b(\bar{\Omega}), a(\bar{\Omega}))$ at the point $J(\bar{\Omega})$ can be computed as

$$\begin{aligned} -\nabla W(\bar{\Omega}) &= -(\lambda_1 \nabla J_1(\bar{\Omega}) + \lambda_2 \nabla J_2(\bar{\Omega})) \\ &= -b(\bar{\Omega}) \nabla J_1(\bar{\Omega}) - a(\bar{\Omega}) \nabla J_2(\bar{\Omega}). \end{aligned} \quad (3.8)$$

Hence, in an iterative algorithm the direction of steepest ascent of the hypervolume indicator (which is to be maximized) corresponds to the direction of steepest descent w.r.t. the weighted sum scalarization (which is to be minimized) with weights $(\lambda_1, \lambda_2) = (b(\bar{\Omega}), a(\bar{\Omega}))$. Note that these *associated weights* are determined by the ratio of the sides of the hypervolume rectangle. The above analysis implies that if an iterative ascent algorithm for the hypervolume scalarization (3.1) is initialized with a strictly feasible starting solution $\Omega^{(0)} \in \mathcal{O}^{\text{ad}}$, i.e., $J(\Omega^{(0)}) < r$, then moving in the direction of steepest ascent ensures that the next iterate $\Omega^{(1)}$ is also strictly feasible irrespective of the step length, i.e., $J_1(\Omega^{(1)}) < r_1$ and $J_2(\Omega^{(1)}) < r_2$. Hence, the constraints $J_1(\Omega) \leq r_1$ and $J_2(\Omega) \leq r_2$ can be omitted in (3.1) when starting with a strictly feasible starting solution. Note that, for any strictly feasible solution $\bar{\Omega} \in \mathcal{O}^{\text{ad}}$, the side lengths of the hypervolume rectangle $a(\bar{\Omega})$ and $b(\bar{\Omega})$ correspond to the slackness of the constraints $J_1(\bar{\Omega}) \leq r_1$ and $J_2(\bar{\Omega}) \leq r_2$. This

is reflected in the gradient direction $\nabla H(\bar{\Omega})$ at $J(\bar{\Omega})$ that is composed of a weighted sum of the gradients of the individual objective functions $\nabla J_1(\bar{\Omega})$ and $\nabla J_2(\bar{\Omega})$. The ratio of the side lengths of the hypervolume rectangle defines the relative contribution of $\nabla J_1(\bar{\Omega})$ and $\nabla J_2(\bar{\Omega})$ to the ascent direction $\nabla H(\bar{\Omega})$. If this ratio is very large (very small, respectively), which is the case when one side of the hypervolume rectangle is considerably larger than the other, then the opposite objective function is emphasized in the computation of $\nabla H(\bar{\Omega})$.

3.2 Hypervolume Scalarization Versus MO Descent Algorithms

Multiobjective gradient descent algorithms were proposed in Fliege and Svaiter (2000); Désidéri (2012); Fliege et al. (2019). In the minimization case, they can be interpreted as iterative optimization processes with search direction

$$d^{(k)} := - \left(\omega_1^{(k)} \nabla J_1(\Omega^{(k)}) + \omega_2^{(k)} \nabla J_2(\Omega^{(k)}) \right)$$

in iteration k with weights $\omega_1^{(k)}, \omega_2^{(k)} \geq 0$. The search direction $d^{(k)}$ is thus induced by a positive linear combination of the gradients of the individual objective functions with weights $\omega_1^{(k)}, \omega_2^{(k)}$. These weights are chosen adaptively to achieve the *steepest* descent w.r.t. all objectives in each iteration. This is in contrast to applying a single-objective gradient descent algorithm to the weighted sum scalarization (3.3), in which the weights are predefined and thus fixed throughout all iterations of the optimization procedure. We argue that performing gradient ascent steps w.r.t. the hypervolume indicator is somewhat similar to applying a multiobjective gradient descent algorithm: The weights $\omega_1^{(k)} = b(\Omega^{(k)})$ and $\omega_2^{(k)} = a(\Omega^{(k)})$ of the gradients of the individual objective functions are chosen adaptively, depending on the side lengths of the hypervolume rectangle induced by the current iterate, c.f. (3.7).

While each iterate $\Omega^{(k)}$ in a multiobjective gradient descent algorithm dominates all previous iterates $\Omega^{(k-\ell)}$ for all $\ell = 1, \dots, k$, i.e., $J(\Omega^{(k)}) \leq J(\Omega^{(k-\ell)})$, this is in general not the case when applying a single-objective ascent algorithm to the hypervolume scalarization or a single-objective descent algorithm to the weighted sum scalarization, respectively. Indeed, only single-objective improvements can be guaranteed with respect to the hypervolume indicator and the weighted sum objective, i.e., $H(\Omega^{(k)}) \geq H(\Omega^{(k-\ell)})$ and $\lambda^\top J(\Omega^{(k)}) \leq \lambda^\top J(\Omega^{(k-\ell)})$, respectively. This can be seen in Figure 3 when comparing the dominance cone $J(\bar{\Omega}) - \mathbb{R}_+^2$ (which contains the image of the next iterate in a multiobjective descent algorithm), the set of improving points for the hypervolume indicator given by $h_{\bar{\Omega}} - \mathbb{R}_+^2 = \{y \in \mathbb{R}^2 : (r_1 - y_1) \cdot (r_2 - y_2) \geq H(\bar{\Omega})\}$, and the set of improving points for the weighted sum objective given by $w_{\bar{\Omega}} - \mathbb{R}_+^2 = \{y \in \mathbb{R}^2 : \lambda^\top y \leq J(\bar{\Omega})\}$ (with $\lambda^\top = (b(\bar{\Omega}), a(\bar{\Omega}))$). Since the dominance cone $J(\bar{\Omega}) - \mathbb{R}_+^2$ is a proper subset of the respective sets of improving points, multiobjective gradient descent algorithms are more restrictive regarding the choice of the next iterate. The weighted sum scalarization is in this

regard the most flexible approach, allowing for the deterioration of individual objective functions as long as the weighted sum value improves. In this respect, we can say that the hypervolume indicator function compromises between the other two approaches.

3.3 Contour Lines of the Hypervolume Indicator

The regions of the Pareto front which can be attained using a scalarization method are determined by the level sets and contour lines of the scalarizing function. Consider, for example, the contour lines of the weighted sum scalarization (hyperplanes) and the weighted Chebyshev scalarization (hyperboxes with weight-dependent aspect ratio), respectively. While the former is restricted to supported nondominated points, the latter has the potential to determine all nondominated points, including the non-supported ones. The contour lines of the hypervolume indicator for a given reference point $r \in \mathbb{R}^2$ are illustrated in Figure 4a. It can be seen that their shape strongly depends on their Hausdorff distance from the reference point. The closer the reference point is to a nondominated outcome vector in at least one component, the more the corresponding contour line resembles the contour line of a weighted l_∞ function, i.e., a weighted Chebyshev scalarization. On the other hand, the further the reference point is from a nondominated outcome vector (in all components), the more the contour line resembles that of an l_1 function, that is, a weighted sum scalarization with equal weights. More formally, considering again the parameterization of the contour line of the hypervolume indicator w.r.t. J_1 , see (3.5), and assuming for now that $J_1 < J_1(\bar{\Omega}) < r_1$, we can see that

$$\lim_{r \rightarrow J(\bar{\Omega})} h_{\bar{\Omega}}(J_1) = \lim_{r \rightarrow J(\bar{\Omega})} \frac{\overbrace{(J_1(\bar{\Omega}) - J_1)}^{\text{constant}} \overbrace{b(\bar{\Omega})}^{\rightarrow 0}}{\underbrace{a(\bar{\Omega})}_{\rightarrow 0} + \underbrace{J_1(\bar{\Omega}) - J_1}_{\text{constant}}} + J_2(\bar{\Omega}) = J_2(\bar{\Omega})$$

This implies that for $r \rightarrow J(\bar{\Omega})$ the contour line $h_{\bar{\Omega}}$ converges point-wise to a horizontal line passing through r when $J_1 < J_1(\bar{\Omega})$. By using a parameterization of $h_{\bar{\Omega}}$ w.r.t. J_2 it can be shown that the contour line $h_{\bar{\Omega}}$ converges point-wise to a vertical line passing through r when $J_2 < J_2(\bar{\Omega})$.

Similarly, when $r = (r_1, r_1)^\top$ (i.e., $r_2 = r_1$) and $r_1 \rightarrow \infty$ we obtain for $J_1 \neq J_1(\bar{\Omega})$ that

$$\lim_{r_1 \rightarrow \infty} h_{\bar{\Omega}}(J_1) = \lim_{r_1 \rightarrow \infty} \frac{\overbrace{(J_1(\bar{\Omega}) - J_1)}^{\text{constant}} \overbrace{b(\bar{\Omega})}^{\rightarrow \infty}}{\underbrace{a(\bar{\Omega})}_{\rightarrow \infty} + \underbrace{J_1(\bar{\Omega}) - J_1}_{\text{constant}}} + \underbrace{J_2(\bar{\Omega})}_{\text{constant}} = -J_1 + J_1(\bar{\Omega}) + J_2(\bar{\Omega}),$$

and hence $h_{\bar{\Omega}}$ converges point-wise to a line with slope -1 and passing through $J(\bar{\Omega})$. The contour lines passing through a given point $J(\bar{\Omega})$, that are generated by different reference points, are illustrated in Figure 4b.

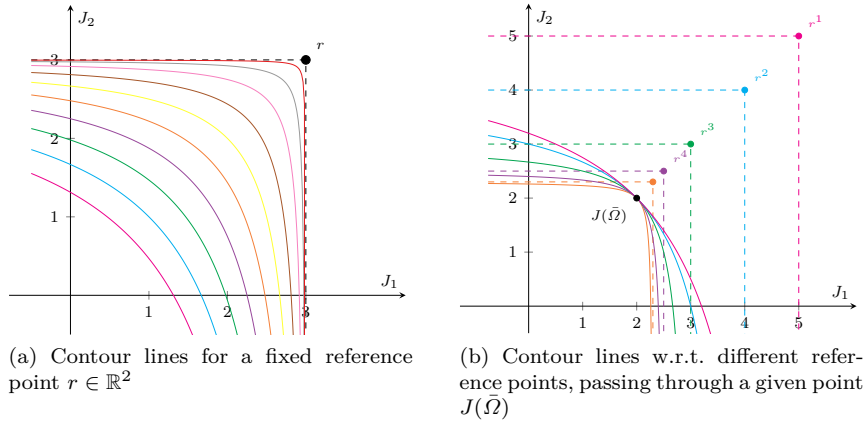


Fig. 4: Contour lines of the hypervolume indicator

4 Numerical Implementation

The hypervolume scalarization (3.1) of the biobjective shape optimization problem (2.5) is solved using an iterative ascent algorithm in the framework of a ‘*first discretize, then optimize*’ approach. The discretization of the shape optimization problem is based on a finite element approach (see, e.g., Braess (2007)). Consequently, the state equation (2.2) and the objective functions given in (2.3) and (2.4) are evaluated on a finite element grid. In this section details of the numerical implementation are presented including the discretization approach using structured finite element grids, the determination of ascent directions for the hypervolume scalarization, and the update scheme for the current shape. Finally, the choice of the reference point is discussed.

4.1 Finite Element Discretization

The component to be optimized (i.e., the current shape) is discretized via finite elements on a structured grid in the following way. Consider a rectangular domain $\hat{\Omega}$ with width $\hat{\Omega}_l$ and height $\hat{\Omega}_h$ that entirely covers the initial shape $\Omega^{(0)}$ (see Figure 2 for an illustration). The domain is discretized by a regular triangular grid denoted by \hat{T} with n_x and n_y grid points in x - and y -direction, respectively. The corresponding mesh sizes are given by $m_x = \hat{\Omega}_l / (n_x - 1)$ and $m_y = \hat{\Omega}_h / (n_y - 1)$. The mesh sizes should be chosen such that $m_x \approx m_y$ in order to avoid distorted finite elements. In a next step, as visualized in Figure 5, the boundary $\partial\Omega^{(0)}$ of the initial shape is superimposed onto the grid and the closest grid points are moved onto the boundary in a controlled way, such that degenerate elements do not occur. More precisely, it is guaranteed that the maximum angle condition (Babuska and Aziz (1976)) is satisfied. The adapted grid is denoted by T_0^∞ . The computations are performed only on

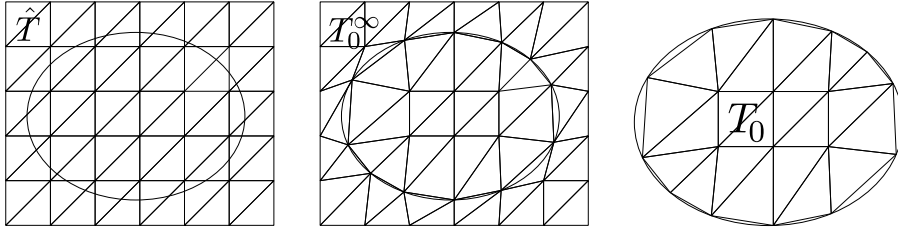


Fig. 5: Exemplary adaption of the finite element grid to a shape $\Omega^{(0)}$ in iteration $k = 0$

that part of the grid that represents $\Omega^{(0)}$ which is called the *active grid* and denoted by T_0 , see Figure 5.

In the iterative optimization process (see Section 4.2 for more details) the active grid T_{k+1} of the next iteration, $k = 0, 1, 2, \dots$, representing the updated shape $\Omega^{(k+1)}$, is generated by starting again from \hat{T} . Consequently, the connectivity is maintained throughout all iterations. Rather than going through all elements (as required in the first iteration) we now incorporate the information from the previous mesh T_k^∞ . In the ‘updating procedure’ it is only necessary to iterate over elements in a certain neighborhood of previous boundary elements, and adapt them if necessary. Note that a complete re-meshing is necessary only when the next iterate $\Omega^{(k+1)}$ differs too much from $\Omega^{(k)}$, which is usually only the case when the update step exceeds the mesh size in at least one grid point. This property and the persistent connectivity lead to a significant speed up compared to standard re-meshing techniques. As stated above, in most iterations the majority of the elements inside the component remain unchanged. This leads to an additional advantage over both standard re-meshing and mesh morphing approaches. In regions where elements are not changed, the entries of the governing PDE are also not changed. Therefore, a full re-assembly of the system matrix and the right hand side is avoided and only entries corresponding to modified elements have to be updated.

In the following, we distinguish between the continuous shape $\Omega^{(k)}$ and its (approximative) representation on the finite element grid. The discretized shape, induced by the active grid T_k , is denoted by $\Omega_{n_x \times n_y}^{(k)}$. This implicates that the gradient of the hypervolume indicator ∇H is evaluated w.r.t. the discretized shape, i.e., $\nabla H(\Omega_{n_x \times n_y}^{(k)})$ is defined on the grid points given by T_k .

4.2 Iterative Ascent Algorithm

Ascent direction and update scheme. In the following we develop an iterative update scheme for the biobjective shape optimization problem (2.5). As stated above, the finite element discretization of the component is based on a regular grid, where only the grid points that are close to the boundary of the current

shape $\Omega^{(k)}$ are adapted to the boundary of the component. All finite elements that are not adjacent to the boundary of the current shape $\Omega^{(k)}$ remain unchanged and form a subset of the regular grid. Whenever the component is updated, we thus want to modify the boundary points of the component without moving inner or outer grid points. For this purpose we consider the grid points lying on the boundary of the shape $\Omega^{(k)}$, or rather of its finite element representation $\Omega_{n_x \times n_y}^{(k)}$. Since these grid points are induced by the active grid T_k , we call them *boundary points* and denote them by ∂T_k .

The discretized shape $\Omega_{n_x \times n_y}^{(k)}$ in iteration k is updated by determining a search direction and calculating a step length. For a given search direction $d^{(k)}$ and step length t_k , we define the update of the component by the expression

$$\Omega_{n_x \times n_y}^{(k+1)} = \Omega_{n_x \times n_y}^{(k)} \oplus t_k \cdot d^{(k)}. \quad (4.1)$$

The operation \oplus is defined as follows: First, the boundary points are moved by $t_k \cdot d^{(k)}$ resulting in

$$\partial T_k' := \partial T_k + t_k \cdot d^{(k)}.$$

Then $\partial T_k'$ is fitted by cubic splines, resulting in the updated shape $\Omega^{(k+1)}$. Finally, the new active grid T_{k+1} is generated based on $\partial \Omega^{(k+1)}$ as detailed in Section 4.1. This defines $\Omega_{n_x \times n_y}^{(k+1)}$ in (4.1). Note that the boundary points ∂T_{k+1} of iteration $(k+1)$ do not have to coincide with $\partial T_k'$ since T_{k+1} is generated based on $\partial \Omega^{(k+1)}$, which is defined by cubic splines.

The search direction $d^{(k)}$ is computed based on the gradient of the hypervolume indicator function $\nabla H(\Omega_{n_x \times n_y}^{(k)})$. However, due to the discretization $\nabla H(\Omega_{n_x \times n_y}^{(k)})$ is usually largely irregular and needs to be smoothed to avoid zigzagging boundaries and overfitted spline representations of the next iterate. To obtain a smooth deformation field for the boundary points, we follow the approach suggested in Schmidt and Schulz (2009) and use a *Dirichlet-to-Neumann* map to compute a smoothed gradient. Let $\tilde{\nabla} H(\Omega_{n_x \times n_y}^{(k)})$ denote the smoothed gradient. Then

$$d^{(k)} := \tilde{\nabla} H(\Omega_{n_x \times n_y}^{(k)}) \Big|_{\partial T_k} \quad (4.2)$$

denotes its values on the boundary points. See Figure 6 for a visual comparison of the gradient $\nabla H(\Omega_{n_x \times n_y}^{(k)})$ and the smoothed gradient $\tilde{\nabla} H(\Omega_{n_x \times n_y}^{(k)})$ of an exemplary shape.

The step length t_k in direction $d^{(k)}$ is determined according to the Armijo-rule (see, e.g., Geiger and Kanzow, 1999; Bazaraa et al., 2006). For given parameters $\sigma, \beta \in (0, 1)$ we set

$$t_k := \max_{\ell \in \mathbb{N}_0} \left\{ \beta^\ell : H(\Omega_{n_x \times n_y}^{(k)} \oplus \beta^\ell d^{(k)}) \geq H(\Omega_{n_x \times n_y}^{(k)}) + \sigma \beta^\ell \nabla H(\Omega_{n_x \times n_y}^{(k)})^\top d^{(k)} \right\}. \quad (4.3)$$

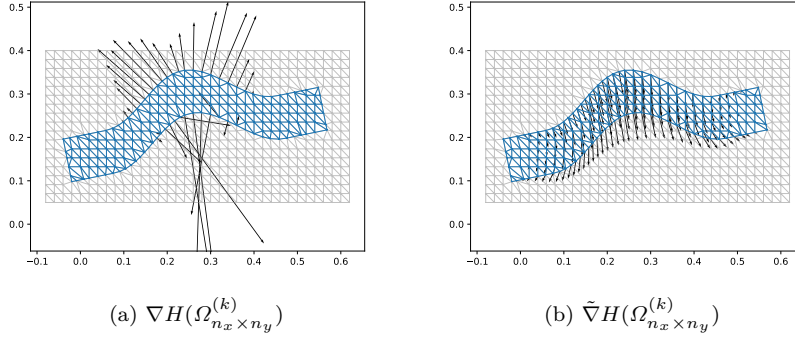


Fig. 6: Unsmoothed (left) and smoothed (right) gradient of the hypervolume indicator

The complete iterative ascent algorithm using smoothed gradients is summarized in Algorithm 1. As stopping conditions, an upper bound K on the number of iterations, an upper bound on the number of iterations ℓ in the evaluation of the Armijo rule (4.3), and a lower bound ε on the Frobenius norm of the search direction are used.

Data: starting solution $\Omega_{n_x \times n_y}^{(0)}$, reference point r ; Armijo parameters $\sigma, \beta \in (0, 1)$, accuracy $\varepsilon > 0$, maximal number of iterations $K \in \mathbb{N}$;

Result: approximate solution of the hypervolume problem $\Omega_{n_x \times n_y}^*$;

Set $k := 0$, calculate $\tilde{\nabla} H(\Omega_{n_x \times n_y}^{(0)})|_{\partial T_k}$;

while $\|\tilde{\nabla} H(\Omega_{n_x \times n_y}^{(k)})|_{\partial T_k}\| < \varepsilon$ and $k \leq K$ **do**

Compute a search direction $d^{(k)} := \tilde{\nabla} H(\Omega_{n_x \times n_y}^{(k)})|_{\partial T_k}$;

Compute a step length $t_k \in (0, 1]$ according to the Armijo-rule (4.3);

$\Omega_{n_x \times n_y}^{(k+1)} := \Omega_{n_x \times n_y}^{(k)} \oplus t_k \cdot d^{(k)}$;

$k := k + 1$;

end

$\Omega_{n_x \times n_y}^* := \Omega_{n_x \times n_y}^{(k)}$;

Algorithm 1: Iterative optimization scheme for hypervolume maximization

Re-scaling the search direction. In order to avoid re-meshing operations whenever possible, it is often advantageous to re-scale the search direction such that

the grid T_{k+1}^∞ can be computed by performing a simple grid adaption (c.f. Section 4.1). Indeed, when for example for a large gradient (e.g., resulting from large given loads) the step $t_k = 1$ is not feasible, this leads to numerical problems when implementing Algorithm 1. To avoid such situations, the search direction $d^{(k)}$ is re-scaled in the following way: First, the norm $\|d_{ij}^{(k)}\|_2$ is calculated for all grid points (i, j) , $i = 1, \dots, n_x$ and $j = 1, \dots, n_y$, where $d_{ij}^{(k)}$ is the entry of $d^{(k)}$ for grid point (i, j) . Since $d^{(k)}$ is only defined at the boundary points, $d_{ij}^{(k)}$ is set to zero for grid points (i, j) that are not related to the boundary points. When $\max_{i,j} \|d_{ij}\|_2$ is larger than a predefined percentage p_{\max} of the minimum mesh size $\min\{m_x, m_y\}$, with $p_{\max} < 1$, then the search direction is re-scaled by the factor $s(p_{\max})$ given by

$$s(p) = \frac{p}{\max_{i,j} \|d_{ij}\|_2} \min\{m_x, m_y\}. \quad (4.4)$$

By computing a step length $t_k \in (0, 1]$ it is then guaranteed that the boundary points do not move more than the mesh size, and a grid adaption can be performed.

Furthermore, numerical tests have shown that near an efficient solution the search direction $d^{(k)}$ becomes very small, thus slowing down the potential improvement while the unsmoothed gradient $\nabla H(\Omega_{n_x \times n_y}^{(k)})$ may still be large. Then the search direction is re-scaled when $\max_{i,j} \|d_{ij}\|_2$ is less than a given percentage $p_{\min} \in (0, 1)$ of the mesh size. In this case, the search direction is multiplied by $s(p_{\min})$.

Grid refinement. In order to accelerate the convergence to the Pareto front, it may be useful to start the optimization on a coarse grid and later continue on a finer grid. Since for numerical reasons each update step is restricted to the mesh size, we expect that a coarse grid allows for larger steps in early stages of the optimization procedure. Moreover, the evaluation of objective function values and gradients on the coarse grid is generally faster. On the other hand, with a finer resolution of the grid the accuracy of the approximation of objective function values and gradients is improved. Thus, near an efficient solution large steps are not necessary and a finer grid may result in a better approximation of the shape.

4.3 Choice and Variation of the Reference Point

For the iterative ascent algorithm (Algorithm 1) a feasible starting solution $\Omega_{n_x \times n_y}^{(0)}$ and a reference point $r \in \mathbb{R}^2$ with $r > J(\Omega_{n_x \times n_y}^{(0)})$ are needed as input. Since in general neither the Pareto front nor efficient shapes are known beforehand, the reference point can only be chosen based on a known feasible shape $\Omega_{n_x \times n_y}^{(0)}$ that may actually be far from the Pareto front. To ensure that the hypervolume scalarization (3.1) is well defined and that the constraints

$J_1(\Omega) \leq r_1$ and $J_2(\Omega) \leq r_2$ are redundant throughout the execution of Algorithm 1 (c.f. Section 3 for a discussion of this issue,) we select the reference point r such that

$$r = \left(J_1(\Omega_{n_x \times n_y}^{(0)}) + \delta_1, J_2(\Omega_{n_x \times n_y}^{(0)}) + \delta_2 \right)^\top \quad (4.5)$$

with positive parameters $\delta_1, \delta_2 > 0$.

The impact of the reference point on the hypervolume indicator was already discussed in Auger et al. (2009). They derive properties of *optimal μ distributions* depending on the choice of the reference point, i.e., of representative subsets of the Pareto front of cardinality μ maximizing the joint hypervolume. Moreover, they derive an explicit lower bound for the reference point guaranteeing that, under appropriate conditions, the extreme solutions are contained in the representation.

In order to take advantage of the fact that, at least theoretically, the complete Pareto front can be generated by varying the reference point in the hypervolume scalarization (3.1) (see Corollary 2), we suggest a systematic approach to approximate the Pareto front of problem (2.5) by repeatedly solving problem (3.1) with *different* reference points. The ultimate goal is the determination of a *Pareto front generating* reference set (PFG reference set for short). In order to formally introduce the concept of a PFG reference set, let $\mathcal{Y}_{\text{ND}} \subset \mathbb{R}^2$ denote the Pareto front in the objective space and let \mathcal{Y}_{PND} denote the set of *properly* nondominated outcome vectors according to Geoffrion (1968) (briefly speaking this is the set of all nondominated outcome vectors that have bounded trade-offs). Moreover, let $\mathcal{Y}_{\text{H}}(r)$ denote the optimal outcome vector obtained from solving problem (3.1), and let $\mathcal{Y}_{\text{H}}(R) := \bigcup_{r \in R} \mathcal{Y}_{\text{H}}(r)$ denote the set of all optimal outcome vectors obtained over all reference points in the set $R \subseteq \mathbb{R}^2$, called *reference set*.

Definition 3 A set $R \subseteq \mathbb{R}^2$ is called Pareto front generating (PFG) reference set if and only if $\mathcal{Y}_{\text{ND}} = \mathcal{Y}_{\text{H}}(R)$. Moreover, we denote a reference set $R \subseteq \mathbb{R}^2$ as pPFG reference set, if and only if it is generating the set of properly nondominated outcome vectors, i.e., $\mathcal{Y}_{\text{PND}} = \mathcal{Y}_{\text{H}}(R)$.

Note that a PFG reference set always exists since, for example, $R = \mathcal{Y}_{\text{ND}}$ is a valid selection, c.f. Corollary 2. Note also that a PFG reference set may contain redundant reference points that can be omitted without losing the PFG property.

From a practical point of view, we can not expect to solve the hypervolume scalarization (3.1) to global optimality and thus only aim at an approximation of the Pareto front \mathcal{Y}_{ND} . Towards this end, we suggest to choose the reference set R based on varying values of $\delta = (\delta_1, \delta_2) > (0, 0)$ in (4.5). More precisely, we suggest to set $(\delta_1, \delta_2) = (\xi, \frac{1}{\xi})$ with $\xi > 0$ in (4.5), i.e.,

$$r(\xi) := \left(J_1(\Omega_{n_x \times n_y}^{(0)}) + \xi, J_2(\Omega_{n_x \times n_y}^{(0)}) + \frac{1}{\xi} \right)^\top \quad (4.6)$$

and set $R_H := \{r(\xi) \in \mathbb{R}^2: \xi > 0\}$. This ensures that the starting solution $\Omega_{n_x \times n_y}^{(0)}$ is strictly feasible for the hypervolume scalarization (3.1) for all choices of $r(\xi) \in R_H$, i.e., $J(\Omega_{n_x \times n_y}^{(0)}) < r(\xi)$ for all $\xi > 0$. This implies that $R_H \subset J(\Omega_{n_x \times n_y}^{(0)}) + \mathbb{R}_{>}^2$ with $\mathbb{R}_{>}^2 = \{y \in \mathbb{R}^2: y_i > 0, i = 1, 2\}$. Hence, $\mathcal{Y}_{\text{ND}} \cap R_H = \emptyset$, i.e., the Pareto front lies completely ‘below’ the reference set.

For reference points from the reference set R_H the result of Theorem 1 can be strengthened.

Lemma 4 *An optimal solution of the hypervolume scalarization (3.1) w.r.t. a reference point from the set R_H is a properly efficient solution of the corresponding biobjective optimization problem (2.5).*

Proof Consider an optimal solution of the hypervolume scalarization (3.1) $y = J(\bar{\Omega}) \in \mathcal{Y}_H(\bar{r})$ for an arbitrary reference point $\bar{r} \in R_H$. Since $J(\Omega_{n_x \times n_y}^{(0)}) < \bar{r}$ we know that for this reference point $H(\Omega_{n_x \times n_y}^{(0)}) > 0$, and hence also $H(\bar{\Omega}) > 0$. Therefore, the hypervolume rectangle induced by y and \bar{r} must have positive side lengths $a(\bar{\Omega}) > 0$ and $b(\bar{\Omega}) > 0$. This implies that y is locally optimal for a weighted sum scalarization with weights $\lambda_1 = b(\bar{\Omega}) > 0$ and $\lambda_2 = a(\bar{\Omega}) > 0$ since this has the same cone of improving directions as the hypervolume indicator in this point (see Section 3.1 for a detailed derivation). It follows that $y \in \mathcal{Y}_{\text{PND}}$. \square

Note that Lemma 4 is in line with the results of Auger et al. (2009) who showed that, irrespective of the choice of the reference point, nondominated points with unbounded trade-off are not part of what they call ‘optimal μ distributions’, that is, of subsets of points on the Pareto front that maximize the joint hypervolume.

Note also that by varying $\xi > 0$ different reference points are obtained that lie along a hyperbola in the objective space, see Figure 7 for an illustration. Indeed, for two parameter values $0 < \xi_1 < \xi_2$ we have that $r_1(\xi_1) < r_1(\xi_2)$ and $r_2(\xi_1) > r_2(\xi_2)$. Hence, the hypervolume rectangle spanned by $J(\Omega_{n_x \times n_y}^{(0)})$ and $r(\xi_1)$ cannot be contained in the hypervolume rectangle spanned by $J(\Omega_{n_x \times n_y}^{(0)})$ and $r(\xi_2)$ and vice versa. However, for all associated hypervolume rectangles it holds that $a(\Omega_{n_x \times n_y}^{(0)}) \cdot b(\Omega_{n_x \times n_y}^{(0)}) = 1$.

While we can not guarantee that the set R_H is a PFG reference set in general, we argue that the chances are high that the reference set R_H induces solutions with varying objective values. This is validated by the numerical experiments described in Section 5 below. From a theoretical perspective, we argue that at least for convex problems the reference set R_H is pPFG, i.e., the properly efficient set can be generated by varying the reference points in R_H .

Theorem 5 *Suppose that the biobjective optimization problem (2.5) is convex, i.e., all efficient solutions of (2.5) are supported efficient solutions. Then $\mathcal{Y}_{\text{PND}} = \mathcal{Y}_H(R_H)$, i.e., R_H is a pPFG reference set.*

Proof First, recall that Lemma 4 implies that $\mathcal{Y}_H(R_H) \subseteq \mathcal{Y}_{\text{PND}}$. Thus, it remains to be shown that $\mathcal{Y}_{\text{PND}} \subseteq \mathcal{Y}_H(R_H)$.

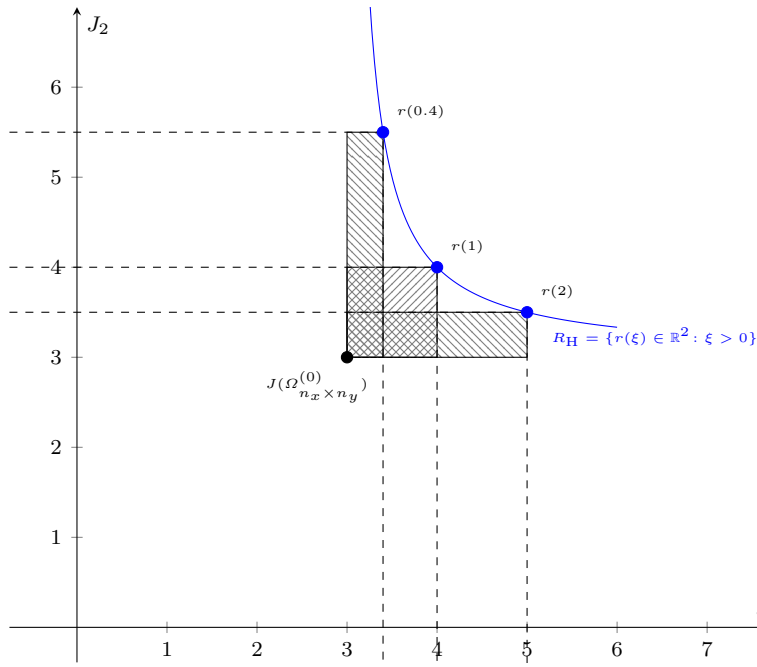


Fig. 7: Reference set R_H defining a hyperbola in the objective space

Now let $y = J(\bar{\Omega}) \in \mathcal{Y}_{\text{PND}}$ be properly efficient. Then there exist weights $\lambda_1, \lambda_2 > 0$ such that $\bar{\Omega}$ is optimal for the associated weighted sum scalarization (3.3), see, e.g., Ehrgott (2005); Geoffrion (1968). Now define a rectangle with sides $a := \lambda_2 > 0$ and $b := \lambda_1 > 0$ and let $r' := J(\bar{\Omega}) + (a, b)^\top$. It is easy to see that the hypervolume rectangle spanned by $J(\bar{\Omega})$ and r' is precisely the rectangle with side lengths a and b . From the discussion in Section 3.1 we can conclude that $\bar{\Omega}$ is optimal for the hypervolume scalarization (3.1) with reference point r' since in this situation the set of improving points w.r.t. the hypervolume indicator is completely contained in that of the weighted sum (which is empty in this case). See Figure 8 for an illustration of the respective contour lines. Moreover, $J(\bar{\Omega})$ is the unique optimal outcome vector of the hypervolume scalarization (3.1) in this case. Since this property only relies on the ratio $\frac{b}{a}$ of the side lengths of the hypervolume rectangle, the reference point can be moved along the half-line starting at $J(\bar{\Omega})$ and passing through r' until it intersects the set R_H . This intersection point exists and can be determined as the positive solution ξ of the system

$$\begin{aligned} J_1(\bar{\Omega}) + \tau a &= J_1(\Omega_{n_x \times n_y}^{(0)}) + \xi \\ J_2(\bar{\Omega}) + \tau b &= J_2(\Omega_{n_x \times n_y}^{(0)}) + \frac{1}{\xi} \end{aligned}$$

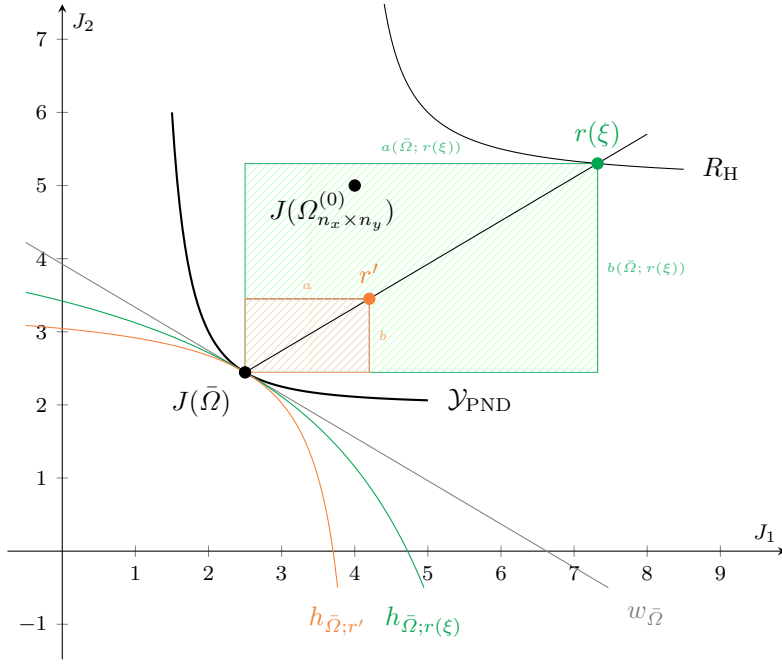


Fig. 8: Identification of a reference point $r(\xi) \in R_H$ that generates a given outcome vector $J(\bar{\Omega})$

which yields

$$\xi^2 - \xi \underbrace{\left((J_1(\bar{\Omega}) - J_1(\Omega_{n_x \times n_y}^{(0)})) - \frac{a}{b}(J_2(\bar{\Omega}) - J_2(\Omega_{n_x \times n_y}^{(0)})) \right)}_{=:c} - \frac{a}{b} = 0.$$

We obtain two solutions for ξ , namely

$$\xi_{1,2} = \frac{c}{2} \pm \sqrt{\frac{c^2}{4} + \frac{a}{b}}.$$

Since $\frac{a}{b} > 0$ we have that $(\frac{c^2}{4} + \frac{a}{b})^{\frac{1}{2}} > |\frac{c}{2}|$ and hence exactly one of the two solutions for ξ is positive, irrespective of the sign of c . This solution yields the sought reference point $r(\xi) \in R_H$, and we can conclude that $\mathcal{Y}_{\text{PND}} \subseteq \mathcal{Y}_H(R_H)$. \square

Note that even though Theorem 5 is restricted to convex problems, the set R_H may be pPFG also for non-convex problems. This depends on the degree of non-convexity on one hand, and on the distance of the reference set R_H from the Pareto front in the objective space on the other hand, where the latter is defined by the image of the starting solution $J(\Omega_{n_x \times n_y}^{(0)})$.

5 Numerical Tests

We consider a 2D test case of a ceramic rod consisting of Beryllium oxide (BeO). The rod is 0.6 m long and has a height of 0.1 m. It is fixed on one side while tensile load is applied on the other side, modeling a horizontal load transfer. On the fixed boundary Ω_D of the rod the displacement field $u(x)$ is zero, i.e., here Dirichlet boundary conditions hold, c.f. (2.2). The part $\Omega_{N_{\text{fixed}}}$ denotes the boundary on which the surface loads g are applied, i.e., where Neumann boundary conditions hold. The remaining parts of the boundary, denoted by $\Omega_{N_{\text{free}}}$, can be modified according to the optimization objectives whereas the boundaries Ω_D and $\Omega_{N_{\text{fixed}}}$ are fixed during the optimization. In particular, the corner points cannot be moved. In Figure 2 the rod and the decomposition of its boundary into Ω_D , $\Omega_{N_{\text{fixed}}}$ and $\Omega_{N_{\text{free}}}$ are illustrated.

As we consider the same ceramic material BeO as in Doganay et al. (2019), we use the same material parameter setting in order to evaluate the state equation (2.2) and the PoF objective (2.3). In particular, Young's modulus is set to $E = 320$ GPa, Poisson's ratio to $\nu = 0.25$, the ultimate tensile strength to 140 MPa and the Weibull parameter to $m = 5$. The surface loads, acting on $\Omega_{N_{\text{fixed}}}$, are set to $g = (10^8, 0)^\top$ Pa and the volume force density is set to $f = (0, 1000)^\top$ Pa. To speed up the numerical evaluation of the state equation, an improved discretization approach based on regular grids is employed, see Section 4.1 for more details and Figure 9 for an illustration of the considered shape. In contrast to Doganay et al. (2019), volume forces that model, e.g., gravity, are included in the numerical evaluation.

We choose a bent rod as shown in Figure 2 as the initial shape (i.e., as the starting solution) for the optimization that is implemented using Algorithm 1. Note that this shape is clearly not efficient. Indeed, in this situation the efficient shapes are straight rods of different width, trading off between small probabilities of failure and high volumes on one hand, and high probabilities of failure and low volumes on the other hand. This simple setting has the advantage that the optimization results can be validated and compared with the 'true' Pareto front.

The rectangular domain $\hat{\Omega}$ that contains all admissible shapes \mathcal{O}^{ad} is given by a rectangle of width $\hat{\Omega}_l = 0.7$ m and height $\hat{\Omega}_h = 0.35$ m. The initial shape is placed in the interior of this rectangle. Moreover, the whole setting is slightly rotated so that the boundaries of the initial shape can be represented by interpolating cubic splines. The domain $\hat{\Omega}$ is discretized by a triangular grid with $n_x \times n_y$ grid points as described in Section 4.1. Hereby, we determine the number of grid points such that a prescribed mesh size of $m_x = m_y \approx 0.02$ m is realized. In our setting, this results in a discretization with 36×18 grid points, see Figure 9. Note that this shape representation has more degrees of freedom than the B-spline representation in Doganay et al. (2019). This significantly increases the flexibility of our approach.

In the numerical tests reported below, we use the Armijo-rule (4.3) with parameters $\beta = 0.5$ and $\sigma = 0.1$. Moreover, the smoothed gradient $d^{(k)}$ is adaptively scaled such that $\max_{i,j} \|d_{ij}^{(k)}\|_2$ is within 10% to 90% of the mesh size,

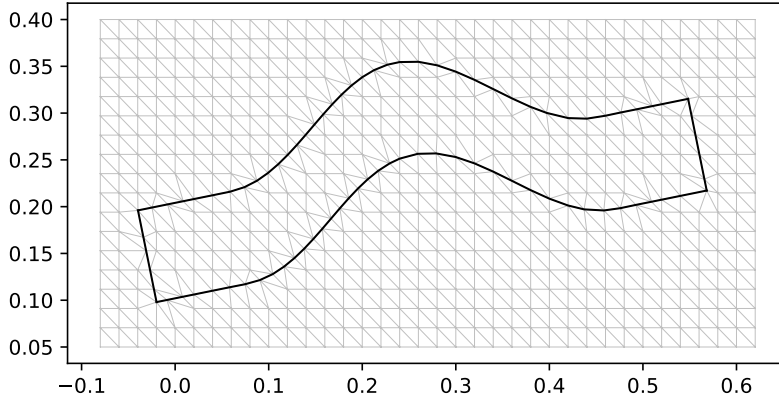


Fig. 9: Illustration of the discretized rod with 36×18 grid points

i.e., $p_{\min} = 0.1$ and $p_{\max} = 0.9$ as explained in Section 4.2. The optimization process is restricted to 500 iterations, i.e., $K = 500$. To avoid tiny step sizes, the Armijo-rule is restricted to 30 iterations, i.e., $\ell \in \{0, \dots, 30\}$. Whenever the Armijo rule does not yield a sufficient improvement for reasonable step sizes, the algorithm stops. Moreover, the algorithm stops when the Frobenius norm of the unscaled search direction $d^{(k)}$ is lower than $\varepsilon = 10^{-10}$.

Twelve exemplary reference points $r(\xi)$ are selected from the pPFG reference set R_H (see (4.6)), with $\xi \in \{0.0005, 0.001, 0.0015, \dots, 0.005, 0.006, 0.007\}$. Algorithm 1 is implemented in Python 3.6 and all tests are run under openuse 15.0 on an IntelCore i7-8700 with 32 GB RAM.

5.1 Results

The twelve optimization runs with varying reference points yield 12 different mutually nondominated outcome vectors. The resulting approximation of the Pareto front is shown in Figure 10. As was to be expected, the optimized shapes resemble a straight rod with varying thickness. The objective vector of the initial shape $\Omega_{36 \times 18}^{(0)}$ is $(18456.28, 0.06)^\top$, which is actually far from the Pareto front.

Several exemplary shapes are shown in Figure 10. In addition to the illustration of the respective shapes, the stresses in the direction of the tensile load acting on the finite elements are depicted as arrows at the respective center points. Note that the diagonal pattern of the arrows is due to the rotated (not axis-parallel) grid structure.

It turns out that the iterates evolve similarly for different optimization runs. For about 50 iterations the hump of the rod decreases quickly, and the Armijo rule often returns a sufficient improvement already for $t_k = 1$. Thus, the shape converges quickly to a nearly straight joint, however, usually with a

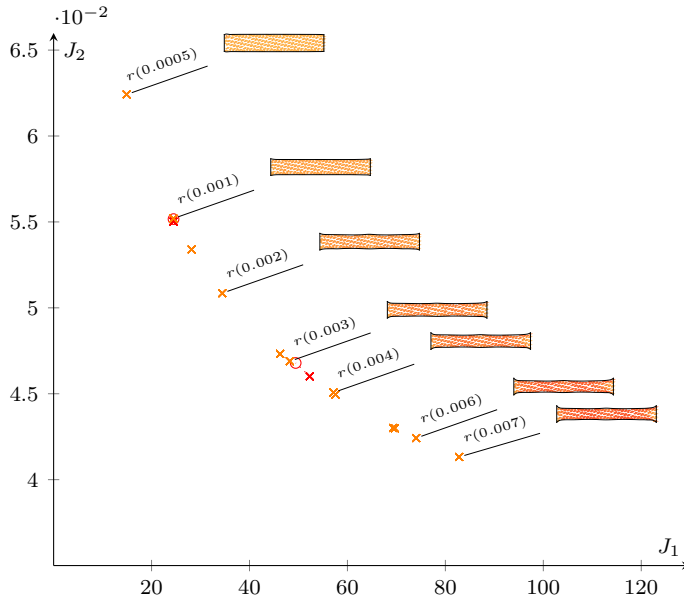


Fig. 10: Coarse approximation on a 36×18 grid (orange) and two exemplary results from the grid refinement (red)

wavy boundary, and its probability of failure improves considerably. In later iterations the changes are less visible and the shape converges slowly to a straight joint, the thickness of which depends on the chosen reference point. During the optimization process, the volume decreases only slowly as the wavy boundary gets smoother.

The optimization runs require between 292 and 486 iterations with one exception: for the reference point $r(0.0005)$ the optimization process stops at the maximum of 500 iterations. In all other cases the algorithm terminated in the Armijo rule.

The solutions from the coarse 36×18 grid were used as starting solutions on a refined 71×36 grid with mesh size $m = 0.01$ m. Exemplary results for the reference points $r(0.001)$ and $r(0.003)$ are shown by red crosses in Figure 10. The respective starting solutions, which are now evaluated on a 71×36 grid, are illustrated by red circles in the same Figure. The results are smoother, see Figure 11, but require significantly more computational time. Nevertheless, in this test case setting the coarse grid approximation obtained already considerably good results, which are not dominated by the corresponding results on the finer grid.

Moreover, numerical tests show that shapes representing very thin or thick straight rods, respectively, induce numerical difficulties due to the fixed boundary parts and, particularly, their corner points. Thick rods suffer from overfit-



(a) Solution $\Omega_{36 \times 18}^*$ evaluated on a 71×36 grid

(b) Solution $\Omega_{71 \times 36}^*$

Fig. 11: Comparison of the coarse and fine grid solution obtained for reference point $r(0.003)$

ted spline representations of their boundaries. Thin rods, on the other hand, lead to distorted finite elements at the spiky corner points.

6 Conclusion

The hypervolume indicator function is a well suited scalarization approach for multiobjective shape optimization problems, since it does not require an explicit handling of additional constraints. In contrast to the weighted sum approach which shares this advantage it is also capable of computing non-supported efficient solutions. The nonlinear dependency on the individual objective functions is not a critical issue in this context since the individual objective functions are already nonlinear. Furthermore, the gradient of the hypervolume can be easily computed based on the gradients of the objective functions.

We show that for convex problems, a non-trivial reference set generating the set of properly nondominated outcome vectors (pPFG reference set) can be determined without prior knowledge on the Pareto front. We validate the approach at a case study from engineering design where the shape of a horizontal bar under tensile load is to be optimized w.r.t. reliability and cost.

Clearly, there are other possibilities for defining pPFG reference sets. We will test and analyze different choices of reference sets in the future. Moreover, the question if and how non-trivial pPFG reference sets can be defined for non-convex problems – without knowing the Pareto front beforehand – will be analyzed. From an engineering point of view, 3D shape optimization are still a challenge that needs further attention.

References

- A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Theory of the hypervolume indicator: optimal μ -distributions and the choice of the reference point. In *Proceedings of the 10th ACM SIGEVO workshop on Foundations of genetic algorithms*, pages 87–102. ACM, 2009.

- I. Babuska and A. Aziz. On the angle condition in the finite element method. *SIAM Journal on Numerical Analysis*, 13, 1976. doi: 10.1137/0713021.
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming – Theory and Algorithms*. Wiley, 3rd edition, 2006.
- M. Bolten, H. Gottschalk, and S. Schmitz. Minimal failure probability for ceramic design via shape control. *J. Optim. Theory Appl.*, 166:983–1001, 2015. doi: 10.1007/s10957-014-0621-8.
- M. Bolten, H. Gottschalk, C. Hahn, and M. Saadi. Numerical shape optimization to decrease failure probability of ceramic structures. *Computing and Visualization in Science*, 2019. doi: 110.1007/s00791-019-00315-z.
- D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 3 edition, 2007. doi: 10.1017/CBO9780511618635.
- K. Bringmann, T. Friedrich, and P. Klitzke. Two-dimensional subset selection for hypervolume and epsilon-indicator. In *Proceedings of the 2014 conference on Genetic and evolutionary computation - GECCO'14*. ACM Press, 2014. doi: 10.1145/2576768.2598276.
- D. V. Chirkov, A. S. Ankudinova, A. E. Kryukov, S. G. Cherny, and V. A. Skorospelov. Multi-objective shape optimization of a hydraulic turbine runner using efficiency, strength and weight criteria. *Structural and Multidisciplinary Optimization*, 58(2):627–640, 2018. doi: 10.1007/s00158-018-1914-6.
- B. Dandurand and M. M. Wiecek. Quadratic scalarization for decomposed multiobjective optimization. *OR Spectrum*, 38(4):1071–1096, 2016. doi: 10.1007/s00291-016-0453-z.
- K. Deb and T. Goel. Multi-objective evolutionary algorithms for engineering shape design. In *Evolutionary Optimization*, volume 48 of *International Series in Operations Research & Management Science*, pages 147–175. Springer, Boston, MA, 2002.
- J.-A. Désidéri. Multiple-gradient descent algorithm (MGDA) for multiobjective optimization. *Comptes Rendus Mathématique*, 350:313–318, 2012. doi: 10.1016/j.crma.2012.03.014.
- O. T. Doganay, C. Hahn, H. Gottschalk, K. Klamroth, J. Schultes, and M. Stiglmayr. Gradient based biobjective shape optimization to improve reliability and cost of ceramic components. *Optimization and Engineering*, 2019. doi: 10.1007/s11081-019-09478-7.
- M. Ehrgott. *Multicriteria Optimization*. Springer, Berlin Heidelberg, 2 edition, 2005.
- J. Fliege. Gap-free computation of Pareto-points by quadratic scalarizations. *Mathematical Methods of Operations Research*, 59:69–89, 2004.
- J. Fliege and B. F. Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51(3):479–494, 2000. doi: 10.1007/s001860000043.
- J. Fliege, A. Vaz, and L. Vicente. Complexity of gradient descent for multi-objective optimization. *Optimization Methods and Software*, 34(5):949–959, 2019. doi: 10.1080/10556788.2018.1510928.

- C. Geiger and C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, 1999.
- A. M. Geoffrion. Proper efficiency and the theory of vector maximization. *Journal of mathematical analysis and applications*, 22(3):618–630, 1968.
- A. P. Guerreiro, C. M. Fonseca, and L. Paquete. Greedy hypervolume subset selection in low dimensions. *Evolutionary Computation*, 24(3):521–544, 2016. doi: 10.1162/EVCO.a.00188.
- J. Haslinger and R. A. E. Mäkinen. *Introduction to Shape Optimization*. SIAM, 2003.
- V. A. S. Hernandez, O. Schütze, H. Wang, A. Deutz, and M. Emmerich. The set-based hypervolume newton method for bi-objective optimization. *IEEE Transactions on Cybernetics*, pages 1–11, 2018. doi: 10.1109/tcyb.2018.2885974.
- L. Iapichino, S. Ulbrich, and S. Volkwein. Multiobjective PDE-constrained optimization using the reduced-basis method. *Advances in Computational Mathematics*, 43(5):945–972, 2017. doi: 10.1007/s10444-016-9512-x.
- T. Kuhn, C. M. Fonseca, L. Paquete, S. Ruzika, M. M. Duarte, and J. R. Figueira. Hypervolume subset selection in two dimensions: Formulations and algorithms. *Evolutionary Computation*, 24(3):411–425, 2016. doi: 10.1162/evco.a.00157.
- K. Miettinen. *Nonlinear Multiobjective Optimization*. Springer, 1998.
- S. Schmidt and V. Schulz. Impulse response approximations of discrete shape Hessians with application in cfd. *SIAM Journal on Control and Optimization*, 48(4):2562–2580, 2009. doi: 10.1137/080719844.
- B. Schulze, M. Stiglmayr, D. Willems, C. M. Fonseca, L. Paquete, and S. Ruzika. On the biobjective cardinality constrained knapsack problem — approximating the hypervolume by quadratic scalarizations. *Mathematical Methods of Operations Research*, 2020. doi: 10.1007/s00186-020-00702-0.
- R. E. Steuer. *Multiple Criteria Optimization: Theory, Computation and Application*. John Wiley, New York, 1986. ISBN 978-0894643934.
- C. Touré, N. Hansen, A. Auger, and D. Brockhoff. Uncrowded hypervolume improvement. In *Proceedings of the Genetic and Evolutionary Computation Conference on - GECCO'19*. ACM Press, 2019. doi: 10.1145/3321707.3321852.
- W. Weibull. *A Statistical Theory of the Strength of Materials*, volume 151 of *Ingenjörsvetenskapsakademiens Handlingar*. Generalstabens Litografiska Anst. Förlag, 1939.
- A. P. Wierzbicki. On the completeness and constructiveness of parametric characterizations to vector optimization problems. *OR-Spektrum*, 8(2):73–87, 1986a. doi: 10.1007/bf01719738.
- A. P. Wierzbicki. A methodological approach to comparing parametric characterizations of efficient solutions. In *Lecture Notes in Economics and Mathematical Systems*, pages 27–45. Springer Berlin Heidelberg, 1986b. doi: 10.1007/978-3-662-02473-7_4.
- K. Yang, M. Emmerich, A. Deutz, and T. Bäck. Efficient computation of expected hypervolume improvement using box decomposition algo-

rithms. *Journal of Global Optimization*, 75(1):3–34, 2019. doi: 10.1007/s10898-019-00798-7.

E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999.