

Multiobjective PDE-Constrained Optimization Using the Reduced-Basis Method

L. Iapichino · S. Ulbrich · S. Volkwein

Received: October 31, 2013/ Accepted: date

Abstract In this paper the reduced basis method is utilized to solve multiobjective optimization problems governed by linear variational equations. These problems often arise in practical applications, where the quality of the system behavior has to be measured by more than one criterium. For the numerical solution the weighting sum method is applied. This approach leads to an algorithm, where many parameterized quadratic optimization problems are solved very efficiently by a appropriate reduced basis approximation. Further, the number of parameter variations is reduced by a sensitivity analysis for the parameterized objective.

1 Introduction

Often in real applications the optimization issue is described by introducing several objective functions which compete with each other. In these cases a good

Laura Iapichino
University of Konstanz
Department of Mathematics and Statistics
Universitätsstraße 10
D-78457 Konstanz, Germany
E-mail: Laura.Iapichino@uni-konstanz.de

Stefan Ulbrich
Darmstadt University of Technology
Department of Mathematics
Dolivostraße 15
Darmstadt, Germany

Stefan Volkwein
University of Konstanz
Department of Mathematics and Statistics
Universitätsstraße 10
D-78457 Konstanz, Germany
E-mail: Stefan.Volkwein@uni-konstanz.de

compromise is the main issue. In the multiobjective optimization we determine efficient or Pareto optimal points which can not be improved simultaneously for all considered objectives. In contrast with the solution of scalar-valued optimization problems, the computation of efficient points is much harder. In particular, if the constraints are given by partial differential equations, we propose reduced-order techniques to handle the computational complexity and to calculate efficient points in reasonable computational times.

In the present paper we study a multiobjective optimization problem [3, 19] with quadratic objectives subject to a linear variational equality, which stands for a weak formulation of an elliptic partial differential equation. For the numerical solution, we apply the weighting method (see [23]), where parameterized scalar-valued quadratic programming problems have to be solved for many parameter values. Thus, after a discretization by standard Galerkin techniques like finite elements, we have to solve many large scale optimization problems. In this work we make use of a reduced basis (RB) approximation [16], which is known to be very efficient for parameterized problem. For the construction of an accurate RB scheme, we make use of the a-posteriori error analysis presented in [14]. Let us refer also to the related work [4, 10] on a-posteriori error estimates for linear-quadratic optimal control problems. It turns out from our numerical experiments that the RB method leads to very fast solutions methods for the parameterized scalar-valued quadratic programming problems. In order to further reduce the numerical complexity of the weighting sum algorithm, we apply a sensitivity analysis for the objective with respect to parameter changes. This strategy leads to a reduction of the parameter variations needed to identify the set of efficient points.

The paper is organized as follows: in Section 2 we introduce our multiobjective optimization problem and recall the definition of efficiency and Pareto optimal points. The optimality conditions for the parameterized scalar-valued minimization problem are formulated as a saddle point problem in Section 3. Here we also introduce a high-fidelity (e.g., finite elements or finite volumes) approximation of the saddle point problem. In Section 4 the RB discretization of the saddle point problem is proposed. Moreover, we discuss the offline and online decomposition of the problem and, finally, the a-posteriori error estimate from [14] is formulated for our application. In Section 5 we explain the sensitivity analysis of the problem. Numerical experiments are shown in Section 6.

2 The multiobjective optimization problem

Let V and H be real, separable Hilbert spaces and suppose that V is dense in H with compact embedding. By $\langle \cdot, \cdot \rangle_H$ and $\langle \cdot, \cdot \rangle_V$ we denote the inner products in H and V , respectively. For $n_c \in \mathbb{N}$ let $\mathcal{D}_c \subset \mathbb{R}^{n_c}$ stand for a parameter set. We suppose that for every parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$ we are given a parameter-

dependent symmetric bilinear form $a(\cdot, \cdot; \boldsymbol{\mu}_c) : V \times V \rightarrow \mathbb{R}$ satisfying:

$$\inf_{\varphi \in V} \frac{a(\varphi, \varphi; \boldsymbol{\mu}_c)}{\|\varphi\|_V^2} \geq \alpha \quad \text{for all } \varphi \in V, \quad (2.1a)$$

$$|a(\varphi, \phi; \boldsymbol{\mu}_c)| \leq \beta \|\varphi\|_V \|\phi\|_V \quad \text{for all } \varphi, \phi \in V, \quad (2.1b)$$

for constants α and $\beta > 0$ independent of $\boldsymbol{\mu}_c$. By identifying H with its dual H' we have $V \hookrightarrow H = H' \hookrightarrow V'$, each embedding being continuous and dense. We suppose that the parameter-dependent inhomogeneity $f(\boldsymbol{\mu}_c)$ belongs to V' for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

For a parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$ and for a control $u \in \mathbb{R}^m$, the state $y \in V$ solves the following linear variational problem

$$a(y, \varphi; \boldsymbol{\mu}_c) = \langle f(\boldsymbol{\mu}_c) + \mathcal{E}u, \varphi \rangle_{V', V} \quad \text{for all } \varphi \in V, \quad (2.2)$$

where $\langle \cdot, \cdot \rangle_{V', V}$ stands for the dual pairing between V and its dual space V' and \mathcal{E} belongs to the Banach space $L(\mathbb{R}^m, V')$ of all bounded, linear operator from \mathbb{R}^m to V' . Since (2.1) holds, it is known that for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$, $u \in \mathbb{R}^m$ and $f(\boldsymbol{\mu}_c) \in V'$ there is a unique weak solution $y = y(\boldsymbol{\mu}_c) \in V$ satisfying (2.2) and

$$\|y\|_V \leq C (\|f(\boldsymbol{\mu}_c)\|_{V'} + \|u\|_{\mathbb{R}^m}), \quad (2.3)$$

for a constant $C > 0$ which is independent of $\boldsymbol{\mu}_c$, u and $f(\boldsymbol{\mu}_c)$. In (2.3) we denote by $\|\cdot\|_{\mathbb{R}^m}$ the Euclidean norm in \mathbb{R}^m .

Remark 2.1 Let $\boldsymbol{\mu}_c \in \mathcal{D}_c$ be chosen arbitrarily. Suppose that $\hat{y} = \hat{y}(\boldsymbol{\mu}_c) \in V$ is the unique solution to

$$a(\hat{y}, \varphi; \boldsymbol{\mu}_c) = \langle f(\boldsymbol{\mu}_c), \varphi \rangle_{V', V} \quad \text{for all } \varphi \in V.$$

Furthermore, we define the parameter-dependent linear mapping $\mathcal{S} = \mathcal{S}(\boldsymbol{\mu}_c) : \mathbb{R}^m \rightarrow V$ as follows: $y = \mathcal{S}(\boldsymbol{\mu}_c)u$ is the unique solution to

$$a(y, \varphi; \boldsymbol{\mu}_c) = \langle \mathcal{E}u, \varphi \rangle_{V', V} \quad \text{for all } \varphi \in V.$$

Then, $y = \hat{y} + \mathcal{S}u$ solves (2.2). It follows from (2.3) that the operator \mathcal{S} is bounded for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$. \diamond

We set $X = V \times \mathbb{R}^m$ and we introduce the vector-valued objective $J : X \rightarrow \mathbb{R}^k$ by

$$J_i(x) = \frac{1}{2} \|\mathcal{C}_i y - w_i\|_{W_i}^2, \quad \text{for } i = 1, \dots, k-1, \quad J_k(x) = \frac{\gamma}{2} \|u\|_U^2,$$

for $x = (y, u) \in X$, where W_1, \dots, W_{k-1} are Hilbert spaces, $\mathcal{C}_i \in L(V, W_i)$ and $w_i \in W_i$ hold for $1 \leq i \leq k-1$. Furthermore, $\gamma > 0$ is a regularization parameter.

Example 2.1 Suppose that Ω is an open and bounded domain in \mathbb{R}^d , $d \in \{1, 2, 3\}$, with Lipschitz-continuous boundary $\Gamma = \partial\Omega$. We set $H = L^2(\Omega)$ and $V = H^1(\Omega)$. For more details on Lebesgue and Sobolev spaces we refer the reader to [5], for instance. Let $k = 3$ and $W_1 = W_2 = H$. The operator \mathcal{C}_1 is the canonical embedding from V into H , the mapping \mathcal{C}_2 is given by $\mathcal{C}_2 y = \|\nabla y\|_{\mathbb{R}^d}^2 \in H$ for $y \in V$ and $w_2 = 0$ holds. Then, for $x = (y, u) \in X$, the first two components of the cost functional are given by

$$J_1(x) = \frac{1}{2} \int_{\Omega} |y(\mathbf{x}) - w_1(\mathbf{x})|^2 \, d\mathbf{x} \quad \text{and} \quad J_2(x) = \frac{1}{2} \int_{\Omega} |\nabla y(\mathbf{x})|_{\mathbb{R}^d}^2 \, d\mathbf{x}.$$

◇

By using Remark 2.1, we define the parameter-dependent reduced cost functional $\hat{J}(\cdot; \boldsymbol{\mu}_c) : \mathbb{R}^m \rightarrow \mathbb{R}^k$ by

$$\hat{J}(u; \boldsymbol{\mu}_c) = J(\hat{y}(\boldsymbol{\mu}_c) + \mathcal{S}(\boldsymbol{\mu}_c)u, u) = \frac{1}{2} \begin{pmatrix} \|\hat{\mathcal{C}}_1(\boldsymbol{\mu}_c)u - \hat{w}_1(\boldsymbol{\mu}_c)\|_{W_1}^2 \\ \vdots \\ \|\hat{\mathcal{C}}_{k-1}(\boldsymbol{\mu}_c)u - \hat{w}_{k-1}(\boldsymbol{\mu}_c)\|_{W_{k-1}}^2 \\ \gamma \|u\|_{\mathbb{R}^m}^2 \end{pmatrix},$$

for $\boldsymbol{\mu}_c \in \mathcal{D}_c$, where we set $\hat{\mathcal{C}}_i(\boldsymbol{\mu}_c) = \mathcal{C}_i \mathcal{S}(\boldsymbol{\mu}_c) \in L(\mathbb{R}^m, W_i)$ and $\hat{w}_i(\boldsymbol{\mu}_c) = w_i - \mathcal{C}_i \hat{y}(\boldsymbol{\mu}_c)$, for $i = 1, \dots, k-1$.

In order to introduce our multiobjective optimization problem, we need the following definitions, which are taken from Chapter 3 in [7].

Definition 2.1 (Order relation) Let (\mathcal{W}, \leq) denote an *order relation* in \mathbb{R}^k , i.e., $\mathcal{W} \subset \mathbb{R}^k \times \mathbb{R}^k$ holds and for any $(z^1, z^2) \in \mathcal{W}$ we have

$$z^1 \leq z^2 \quad \Leftrightarrow \quad z^2 - z^1 \in \mathbb{R}_+^k = \{z \in \mathbb{R}^k \mid z_i \geq 0 \text{ for } i = 1, 2, 3\}.$$

Definition 2.2 (Pareto optimal) Let $Z = \hat{J}(\mathbb{R}^m; \boldsymbol{\mu}_c) \subset \mathbb{R}^k$ the image set of \mathbb{R}^m under the cost functional $\hat{J}(\cdot; \boldsymbol{\mu}_c)$ for a given $\boldsymbol{\mu}_c \in \mathcal{D}_c$. We call a point $\bar{z} \in Z$ (globally) *efficient* with respect to the order relation \leq , if there exists no $z \in Z \setminus \{\bar{z}\}$ with $z \leq \bar{z}$. If \bar{z} is efficient and $\bar{u} \in \mathbb{R}^m$ satisfies $\bar{z} = \hat{J}(\bar{u}; \boldsymbol{\mu}_c)$, we call \bar{u} *Pareto optimal*. If $\bar{x} \in \mathbb{R}^m$ holds and if there exists a neighborhood $\mathcal{N}(\bar{u}) \subset \mathbb{R}^m$ of \bar{u} so that $\bar{y} = \hat{J}(\bar{u}; \boldsymbol{\mu}_c)$ is efficient for the (local) image set $\hat{J}(\mathcal{N}(\bar{u}); \boldsymbol{\mu}_c) \subset Z$, the point \bar{u} is called *locally Pareto optimal* and \bar{y} is said to be *locally efficient*.

Our goal is to find Pareto optimal points for the reduced cost functional $\hat{J}(\cdot; \boldsymbol{\mu}_c)$. Notice that $\hat{J}(\cdot; \boldsymbol{\mu}_c)$ is continuously differentiable for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$. First-order necessary optimality conditions for Pareto optimality are presented in the next theorem which is proved in [6]. The proof is based on the result of Kuhn-Tucker [13].

Theorem 2.1 *Suppose that $\bar{u} \in \mathbb{R}^m$ is Pareto optimal. Then, there is a parameter $\bar{\mu}_o \in \mathbb{R}^k$ satisfying*

$$\bar{\mu}_{o,i} \geq 0, \quad \sum_{i=1}^k \bar{\mu}_{o,i} = 1 \quad \text{and} \quad \sum_{i=1}^k \bar{\mu}_{o,i} \hat{J}'_i(\bar{u}; \boldsymbol{\mu}_c) = 0. \quad (2.4)$$

Let us introduce the parameter set

$$\mathcal{D}_o = \left\{ \boldsymbol{\mu}_o = (\mu_{o,i}) \in \mathbb{R}_+^k \mid \sum_{i=1}^k \mu_{o,i} = 1 \right\} \subset \underbrace{[0, 1] \times \dots \times [0, 1]}_{k\text{-times}} = [0, 1]^k,$$

for the optimization parameters occurring in Theorem 2.1. Then we combine the optimization parameters with the parameters in the state constraints (2.2) by setting $\mathcal{D} = \mathcal{D}_o \times \mathcal{D}_c \subset \mathbb{R}^n$ with $n = k + n_c$. For any $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c) \in \mathcal{D}$ we define the parameter-dependent, scalar-valued cost functional

$$\hat{\mathcal{J}}(u; \boldsymbol{\mu}) = \sum_{i=1}^k \mu_{o,i} \hat{J}_i(u; \boldsymbol{\mu}_c) \quad \text{for } u \in \mathbb{R}^m.$$

Then, (2.4) are the first-order necessary optimality conditions for a local solution \bar{u} to the parameter-dependent optimization problem

$$\min \hat{\mathcal{J}}(u; \bar{\boldsymbol{\mu}}) \quad \text{subject to (s.t.) } u \in \mathbb{R}^m, \quad (\hat{\mathbf{P}}_{\boldsymbol{\mu}})$$

for the parameter $\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}$. We note that $\mu_{o,k} = 1 - \sum_{i=1}^{k-1} \mu_{o,i}$, nevertheless, for sake of notation simplicity, we carry on our theory denoting $\mu_{o,1}, \dots, \mu_{o,k}$ as independent parameters. To solve $(\hat{\mathbf{P}}_{\boldsymbol{\mu}})$ we can apply methods from quadratic programming; see, e.g., [15]. However, the optimal parameter $\bar{\boldsymbol{\mu}}$ is not known a-priorily. In the weighting method – first introduced by Zadeh [23] – Pareto optimal points are computed by solving $(\hat{\mathbf{P}}_{\boldsymbol{\mu}})$ for various $\boldsymbol{\mu}_o \in \mathcal{D}_o$. However, in general, this does not yield all Pareto optimal points for $\hat{\mathcal{J}}(\cdot; \boldsymbol{\mu}_c)$, due to the fact that second-order conditions for an optimal solution to $(\hat{\mathbf{P}}_{\boldsymbol{\mu}})$ are not necessary for a Pareto optimal point. The missing (second-order) optimality conditions leads to the differences in scalar-valued and multiobjective optimization; see, e.g., the discussion in [7, Section 4.4]. In this paper we will not address this issue, but we compute an approximation of the set of Pareto optimal points by applying the weighting method.

3 Optimality system for the scalar-valued linear-quadratic problem

Let $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c) \in \mathcal{D}$ be arbitrarily given. The solution of scalar-valued minimization problem $(\hat{\mathbf{P}}_{\boldsymbol{\mu}})$ can be characterized by its Karush-Kuhn-Tucker conditions, which leads to a system of variational problems having the structure of a saddle-point problem and containing the state equation, the adjoint equation and the optimality condition. This saddle-point structure is particularly advantageous, since its solution can be efficiently solved by the reduced basis problem.

3.1 Saddle point formulation for $(\hat{\mathbf{P}}_\mu)$

We introduce the μ -dependent Lagrangian for $(\hat{\mathbf{P}}_\mu)$ as

$$\mathcal{L}(y, u, p; \mu) = \sum_{i=1}^k \mu_{o,i} J(y, u; \mu_c) + a(y, p; \mu_c) - \langle f(\mu_c) + \mathcal{E}u, p \rangle_{V',V}$$

for $(y, u) \in X$, $p \in V$ and $\mu \in \mathcal{D}$. Notice that $(\hat{\mathbf{P}}_\mu)$ is a linear-quadratic optimal control problem. Therefore, the existence of a unique solution $\bar{u} = \bar{u}(\mu)$ follows by standard arguments for any $\mu \in \mathcal{D}$; see, e.g., [8, 20]. Then, \bar{u} satisfies together with the unique associated optimal state $\bar{y} = \bar{y}(\mu)$ and optimal adjoint $\bar{p} = \bar{p}(\mu)$ the following first-order necessary optimality conditions:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial y}(\bar{x}, \bar{p}; \mu) y = a(y, \bar{p}; \mu_c) + \sum_{i=1}^{k-1} \mu_{o,i} \langle \mathcal{C}_i \bar{y} - w_i, \mathcal{C}_i y \rangle_{W_i}, \\ 0 &= \frac{\partial \mathcal{L}}{\partial u}(\bar{x}, \bar{p}; \mu) u = \mu_{o,k} \gamma \bar{u}^\top u - \langle \mathcal{E}u, \bar{p} \rangle_{V',V}, \\ 0 &= \frac{\partial \mathcal{L}}{\partial p}(\bar{x}, \bar{p}; \mu) p = a(\bar{y}, p; \mu_c) - \langle f(\mu_c) + \mathcal{E}\bar{u}, p \rangle_{V',V}, \end{aligned}$$

for directions $x = (y, u) \in X$, $p \in V$ and for $\bar{x} = (\bar{y}, \bar{u})$. We define the following bilinear forms:

$$\begin{aligned} \mathcal{A}(x, \tilde{x}; \mu) &= \sum_{i=1}^{k-1} \mu_{o,i} \langle \mathcal{C}_i y, \mathcal{C}_i \tilde{y} \rangle_{W_i} + \mu_{o,k} \gamma u^\top \tilde{u}, \\ \mathcal{B}(x, \tilde{p}; \mu) &= a(y, \tilde{p}; \mu_c) - \langle \mathcal{E}\bar{u}, \tilde{p} \rangle_{V',V}, \end{aligned}$$

and linear forms:

$$\mathcal{F}(\tilde{x}; \mu) = \sum_{i=1}^{k-1} \mu_{o,i} \langle w_i, \mathcal{C}_i \tilde{y} \rangle_{W_i}, \quad \mathcal{G}(\tilde{p}; \mu) = \langle f(\mu_c), \tilde{p} \rangle_{V',V},$$

for all $x = (y, u)$, $\tilde{x} = (\tilde{y}, \tilde{u}) \in X$ and $\tilde{p} \in V$. Then, the first-order optimality conditions can be written as the following linear saddle point system:

$$\begin{aligned} \mathcal{A}(\bar{x}, x; \mu) + \mathcal{B}(x, \bar{p}; \mu) &= \mathcal{F}(x; \mu) \quad \text{for all } x = (y, u) \in X, \\ \mathcal{B}(\bar{x}, p; \mu) &= \mathcal{G}(p; \mu) \quad \text{for all } p \in V. \end{aligned} \tag{3.1}$$

Since the bilinear form $\mathcal{A}(\cdot, \cdot; \mu)$ is continuous and coercive and $\mathcal{B}(\cdot, \cdot; \mu)$ is continuous and satisfies the inf-sup condition, thanks to the Brezzi theorem [2], problem (3.1) has a unique solution for any $\mu \in \mathcal{D}$.

3.2 High-fidelity Galerkin approximation

Now we introduce a high-fidelity Galerkin approximation for the infinite dimensional saddle point problem (3.1). For that purpose we assume that a typically large $\mathcal{N}_1 \in \mathbb{N}$ and an $\mathcal{N}_2 \in \mathbb{N}$ with $\mathcal{N}_2 \leq m$ are given and $\mathcal{N}_{12} = \mathcal{N}_1 + \mathcal{N}_2$. Let $\{\varphi_i\}_{i=1}^{\mathcal{N}_1}$ and $\{\phi_i\}_{i=1}^{\mathcal{N}_2}$ denote a set of linearly independent basis functions in V and \mathbb{R}^m , respectively. We set the finite dimensional ansatz space

$$V^{\mathcal{N}} = \text{span}\{\varphi_1, \dots, \varphi_{\mathcal{N}_1}\} \subset V, \quad U^{\mathcal{N}} = \text{span}\{\phi_1, \dots, \phi_{\mathcal{N}_2}\} \subset \mathbb{R}^m.$$

Then, the high-fidelity Galerkin approximation of the optimality system (3.1) reads as follows: find $(x^{\mathcal{N}}, p^{\mathcal{N}}) \in X^{\mathcal{N}} \times V^{\mathcal{N}}$ such that

$$\begin{aligned} \mathcal{A}(\bar{x}^{\mathcal{N}}, x; \boldsymbol{\mu}) + \mathcal{B}(x, \bar{p}^{\mathcal{N}}, \boldsymbol{\mu}) &= \mathcal{F}(x; \boldsymbol{\mu}) \quad \text{for all } x = (y, u) \in X^{\mathcal{N}}, \\ \mathcal{B}(\bar{x}^{\mathcal{N}}, p; \boldsymbol{\mu}) &= \mathcal{G}(p; \boldsymbol{\mu}) \quad \text{for all } p \in V^{\mathcal{N}}, \end{aligned} \quad (3.2)$$

where we have set $X^{\mathcal{N}} = V^{\mathcal{N}} \times U^{\mathcal{N}}$. We use the notations: $\bar{x}^{\mathcal{N}} = (\bar{y}^{\mathcal{N}}, \bar{u}^{\mathcal{N}})$ and

$$\bar{y}^{\mathcal{N}} = \sum_{i=1}^{\mathcal{N}_1} \bar{x}_i \varphi_i = \sum_{i=1}^{\mathcal{N}_1} \bar{y}_i \varphi_i, \quad \bar{u}^{\mathcal{N}} = \sum_{i=1}^{\mathcal{N}_2} \bar{x}_{\mathcal{N}_1+i} \phi_i = \sum_{i=1}^{\mathcal{N}_2} \bar{u}_i \phi_i, \quad \bar{p}^{\mathcal{N}} = \sum_{i=1}^{\mathcal{N}_1} \bar{p}_i \varphi_i.$$

The bilinear form $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ is continuous in $X^{\mathcal{N}} \times X^{\mathcal{N}}$ and coercive over

$$X_0^{\mathcal{N}} = \{x \in X^{\mathcal{N}} \mid \mathcal{B}(x, p; \boldsymbol{\mu}) = 0 \text{ for all } p \in V^{\mathcal{N}}\}.$$

Furthermore, $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ continuous and inf-sup stable in $X^{\mathcal{N}} \times V^{\mathcal{N}}$. The existence, uniqueness and stability of problem (3.2) is ensured by the Brezzi theorem [2]. Indeed, stability is prescribed by imposing that the coercivity and inf-sup conditions are still valid at the discrete level. Problem (3.2) admits a unique solution providing that $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ is coercive over $X^{\mathcal{N}}$:

$$\exists \alpha_{LB}^{\mathcal{N}}(\boldsymbol{\mu}) > 0 : \alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{x \in X^{\mathcal{N}}} \frac{\mathcal{A}(x, x; \boldsymbol{\mu})}{\|x\|_X^2} \geq \alpha_{LB}^{\mathcal{N}}(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (3.3)$$

and $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ is inf-sup stable over $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, so that the following discrete Brezzi inf-sup condition [2]:

$$\exists \beta_{LB}^{\mathcal{N}}(\boldsymbol{\mu}) > 0 : \beta^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{p \in V^{\mathcal{N}}} \sup_{x \in X^{\mathcal{N}}} \frac{\mathcal{B}(x, p; \boldsymbol{\mu})}{\|x\|_X \|p\|_V} \geq \beta_{LB}^{\mathcal{N}}(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (3.4)$$

is satisfied. The numerical solution of problem (3.2) yields to a linear system of dimension $2\mathcal{N}_{12} + \mathcal{N}_1$ of the following structure:

$$\begin{aligned} \sum_{j=1}^{\mathcal{N}_{12}} A_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) \bar{x}_j^{\mathcal{N}} + \sum_{j=1}^{\mathcal{N}_1} B_{ji}^{\mathcal{N}}(\boldsymbol{\mu}) \bar{p}_j^{\mathcal{N}} &= F_i^{\mathcal{N}}(\boldsymbol{\mu}), \quad i = 1, \dots, \mathcal{N}_{12}, \\ \sum_{j=1}^{\mathcal{N}_{12}} B_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) \bar{x}_j^{\mathcal{N}} &= G_j^{\mathcal{N}}(\boldsymbol{\mu}), \quad i = 1, \dots, \mathcal{N}_1, \end{aligned} \quad (3.5)$$

where the matrices $A^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}_{12} \times \mathcal{N}_{12}}$, $B^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}_{12} \times \mathcal{N}_1}$ and the vectors $F^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}_{12}}$, $G^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}_1}$ are given by

$$\begin{aligned}
A_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) &= \begin{cases} \mathcal{A}((\varphi_j, 0), (\varphi_i, 0); \boldsymbol{\mu}) & \text{for } i, j = 1, \dots, \mathcal{N}_1, \\ \mathcal{A}((0, \phi_{j-\mathcal{N}_1}), (0, \phi_{i-\mathcal{N}_1}); \boldsymbol{\mu}) & \text{for } i, j = \mathcal{N}_1 + 1, \dots, \mathcal{N}_{12}, \\ 0 & \text{otherwise,} \end{cases} \\
B_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) &= \begin{cases} \mathcal{B}((\varphi_j, 0), \varphi_i; \boldsymbol{\mu}) & \text{for } i, j = 1, \dots, \mathcal{N}_1, \\ \mathcal{B}((0, \phi_{j-\mathcal{N}_1}), \varphi_i; \boldsymbol{\mu}) & \text{for } i = 1, \dots, \mathcal{N}_1, j = \mathcal{N}_1 + 1, \dots, \mathcal{N}_{12}, \end{cases} \\
F_i^{\mathcal{N}}(\boldsymbol{\mu}) &= \begin{cases} \mathcal{F}((\varphi_j, 0); \boldsymbol{\mu}) & \text{for } i = 1, \dots, \mathcal{N}_1, \\ \mathcal{F}((0, \phi_{j-\mathcal{N}_1}); \boldsymbol{\mu}) & \text{for } i = \mathcal{N}_1 + 1, \dots, \mathcal{N}_{12}, \end{cases} \\
G_i^{\mathcal{N}}(\boldsymbol{\mu}) &= \mathcal{G}(\varphi_i; \boldsymbol{\mu}) \quad \text{for } i = 1, \dots, \mathcal{N}_1.
\end{aligned}$$

In order to compute an accurate approximation of the saddle-point problem (3.1), \mathcal{N}_1 and \mathcal{N}_{12} are typically large. Consequently, (3.5) has a big dimension and requires long computational time, specially if we want to compute a large set of Pareto optimal control points, that means the solution of (3.5) for many different values of the parameters.

4 Reduced basis approximation

The idea of the reduced-basis (RB) method is to reduce the complexity of the numerical solution of problem (3.2). The main basic idea is to exploit the smooth parametric dependence of the problem and, thanks to this assumption, to define a small and suitable basis space for defining the approximated solution of the problem. The RB method consists in two main stages, during the first one, so called *offline*, we compute the basis functions that are more complex in comparison with the high-fidelity ones and require a bigger effort to compute. Nevertheless, once the first expensive offline stage is computed, for every new value of the parameter $\boldsymbol{\mu}$, the RB method provides a dataset very effective for the computation of an accurate, reliable and fast approximation of the problem during the *online* phase.

4.1 Offline phase

To define a suitable basis spaces, we build a set of parameter samples $S_N^\mu = \{\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N\}$ and correspondingly a set of pairs $\{(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}^i), \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}^i))\}_{i=1}^N$ which are solutions of the saddle point problem (3.2) for parameters $\boldsymbol{\mu} \in S_N^\mu$. The choice of the parameter set S_N^μ can be done by using both the classical greedy algorithm, proposed in [17], and the optimization greedy, recently proposed in [22], particularly useful for problems involving a large number of parameters (i.e. a large number of cost functionals) or distributed parameters functions.

The RB approximation of (3.2) consists in a Galerkin projection onto low dimensional subspaces containing the solution pairs $\{(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}^i), \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}^i))\}_{i=1}^N$. In

order to guarantee the approximation stability of the reduced basis method for the saddle point problem, we fulfill the inf-sup condition by defining the following spaces:

$$V^N = \text{span} \{\bar{y}^N(\boldsymbol{\mu}^1), \dots, \bar{y}^N(\boldsymbol{\mu}^i)\} \cup \text{span} \{\bar{p}^N(\boldsymbol{\mu}^1), \dots, \bar{p}^N(\boldsymbol{\mu}^N)\}, \quad (4.1)$$

$$U^N = \text{span} \{\bar{u}^N(\boldsymbol{\mu}^1), \dots, \bar{u}^N(\boldsymbol{\mu}^N)\}. \quad (4.2)$$

Let $\{\psi_i\}_{i=1}^{2N}$ and $\{\zeta_i\}_{i=1}^N$ denote orthonormal bases for V^N and U^N , respectively. We set $X^N = V^N \times U^N$. The constant

$$\beta^N(\boldsymbol{\mu}) := \inf_{p \in V^N} \sup_{x \in X^N} \frac{\mathcal{B}(x, p; \boldsymbol{\mu})}{\|x\|_X \|p\|_V} \quad \forall \boldsymbol{\mu} \in \mathcal{D} \quad (4.3)$$

fulfills a Brezzi RB inf-sup condition (see [14])

$$\beta^N(\boldsymbol{\mu}) \geq \alpha^N(\boldsymbol{\mu}) > 0 \quad \forall \boldsymbol{\mu} \in \mathcal{D},$$

where $\alpha(\boldsymbol{\mu})$ has been introduced in (3.3). The RB approximation of problem (3.2) reads: find the pair $(\bar{x}^N, \bar{p}^N) \in X^N \times V^N$ such that

$$\begin{aligned} \mathcal{A}(\bar{x}^N, x; \boldsymbol{\mu}) + \mathcal{B}(x, \bar{p}^N; \boldsymbol{\mu}) &= \mathcal{F}(x; \boldsymbol{\mu}) \quad \text{for all } x \in X^N, \\ \mathcal{B}(\bar{x}^N, p; \boldsymbol{\mu}) &= \mathcal{G}(p; \boldsymbol{\mu}) \quad \text{for all } p \in V^N. \end{aligned} \quad (4.4)$$

In addition to the selection of S_N^μ and the computation of the basis functions, during the offline phase we compute the parameter independent parts of the matrices and vectors that define the algebraic formulation of problem (4.4). For this purpose, we have to deal with an affine decomposition of the linear and bilinear forms. In particular, we note that $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ and $\mathcal{F}(\cdot; \boldsymbol{\mu})$ are affine by definition. Thus, we require the affine parametric dependence of the $\boldsymbol{\mu}$ -dependent expressions in the state equation (2.2), i.e. of the forms $a(\cdot, \cdot; \boldsymbol{\mu}_c)$ and $f(\boldsymbol{\mu}_c)$ with $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c) \in \mathcal{D}$. However, if they are not affine, it is possible to approximate them into affine linear and bilinear forms through the empirical interpolation method [1]. So that it is possible to decouple the forms as follows:

$$\mathcal{A}(x, \tilde{x}; \boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{A}}} \Theta_{\mathcal{A}}^i(\boldsymbol{\mu}) \mathcal{A}^i(x, \tilde{x}), \quad \mathcal{F}(x; \boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{F}}} \Theta_{\mathcal{F}}^i(\boldsymbol{\mu}) \mathcal{F}^i(\tilde{x})$$

and to approximate

$$\mathcal{B}(x, p; \boldsymbol{\mu}) \approx \sum_{i=1}^{\ell_{\mathcal{B}}} \Theta_{\mathcal{B}}^i(\boldsymbol{\mu}) \mathcal{B}^i(x, p), \quad \mathcal{G}(p; \boldsymbol{\mu}) \approx \sum_{i=1}^{\ell_{\mathcal{G}}} \Theta_{\mathcal{G}}^i(\boldsymbol{\mu}) \mathcal{G}^i(p)$$

for $x = (y, u)$, $\tilde{x} = (\tilde{y}, \tilde{u}) \in X$, $p \in V$, $\boldsymbol{\mu} \in \mathcal{D}$. We note that this affine decoupling of the problem is very effective for the computational time of the RB

solution, since we can compute in the offline phase the parameter independent parts of the problem (4.4). Especially, we determine once in this step the following matrices:

$$\begin{aligned}
A_{ij}^{i_1} &= \begin{cases} \mathcal{A}^{i_1}((\psi_j, 0), (\psi_i, 0)) & \text{for } i, j = 1, \dots, 2N, \\ \mathcal{A}^{i_1}((0, \zeta_{j-2N}), (0, \zeta_{i-2N})) & \text{for } i, j = 2N + 1, \dots, 3N, \\ 0 & \text{otherwise,} \end{cases} \\
B_{ij}^{i_2} &= \begin{cases} \mathcal{B}^{i_2}((\psi_j, 0), \psi_i) & \text{for } i, j = 1, \dots, 2N, \\ \mathcal{B}^{i_2}((0, \zeta_{j-2N}), \psi_i) & \text{for } i = 1, \dots, 2N, j = 2N + 1, \dots, 3N, \end{cases} \\
F_i^{i_3} &= \begin{cases} \mathcal{F}^{i_3}((\psi_j, 0)) & \text{for } i = 1, \dots, 2N, \\ \mathcal{F}^{i_3}((0, \zeta_{j-2N})) & \text{for } i = 2N + 1, \dots, 3N, \end{cases} \\
G_i^{i_4} &= \mathcal{G}^{i_4}(\psi_i) \quad \text{for } i = 1, \dots, 2N.
\end{aligned}$$

for $1 \leq i_1 \leq \ell_{\mathcal{A}}$, $1 \leq i_2 \leq \ell_{\mathcal{B}}$, $1 \leq i_3 \leq \ell_{\mathcal{F}}$, and $1 \leq i_4 \leq \ell_{\mathcal{G}}$.

4.2 Online phase

In the online stage, for each new parameter value the parametric coefficients of the system can be quickly evaluated. Finally, a small linear system can be solved efficiently during the online stage many times to find the coefficients $\bar{x} = (\bar{x}_i^N) \in \mathbb{R}^{3N}$ and $\bar{p}^N = (\bar{p}_i^N) \in \mathbb{R}^{2N}$ that will give the final reduced basis solution for each new value of μ . The RB linear system is the following:

$$\begin{aligned}
\sum_{j=1}^{3N} A_{ji}^N(\mu) \bar{x}_j^N(\mu) + \sum_{j=1}^{2N} B_{ji}^N(\mu) \bar{p}_j^N(\mu) &= F_i^N(\mu), \quad i = 1, \dots, 3N, \\
\sum_{j=1}^{2N} B_{ij}^N(\mu) \bar{x}_j^N &= G_i^N(\mu), \quad i = 1, \dots, 2N,
\end{aligned} \tag{4.5}$$

where

$$\begin{aligned}
A^N(\mu) &= \sum_{i=1}^{\ell_{\mathcal{A}}} \Theta_{\mathcal{A}}^i(\mu) A^i, & F^N(\mu) &= \sum_{i=1}^{\ell_{\mathcal{F}}} \Theta_{\mathcal{F}}^i(\mu) F^i, \\
B^N(\mu) &= \sum_{i=1}^{\ell_{\mathcal{B}}} \Theta_{\mathcal{B}}^i(\mu) B^i, & G^N(\mu) &= \sum_{i=1}^{\ell_{\mathcal{G}}} \Theta_{\mathcal{G}}^i(\mu) G^i.
\end{aligned}$$

4.3 Error bounds estimation

The rigorous error bound is one of the most important ingredients of the RB method. It allows to define a suitable, efficient and relatively fast selection

of the parameter set S_N^μ and it provides a certificate level of accuracy of the approximate solution compared with the high-fidelity FE one. Thanks to the saddle-point formulation of the problem, we exploit the error bounds proposed in [18] for Stokes problem and more recently in [14] for elliptic linear-quadratic optimal control problems. Thus, we have a rigorous and inexpensive bound for error between the high-fidelity solution of (3.2) and the RB solution of (4.4).

$$\left(\|\bar{x}^N(\boldsymbol{\mu}) - \bar{x}^N(\boldsymbol{\mu})\|_X^2 + \|\bar{p}^N(\boldsymbol{\mu}) - \bar{p}^N(\boldsymbol{\mu})\|_V^2\right)^{1/2} \leq \Delta_N(\boldsymbol{\mu})$$

for any $\boldsymbol{\mu} \in \mathcal{D}$. Moreover, we have a rigorous and inexpensive bound for the error of the values cost functional computed with the high-fidelity solution and the RB one:

$$|\hat{\mathfrak{J}}(\bar{u}^N(\boldsymbol{\mu}); \boldsymbol{\mu}) - \hat{\mathfrak{J}}(\bar{u}^N(\boldsymbol{\mu}); \boldsymbol{\mu})| \leq \Delta_N^J(\boldsymbol{\mu})$$

for any $\boldsymbol{\mu} \in \mathcal{D}$. We note that the offline-online computational decomposition can be adopted also for the computation of the error bounds in order to be able to efficiently compute it online together with the RB solution of the problem.

The offline stage, performed once, is very expensive and \mathcal{N} -dependent, while the online evaluation, performed many times for each new desired $\boldsymbol{\mu}$, is very inexpensive and \mathcal{N} -independent. This efficient and reliable error estimation permits to predict the RB error with respect to the FE solution without computing the latter and it is crucial during the greedy algorithm to speed up the efficient selection of the snapshots.

The a posteriori error estimation takes advantage of two quantities: the dual norm of residuals and an effective lower bound of the (parametric) stability factor, given in this case by the Babuška inf-sup constant $\beta^N(\boldsymbol{\mu})$ defined in (3.4). Let us define the residuals $r_{\text{du}}^N(\cdot; \boldsymbol{\mu})$ and $r_{\text{pr}}^N(\cdot; \boldsymbol{\mu})$ by

$$\begin{aligned} r_{\text{du}}^N(x; \boldsymbol{\mu}) &= \mathcal{F}^N(x; \boldsymbol{\mu}) - \mathcal{A}^N(\bar{x}^N(\boldsymbol{\mu}), x; \boldsymbol{\mu}) - \mathcal{B}(x, \bar{p}^N(\boldsymbol{\mu}); \boldsymbol{\mu}), & \forall x \in X^N, \\ r_{\text{pr}}(p; \boldsymbol{\mu}) &= G(q; \boldsymbol{\mu}) - \mathcal{B}(\bar{x}^N(\boldsymbol{\mu}), p; \boldsymbol{\mu}), & \forall p \in V^N. \end{aligned}$$

Then, we obtain the following upper bounds for the a-posteriori error estimators (see, e.g., [11])

$$\begin{aligned} \Delta_N(\boldsymbol{\mu}) &\leq \frac{1}{\beta_{LB}^N(\boldsymbol{\mu})^2} \left(\|r_{\text{du}}^N(\cdot; \boldsymbol{\mu})\|_{(X^N)'}^2 + \|r_{\text{pr}}^N(\cdot; \boldsymbol{\mu})\|_{(V^N)'}^2 \right), \\ \Delta_N^J(\boldsymbol{\mu}) &\leq \frac{1}{2} \frac{\|r(\cdot; \boldsymbol{\mu})\|_{(X^N)'}}{\beta_{LB}^N(\boldsymbol{\mu})} \end{aligned}$$

for $\boldsymbol{\mu} \in \mathcal{D}$, where

$$\|r_{\text{du}}(\cdot; \boldsymbol{\mu})\|_{(X^N)'} = \sup_{x \in X^N} \frac{r_{\text{du}}(x; \boldsymbol{\mu})}{\|x\|_X}, \quad \|r_{\text{pr}}(\cdot; \boldsymbol{\mu})\|_{(V^N)'} = \sup_{p \in V^N} \frac{r_{\text{pr}}(p; \boldsymbol{\mu})}{\|p\|_V}$$

and $r^N(x, p; \boldsymbol{\mu}) = r_{\text{du}}(x; \boldsymbol{\mu}) + r_{\text{pr}}(p; \boldsymbol{\mu})$ for $(x, p) \in X^N \times V^N$ hold. The computation of dual norms of residuals is based instead on the Riesz representation of the residuals and on the affine decomposition of the parametric operators.

Moreover, the approximation of the lower bound of $\beta^{\mathcal{N}}(\boldsymbol{\mu})$ plays an important role for a rigorous error estimation. It can be computed by the Natural Norm Successive Constraint Method, that represent an improvement of the SCM, see [9]. However, since this approximation of the lower bound can be very time consuming, in this paper we adopt an alternative strategy recently proposed and compared with the previous one in [14]. It consists in defining a surrogate $\beta_s^{\mathcal{N}}(\boldsymbol{\mu})$ obtained by computing $\beta^{\mathcal{N}}(\boldsymbol{\mu})$ for a small set of parameter values and by using these computations to define, by interpolation, the surrogate $\beta_s^{\mathcal{N}}(\boldsymbol{\mu}) \forall \boldsymbol{\mu} \in \mathcal{D}$. Despite this surrogate interpolation can not be seen as a rigorous lower bound, it represents a sharp approximation and, at a much lower computational cost, it provides a suitable and efficient error bound estimation.

5 Sensitivity analysis for the computation of the Pareto optimal points

We recall that the parameter $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c) \in \mathcal{D}$ is given by components of $\boldsymbol{\mu}_o$, which are chosen for the weighted sum of the objective, and by components of $\boldsymbol{\mu}_c$ that stand for physical and/or geometrical parameters in the state equation. In this section we present a strategy which is used for the weighting sum method. To reduce the number of possible optimization parameter variations $\boldsymbol{\mu}_o \in \mathcal{D}_o$, we apply a sensitivity analysis for the reduced cost functional with respect to the optimization parameter $\boldsymbol{\mu}_o$. Suppose that we have computed the optimal solution $\bar{u} = \bar{u}(\bar{\boldsymbol{\mu}})$ to $(\hat{\mathbf{P}}_{\boldsymbol{\mu}})$ for a chosen parameter $\bar{\boldsymbol{\mu}} = (\bar{\boldsymbol{\mu}}_o, \boldsymbol{\mu}_c)$. By $\bar{y} = \bar{y}(\bar{\boldsymbol{\mu}})$ we denote the associated optimal state and by $\bar{p} = \bar{p}(\bar{\boldsymbol{\mu}})$ the associated optimal Lagrange multiplier and we set $\bar{x} = (\bar{y}, \bar{u})$. In the weighting sum method, we are interested to choose weights $\boldsymbol{\mu}_o$ leading to significant changes in the cost functional. For that reason we look at the Taylor expansion of the reduced objective with respect to changes in $\boldsymbol{\mu}_o$:

$$\hat{\mathfrak{J}}(u; \boldsymbol{\mu}) = \hat{\mathfrak{J}}(\bar{u}; \bar{\boldsymbol{\mu}}) + \frac{\partial \hat{\mathfrak{J}}}{\partial \boldsymbol{\mu}_o}(\bar{u}; \bar{\boldsymbol{\mu}})(\boldsymbol{\mu}_o - \bar{\boldsymbol{\mu}}_o) + \mathcal{O}(\|\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}\|_{\mathbb{R}^n}^2),$$

where $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c)$, i.e., $\boldsymbol{\mu}$ and $\bar{\boldsymbol{\mu}}$ only differ in the first k components. Hence, we have to compute the partial derivatives of $\hat{\mathfrak{J}}$ at $(\bar{u}; \bar{\boldsymbol{\mu}})$ with respect to $\boldsymbol{\mu}_o$ in order to get the requested information. We recall that

$$\hat{\mathfrak{J}}(\bar{u}; \bar{\boldsymbol{\mu}}) = \frac{1}{2} \left(\sum_{i=1}^{k-1} \bar{\mu}_{o,i} \|\mathcal{C}_i \bar{y} - w_i\|_{W_i}^2 + \mu_{o,k} \gamma \|\bar{u}\|_{\mathbb{R}^n}^2 \right).$$

Where, by considering that $\mu_{o,k} = \sum_{i=1}^{k-1} \mu_{o,i}$, we compute $k-1$ derivatives of the cost functional. We have for $i = 1, \dots, k-1$:

$$\frac{\partial \hat{\mathfrak{J}}}{\partial \mu_i}(\bar{u}; \bar{\boldsymbol{\mu}}) = \frac{1}{2} \|\mathcal{C}_i \bar{y} - w_i\|_{W_i}^2 + \bar{\mu}_i \langle \mathcal{C}_i \bar{y} - w_i, \bar{y}_{\mu_i} \rangle_{W_i} + \frac{\gamma}{2} \|\bar{u}\|_{\mathbb{R}^n}^2 - \bar{\mu}_i \gamma \langle \bar{u}, \bar{u}_{\mu_i} \rangle_{\mathbb{R}^n},$$

where the sensitivities \bar{y}_{μ_i} and \bar{u}_{μ_i} have to be computed. The first-order optimality conditions can be expressed as the linear system (3.1). Thus, we can define an operator $\mathcal{K}(\bar{\boldsymbol{\mu}}) \in L(V \times \mathbb{R}^m \times V, V' \times \mathbb{R}^m \times V')$ and a right-hand side $\mathcal{R}(\bar{\boldsymbol{\mu}}) \in V' \times \mathbb{R}^m \times V'$ so that (3.1) can be written shortly as

$$\mathcal{K}(\bar{\boldsymbol{\mu}})(\bar{y}, \bar{u}, \bar{p}) = \mathcal{R}(\bar{\boldsymbol{\mu}}) \quad \text{in } V' \times \mathbb{R}^m \times V'. \quad (5.1)$$

Now we differentiate (5.1) with respect to $\boldsymbol{\mu}_o$. This gives for $i = 1, \dots, k-1$:

$$\mathcal{K}(\bar{\boldsymbol{\mu}})(\bar{y}_{\mu_i}, \bar{u}_{\mu_i}, \bar{p}_{\mu_i}) = \mathcal{R}_{\mu_i}(\bar{\boldsymbol{\mu}}) - \mathcal{K}_{\mu_i}(\bar{\boldsymbol{\mu}})(\bar{y}, \bar{u}, \bar{p}) \quad \text{in } V' \times \mathbb{R}^m \times V', \quad (5.2)$$

where

$$\begin{aligned} \left\langle \mathcal{K}_{\mu_i}(\bar{\boldsymbol{\mu}})(\bar{y}, \bar{u}, \bar{p}), (y^\delta, u^\delta, p^\delta) \right\rangle_{V' \times \mathbb{R}^m \times V', V \times \mathbb{R}^m \times V} &= \left\langle \langle \mathcal{C}_i \cdot, \mathcal{C}_i y^\delta \rangle_{W_i}, \bar{y} \right\rangle_{V', V} \\ &\quad - \left\langle \gamma u^\delta, \bar{u} \right\rangle_{\mathbb{R}^m} \\ \left\langle \mathcal{R}_{\mu_i}(\bar{\boldsymbol{\mu}}), (y^\delta, u^\delta, p^\delta) \right\rangle_{V' \times \mathbb{R}^m \times V', V \times \mathbb{R}^m \times V} &= \langle w_i, \mathcal{C}_i y^\delta \rangle_{W_i}. \end{aligned}$$

Now, we can compute the sensitivities $\bar{y}_{\mu_i} \in V$, $\bar{u}_{\mu_i} \in \mathbb{R}^m$ from (5.2).

The advantageous feature of the explained sensitivity theory is its efficient online-offline computational decoupling coming from the RB precomputed structures.

At a very small computational effort, we are able to define a suitable parameter set able to provide an approximation of the Pareto optimal set. In the following we summarize the steps we adopt for defining the suitable Pareto points approximation. Starting with the parametric saddle point formulation of the problem where the parameter vector comes from both the weighted sum of the cost functionals and the PDE constraints:

- i) we compute the RB offline step for the RB approximation of the saddle point problem;
- ii) we compute the RB online step for defining the RB optimal solutions corresponding to one or more initial parameter guesses;
- iii) we use the sensitivity analysis for defining a suitable parameter set Ξ_s that leads to significant variations of the cost functional;
- iv) we compute the RB Pareto optimal solutions for the parameter set Ξ_s defined in the previous step.

We note that the third step allows us to drastically reduce the number of online RB computations needed to recover a suitable distribution of the Pareto optimal solutions at a very low computational effort. Indeed, since we have the RB optimal solution for an initial parameter guess (one or more), we predict the value of the cost functional for every new parameter value. We point out that, by recalling the Taylor expansion of the cost functional, the main computations of this prediction depend on the initial parameter guess and mainly come from some precomputed quantity of its RB approximation.

6 Numerical examples

In this section we present numerical examples illustrating the efficiency of our proposed strategy. We consider different control input spaces and different geometric parameters. We start with a non parametric PDE constrain focusing on the optimal control solutions defined by varying the parameter weights of the cost functional. In the second example, we introduce a geometrical parameter leading to a parameter on the PDE constraint and we reduce the control space dimension in order to be able to show a feasible set of cost functional values (by varying the control) and the effectiveness (together with the sensitivity analysis) of the RB method for defining the Pareto optimal solutions. The numerical computations are performed in Matlab and the high-fidelity Galerkin approximation used is a finite element scheme with \mathbb{P}_1 piecewise elements.

6.1 Distributed multiobjective optimal control for the Laplace equation

We consider the domain Ω given by a rectangle separated into two subdomains Ω_1 and Ω_2 and represented in Figure 6.1.

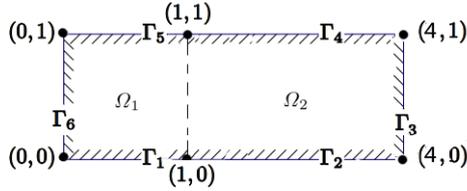


Fig. 6.1 Domain representation

We introduce the multiobjective optimization problem in which the vector cost functional is defined as follows:

$$J_1(y) = \frac{1}{2} \|y - w_1\|_{L^2(\Omega)}^2, \quad J_2(y) = \frac{1}{2} \|\nabla y\|_{L^2(\Omega)}^2, \quad J_3(u) = \frac{\alpha}{2} \|u\|_U^2 \quad (6.1)$$

where $w_1 = 1$ in Ω_1 and $w_1 = 0.6$ in Ω_2 . The state function $y \in Y = H_0^1(\Omega)$ solves the following Laplace problem:

$$\begin{cases} -\Delta y = u & \text{in } \Omega, \\ y = 1 & \text{on } \Gamma = \partial\Omega, \end{cases} \quad (6.2)$$

where $u \in L^2(\Omega)$ is the control function. We note that in this case we have an infinite dimensional control, that leads of course to a discrete control space $U^{\mathcal{N}} \subset U$ in the FE model, bu leading to our theoretical framework.

In order to apply the Pareto optimal theory we introduce the following cost functional:

$$\hat{J}(y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), \boldsymbol{\mu}) = \mu_1 J_1(y(\boldsymbol{\mu})) + \mu_2 J_2(y(\boldsymbol{\mu})) + (1 - \mu_1 - \mu_2) J_3(u(\boldsymbol{\mu})),$$

and the parametrized optimal control problem reads:

$$\min_{y,u} \hat{J}(y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), \boldsymbol{\mu}) \text{ s. t. } (y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U \text{ solves (6.2)}. \quad (6.3)$$

We apply the RB method to the saddle-point formulation of problem (6.3). In order to define the problem we introduce the space $Q = Y$. The numerical approximation of the reduced basis functions (state, control and adjoint variables) are done by using \mathbb{P}_1 finite elements. The dimension of the FE spaces $Y^{\mathcal{N}}, U^{\mathcal{N}}, Q^{\mathcal{N}}$ is $\mathcal{N}_1 = \mathcal{N}_2 = 11441$ obtained by using a mesh of 22528 elements. The parameters involved in the optimal control problem are defined in the following set $\mu_1 \in [0, 1]$ and $\mu_2 \in [0, 1 - \mu_1]$.

We consider three different parameter values leading to different cost functionals and different optimal solutions. The first one (Figure 6.2) induces a dominating first cost functional ($\boldsymbol{\mu}^1 = (0.9, 0)$), the second one (Figure 6.3) a dominating second cost functional ($\boldsymbol{\mu}^2 = (0.11, 0.83)$) and in the last one (Figure 6.4) we consider an equally distributed dominance between the three cost functionals ($\boldsymbol{\mu}^3 = (0.3, 0.3)$).

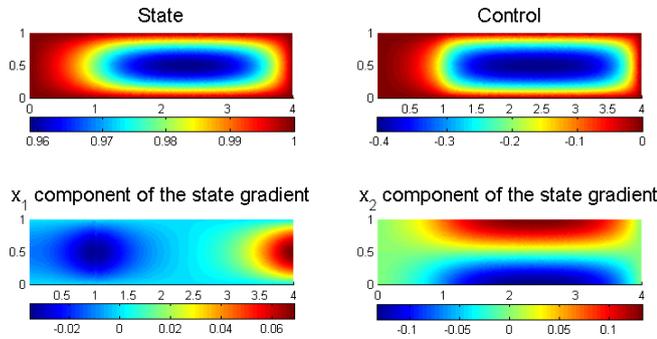


Fig. 6.2 Optimal control solution with $\boldsymbol{\mu} = (0.9, 0)$.

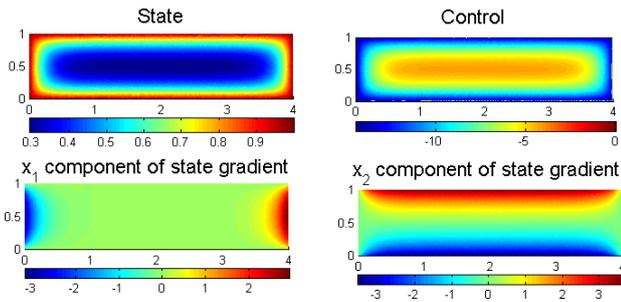


Fig. 6.3 Optimal control solution with $\boldsymbol{\mu} = (0.11, 0.83)$.

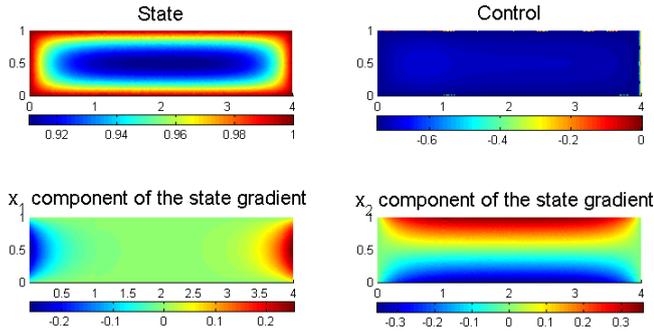


Fig. 6.4 Optimal control solution with $\boldsymbol{\mu} = (0.3, 0.3)$.

In Figure 6.5 we show the lower bound for the Babuska inf-sup constant $\beta^{\mathcal{N}}(\boldsymbol{\mu})$ obtained by varying the values of the parameter $\boldsymbol{\mu}$. Due to the smooth parameter dependence of Babuska inf-sup constant, we use in our RB computations, a linear interpolant surrogate of $\beta^{\mathcal{N}}(\boldsymbol{\mu})$.

Finally in Figure 6.6 we consider a set of 100 randomly selected parameters values Ξ_{test} , we compute the average and the maximum of the errors between the FE approximation of the solution and the RB one $\|x^{\mathcal{N}}(\boldsymbol{\mu}) - x^N(\boldsymbol{\mu})\|_X$, $\boldsymbol{\mu} \in \Xi_{test}$ and we compare them with the posteriori error bound $\Delta_N(\boldsymbol{\mu})$. We do the same comparison between the error bound $\Delta_N^J(\boldsymbol{\mu})$ and the error between the FE approximation of the cost functional and the RB one $|\hat{J}^{\mathcal{N}}(\boldsymbol{\mu}) - \hat{J}^N(\boldsymbol{\mu})|$, $\boldsymbol{\mu} \in \Xi_{test}$.

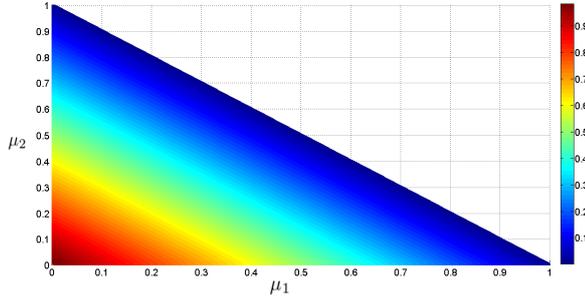


Fig. 6.5 Lower bound for the Babuska inf-sup constant $\beta^{\mathcal{N}}(\boldsymbol{\mu})$.

As regards the computational performances, the offline computational time is about 21 minutes, the online evaluation time by using $N = 15$ basis functions (in total: $2N$ for the state, N for the control and $2N$ for the adjoint) and including the evaluation of the a posteriori error bound is 0.016 seconds; while the evaluation of the FE solution requires about 1.26 seconds, by obtaining a speedup equal to 88,32. We show in Figures 6.7 and 6.8, respectively the RB computational time and the speedup with respect to a FE computational time by varying the number of basis functions.

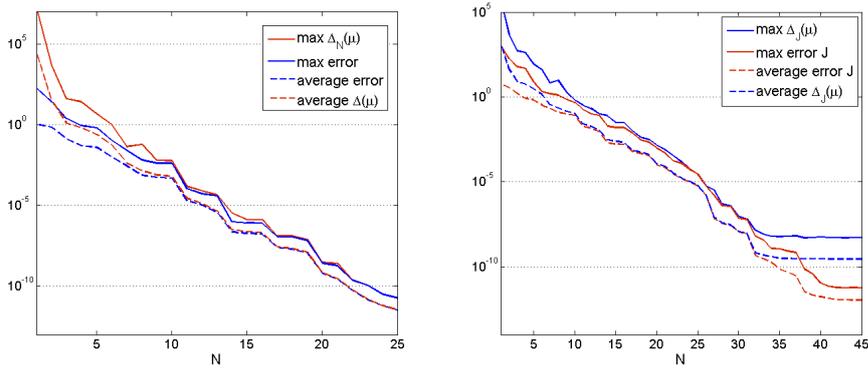


Fig. 6.6 Average and maximum errors and error bounds regarding the solution of the problem (left) and the cost functional (right) between the FE and RB approximations.

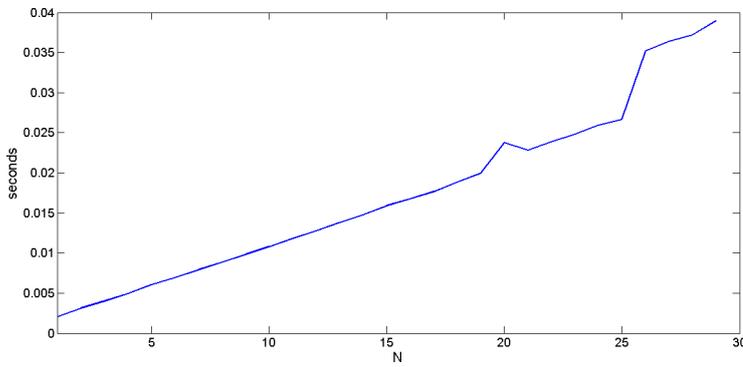


Fig. 6.7 Online evaluation time (in seconds) by varying the number N of basis functions and including the evaluation of the a posteriori error bound.

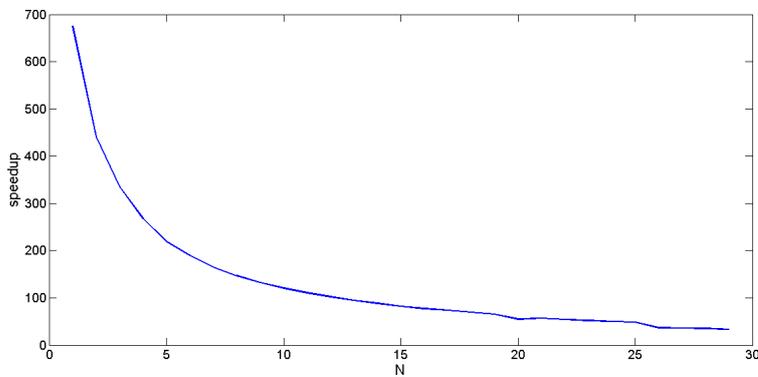


Fig. 6.8 Speedup between the FE computational time required for a single parameter value and the RB computational time required for a single parameter value by varying the number N of basis functions and including the evaluation of the a posteriori error bound .

We note that even if for this specific problem the FE computation is not particularly expensive, in order to define a suitable Pareto optimal point set, we have to find the optimal solution of the parametric problem several times (for many different parameter values) and the RB method permits to, for instance, find 88 Pareto points at the time of only one possible FE one. In the following example we consider a simple PDE in order to show this computational gain, particularly effective if added to a sensitivity analysis, that allows to further reduce the number of computations.

6.2 Multiobjective optimal control with smaller control space and a geometrical parameter

We consider the same problem of Section 6.1, this time by introducing a geometrical parameter $\mu_3 \in [1, 3.5]$ that defines the length of the domain Ω_{μ_3} still given by a rectangle separated in two subdomains Ω_1 and Ω_2 and represented in Figure 6.9.

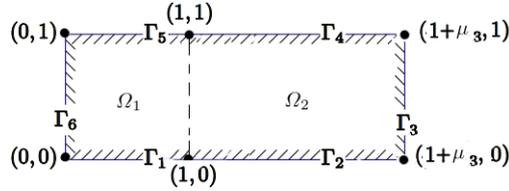


Fig. 6.9 Domain representation of Ω_{μ_3}

We consider a small control space, $U = \mathbb{R}^2$. We recall that the vector cost functional is defined as follows:

$$J_1(y) = \frac{1}{2} \|y - w_1\|_{L^2(\Omega_{\mu_3})}^2, \quad J_2(y) = \frac{1}{2} \|\nabla y\|_{L^2(\Omega_{\mu_3})}^2, \quad J_3(u) = \frac{\alpha}{2} \|u\|_U^2 \quad (6.4)$$

where $w_1 = 1$ in Ω_1 and $w_1 = 0.6$ in Ω_2 . The state function $y \in Y = H_0^1(\Omega_{\mu_3})$ solves the following Laplace problem:

$$\begin{cases} -\Delta y = \sum_{i=1}^2 u_i b_i & \text{in } \Omega_{\mu_3}, \\ y = 1 & \text{on } \Gamma_D = \partial\Omega_{\mu_3}, \end{cases} \quad (6.5)$$

where $u_i \in \mathbb{R}^2$ define the control function and b_1 and $b_2 \in L^\infty(\Omega)$ are the characteristic functions of Ω_1 and Ω_2 respectively.

As in the previous example, we apply the Pareto optimal theory by introducing the following cost functional:

$$\hat{J}(y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), \boldsymbol{\mu}) = \mu_1 J_1(y(\boldsymbol{\mu})) + \mu_2 J_2(y(\boldsymbol{\mu})) + (1 - \mu_1 - \mu_2) J_3(u(\boldsymbol{\mu})),$$

and the parametrized optimal control problem:

$$\min_{y,u} \hat{J}(y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), \boldsymbol{\mu}) \text{ s. t. } (y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U \text{ solves (6.5)}. \quad (6.6)$$

As in the previous example, we show some results regarding the using of the reduced basis method for solving the optimal control problem. In Figure 6.10 and 6.11 we show the RB solutions of the problem (6.6) for $\boldsymbol{\mu} = (0.2, 0.7, 3)$ and $\boldsymbol{\mu} = (0.2, 0.7, 1)$ respectively, the plots include the state function, the control function, the adjoint function and the gradient of the state function.

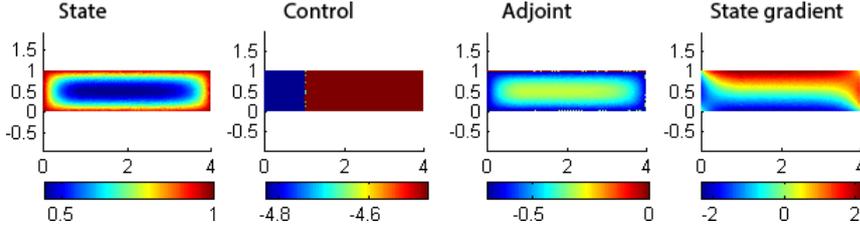


Fig. 6.10 Solution of the problem (6.6) for $\boldsymbol{\mu} = (0.2, 0.7, 3)$, the plot includes the state function, the control function, the adjoint function and the gradient of the state function.

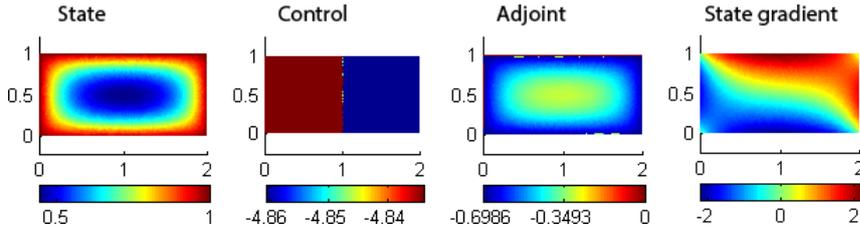


Fig. 6.11 Solution of the problem (6.6) for $\boldsymbol{\mu} = (0.2, 0.7, 1)$, the plot includes the state function, the control function, the adjoint function and the gradient of the state function.

In Figure 6.12 we show the error between the RB and FE solution of the optimal control problem over a set of 500 random samples by varying the number of basis functions.

From now we set $\mu_3 = 3$ and we focus our attention on the variation of the multiobjective parameters of the problem. In Figure 6.13 we show the values of the coefficients u_1, u_2 corresponding to the optimal control of the problem (6.6) by varying the parameters values μ_1 and μ_2 .

We consider a subset of the possible control functions such that $-30 \leq u_i \leq 10$ and we plot in Figure 6.14 the following set in \mathbb{R}^2 :

$$S = \{(w_1, w_2) \in \mathbb{R}^2 : w_1 = J_1(u), w_2 = J_2(u), u = \sum_{i=1}^2 u_i b_i, -30 \leq u_i \leq 10\} \quad (6.7)$$

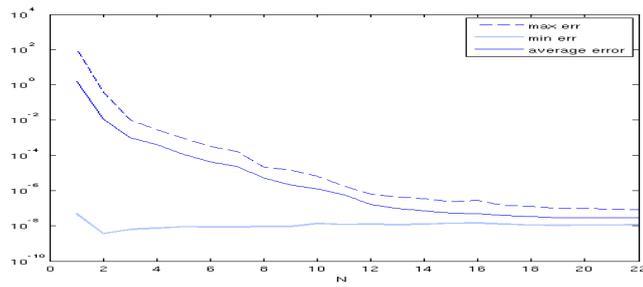


Fig. 6.12 Error (maximum, minimum and average) between the RB and FE solution of the optimal control problem over a set of 500 random samples by varying the number of basis functions.

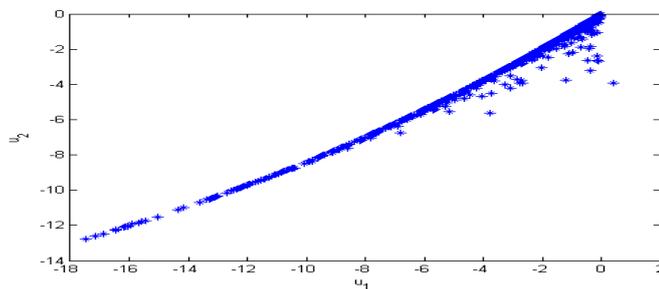


Fig. 6.13 Coefficients values of the optimal control functions solution of the problem (6.6) by varying the parameters values.

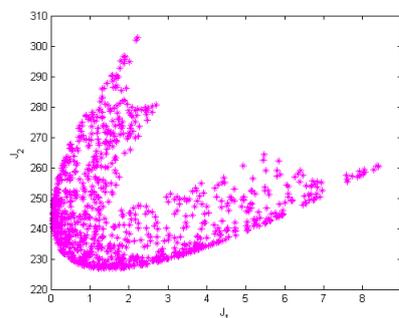


Fig. 6.14 Set of the possible values of the cost functionals J_1 and J_2 , by varying the function u .

Then, we solve with the reduced basis method the multiobjective optimal control by varying the parameter values μ_1 and μ_2 . We compute the Pareto optimal points and the corresponding efficient points of problem (6.6). In Figure 6.15 we plot the set S and the set of efficient points. In Figure 6.16 we

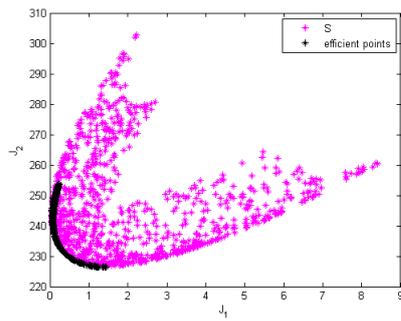


Fig. 6.15 Set of the possible values of the cost functionals J_1 and J_2 , by varying the function u and the subset of the efficient Pareto points.

show a similar plot, this time including the third cost functional and the corresponding efficient points.

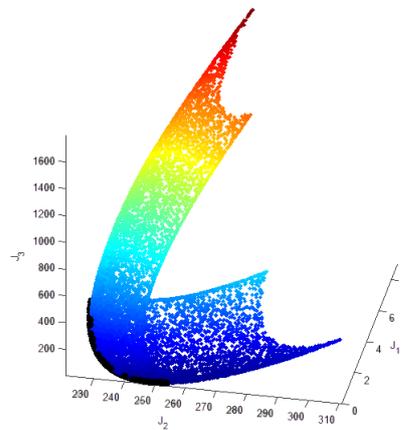


Fig. 6.16 Set of the possible values of the cost functionals J_1 , J_2 , J_3 by varying the function u and the subset of the efficient points.

6.2.1 Sensitivity analysis

We apply now the sensitivity analysis introduced in Section 6.18 with the purpose of showing how it helps in simplifying the computation of the set of Pareto optimal points.

Thanks to the inexpensive prediction of the value cost functional by varying the parameter μ_o , we are able to define a very small set of parameter values able to span the whole set of Pareto optimal points. In Figure 6.17 we show the interpolation of Pareto optimal solutions by selecting a random set of parameter values. We note that even with 100 parameter values we are not able

to cover the range of efficient points (see Figure 6.15 for a comparison). In Figure 6.18 we show the interpolation of Pareto optimal solutions in correspondence of the set of parameter values computed by the sensitivity analysis. We note that with only 20 RB solutions of the problem we are able to define the whole range of Pareto optimal solutions and a suitable prediction of the efficient points.

In conclusion, we point out that the RB solutions of the multiobjective problems are much faster than the FE ones, but thanks to sensitivity analysis we are able to further drastically reduce the number of the RB computations needed to define a suitable set of Pareto optimal solution of the problem.

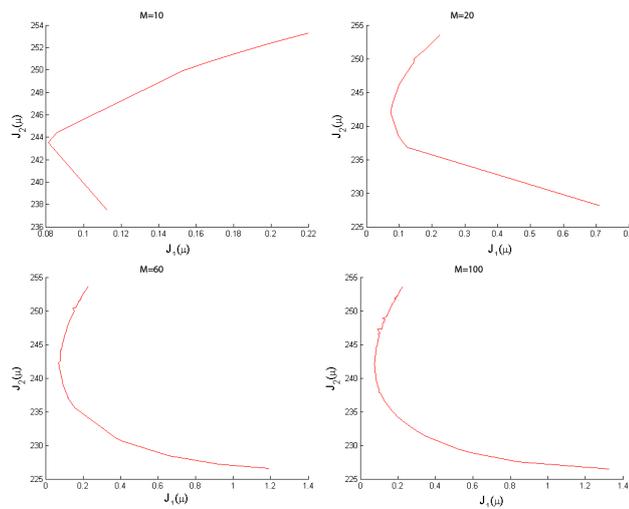


Fig. 6.17 Piecewise linear interpolation of M Pareto optimal points computed randomly.

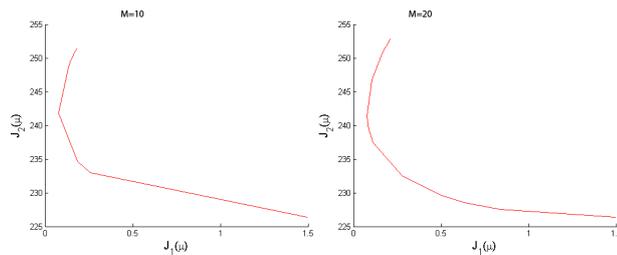


Fig. 6.18 Piecewise linear interpolation of M Pareto optimal points with the help of the sensitivity analysis.

7 Conclusions

In this work, a model order reduction strategy is proposed in the framework of the multiobjective optimal problems, characterized by more than one cost functional. Through the use of a parameter vector, the multiobjective problem leads to a parametric optimal control problem involving a new parametric cost functional defined as a weighted sum of the original cost functionals. For solving this problem we efficiently use the reduced basis method with the aim of provide a low order method for the computation of the set of Pareto solutions of the problem. Each Pareto optimal point represents the optimal controls corresponding to the problem considering a different weighted sum of the cost functionals, such that it is possible to predict the global set of solutions without compromising, a-priori, any cost functional. The use of the RB method, together with an useful and inexpensive sensitive analysis, allows to drastically reduce the computational times compared with other classical numerical techniques (e.g. finite elements). Moreover, a rigorous error bound analysis permits to ensure a certain level of accuracy of the solution.

Acknowledgments

This research was kindly supported by the EU FP7 Marie Curie Zukunftskolleg Incoming Fellowship Programme, University of Konstanz (grant no. 291784) and by the project *Reduced Basis Methods for Model Reduction and Sensitivity Analysis of Complex Partial Differential Equations with Applications to Lithium-Ion Batteries* funded by the Adam Opel AG.

References

1. M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 339(9):667–672, 2004.
2. F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Elements Methods*. In Springer-Verlag, New York, 1991.
3. I. Das. *newblock Nonlinear Multicriteria Optimization and Robust Optimality*. Ph.D thesis, Rice University, Houston, Texas, 1997.
4. L. Dede. Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. *SIAM Journal on Scientific Computing*, 32:997-1019, 2010.
5. L.C. Evans. *Partial Differential Equations*. American Math. Society, Providence, Rhode Island, 2008.
6. A. Göpfert and R. Nehse. *Vektoroptimierung*. BSB Teubner Verlagsgesellschaft, Leipzig, 1990.
7. C. Hillermeier. *Nonlinear Multiobjective Optimization. A Generalized Homotopy Approach*. Birkhäuser Verlag, Basel, 2001.
8. M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer, 2009.
9. D.B.P. Huynh, D.J. Knezevic, Y. Chen, J.S. Hesthaven and A.T. Patera. A natural-norm Successive Constraint Method for inf-sup lower bounds. *Computer Methods in Applied Mechanics and Engineering*, 199:1963-1975, 2010.

10. M. Grepl and M. Kärcher. A posteriori error estimation for reduced order solutions of parametrized parabolic optimal control problems. Submitted, 2013.
11. L. Iapichino. *Reduced basis methods for the solution of parametrized PDEs in repetitive and complex networks with application to CFD*. Ph.D thesis, École Polytechnique Fédérale de Lausanne, 2012. N. 5529, <http://infoscience.epfl.ch>.
12. L. Iapichino and S. Volkwein. Greedy sampling of distributed parameters in the reduced-basis method by numerical optimization. Submitted, 2013.
13. H. Kuhn and A. Tucker. Nonlinear programming. In Newman, J., editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 481-492, 1951.
14. F. Negri, G. Rozza, A. Manzoni, and A. Quateroni. Reduced basis method for parametrized elliptic optimal control problems. To appear in *SIAM Journal on Scientific Computing*, 2012.
15. J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Series in Operation Research, second edition, 2006.
16. A.T. Patera and G. Rozza. Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations. MIT Pappalardo Graduate Monographs in Mechanical Engineering, 2006.
17. G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Arch. Comput. Methods Engrg.*, 15, 229-275, 2008.
18. G. Rozza and K. Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *Comput. Meth. Appl. Mech. Engrg.*, 7:1244-1260, 2007.
19. W. Stadler. *Multicriteria Optimization in Engineering and in the Sciences*. Plenum Press, New York, 1988.
20. F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*. American Math. Society, Providence, volume 112, 2010.
21. F. Tröltzsch and S. Volkwein. POD a-posteriori error estimates for linear-quadratic optimal control problems. *Computational Optimization and Applications*, 44:83-115, 2009.
22. K. Urban, S. Volkwein and O. Zeeb. Greedy sampling using nonlinear optimization. To appear in *Proceedings of the CECAM Workshop on Reduced Basis, POD and Reduced Order Methods for model and computational reduction: towards real-time computing and visualization?*, Lausanne, 2012.
23. L. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, 8, 1963.