Multiobjective PDE-constrained optimization using the reduced-basis method

Iapichino, L.; Ulbrich, S.; Volkwein, S.

DOI
10.1007/s10444-016-9512-x

Publication date
2017

Document Version
Final published version

Published in
Advances in Computational Mathematics

Citation (APA)
Iapichino, L., Ulbrich, S., & Volkwein, S. (2017). Multiobjective PDE-constrained optimization using the reduced-basis method. *Advances in Computational Mathematics*, *43*(5), 945-972. https://doi.org/10.1007/s10444-016-9512-x

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Adv Comput Math (2017) 43:945–972
DOI 10.1007/s10444-016-9512-x

Multiobjective PDE-constrained optimization using the reduced-basis method

L. Iapichino[1] · S. Ulbrich[2] · S. Volkwein[3]

Received: 15 April 2015 / Accepted: 28 December 2016 /
Published online: 24 January 2017

Abstract In this paper the reduced basis (RB) method is applied to solve quadratic multiobjective optimal control problems governed by linear parametrized variational equations. These problems often arise in applications, where the quality of the system behavior has to be measured by more than one criterium. The weighted sum method is exploited for defining scalar-valued linear-quadratic optimal control problems built by introducing additional optimization parameters. The optimal controls corresponding to specific choices of the optimization parameters are efficiently computed by the RB method. The accuracy is guaranteed by an a-posteriori error estimate. An effective sensitivity analysis allows to further reduce the computational times for identifying a suitable and representative set of optimal controls.

Keywords Multiobjective PDE-constrained optimization · Weighted sum method · Reduced basis method · A-posteriori error · Sensitivity analysis

Mathematics Subject classification (2010) 35J20 · 49N10 · 65N30 · 90C29

Communicated by: Karsten Urban

✉ S. Volkwein
Stefan.Volkwein@uni-konstanz.de

L. Iapichino
L.Iapichino@tudelft.nl

[1] Department of Precision and Microsystems Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

[2] Department of Mathematics, Darmstadt University of Technology, Dolivostraße 15, Darmstadt, Germany

[3] Department of Mathematics and Statistics, University of Konstanz, Universitätsstraße 10, D-78457 Konstanz, Germany

1 Introduction

In real applications, optimization problems are often described by introducing several objective functions conflicting with each other. This leads to *multiobjective* or *multicriterial* optimization problems; see, e.g., [7, 9, 11, 31]. Finding the optimal control that represents a good compromise is the main issue in these problems. For that reason the concept of Pareto optimal or efficient points is developed. In contrast to scalar-valued optimization problems, the computation of a set of Pareto optimal points is required. Consequently, many scalar-valued constrained optimization problems have to be solved.

In the present paper, we study infinite dimensional multiobjective optimization problems governed by $k \geq 2$ quadratic objectives and by linear parametrized variational constraints described by a weak formulation of a linear parametrized elliptic partial differential equation (PDE). For the numerical solution, we apply the weighted sum method [35], where parametrized scalar-valued quadratic programming problems have to be solved for many parameter values in order to define a sufficiently accurate approximation of the set of Pareto optimal points. Note that we have two sorts of parameters in our problem: parameters related to the variational constraints and parameters required by the scalarization of the multiobjective problem through the weighted sum method. Our scalar-valued parametrized optimal control problems are infinite dimensional optimization problems (see [16, 32] for PDE constrained problems). After a discretization by a high-fidelity (HF) discretization technique like finite elements or finite volumes, the Pareto points can be computed by solving many large scale parametrized optimization problems. Therefore, we make use of a reduced basis (RB) approximation [26], which is known to be very efficient for parametrized linear-quadratic optimal control problems [8, 12]. Let us also refer to [33] for POD based Galerkin schemes for time-dependent linear-quadratic optimal control problems, the latter work is extended to nonlinear problems in [22].

The contribution of the present work is a successful combination of the greedy approach for parametrized linear-quadratic optimal control problems [24] with the weighted sum method. In each iteration of the greedy method appropriate Pareto points to the HF multiobjective optimization problem are computed and used as basis functions in the reduced-order discretization. For the construction of an accurate RB scheme, we apply the a-posteriori error analysis presented in [24]. For each value of the parameters, the RB method allows to drastically reduce the computational times required for the optimal solution of the corresponding scalar-valued quadratic programming problem. The complete set of Pareto points is defined by solving the problem for all the parameter values. A sensitivity analysis for the objectives allows us to reduce the number of scalar-valued optimization solutions in the weighted sum method. To sum up, our strategy allows to compute a sufficiently accurate RB approximation of the set of Pareto optimal points for an arbitrarily chosen parameter in the variation constraints. The present approach is already utilized for more general problems including time-dependence and semilinear state equations in [21]. Further preliminary results combining reduced-order modeling and multiobjective PDE-constrained optimization are recently derived in [2, 27], where different optimization methods are used.

The paper is organized as follows: in Section 2 we introduce the multiobjective optimal control problem and we recall the definition of efficient points and Pareto optimal points. The first-order optimality conditions for the scalar-valued parametrized minimization problems are formulated as a saddle point problem in Section 3. Here we also introduce a HF 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 RB method and, finally, the a-posteriori error estimate from [24] is formulated for our problem. In Section 5 the sensitivity analysis is described. Numerical experiments are shown in Section 6. Finally we draw some conclusions in Section 7.

2 The multiobjective optimization problem

In this section we introduce a constrained multiobjective optimal control problem. The equality constraint is given by a parametrized affine variational equation, which stands for a weak formulation of a parameter-dependent linear elliptic partial differential equation. The parameter can refer to geometrical or model features of the variational equation.

2.1 The state equation

Let V and H be real, separable Hilbert spaces and suppose that V is dense in H with compact embedding. We denote by $\langle \cdot\,, \cdot \rangle_H$ and $\langle \cdot\,, \cdot \rangle_V$ the inner products in H and V, respectively.

The set $\mathcal{D}_c \subset \mathbb{R}^{n_c}$, $n_c \in \mathbb{N}$, stands for all considered *geometrical* and/or *model parameters*. We suppose that for every parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$ the parameter-dependent symmetric bilinear form $a(\cdot\,, \cdot\,; \boldsymbol{\mu}_c) : V \times V \to \mathbb{R}$ satisfies

$$\begin{aligned} &\inf_{\varphi \in V} \frac{a(\varphi, \varphi; \boldsymbol{\mu}_c)}{\|\varphi\|_V^2} \geq \eta_1 && \text{for all } \varphi \in V, \\ &\left| a(\varphi, \phi; \boldsymbol{\mu}_c) \right| \leq \eta_2 \, \|\varphi\|_V \|\phi\|_V && \text{for all } \varphi, \phi \in V, \end{aligned} \tag{1}$$

where $\eta_1 > 0$ and $\eta_2 \geq 0$ are 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. The parameter-dependent inhomogeneity $f(\boldsymbol{\mu}_c)$ is supposed to belong to V' for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

We assume that the set U of admissible controls is a real, separable Hilbert space endowed with the inner product $\langle \cdot\,, \cdot \rangle_U$ and the induced norm $\| \cdot \|_U = \langle \cdot\,, \cdot \rangle_U^{1/2}$. For a parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$ and for a control $u \in U$, the state $y \in V$ solves the following linear elliptic and coercive variational problem

$$a(y, \varphi; \boldsymbol{\mu}_c) = \langle f(\boldsymbol{\mu}_c) + \mathcal{E} u, \varphi \rangle_{V',V} \quad \text{for all } \varphi \in V, \tag{2}$$

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 $L(U, V')$ of all bounded, linear operators from U to V' equipped with the norm

$$\|\mathcal{E}\|_{L(U,V')} = \sup \left\{ \|\mathcal{E}u\|_{V'} \,\middle|\, u \in U \text{ and } \|u\|_U = 1 \right\}.$$

The following result follows from the standard variational theory; see, e.g., [10, Chapter 6].

Proposition 2.1 *Suppose that* Eq. 1 *and* $\mathcal{E} \in L(U, V')$ *hold. Then, for every* $\boldsymbol{\mu}_c \in \mathcal{D}_c$, $u \in U$ *and* $f(\boldsymbol{\mu}_c) \in V'$ *there is a unique weak solution* $y = y(\boldsymbol{\mu}_c) \in V$ *satisfying* (2) *and*

$$\|y\|_V \leq \eta \left(\|f(\boldsymbol{\mu}_c)\|_{V'} + \|u\|_U \right) \tag{3}$$

for the constant $\eta = (1 + \|\mathcal{E}\|_{L(U,V')})/\eta_1 > 0$.

Remark 2.2 (Control-to-state operator) Let $\boldsymbol{\mu}_c \in \mathcal{D}_c$ be chosen arbitrarily. Due to Proposition 2.1 there exists a unique solution $\hat{y} = \hat{y}(\boldsymbol{\mu}_c) \in V$ 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 $\mathcal{S} = \mathcal{S}(\boldsymbol{\mu}_c) : U \to V$ as follows: $y = \mathcal{S}(\boldsymbol{\mu}_c)u$, thanks again to Proposition 2.1, 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). It follows from Eq. 3 that the operator $\mathcal{S}$ is bounded for every $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

2.2 The vector-valued cost functional

We set $X = V \times U$ and introduce the following vector-valued objective $J : X \to \mathbb{R}^k$

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

where $x = (y, u) \in X$, $W_1, \ldots, W_{k-1}$ are (real) 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 2.3 Let us give an application which is utilized in our numerical experiments carried out in Section 6. Suppose that Ω is an open and bounded domain in $\mathbb{R}^2$ with Lipschitz-continuous boundary $\Gamma = \partial\Omega$. We set $H = L^2(\Omega)$, $H^2 = H \times H$ and $V = H^1(\Omega)$. For more details on Lebesgue and Sobolev spaces we refer the reader to [10, Chapter 5], for instance. Let $k = 3$ and $W_1 = H$, $W_2 = H \times H$. The operator $\mathcal{C}_1$ is the canonical embedding from V into H, the mapping $\mathcal{C}_2$ is given by $\mathcal{C}_2 y = \nabla y \in W_2$ 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

$$J_1(x) = \frac{1}{2} \int_\Omega \left| y(\boldsymbol{x}) - w_1(\boldsymbol{x}) \right|^2 \mathrm{d}\boldsymbol{x} \quad \text{and} \quad J_2(x) = \frac{1}{2} \int_\Omega \left| \nabla y(\boldsymbol{x}) \right|^2_2 \mathrm{d}\boldsymbol{x},$$

where $|\cdot|_2$ denotes the Euclidean norm in $\mathbb{R}^2$.

Due to Remark 2.2 the state $y = \mathcal{S}(\boldsymbol{\mu}_c)u$ is uniquely determined by a control $u \in U$ for any parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$. This is an intrinsic property of optimal control problems. Motivated by this fact we reduce the number of optimization variables by eliminating the control-dependent state variable in the objective: We define the parameter-dependent *reduced cost functional* $\hat{J}(\cdot\,; \boldsymbol{\mu}_c) : U \to \mathbb{R}^k$ for any $\boldsymbol{\mu}_c \in \mathcal{D}_c$ by

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

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

2.3 The multiobjective optimal control problem

To define our multiobjective optimization problem the concepts of order relation and Pareto optimality is needed; see, e.g., Chapter 3 in [15].

Definition 2.4 (Order relation) Let $(\mathbb{R}^k, \leq)$ denote the *order relation in* $\mathbb{R}^k$ defined by

$$z^1 \leq z^2 \quad \Leftrightarrow \quad z^2 - z^1 \in \mathbb{R}^k_+ = \left\{ z \in \mathbb{R}^k \,\middle|\, z_i \geq 0 \text{ for } i = 1, \ldots, k \right\}$$

for all $z^1, z^2 \in \mathbb{R}^k$.

Definition 2.5 (Pareto optimal) Let $\mathcal{Z} = \hat{J}(U; \boldsymbol{\mu}_c) \subset \mathbb{R}^k$ be the image set of U under the cost functional $\hat{J}(\cdot\,; \boldsymbol{\mu}_c)$ for a given $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

1) We call a point $\bar{z} \in \mathcal{Z}$ *globally efficient* with respect to the order relation $\leq$, if there exists no $z \in \mathcal{Z} \setminus \{\bar{z}\}$ with $z \leq \bar{z}$.
2) If $\bar{z}$ is efficient and $\bar{u} \in U$ satisfies $\bar{z} = \hat{J}(\bar{u}; \boldsymbol{\mu}_c)$, we call $\bar{u}$ *Pareto optimal*.
3) Let $\bar{u} \in U$ hold. If there exists a neighborhood $\mathcal{N}(\bar{u}) \subset U$ of $\bar{u}$ so that $\bar{z} = \hat{J}(\bar{u}; \boldsymbol{\mu}_c)$ is efficient for the (local) image set $\hat{J}(\mathcal{N}(\bar{u}); \boldsymbol{\mu}_c) \subset \mathcal{Z}$, the point $\bar{u}$ is called *locally Pareto optimal*. Moreover, $\bar{z}$ is said to be *locally efficient*.

The parametrized *multiobjective optimal control problem* can be defined as follows: Find Pareto optimal points for the vector-valued reduced cost functional $\hat{J}(\cdot\,; \boldsymbol{\mu}_c)$ for any $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

2.4 First-order optimality conditions

The cost functional $\hat{J}(\cdot\,; \boldsymbol{\mu}_c)$ 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 [9, Theorem 3.21 and Corollary 3.23]. The proof is based on the result of Kuhn-Tucker [23].

Theorem 2.6 *Suppose that $\bar{u} \in U$ is Pareto optimal. Then, there exists a parameter vector $\bar{\boldsymbol{\mu}}_o = (\bar{\mu}_{0,1}, \ldots, \bar{\mu}_{o,k}) \in \mathbb{R}^k$ satisfying the Karush-Kuhn-Tucker conditions*

$$0 \leq \bar{\mu}_{o,i} \leq 1, \quad \sum_{i=1}^{k} \bar{\mu}_{o,i} = 1 \quad \text{and} \quad \sum_{i=1}^{k} \bar{\mu}_{o,i} \hat{J}_i'(\bar{u}; \boldsymbol{\mu}_c) = 0. \tag{4}$$

Since all k components of $\hat{J}(\cdot\,; \boldsymbol{\mu}_c)$ are convex on U, Eq. 4 *is also a sufficient condition for $\bar{u}$ to be Pareto optimal.*

Motivated by Theorem 2.6, let us choose $0 < \mu_{lb} \ll 1$ and define the set

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

for the *optimization parameters* in Eq. 4. Let us mention that the condition $\mu_{o,k} \geq \mu_{lb}$ will be necessary for the well-posedness of the scalar-valued optimal problem ($\hat{\mathbf{P}}_{\boldsymbol{\mu}}$) introduced below. Moreover, we combine the optimization parameters with the parameters involved in the state Eq. 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, quadratic, scalar-valued cost functional as

$$\hat{\mathfrak{J}}(u; \boldsymbol{\mu}) = \sum_{i=1}^{k} \mu_{o,i} \hat{J}_i(u; \boldsymbol{\mu}_c) = \boldsymbol{\mu}_o^\top \hat{J}(u; \boldsymbol{\mu}_c) \quad \text{for } u \in U,$$

where the symbol '$\top$' stands for the transpose of a vector or matrix. Then, Eq. 4 are the first-order suffiient optimality conditions for a local solution $\bar{u} = \bar{u}(\boldsymbol{\mu})$ to the parameter-dependent quadratic optimization problem

$$\min \hat{\mathfrak{J}}(u; \boldsymbol{\mu}) \quad \text{subject to (s.t.)} \quad u \in U \tag{$\hat{\mathbf{P}}_{\boldsymbol{\mu}}$}$$

for the parameter $\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}$. In the weighted sum method – first introduced by Zadeh [35] – Pareto optimal points are computed by solving ($\hat{\mathbf{P}}_{\boldsymbol{\mu}}$) for various $\boldsymbol{\mu}_o \in \mathcal{D}_o$; see [9, Chapter 3], for instance. To solve ($\hat{\mathbf{P}}_{\boldsymbol{\mu}}$) we can apply methods from quadratic programming; see, e.g., [25, Chapter 16].

Remark 2.7 1) We apply the weighted sum method in this paper to compute Pareto optimal points for a given parameter $\boldsymbol{\mu}_c \in \mathcal{D}_c$. The computation of the Pareto optimal points requires the solution of many scalar-valued optimization problems. Since we are interested in the Pareto optimal points for any $\boldsymbol{\mu}_c \in \mathcal{D}_c$, we have to solve a very large number of optimal control problems. For this reason we propose a reduced-order approach using the reduced-basis method.

2) The reduced basis approximation allows a very fast computation of an approximate (i.e., suboptimal) solution to Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$. Moreover, we can estimate the error between the suboptimal and the (unknown) optimal solution to Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$. Therefore, we can ensure that our computed (suboptimal) Pareto optimal points are sufficiently accurate.

3) Note that $\mu_{o,k} = 1 - \sum_{i=1}^{k-1} \mu_{o,i}$ holds, which is utilized in our numerical experiments. Nevertheless, for sake of notation simplicity $\mu_{o,1}, \ldots, \mu_{o,k}$ will be considered as independent parameters.

3 Optimality system for the scalar-valued 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 RB method.

3.1 Saddle point formulation for Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$

First we mention that Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$ is a quadratic programming problem. We set $Z = X \times V$. Let us define the bilinear forms

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

for all $x = (y, u) \in X$ and $(\tilde{x}, \tilde{p}) \in Z$ with $\tilde{x} = (\tilde{y}, \tilde{u})$.

Lemma 3.1 *Suppose that* Eq. 1, $\mathcal{E} \in L(U, V')$, $\gamma > 0$, $\boldsymbol{\mu} \in \mathcal{D}$ *and* $\mathcal{C}_i \in L(V, W_i)$ *for* $1 \le i \le k-1$. *Then:*

1) *The mapping* $X \ni x \mapsto \mathcal{B}(x, \cdot\,; \boldsymbol{\mu}) \in V'$ *is continuous and surjective.*
2) *For the constant* $\eta > 0$ *introduced in* Eq. 3 *we have*

$$\|\tilde{y}\|_V \le \eta \, \|\tilde{u}\|_U \quad \text{for all } \ (\tilde{y}, \tilde{u}) \in X_0, \tag{1}$$

where we define $X_0 = \{x \in X \mid \mathcal{B}(x, \tilde{p}; \boldsymbol{\mu}) = 0 \text{ for all } \tilde{p} \in V\} \subset X$.
3) *The bilinear form* $\mathcal{A}(\cdot\,, \cdot\,; \boldsymbol{\mu})$ *is continuous on* $X \times X$ *and coercive on* X_0. *In particular, we have*

$$\mathcal{A}(\tilde{x}, \tilde{x}; \boldsymbol{\mu}) \ge \alpha \, \|\tilde{x}\|_X^2 \quad \text{for all } \ \tilde{x} \in X_0, \tag{2}$$

where $\alpha = \mu_{lb} \gamma \min(1/\eta, 1)/2$ *is independent of* $\boldsymbol{\mu}$.

Proof 1) The continuity follows directly from Eq. 1 and $\mathcal{E} \in L(U, V')$. To verify that $x \mapsto \mathcal{B}(x, \cdot\,; \boldsymbol{\mu})$ is surjective we have to show that there exists an element $x = (y, u) \in X$ such that $\mathcal{B}(x, \cdot\,; \boldsymbol{\mu}) = F$ holds in V' for any $F \in V'$. From $\mathcal{B}(x, \cdot\,; \boldsymbol{\mu}) = F$ it follows that

$$a(y, p; \boldsymbol{\mu}_c) - \langle \mathcal{E} u, p \rangle_{V',V} = \langle F, p \rangle_{V',V} \quad \text{for all } p \in V. \tag{3}$$

Equation 3 coincides – after replacing F by $f(\boldsymbol{\mu}_c)$ – with the state (2), which is uniquely solvable. Hence, for any $u \in U$ there exists a unique state $y(u) \in V$ satisfying $\mathcal{B}(x, \cdot\,; \boldsymbol{\mu}) = F$ in V' with $x = (y(u), u)$, i.e., the mapping $x \mapsto \mathcal{B}(x, \cdot\,; \boldsymbol{\mu})$ is surjective from X to V'.

2) Let $p \in V$ be chosen arbitrarily and $(\tilde{y}, \tilde{u}) \in X_0$. Then, $a(\tilde{y}, p; \boldsymbol{\mu}_c) = \langle \mathcal{E}\tilde{u}, p\rangle_{V',V}$ holds, which is the state equation for the choice $f(\boldsymbol{\mu}_c) = 0$. Thus, Eq. 1 follows directly from Eq. 3.

3) The continuity of $\mathcal{A}(\cdot\,, \cdot\,; \boldsymbol{\mu})$ follows directly from $\mathcal{C}_i \in L(V, W_i)$, $1 \le i \le k-1$. Let $(x, p) \in Z$ hold. By definition of the parameter set $\mathcal{D}_o$ we have $\mu_{o,k} \ge \mu_{lb} > 0$. Moreover, γ is positive. Utilizing (1), it follows that

$$\begin{aligned}\mathcal{A}(\tilde{x}, \tilde{x}; \boldsymbol{\mu}) &= \sum_{i=1}^{k-1} \mu_{o,i} \, \|\mathcal{C}_i \tilde{y}\|^2_{W_i} + \mu_{o,k}\gamma \, \|\tilde{u}\|^2_U \ge \mu_{o,k}\gamma \, \|\tilde{u}\|^2_U \ge \mu_{lb}\gamma \, \|\tilde{u}\|^2_U \\ &\ge \frac{\mu_{lb}\gamma}{2}\left(\frac{1}{\eta}\,\|\tilde{y}\|^2_V + \|\tilde{u}\|^2_U\right) \quad \text{for all } \tilde{x} = (\tilde{y}, \tilde{u}) \in X_0,\end{aligned}$$

which imply (2).

□

Remark 3.2 From closed range theory [4, Chapter 2] and Lemma 3.1-1) we infer that

$$\beta(\boldsymbol{\mu}) := \inf_{p \in V} \sup_{x \in X} \frac{\mathcal{B}(x, p; \boldsymbol{\mu})}{\|x\|_X \|p\|_V} > 0 \quad \text{for all } \boldsymbol{\mu} \in \mathcal{D}. \tag{4}$$

The condition (4) is called the *Brezzi inf-sup-condition* and $\beta(\boldsymbol{\mu})$ the *Brezzi inf-sup constant*; see [5]. It follows from [24, Lemma 2.1] that $\beta(\boldsymbol{\mu}) \ge \alpha$ is valid for all $\boldsymbol{\mu} \in \mathcal{D}$. Hence the Brezzi inf-sup constant is bounded from below by a positive constant which is independent of the parameter $\boldsymbol{\mu}$.

Utilizing Lemma 3.1 the existence of a unique solution $\bar{u} = \bar{u}(\boldsymbol{\mu})$ for Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$ can be proved in a standard way for any $\boldsymbol{\mu} \in \mathcal{D}$; see, e.g., [16, 32]. We introduce the $\boldsymbol{\mu}$-dependent Lagrangian functional for Eq. $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$ as

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

where $x = (y, u) \in X$ stands for the *primal variable*, $p \in V$ is the *Lagrange multiplier* (or *adjoint variable*) associated with the equality constraint (2) and $\boldsymbol{\mu} \in \mathcal{D}$ holds.

The optimal solution $\bar{u}$ can be characterized by first-order sufficient optimality conditions: $\bar{u}$ satisfies together with the unique associated optimal state $\bar{y} = \bar{y}(\boldsymbol{\mu})$ and optimal adjoint $\bar{p} = \bar{p}(\boldsymbol{\mu})$ the coupled linear equation system

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

for all directions $(x, p) \in Z$ with $x = (y, u)$. In order to write the first-order optimality conditions in a more compact form, we define the two linear forms

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

for any $x = (y, u) \in X$ and $(\tilde{x}, \tilde{p}) \in Z$ with $\tilde{x} = (\tilde{y}, \tilde{u})$. The first-order optimality conditions can be expressed as follows: find $(\bar{x}(\boldsymbol{\mu}), \bar{p}(\boldsymbol{\mu})) \in Z$ such that

$$
\begin{aligned}
\mathcal{A}(\bar{x}(\boldsymbol{\mu}), x; \boldsymbol{\mu}) + \mathcal{B}(x, \bar{p}(\boldsymbol{\mu}), \boldsymbol{\mu}) &= \mathcal{F}(x; \boldsymbol{\mu}) \quad \text{for all } x \in X, \\
\mathcal{B}(\bar{x}(\boldsymbol{\mu}), p; \boldsymbol{\mu}) &= \mathcal{G}(p; \boldsymbol{\mu}) \quad \text{for all } p \in V.
\end{aligned} \tag{5}
$$

Proposition 3.3 *Let all hypotheses of Lemma* 3.1 *be satisfied. Then,* Eq. 5 *admits a unique solution* $(\bar{x}(\boldsymbol{\mu}), \bar{p}(\boldsymbol{\mu})) \in Z$ *with* $\bar{x}(\boldsymbol{\mu}) = (\bar{y}(\boldsymbol{\mu}), \bar{u}(\boldsymbol{\mu}))$ *for any parameter* $\boldsymbol{\mu} \in \mathcal{D}$.

Proof The claim follows from the Brezzi theorem [6, Chapter II.1.1] and Lemma 3.1. □

Remark 3.4 Note that Eq. 5 involves optimization parameters $\boldsymbol{\mu}_o \in \mathcal{D}_o$ as well as geometrical or model parameters $\boldsymbol{\mu}_c \in \mathcal{D}_c$.

3.2 High-fidelity (HF) Galerkin approximation

The parametrized problem (5) is an infinite-dimensional saddle point problem which has to be discretized for computing its numerical solution. Hence, we introduce a HF Galerkin approximation of Eq. 5. Let us assume that $\{\varphi_i\}_{i=1}^{\mathcal{N}_1}$ and $\{\phi_i\}_{i=1}^{\mathcal{N}_2}$ denote sets of linearly independent basis functions in V and U, respectively, where $\mathcal{N}_1 \in \mathbb{N}$ and $\mathcal{N}_2 \in \mathbb{N}$ are typically very large. We set $\mathcal{N}_{12} = \mathcal{N}_1 + \mathcal{N}_2$. We introduce the finite dimensional spaces:

$$
V^{\mathcal{N}} = \operatorname{span}\left\{\varphi_1, \ldots, \varphi_{\mathcal{N}_1}\right\} \subset V, \quad U^{\mathcal{N}} = \operatorname{span}\left\{\phi_1, \ldots, \phi_{\mathcal{N}_2}\right\} \subset U.
$$

The high-fidelity Galerkin approximation of the optimality system (5) reads as follows: find $(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}), \bar{p}^{\mathcal{N}}(\boldsymbol{\mu})) \in Z^{\mathcal{N}}$ such that

$$\begin{aligned} \mathcal{A}(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}), x^{\mathcal{N}}; \boldsymbol{\mu}) + \mathcal{B}(x^{\mathcal{N}}, \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}), \boldsymbol{\mu}) &= \mathcal{F}(x^{\mathcal{N}}; \boldsymbol{\mu}) \quad \text{for all } x^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}; \boldsymbol{\mu}) &= \mathcal{G}(p^{\mathcal{N}}; \boldsymbol{\mu}) \quad \text{for all } p^{\mathcal{N}} \in V^{\mathcal{N}}, \end{aligned} \tag{6}$$

where we have set $Z^{\mathcal{N}} = X^{\mathcal{N}} \times V^{\mathcal{N}}$, $X^{\mathcal{N}} = V^{\mathcal{N}} \times U^{\mathcal{N}}$, $\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}) = (\bar{y}^{\mathcal{N}}(\boldsymbol{\mu}), \bar{u}^{\mathcal{N}}(\boldsymbol{\mu}))$ and

$$\begin{aligned} \bar{y}^{\mathcal{N}}(\boldsymbol{\mu}) &= \sum_{i=1}^{\mathcal{N}_1} \bar{\mathrm{x}}_i^{\mathcal{N}}(\boldsymbol{\mu})\varphi_i = \sum_{i=1}^{\mathcal{N}_1} \bar{\mathrm{y}}_i^{\mathcal{N}}(\boldsymbol{\mu})\varphi_i, \quad \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_1} \bar{\mathrm{p}}_i^{\mathcal{N}}(\boldsymbol{\mu})\varphi_i, \\ \bar{u}^{\mathcal{N}}(\boldsymbol{\mu}) &= \sum_{i=1}^{\mathcal{N}_2} \bar{\mathrm{x}}_{\mathcal{N}_1+i}^{\mathcal{N}}(\boldsymbol{\mu})\phi_i = \sum_{i=1}^{\mathcal{N}_2} \bar{\mathrm{u}}_i^{\mathcal{N}}(\boldsymbol{\mu})\phi_i. \end{aligned}$$

The following results can be derived by the same arguments in the proof of Lemma 3.1.

Lemma 3.5 *Suppose that* Eq. 1, $\mathcal{E} \in L(U, V')$, $\gamma > 0$, $\boldsymbol{\mu} \in \mathcal{D}$ *and* $\mathcal{C}_i \in L(V, W_i)$ *for* $1 \le i \le k-1$. *Then it follows:*

1) *The mapping* $X^{\mathcal{N}} \ni x^{\mathcal{N}} \mapsto \mathcal{B}(x^{\mathcal{N}}, \cdot\,; \boldsymbol{\mu}) \in (V^{\mathcal{N}})'$ *is continuous and surjective.*
2) *The bilinear form* $\mathcal{A}(\cdot\,, \cdot\,; \boldsymbol{\mu})$ *is continuous on* $X^{\mathcal{N}} \times X^{\mathcal{N}}$ *and coercive on the subspace* $X_0^{\mathcal{N}} = \{x^{\mathcal{N}} \in X^{\mathcal{N}} \,\big|\, \mathcal{B}(x^{\mathcal{N}}, p^{\mathcal{N}}; \boldsymbol{\mu}) = 0 \text{ for all } p^{\mathcal{N}} \in V^{\mathcal{N}}\}$, *i.e.,*

$$\mathcal{A}(\tilde{x}^{\mathcal{N}}, \tilde{x}^{\mathcal{N}}; \boldsymbol{\mu}) \ge \alpha \, \|\tilde{x}^{\mathcal{N}}\|_X^2 \quad \text{for all } \tilde{x}^{\mathcal{N}} \in X_0^{\mathcal{N}},$$

where the coercivity constant α *is the same as in* Eq. 2.

Remark 3.6 As in Remark 3.2 it follows from closed range theory and Lemma 3.5-1) that the *Brezzi HF inf-sup condition* hold:

$$\beta^{\mathcal{N}}(\boldsymbol{\mu}) := \inf_{p^{\mathcal{N}} \in V^{\mathcal{N}}} \sup_{x^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{B}(x^{\mathcal{N}}, p^{\mathcal{N}}; \boldsymbol{\mu})}{\|x^{\mathcal{N}}\|_X \|p^{\mathcal{N}}\|_V} > 0 \quad \text{for all } \boldsymbol{\mu} \in \mathcal{D}.$$

Since the state $y^{\mathcal{N}}$ and the adjoint $p^{\mathcal{N}}$ belong to the same subspace $V^{\mathcal{N}}$ we derive analogously to the proof of Lemma 2.1 in [24] that $\beta^{\mathcal{N}}(\boldsymbol{\mu}) \ge \alpha > 0$, where $\alpha = \mu_{lb}\gamma \min(1/\eta, 1)/2$ is independent of $\mathcal{N}$ and $\boldsymbol{\mu}$.

Proposition 3.7 *Let all hypotheses of Lemma* 3.5 *hold. Then,* Eq. 6 *possesses a unique solution* $(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}), \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}))$ *for any parameter* $\boldsymbol{\mu} \in \mathcal{D}$.

Proof The existence of a unique solution to Eq. 6 is ensured by the Brezzi theorem [6, Chapter II.1.1] and Lemma 3.5. □

The numerical solution of problem (6) leads to a linear algebraic system of dimension $\mathcal{N}_{12} + \mathcal{N}_1$ of the following structure:

$$\begin{aligned} \sum_{j=1}^{\mathcal{N}_{12}} A_{ij}^{\mathcal{N}}(\boldsymbol{\mu})\bar{\mathrm{x}}_j^{\mathcal{N}}(\boldsymbol{\mu}) + \sum_{j=1}^{\mathcal{N}_1} B_{ji}^{\mathcal{N}}(\boldsymbol{\mu})\bar{\mathrm{p}}_j^{\mathcal{N}}(\boldsymbol{\mu}) &= F_i^{\mathcal{N}}(\boldsymbol{\mu}), \quad i = 1, \ldots, \mathcal{N}_{12}, \\ \sum_{j=1}^{\mathcal{N}_{12}} B_{ij}^{\mathcal{N}}(\boldsymbol{\mu})\bar{\mathrm{x}}_j^{\mathcal{N}}(\boldsymbol{\mu}) &= G_i^{\mathcal{N}}(\boldsymbol{\mu}), \quad i = 1, \ldots, \mathcal{N}_1, \end{aligned} \tag{7}$$

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

$$A_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) = \begin{cases} \mathcal{A}\big((\varphi_j, 0), (\varphi_i, 0); \boldsymbol{\mu}\big), & i, j = 1, \ldots, \mathcal{N}_1, \\ \mathcal{A}\big((0, \phi_{j-\mathcal{N}_1}), (0, \phi_{i-\mathcal{N}_1}); \boldsymbol{\mu}\big), & i, j = \mathcal{N}_1 + 1, \ldots, \mathcal{N}_{12}, \\ 0 & \text{otherwise}, \end{cases}$$

$$B_{ij}^{\mathcal{N}}(\boldsymbol{\mu}) = \begin{cases} \mathcal{B}\big((\varphi_j, 0), \varphi_i; \boldsymbol{\mu}\big), & i, j = 1, \ldots, \mathcal{N}_1, \\ \mathcal{B}\big((0, \phi_{j-\mathcal{N}_1}), \varphi_i; \boldsymbol{\mu}\big), & i = 1, \ldots, \mathcal{N}_1,\ j = \mathcal{N}_1 + 1, \ldots, \mathcal{N}_{12}, \end{cases}$$

$$F_i^{\mathcal{N}}(\boldsymbol{\mu}) = \begin{cases} \mathcal{F}\big((\varphi_i, 0); \boldsymbol{\mu}\big), & i = 1, \ldots, \mathcal{N}_1, \\ \mathcal{F}\big((0, \phi_{i-\mathcal{N}_1}); \boldsymbol{\mu}\big), & i = \mathcal{N}_1 + 1, \ldots, \mathcal{N}_{12}, \end{cases}$$

$$G_i^{\mathcal{N}}(\boldsymbol{\mu}) = \mathcal{G}(\varphi_i; \boldsymbol{\mu}), \qquad i = 1, \ldots, \mathcal{N}_1.$$

Remark 3.8 (Motivation for the reduced-order approach) In order to compute a sufficiently accurate approximation $(\bar{x}^{\mathcal{N}}, \bar{p}^{\mathcal{N}}) \in Z^{\mathcal{N}}$ of the solution $(\bar{x}, \bar{p}) \in Z$ to the infinite dimensional saddle point problem (5), we generally have to choose large $\mathcal{N}_1$ and $\mathcal{N}_2$. Consequently, the solution of system (7) could require long computational times. Moreover, to compute approximations of the Pareto optimal points for various geometrical and/or model parameters we have to solve (7) for many different parameters $\boldsymbol{\mu} \in \mathcal{D}$. The following reduced-basis approach is proposed for allowing very fast and accurate solutions of the saddle point problem.

4 Reduced basis (RB) approximation

The basic idea of the RB method is to exploit the smooth parametric dependence of the saddle point problem (6) and to define small and suitable basis spaces where the problem is defined. The RB method consists in two main stages. During the first one, so called *offline* phase, we define the reduced basis functions that are solution of the problem system for a properly chosen set of parameter values. After the computationally expensive offline stage, during the *online* phase, for any new value of the parameter $\boldsymbol{\mu} \in \mathcal{D}$, the RB method provides a very effective dataset for the computation of a accurate, reliable and fast approximation of the problem solution.

4.1 The offline phase

In order to define the reduced basis spaces, we build a set of parameter samples $S_N^{\boldsymbol{\mu}} = \{\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^N\} \subset \mathcal{D}$ and correspondingly the pairs $\{(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}^i), \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}^i))\}_{i=1}^N \subset Z^{\mathcal{N}}$ which are solutions to Eq. 6 for parameters $\boldsymbol{\mu}^i \in S_N^{\boldsymbol{\mu}}, i = 1, \ldots, N$. The choice of the parameter set S_N^{μ} can be done by using both the classical greedy algorithm, reviewed in [29], and the optimization greedy, recently proposed in [34], particularly useful for problems involving a large number of parameters, i.e., a large number of cost functionals or distributed parameter functions [18, 20].

The RB approximation of Eq. 6 consists in a Galerkin projection onto low dimensional subspaces spanned by 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 RB method for the saddle point problem, we fulfill the inf-sup condition by defining the following spaces:

$$V^N = \text{span}\left\{\bar{y}^{\mathcal{N}}(\boldsymbol{\mu}^1), \ldots, \bar{y}^{\mathcal{N}}(\boldsymbol{\mu}^N)\right\} \oplus \text{span}\left\{\bar{p}^{\mathcal{N}}(\boldsymbol{\mu}^1), \ldots, \bar{p}^{\mathcal{N}}(\boldsymbol{\mu}^N)\right\},$$
$$U^N = \text{span}\left\{\bar{u}^{\mathcal{N}}(\boldsymbol{\mu}^1), \ldots, \bar{u}^{\mathcal{N}}(\boldsymbol{\mu}^N)\right\}.$$

Let $\{\psi_i\}_{i=1}^{2N}$ and $\{\zeta_i\}_{i=1}^N$ denote orthonormal bases for V^N and U^N, respectively. We set $X^N = V^N \times U^N$ and $Z^N = X^N \times V^N$. The use of the same subspace for the state and ajoint is cruical for stability of the reduced basis method. The RB approximation of problem (6) reads: find the pair $(\bar{x}^N(\boldsymbol{\mu}), \bar{p}^N(\boldsymbol{\mu})) \in Z^N$ such that

$$\begin{aligned} \mathcal{A}(\bar{x}^N(\boldsymbol{\mu}), x^N; \boldsymbol{\mu}) + \mathcal{B}(x^N, \bar{p}^N(\boldsymbol{\mu}); \boldsymbol{\mu}) &= \mathcal{F}(x^N; \boldsymbol{\mu}) \quad \text{for all } x^N \in X^N, \\ \mathcal{B}(\bar{x}^N(\boldsymbol{\mu}), p^N; \boldsymbol{\mu}) &= \mathcal{G}(p^N; \boldsymbol{\mu}) \quad \text{for all } p^N \in V^N. \end{aligned} \tag{1}$$

Proposition 4.1 *Let all hypotheses of Lemma* 3.5 *hold. Then, there exists a unique solution* $(\bar{x}^N(\boldsymbol{\mu}), \bar{p}^N(\boldsymbol{\mu}))$ *to* Eq. 1.

Proof As in Lemma 3.5-1) and Remark 3.6 we find that

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

fulfills a *Brezzi RB inf-sup condition*

$$\beta^N(\boldsymbol{\mu}) \geq \alpha > 0 \quad \text{for all} \boldsymbol{\mu} \in \mathcal{D}, \tag{2}$$

where $\alpha = \mu_{lb} \min(1/\eta, 1)/2$ has been introduced in Eq. 2. Since X^N is a subspace of $X^{\mathcal{N}}$ we infer as in Lemma 3.5-2) that

$$\mathcal{A}(\tilde{x}^N, \tilde{x}^N; \boldsymbol{\mu}) \geq \alpha \|\tilde{x}^N\|_X^2 \quad \text{for all } \tilde{x}^N \in X_0^N, \tag{3}$$

where $X_0^N = \{x^N \in X^N \mid \mathcal{B}(x^N, p^N; \boldsymbol{\mu}) = 0 \text{ for all } p^N \in V^N\}$. From Eqs. 2 and 3 it follows that Eq. 1 admits a unique solution. □

We have already mentioned that in the offline phase the selection of $S_N^{\boldsymbol{\mu}}$ and the computation of the basis functions are carried out. In the offline phase, we also compute the parameter independent parts of the coefficient matrices and vectors,

occurring in the algebraic formulation of Eq. 1. This allows us to accelerate our computational effort in the online phase. For this purpose, an affine decomposition of the linear and bilinear forms is required. 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 involved in the state (2), i.e., the forms $a(\cdot\,,\cdot\,;\boldsymbol{\mu}_c)$ and $f(\boldsymbol{\mu}_c)$ with $\boldsymbol{\mu}_c \in \mathcal{D}_c$. However, if they are not affine, it is possible to approximate them by affine linear and bilinear forms through the empirical interpolation method [3]. So that it is possible to decouple the forms for any parameter $\boldsymbol{\mu} \in \mathcal{D}$ as follows:

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

and to approximate

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

for $x = (y,u) \in X$ and $(\tilde{x},\tilde{p}) \in Z$ with $\tilde{x} = (\tilde{y},\tilde{u})$. Thanks to these assumptions, the following low dimensional matrices can be computed only once and during the offline phase:

$$A^{i_1}_{ij} = \begin{cases} \mathcal{A}^{i_1}\big((\psi_j,0),(\psi_i,0)\big), & i,j = 1,\dots,2N, \\ \mathcal{A}^{i_1}\big((0,\zeta_{j-2N}),(0,\zeta_{i-2N})\big), & i,j = 2N+1,\dots,3N, \\ 0 & \text{otherwise}, \end{cases}$$

$$B^{i_2}_{ij} = \begin{cases} \mathcal{B}^{i_2}\big((\psi_j,0),\psi_i\big), & i,j = 1,\dots,2N, \\ \mathcal{B}^{i_2}\big((0,\zeta_{j-2N}),\psi_i\big), & i = 1,\dots,2N,\ j = 2N+1,\dots,3N, \end{cases}$$

$$F^{i_3}_{i} = \begin{cases} \mathcal{F}^{i_3}\big((\psi_i,0)\big), & i = 1,\dots,2N, \\ \mathcal{F}^{i_3}\big((0,\zeta_{i-2N})\big), & i = 2N+1,\dots,3N, \end{cases}$$

$$G^{i_4}_{i} = \mathcal{G}^{i_4}(\psi_i), \qquad i = 1,\dots,2N$$

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

4.2 The online phase

In the online phase the parameter dependent part of the system can be rapidly evaluated for each new parameter value. Finally, a low dimensional linear system can be assembled and solved efficiently during the online stage for any new value of $\boldsymbol{\mu} \in \mathcal{D}$. Analogously to Eq. 7, the algebraic formulation of Eq. 1 is the following:

$$\begin{aligned} \sum_{j=1}^{3N} A^N_{ij}(\boldsymbol{\mu})\bar{\mathrm{x}}^N_j(\mu) + \sum_{j=1}^{2N} B^N_{ji}(\boldsymbol{\mu})\bar{\mathrm{p}}^N_j(\mu) = F^N_i(\boldsymbol{\mu}), \quad i = 1,\dots,3N, \\ \sum_{j=1}^{3N} B^N_{ij}(\boldsymbol{\mu})\bar{\mathrm{x}}^N_j(\mu) = G^N_i(\boldsymbol{\mu}), \quad i = 1,\dots,2N, \end{aligned} \tag{4}$$

where

$$A^N(\boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{A}}} \Theta^i_{\mathcal{A}}(\boldsymbol{\mu}) A^i, \qquad F^N(\boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{F}}} \Theta^i_{\mathcal{F}}(\boldsymbol{\mu}) F^i,$$
$$B^N(\boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{B}}} \Theta^i_{\mathcal{B}}(\boldsymbol{\mu}) B^i, \qquad G^N(\boldsymbol{\mu}) = \sum_{i=1}^{\ell_{\mathcal{G}}} \Theta^i_{\mathcal{G}}(\boldsymbol{\mu}) G^i.$$

The solution vectors $\bar{\mathrm{x}} = (\bar{\mathrm{x}}_i^N) \in \mathbb{R}^{3N}$ and $\bar{\mathrm{p}}^N = (\bar{\mathrm{p}}_i^N) \in \mathbb{R}^{2N}$ to Eq. 4 define the final RB solutions:

$$\bar{y}^N = \sum_{i=1}^{2N} \bar{\mathrm{x}}_i^N \psi_i, \quad \bar{u}^N = \sum_{i=1}^{N} \bar{\mathrm{x}}_{i+2N}^N \zeta_i \quad \text{and} \quad \bar{p}^N = \sum_{i=1}^{2N} \bar{\mathrm{p}}_i^N \psi_i.$$

4.3 A-posteriori error estimates

A rigorous error estimate 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 $S_N^{\boldsymbol{\mu}}$ in the offline phase and it provides a certified level of accuracy of the approximate solution compared with the high-fidelity solution in the online phase. Thanks to the saddle point formulation of the problem, we exploit the error estimates proposed in [30] for Stokes problem and more recently in [24] for elliptic linear-quadratic optimal control problems. Thus, we have a rigorous and inexpensive estimate for the error between the HF solution of Eq. 6 and the RB solution of Eq. 1:

$$\left(\|\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}) - \bar{x}^N(\boldsymbol{\mu})\|_X^2 + \|\bar{p}^{\mathcal{N}}(\boldsymbol{\mu}) - \bar{p}^N(\boldsymbol{\mu})\|_V^2\right)^{1/2} \leq \Delta_N(\boldsymbol{\mu}) \quad \text{for any } \boldsymbol{\mu} \in \mathcal{D}.$$

Moreover, we have a rigorous and inexpensive estimate for the error on the cost functional evaluated by using the HF solution and the RB solution:

$$\left|\hat{\mathfrak{J}}(\bar{u}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) - \hat{\mathfrak{J}}(\bar{u}^N(\boldsymbol{\mu}); \boldsymbol{\mu})\right| \leq \Delta_N^J(\boldsymbol{\mu}) \quad \text{for any } \boldsymbol{\mu} \in \mathcal{D}.$$

In Eq. 7 we will quantify the estimators $\Delta_N(\boldsymbol{\mu})$ and $\Delta_N^J(\boldsymbol{\mu})$. Note that the offline-online computational decomposition can be adopted also for the computation of the error estimates in order to be able to efficiently compute it online together with the RB solution of the problem.

During the offline stage, performed once, the parameter independent parts of the error estimates can be computed, while during the inexpensive online evaluation, performed for any desired $\boldsymbol{\mu}$, the parameter dependent parts can be rapidly evaluated. The fast evaluation of the error estimates permits to predict the RB error with respect to the HF solution without computing the latter and it is crucial during the greedy algorithm to speed up the efficient selection of the snapshots. For more details we refer to [26, Chapter 3.3] and the recent works [14, 28].

In order to formulate the a-posteriori error estimates, we write (7) as a single equation; compare with [24, Section 4]. For any parameter $\boldsymbol{\mu} \in \mathcal{D}$ let us define the linear parameter-dependent bilinear form

$$\mathcal{K}(z, \tilde{z}; \boldsymbol{\mu}) = \mathcal{A}(x, \tilde{x}; \boldsymbol{\mu}) + \mathcal{B}(\tilde{x}, p; \boldsymbol{\mu}) + \mathcal{B}(x, \tilde{p}; \boldsymbol{\mu}) \quad \forall z = (x, p), \tilde{z} = (\tilde{x}, \tilde{p}) \in Z \tag{5}$$

and the parameter-dependent linear functional

$$\mathcal{R}(z;\boldsymbol{\mu}) = \mathcal{F}(x;\boldsymbol{\mu}) + \mathcal{G}(p;\boldsymbol{\mu}) \quad \forall z = (x,\, p) \in Z.$$

Then, Eq. 5 is equivalent with

$$\mathcal{K}(\bar{z}^{\mathcal{N}}(\boldsymbol{\mu}), z^{\mathcal{N}};\boldsymbol{\mu}) = \mathcal{R}(z^{\mathcal{N}};\boldsymbol{\mu}) \quad \forall z^{\mathcal{N}} = (x^{\mathcal{N}},\, p^{\mathcal{N}}) \in Z^{\mathcal{N}}, \tag{6}$$

where $\bar{z}^{\mathcal{N}}(\boldsymbol{\mu}) = (\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}),\, \bar{p}^{\mathcal{N}}(\boldsymbol{\mu})) \in Z^{\mathcal{N}}$.

Proposition 4.2 *Let all hypotheses of Lemma* 3.5 *hold. Then,* Eq. 6 *has a unique solution* $\bar{z}^{\mathcal{N}}(\boldsymbol{\mu})$ *for any parameter* $\boldsymbol{\mu} \in \mathcal{D}$.

Proof We introduce the *Babuška HF inf-sup constant* $\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu})$ [1] associated with Eq. 6 by

$$\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{\tilde{z}^{\mathcal{N}} \in Z^{\mathcal{N}}} \sup_{z^{\mathcal{N}} \in Z^{\mathcal{N}}} \frac{\mathcal{K}(z^{\mathcal{N}}, \tilde{z}^{\mathcal{N}};\boldsymbol{\mu})}{\|z^{\mathcal{N}}\|_Z \|\tilde{z}^{\mathcal{N}}\|_Z}.$$

Since $\mathcal{A}(\cdot\,,\cdot\,;\boldsymbol{\mu})$ and $\mathcal{B}(\cdot\,,\cdot\,;\boldsymbol{\mu})$ satisfy the hypotheses of the Brezzi theorem, the compound form $\mathcal{K}(\cdot\,,\cdot\,;\boldsymbol{\mu})$ is bounded and $\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) > 0$ holds for all $\boldsymbol{\mu} \in \mathcal{D}$. Therefore, Eq. 6 has a uique solution. □

We assume that we can bound the Babuška HF inf-sup constant from below by a $\boldsymbol{\mu}$-dependent positive constant $\hat{\beta}^{\mathcal{N}}_{lb}(\boldsymbol{\mu})$. Moreover, let $\hat{\beta}_0$ be a $\boldsymbol{\mu}$- and $\mathcal{N}$-independent lower bound for $\hat{\beta}_{lb}(\boldsymbol{\mu})$ (for stability reasons) so that we have

$$\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) \geq \hat{\beta}^{\mathcal{N}}_{lb}(\boldsymbol{\mu}) \geq \hat{\beta}_0 > 0 \quad \text{for all } \boldsymbol{\mu} \in \mathcal{D}.$$

Remark 4.3 (Estimation of $\hat{\beta}_{lb}(\boldsymbol{\mu})$) An effective computation of a lower bound $\hat{\beta}^{\mathcal{N}}_{lb}(\boldsymbol{\mu})$ for the constant $\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu})$ 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 [17]. However, since this approximation of the lower bound can be very time consuming, we adopt an alternative strategy recently proposed and compared with the previous one in [24]. It consists in defining a surrogate $\beta^{\mathcal{N}}_s(\boldsymbol{\mu})$ obtained by computing the expensive $\hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu})$ for a small set of parameter values equally distributed in $\mathcal{D}$ and by using these computations to define, by interpolation, the surrogate $\hat{\beta}^{\mathcal{N}}_s(\boldsymbol{\mu})$ for all $\boldsymbol{\mu} \in \mathcal{D}$. 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 estimate.

Suppose that we have determined an RB solution $(\bar{x}^N(\boldsymbol{\mu}), \bar{p}^N(\boldsymbol{\mu}))$ to Eq. 1. Let us define the residuals $r^N_{\mathsf{du}}(\cdot\,;\boldsymbol{\mu}) \in (X^{\mathcal{N}})'$ and $r^N_{\mathsf{pr}}(\cdot\,;\boldsymbol{\mu}) \in (V^{\mathcal{N}})'$ by

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

Then, we obtain the following a-posteriori error estimates (see, e.g., [19])

$$\begin{aligned}\Delta_N(\boldsymbol{\mu}) &= \frac{1}{\hat{\beta}_{lb}^{\mathcal{N}}(\boldsymbol{\mu})}\left(\|r_{\mathsf{du}}^N(\cdot\,;\boldsymbol{\mu})\|_{(X^{\mathcal{N}})'}^2 + \|r_{\mathsf{pr}}^N(\cdot\,;\boldsymbol{\mu})\|_{(V^{\mathcal{N}})'}^2\right)^{1/2},\\ \Delta_N^J(\boldsymbol{\mu}) &= \frac{1}{2\hat{\beta}_{lb}^{\mathcal{N}}(\boldsymbol{\mu})}\left(\|r_{\mathsf{du}}^N(\cdot\,;\boldsymbol{\mu})\|_{(X^{\mathcal{N}})'}^2 + \|r_{\mathsf{pr}}^N(\cdot\,;\boldsymbol{\mu})\|_{(V^{\mathcal{N}})'}^2\right)\end{aligned} \tag{7}$$

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

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

Remark 4.4 (Evaluation of dual norms) The computation of the dual norms of the residuals is based on the Riesz representation of the residuals and on the affine decomposition of the parametric operators. See, e.g., [29] for the offline-online efficient procedure adopted for the computational decomposition of the residuals norms.

5 Computation of Pareto optimal points by sensitivity analysis

We recall that the parameter $\boldsymbol{\mu} = (\boldsymbol{\mu}_o, \boldsymbol{\mu}_c) \in \mathcal{D}$ is given by two components: $\boldsymbol{\mu}_o$, which can be chosen in the weighted sum defining the objective, and $\boldsymbol{\mu}_c$ that stands for physical and/or geometrical parameters involved in the state equation. In this section we present a criterium that can be used for the weighted sum method, in order to reduce significantly the number of computations required for identifying a relevant set of Pareto solutions that is sufficient to interpolate the complete set. To reduce the number of 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 $\boldsymbol{\mu}_o$. For that purpose we utilize the notation $\hat{\mathfrak{J}}_{\boldsymbol{\mu}_o}$ for the partial derivative of the cost $\hat{\mathfrak{J}}$ with respest to the parameter $\boldsymbol{\mu}_o$.

Suppose that we have computed the RB solution $\bar{z}^N = \bar{z}^N(\boldsymbol{\mu}^0)$ to Eq. 1 for an initial parameter $\boldsymbol{\mu}^0 = (\boldsymbol{\mu}_o^0, \boldsymbol{\mu}_c^0)$. By $\bar{y}^N = \bar{y}^N(\boldsymbol{\mu}^0)$ we denote the associated optimal state and by $\bar{p}^N = \bar{p}^N(\boldsymbol{\mu}^0)$ the associated Lagrange multiplier. We are interested in choosing only the optimization parameters $\boldsymbol{\mu}_o \in \mathcal{D}_o$ leading to significant changes in the cost functional that can provide a relevant optimal solution of the problem. For that reason we introduce the Taylor expansion of the reduced objective with respect to changes in $\boldsymbol{\mu}_o$:

$$\hat{\mathfrak{J}}(u;\boldsymbol{\mu}^+) = \hat{\mathfrak{J}}(\bar{u}^N;\boldsymbol{\mu}^0) + \hat{\mathfrak{J}}_{\boldsymbol{\mu}_o}(\bar{u}^N;\boldsymbol{\mu}^0)(\boldsymbol{\mu}_o^+ - \boldsymbol{\mu}_o^0) + \mathcal{O}\big(|\boldsymbol{\mu}_o^+ - \boldsymbol{\mu}_o^0|_2\big), \tag{1}$$

where $\boldsymbol{\mu}^+ = (\boldsymbol{\mu}_o^+, \boldsymbol{\mu}_c^0)$, i.e., $\boldsymbol{\mu}^+$ and $\boldsymbol{\mu}^0$ only differ in the first k components. Hence, we have to compute $\hat{\mathfrak{J}}(\bar{u}^N; \boldsymbol{\mu}^0)$ as well as the partial derivative of $\hat{\mathfrak{J}}_{\boldsymbol{\mu}_o}$ in order to get the requested information. Utilizing $\mu_{o,k} = 1 - \sum_{i=1}^{k-1} \mu_{o,i}$ we have

$$\begin{aligned}\hat{\mathfrak{J}}(\bar{u};^N \boldsymbol{\mu}^0) &= \frac{1}{2}\left(\sum_{i=1}^{k-1} \mu_{o,i}^0 \left\|\mathcal{C}_i \bar{y}^N - w_i\right\|_{W_i}^2 + \mu_{o,k}^0 \gamma \left\|\bar{u}^N\right\|_U^2\right) \\ &= \frac{1}{2}\left(\sum_{i=1}^{k-1} \mu_{o,i}^0 \left\|\mathcal{C}_i \bar{y}^N - w_i\right\|_{W_i}^2 + \left(1 - \sum_{i=1}^{k-1} \mu_{o,i}^0\right)\gamma \left\|\bar{u}\right\|_U^2\right).\end{aligned}$$

Now we compute the derivatives of the cost functional with respect to $\mu_{o,j}$ for $j = 1, \ldots, k-1$:

$$\begin{aligned}\hat{\mathfrak{J}}_{\mu_{o,j}}(\bar{u}^N; \boldsymbol{\mu}^0) = \frac{1}{2}\left\|\mathcal{C}_j \bar{y}^N - w_j\right\|_{W_j}^2 + \sum_{i=1}^{k-1} \mu_{o,i}^0 \left\langle \mathcal{C}_i \bar{y}^N - w_i, \mathcal{C}_i \bar{y}_{\mu_{o,j}}^N \right\rangle_{W_i} \\ - \frac{\gamma}{2}\left\|\bar{u}^N\right\|_U^2 + \mu_{o,k}^0 \gamma \left\langle \bar{u}^N, \bar{u}_{\mu_{o,j}}^N \right\rangle_U,\end{aligned}$$

where the $k-1$ sensitivities $\bar{z}_{\mu_{o,j}}^N = (\bar{x}_{\mu_{o,j}}^N, \bar{p}_{\mu_{o,j}}^N) \in Z^N$ with $\bar{x}_{\mu_{o,j}}^N = (\bar{y}_{\mu_{o,j}}^N, \bar{u}_{\mu_{o,j}}^N)$ are computed as follows, see, e.g., [13]. Utilizing the parametrized bilinear form $\mathcal{K}(\cdot;\cdot;\boldsymbol{\mu})$ introduced in Eq. 5, the first-order optimality conditions (1) can be expressed as

$$\mathcal{K}(\bar{z}^N; z^N; \bar{\boldsymbol{\mu}}) = \mathcal{R}(z^N; \bar{\boldsymbol{\mu}}) \quad \text{for all } z^N \in Z^N. \tag{2}$$

We differentiate (2) with respect to the optimization parameter $\boldsymbol{\mu}_{o,j}$ for $j = 1, \ldots, k-1$:

$$\mathcal{K}(\bar{z}_{\mu_{o,j}}^N, z^N; \boldsymbol{\mu}^0) = \mathcal{R}_{\mu_{o,j}}(z^N; \boldsymbol{\mu}^0) - \mathcal{K}_{\mu_{o,j}}(\bar{z}^N, z^N; \boldsymbol{\mu}^0) \quad \forall z^N \in Z^N, \tag{3}$$

where

$$\begin{aligned}\mathcal{K}_{\mu_{o,j}}(\bar{z}^N, z^N; \boldsymbol{\mu}^0) &= \left\langle \mathcal{C}_j \bar{y}^N, \mathcal{C}_j y^N \right\rangle_{W_j} - \gamma \left\langle \bar{u}^N, u^N \right\rangle_U, \\ \mathcal{R}_{\mu_{o,j}}(z^N; \boldsymbol{\mu}^0) &= \left\langle w_j, \mathcal{C}_j y^N \right\rangle_{W_j}\end{aligned}$$

for $j = 1, \ldots, k-1$ and for $z^N = (x^N, p^N) \in Z^N$ with $x^N = (y^N, u^N)$. Now, we can rapidly compute the sensitivities $\bar{z}_{\mu_{o,j}}^N$, $1 \le j \le k-1$, from the linear system (3), where the coefficient matrix has been already defined for the computation of $\bar{z}^N$. The advantageous feature of the explained sensitivity theory is its efficient online-offline computational decoupling coming from the RB precomputed structures. Note that the partial derivative of the bilinear and linear forms are readily computable thanks to the affine parameter decomposition (assuming the parameter-dependent functions are easily differentiable). Thus, at a very small computational effort, we are able to define a suitable parameter set that is useful for computing specific optimal solutions relevant for identifing the entire Pareto optimal set. Starting by considering the parametric saddle point formulation of the problem where the parameter vector is defined by the parameters used in the weighted sum of the cost functionals and

the ones involved in the state equation, in the following listed steps we describe the entire strategy we adopt for defining the suitable Pareto points approximation.

1) The offline phase for the RB approximation is carried out as described in Section 4.1;
2) The online phase is used to compute a Pareto optimal point corresponding to an initial optimization parameter $\boldsymbol{\mu}^0 \in \mathcal{D}$ as explained in Section 4.2;
3) Thanks to the sensitivity analysis (see above) we rapidly compute a prediction of the cost functional value corresponding to a suitably large set of parameter values. Among this set, we select the ones that span the entire cost functional value range. This apprach provides a suitable discrete parameter set $\Xi_s \subset \mathcal{D}_o$, that leads to different and not close variations of the (reduced) cost functional;
4) The set of Pareto optimal solutions is computed by using the online step of the RB method corresponding to the parameter set Ξ_s determined in step 3).

Remark 5.1 Step 3) allows us to drastically reduce the number of online RB computations needed to recover a suitable distribution of the Pareto optimal points. Let us summarize the main computations required by the sensitivity analysis: (i) the computation of the cost value $\hat{\mathfrak{J}}(\bar{u}; {}^N\boldsymbol{\mu}^0)$ by utilizing the RB solution $\bar{z}^N$ and (ii) the solution of the reduced-order system (3) to determine the sensitivities $\bar{z}^N_{\boldsymbol{\mu}_o}$.

6 Numerical examples

In this section we present numerical examples illustrating the efficiency of our proposed strategy. Different control input spaces and different geometric parameters are considered. We start with a non parametric PDE constraint focusing on the optimal control solutions defined by varying the parameters representing the weights involved in the cost functional. In the second example, we introduce a geometrical parameter leading to a parameter in the PDE constraint. Moreover, as last numerical test, we reduce the control space dimension in order to be able to show the feasible set of cost functional values (by varying arbitrarily 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. For the HF Galerkin approximation we utilized a finite element (FE) scheme with piecewise $\mathbb{P}_1$ elements.

Run 1 ($\boldsymbol{\mu}_o \in \mathbb{R}^3$ and $\boldsymbol{\mu}_c \equiv 0$) In our first test we choose $k = 3$ optimization parameters, but no parameters in the state equation, i.e., $\boldsymbol{\mu}_o \in \mathbb{R}^3$ and $\boldsymbol{\mu}_c \equiv 0$. We consider the domain $\Omega \in \mathbb{R}^2$ given by a rectangle separated into two disjunct subdomains $\Omega_1 = (0, 1) \times (0, 1)$ and $\Omega_2 = (1, 4) \times (0, 1)$ and represented in the left plot of Fig. 1.

Let $U = L^2(\Omega)$ be the space of admissible controls. We introduce the multiobjective optimization problem in which the vector-valued cost functional is defined as follows:

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

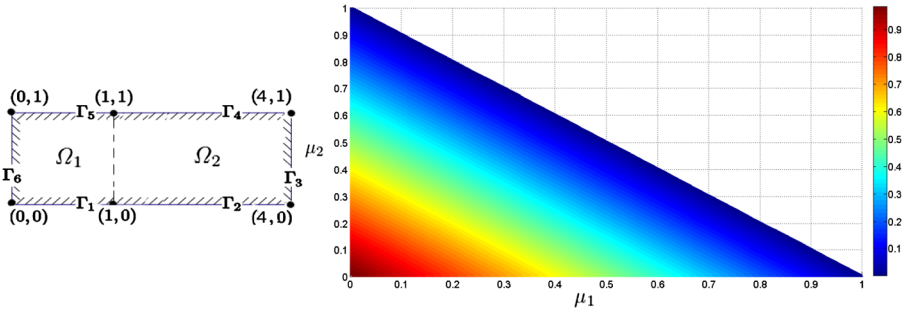

Fig. 1 Run 1: Domain Ω separated into the two subdomains Ω_1, Ω_2 (*left*) and lower estimate for the Babuska (FE) inf-sup constant $\beta^{\mathcal{N}}(\boldsymbol{\mu})$ (*right*)

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

$$-\Delta y = u \text{ in } \Omega, \quad y = 1 \text{ on } \Gamma = \partial\Omega. \tag{1}$$

In order to apply the weighed sum method for the computation of the Pareto optimal points, we consider the weighted sum of the cost vector:

$$J(x(\boldsymbol{\mu}_o); \boldsymbol{\mu}_o) = \mu_{o,1} J_1(y(\boldsymbol{\mu}_o)) + \mu_{o,2} J_2(y(\boldsymbol{\mu}_o)) + (1 - \mu_{o,1} - \mu_{o,2}) J_3(u(\boldsymbol{\mu}_o))$$

for $x(\boldsymbol{\mu}_o) = (y(\boldsymbol{\mu}_o), u(\boldsymbol{\mu}_o)) \in X$ and $\boldsymbol{\mu}_o = (\mu_{o,1}, \mu_{o,2}, \mu_{o,3}) \in \mathcal{D}_o$. Thus, for any $\boldsymbol{\mu} \in \mathcal{D}_o$ the parametrized optimal control problem reads

$$\min J(x(\boldsymbol{\mu}_o); \boldsymbol{\mu}_o) \quad \text{s.t.} \quad x(\boldsymbol{\mu}_o) = (y(\boldsymbol{\mu}_o), u(\boldsymbol{\mu}_o)) \in X \text{ solves (1)}.$$

The numerical approximation of the RB functions (state, control and adjoint variables) is based on the FE discretization. The dimension of the FE space $V^{\mathcal{N}}$ is $\mathcal{N}_1 = 11441$ obtained by using a mesh of 22528 elements. For the control space we choose $U^{\mathcal{N}} = V^{\mathcal{N}}$, so that we get $\mathcal{N}_2 = \mathcal{N}_1$ and $\mathcal{N}_{12} = 22882$. Since the condition $\mu_{o,3} = 1 - \mu_{o,1} - \mu_{o,2}$ holds, we consider only the two parameters in our numerical implementation. The two parameters are defined in the set $\mu_{o,1} \in [0, 1]$ and $\mu_{o,2} \in [0, 1 - \mu_{o,1}]$. In order to illustrate the dependence of the optimal solution on the choice of the weighting parameters we consider three different parameter values leading to different cost functionals and therefore to different optimal solutions. In Fig. 2 we study a dominating first cost functional ($\boldsymbol{\mu}_o^1 = (0.9, 0)$), in Fig. 3 a dominating second cost functional ($\boldsymbol{\mu}_o^2 = (0.11, 0.83)$) and in Fig. 4 we consider an equally distributed dominance between the three cost functionals ($\boldsymbol{\mu}_o^3 = (0.3, 0.3)$). Due to the smooth parameter dependence of the Babuska FE inf-sup constant, we use a linear interpolant surrogate of $\boldsymbol{\mu}_o \mapsto \beta^{\mathcal{N}}(\boldsymbol{\mu}_o)$ in our a-posteriori error computations. In the right plot of Fig. 1 we show the surrogate $\beta_s^{\mathcal{N}}(\boldsymbol{\mu}_o)$ obtained by using values of the parameter $\boldsymbol{\mu}_o = (\mu_{o,1}, \mu_{o,2})$ equally distributed in $\mathcal{D}_o$. The errors and the a-posteriori error estimates, computed as described in Section 4.3, are presented

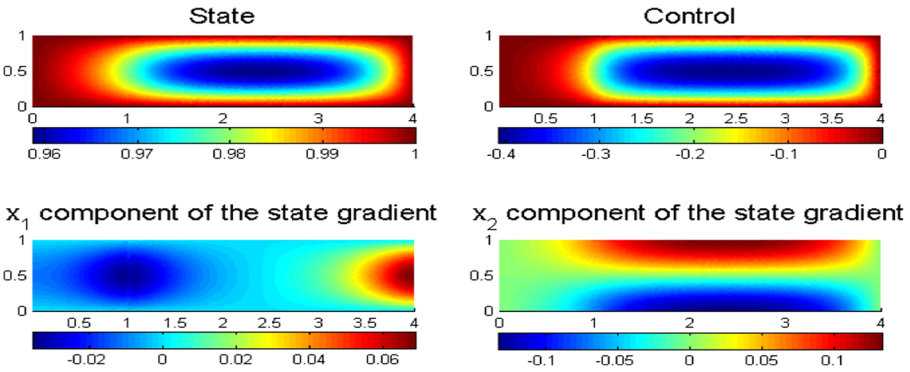

Fig. 2 Run 1: Optimal FE state $\bar{y}^{\mathcal{N}}$ (*top left*), optimal FE control $\bar{u}^{\mathcal{N}}$ (*top right*), FE partial derivative $\bar{y}^{\mathcal{N}}_{x_1}$ (*bottom left*) and FE partial derivative $\bar{y}^{\mathcal{N}}_{x_2}$ (*bottom right*) for $\boldsymbol{\mu}_o^1 = (0.9, 0)$

in Fig. 5. In the left plot 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 solution, i.e.,

$$\left(\|\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}_o) - \bar{x}^{N}(\boldsymbol{\mu}_o)\|_X^2 + \|\bar{p}^{\mathcal{N}}(\boldsymbol{\mu}_o) - \bar{p}^{N}(\boldsymbol{\mu}_o)\|_V^2 \right)^{1/2} \quad \text{for } \boldsymbol{\mu}_o \in \Xi_{test} \subset \mathcal{D}_o,$$

and we compare them with the a-posteriori error estimate $\Delta_N(\boldsymbol{\mu}_o)$. In the right plot of Fig. 5 we do the same comparison for the error estimate $\Delta_N^J(\boldsymbol{\mu}_o)$ as well as the associated difference between the FE and the RB cost, i.e.,

$$\left| J(\bar{x}^{\mathcal{N}}(\boldsymbol{\mu}_o); \boldsymbol{\mu}_o) - J(\bar{x}^{N}(\boldsymbol{\mu}_o); \boldsymbol{\mu}_o) \right| \quad \text{for } \boldsymbol{\mu}_o \in \Xi_{test} \subset \mathcal{D}_o.$$

Now let us comment on the computational effort. The offline phase lasts about 21 minutes. The online evaluation time by using $N = 15$ basis functions (in total: $2N$ for

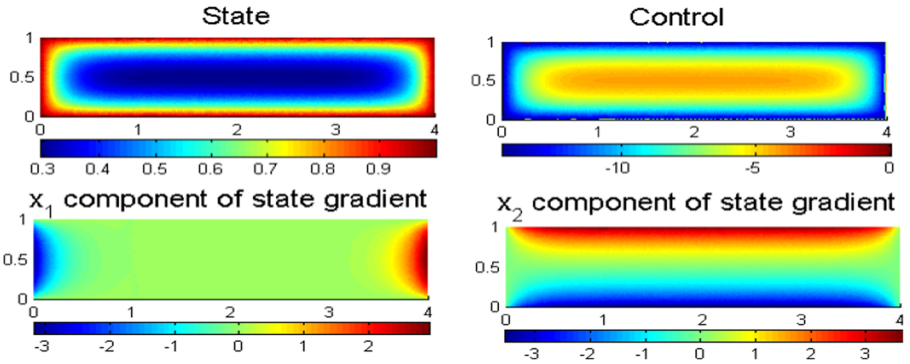

Fig. 3 Run 1: Optimal FE state $\bar{y}^{\mathcal{N}}$ (*top left*), optimal FE control $\bar{u}^{\mathcal{N}}$ (*top right*), FE partial derivative $\bar{y}^{\mathcal{N}}_{x_1}$ (*bottom left*) and FE partial derivative $\bar{y}^{\mathcal{N}}_{x_2}$ (*bottom right*) for $\boldsymbol{\mu}_o^2 = (0.11, 0.83)$

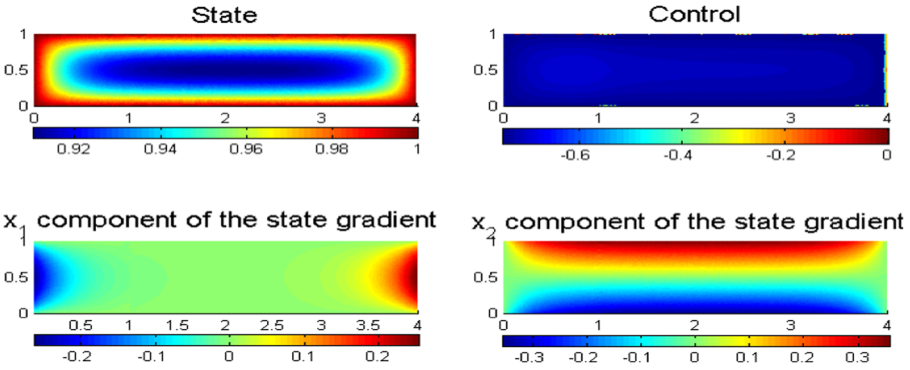

Fig. 4 Run 1: Optimal state $\bar{y}^{\mathcal{N}}$ (*top left*), optimal control $\bar{u}^{\mathcal{N}}$ (*top right*), partial derivative $\bar{y}^{\mathcal{N}}_{x_1}$ (*bottom left*) and partial derivative $\bar{y}^{\mathcal{N}}_{x_2}$ (*bottom right*) for $\boldsymbol{\mu}^3_o = (0.3, 0.3)$

the state, N for the control and $2N$ for the adjoint) and including the evaluation of the a-posteriori error estimate is 0.016 seconds; while the evaluation of the FE solution requires about 1.26 seconds, by obtaining a speed-up equal to 88. We show in Fig. 6 the RB computational time and the speedup with respect to a FE computational time by varying the number of basis functions. 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, for instance, to find 88 Pareto points at the time of only one possible FE one.

Remark 6.1 (Justification of the RB approach) Let us note that the RB online CPU time is independent on the discretization mesh used, the size of the linear system depends only on the number of basis functions used in the scheme, so in general for finer FE scheme or larger computational domain, we could achieve even a larger speed-up. The offline RB computations requires long times (specially compared with

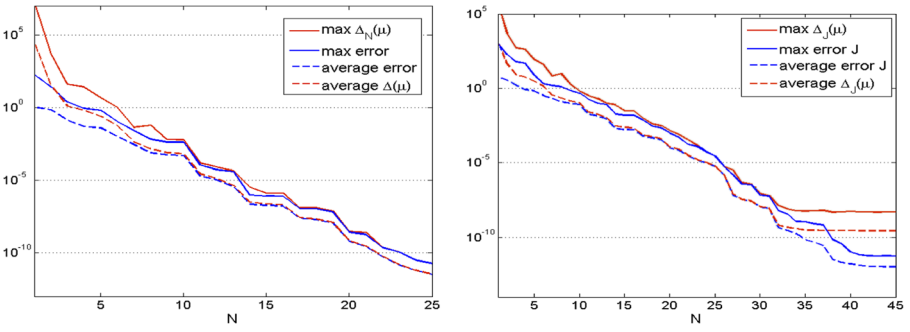

Fig. 5 Run 1: Average errors, maximum errors and error estimates regarding the solution of the problem (*left*) and the cost functional (*right*) between the FE and RB approximations

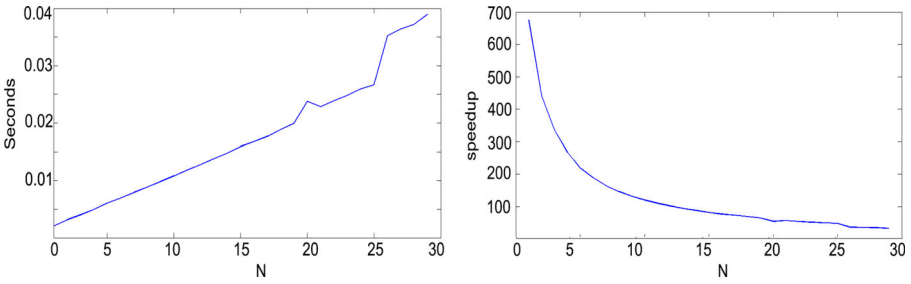

Fig. 6 Run 1: Online evaluation time (in seconds) by varying the number N of basis functions and including the evaluation of the a-posteriori error estimate (*left*) and speed-up between the FE computational time required for a single parameter value and the RB computational time required for a single parameter value by varying the number N of basis functions and including the evaluation of the a-posteriori error estimate (*right*)

the a single FE solution), nevertheless this step is performed only once and allows to provide fast numerical solutions for every parameters values. In a more general framework, the number of parameters can be much larger and the number of FE basis functions to cover all the parameters domain, as well as the offline computations required, could be much more expensive than the RB offline time. In the following test runs we consider a simpler PDE in order to show that the computational gain can be even more effective if we exploit also the proposed sensitivity analysis, which allows to further reduce the number of computations.

Run 2 (Control problem with geometrical parameter) Let us extend Run 1 by introducing a geometrical parameter $\mu_3 \in [1, 3.5]$ that defines the length of the spatial domain Ω_{μ_3} still given by a rectangle separated in two subdomains Ω_1 and Ω_2 and

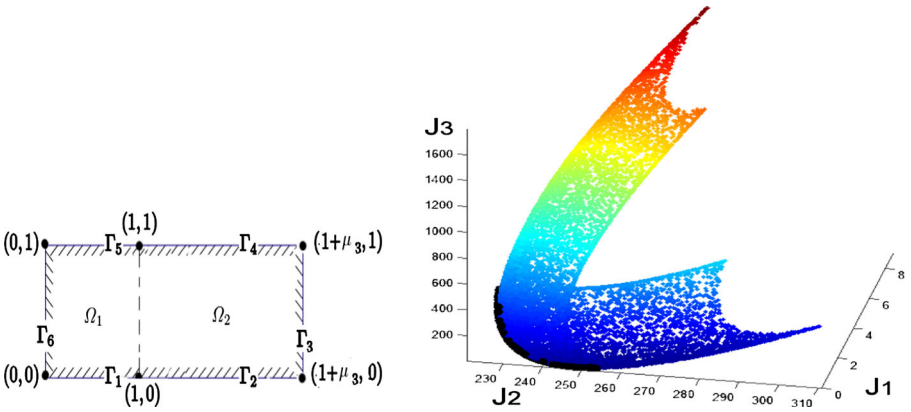

Fig. 7 Run 2: Domain representation of Ω_{μ_3} (*left*) and the set of the possible values of the cost functionals $\hat{J}_1(u;\mu_3)$, $\hat{J}_2(u;\mu_3)$, $\hat{J}_3(u;\mu_3)$ by varying the control function $u = (u_1, u_2)$ and the subset of the efficient points (*right*)

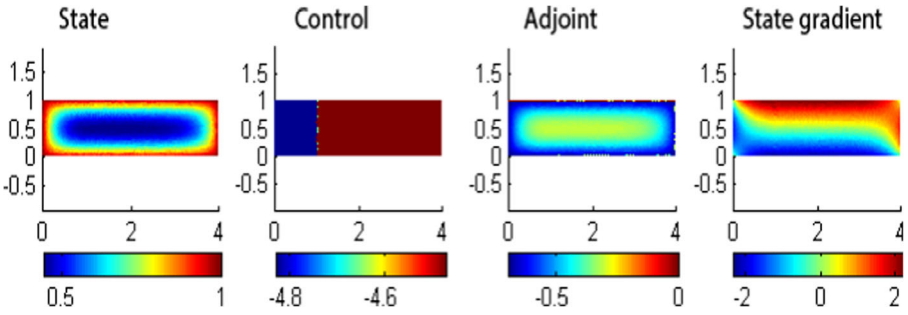

Fig. 8 Run 2: Optimal FE state $\bar{y}^{\mathcal{N}}$, optimal FE control $\bar{u}^{\mathcal{N}}$, associated FE adjoint $\bar{p}^{\mathcal{N}}$ and the sum $\bar{y}^{\mathcal{N}}_{x_1} + \bar{y}^{\mathcal{N}}_{x_2}$ for $\boldsymbol{\mu}^1 = (0.2, 0.7, 3)$

represented in the left plot of Fig. 7. We consider the two-dimensional control space $U = \mathbb{R}^2$. Let us recall that the cost functional vector is defined as follows:

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

where $w_1 = 1$ in Ω_1 and $w_1 = 0.6$ in Ω_2. Note that the parameter $\boldsymbol{\mu}$ belongs to the subset $\mathcal{D} \subset \mathbb{R}^3$ with $\boldsymbol{\mu}_o = (\mu_1, \mu_2)$ and $\boldsymbol{\mu}_c = \mu_3$. The state function $y \in V = H^1_0(\Omega_{\mu_3})$ solves the following Laplace problem:

$$-\Delta y = u_1 b_1 + u_2 b_2 \text{ in } \Omega_{\mu_3}, \quad y = 1 \text{ on } \Gamma_D = \partial\Omega_{\mu_3}, \tag{2}$$

where $u = (u_1, u_2) \in \mathbb{R}^2$ is the control function and $b_1, b_2 \in L^\infty(\Omega)$ are the characteristic functions of Ω_1, Ω_2 respectively. As in Run 1, we apply the weighted sum method to generate the Pareto optimal set. We introduce

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

and the parametrized optimal control problem:

$$\min J(x(\boldsymbol{\mu}); \boldsymbol{\mu}) \text{ s.t.} x(\boldsymbol{\mu}) = (y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in X \text{ solves}(2). \tag{3}$$

In Figs. 8 and 9 we show the RB solutions to Eq. 3 for $\boldsymbol{\mu}^1 = (0.2, 0.7, 3)$ and $\boldsymbol{\mu}^2 = (0.2, 0.7, 1)$ respectively, the plots include the optimal FE state function, the optimal

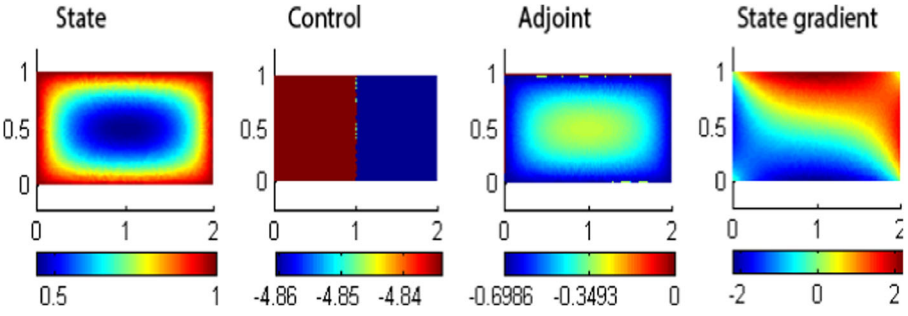

Fig. 9 Run 2: Optimal FE state $\bar{y}^{\mathcal{N}}$, optimal FE control $\bar{u}^{\mathcal{N}}$, associated FE adjoint $\bar{p}^{\mathcal{N}}$ and the sum $\bar{y}^{\mathcal{N}}_{x_1} + \bar{y}^{\mathcal{N}}_{x_2}$ for $\boldsymbol{\mu}^2 = (0.2, 0.7, 1)$

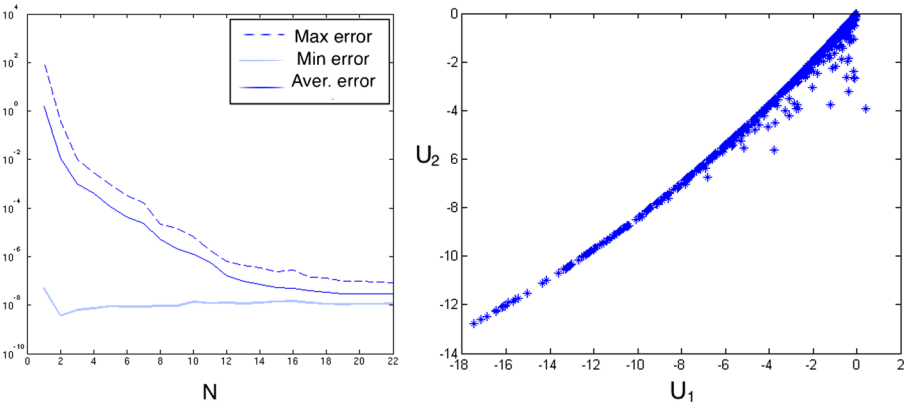

Fig. 10 Run 2: Error (maximum, minimum and average) between the optimal RB and FE solution over a set of 500 random samples by varying the number of basis functions (*left*) and the optimal controls $\bar{u}^{\mathcal{N}} = (u_1, u_2)$ by varying the parameter values (*right*)

FE control function, the FE adjoint function and the sum of the two components of the gradient of the state function. In the left plot of Fig. 10 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.

Now we set $\boldsymbol{\mu}_c = \mu_3 = 3$ and we focus our attention on the variation of the multiobjective parameters of the problem. In the right plot of Fig. 10 we show the values of the coefficients u_1, u_2 corresponding to the optimal control of the problem (3) by varying the parameters values μ_1 and μ_2.

We consider a subset of the possible control functions such that $-30 \leq u_i \leq 10$. In the left plot of Fig. 11 the set

$$S = \left\{ (\hat{J}_1(u; \mu_3), \hat{J}_2(u; \mu_3)) \,\middle|\, u = (u_1, u_2) \text{ with } -30 \leq u_i \leq 10, i = 1, 2 \right\}$$

is presented. Then, we solve with the RB method the multiobjective optimal control problem by choosing randomly a large set of optimization parameters $\boldsymbol{\mu}_o = (\mu_{o,1}, \mu_{o,2})$. In this way we obtain the Pareto optimal points and the corresponding

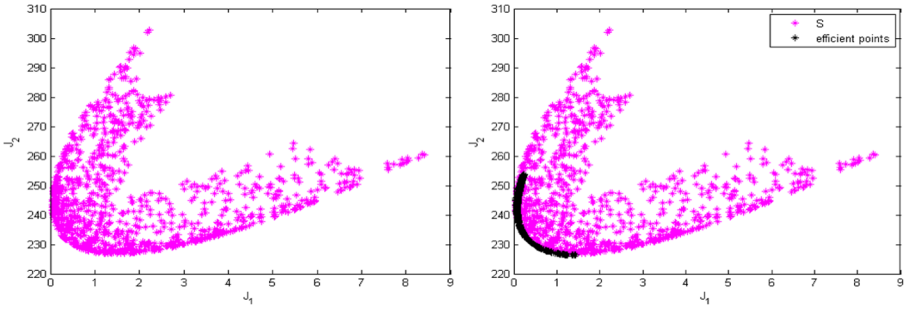

Fig. 11 Run 2: Set S of the possible values of the components $\hat{J}_1(\bar{u}; \mu_3)$ and $\hat{J}_2(\bar{u}; \mu_3)$ by varying the function u and the subset of the efficient Pareto points

efficient points of Eq. 3, i.e., the cost components J_1 and J_2 evaluated at the optimal solutions. In the right plot of Fig. 11 we plot the set S and the set of efficient points. We note that for $J_2 \geq 245$, the algorithm appears not working correctly, because for some Pareto points, both J_1 and J_2 have not minimum values. For that reason, we show in the right plot of Fig. 7 a similar plot, this time including the third cost functional and the corresponding efficient points. We observe that all the Pareto optimal points defined by the proposed RB strategy are correctly computed, since all the three cost functionals are considered in the multiobjective problem.

Run 3 (Sensitivity analysis) In this test we apply the sensitivity analysis introduced in Section 5 to the multiobjective control problem of Run 2. Our purpose is to show how the sensitivity analysis improves significantly the efficiency of our RB approach. Thanks to the inexpensive prediction of the cost functional value by varying the optimization parameter $\boldsymbol{\mu}_o$, we are able to span the whole set of Pareto optimal points by computing the optimal solution to a very small set of parameter values.

In Fig. 12 the interpolation of the Pareto optimal solutions obtained by varying randomly the value of the parameter $\boldsymbol{\mu}_o$ is presented. We note that even with 100 parameter values we are not able to cover the range of efficient points (see right plot of Fig. 11 for a comparison). In Fig. 13 we show the interpolation of the Pareto optimal solutions in correspondence of a smaller set of parameter values computed

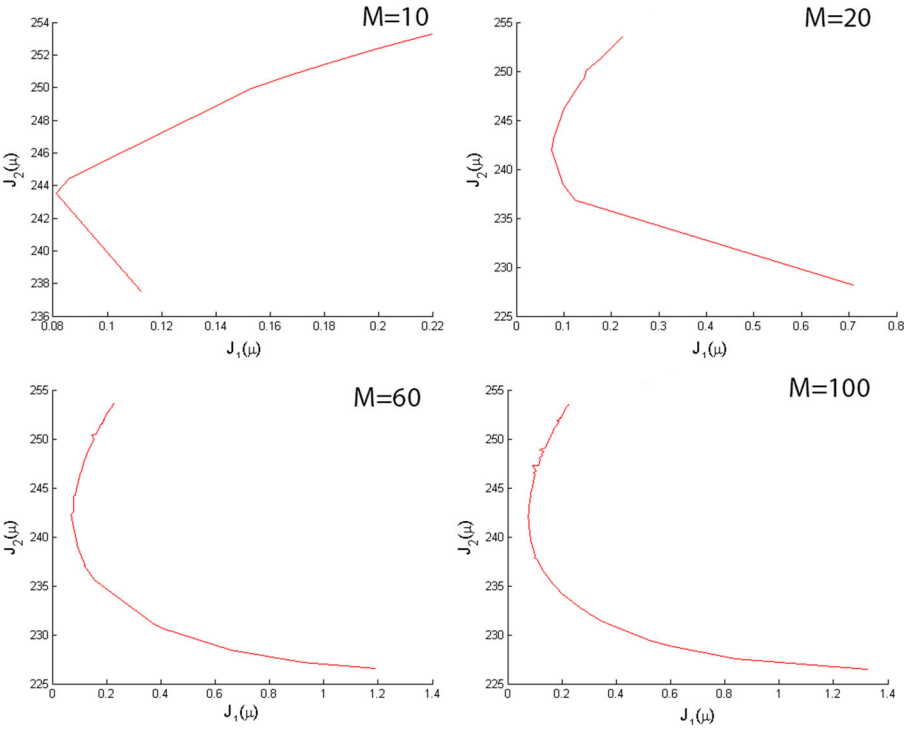

Fig. 12 Run 3: Piecewise linear interpolation of M Pareto optimal points computed randomly

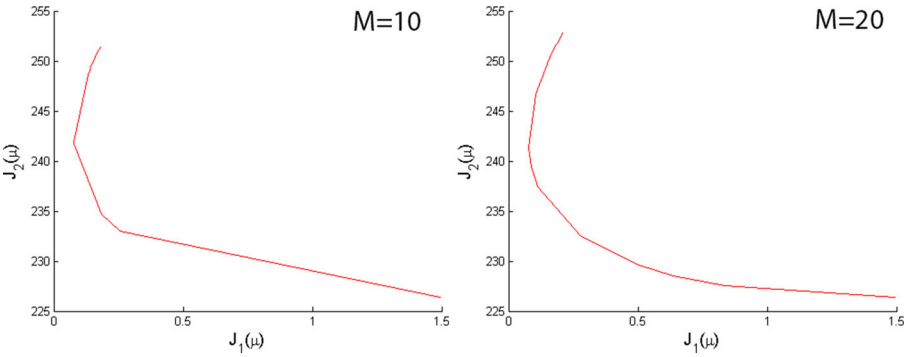

Fig. 13 Run 3: Piecewise linear interpolation of M Pareto optimal points with the help of the sensitivity analysis

by the sensitivity analysis. Note that by using only 20 values of the parameter $\boldsymbol{\mu}_o$, efficiently selected with the sensitivity analysis, and consequently 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 distribution. Regarding the computational times, for this test case, the FE method requires about 1.1 seconds to compute a single Pareto solution. The complete online computation of 20 Pareto solutions, together with the error estimates and with the use of the sensitivity analysis, requires only about 0.3 seconds. Note that a sensitivity analysis for the FE problem requires to solve (3) in the high-dimensional FE spaces. In conclusion, we point out that the RB solutions of the multiobjective problems are much faster than the FE ones, but thanks to an inexpensive 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.

7 Conclusions

We consider multiobjective optimal problems governed by linear variational equations, which depend on geometrical and/or model parameter vector $\boldsymbol{\mu}_c$. The goal is to propose a numerical strategy which allows to quickly determine a sufficiently accurate approximation of the set of Pareto optimal points for an arbitrarily chosen $\boldsymbol{\mu}_c$. To compute Pareto optimal points we apply the weighted sum method which requires to solve a very large number of scalar-valued optimal control problems, where the cost functional is built through an additional optimization parameter $\boldsymbol{\mu}_o$. These problems can be solved very efficiently by the proposed RB strategy. By applying a sensitivity analysis we are also able to reduce significantly the number of different values $\boldsymbol{\mu}_o$ that have to be chosen in the weighted sum method to identifying the Pareto optimal solutions. To sum up, the use of the RB method – together with an useful and inexpensive sensitive analysis – allows to solve multiobjective problems, at a very low 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. The second author was supported by the DFG within SFB 666 and SFB 805 and by the BMBF within SIMUROM.

References

1. Babuška, I.: The finite element method with Lagrange multipliers. Numer. Math. **20**, 179–192 (1973)
2. Banholzer, S., Beermann, D., Volkwein, S.: POD-based bicriterial optimal control by the reference point method. IFAC-PapersOnLine **49**(8), 210–215 (2016)
3. Barrault, M., Maday, Y., Nguyen, N.C., Patera, A.T.: 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)
4. Brezis, H.: Functional analysis, Sobolev spaces and partial differential equations. Springer, New York (2011)
5. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. ESAIM: Math. Modell. Numer. Anal. **8**, 129–151 (1974)
6. Brezzi, F., Fortin, M.: Mixed and hybrid finite elements methods. Springer, New York (1991)
7. Das, I.: newblock nonlinear multicriteria optimization and robust optimality. Ph.D thesis. Rice University, Houston, Texas (1997)
8. Dede, L.: Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. SIAM J. Sci. Comput. **32**, 997–1019 (2010)
9. Ehrgott, M.: Multicriteria optimization. Springer, Berlin (2005)
10. Evans, L.C.: Partial differential equations. American Mathematics of Society Providence, Rhode, Island (2008)
11. Göpfert, A., Nehse, R.: Vektoroptimierung. BSB Teubner Verlagsgesellschaft, Leibzig (1990)
12. Grepl, M., Kärcher, M.: A posteriori error estimation for reduced order solutions of parametrized parabolic optimal control problems. To appear in ESAIM: Mathematical Modelling and Numerical Analysis (2014)
13. Griesse, R., Vexler, B.: Numerical sensitivity analysis for the quantity of interest in PDE-constrained optimization. SIAM J. Sci. Comput. **29**, 22–48 (2007)
14. Hesthaven, J., Rozza, G., Stamm, B.: Certified reduced basis methods for parametrized partial differential equations. Springer Briefs in Mathematics, Springer Cham (2016)
15. Hillermeier, C.: Nonlinear multiobjective optimization. A generalized homotopy approach. Basel, Birkhäuser Verlag (2001)
16. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints. Springer (2009)
17. Huynh, D.B.P., Knezevic, D.J., Chen, Y., Hesthaven, J.S., Patera, A.T.: A natural-norm successive constraint method for inf-sup lower bounds. Comput. Methods Appl. Mech. Eng. **199**, 1963–1975 (2010)
18. Iapichino, L., Volkwein, S.: Optimization strategy for parameter sampling in the reduced basis method. IFAC-PapersOnLine **48**, 707–712 (2015)
19. Iapichino, L. 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
20. Iapichino, L., Volkwein, S. Greedy sampling of distributed parameters in the reduced-basis method by numerical optimization Konstanzer Schriften in Mathematik No. 308 University of Konstanz, 2013. http://nbn-resolving.de/urn:nbn:de:bsz:352-228388
21. Iapichino, L., Trenz, S., Volkwein, S.: Reduced-order multiobjective optimal control of semilinear parabolic problems. In: Numerical Mathematics and Advanced Applications (ENUMATH

2015), Lecture Notes in Computational Science and Engineering, vol. 112, pp. 389–397 (2016)
22. Kammann, E., Tröltzsch, F., Volkwein, S.: A method of a-posteriori error estimation with application to proper orthogonal decomposition. Math. Modell. Numer. Anal. **47**(933–934), 218–225 (2013)
23. Kuhn, H., Tucker, A.: Nonlinear programming. In: Newman, J. (ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 481–492. University of California Press, Berkeley (1951)
24. Negri, F., Rozza, G., Manzoni, A., Quateroni, A.: Reduced basis method for parametrized elliptic optimal control problems. SIAM J. Sci. Comput. **35**(5), A2316–A2340
25. Nocedal, J., Wright, S.J. Numerical optimization, 2nd edn. Springer Series in Operation Research (2006)
26. Patera, A.T., Rozza, G.: Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations. MIT Pappalardo Graduate Monographs in Mechanical Engineering (2006)
27. Peitz, S., Oder-Blöbaum, S., Dellnitz, M.: Multiobjective optimal control methods for fluid flow using reduced order modeling. Submitted (2015)
28. Quarteroni, A., Manzoni, A., Negri, F.: Reduced basis methods for partial differential equatiuons an introduction. Springer Cham (2016)
29. Rozza, G., Huynh, D.B.P., Patera, A.T.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Arch. Comput. Methods Eng. **15**, 229–275 (2008)
30. Rozza, G., Veroy, K.: On the stability of the reduced basis method for Stokes equations in parametrized domains. Comput. Meth. Appl. Mech. Engr. **7**, 1244–1260 (2007)
31. Stadler, W.: Multicriteria optimization in engineering and in the sciences. Plenum Press, New York (1988)
32. Tröltzsch, F.: Optimal control of partial differential equations. Theory, Methods and applications, vol. 112. American Math. Society, Providence (2010)
33. Tröltzsch, F., Volkwein, S.: POD a-posteriori error estimates for linear-quadratic optimal control problems. Comput. Optim. Appl. **44**, 83–115 (2009)
34. Urban, K., Volkwein, S., Zeeb, O.: Greedy sampling using nonlinear optimization. In: Quarteroni, A., Rozza, R. (eds.) Reduced Order Methods for Modeling and Computational Reduction, pp. 137—157. Springer (2014)
35. Zadeh, L.: OptiMality and non-scalar-valued performance criteria. IEEE Transactions on Automatic Control, 8 (1963)