

# A NON-CONFORMING DUAL APPROACH FOR ADAPTIVE TRUST-REGION REDUCED BASIS APPROXIMATION OF PDE-CONSTRAINED PARAMETER OPTIMIZATION

TIM KEIL<sup>1,\*</sup>, LUCA MECHELLI<sup>2</sup>, MARIO OHLBERGER<sup>1</sup>, FELIX SCHINDLER<sup>1</sup> AND STEFAN VOLKWEIN<sup>2</sup>

**Abstract.** In this contribution we propose and rigorously analyze new variants of adaptive Trust-Region methods for parameter optimization with PDE constraints and bilateral parameter constraints. The approach employs successively enriched Reduced Basis surrogate models that are constructed during the outer optimization loop and used as model function for the Trust-Region method. Each Trust-Region sub-problem is solved with the projected BFGS method. Moreover, we propose a non-conforming dual (NCD) approach to improve the standard RB approximation of the optimality system. Rigorous improved *a posteriori* error bounds are derived and used to prove convergence of the resulting NCD-corrected adaptive Trust-Region Reduced Basis algorithm. Numerical experiments demonstrate that this approach enables to reduce the computational demand for large scale or multi-scale PDE constrained optimization problems significantly.

**Mathematics Subject Classification.** 49M20, 49K20, 35J20, 65N30, 90C06.

Received June 16, 2020. Accepted April 12, 2021.

## 1. INTRODUCTION

We are concerned with the development and rigorous analysis of novel efficient model order reduction methods for parameter optimization constrained by coercive variational state equations using the *first optimize, then discretize* approach. The methods are based on successive enrichment of the underlying reduced order models within the framework of Trust-Region optimization. Optimization problems constrained by partial differential equations (PDEs) arise in many fields of application in engineering and across all sciences. Examples of such problems include optimal (material) design or optimal control of processes and inverse problems, where parameters of a PDE model are unknown and need to be estimated from measurements. The numerical solution of such problems is very challenging as the underlying PDEs have to be solved repeatedly within outer optimization algorithms and the dimension of the parameters that need to be optimized might be very high or even infinite dimensional. PDE constrained optimization problems have been of interest for many decades. Classically, the

---

*Keywords and phrases.* PDE constrained optimization, Trust-Region method, error analysis, Reduced Basis method, model order reduction, parametrized systems, large scale problems.

<sup>1</sup> Mathematics Münster, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, D-48149 Münster, Germany.

<sup>2</sup> Department of Mathematics and Statistics, University of Konstanz, D-78457 Konstanz, Germany.

\*Corresponding author: [tim.keil@wwu.de](mailto:tim.keil@wwu.de)

underlying PDE (forward problem) is approximated by a high dimensional full order model (FOM) that results from discretization, *e.g.*, by the Finite Element or Finite Volume method. Hence, the complexity of the optimization problem directly depends on the numbers of degrees of freedom (DOF) of the FOM. Mesh adaptivity has been advised to minimize the number of DOFs; see, *e.g.*, [3, 4, 15, 33, 42, 54] and the references therein.

### Model order reduction for PDE constrained optimization and optimal control

A more recent approach is the usage of model order reduction (MOR) methods in order to replace the FOM by a surrogate reduced order model (ROM) of possibly very low dimension. MOR is a very active research field that has seen tremendous development in recent years, both from a theoretical and application point of view. For an introduction and overview we refer to the monographs and collections [5, 6, 30, 51]. A particular promising model reduction approach for parameterized partial differential equations (pPDEs) is the Reduced Basis (RB) method that relies on the approximation of the solution manifold of pPDEs by low dimensional linear approximation spaces that are spanned from suitably selected particular solutions, called snapshots. *A posteriori* error estimation for solutions of the ROM with respect to the FOM is the basis for efficient Greedy algorithms to select the snapshots in a quasi-optimal way [9, 26]. Alternatively, construction of reduced bases using proper orthogonal decomposition (POD) may be used [25]. The construction of a reduced basis and the respective projected ROM is generally called the offline phase, whereas evaluating the ROM is called online phase.

There exists a large amount of literature using such reduced order surrogate models for optimization methods. *A posteriori* error estimates for reduced approximation of linear-quadratic optimization problems and parametrized optimal control problems with control constraints were studied, *e.g.*, in [17, 24, 37, 44, 49]. In [18], an RB approach is proposed which also enables an estimation on the actual error on the control variable and not only on the gradient of the output functional. Certified Reduced Basis methods for parametrized elliptic optimal control problems with distributed controls were studied in [38]. With the help of an *a posteriori* error estimator, ROMs can be constructed with respect to a desired accuracy but also with respect to a local area in the parameter set [20, 28]. For very high dimensional parameter sets, simultaneous parameter and state reduction has been advised [31, 32, 41]. However, constructing a reduced order surrogate for a prohibitively expensive forward problem can also take a significant amount of computational resources. To remedy this, it is beneficial to use optimization methods that optimize on a local level of the control variable, assuming the surrogate only to be accurate enough in the respective parameter region. Hence, we require an approach which goes beyond the classical offline/online decomposition. Recently, RB methods have been advised with a progressive construction of ROMs [8, 21, 58]. Also localized RB methods that are based on efficient localized *a posteriori* error control and online enrichment [11, 46] overcome traditional offline/online splitting and are thus particularly well suited for applications in optimization or inverse problems [47, 48].

### Trust-Region reduced order models for second-order methods

Trust-Region (TR) approaches are a class of optimization methods that are advantageous for the usage of locally accurate surrogate models. The main idea is to solve optimization sub-problems only in a local area of the parameter set which resolves the burden of constructing a global RB space. The problem that might occur is the fact that during this minimization one usually moves away from the original parameters on which the reduced order model was built, and the quality of the reduced model cannot be guaranteed anymore. For that reason, *a priori* and *a posteriori* error analysis are required to ensure accurate reduced order approximations for the optimization problem; *cf.* [25, 34, 37]. In [1, 56], a TR approach was proposed to control the quality of the (POD) reduced order model, referred to as TR-POD, a meanwhile well-established method in applications; *cf.* [7, 14].

TR methods ensure global convergence for locally convergent methods. In each iteration of the TR algorithm the nonlinear objective is replaced by a model function which can be optimized with much less effort; *cf.* [16, 45]. One suitable choice for the model is a reduced order discretization of the objective (*e.g.*, by utilizing a second-order Taylor approximation). To ensure convergence to stationary points the accuracy of the model function and

of its gradient have to be monitored. In [53], *a posteriori* error bounds are utilized to monitor the approximation quality of the gradient. We also refer to [23], where the authors utilize basis update strategies to improve the reduced order approximation scheme with respect to the optimization goal. The TR strategy can be combined with second-order methods for nonlinear optimization: with the Newton method to solve the reduced problem and with the SQP method for the all-at-once approach; cf. [29].

Constraints on the control and the metric for the Trust-Region radius can affect the convergence of the method. For an error-aware TR method, the TR radius is directly characterized by the *a posteriori* error estimator for the cost functional of the surrogate model. Thus, the offline phase of the RB method can completely be omitted since the RB model can be adaptively enriched during the outer optimization loop. With this procedure the surrogate model eventually will have a high accuracy around the optimum of the optimization problem, ignoring the accuracy of the part which the outer (and inner) optimization loop does not approach at all. Error aware TR-RB methods can be utilized in many different ways. One possible TR-RB approach has been extensively studied in [50] for linear parametric elliptic equations, which ensures convergence of the nonlocal TR-RB. Note that the experiments in [50] are for up to six dimensional parameter sets without inequality constraints. In [57], the TR framework is combined with an efficient RB error bound for defining the Trust-Region in the design optimization of vibrating structures using frequency domain formulations.

### Main results

In this contribution we present several significant advances for adaptive Trust-Region Reduced Basis optimization methods for parameterized partial differential equations:

- For the model function in the TR-RB approach, we follow a non-conforming dual (NCD) approach by choosing as model function the Lagrangian associated to the optimization problem. This permits more accurate results in terms of approximation of the optimal solution;
- we provide efficiently computable *a posteriori* error estimates for all reduced quantities for different choices of the cost functional and its (approximate) gradient;
- we rigorously prove the convergence of the TR-RB method with bilateral inequality constraints on the parameters;
- we devise several new adaptive enrichment strategies for the progressive construction of the Reduced Basis spaces;
- we demonstrate in numerical experiments that our new TR-RB methods outperform existing model reduction approaches for large scale optimization problems in well defined benchmark problems.

### Organization of the article

In Section 2, we introduce the PDE constrained optimization problem and state first- and second-order optimality conditions. These serve as a basis for the full order discretization derived in Section 3. Moreover, in Section 3 we introduce different strategies of model order reduction for the full order model and derive rigorous *a posteriori* error estimates for all equations, functionals, and gradient information. Section 4 is devoted to the derivation of Trust-Region – Reduced Basis methods and the presentation of the convergence analysis of the adaptive TR-RB algorithm. In addition, we discuss in detail several variants of new TR-RB algorithms that differ in their respective reduced gradient information as well as in the enrichment strategies for the construction of the corresponding reduced models. All variants are thoroughly analyzed numerically in Section 5, where we consider three well defined benchmark problems. We also compare with selected state of the art optimization methods from the literature.

## 2. PROBLEM FORMULATION

Given  $\mu_a, \mu_b \in \mathbb{R}^P$  with  $P \in \mathbb{N}$  we consider the compact and convex admissible parameter set

$$\mathcal{P} := \{\mu \in \mathbb{R}^P \mid \mu_a \leq \mu \leq \mu_b\} \subset \mathbb{R}^P,$$

where  $\leq$  is understood component-wise. Let  $V$  be a real-valued Hilbert space with inner product  $(\cdot, \cdot)$  and induced norm  $\|\cdot\|$ . We are interested in efficiently approximating PDE-constrained parameter optimization problems with the quadratic continuous cost functional

$$\mathcal{J} : V \times \mathcal{P} \rightarrow \mathbb{R}, \quad (u, \mu) \mapsto \mathcal{J}(u, \mu) = \Theta(\mu) + j_\mu(u) + k_\mu(u, u),$$

where  $\Theta : \mathcal{P} \rightarrow \mathbb{R}$  denotes a parameter function and, for each  $\mu \in \mathcal{P}$ ,  $j_\mu \in V'$  is a parameter-dependent continuous linear functional and  $k_\mu : V \times V \rightarrow \mathbb{R}$  a continuous symmetric bilinear form. To be more precise, we consider the following constrained minimization problem:

$$\min_{(u, \mu) \in V \times \mathcal{P}} \mathcal{J}(u, \mu), \tag{P.a}$$

subject to  $(u, \mu)$  satisfying the *state – or primal – equation*

$$a_\mu(u, v) = l_\mu(v) \quad \text{for all } v \in V, \tag{P.b}$$

where, for each  $\mu \in \mathcal{P}$ ,  $a_\mu : V \times V \rightarrow \mathbb{R}$  denotes a continuous and coercive bilinear form and  $l_\mu \in V'$  denotes a continuous linear functional. For given  $u \in V$ ,  $\mu \in \mathcal{P}$ , we introduce the primal residual  $r_\mu^{\text{pr}}(u) \in V'$  associated with (P.b) by

$$r_\mu^{\text{pr}}(u)[v] := l_\mu(v) - a_\mu(u, v) \quad \text{for all } v \in V. \tag{2.1}$$

The primal residual plays a crucial role for *a posteriori* error analysis and for sensitivities of solution maps.

**Remark 2.1.** The Lagrange functional for (P) is given by  $\mathcal{L}(u, \mu, p) = \mathcal{J}(u, \mu) + r_\mu^{\text{pr}}(u)[p]$  for  $(u, \mu) \in V \times \mathcal{P}$  and for  $p \in V$ .

To apply RB methods efficiently, we require the parametrization of the problem to be separable from  $V$  throughout the work. This separability is a standard assumption for RB methods and can be circumvented by using empirical interpolation techniques [2, 13, 19].

**Assumption 2.2** (Parameter-separability). *We assume  $a_\mu$ ,  $l_\mu$ ,  $j_\mu$ ,  $k_\mu$  to be parameter separable with  $\Xi^a, \Xi^l, \Xi^j, \Xi^k \in \mathbb{N}$  non-parametric components  $a_\xi : V \times V \rightarrow \mathbb{R}$  for  $1 \leq \xi \leq \Xi^a$ ,  $l_\xi \in V'$  for  $1 \leq \xi \leq \Xi^l$ ,  $j_\xi \in V'$  for  $1 \leq \xi \leq \Xi^j$  and  $k_\xi : V \times V \rightarrow \mathbb{R}$  for  $1 \leq \xi \leq \Xi^k$ , and respective parameter functions  $\theta_\xi^a, \theta_\xi^l, \theta_\xi^j, \theta_\xi^k : \mathcal{P} \rightarrow \mathbb{R}$ , such that*

$$a_\mu(u, v) = \sum_{i=1}^{\Xi^a} \theta_i^a(\mu) a_i(u, v), \quad l_\mu(v) = \sum_{i=1}^{\Xi^l} \theta_i^l(\mu) l_i(v),$$

and analogously for  $j_\mu$  and  $k_\mu$ .

Due to Assumption 2.2, all quantities which linearly depend on  $a_\mu$ ,  $l_\mu$ ,  $j_\mu$  and  $k_\mu$  (such as  $\mathcal{J}$  or the primal residual) are also separable w.r.t. the parameter. Since we will use a Lagrangian ansatz for an explicit computation of derivatives, we require some notation that we use throughout this paper.

## 2.1. A note on differentiability

If  $\mathcal{J} : V \times \mathcal{P} \rightarrow \mathbb{R}$  is Fréchet differentiable w.r.t. each  $\mu \in \mathcal{P}$ , for each  $u \in V$  and each  $\mu \in \mathcal{P}$  there exists a bounded linear functional  $\partial_\mu \mathcal{J}(u, \mu) \in \mathbb{R}^P$ , such that the Fréchet derivative of  $\mathcal{J}$  w.r.t. its second argument in the direction of  $\nu \in \mathbb{R}^P$  is given by  $\partial_\mu \mathcal{J}(u, \mu) \cdot \nu$  (noting that the dual space of  $\mathbb{R}^P$  is itself). We refer to  $\partial_\mu \mathcal{J}(u, \mu)$  as the derivative w.r.t.  $\mu$ . In addition, for  $u \in V$ ,  $\mu \in \mathcal{P}$  we denote the *partial derivative* of  $\mathcal{J}(u, \mu)$  w.r.t. the  $i$ th component of  $\mu$  by  $\partial_{\mu_i} \mathcal{J}(u, \mu)$  for  $1 \leq i \leq P$ . Note that  $\partial_{\mu_i} \mathcal{J}(u, \mu) = \partial_\mu \mathcal{J}(u, \mu) \cdot e_i$ , where  $e_i \in \mathbb{R}^P$  denotes the  $i$ th canonical unit vector. Furthermore, we denote the *gradient* w.r.t. its second argument – the

vector of components  $\partial_{\mu_i} \mathcal{J}(u, \mu)$  – by the operator  $\nabla_{\mu} \mathcal{J} : V \times \mathcal{P} \rightarrow \mathbb{R}^P$ . Similarly, if  $\mathcal{J}$  is Fréchet differentiable w.r.t. each  $u \in V$ , for each  $u \in V$  and each  $\mu \in \mathcal{P}$  there exists a bounded linear functional  $\partial_u \mathcal{J}(u, \mu) \in V'$ , such that the Fréchet derivative of  $\mathcal{J}$  w.r.t. its first argument in any direction  $v \in V$  is given by  $\partial_u \mathcal{J}(u, \mu)[v]$ . We refer to  $\partial_u \mathcal{J}(u, \mu)$  simply as the derivative w.r.t.  $u$ . If  $\mathcal{J}$  is twice Fréchet differentiable w.r.t. each  $\mu \in \mathcal{P}$ , we denote its *hessian* w.r.t. its second argument by the operator  $\mathcal{H}_{\mu} \mathcal{J} : V \times \mathcal{P} \rightarrow \mathbb{R}^{P \times P}$ .

We treat  $a$ ,  $l$ ,  $j$  and  $k$  in a similar manner, although, for notational compactness, we indicate their parameter-dependency differently. For instance, interpreting the bilinear form  $a$  as a map  $a : V \times V \times \mathcal{P} \rightarrow \mathbb{R}$ ,  $(u, v, \mu) \mapsto a_{\mu}(u, v)$ , we denote the Fréchet derivatives of  $a$  w.r.t. the first, second and third argument of said map in the direction of  $w \in V$ ,  $\nu \in \mathbb{R}^P$  by  $\partial_u a_{\mu}(u, v)[w] \in \mathbb{R}$ ,  $\partial_v a_{\mu}(u, v)[w] \in \mathbb{R}$  and  $\partial_{\mu} a_{\mu}(u, v) \cdot \nu \in \mathbb{R}$ , respectively. Similarly, interpreting the linear functional  $l$  as a map  $l : V \times \mathcal{P} \rightarrow \mathbb{R}$ ,  $(v, \mu) \mapsto l_{\mu}(v)$ , we denote the Fréchet derivatives of  $l$  w.r.t. the first and second argument of said map in the direction of  $w \in V$ ,  $\nu \in \mathbb{R}^P$  by  $\partial_v l_{\mu}(v)[w] \in \mathbb{R}$  and  $\partial_{\mu} l_{\mu}(v) \cdot \nu \in \mathbb{R}$ , respectively. We omit the word Fréchet when referring to the derivatives of  $\mathcal{J}$ ,  $a$ ,  $l$ ,  $j$  and  $k$ , in order to simplify the notation, unless it is strictly necessary to specify it.

We apply this notation for Fréchet and partial derivatives for functionals and bilinear forms throughout this manuscript. Note that we denote the derivatives w.r.t. the symbol of the argument in the original definition of the functional or bilinear form, not w.r.t. the symbol of the actual argument, *i.e.*, we use  $\partial_u \mathcal{J}(u_{\mu}, \mu)$  for the derivative w.r.t. the first argument, not  $\partial_{u_{\mu}} \mathcal{J}(u_{\mu}, \mu)$  or  $\partial_v a_{\mu}(u, p)$  for the derivative w.r.t. the second argument, not  $\partial_p a_{\mu}(u, p)$ . In case  $u$  is also dependent on  $\mu$ , we denote the total derivative w.r.t.  $\mu_i$  by  $d_{\mu_i}$ , *i.e.*,  $d_{\mu_i} \mathcal{J}(u_{\mu}, \mu) = \partial_{\mu_i} \mathcal{J}(u_{\mu}, \mu) + \partial_u \mathcal{J}(u_{\mu}, \mu)[d_{\mu_i} u_{\mu}]$ . Note also that, due to Assumption 2.2, we can exchange the order of differentiation w.r.t.  $V$  and  $\mathbb{R}^P$ , *i.e.*,  $\partial_u(\partial_{\mu} \mathcal{J}(u, \mu) \cdot \nu)[w] = \partial_{\mu}(\partial_u \mathcal{J}(u, \mu)[w]) \cdot \nu$ .

**Assumption 2.3** (Differentiability of  $a$ ,  $l$  and  $\mathcal{J}$ ). *We assume  $a_{\mu}$ ,  $l_{\mu}$  and  $\mathcal{J}$  to be twice Fréchet differentiable w.r.t.  $\mu$ . This obviously requires that all parameter-dependent coefficient functions in Assumption 2.2 are twice differentiable as well.*

**Remark 2.4** (Derivatives w.r.t.  $V$ ). Due to the (bi-)linearity of  $a$ ,  $l$ ,  $j$  and  $k$ , we can immediately compute their derivatives w.r.t. arguments in  $V$ . For  $u, v \in V$ ,  $\mu \in \mathcal{P}$ , the derivatives of  $a$ ,  $l$  and  $\mathcal{J}$  w.r.t. arguments in  $V$  in the direction of  $w \in V$  are given, respectively, by

$$\partial_u a_{\mu}(u, v)[w] = a_{\mu}(w, v), \quad \partial_v a_{\mu}(u, v)[w] = a_{\mu}(u, w), \quad \partial_v l_{\mu}(v)[w] = l_{\mu}(w), \quad \partial_u \mathcal{J}(u, \mu)[w] = j_{\mu}(w) + 2k_{\mu}(w, u).$$

We compute the partial derivatives of  $a$  and  $l$  w.r.t. the parameter by means of their separable decomposition.

**Remark 2.5** (Derivatives w.r.t.  $\mathcal{P}$ ). For  $\mu \in \mathcal{P}$ ,  $u, v \in V$  the derivatives of  $a$  and  $l$  w.r.t.  $\mu$  in the direction of  $\nu \in \mathbb{R}^P$  are given by

$$\partial_{\mu} a_{\mu}(u, v) \cdot \nu = \sum_{\xi=1}^{\Xi^a} (\partial_{\mu} \theta_{\xi}^a(\mu) \cdot \nu) a_{\xi}(u, v) \quad \text{and} \quad \partial_{\mu} l_{\mu}(v) \cdot \nu = \sum_{\xi=1}^{\Xi^l} (\partial_{\mu} \theta_{\xi}^l(\mu) \cdot \nu) l_{\xi}(v),$$

respectively, if  $u, v$  do not depend on  $\mu$ . We also introduce the following shorthand notation for the derivative of functionals and bilinear forms w.r.t. the parameter in the direction of  $\nu \in \mathbb{R}^P$ , *e.g.*, for  $\mu \in \mathcal{P}$  we introduce

$$\begin{aligned} \partial_{\mu} l_{\mu} \cdot \nu \in V' & & v \mapsto (\partial_{\mu} l_{\mu} \cdot \nu)(v) & := \partial_{\mu} l_{\mu}(v) \cdot \nu & \text{and} \\ \partial_{\mu} a_{\mu} \cdot \nu \in V \times V \rightarrow \mathbb{R} & & u, v \mapsto (\partial_{\mu} a_{\mu} \cdot \nu)(u, v) & := \partial_{\mu} a_{\mu}(u, v) \cdot \nu, \end{aligned}$$

and note that  $\partial_{\mu} l_{\mu}$  and  $\partial_{\mu} a_{\mu}$  are continuous and separable w.r.t. the parameter, owing to Assumption 2.2.

The bilinear form  $a_{\mu}(\cdot, \cdot)$  is continuous and coercive for all  $\mu \in \mathcal{P}$ . Thus we can define the bounded solution map  $\mathcal{S} : \mathcal{P} \rightarrow V$ ,  $\mu \mapsto u_{\mu} := \mathcal{S}(\mu)$ , where  $u_{\mu}$  is the unique solution to (P.b) for a given  $\mu$ . The Fréchet derivatives of  $\mathcal{S}$  are a common tool for RB methods and optimization, *e.g.*, for constructing Taylor RB spaces that consist of the primal solution as well as their sensitivities (see [27]) or for deriving optimality conditions for (P) (see [35]).

**Proposition 2.6** (Fréchet derivative of the solution map). *Considering the solution map  $\mathcal{S} : \mathcal{P} \rightarrow V$ ,  $\mu \mapsto u_\mu$ , its Fréchet derivative  $d_\nu u_\mu \in V$  w.r.t. a direction  $\nu \in \mathbb{R}^P$  is the unique solution of*

$$a_\mu(d_\nu u_\mu, v) = \partial_\mu r_\mu^{\text{Pr}}(u_\mu)[v] \cdot \nu \quad \text{for all } v \in V. \tag{2.2}$$

*Proof.* We refer to [35] for the proof of this result. □

### 2.2. Optimal solution and optimality conditions

In this section, we discuss the existence of an optimal solution for problem (P). Then, we characterize a locally optimal solution through first- and second-order optimality conditions. Throughout the paper, a bar indicates optimality.

**Theorem 2.7** (Existence of an optimal solution). *Problem (P) admits an optimal solution pair  $(\bar{u}, \bar{\mu}) \in V \times \mathcal{P}$ , where  $\bar{u} := u_{\bar{\mu}}$  is the solution of (P.b) for the parameter  $\bar{\mu}$ .*

*Proof.* Note that the quantities involved in problem (P) satisfies Assumption 1.44 in [35]. Thus the existence follows from Theorem 1.45 of [35]. □

Let us introduce the reduced cost functional  $\hat{\mathcal{J}} : \mathcal{P} \mapsto \mathbb{R}$ ,  $\mu \mapsto \hat{\mathcal{J}}(\mu) := \mathcal{J}(u_\mu, \mu) = \mathcal{J}(\mathcal{S}(\mu), \mu)$ . Then problem (P) is equivalent to the so-called reduced problem

$$\min_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu). \tag{\hat{P}}$$

**Remark 2.8.** (1) Since  $r_\mu^{\text{Pr}}(u_\mu)[p] = 0$  for any  $p \in V$ , it follows that  $\hat{\mathcal{J}}(\mu) = \mathcal{L}(u_\mu, \mu, p)$  for any  $p \in V$ .

- (2) The cost functional  $\hat{\mathcal{J}}$  is in general non-convex, thus the existence of a unique minimum for  $\hat{\mathcal{J}}$  (and thus of  $\mathcal{J}$ ) can not be guaranteed.
- (3) Let  $(\bar{u}, \bar{\mu}) \in V \times \mathcal{P}$  be a local optimal solution to (P) with  $\bar{u} := u_{\bar{\mu}}$  the solution of the primal equation (P.b) for the parameter  $\bar{\mu}$ . Then the following constraint qualification holds true: For any  $f \in V'$  there exists a pair  $(u, \mu) \in V \times \mathbb{R}^P$  solving

$$a_{\bar{\mu}}(u, v) - \partial_\mu r_{\bar{\mu}}^{\text{Pr}}(\bar{u})[v] \cdot \mu = f(v) \quad \text{for all } v \in V.$$

- (4) Theorem 2.7 does not provide any solution method.

One can derive first-order necessary optimality conditions in order to compute candidates for a local optimal solution of (P). We refer to Corollary 1.3 of [35] for a proof of the following result:

**Proposition 2.9** (First-order necessary optimality conditions for (P)). *Let  $(\bar{u}, \bar{\mu}) \in V \times \mathcal{P}$  be a local optimal solution to (P). Moreover, let Assumption 2.3 hold true. Then there exists a unique Lagrange multiplier  $\bar{p} \in V$  such that the following first-order necessary optimality conditions hold:*

$$r_{\bar{\mu}}^{\text{Pr}}(\bar{u})[v] = 0 \quad \text{for all } v \in V, \tag{2.3a}$$

$$\partial_u \mathcal{J}(\bar{u}, \bar{\mu})[v] - a_{\bar{\mu}}(v, \bar{p}) = 0 \quad \text{for all } v \in V, \tag{2.3b}$$

$$(\nabla_\mu \mathcal{J}(\bar{u}, \bar{\mu}) + \nabla_\mu r_{\bar{\mu}}^{\text{Pr}}(\bar{u})[\bar{p}]) \cdot (\nu - \bar{\mu}) \geq 0 \quad \text{for all } \nu \in \mathcal{P}. \tag{2.3c}$$

Note that (2.3a) resembles the state equation (P.b). From (2.3b) we deduce the *adjoint – or dual – equation* with unique solution  $p_\mu \in V$  for a fixed  $\mu \in \mathcal{P}$ , i.e.,

$$a_\mu(v, p_\mu) = \partial_u \mathcal{J}(u_\mu, \mu)[v] = j_\mu(v) + 2k_\mu(v, u_\mu) \quad \text{for all } v \in V, \tag{2.4}$$

given the solution  $u_\mu \in V$  to the state equation (P.b). From (2.3b) we observe that the variable  $\bar{p}$  of the optimal triple solves the dual equation (2.4) for  $\bar{\mu}$ . Similarly to the primal solution, we can consider the dual solution

map  $\mathcal{A} : \mathcal{P} \rightarrow V$ ,  $\mu \mapsto \mathcal{A}(\mu) := p_\mu$ , where  $p_\mu$  is the solution of (2.4) for the parameter  $\mu$ . In particular,  $\bar{p} = p_{\bar{\mu}}$ . For given  $u, p \in V$ , we also introduce the dual residual  $r_\mu^{\text{du}}(u, p) \in V'$  associated with (2.4) by

$$r_\mu^{\text{du}}(u, p)[v] := j_\mu(v) + 2k_\mu(v, u) - a_\mu(v, p) \quad \text{for all } v \in V. \quad (2.5)$$

In addition, from the dual equation (2.4), we obtain the following formulation for the dual sensitivities.

**Proposition 2.10** (Fréchet derivative of the dual solution map). *Considering the dual solution map  $\mathcal{A} : \mathcal{P} \rightarrow V$ ,  $\mu \mapsto p_\mu$ , its directional derivative  $d_\eta p_\mu \in V$  w.r.t. a direction  $\eta \in \mathcal{P}$  is the solution of*

$$a_\mu(q, d_\eta p_\mu) = -\partial_\mu a_\mu(q, p_\mu) \cdot \eta + d_\mu \partial_u \mathcal{J}(u_\mu, \mu)[q] \cdot \eta = \partial_\mu r_\mu^{\text{du}}(u_\mu, p_\mu)[q] \cdot \eta + 2k_\mu(q, d_\eta u_\mu) \quad (2.6)$$

for all  $q \in V$ , where the latter equality holds for quadratic  $\mathcal{J}$  as in (P.a).

*Proof.* Note that  $\mathcal{A}$  is well defined because the bilinear form  $a_\mu(\cdot, \cdot)$  is continuous and coercive. For a proof of the other claims we refer to [35], for instance.  $\square$

Furthermore, we can compute first-order derivatives of  $\hat{\mathcal{J}}$ .

**Proposition 2.11** (Gradient of  $\hat{\mathcal{J}}$ ). *For given  $\mu \in \mathcal{P}$ , the gradient of  $\hat{\mathcal{J}}$ ,  $\nabla_\mu \hat{\mathcal{J}} : \mathcal{P} \rightarrow \mathbb{R}^P$ , is given by*

$$\nabla_\mu \hat{\mathcal{J}}(\mu) = \nabla_\mu \mathcal{J}(u_\mu, \mu) + \nabla_\mu r_\mu^{\text{pr}}(u_\mu)[p_\mu] = \nabla_\mu \Theta(\mu) + \nabla_\mu j_\mu(u_\mu) + \nabla_\mu k_\mu(u_\mu, u_\mu) + \nabla_\mu l_\mu(p_\mu) - \nabla_\mu a_\mu(u_\mu, p_\mu).$$

*Proof.* This follows from (2.1), (2.2), (2.4) and (P.a), cf. [35].  $\square$

**Remark 2.12.** The proof of Proposition 2.11 relies on the fact that both  $u_\mu$  and  $p_\mu$  belong to the same space  $V$ ; cf. [35]. In particular, for any  $\mu \in \mathcal{P}$ , we have  $\nabla_\mu \hat{\mathcal{J}}(\mu) = \nabla_\mu \mathcal{L}(u_\mu, \mu, p_\mu)$ .

For  $\bar{\mu}$  satisfying the first-order necessary optimality conditions, we have that  $\bar{\mu}$  is a stationary point of the cost functional  $\hat{\mathcal{J}}$ . Thus,  $\bar{\mu}$  can be either a local minimum, a saddle point or a local maximum of the cost functional  $\hat{\mathcal{J}}$  (and obviously the same relationship occurs between  $(\bar{u}, \bar{\mu})$  and  $\mathcal{J}$ ). We thus consider second-order sufficient optimality conditions in order to characterize local minima of the functional  $\hat{\mathcal{J}}$ , requiring its hessian.

**Proposition 2.13** (Hessian of  $\hat{\mathcal{J}}$ ). *The hessian of  $\hat{\mathcal{J}}$ ,  $\hat{\mathcal{H}}_\mu := \mathcal{H}_\mu \hat{\mathcal{J}} : \mathcal{P} \rightarrow \mathbb{R}^{P \times P}$ , is determined by its application to a direction  $\nu \in \mathbb{R}^P$ , given by*

$$\begin{aligned} \hat{\mathcal{H}}_\mu(\mu) \cdot \nu &= \nabla_\mu \left( \partial_u \mathcal{J}(u_\mu, \mu)[d_\nu u_\mu] + l_\mu(d_\nu p_\mu) - a_\mu(d_\nu u_\mu, p_\mu) - a_\mu(u_\mu, d_\nu p_\mu) \right. \\ &\quad \left. + (\partial_\mu \mathcal{J}(u_\mu, \mu) + \partial_\mu l_\mu(p_\mu) - \partial_\mu a_\mu(u_\mu, p_\mu)) \cdot \nu \right), \end{aligned}$$

where  $u_\mu, p_\mu \in V$  denote the primal and dual solutions, respectively. For a quadratic  $\mathcal{J}$  as in (P.a) the above formula simplifies to

$$\begin{aligned} \hat{\mathcal{H}}_\mu(\mu) \cdot \nu &= \nabla_\mu \left( j_\mu(d_\nu u_\mu) + 2k_\mu(d_\nu u_\mu, u_\mu) + l_\mu(d_\nu p_\mu) - a_\mu(d_\nu u_\mu, p_\mu) - a_\mu(u_\mu, d_\nu p_\mu) \right. \\ &\quad \left. + (\partial_\mu \mathcal{J}(u_\mu, \mu) + \partial_\mu l_\mu(p_\mu) - \partial_\mu a_\mu(u_\mu, p_\mu)) \cdot \nu \right). \end{aligned}$$

*Proof.* See, e.g., [35] for the first part. The second one follows from Remark 2.4.  $\square$

**Proposition 2.14** (Second-order sufficient optimality conditions). *Let Assumption 2.3 hold true. Suppose that  $\bar{\mu} \in \mathcal{P}$  satisfies the first-order necessary optimality conditions (2.3). If  $\hat{\mathcal{H}}_\mu(\bar{\mu})$  is positive definite on the critical cone  $\mathcal{C}(\bar{\mu})$  at  $\bar{\mu} \in \mathcal{P}$ , i.e., if  $\nu \cdot (\hat{\mathcal{H}}_\mu(\bar{\mu}) \cdot \nu) > 0$  for all  $\nu \in \mathcal{C}(\bar{\mu}) \setminus \{0\}$ , with*

$$\mathcal{C}(\bar{\mu}) := \{ \nu \in \mathbb{R}^P \mid \exists \mu \in \mathcal{P}, c_1 > 0 : \nu = c_1(\mu - \bar{\mu}), \nabla_\mu \hat{\mathcal{J}}(\bar{\mu}) \cdot \nu = 0 \},$$

then  $\bar{\mu}$  is a strict local minimum of  $(\hat{\mathcal{P}})$ .

*Proof.* For this result we refer to [12, 45], for instance.  $\square$

### 3. HIGH DIMENSIONAL DISCRETIZATION AND MODEL ORDER REDUCTION

We first discretize the optimization problem (P) as well as the corresponding optimality conditions using the classical Ritz–Galerkin projection onto a possibly high dimensional approximation space  $V_h \subset V$ , such as conforming Finite Elements. Note that we restrict ourselves to a conforming approximation for simplicity and that we do not further specify the choice of  $V_h$ , as neither impacts the analysis below. Based on this idea, we then derive different ways for the ROM using the Reduced Basis method with possibly different reduced primal and dual state spaces. Thus, the resulting ROM optimality system will in general not be equivalent to a Ritz–Galerkin projection of the FOM one onto a reduced space  $V_r \subset V_h$ . For this reason, we will introduce a non-conforming dual-corrected (NCD-corrected) approach; *cf.* Section 3.3.

#### 3.1. FOM for the optimality system

For the discretization of the optimization problem we assume that a finite-dimensional subspace  $V_h \subset V$  is given and obtain the FOM for the optimality system of (P) by Ritz–Galerkin projection of equations (2.3) onto  $V_h$ . In particular, we have for each  $\mu \in \mathcal{P}$  the solution  $u_{h,\mu} \in V_h$  of the *discrete primal equation*

$$a_\mu(u_{h,\mu}, v_h) = l_\mu(v_h) \quad \text{for all } v_h \in V_h, \quad (3.1)$$

and hence  $r_\mu^{\text{pr}}(u_{h,\mu})[v_h] = 0$  for all  $v_h \in V_h$ ,  $\mu \in \mathcal{P}$ . We also have for each  $\mu \in \mathcal{P}$  the solution  $p_{h,\mu} \in V_h$  of the *discrete dual equation*

$$a_\mu(v_h, p_{h,\mu}) = \partial_u \mathcal{J}(u_{h,\mu}, \mu)[v_h] = j_\mu(v_h) + 2k_\mu(v_h, u_{h,\mu}) \quad \text{for all } v_h \in V_h, \quad (3.2)$$

and hence  $r_\mu^{\text{du}}(u_{h,\mu}, p_{h,\mu})[v_h] = 0$  for all  $v_h \in V_h$ ,  $\mu \in \mathcal{P}$ . Similarly, the *discrete primal sensitivity equations* for solving for  $d_\nu u_{h,\mu} \in V_h$  as well as *discrete dual sensitivity equations* for solving for  $d_\nu p_{h,\mu} \in V_h$  at any direction  $\nu \in \mathbb{R}^P$  follow analogously to Propositions 2.6 and 2.10. Furthermore,  $\hat{\mathcal{J}}$  is approximated by the *discrete reduced functional*

$$\hat{\mathcal{J}}_h(\mu) := \mathcal{J}(u_{h,\mu}, \mu) = \mathcal{L}(u_{h,\mu}, \mu, p_h) \quad \text{for all } p_h \in V_h, \quad (3.3)$$

where  $u_{h,\mu} \in V_h$  is the solution of (3.1) and we formulate the discrete optimization problem

$$\min_{\mu \in \mathcal{P}} \hat{\mathcal{J}}_h(\mu). \quad (\hat{\text{P}}_h)$$

Further,  $\bar{\mu}_h$  denotes a locally optimal solution to  $(\hat{\text{P}}_h)$  satisfying the first- and second-order optimality conditions.

**Remark 3.1.** Since  $u_{h,\mu}$  and  $p_{h,\mu}$  belong to the same space  $V_h$ , Propositions 2.9–2.11, 2.13, 2.14 from Section 3.1 hold for the FOM as well, with all quantities replaced by their discrete counterparts.

As usual in the context of RB methods, we eliminate the issue of “truth” by assuming that the high dimensional space  $V_h$  is accurate enough to approximate the true solution.

**Assumption 3.2** (This is the “truth”). *We assume that the primal discretization error  $\|u_\mu - u_{h,\mu}\|$ , the dual error  $\|p_\mu - p_{h,\mu}\|$ , the primal sensitivity errors  $\|d_{\mu_i} u_\mu - d_{\mu_i} u_{h,\mu}\|$  and the dual sensitivity errors  $\|d_{\mu_i} p_\mu - d_{\mu_i} p_{h,\mu}\|$  are negligible for all  $\mu \in \mathcal{P}$ ,  $1 \leq i \leq P$ .*

To define suitable ROMs, in what follows, we assume we are given problem adapted RB spaces  $V_r^{\text{pr}}, V_r^{\text{du}} \subset V_h$ , the construction of which is detailed in Section 4.3. We stress here that  $V_r^{\text{pr}}$  and  $V_r^{\text{du}}$  might not coincide, which will imply further discussions of the RB approximation of the optimality system (2.3).



### 3.2. ROM for the optimality system – Standard approach

Given a RB space  $V_r^{\text{pr}} \subset V_h$  of low dimension  $n := \dim V_r^{\text{pr}}$  and dual RB space  $V_r^{\text{du}} \subset V_h$  of low dimension  $m := \dim V_r^{\text{du}}$ , we obtain the RB approximation of state and adjoint equations as follows:

- RB approximation for (2.3a): For each  $\mu \in \mathcal{P}$  the primal variable  $u_{r,\mu} \in V_r^{\text{pr}}$  of the *RB approximate primal equation* is defined through

$$a_\mu(u_{r,\mu}, v_r) = l_\mu(v_r) \quad \text{for all } v_r \in V_r^{\text{pr}}. \quad (3.4a)$$

- RB approximation for (2.3b): For each  $\mu \in \mathcal{P}$ ,  $u_{r,\mu} \in V_r^{\text{pr}}$  the dual/adjoint variable  $p_{r,\mu} \in V_r^{\text{du}}$  satisfies the *RB approximate dual equation* through

$$a_\mu(q_r, p_{r,\mu}) = \partial_u \mathcal{J}(u_{r,\mu}, \mu)[q_r] = j_\mu(q_r) + 2k_\mu(q_r, u_{r,\mu}) \quad \text{for all } q_r \in V_r^{\text{du}}. \quad (3.4b)$$

Analogously to Proposition 2.6, we define the *RB solution map*  $\mathcal{S}_r : \mathcal{P} \rightarrow V_r^{\text{pr}}$  by  $\mu \mapsto u_{r,\mu}$  and analogously to Proposition 2.10 the *RB dual solution map*  $\mathcal{A}_r : \mathcal{P} \rightarrow V_r^{\text{du}}$  by  $\mu \mapsto p_{r,\mu}$ , where  $u_{r,\mu}$  and  $p_{r,\mu}$  denote the primal and dual reduced solutions of (3.4a) and (3.4b), respectively.

To approximate  $(\hat{\mathbf{P}}_h)$ , we introduce the *RB reduced functional* by

$$\hat{J}_r(\mu) := \mathcal{J}(u_{r,\mu}, \mu) = \mathcal{J}(\mathcal{S}_r(\mu), \mu), \quad \text{where } u_{r,\mu} \in V_r^{\text{pr}} \text{ is the solution of (3.4a)} \quad (3.5)$$

instead of  $\hat{J}_h$  and the problem of finding a locally optimal solution  $\bar{\mu}_r$  of

$$\min_{\mu \in \mathcal{P}} \hat{J}_r(\mu). \quad (3.6)$$

Now, a solution to the optimality system (2.3) is approximated by the RB triple  $(u_{r,\bar{\mu}_r}, \bar{\mu}_r, p_{r,\bar{\mu}_r})$ .

As proposed in [50], for computing an approximation of the gradient of  $\hat{J}_r$ , the gradient from Proposition 2.11 can be utilized by replacing  $u_\mu$  and  $p_\mu$  with their RB counterparts. However, it can not be guaranteed in general that the computed gradient is the actual gradient of  $\hat{J}_r$ , if  $V_r^{\text{pr}}$  and  $V_r^{\text{du}}$  are chosen to be different. To see this, we consider first the Lagrangian and note that, for  $1 \leq i \leq P$  and all  $p \in V_r^{\text{pr}}$ , it holds

$$\hat{J}_r(\mu) = \mathcal{L}(u_{r,\mu}, \mu, p), \quad (\nabla_\mu \hat{J}_r(\mu))_i = \partial_u \mathcal{L}(u_{r,\mu}, \mu, p)[d_{\mu_i} u_{r,\mu}] + d_{\mu_i} \mathcal{L}(u_{r,\mu}, \mu, p). \quad (3.7)$$

Now, following [50], we define the *inexact gradient*  $\tilde{\nabla}_\mu \hat{J}_r : \mathcal{P} \rightarrow \mathbb{R}^P$  by

$$(\tilde{\nabla}_\mu \hat{J}_r(\mu))_i := \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu})[p_{r,\mu}] = \partial_{\mu_i} \mathcal{L}(u_{r,\mu}, \mu, p_{r,\mu}) \quad (3.8)$$

for all  $1 \leq i \leq P$  and  $\mu \in \mathcal{P}$ , where  $u_{r,\mu} \in V_r^{\text{pr}}$  and  $p_{r,\mu} \in V_r^{\text{du}}$  denote the primal and approximate dual reduced solutions of (3.4a) and (3.4b), respectively. With the superscript  $\sim$  we stress that  $\tilde{\nabla}_\mu \hat{J}_r(\mu)$  is not the actual gradient of  $\hat{J}_r$ , but its approximation. Choosing in contrast  $p = p_{r,\mu} \in V_r^{\text{du}}$  in (3.7) and considering (3.8) lead to

$$(\nabla_\mu \hat{J}_r(\mu))_i = \partial_u \mathcal{L}(u_{r,\mu}, \mu, p_{r,\mu})[d_{\mu_i} u_{r,\mu}] + (\tilde{\nabla}_\mu \hat{J}_r(\mu))_i.$$

Note that, in general, it does not hold that  $\partial_u \mathcal{L}(u_{r,\mu}, \mu, p_{r,\mu}) = 0$ , since (3.4b) is not the dual equation with respect to the optimization problem (3.6), cf. Section 1.6.4 of [35], which would only be true if  $V_r^{\text{du}} \subseteq V_r^{\text{pr}}$ . Thus, (3.8) defines only an approximation of the true gradient of  $\hat{J}_r$  with the choice made in [50]. This introduces an additional approximation error in reconstructing the solution of the optimality system (2.3), which is well visible in our numerical experiments (see Sect. 5.3): the standard RB approach leads to a significant lack in accuracy, requiring additional steps to enrich the RB space and cover this gap. We therefore propose to add a correction term to  $\hat{J}_r$ , based on the previous remarks.

### 3.3. ROM for the optimality system – NCD-corrected approach

Following the primal-dual RB approach for linear output functionals ([27], Sect. 2.4), it is more suitable to add a correction term to the output functional for which improved error estimates are available. We seek to minimize the Lagrangian corresponding to problem (P). A similar approach, in the context of adaptive finite elements, can be found in [3, 52]. We utilize (3.4b) to extend the primal-dual RB approach of Section 2.4 from [27] to quadratic output functionals and define the *NCD-corrected RB reduced functional*

$$\hat{\mathcal{J}}_r(\mu) := \mathcal{L}(u_{r,\mu}, \mu, p_{r,\mu}) = \hat{\mathcal{J}}_r(\mu) + r_\mu^{\text{pr}}(u_{r,\mu})[p_{r,\mu}] \quad (3.9)$$

with  $u_{r,\mu} \in V_r^{\text{pr}}$  and  $p_{r,\mu} \in V_r^{\text{du}}$  the solutions of (3.4a) and (3.4b), respectively. Note that  $\hat{\mathcal{J}}_r$  coincides with the functional  $\hat{\mathcal{J}}_r$  in (3.5) if  $V_r^{\text{du}} = V_r^{\text{pr}}$ . We then consider the *RB reduced optimization problem* of finding a locally optimal solution  $\bar{\mu}_r$  of

$$\min_{\mu \in \mathcal{P}} \hat{\mathcal{J}}_r(\mu). \quad (\hat{\text{P}}_r)$$

Computing the actual gradient of  $\hat{\mathcal{J}}_r$  results in the next proposition, proved following ([35], Sect. 1.6.2).

**Proposition 3.3** (Gradient of the NCD-corrected RB reduced functional). *The  $i$ th component of the true gradient of  $\hat{\mathcal{J}}_r$  is given by*

$$(\nabla_\mu \hat{\mathcal{J}}_r(\mu))_i = \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu})[p_{r,\mu} + w_{r,\mu}] - \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})[z_{r,\mu}]$$

where  $u_{r,\mu} \in V_r^{\text{pr}}$  and  $p_{r,\mu} \in V_r^{\text{du}}$  denote the RB approximate primal and dual solutions of (3.4a) and (3.4b),  $z_{r,\mu} \in V_r^{\text{du}}$  solves

$$a_\mu(z_{r,\mu}, q) = -r_\mu^{\text{pr}}(u_{r,\mu})[q] \quad \forall q \in V_r^{\text{du}} \quad (3.10)$$

and  $w_{r,\mu} \in V_r^{\text{pr}}$  solves

$$a_\mu(v, w_{r,\mu}) = r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})[v] - 2k_\mu(z_{r,\mu}, v), \quad \forall v \in V_r^{\text{pr}}. \quad (3.11)$$

### 3.4. A posteriori error analysis

A *posteriori* error estimates are required for controlling the accuracy of the reduced order model. In addition, we also use them for the error aware TR method (which is explained in Sect. 4.1). We derive a *posteriori* error estimates for all reduced terms that we need for the TR method. Moreover, we suggest further advances for the reduction of sensitivities and gradients. From a model reduction perspective, these error estimates need to be computed efficiently such that the time for the evaluation for many parameters can be neglected. Note that Assumption 2.2 is crucial for this, since it allows to precompute most of the required terms. For any functional  $l \in V_h'$  or bilinear form  $a : V_h \times V_h \rightarrow \mathbb{R}$ , we denote their dual or operator norms  $\|l\|$  and  $\|a\|$  by the continuity constants  $\gamma_l$  and  $\gamma_a$ , respectively. We also use this notation for the continuity constants of all appearing (bi-)linear forms  $k_\mu$ ,  $\partial_{\mu_i} k_\mu$ ,  $\partial_{\mu_i} j_\mu$  and  $\partial_{\mu_i} a_\mu$ . The same consideration applies for the norm  $\|\cdot\|$  in  $V_h'$  of the residuals. For  $\mu \in \mathcal{P}$ , we denote the coercivity constant of  $a_\mu$  w.r.t. the  $V_h$ -norm by  $\underline{a}_\mu > 0$ .

#### 3.4.1. Standard RB estimates for the optimality system

We start with the residual based *a posteriori* error estimation for the primal variable, which is a standard result from RB theory and has extensively been used in the literature. For a proof, we refer to [55].

**Proposition 3.4** (Upper bound on the primal model reduction error). *For  $\mu \in \mathcal{P}$  let  $u_{h,\mu} \in V_h$  be the solution of (3.1) and  $u_{r,\mu} \in V_r^{\text{pr}}$  the solution of (3.4a). Then it holds*

$$\|u_{h,\mu} - u_{r,\mu}\| \leq \Delta_{\text{pr}}(\mu) := \underline{a}_\mu^{-1} \|r_\mu^{\text{pr}}(u_{r,\mu})\|.$$

For the reduced dual problem, a similar idea can be used to derive the following estimation, accounting for the fact that  $p_{r,\mu}$  is not a Galerkin projection of  $p_{h,\mu}$ . For a proof, we refer to Lemma 3 from [50].

**Proposition 3.5** (Upper bound on the dual model reduction error). *For  $\mu \in \mathcal{P}$ , let  $p_{h,\mu} \in V_h$  be the solution of (3.2) and  $p_{r,\mu} \in V_r^{\text{du}}$  the solution of (3.4b). Then it holds*

$$\|p_{h,\mu} - p_{r,\mu}\| \leq \Delta_{\text{du}}(\mu) := \underline{a}_\mu^{-1} \left( 2\gamma_{k_\mu} \Delta_{\text{pr}}(\mu) + \|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| \right).$$

In the next proposition we state the result of the standard approach from Theorem 4 of [50]. Furthermore we show an improved version by using, in contrast to [50], the NCD-corrected reduced functional, which results in an optimal higher order *a posteriori* upper bound without lower order terms.

**Proposition 3.6** (Upper bound on the model reduction error of the reduced output).

(i) *With the notation from above, we have for the standard RB reduced cost functional*

$$|\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_r(\mu)| \leq \Delta_{\hat{\mathcal{J}}_r}(\mu) := \Delta_{\text{pr}}(\mu) \|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| + \Delta_{\text{pr}}(\mu)^2 \gamma_{k_\mu} + |r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}]|.$$

(ii) *Furthermore, we have for the NCD-corrected RB reduced cost functional (or equivalently for the Lagrangian for any  $p \in V_h$ )*

$$|\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_r(\mu)| = |\mathcal{L}(u_{h,\mu}, \mu, p) - \mathcal{L}(u_{r,\mu}, \mu, p)| \leq \Delta_{\hat{\mathcal{J}}_r}(\mu) := \Delta_{\text{pr}}(\mu) \|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| + \Delta_{\text{pr}}(\mu)^2 \gamma_{k_\mu}.$$

*Proof.* We refer to Theorem 4 of [50] for a proof of (i). Regarding (ii), using the shorthand  $e_{h,\mu}^{\text{pr}} := u_{h,\mu} - u_{r,\mu}$  and  $a_\mu(e_{h,\mu}^{\text{pr}}, p_{r,\mu}) = r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}]$  lead us to

$$\begin{aligned} |\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_r(\mu)| &= |j_{h,\mu}(e_{h,\mu}^{\text{pr}}) + k_\mu(u_{h,\mu}, u_{h,\mu}) - k_\mu(u_{r,\mu}, u_{r,\mu}) - a_\mu(e_{h,\mu}^{\text{pr}}, p_{r,\mu})| \\ &= |r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [e_{h,\mu}^{\text{pr}}] - 2k_\mu(u_{r,\mu}, e_{h,\mu}^{\text{pr}}) + k_\mu(u_{h,\mu}, u_{h,\mu}) - k_\mu(u_{r,\mu}, u_{r,\mu})| \\ &\leq \|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| \|e_{h,\mu}^{\text{pr}}\| + \gamma_{k_\mu} \|e_{h,\mu}^{\text{pr}}\|^2, \end{aligned}$$

where we used the definition of the dual residual in the second equality and Cauchy–Schwarz for the inequality. The assertion follows by using Proposition 3.4.  $\square$

**Remark 3.7.** The estimator  $\Delta_{\hat{\mathcal{J}}_r}(\mu)$  is continuous w.r.t.  $\mu$ , since the Riesz-representative of the residual is continuous.

For the inexact and NCD-corrected gradient, we derive the following *a posteriori* estimators.

**Proposition 3.8** (Upper bound on the model reduction error of the gradient of reduced output).

(i) *For the inexact gradient  $\tilde{\nabla}_\mu \hat{\mathcal{J}}_r(\mu)$  from the standard-RB approach (3.8), we have*

$$\begin{aligned} \|\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \tilde{\nabla}_\mu \hat{\mathcal{J}}_r(\mu)\|_2 &\leq \Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu) = \|\underline{\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}}(\mu)\|_2 \quad \text{with} \\ (\underline{\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}}(\mu))_i &:= 2\Delta_{\text{pr}}(\mu) \|u_{r,\mu}\| \gamma_{\partial_{\mu_i} k_\mu} + \Delta_{\text{pr}}(\mu) (\gamma_{\partial_{\mu_i} j_\mu} + \gamma_{\partial_{\mu_i} a_\mu} \|p_{r,\mu}\|) \\ &\quad + \Delta_{\text{du}}(\mu) (\gamma_{\partial_{\mu_i} l_\mu} + \gamma_{\partial_{\mu_i} a_\mu} \|u_{r,\mu}\|) \\ &\quad + \Delta_{\text{pr}}(\mu) \Delta_{\text{du}}(\mu) \gamma_{\partial_{\mu_i} a_\mu} + (\Delta_{\text{pr}})^2(\mu) \gamma_{\partial_{\mu_i} k_\mu}. \end{aligned}$$

(ii) *For the gradient  $\nabla_\mu \hat{\mathcal{J}}_r(\mu)$  of the NCD-corrected reduced functional, computed with the adjoint approach from Definition 3.3, we have*

$$\|\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu)\|_2 \leq \Delta_{\nabla \hat{\mathcal{J}}_r}^*(\mu) = \|\underline{\Delta_{\nabla \hat{\mathcal{J}}_r}^*}(\mu)\|_2 \quad \text{with}$$

$$\begin{aligned}
(\Delta_{\nabla \hat{\mathcal{J}}_r}^*(\mu))_i &:= 2\Delta_{\text{pr}}(\mu)\|u_{r,\mu}\| \gamma_{\partial_{\mu_i} k_\mu} + \Delta_{\text{pr}}(\mu)(\gamma_{\partial_{\mu_i} j_\mu} + \gamma_{\partial_{\mu_i} a_\mu} \|p_{r,\mu}\|) \\
&\quad + \Delta_{\text{du}}(\mu)(\gamma_{\partial_{\mu_i} l_\mu} + \gamma_{\partial_{\mu_i} a_\mu} \|u_{r,\mu}\|) + \Delta_{\text{pr}}(\mu) \Delta_{\text{du}}(\mu) \gamma_{\partial_{\mu_i} a_\mu} + (\Delta_{\text{pr}})^2(\mu) \gamma_{\partial_{\mu_i} k_\mu} \\
&\quad + (\gamma_{\partial_{\mu_i} l_\mu} + \gamma_{\partial_{\mu_i} a_\mu} \|u_{r,\mu}\|) \underline{a}_\mu^{-1} (\|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| + 2\gamma_{k_\mu} \underline{a}_\mu^{-1} \|r_\mu^{\text{pr}}(u_{r,\mu})\|) \\
&\quad + \underline{a}_\mu^{-1} \|r_\mu^{\text{pr}}(u_{r,\mu})\| (\gamma_{\partial_{\mu_i} j} + 2\gamma_{\partial_{\mu_i} k} \|u_{r,\mu}\| + \gamma_{\partial_{\mu_i} a} \|p_{r,\mu}\|).
\end{aligned}$$

*Proof.* (i) For  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)$ , we have

$$(\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \tilde{\nabla}_\mu \hat{\mathcal{J}}_r(\mu))_i = \partial_{\mu_i} \mathcal{J}(u_{h,\mu}, \mu) - \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}].$$

Regarding the first contribution, we obtain with  $\|u_{h,\mu}\|_h \leq \|e_{h,\mu}^{\text{pr}}\|_h + \|u_{r,\mu}\|_h$

$$\begin{aligned}
|\partial_{\mu_i} \mathcal{J}(u_{h,\mu}, \mu) - \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu)| &= \left| \partial_{\mu_i} j_{h,\mu} \left( e_{h,\mu}^{\text{pr}} \right) + \partial_{\mu_i} k_\mu \left( e_{h,\mu}^{\text{pr}}, u_{r,\mu} \right) + \partial_{\mu_i} k_\mu \left( u_{h,\mu}, e_{h,\mu}^{\text{pr}} \right) \right| \\
&\leq \Delta_{\text{pr}}(\mu) (\gamma_{\partial_{\mu_i} j_\mu} + \gamma_{\partial_{\mu_i} k_\mu} (2\|u_{r,\mu}\| + \Delta_{\text{pr}}(\mu))).
\end{aligned}$$

For the other contributions we refer to Theorem 5 of [50].

(ii) For the adjoint estimator  $\Delta_{\nabla_\mu \hat{\mathcal{J}}_r}^*$ , we have

$$\begin{aligned}
(\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu))_i &= \partial_{\mu_i} \mathcal{J}(u_{h,\mu}, \mu) - \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] \\
&\quad - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [w_{r,\mu}] + \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [z_{r,\mu}].
\end{aligned}$$

The first line is equal to the estimator  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)$ , the first term of the second line can be estimated by

$$\partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [w_{r,\mu}] \leq \gamma_{\partial_{\mu_i} l_\mu} \|w_{r,\mu}\| + \gamma_{\partial_{\mu_i} a_\mu} \|u_{r,\mu}\| \|w_{r,\mu}\|.$$

The second term can analogously be estimated by

$$\partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [z_{r,\mu}] \leq \gamma_{\partial_{\mu_i} j} \|z_{r,\mu}\| + 2\gamma_{\partial_{\mu_i} k} \|z_{r,\mu}\| \|u_{r,\mu}\| + \gamma_{\partial_{\mu_i} a} \|z_{r,\mu}\| \|p_{r,\mu}\|.$$

We also have

$$\begin{aligned}
\underline{a}_\mu \|w_{r,\mu}\|^2 &\leq a_\mu(w_{r,\mu}, w_{r,\mu}) = r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [w_{r,\mu}] - 2k_\mu(z_{r,\mu}, w_{r,\mu}) \\
&\leq \|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| \|w_{r,\mu}\| + 2\gamma_{k_\mu} \|z_{r,\mu}\| \|w_{r,\mu}\|
\end{aligned}$$

which gives

$$\|w_{r,\mu}\| \leq \underline{a}_\mu^{-1} (\|r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})\| + 2\gamma_{k_\mu} \|z_{r,\mu}\|).$$

For  $z_\mu$  we estimate

$$\underline{a}_\mu \|z_{r,\mu}\|^2 \leq a_\mu(z_{r,\mu}, z_{r,\mu}) = -r_\mu^{\text{pr}}(u_{r,\mu}) [z_\mu] \leq \|r_\mu^{\text{pr}}(u_{r,\mu})\| \|z_{r,\mu}\|.$$

Summing all together gives the assertion.  $\square$

In a view of Section 3.3, we emphasize that the estimator for the NCD-corrected gradient does not show a better approximation of the FOM gradient since more terms are added to the standard estimate. Proposition 3.6(ii) suggests that there exist an estimator of higher order which we derive in the following section.

### 3.4.2. Sensitivity based approximation and estimation

We elaborate a better estimator for the NCD-corrected gradient by using sensitivities of the reduced primal and dual solutions. In addition, approximated sensitivities that are computed from the FOM sensitivities suggest an even better approximation of the FOM gradient.

We define the derivatives of the primal and dual solution maps associated with (3.4) in direction  $\nu \in \mathbb{R}^P$  as the solutions  $d_\nu u_{r,\mu} \in V_r^{\text{pr}}$  and  $d_\nu p_{r,\mu} \in V_r^{\text{du}}$  of

$$a_\mu(d_\nu u_{r,\mu}, v_r) = \partial_\mu r_\mu^{\text{pr}}(u_{r,\mu})[v_r] \cdot \nu \quad \text{for all } v_r \in V_r^{\text{pr}} \text{ and} \quad (3.12)$$

$$\begin{aligned} a_\mu(q_r, d_\nu p_{r,\mu}) &= -\partial_\mu a_\mu(q_r, p_{r,\mu}) \cdot \nu + d_\mu \partial_u \mathcal{J}(u_{r,\mu}, \mu)[q_r] \cdot \nu \\ &= \partial_\mu r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})[q_r] \cdot \nu + 2k_\mu(q_r, d_\nu u_{r,\mu}) \end{aligned} \quad \text{for all } q_r \in V_r^{\text{du}}, \quad (3.13)$$

respectively, analogously to Propositions 2.6 and 2.10, where the last equality holds for quadratic functionals as in (P.a). With these sensitivities we can compute the same gradient of the NCD-corrected RB reduced functional from Proposition 3.3 in a different manner.

**Proposition 3.9** (Gradient of the NCD-corrected RB reduced functional – Sensitivity approach). *The  $i$ th component of the true gradient of  $\hat{\mathcal{J}}_r$ ,  $\nabla_\mu \hat{\mathcal{J}}_r : \mathcal{P} \rightarrow \mathbb{R}^P$ , is given by*

$$(\nabla_\mu \hat{\mathcal{J}}_r(\mu))_i = \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu})[p_{r,\mu}] + r_\mu^{\text{pr}}(u_{r,\mu})[d_{\mu_i} p_{r,\mu}] + r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})[d_{\mu_i} u_{r,\mu}]$$

for all  $1 \leq i \leq P$  and  $\mu \in \mathcal{P}$ , where  $u_{r,\mu} \in V_r^{\text{pr}}$  and  $p_{r,\mu} \in V_r^{\text{du}}$  solve (3.4),  $d_{\mu_i} u_{r,\mu} \in V_r^{\text{pr}}$  and  $d_{\mu_i} p_{r,\mu} \in V_r^{\text{du}}$  denote the derivatives of RB primal and dual solution maps from (3.12) and (3.13).

*Proof.* It follows from the chain rule and Remark 2.4; see Section 1.6.1 of [35].  $\square$

Note that the sensitivity based gradient is mathematically equivalent to the one in Proposition 3.3, but the second only requires to solve (3.10) and (3.11) once, because they can be reused for every component  $\mu_i$ , whereas the computation of the gradient in Proposition 3.9 requires to solve (3.12) and (3.13) for each  $1 \leq i \leq P$ ; cf. [35].

In terms of numerical approximation w.r.t. the FOM functional, we note that, e.g., a solution  $d_{\mu_i} u_{r,\mu} \in V_r^{\text{pr}}$  of (3.12) does not necessarily need to be a good approximation of the FOM version  $d_{\mu_i} u_{h,\mu} \in V_h$  even though  $u_{h,\mu}$  is contained in  $V_r^{\text{pr}}$  since the high dimensional sensitivities are not generally contained in the respective reduced space (cf. Prop. 3.12).

To remedy this we could compute the FOM sensitivities for all canonical directions and either include them in the respective primal and dual space (thus forming Taylor RB spaces) or distribute all directional sensitivities to problem adapted RB spaces for the primal and dual sensitivities w.r.t. all canonical directions:  $V_r^{\text{pr}, d_{\mu_i}}, V_r^{\text{du}, d_{\mu_i}} \subset V_h$ . Thus, we again commit a variational crime.

**Definition 3.10** (Approximate partial derivatives of the RB primal and dual solution maps). Considering the reduced primal and dual solution maps  $\mathcal{P} \rightarrow V_r^{\text{pr}}$ ,  $\mu \mapsto u_{r,\mu}$  and  $\mathcal{P} \rightarrow V_r^{\text{du}}$ ,  $\mu \mapsto p_{r,\mu}$ , respectively, where  $u_{r,\mu}$  and  $p_{r,\mu}$  are the solutions of (3.4a) and (3.4b), we define their approximate partial derivatives w.r.t. the  $i$ th component of  $\mu$  by  $\tilde{d}_{\mu_i} u_{r,\mu} \in V_r^{\text{pr}, d_{\mu_i}}$  and  $\tilde{d}_{\mu_i} p_{r,\mu} \in V_r^{\text{du}, d_{\mu_i}}$ , respectively, as solutions of the sensitivity equations

$$a_\mu(\tilde{d}_{\mu_i} u_{r,\mu}, v_r) = \partial_\mu r_\mu^{\text{pr}}(u_{r,\mu})[v_r] \cdot e_i \quad \text{for all } v_r \in V_r^{\text{pr}, d_{\mu_i}}, \quad (3.14)$$

$$a_\mu(q_r, \tilde{d}_{\mu_i} p_{r,\mu}) = \partial_\mu r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu})[q_r] \cdot e_i + 2k_\mu(q_r, \tilde{d}_{\mu_i} u_{r,\mu}) \quad \text{for all } q_r \in V_r^{\text{du}, d_{\mu_i}}. \quad (3.15)$$

Similarly, we denote the approximate partial derivatives in direction  $\nu \in \mathbb{R}^P$  by  $\tilde{d}_\nu u_{r,\mu}$  and  $\tilde{d}_\nu p_{r,\mu}$ , respectively, defined by substituting  $e_i$  with  $\nu$  above.

Following Propositions 2.6 and 2.10 we would obtain  $\tilde{d}_{\mu_i} u_{r,\mu} = d_{\mu_i} u_{r,\mu}$ , if  $V_r^{\text{pr}, d_{\mu_i}} = V_r^{\text{pr}}$  and  $\tilde{d}_{\mu_i} p_{r,\mu} = d_{\mu_i} p_{r,\mu}$ , if  $V_r^{\text{du}, d_{\mu_i}} = V_r^{\text{du}}$ . Moreover, the approximate partial derivatives depend on the choice of the corresponding reduced approximation spaces.

**Definition 3.11** (Approximate gradient of the NCD-corrected RB reduced functional). We define the approximate gradient  $\tilde{\nabla}_\mu \hat{\mathcal{J}}_r : \mathcal{P} \rightarrow \mathbb{R}^P$  of  $\hat{\mathcal{J}}_r$  by

$$\left(\tilde{\nabla}_\mu \hat{\mathcal{J}}_r(\mu)\right)_i := \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] + r_\mu^{\text{pr}}(u_{r,\mu}) [\tilde{d}_{\mu_i} p_{r,\mu}] + r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [\tilde{d}_{\mu_i} u_{r,\mu}], \quad (3.16)$$

for  $1 \leq i \leq P$ , where  $u_{r,\mu} \in V_r^{\text{pr}}$ ,  $p_{r,\mu} \in V_r^{\text{du}}$  denote the reduced primal and dual solutions and  $\tilde{d}_{\mu_i} u_{r,\mu} \in V_r^{\text{pr}, d_{\mu_i}}$  and  $\tilde{d}_{\mu_i} p_{r,\mu} \in V_r^{\text{du}, d_{\mu_i}}$  denote the solutions of (3.14) and (3.15).

Both gradients from Definition 3.9 and Proposition 3.11 yield higher order estimate. To show this, we first derive error estimates for the reduction error of the reduced sensitivities from (3.12) and (3.13) as well as for (3.14) and (3.15). For  $v_h \in V_h$ , the residuals of the equation in Propositions 2.6 and 2.10 for the canonical directions are respectively given by

$$r_\mu^{\text{pr}, d_{\mu_i}}(u_{h,\mu}, d_{\mu_i} u_{h,\mu}) [v_h] := \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [v_h] - a_\mu(d_{\mu_i} u_{h,\mu}, v_h), \quad (3.17)$$

$$r_\mu^{\text{du}, d_{\mu_i}}(u_{h,\mu}, p_{h,\mu}, d_{\mu_i} u_{h,\mu}, d_{\mu_i} p_{h,\mu}) [v_h] := \partial_{\mu_i} r_\mu^{\text{du}}(u_{h,\mu}, p_{h,\mu}) [v_h] + 2k_\mu(v_h, d_{\mu_i} u_{h,\mu}) - a_\mu(v_h, d_{\mu_i} p_{h,\mu}). \quad (3.18)$$

**Proposition 3.12** (Residual based upper bound on the model reduction error of the sensitivity of the primal solution map). For  $\mu \in \mathcal{P}$  and  $1 \leq i \leq P$ , let  $d_{\mu_i} u_{h,\mu} \in V_h$  be the solution of the discrete version of (2.2) and  $d_{\mu_i} u_{r,\mu} \in V_r^{\text{pr}, d_{\mu_i}}$  be the solution of (3.12). We then have

$$\|d_{\mu_i} u_{h,\mu} - d_{\mu_i} u_{r,\mu}\| \leq \Delta_{d_{\mu_i} \text{pr}}(\mu) := \underline{a}_\mu^{-1} \left( \gamma_{\partial_{\mu_i} a_\mu} \Delta_{\text{pr}}(\mu) + \|r_\mu^{\text{pr}, d_{\mu_i}}(u_{r,\mu}, d_{\mu_i} u_{r,\mu})\| \right).$$

*Proof.* Using the shorthand  $d_{\mu_i} e_{h,\mu}^{\text{pr}} := d_{\mu_i} u_{h,\mu} - d_{\mu_i} u_{r,\mu}$ , we obtain

$$\begin{aligned} \underline{a}_\mu \|d_{\mu_i} e_{h,\mu}^{\text{pr}}\|^2 &\leq a_\mu \left( d_{\mu_i} e_{h,\mu}^{\text{pr}}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) = a_\mu \left( d_{\mu_i} u_{h,\mu}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) - a_\mu \left( d_{\mu_i} u_{r,\mu}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) \\ &= \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) \left[ d_{\mu_i} e_{h,\mu}^{\text{pr}} \right] - a_\mu \left( d_{\mu_i} u_{r,\mu}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) \\ &= \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) \left[ d_{\mu_i} e_{h,\mu}^{\text{pr}} \right] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) \left[ d_{\mu_i} e_{h,\mu}^{\text{pr}} \right] + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) \left[ d_{\mu_i} e_{h,\mu}^{\text{pr}} \right] - a_\mu \left( d_{\mu_i} u_{r,\mu}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) \\ &= -\partial_{\mu_i} a_\mu \left( e_{h,\mu}^{\text{pr}}, d_{\mu_i} e_{h,\mu}^{\text{pr}} \right) + r_\mu^{\text{pr}, d_{\mu_i}}(u_{r,\mu}, d_{\mu_i} u_{r,\mu}) \left[ d_{\mu_i} e_{h,\mu}^{\text{pr}} \right] \\ &\leq \gamma_{\partial_{\mu_i} a_\mu} \|e_{h,\mu}^{\text{pr}}\| \|d_{\mu_i} e_{h,\mu}^{\text{pr}}\| + \|r_\mu^{\text{pr}, d_{\mu_i}}(u_{r,\mu}, d_{\mu_i} u_{r,\mu})\| \|d_{\mu_i} e_{h,\mu}^{\text{pr}}\| \end{aligned}$$

using the coercivity of  $a_\mu$  in the first inequality, the definition  $d_{\mu_i} e_{h,\mu}^{\text{pr}}$  in the first equality, Proposition 2.6 applied to  $u_{h,\mu}$  in the second equality, the definition of the discrete sensitivity primal residual (3.17) in the third equality and the continuity of  $d_{\mu_i} a_\mu$  in the last inequality.  $\square$

We emphasize that the same result can be shown for  $\tilde{d}_{\mu_i} u_{r,\mu}$  by replacing  $d_{\mu_i} u_{r,\mu}$  and using the equation (3.14) instead of (3.12). We call the resulting error estimator  $\Delta_{\tilde{d}_{\mu_i} \text{pr}}(\mu)$ .

**Proposition 3.13** (Residual based upper bound on the model reduction error of the sensitivity of the dual solution map). For  $\mu \in \mathcal{P}$  and  $1 \leq i \leq P$ , let  $d_{\mu_i} p_{h,\mu} \in V_h$  be the solution of the discrete version of (2.6) and  $d_{\mu_i} p_{r,\mu} \in V_r^{\text{pr}, d_{\mu_i}}$  be the solution of (3.13). We then obtain

$$\begin{aligned} \|d_{\mu_i} p_{h,\mu} - d_{\mu_i} p_{r,\mu}\| &\leq \Delta_{d_{\mu_i} \text{du}}(\mu) \quad \text{with} \\ \Delta_{d_{\mu_i} \text{du}}(\mu) &:= \underline{a}_\mu^{-1} \left( 2\gamma_{\partial_{\mu_i} k_\mu} \Delta_{\text{pr}}(\mu) + \gamma_{\partial_{\mu_i} a_\mu} \Delta_{\text{du}}(\mu) + 2\gamma_{k_\mu} \Delta_{d_{\mu_i} \text{pr}}(\mu) \right. \\ &\quad \left. + \|r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu})\| \right). \end{aligned}$$

*Proof.* Using the shorthand  $d_{\mu_i} e_{h,\mu}^{\text{du}} := d_{\mu_i} p_{h,\mu} - d_{\mu_i} p_{r,\mu}$  and  $e_{h,\mu}^{\text{du}} := p_{h,\mu} - p_{r,\mu}$ , we obtain

$$\begin{aligned}
\underline{a}_\mu \|d_{\mu_i} e_{h,\mu}^{\text{du}}\|^2 &\leq a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} e_{h,\mu}^{\text{du}}) = \underbrace{a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} p_{h,\mu})}_{= \partial_{\mu_i} r_\mu^{\text{du}}(u_{h,\mu}, p_{h,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] + 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{h,\mu})} - a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} p_{r,\mu}) \\
&= \partial_{\mu_i} r_\mu^{\text{du}}(u_{h,\mu}, p_{h,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] + 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{h,\mu}) - \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] \\
&\quad - 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{r,\mu}) + \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] + 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{r,\mu}) - a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} p_{r,\mu}) \\
&= \partial_{\mu_i} j_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}) + 2\partial_{\mu_i} k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, u_{h,\mu}) - \partial_{\mu_i} a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, p_h) - \partial_{\mu_i} j_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}) + 2\partial_{\mu_i} k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, u_{r,\mu}) \\
&\quad - \partial_{\mu_i} a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, p_r) + 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{h,\mu}) - 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} u_{r,\mu}) \\
&\quad + r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] \\
&= 2\partial_{\mu_i} k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, e_{h,\mu}^{\text{pr}}) - \partial_{\mu_i} a_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, e_{h,\mu}^{\text{du}}) + 2k_\mu (d_{\mu_i} e_{h,\mu}^{\text{du}}, d_{\mu_i} e_{h,\mu}^{\text{pr}}) \\
&\quad + r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) [d_{\mu_i} e_{h,\mu}^{\text{du}}] \\
&\leq \left( 2\gamma_{\partial_{\mu_i} k_\mu} \|e_{h,\mu}^{\text{pr}}\| + \gamma_{\partial_{\mu_i} a_\mu} \|e_{h,\mu}^{\text{du}}\| \right) \|d_{\mu_i} e_{h,\mu}^{\text{du}}\| \\
&\quad + 2\gamma_{k_\mu} \|d_{\mu_i} e_{h,\mu}^{\text{pr}}\| \|d_{\mu_i} e_{h,\mu}^{\text{du}}\| + \|r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu})\| \|d_{\mu_i} e_{h,\mu}^{\text{du}}\|
\end{aligned}$$

using the coercivity of  $a_\mu$  in the first inequality, the definition of  $d_{\mu_i} e_{h,\mu}^{\text{du}}$  in the first equality, Proposition 2.10 applied to  $p_{h,\mu}$  in the second equality, the definition of the dual residual in (2.5) in the third equality and continuity of all parts in the last inequality.  $\square$

Again, the same result holds for  $\tilde{d}_{\mu_i} p_{r,\mu}$  if we replace  $d_{\mu_i} p_{r,\mu}$  and use (3.15) instead of (3.13). The resulting error estimator is then called  $\Delta_{\tilde{d}_{\mu_i} \text{du}}(\mu)$ .

Using the residual based *a posteriori* error estimates for the primal sensitivities, we are able to state two *a posteriori* error bounds on the model reduction error of the true gradient and the approximated gradient of the NCD-corrected functional.

**Proposition 3.14** (Upper bound on the model reduction error of the gradient of the reduced output – sensitivity approach).

- (i) For the gradient  $\nabla_\mu \hat{\mathcal{J}}_r(\mu)$  of the NCD-corrected RB reduced functional, computed with sensitivities according to Proposition 3.9, we have

$$\begin{aligned}
\|\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu)\|_2 &\leq \Delta_{\nabla \hat{\mathcal{J}}_r}(\mu) = \|\underline{\Delta}_{\nabla \hat{\mathcal{J}}_r}(\mu)\|_2 \quad \text{with} \\
(\underline{\Delta}_{\nabla \hat{\mathcal{J}}_r}(\mu))_i &:= \gamma_{\partial_{\mu_i} k_\mu} (\Delta_{\text{pr}}(\mu))^2 + \gamma_{a_\mu} \Delta_{d_{\mu_i} \text{pr}}(\mu) \Delta_{\text{du}}(\mu) \\
&\quad + \|r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu})\| \Delta_{\text{pr}}(\mu).
\end{aligned}$$

- (ii) Furthermore, we have for the approximate gradient from Definition 3.11

$$\begin{aligned}
\|\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \tilde{\nabla}_\mu \hat{\mathcal{J}}_r(\mu)\|_2 &\leq \Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu) = \|\underline{\Delta}_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)\|_2 \quad \text{with} \\
(\underline{\Delta}_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu))_i &:= \gamma_{\partial_{\mu_i} k_\mu} (\Delta_{\text{pr}}(\mu))^2 + \gamma_{a_\mu} \Delta_{\tilde{d}_{\mu_i} \text{pr}}(\mu) \Delta_{\text{du}}(\mu) \\
&\quad + \|r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, \tilde{d}_{\mu_i} u_{r,\mu}, \tilde{d}_{\mu_i} p_{r,\mu})\| \Delta_{\text{pr}}(\mu).
\end{aligned}$$

*Proof.* (i) To prove the first assertion, we use  $r_\mu^{\text{pr}}(u_{h,\mu}) [d_{\mu_i} p_{r,\mu}] = 0$  and  $r_\mu^{\text{du}}(u_{h,\mu}, p_{h,\mu}) [d_{\mu_i} u_{r,\mu}] = 0$  to obtain

$$(\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu))_i = \partial_{\mu_i} \mathcal{J}(u_{h,\mu}, \mu) - \partial_{\mu_i} \mathcal{J}(u_{r,\mu}, \mu) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}]$$

$$\begin{aligned}
 & -r_\mu^{\text{pr}}(u_{r,\mu}) [d_{\mu_i} p_{r,\mu}] - r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) [d_{\mu_i} u_{r,\mu}] \\
 = & \partial_{\mu_i} j_\mu \left( e_{h,\mu}^{\text{pr}} \right) + \partial_{\mu_i} k_\mu (u_{h,\mu}, u_{h,\mu}) - \partial_{\mu_i} k_\mu (u_{r,\mu}, u_{r,\mu}) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] \\
 & - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] + \underbrace{r_\mu^{\text{pr}} \left( e_{h,\mu}^{\text{pr}} \right) [d_{\mu_i} p_{r,\mu}]}_{= (*)} + \underbrace{r_\mu^{\text{du}} \left( e_{h,\mu}^{\text{pr}}, e_{h,\mu}^{\text{du}} \right) [d_{\mu_i} u_{r,\mu}]}_{= (**)}.
 \end{aligned}$$

For the last two residual terms we have

$$\begin{aligned}
 (*) & = l_\mu (d_{\mu_i} p_{r,\mu}) - l_\mu (d_{\mu_i} p_{r,\mu}) - a_\mu \left( e_{h,\mu}^{\text{pr}}, d_{\mu_i} p_{r,\mu} \right) \\
 & = -a_\mu \left( e_{h,\mu}^{\text{pr}}, d_{\mu_i} p_{r,\mu} \right) + \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] + 2k_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{pr}} \right) \\
 & \quad - \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] - 2k_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{pr}} \right) \\
 & = r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] - \partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] - 2k_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{pr}} \right)
 \end{aligned}$$

and

$$(***) = j_\mu (d_{\mu_i} u_{r,\mu}) - j_\mu (d_{\mu_i} u_{r,\mu}) + 2k_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{pr}} \right) - a_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{du}} \right).$$

Thus, by summing both terms we have

$$(*) + (***) = r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] - \underbrace{\partial_{\mu_i} r_\mu^{\text{du}}(u_{r,\mu}, p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right]}_{= (***)} - a_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{du}} \right)$$

and for (\*\*\*) it holds

$$(***) = \partial_{\mu_i} j_\mu \left( e_{h,\mu}^{\text{pr}} \right) + 2\partial_{\mu_i} k_\mu \left( e_{h,\mu}^{\text{pr}}, u_{r,\mu} \right) - \partial_{\mu_i} a_\mu \left( e_{h,\mu}^{\text{pr}}, p_{r,\mu} \right).$$

Combining (\*), (\*\*\*) and (\*\*\*) with the previous result, we have

$$\begin{aligned}
 \left( \nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu) \right)_i & = \partial_{\mu_i} j_\mu \left( e_{h,\mu}^{\text{pr}} \right) + \partial_{\mu_i} k_\mu (u_{h,\mu}, u_{h,\mu}) - \partial_{\mu_i} k_\mu (u_{r,\mu}, u_{r,\mu}) \\
 & \quad - \partial_{\mu_i} j_\mu \left( e_{h,\mu}^{\text{pr}} \right) - 2\partial_{\mu_i} k_\mu (e_{h,\mu}^{\text{pr}}, u_{r,\mu}) + \partial_{\mu_i} a_\mu (e_{h,\mu}^{\text{pr}}, p_{r,\mu}) + \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] \\
 & \quad - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] + r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] - a_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{du}} \right) \\
 = & \partial_{\mu_i} k_\mu \left( e_{h,\mu}^{\text{pr}}, e_{h,\mu}^{\text{pr}} \right) + r_\mu^{\text{du}, d_{\mu_i}}(u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right] \\
 & \quad + \underbrace{\partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] + \partial_{\mu_i} a_\mu \left( e_{h,\mu}^{\text{pr}}, p_{r,\mu} \right) - a_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{du}} \right)}_{= (****)}.
 \end{aligned}$$

Further, we have

$$\begin{aligned}
 \partial_{\mu_i} r_\mu^{\text{pr}}(u_{h,\mu}) [p_{h,\mu}] - \partial_{\mu_i} r_\mu^{\text{pr}}(u_{r,\mu}) [p_{r,\mu}] & = \partial_{\mu_i} l_\mu \left( e_{h,\mu}^{\text{du}} \right) - \partial_{\mu_i} a_\mu (u_{h,\mu}, p_{h,\mu}) + \partial_{\mu_i} a_\mu (u_{r,\mu}, p_{r,\mu}) \\
 & = a_\mu \left( d_{\mu_i} u_h, e_{h,\mu}^{\text{du}} \right) + \partial_{\mu_i} a_\mu (u_h, e_{h,\mu}^{\text{du}}) - \partial_{\mu_i} a_\mu (u_{h,\mu}, p_{h,\mu}) + \partial_{\mu_i} a_\mu (u_{r,\mu}, p_{r,\mu}),
 \end{aligned}$$

where we used the discretized version of (2.2) in the second equality. Inserting this into (\*) gives

$$(****) = a_\mu \left( d_{\mu_i} u_h, e_{h,\mu}^{\text{du}} \right) - a_\mu \left( d_{\mu_i} u_{r,\mu}, e_{h,\mu}^{\text{du}} \right)$$



$$\begin{aligned}
 & + \underbrace{\partial_{\mu_i} a_\mu (u_h, e_{h,\mu}^{\text{du}}) - \partial_{\mu_i} a_\mu (u_{h,\mu}, p_{h,\mu}) + \partial_{\mu_i} a_\mu (u_{r,\mu}, p_{r,\mu}) + \partial_{\mu_i} a_\mu (e_{h,\mu}^{\text{pr}}, p_{r,\mu})}_{=0} \\
 & = a_\mu \left( d_{\mu_i} e_{h,\mu}^{\text{pr}}, e_{h,\mu}^{\text{du}} \right).
 \end{aligned}$$

In total, we have

$$\begin{aligned}
 \left( \nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_r(\mu) \right)_i & = \partial_{\mu_i} k_\mu \left( e_{h,\mu}^{\text{pr}}, e_{h,\mu}^{\text{pr}} \right) + a_\mu \left( d_{\mu_i} e_{h,\mu}^{\text{pr}}, e_{h,\mu}^{\text{du}} \right) \\
 & \quad + r_\mu^{\text{du}, d_{\mu_i}} (u_{r,\mu}, p_{r,\mu}, d_{\mu_i} u_{r,\mu}, d_{\mu_i} p_{r,\mu}) \left[ e_{h,\mu}^{\text{pr}} \right]
 \end{aligned}$$

which proves the assertion.

- (ii) The estimate follows analogously to (i), by replacing  $d_{\mu_i} u_{r,\mu}$  and  $d_{\mu_i} p_{r,\mu}$  with  $\tilde{d}_{\mu_i} u_{r,\mu}$  and  $\tilde{d}_{\mu_i} p_{r,\mu}$ , respectively.

□

To conclude,  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)$  and  $\Delta_{\nabla \hat{\mathcal{J}}_r}(\mu)$  both decay with second order (*cf.* Sect. 5.3.1). We also point out, that  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)$  is an improved estimator which can be used to replace the poor estimator  $\Delta_{\nabla \mu}^{r,*}(\mu)$ . However both higher order estimators  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu)$  and  $\Delta_{\nabla \hat{\mathcal{J}}_r}(\mu)$  come with the price of computing the dual norm of the sensitivity residuals in (3.17) and (3.18) for each direction which aggravates the computational complexity.

#### 4. THE TRUST-REGION METHOD AND ADAPTIVE ENRICHMENT STRATEGIES

To solve problem (P) we apply the TR method, which iteratively computes a first-order critical point of (P). At each iteration  $k \geq 0$ , we consider a so-called model function  $m^{(k)}$ , which is a cheaply computable approximation of the quadratic cost functional  $\mathcal{J}$  in a neighborhood of the parameter  $\mu^{(k)}$ , *i.e.*, the Trust-Region. Therefore, for  $k \geq 0$ , given a TR radius  $\delta^{(k)}$ , we consider the TR minimization sub-problem

$$\min_{s \in \mathbb{R}^P} m^{(k)}(s) \text{ subject to } \|s\|_2 \leq \delta^{(k)}, \tilde{\mu} := \mu^{(k)} + s \in \mathcal{P} \text{ and } r_{\tilde{\mu}}^{\text{pr}}(u_{\tilde{\mu}})[v] = 0 \text{ for all } v \in V. \tag{4.1}$$

Under suitable assumptions on  $m^{(k)}$ , problem (4.1) admits a unique solution  $\bar{s}^{(k)}$ , which is used to compute the next iterate  $\mu^{(k+1)} = \mu^{(k)} + \bar{s}^{(k)}$ .

##### 4.1. The Trust-Region Reduced Basis method

Slightly different from [1, 50], we choose as model function the NCD-corrected RB reduced functional  $\hat{\mathcal{J}}_r^{(k)}$  defined in (3.9), *i.e.*,  $m^{(k)}(\cdot) = \hat{\mathcal{J}}_r^{(k)}(\mu^{(k)} + \cdot)$  for  $k \geq 0$ , where the super-index  $(k)$  indicates that we use different RB spaces  $V_r^{*,(k)}$  in each iteration. As indicated in Proposition 3.6 and shown in our numerical experiments below,  $\hat{\mathcal{J}}_r^{(k)}$  converges to  $\hat{\mathcal{J}}$  with higher order in comparison to the standard RB reduced functional from (3.5), which has been considered in [50]. We initialize the RB spaces using the initial guess  $\mu^{(0)}$ , *i.e.*, setting  $V_r^{\text{pr},0} = \{u_{h,\mu^{(0)}}\}$  and  $V_r^{\text{du},0} = \{p_{h,\mu^{(0)}}\}$ . At every iteration  $k$  we may – depending on the *a posteriori* estimates – enrich the obtained space using the computed parameter  $\mu^{(k+1)}$ ; for further details see Section 4.3. Possible sufficient and necessary conditions for convergence, dependent on the approximate generalized Cauchy point (AGC)  $\mu_{\text{AGC}}^{(k)}$  (see Definition 4.1), are given in [50]. In contrast to [50], we consider additional bilateral parameter constraints in (4.1). In particular, the presence of these inequality constraints requires a review of the proof of convergence for the TR-RB algorithm. In [50], the convergence is based on the results contained in [57], where the authors consider an equality-constrained optimization problem. We state first how our method differs from the one in [50], then we prove the convergence of this modified algorithm. According to [50], the inexact RB version of problem (4.1) is

$$\min_{\tilde{\mu} \in \mathcal{P}} \hat{\mathcal{J}}_r^{(k)}(\tilde{\mu}) \text{ s.t. } \frac{\Delta_{\hat{\mathcal{J}}_r}(\tilde{\mu})}{\hat{\mathcal{J}}_r^{(k)}(\tilde{\mu})} \leq \delta^{(k)}, \tag{4.2}$$

where  $\tilde{\mu} := \mu^{(k)} + s$ , the equality constraint  $r_{\tilde{\mu}}^{\text{pr}}(u_{\tilde{\mu}})[v] = 0$  is hidden in the definition of  $\hat{\mathcal{J}}_r$  and the inequality ones are concealed in the request  $\tilde{\mu} \in \mathcal{P}$ . As also remarked in [50], the projected BFGS method [40], which we use in order to solve (4.2), computes the AGC point  $\mu_{\text{AGC}}^{(k)}$  in the first iterate and generates a sequence  $\{\mu^{(k,\ell)}\}_{\ell=1}^L$  where  $L$  is the last BFGS iteration. In what follows,  $\mu^{(k,1)} := \mu_{\text{AGC}}^{(k)}$  and the TR iterate  $\mu^{(k+1)} := \mu^{(k,L)}$ . Throughout the paper the index  $k$  refers to the current outer TR iteration,  $\ell$  refers instead to the inner BFGS iteration. Note that  $L$  may be different for each iteration  $k$ , but we will indicate it only when strictly necessary in order to simplify the notation. To describe the projected BFGS method in details, we define

$$\mu^{(k,\ell)}(j) := \text{P}_{\mathcal{P}}(\mu^{(k,\ell)} + \kappa^j d^{(k,\ell)}) \in \mathcal{P} \quad \text{for } j \geq 0, \quad (4.3)$$

where  $\kappa \in (0, 1)$ ,  $d^{(k,\ell)} \in \mathbb{R}^P$  is the chosen descent direction at the iteration  $(k, \ell)$  and the projection operator  $\text{P}_{\mathcal{P}} : \mathbb{R}^P \rightarrow \mathcal{P}$  is defined as

$$(\text{P}_{\mathcal{P}}(\mu))_i := \begin{cases} (\mu_{\mathbf{a}})_i & \text{if } (\mu)_i \leq (\mu_{\mathbf{a}})_i, \\ (\mu_{\mathbf{b}})_i & \text{if } (\mu)_i \geq (\mu_{\mathbf{b}})_i, \\ (\mu)_i & \text{otherwise,} \end{cases} \quad \text{for } i = 1, \dots, P.$$

Note that the operator  $\text{P}_{\mathcal{P}}$  is Lipschitz continuous with constant one; *cf.* [40]. For computing the descent direction  $d^{(k,\ell)}$  we follow the projected BFGS algorithm reported in Section 5.5.3 of [40]. Furthermore, we enforce respectively an Armijo-type condition and the additional TR constraint on  $\hat{\mathcal{J}}_r^{(k)}$

$$\hat{\mathcal{J}}_r^{(k)}(\mu^{(k,\ell)}(j)) - \hat{\mathcal{J}}_r^{(k)}(\mu^{(k,\ell)}) \leq -\frac{\kappa_{\text{arm}}}{\kappa^j} \|\mu^{(k,\ell)}(j) - \mu^{(k,\ell)}\|_2^2, \quad (4.4)$$

$$\frac{\Delta_{\hat{\mathcal{J}}_r}(\mu^{(k,\ell)}(j))}{\hat{\mathcal{J}}_r^{(k)}(\mu^{(k,\ell)}(j))} \leq \delta^{(k)}, \quad (4.5)$$

by selecting  $\mu^{(k,\ell+1)} = \mu^{(k,\ell)}(j^{(k,\ell)})$  for  $\ell \geq 1$ , where  $j^{(k,\ell)} < \infty$  is the smallest index for which the conditions (4.4) and (4.5) hold for some  $\kappa_{\text{arm}} \in (0, \frac{1}{2})$ , generally  $\kappa_{\text{arm}} = 10^{-4}$ ; *cf.* [50]. Moreover, we use as termination criteria for the optimization sub-problem

$$\|\mu^{(k,\ell)} - \text{P}_{\mathcal{P}}(\mu^{(k,\ell)} - \nabla_{\mu} \hat{\mathcal{J}}_r^{(k)}(\mu^{(k,\ell)}))\|_2 \leq \tau_{\text{sub}} \quad (4.6a) \quad \text{or} \quad \beta_2 \delta^{(k)} \leq \frac{\Delta_{\hat{\mathcal{J}}_r}(\mu)}{\hat{\mathcal{J}}_r^{(k)}(\mu)} \leq \delta^{(k)}, \quad (4.6b)$$

where  $\tau_{\text{sub}} \in (0, 1)$  is a predefined tolerance and  $\beta_2 \in (0, 1)$ , generally close to one. Condition (4.6b) is used to prevent that the optimizer spends much time close to the boundary of the Trust-Region, where the model is poor in approximation; *cf.* [50]. Note that, without the projection operator  $\text{P}_{\mathcal{P}}$ , conditions (4.4)–(4.6) coincide with the ones in [50], apart from using the NCD-corrected RB reduced functional. Furthermore, in addition to [50], we consider a condition which allows enlarging the TR radius. A drawback of the TR algorithm proposed in [50] is that the TR radius may be significantly shrunk at the beginning, *i.e.*, when the TR model is poor in approximation. Afterwards, even if the RB space is enriched, *i.e.*, the approximation of the TR model function is improved, the TR radius is kept small. Thus, one misses the local second-order rate of convergence of the BFGS method. More precisely, if  $\mu^{(k,\ell)}$  is close to the locally optimal solution  $\bar{\mu}^{(k)}$  of the TR sub-problem, we want to make full BFGS steps, which gives us faster convergence. The possibility to enlarge the TR radius at each iteration will also decrease the number of outer iterations needed to converge. As a condition for enlarging the radius we check whether the sufficient reduction predicted by the model function  $\hat{\mathcal{J}}_r^{(k)}$  is realized by the objective function, *i.e.*, we check if

$$\varrho^{(k)} := \frac{\hat{\mathcal{J}}_h(\mu^{(k)}) - \hat{\mathcal{J}}_h(\mu^{(k+1)})}{\hat{\mathcal{J}}_r^{(k)}(\mu^{(k)}) - \hat{\mathcal{J}}_r^{(k)}(\mu^{(k+1)})} \geq \eta_e \quad (4.7)$$

for a tolerance  $\eta_\varrho \in [3/4, 1)$ . Condition (4.7) seems costly because of the evaluation of the FOM cost functional  $\hat{\mathcal{J}}_h$ , but, after the enrichment of the RB space, the quantities in the numerator of (4.7) are cheaply accessible, since one has already solved the FOM to generate the new snapshots for the RB space enrichment. Note that this also implies that we can cheaply evaluate the FOM gradient  $\nabla_\mu \hat{\mathcal{J}}_h(\mu^{k+1})$  in case of an enrichment. This knowledge will be used for the stopping criterium in the outer loop of the algorithm. Finally, let us define the AGC point for our constrained case.

**Definition 4.1** (AGC point for simple bounds). At iteration  $k$ , we define the AGC point as

$$\mu_{\text{AGC}}^{(k)} := \mu^{(k,0)} \left( j_c^{(k)} \right) = \text{P}_{\mathcal{P}} \left( \mu^{(k,0)} + \kappa^{j_c^{(k)}} d^{(k,0)} \right),$$

where  $\mu^{(k,0)} := \mu^{(k)}$ ,  $d^{(k,0)} := -\nabla_\mu \hat{\mathcal{J}}_r^{(k)}(\mu^{(k,0)})$  and  $j_c^{(k)}$  is the smallest non-negative integer  $j$  for which  $\mu^{(k,0)}(j)$  satisfies (4.4) and (4.5) for  $\ell = 0$ .

We refer to Algorithm 1 for the proposed TR-RB algorithm.

### 4.2. Convergence analysis

In order to guarantee the well-posedness (because of (4.5)) and the convergence of the method, we make the following assumption

**Assumption 4.2.** *The cost functional  $\mathcal{J}(u, \mu)$  is strictly positive for all  $u \in V$  and all parameters  $\mu \in \mathcal{P}$ .*

Note that this assumption is not too restrictive, since the boundedness from below is a standard assumption in optimization to guarantee the existence of a solution for the minimization problem. If a global lower bound for the cost functional is also known, one can add a sufficiently large constant, without changing the position of its local minima and maxima. Another important request, pointed out in [50, 57], is that an error-aware sufficient decrease condition

$$\hat{\mathcal{J}}_r^{(k+1)} \left( \mu^{(k+1)} \right) \leq \hat{\mathcal{J}}_r^{(k)} \left( \mu_{\text{AGC}}^{(k)} \right) \quad \text{for all } k \in \mathbb{N} \tag{4.8}$$

is fulfilled at each iteration  $k$  of the TR-RB algorithm. As in [50, 57], we consider cheaply computable sufficient and necessary conditions for (4.8) in Algorithm 1 (Steps 4 and 11, respectively). The TR-RB algorithm rejects, then, any computed point which does not satisfy (4.8). One may be concerned of the fact that Algorithm 1 may be trapped in an infinite loop where every computed point is rejected and the TR radius is shrunk all time. We point out that this never happened in our numerical tests. Anyway, we consider a safety termination criteria, which is triggered when the TR radius is smaller than the double machine precision. To prove convergence of Algorithm 1, in what follows, we then assume that this situation can not occur.

**Assumption 4.3.** *For each  $k \geq 0$ , there exists a radius  $\tilde{\delta}^{(k)} > \tau_{\text{mac}}$  for which a solution of (4.2) is such that (4.8) is verified, where  $\tau_{\text{mac}} = 2.22 \times 10^{-16}$  is the double machine precision.*

**Lemma 4.4.** *Let Assumptions 2.2–4.3 hold true. The search of the AGC point defined in Definition 4.1 takes finitely many iterations at each step  $k$  of the TR-RB Algorithm.*

*Proof.* We want to prove that there exists an index  $j_c^{(k)} < \infty$  for each  $k \geq 0$ , for which  $\mu_{\text{AGC}}^{(k)} = \mu^{(k,0)} \left( j_c^{(k)} \right)$  satisfies (4.4) and (4.5) for  $\ell = 0$ . From Theorem 5.4.5 of [40] (and the subsequent discussion) we conclude that for all  $k \in \mathbb{N}$  there exists a strictly positive index  $j_1^{(k)} \in \mathbb{N}$  such that  $\mu^{(k,0)}(j)$  satisfies (4.4) for  $j \geq j_1^{(k)}$  and  $\ell = 0$ . If  $k = 0$ , by construction we have that  $\Delta_{\hat{\mathcal{J}}_r^{(0)}}(\mu^{(0)}) = 0$ . Therefore, there exists a sufficiently large (but finite) index  $j_2^{(0)} \in \mathbb{N}$  such that  $\mu^{(0,0)}(j)$  satisfies (4.5) for all  $j \geq j_2^{(0)}$  and  $\ell = 0$ . The reason relies on the continuity w.r.t.  $\mu$  of the error estimator  $\Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu)$  (cf. Rem. 3.7) and of the cost functional  $\hat{\mathcal{J}}_r^{(k)}(\mu)$  for all

**Algorithm 1:** TR-RB algorithm.

---

**Data:** Initial TR radius  $\delta^{(0)}$ , TR shrinking factor  $\beta_1 \in (0, 1)$ , tolerance for enlarging the TR radius  $\eta_e \in [\frac{3}{4}, 1)$ , initial parameter  $\mu^{(0)}$ , stopping tolerance for the sub-problem  $\tau_{\text{sub}} \ll 1$ , stopping tolerance for the first-order critical condition  $\tau_{\text{FOC}}$  with  $\tau_{\text{sub}} \leq \tau_{\text{FOC}} \ll 1$ , safeguard for TR boundary  $\beta_2 \in (0, 1)$ .

- 1 Set  $k = 0$  and Loop\_flag=True;
- 2 while Loop\_flag do
- 3    Compute  $\mu^{(k+1)}$  as solution of (4.2) with termination criteria (4.6);
- 4    if  $\hat{\mathcal{J}}_r^{(k)}(\mu^{(k+1)}) + \Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu^{(k+1)}) < \hat{\mathcal{J}}_r^{(k)}(\mu_{\text{AGC}}^{(k)})$  then
- 5        Accept  $\mu^{(k+1)}$ , update the RB model at  $\mu^{(k+1)}$  and compute  $\varrho^{(k)}$  from (4.7);
- 6        if  $\varrho^{(k)} \geq \eta_e$  then
- 7            Enlarge the TR radius  $\delta^{(k+1)} = \beta_1^{-1}\delta^{(k)}$ ;
- 8        else
- 9            Set  $\delta^{(k+1)} = \delta^{(k)}$ ;
- 10       end
- 11    else if  $\hat{\mathcal{J}}_r^{(k)}(\mu^{(k+1)}) - \Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu^{(k+1)}) > \hat{\mathcal{J}}_r^{(k)}(\mu_{\text{AGC}}^{(k)})$  then
- 12        Reject  $\mu^{(k+1)}$ , shrink the TR radius  $\delta^{(k+1)} = \beta_1\delta^{(k)}$  and go to 3;
- 13    else
- 14        Update the RB model at  $\mu^{(k+1)}$  and compute  $\varrho^{(k)}$  from (4.7);
- 15        if  $\hat{\mathcal{J}}_r^{(k+1)}(\mu^{(k+1)}) \leq \hat{\mathcal{J}}_r^{(k)}(\mu_{\text{AGC}}^{(k)})$  then
- 16            Accept  $\mu^{(k+1)}$ ;
- 17            if  $\varrho^{(k)} \geq \eta_e$  then
- 18                Enlarge the TR radius  $\delta^{(k+1)} = \beta_1^{-1}\delta^{(k)}$ ;
- 19            else
- 20                Set  $\delta^{(k+1)} = \delta^{(k)}$ ;
- 21            end
- 22        else
- 23            Reject  $\mu^{(k+1)}$ , shrink the TR radius  $\delta^{(k+1)} = \beta_1\delta^{(k)}$  and go to 3;
- 24        end
- 25    end
- 26    if  $\|\mu^{(k+1)} - \text{P}_{\mathcal{P}}(\mu^{(k+1)} - \nabla_{\mu}\hat{\mathcal{J}}_h(\mu^{(k+1)}))\|_2 \leq \tau_{\text{FOC}}$  then
- 27        Set Loop\_flag=False;
- 28    end
- 29    Set  $k = k + 1$ ;
- 30 end

---

$k \in \mathbb{N}$ . Hence there exists  $j_c^{(0)} = \max(j_1^{(0)}, j_2^{(0)}) < \infty$ , for which  $\mu^{(0,0)}(j)$  satisfies (4.4) and (4.5) for  $\ell = 0$ . If  $k \geq 1$ , since the model has been enriched, i.e.,  $\Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu^{(k)}) = 0$ , we can show the claim arguing as we did for  $k = 0$ . Note, in fact, that we increase the iteration counter only when  $\mu^{(k)}$  is accepted at iteration  $k - 1$  and, thus, when the RB model is enriched at this parameter.  $\square$

**Theorem 4.5.** *Let the hypotheses of Lemma 4.4 be verified. Then every accumulation point  $\bar{\mu}$  of the sequence  $\{\mu^{(k)}\}_{k \in \mathbb{N}} \subset \mathcal{P}$  generated by the TR-RB algorithm is an approximate first-order critical point for  $\hat{\mathcal{J}}_h$  (up to the chosen tolerance  $\tau_{\text{sub}}$ ), i.e., it holds*

$$\|\bar{\mu} - \text{P}_{\mathcal{P}}(\bar{\mu} - \nabla_{\mu}\hat{\mathcal{J}}_h(\bar{\mu}))\|_2 \leq \tau_{\text{sub}}. \quad (4.9)$$

*Proof.* The set  $\mathcal{P} \subset \mathbb{R}^P$  is compact. Therefore there exists a sequence of indices  $\{k_i\}_{i \in \mathbb{N}}$  such that the sub-sequence  $\{\mu^{(k_i)}\}_{i \in \mathbb{N}}$  converges to a point  $\bar{\mu} \in \mathcal{P}$ . It remains to show that  $\bar{\mu}$  is an approximate first-order critical

point. At first, note that once the RB space is enriched at a point  $\mu^{(k)}$ , we have  $\Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu^{(k)}) = 0$ . Hence, also  $q^{(k)}(\mu^{(k)}) = 0$  holds, where

$$q^{(k)}(\mu) := \frac{\Delta_{\hat{\mathcal{J}}_r^{(k)}}(\mu)}{\hat{\mathcal{J}}_r^{(k)}(\mu)} \quad \text{for all } k \in \mathbb{N}, \mu \in \mathcal{P}.$$

Note also that  $V_h$  is a finite dimensional space. This implies that at most after  $\dim V_h \leq I < +\infty$  enrichment steps the RB approximation error and the *a posteriori* error estimator are zero for each  $\mu \in \mathcal{P}$ . In particular  $q^{(k_i)}(\mu) = 0$  for all  $\mu \in \mathcal{P}$  and  $i \geq I$ . For this reason, the stopping criterium (4.6b) is not triggered. Hence, we have proved that each  $\mu^{(k_{i+1})}$  is an approximate first-order critical point for  $\hat{\mathcal{J}}_r^{(k_{i+1}-1)}$  (up to the chosen tolerance  $\tau_{\text{sub}}$ ) for all  $i \geq I$ , which yields to

$$\left\| \mu^{(k_{i+1})} - \text{P}_{\mathcal{P}} \left( \mu^{(k_{i+1})} \right) - \nabla_{\mu} \hat{\mathcal{J}}_r^{(k_{i+1}-1)} \left( \mu^{(k_{i+1})} \right) \right\|_2 \leq \tau_{\text{sub}}, \quad \text{for all } i \geq I.$$

Moreover, taking again into account the RB method properties and the fact that  $V_h$  is a finite dimensional space, there exists a constant  $I_{\nabla} > 0$  sufficiently large, such that  $\nabla_{\mu} \hat{\mathcal{J}}_r^{(k_i)}(\mu) = \nabla_{\mu} \hat{\mathcal{J}}_h(\mu) + \epsilon^{(k_i)}$  for all  $\mu$  in a neighborhood of  $\bar{\mu}$  and for  $i \geq I_{\nabla}$ , with  $\epsilon^{(k_i)} \rightarrow 0$  as  $i \rightarrow \infty$ . Thus, exploiting the continuity of the projection operator and assuming  $i \geq \max(I, I_{\nabla})$ , we have that

$$\begin{aligned} \tau_{\text{sub}} &\geq \left\| \mu^{(k_{i+1})} - \text{P}_{\mathcal{P}} \left( \mu^{(k_{i+1})} \right) - \nabla_{\mu} \hat{\mathcal{J}}_r^{(k_{i+1}-1)} \left( \mu^{(k_{i+1})} \right) \right\|_2 \\ &= \left\| \mu^{(k_{i+1})} - \text{P}_{\mathcal{P}} \left( \mu^{(k_{i+1})} \right) - \nabla_{\mu} \hat{\mathcal{J}}_h \left( \mu^{(k_{i+1})} \right) + \left( \epsilon^{(k_{i+1}-1)} \right) \right\|_2 \rightarrow \left\| \bar{\mu} - \text{P}_{\mathcal{P}} \left( \bar{\mu} - \nabla_{\mu} \hat{\mathcal{J}}_h(\bar{\mu}) \right) \right\|_2. \end{aligned}$$

Hence, the accumulation point  $\bar{\mu}$  is an approximate first-order critical point (up to the tolerance  $\tau_{\text{sub}}$ ).  $\square$

**Remark 4.6.** What remains to prove is that  $\bar{\mu}$  is a local minimum of  $\hat{\mathcal{J}}_h$  (or rather a sufficiently close approximation of a local minimum). Exploiting the sufficient decrease condition, one can easily show by contradiction that  $\bar{\mu}$  is not a maximum of  $\hat{\mathcal{J}}_h$ . It can still be a saddle point as well as a local minimum. In the numerical experiments, to verify that the computed point  $\bar{\mu}$  is actually a local minimum, we employ the second-order sufficient optimality conditions after the algorithm terminates.

Moreover, the proof of Theorem 4.5 relies on the fact that  $V_h$  is a finite-dimensional space and that at most after  $\dim V_h \leq I < +\infty$  iterations the RB space is exact. On a practical point of view, having a RB space of the same dimension of the full-order model will not give any speed-up. Since the optimization sequence will fast accumulate around  $\bar{\mu}$ , we expect that the RB model will be accurate enough to not trigger (4.6b) thus implying the TR method convergence in a much smaller number of iterations. The numerical tests in Section 5 confirms this expectation.

### 4.3. Construction of RB spaces

In an enrichment step of the outer loop of the TR-algorithm 1 for  $\mu \in \mathcal{P}$ , we assume to have access to the primal and dual solutions  $u_{h,\mu}, p_{h,\mu} \in V_h$  and consider two strategies to enrich the RB spaces.

- (a) *Lagrangian RB spaces*: we add each FOM solution to the RB space that is directly related to its respective reduced formulation, *i.e.*, for a given  $\mu \in \mathcal{P}$ , we enrich by  $V_r^{\text{pr},k} = V_r^{\text{pr},k-1} \cup \{u_{h,\mu}\}$ ,  $V_r^{\text{du},k} = V_r^{\text{du},k-1} \cup \{p_{h,\mu}\}$ .
- (b) *Aggregated RB space*: we add all available information into a single RB space, *i.e.*,  $V_r^{\text{pr},k} = V_r^{\text{pr},k} = V_r^{\text{pr},k-1} \cup \{u_{h,\mu}, p_{h,\mu}\}$ . According to Section 3 this results in  $\hat{\mathcal{J}}_r(\mu)$  being equal to the standard RB reduced functional from (3.5).

These strategies for constructing RB spaces have a significant impact on the performance and accuracy of the TR-RB method. Note that offline computations for the construction of RB models scale quadratically with the number of basis functions in the RB space. Thus, Lagrange RB spaces in (a) are computationally beneficial compared to (b) at a potential loss of accuracy of the corresponding RB models (since less information is added). Moreover, different spaces as in (a) destroy the duality of state and adjoint equations, *cf.* Section 3.2.

#### 4.4. Trust-Region variants based on adaptive enrichment strategies

A major contribution of this article is to introduce and analyze variants of adaptive TR-RB methods with projected BFGS as sub-problem solver for efficiently computing a solution of the optimization problem (P). In terms of performance we need to account for all computational costs, including traditional offline and online costs of the algorithms. The proposed methods mainly differ in terms of the model function and its gradient information. Following Section 4.1, we propose a TR method which adaptively builds an RB space along the path of optimization (see Algorithm 1). From a MOR perspective this diminishes the offline time of the ROM significantly since no global RB space (with respect to the parameter domain) has to be built in advance. We enrich the model after the sub-problem (4.2) of the TR method has been solved. We distinguish three different approaches:

- (1) *Standard approach*: following Section 3.2, the standard approach for the functional is to replace the FOM quantities by their respective ROM counterpart, *i.e.*, we consider the map  $\mu \mapsto \hat{J}_r(\mu)$  from (3.5). Gradient information can be computed by reducing the corresponding FOM gradient which results in  $\tilde{\nabla}_\mu \mathcal{J}(u_{r,\mu}, \mu)$  from (3.8). Consequently this approach does not allow for using a higher order estimate but  $\Delta_{\hat{J}_r}(\mu)$ .
- (2) *Semi NCD-corrected approach*: a first correction strategy is to replace the functional by the NCD-corrected RB reduced functional  $\hat{\mathcal{J}}_r$  from (3.9) but stick with the inexact gradient of the standard approach. This allows to use the higher order estimator for the functional, *i.e.*,  $\Delta_{\hat{\mathcal{J}}_r}(\mu)$ .
- (3) *NCD-corrected approach*: we propose to consider the NCD-corrected RB reduced functional  $\hat{\mathcal{J}}_r$  from (3.9) and its actual gradient according from Proposition 3.3. Note that we only need to solve two additional equations, independently of the dimension  $P$  of the parameter space.

For the basis construction, we may use variants (a) or (b) from Section 4.3. Note however, that by using the basis enrichment (b), all approaches 1–3 are equivalent. Using variant (a) with BFGS is inspired from [50]. However, our algorithms differ from the TR-RB approach in [50] since we are working with the NCD-corrected reduced cost functional (in 2) and its actual gradient (in 2 and 3). Note that the presence of inequality constraints, which are missing in [50], implies a projection-based optimization algorithm. In addition, we stress that, differently from [50], we take advantage of the proposed condition for enlarging the TR radius and of a stopping criterium independent from the RB *a posteriori* estimates, as presented in Section 4.1.

**Remark 4.7.** Note that we do not use the sensitivity based quantities from Section 3.4.2 although they suggest the highest numerical accuracy w.r.t. the FOM optimality system. However, for the experiments in Section 5, additional computational cost for computing FOM sensitivities will not pay off in the TR-RB algorithm, especially for high-dimensional parameter spaces.

## 5. NUMERICAL EXPERIMENTS

We present numerical experiments to demonstrate the adaptive TR-RB variants from Section 4.4 with both RB constructions from Section 4.3 for quadratic objective functionals with elliptic PDE constraints as in (P), and compare them to state-of-the art algorithms from the literature. We also validate the higher-order *a posteriori* error estimates from Section 3.4 numerically. We consider two setups: first, the elliptic thermal fin problem from Section 5.1.1 of [50] (where the correction term of the proposed NCD-corrected approach vanishes) in Section 5.2. Second, we consider a more challenging optimization problem in Section 5.3, including a detailed analysis of the *a posteriori* error estimates from Section 3.4. All simulations have been performed with a pure Python implementation based on the open source MOR library pyMOR [43], making use of pyMORs builtin vectorized numpy/scipy-based discretizer for the FOM and generic MOR algorithms for projection and orthonormalization (such as a stabilized Gram-Schmidt algorithm) to effortlessly obtain efficient ROMs. The source code to reproduce all results (including detailed interactive jupyter-notebooks<sup>1</sup>) is available at [39]. All

<sup>1</sup>Available at <https://github.com/TiKeil/NCD-corrected-TR-RB-approach-for-pde-opt>.

experiments are based on the same implementation (including a reimplementaion of [50]) and were performed on the same machine multiple times to avoid caching or multi-query effects. Timings may thus be used to compare and judge the computational efficiency of the different algorithms.

We consider stationary heat transfer in a bounded connected spatial domain  $\Omega \subset \mathbb{R}^2$  with polygonal boundary  $\partial\Omega$  partitioned into a non-empty Robin boundary  $\Gamma_R \subset \partial\Omega$  and possibly empty distinct Neumann boundary  $\Gamma_N = \partial\Omega \setminus \Gamma_R$ , and unit outer normal  $n : \partial\Omega \rightarrow \mathbb{R}^2$ . We consider the Hilbert space  $V = H^1(\Omega) := \{v \in L^2(\Omega) \mid \nabla v \in L^2(\Omega)\}$  of weakly differentiable functions and, for an admissible parameter  $\mu \in \mathcal{P}$ , we seek the temperature  $u_\mu \in V$  as the solution of

$$-\nabla \cdot (\kappa_\mu \nabla u_\mu) = f_\mu \text{ in } \Omega, \quad \kappa_\mu \nabla u_\mu \cdot n = c_\mu (u_{\text{out}} - u_\mu) \text{ on } \Gamma_R, \quad \kappa_\mu \nabla u_\mu \cdot n = g_N \text{ on } \Gamma_N \quad (5.1)$$

in the weak sense, with the admissible parameter set, the spatial domain and its boundaries and the data functions  $\kappa_\mu \in L^\infty(\Omega)$ ,  $f_\mu \in L^2(\Omega)$ ,  $c_\mu \in L^\infty(\Gamma_R)$  and  $u_{\text{out}} \in L^2(\Gamma_R)$  defined in the respective experiment. The bilinear form  $a$  and linear functional  $l$  in (P.b) are thus given for all  $\mu \in \mathcal{P}$  and  $v, w \in V$  by

$$a_\mu(v, w) := \int_\Omega \kappa_\mu \nabla v \cdot \nabla w \, dx + \int_{\Gamma_R} c_\mu v w \, ds \quad \text{and} \quad l_\mu(v) := \int_\Omega f_\mu v \, dx + \int_{\Gamma_R} c_\mu u_{\text{out}} v \, ds + \int_{\Gamma_N} g_N v \, ds. \quad (5.2)$$

For the FOM we fix a fine enough reference simplicial or cubic mesh and define  $V_h \subset V$  as the respective space of continuous piecewise (bi-)linear Finite Elements.

Since the inner product and norm have a big influence on the computational efficiency of the *a posteriori* error estimates as well as their sharpness, we use the mesh-independent energy-product  $(u, v) := a_{\check{\mu}}(u, v)$  for a fixed parameter  $\check{\mu} \in \mathcal{P}$ . Since the bilinear form  $a$  in our experiments is symmetric, continuous and coercive, this energy-product indeed is a product over  $V$ . Owing to this choice of the product, we may use the min-theta approach from Proposition 2.35 of [27] to obtain lower bounds on coercivity constants and the max-theta approach from Example 5.12 of [27] to obtain upper bounds on continuity constants, each required for the *a posteriori* error estimates. Compared to the more general Successive Constraint Method [36], this approach yields quite sharp estimates and is computationally more efficient, both offline and online. Due to Assumption 2.2 and the bi-linearity of the objective functional, we may carry out the preassembly of all high-dimensional quantities after each enrichment, which is well-known for RB methods ([27], Sect. 2.5). We would like to point out that while the more accurate and stable preassembly of the estimates from [10] is readily available in pyMOR, the slightly cheaper standard preassembly of the estimates was sufficient for our experiments.

For all experiments, we use an initial TR radius of  $\delta^0 = 0.1$ , a TR shrinking factor  $\beta_1 = 0.5$ , an Armijo step-length  $\kappa = 0.5$ , a truncation of the TR boundary of  $\beta_2 = 0.95$ , a tolerance for enlarging the TR radius of  $\eta_\varrho = 0.75$ , a stopping tolerance for the TR sub-problems of  $\tau_{\text{sub}} = 10^{-8}$ , a maximum number of TR iteration  $K = 40$ , a maximum number of sub-problem iteration  $K_{\text{sub}} = 400$  and a maximum number of Armijo iteration of 50. We also point out that the stopping tolerance for the FOC condition  $\tau_{\text{FOC}}$  is specified in each experiment.

### 5.1. State of the art optimization methods

We compare our proposed methods to the following ones from the literature:

**Adaptive TR-RB with BFGS sub-problem solver and Lagrangian RBs [50]:** we consider the same method as in [50], where the authors used the standard functional and gradient from Section 3.2. Furthermore, no enlarging strategy has been used for the TR-radius and no projection for parameter constraints has been considered. Importantly, the authors did not take advantage of the fact, that the full order FOC condition in line 23 of Algorithm 1 is cheaply available after an enrichment step. Instead they used the reduced FOC condition plus the estimator for the gradient of the cost functional  $\|\tilde{\nabla}_{\check{\mu}} \hat{J}_r(\mu^{(k+1)})\|_2 + \Delta_{\check{\mu}} \hat{J}_r(\mu) \leq \tau_{\text{FOC}}$  in line 23. Note that this approach has multiple drawbacks. First, the evaluation is more costly due to the estimator. Second, it is less accurate and third, it can prevent the TR-RB from converging in case the estimator is not able to be small enough (for instance governed by large constants or numerical issues in the estimator).

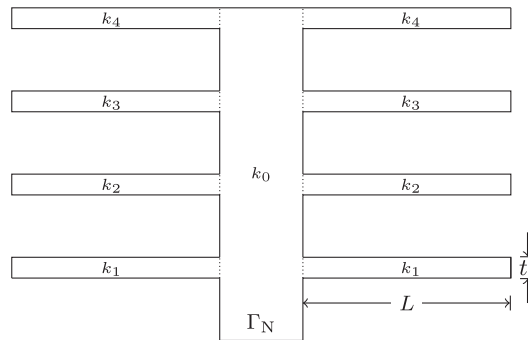


FIGURE 1. Problem definition of the thermal fin example from Section 5.2. Depicted is the spatial domain  $\Omega$  (with  $L = 2.5$  and  $t = 0.25$ ) with Neumann boundary at the bottom with  $|\Gamma_N| = 1$  and Robin boundary  $\Gamma_R := \partial\Omega \setminus \Gamma_N$ , as well as the values  $k_0, \dots, k_4 > 0$  of the diffusion  $k_\mu$ , which is piecewise constant in the respective indicated part of the domain. The other data functions in (5.1) are given by  $f_\mu = 0$ ,  $g_N = -1$ ,  $u_{\text{out}} = 0$  and  $c_\mu = \text{Bi} \in \mathbb{R}$ , the Biot number. We allow to vary the six parameters  $(k_0, \dots, k_4, \text{Bi})$  and define the set of admissible parameters as  $[0.1, 10]^5 \times [0.01, 1] \subset \mathbb{R}^P$  with  $P = 6$ . We choose  $\check{\mu} = (1, 1, 1, 1, 1, 0.1)$  for the energy product.

**FOM projected BFGS:** we consider a standard projected BFGS method, which uses FOM evaluations of the forward model for the reduced cost functional and its gradient. We restrict the number of iterations by 400.

## 5.2. Model problem 1: Elliptic thermal fin model problem

We consider the six-dimensional elliptic thermal fin example from Section 5.1.1 of [50] and refer to Figure 1 for the problem definition. The purpose of this experiment is to show the applicability of the proposed algorithms and to compare them to the one proposed in [50]. For all runs we prescribe the same desired parameter  $\mu^d \in \mathcal{P}$  by randomly drawing  $k_1, \dots, k_4$  strictly within  $\mathcal{P}$  and by setting  $k_0 = 0.1$  and  $\text{Bi} = 0.01$ , to artificially mimic the situation where parameter constraints have to be tackled. Defining  $T^d := q(u_{\mu^d})$  where  $u_{\mu^d} \in V$  is the solution of (P.b) associated with the desired parameter and where  $q(v) := \int_{\Gamma_N} v \, ds$  for  $v \in V$  denotes the mean temperature at the root of the fin, we consider a cost functional  $\mathcal{J}(u, \mu) = \Theta(\mu) + j_\mu(u) + k_\mu(u, u)$  as in (P.a) with  $\Theta(\mu) := (\|\mu^d - \mu\|/\|\mu^d\|)^2 + T^{d^2} + 1$ ,  $j_\mu(v) := -T^d q(v)$  and  $k_\mu(v, w) := 1/2 q(v) q(w)$ . We would like to point out that the authors in [50] dropped the  $T^{d^2} + 1$  term from the definition of  $\Theta$ , which we re-add to ensure Assumption 4.2. This constant term does not change the position of local minima and the derivatives of the cost functional. However, it makes the Trust-Region radius shrink especially at the beginning, slowing down the TR-RB methods. This does not affect the comparison among the TR-RB methods, since all suffer from this issue. Note that for this particular example, the proposed NCD-correction term vanishes, see Remark 5.1. For the FOM, we generate an unstructured simplicial mesh using pyMORs `gmsh` (see [22]) bindings, resulting in  $\dim V_h = 77537$ .

Starting with ten different randomly drawn initial parameters  $\mu^{(0)}$ , we measure the total computational runtime, the number of TR iterations  $k$  and the error in the optimal parameter for all combinations of adaptive TR algorithms from Section 4 and choice of RB spaces from Section 4.3, as well as for the state of the art methods from the literature from Section 5.1.

All considered optimization methods converged (up to a tolerance), but we restrict the presentation to the most informative ones (all results can be found in the accompanying code). As we observe from Table 1, the ROM based adaptive TR-RB algorithms vastly outperform the FOM variant, noting that the computational time of the ROM variants includes all offline and online computations. Figure 2 details the decay of the error decay in the optimal parameter during the optimization for a selected random initial guess. We observe that the



TABLE 1. Performance and accuracy of selected algorithms for the example from Section 5.2 for ten optimization runs with randomly initial guesses  $\mu^{(0)}$ : averaged, minimum and maximum total computational time (column 2) and speed-up compared to the FOM variant (column 3); average, minimum and maximum number of iterations  $k$  required until convergence (column 4), average relative error in the parameter (column 5) and average FOC condition (column 6).

	av. (min/max) runtime[s]	Speed-up	av. (min/max) iter.	rel. error	FOC cond.
FOM proj. BFGS	967.86 (176.69/3401.06)	–	111.20 (25/400)	$3.13 \times 10^{-3}$	$1.19 \times 10^{-2}$
TR-RB from [50]	68.06 (43.28/88.21)	10.40	7.20 (8/13)	$1.34 \times 10^{-6}$	$4.31 \times 10^{-5}$
1(a) TR-RB with $V_r^{\text{pr}} \neq V_r^{\text{du}}$	44.56 (34.22/74.96)	21.72	8.80 (8/11)	$3.08 \times 10^{-6}$	$4.64 \times 10^{-5}$
1(b) TR-RB with $V_r^{\text{pr}} = V_r^{\text{du}}$	43.86 (34.09/74.35)	22.07	8.70 (8/10)	$3.37 \times 10^{-6}$	$6.40 \times 10^{-5}$

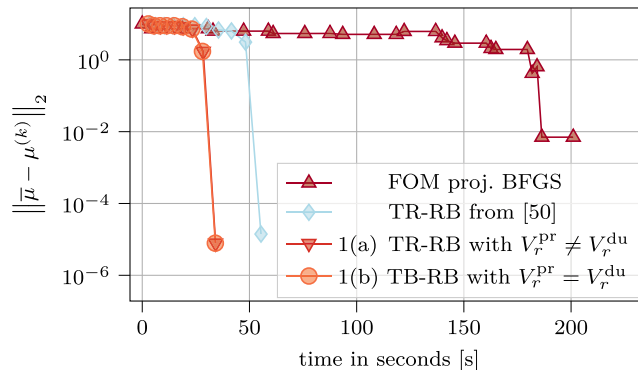


FIGURE 2. Error decay and performance of selected algorithms for the example from Section 5.2 for a single optimization run with random initial guess  $\mu^{(0)}$  for  $\tau_{\text{FOC}} = 5 \times 10^{-4}$ : for each algorithm each marker corresponds to one (outer) iteration of the optimization method and indicates the absolute error in the current parameter, measured against the known desired optimum  $\bar{\mu} = \mu^{\text{d}}$ . In all except the FOM variant, the ROM is enriched in each iteration corresponding to Algorithm 1, depending on the variant in question.

choice of the RB enrichment does not impact the performance of the algorithm for this example too much, see Remark 5.1. Also methods 2(a) and 3(a) show a comparable computational speed (not shown). We also observe that the method from [50] requires more time and more iterations on average, variants 1 are still faster due to the enlarging of the TR radius and of the use of a termination criterium which does not depend on *a posteriori* estimates, which may force additional TR iterations.

**Remark 5.1** (Vanishing NCD-correction for the fin problem). It is important to notice that this model problem is not suitable to fully demonstrate the capabilities of the NCD-corrected approach. The reason is that the choice of the functional is a misfit on only the root edge of the thermal fin, plus a Tikhonov regularization term. Since the root of the thermal fin is also the source of the primal problem, the dual solutions  $p_{r,\mu}$  of the reduced dual equation (3.4b) are thus linearly dependent on the respective primal solutions  $u_{r,\mu}$  and the correction term  $r_{\mu}^{\text{pr}}(u_{r,\mu})[p_{r,\mu}]$  for the NCD-corrected RB reduced functional from (3.9) vanishes. In general, for quadratic objective functionals, this is not the case and all variants with correction terms thus waste unnecessary computational time.

### 5.3. Model problem 2: stationary heat distribution in a building

For these experiments we consider as objective functional a weighted  $L^2$ -misfit on a domain of interest  $D \subseteq \Omega$  and a weighted Tikhonov term comparable to design optimization, optimal control or inverse problems,

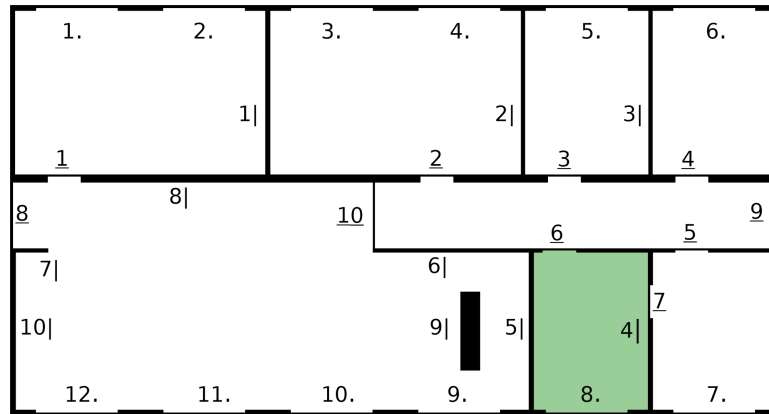


FIGURE 3. Definition of Model problem 2: with  $\Omega := [0, 2] \times [0, 1] \subset \mathbb{R}^2$ . Numbers indicate affine components, where  $i$  is a window,  $\bar{i}$  are doors, and  $i|$  are walls. The  $i$ th heater is located under the  $i$ th window. With respect to (5.1), we consider  $\Gamma_R := \partial\Omega$ , where  $c_\mu$  contains outside wall  $10|$ , outside doors  $\bar{8}$  and  $\bar{9}$  and all windows. All other diffusion components enter the coefficient  $\kappa_\mu$ , whereas the heaters enter into the source term  $f_\mu$ . Furthermore, we set  $u_{\text{out}} = 5$  and the green region illustrates the domain of interest  $D$ .

*i.e.*,

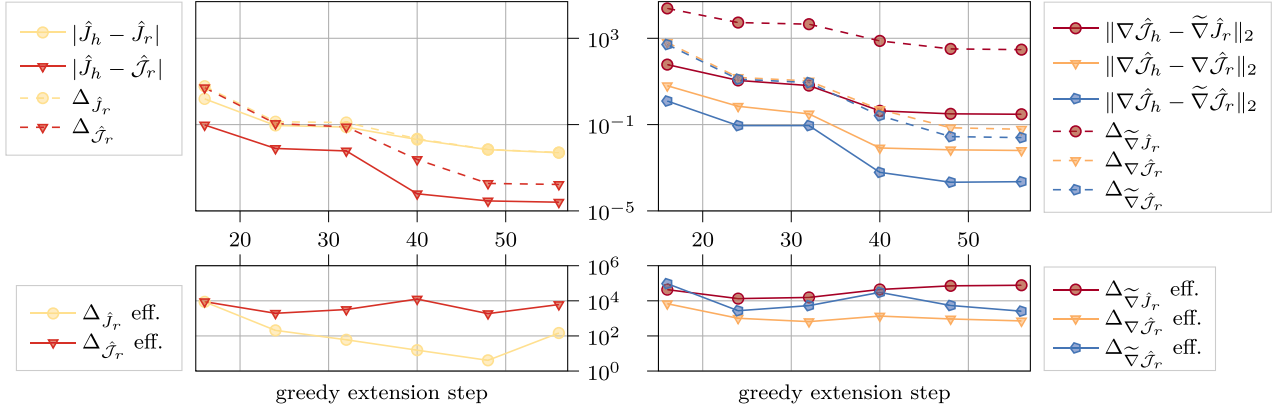
$$\mathcal{J}(v, \mu) = \frac{\sigma_D}{2} \int_D (v - u^d)^2 + \frac{1}{2} \sum_{i=1}^M \sigma_i (\mu_i - \mu_i^d)^2 + 1, \quad (5.3)$$

with given desired state  $u^d \in V$  and parameter  $\mu^d \in \mathcal{P}$  and weights  $\sigma_D, \sigma_i$  specified further below. With respect to (P.a), we thus have  $\Theta(\mu) = \frac{1}{2} \sum_{i=1}^M \sigma_i (\mu_i - \mu_i^d)^2 + \frac{\sigma_D}{2} \int_D u^d u^d + 1$ ,  $j_\mu(u) = -\sigma_D \int_D u^d u$  and  $k_\mu(u, v) = \frac{\sigma_D}{2} \int_D uv$ .

Motivated by ensuring a desired temperature in a single room of a building floor, we consider blueprints with windows, heaters, doors and walls, yielding parameterized diffusion, forces and boundary values as sketched in<sup>2</sup> Figure 3. For simplicity we omit a realistic modeling of temperature and restrict ourselves to academic numbers of the diffusion and heat source quantities. We seek to ensure a desired temperature  $u^d = 18$  and set  $\mu_i^d = 0$ . For the FOM discretization we choose a cubic mesh which resolves all features of the data functions derived from Figure 3, resulting in  $\dim V_h = 80\,601$  degrees of freedom. We consider a ten-dimensional parameter example with three wall sets  $\{1|, 2|, 3|, 8|\}$ ,  $\{4|, 5|, 6|, 7|\}$  and  $\{9|\}$  and seven heater sets,  $\{1, 2\}$ ,  $\{3, 4\}$  and  $\{5\}$ ,  $\{6\}$ ,  $\{7\}$ ,  $\{8\}$  and  $\{9, 10, 11, 12\}$  (each set governed by a single parameter component). The set of admissible parameters is given by  $\mathcal{P} = [0.025, 0.1]^3 \times [0, 100]^7$  and we choose  $\sigma_D = 100$  and  $(\sigma_i)_{1 \leq i \leq 10} = (10\sigma_w, 5\sigma_w, \sigma_w, 2\sigma_h, 2\sigma_h, \sigma_h, \sigma_h, \sigma_h, \sigma_h, 4\sigma_h)$  in (5.3), with  $\sigma_w = 0.05$  and  $\sigma_h = 0.001$ . The choice of  $\sigma_i$  is related to the measure of the walls and how many heaters are considered in each group. The other components of the data functions are fixed and thus not directly involved in the optimization process. Briefly, the diffusion coefficient of air and inside doors is set to 0.5, of the outside wall to 0.001, of outside doors  $\bar{8}$  and  $\bar{9}$  to 0.01 and of windows to 0.025. For the energy product, we choose  $\tilde{\mu} = (0.05, 0.05, 0.05, 10, 10, 10, 10, 10, 10, 10)$ .

We use this setup to inspect different TR-RB algorithms in Section 5.3.2, but also to study the *a posteriori* error estimates from Section 3.4 in the following section.

<sup>2</sup>See <https://github.com/TiKeil/NCD-corrected-TR-RB-approach-for-pde-opt> for the definition of the data functions.



(A) approximating the reduced functional  $\hat{\mathcal{J}}_h$       (B) approximating the gradient of the reduced functional

FIGURE 4. Evolution of the true and estimated model reduction error (*top*) in the reduced functional and its approximations (A) and the gradient of the reduced functional and its approximations (B), as well as error estimator efficiencies (*bottom*), during adaptive greedy basis generation for the experiment from Section 5.3.1. *Top*: depicted is the  $L^\infty(\mathcal{P}_{\text{val}})$ -error for a validation set  $\mathcal{P}_{\text{val}} \subset \mathcal{P}$  of 100 randomly selected parameters, *i.e.*,  $|\hat{\mathcal{J}}_h - \hat{\mathcal{J}}_r|$  corresponds to  $\max_{\mu \in \mathcal{P}_{\text{val}}} |\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_r(\mu)|$ ,  $\Delta_{\hat{\mathcal{J}}_r}$  corresponds to  $\max_{\mu \in \mathcal{P}_{\text{val}}} \Delta_{\hat{\mathcal{J}}_r}(\mu)$ ,  $\|\nabla \hat{\mathcal{J}}_h - \nabla \hat{\mathcal{J}}_r\|_2$  corresponds to  $\max_{\mu \in \mathcal{P}_{\text{val}}} \|\nabla \hat{\mathcal{J}}_h(\mu) - \nabla \hat{\mathcal{J}}_r(\mu)\|_2$ , and so forth. *Bottom*: depicted is the worst effectivity of the respective error estimate (lower: better), *i.e.*, “ $\Delta_{\hat{\mathcal{J}}_r}$  eff.” corresponds to  $\max_{\mu \in \mathcal{P}_{\text{val}}} \Delta_{\hat{\mathcal{J}}_r}(\mu) / |\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_r(\mu)|$ , and so forth.

### 5.3.1. Numerical validation of the a posteriori error estimates

To study the performance of the *a posteriori* error estimates proposed in Section 3.4, we neglect the outer-loop optimization and simply use a goal oriented adaptive greedy algorithm [28] with basis extension (a) from Section 4.3 to generate a ROM which ensures that the worst relative estimated error for the reduced functional and its gradient over the adaptively generated training set and a randomly chosen validation set is below a prescribed tolerance of  $\tau_{\text{FOC}} = 5 \times 10^{-4}$ . In particular we first ensure  $\Delta_{\hat{\mathcal{J}}_r}(\mu) / \hat{\mathcal{J}}_r(\mu) < \tau_{\text{FOC}}$  for  $\Delta_{\hat{\mathcal{J}}_r}$  from Proposition 3.6(i) and continue with  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu) / \|\tilde{\nabla} \hat{\mathcal{J}}_r(\mu)\|_2 < \tau_{\tilde{\nabla} \hat{\mathcal{J}}_r}$  for  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}$  from Proposition 3.8(i), *cf.* Algorithm 2 of [50]. Let us mention that the goal for  $\Delta_{\hat{\mathcal{J}}_r}$  is fulfilled after 24 basis enrichments and we have  $\Delta_{\tilde{\nabla} \hat{\mathcal{J}}_r}(\mu) / \|\tilde{\nabla} \hat{\mathcal{J}}_r(\mu)\|_2 < 4.84$  after 56 basis enrichments, where we artificially stop the algorithm since the associated computational effort is already roughly 17h, demonstrating the need for the proposed adaptive TR-RB algorithm studied in the next section.

As we observe from Figure 4, the error of the NCD-corrected terms is of several orders of magnitude smaller than the corresponding terms of the standard approach. It can also be seen that the (computationally more costly) sensitivity bases quantities, *i.e.*,  $\tilde{\nabla} \hat{\mathcal{J}}_r$ , show the best error. However, all estimators for the corrected and sensitivity based quantities show a worse effectivity, hinting that there is still room for improvement.

### 5.3.2. Optimization results

Similar to Section 5.2, starting with ten different randomly drawn initial parameters  $\mu^{(0)}$ , we measure the total computational runtime, the number of TR iterations  $k$  and the error in the optimal parameter for all combinations of adaptive TR algorithms from Section 4 and choice of RB spaces from Section 4.3, as well as for the state of the art methods from the literature from Section 5.1.

TABLE 2. Performance and accuracy of selected algorithms for two choices of  $\tau_{\text{FOC}}$  for the example from Section 5.3.2 for ten optimization runs with random initial guess, compare Table 1.

(A) Result for $\tau_{\text{FOC}} = 5 \times 10^{-4}$	av. (min/max) runtime[s]	speed-up	av. (min/max) iter.	rel. error	FOC cond.
FOM proj. BFGS	332.57 (196.51/591.85)	–	44.30 (30/60)	$1.40 \times 10^{-3}$	$1.80 \times 10^{-4}$
TR-RB from [50]	117.87 (70.29/166.31)	2.82	10.10 (6/14)	$5.46 \times 10^{-4}$	$1.41 \times 10^{-4}$
1(a) TR-RB with $V_r^{\text{pr}} \neq V_r^{\text{du}}$	91.50 (47.07/230.29)	3.63	8.30 (5/10)	$2.01 \times 10^{-3}$	$2.04 \times 10^{-4}$
1(b) TR-RB with $V_r^{\text{pr}} = V_r^{\text{du}}$	78.65 (54.69/114.36)	4.23	6.90 (5/9)	$2.53 \times 10^{-4}$	$8.23 \times 10^{-5}$
2(a) TR-RB semi NCD-corrected	79.47 (63.38/94.28)	4.18	8.50 (7/10)	$5.98 \times 10^{-5}$	$1.02 \times 10^{-4}$
3(a) TR-RB NCD-corrected	71.84 (50.38/87.16)	4.63	7.40 (5/9)	$1.09 \times 10^{-3}$	$6.12 \times 10^{-5}$
(B) Result for $\tau_{\text{FOC}} = 10^{-6}$	av. (min/max) runtime[s]	speed-up	av. (min/max) iter.	rel. error	FOC cond.
FOM proj. BFGS	409.28 (317.25/637.55)	–	57.00 (49/71)	$2.82 \times 10^{-6}$	$3.35 \times 10^{-7}$
TR-RB from [50]	614.81 (566.66/671.97)	0.66	40.00 (40/40)	$8.46 \times 10^{-7}$	$8.44 \times 10^{-8}$
1(a) TR-RB with $V_r^{\text{pr}} \neq V_r^{\text{du}}$	165.48 (92.26/417.24)	2.47	15.30 (10/40)	$3.29 \times 10^{-6}$	$5.43 \times 10^{-7}$
1(b) TR-RB with $V_r^{\text{pr}} = V_r^{\text{du}}$	86.39 (62.68/124.43)	4.74	7.80 (6/10)	$3.52 \times 10^{-6}$	$3.03 \times 10^{-7}$
2(a) TR-RB semi NCD-corrected	90.37 (80.97/102.60)	4.53	9.80 (9/11)	$8.12 \times 10^{-7}$	$2.26 \times 10^{-7}$
3(a) TR-RB NCD-corrected	88.24 (58.18/108.90)	4.64	8.90 (6/10)	$2.65 \times 10^{-6}$	$2.73 \times 10^{-7}$

All algorithms converged (up to a tolerance) to the same point  $\bar{\mu}$  and it was verified *a posteriori* that this point is a local minimum of  $\hat{\mathcal{J}}$ , *i.e.*, it satisfies the second-order sufficient optimality conditions. The value of  $\bar{\mu}$  in order to compute the relative error was calculated with the FOM projected Newton method for a FOC condition tolerance of  $10^{-12}$  and, thanks to the choice of the cost functional weights, the target  $u^{\text{d}}$  is approximate by  $\bar{u}$  with a relative error of  $1.7 \times 10^{-6}$  in  $D$ . We consider the same setup for two different stopping tolerances  $\tau_{\text{FOC}} = 5 \times 10^{-4}$  and  $\tau_{\text{FOC}} = 10^{-6}$  to demonstrate that the performance (both in terms of time and convergence) of the methods vastly depends on the choice of  $\tau_{\text{FOC}}$ .

From Table 2, we observe that all proposed TR-RB methods speed up the FOM projected BFGS method with the NCD-corrected approach outperforming the others, since the gradient used is the true one of the model function  $\hat{\mathcal{J}}_r$ . Moreover, independently of the model function, the algorithm from [50] is much slower, demonstrating the positive impact of the suggested improvements on enlarging the TR radius and on the termination criterium based on cheaply available FOM information (instead of relying on an *a posteriori* estimate), also visible in the number of outer TR iterations. Comparing our proposed TR variants in terms of iterations, it is more beneficial to consider an aggregated RB space, *i.e.*,  $V_r^{\text{pr}} = V_r^{\text{du}}$ . While enrichment (a) is more costly and the time-to-ROM-solution is slightly larger, the richer space seems to allow for better approximations of  $\hat{\mathcal{J}}_h$ .

All methods approximate the optimal parameter  $\bar{\mu}$  with a small relative error and reach the desired tolerance for the FOC condition. However, in view of the resulting relative error in Table 2 and Figure 5, we observe that the choice  $\tau_{\text{FOC}} = 5 \times 10^{-4}$  is not sufficiently small for this model problem. In fact, we observe for most of the variants, that this choice for the tolerance  $\tau_{\text{FOC}}$  does not guarantee an adequately low relative error in approximating  $\bar{\mu}$  and affects the timings by stopping the method too early. We conclude that the choice  $\tau_{\text{FOC}} = 10^{-6}$  instead results in a valid optimum of all variants (up to a tolerance of  $10^{-6}$ ). Importantly, for this choice of  $\tau_{\text{FOC}}$ , we point out that the variant from [50] only stopped because we restricted the maximum number of iterations to 40, although the FOC condition dropped under the depicted tolerance of  $10^{-6}$ . This is caused by the fact that in [50] the *a posteriori* estimate, which is summed to the FOC condition, cannot get numerically small enough, showing the limit of the proposed stopping criterium in [50].

From Figure 5B we conclude that the NCD-corrected approaches 2(a) and 3(a) outperform the standard ROM variant 1(a), which also reached the maximum number of iterations for one of the ten samples. Consequently, the NCD-correction entirely resolves the issue of the variational crime (introduced by splitting the reduced spaces), since it shows roughly the same performance as variant 1(b). However, looking at the minimum and maximum number of computational time in Table 2, variant 3(a) shows a less volatile and more robust behavior.

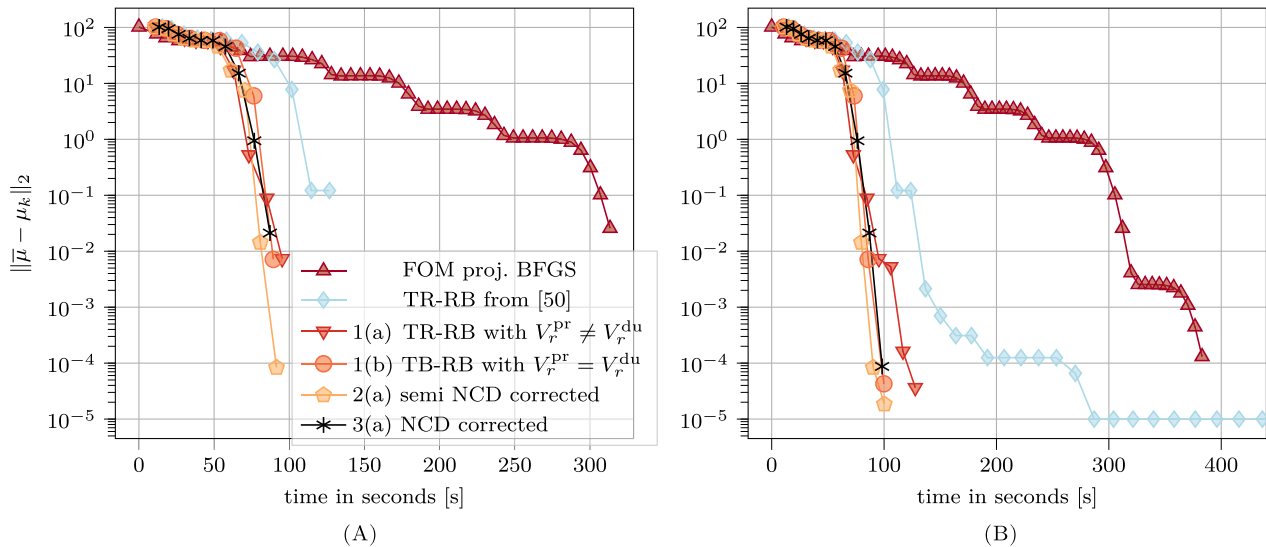


FIGURE 5. Error decay and performance of selected algorithms for two choices of  $\tau_{\text{FOC}}$  for the example from Section 5.3.2 for a single optimization run with random initial guess, compare Figure 2. (A) Result for  $\tau_{\text{FOC}} = 5 \cdot 10^{-4}$ . (B) Result for  $\tau_{\text{FOC}} = 10^{-6}$ .

## 6. CONCLUSION

In this work we proposed and analyzed several variants of new adaptive Trust-Region Reduced Basis methods for parameterized partial differential equations. First, we proved convergence of the modified algorithm in case of additional bilateral constraints on the parameter set, making this method more appealing for real-world applications. Second, the use of a NCD-corrected RB reduced functional improves the RB approximation compared to the standard approach, and enables the possibility of using an exact gradient in the case of separate RB spaces (each variant accompanied by rigorous *a posteriori* error estimates). This approach turns out to be the most reliable in terms of computational time and accuracy, outperforming the existing TR-RB method. Furthermore, the proposed cheaply-computable criteria for enlarging the TR radius and for terminating the iterations ensure a faster convergence. In future works we are interested in considering the projected Newton method to replace the projected BFGS method used in this contribution. This leads to additional effort on developing *a posteriori* estimates for the RB approximation of the hessian and of the optimal parameter. In addition, we are interested in combining the proposed TR-RB algorithm with localized RB methods for large-scale applications.

*Acknowledgements.* The authors acknowledge funding by the Deutsche Forschungsgemeinschaft for the project *Localized Reduced Basis Methods for PDE-constrained Parameter Optimization* under contracts OH 98/11-1; SCHI 1493/1-1; VO 1658/6-1. Tim Keil, Mario Ohlberger and Felix Schindler acknowledge funding by the Deutsche Forschungsgemeinschaft under Germany’s Excellence Strategy EXC 2044 390685587, Mathematics Münster: Dynamics – Geometry – Structure.

## REFERENCES

- [1] E. Arian, M. Fahl and E.W. Sachs, Trust-region proper orthogonal decomposition for flow control. Technical Report 2000-25, ICASE (2000). <http://www.dtic.mil/docs/citations/ADA377382>.
- [2] 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** (2004) 667–672.

- [3] R. Becker, H. Kapp and R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concept. *SIAM J. Control Optim.* **39** (2000) 113–132.
- [4] O. Benedix and B. Vexler, A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Comput. Optim. Appl.* **44** (2009) 3–25.
- [5] P. Benner, A. Cohen, M. Ohlberger and K. Willcox, editors. Model Reduction and Approximation. In: Vol. 15 of *Computational Science & Engineering*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2017). Theory and algorithms.
- [6] P. Benner, M. Ohlberger, A. Patera, G. Rozza and K. Urban, editors. Model Reduction of Parametrized Systems. In: Vol. 17 of *Modeling, Simulation and Applications*. Springer-Verlag GmbH (2017).
- [7] M. Bergmann and L. Cordier, Optimal control of the cylinder wake in the laminar regime by trust-region methods and POD reduced-order models. *J. Comput. Phys.* **227** (2008) 7813–7840.
- [8] M. Bernreuther, G. Müller and S. Volkwein, Reduced basis model order reduction in optimal control of a nonsmooth semilinear elliptic PDE. In: To Appear in *New Trends in PDE Constrained Optimization* (2020).
- [9] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova and P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods. *SIAM J. Math. Anal.* **43** (2011) 1457–1472.
- [10] A. Buhr, C. Engwer, M. Ohlberger and S. Rave, A numerically stable a posteriori error estimator for reduced basis approximations of elliptic equations. In: 11th World Congress on Computational Mechanics, WCCM 2014, 5th European Conference on Computational Mechanics, ECCM 2014 and 6th European Conference on Computational Fluid Dynamics, ECFD 2014 (2014) 4094–4102.
- [11] A. Buhr, C. Engwer, M. Ohlberger and S. Rave, ArbiLoMod, a simulation technique designed for arbitrary local modifications. *SIAM J. Sci. Comput.* **39** (2017) A1435–A1465.
- [12] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control. *Jahresber. Dtsch. Math. Ver.* **117** (2015) 3–44.
- [13] S. Chaturantabut and D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.* **32** (2010) 2737–2764.
- [14] X. Chen, S. Akella and I.M. Navon, A dual-weighted trust-region adaptive POD 4-D Var applied to a finite-volume shallow water equations model on the sphere. *Int. J. Numer. Methods Fluids* **68** (2012) 377–402.
- [15] D. Clever, J. Lang, S. Ulbrich and C. Ziem, Generalized Multilevel SQP-methods for PDAE-constrained Optimization Based on Space-Time Adaptive PDAE Solvers. Springer Basel, Basel (2012) 51–74.
- [16] A.R. Conn, N.I.M. Gould and P.L. Toint, Trust Region Methods. *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (2000).
- [17] L. Dedè, Reduced basis method and error estimation for parametrized optimal control problems with control constraints. *J. Sci. Comput.* **50** (2012) 287–305.
- [18] M.A. Dihlmann and B. Haasdonk, Certified PDE-constrained parameter optimization using reduced basis surrogate models for evolution problems. *Comput. Optim. Appl.* **60** (2015) 753–787.
- [19] M. Drohmann, B. Haasdonk and M. Ohlberger, Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation. *SIAM J. Sci. Comput.* **34** (2012) A937–A969.
- [20] J.L. Eftang, A.T. Patera and E.M. Rønquist, An “hp” certified reduced basis method for parametrized elliptic partial differential equations. *SIAM J. Sci. Comput.* **32** (2010) 3170–3200.
- [21] D. Garmatter, B. Haasdonk and B. Harrach, A reduced basis Landweber method for nonlinear inverse problems. *Inverse Prob.* **32** (2016) 035001.
- [22] C. Geuzaine and J.-F. Remacle, Gmsh: a 3-d finite element mesh generator with built-in pre-and post-processing facilities. *Int. J. Numer. Methods Eng.* **79** (2009) 1309–1331.
- [23] C. Gräßle, M. Gubisch, S. Metzdorf, S. Rogg and S. Volkwein, POD basis updates for nonlinear PDE control. *at – Automatisierungstechnik* **65** (2017) 298–307.
- [24] M.A. Grepl and M. Kärcher, Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems. *C. R. Math. Acad. Sci. Paris* **349** (2011) 873–877.
- [25] M. Gubisch and S. Volkwein, Proper orthogonal decomposition for linear-quadratic optimal control, edited by P. Benner, A. Cohen, M. Ohlberger and K. Willcox. In: *Model Reduction and Approximation: Theory and Algorithms*. SIAM, Philadelphia, PA (2017) 3–63.
- [26] B. Haasdonk, Convergence rates of the POD-greedy method. *ESAIM: M2AN* **47** (2013) 859–873.
- [27] B. Haasdonk, Reduced basis methods for parametrized PDEs: a tutorial introduction for stationary and instationary problems. *Model Reduction Approximation: Theory Algorithms* **15** (2017) 65.
- [28] B. Haasdonk, M. Dihlmann and M. Ohlberger, A training set and multiple bases generation approach for parameterized model reduction based on adaptive grids in parameter space. *Math. Comput. Model. Dyn. Syst.* **17** (2011) 423–442.
- [29] M. Heinkenschloss and L.N. Vicente, Analysis of inexact trust-region SQP algorithms. *SIAM J. Optim.* **12** (2002) 283–302.
- [30] J.S. Hesthaven, G. Rozza and B. Stamm, Certified Reduced Basis Methods for Parametrized Partial Differential Equations. SpringerBriefs in Mathematics. Springer International Publishing (2016).
- [31] C. Himpe and M. Ohlberger, Cross-gramian-based combined state and parameter reduction for large-scale control systems. *Math. Probl. Eng.* **13** (2014) 843869.
- [32] C. Himpe and M. Ohlberger, Data-driven combined state and parameter reduction for inverse problems. *Adv. Comput. Math.* **41** (2015) 1343–1364.

- [33] M. Hintermüller, M. Hinze and R.H.W. Hoppe, Weak-duality based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state-constraints. *J. Comput. Math.* **30** (2012) 101–123.
- [34] M. Hinze and S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. *Comput. Optim. App.* **39** (2008) 319–345.
- [35] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich, Optimization with PDE Constraints. Springer Netherlands (2009).
- [36] D.B.P. Huynh, G. Rozza, S. Sen and A.T. Patera, A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C. R. Math.* **345** (2007) 473–478.
- [37] E. Kammann, F. Tröltzsch and S. Volkwein, A posteriori error estimation for semilinear parabolic optimal control problems with application to model reduction by POD. *ESAIM: M2AN* **47** (2013) 555–581.
- [38] M. Kärcher, Z. Tokoutsis, M.A. Grepl and K. Veroy, Certified reduced basis methods for parametrized elliptic optimal control problems with distributed controls. *J. Sci. Comput.* **75** (2018) 276–307.
- [39] T. Keil, L. Mechelli, M. Ohlberger, F. Schindler and S. Volkwein, Software for NCD corrected TR-RB approach for PDE constrained optimization. (2020). DOI: [10.5281/zenodo.3897230](https://doi.org/10.5281/zenodo.3897230).
- [40] C.T. Kelley, Iterative Methods for Optimization. Society for Industrial and Applied Mathematics (1999).
- [41] C. Lieberman, K. Willcox and O. Ghattas, Parameter and state model reduction for large-scale statistical inverse problems. *SIAM J. Sci. Comput.* **32** (2010) 2523–2542.
- [42] W. Liu and N. Yan, A posteriori error estimates for distributed convex optimal control problems. *Adv. Comput. Math.* **15** (2001) 285–309.
- [43] R. Milk, S. Rave and F. Schindler, pyMOR – Generic Algorithms and Interfaces for Model Order Reduction. *SIAM J. Sci. Comput.* **38** (2016) S194–S216.
- [44] F. Negri, G. Rozza, A. Manzoni and A. Quarteroni, Reduced basis method for parametrized elliptic optimal control problems. *SIAM J. Sci. Comput.* **35** (2013) A2316–A2340.
- [45] J. Nocedal and S.J. Wright, Numerical Optimization, 2nd edition. *Springer Series in Operations Research and Financial Engineering*. Springer New York (2006).
- [46] M. Ohlberger and F. Schindler, Error control for the localized reduced basis multiscale method with adaptive on-line enrichment. *SIAM J. Sci. Comput.* **37** (2015) A2865–A2895.
- [47] M. Ohlberger and F. Schindler, Non-conforming localized model reduction with online enrichment: towards optimal complexity in PDE constrained optimization. In: Vol. 200 of *Finite Volumes for Complex Applications VIII – Hyperbolic, Elliptic and Parabolic Problems*. Springer Proc. Math. Stat. Springer, Cham (2017) 357–365.
- [48] M. Ohlberger, M. Schaefer and F. Schindler, Localized model reduction in PDE constrained optimization. *Int. Ser. Numer. Math.* **169** (2018) 143–163.
- [49] I.B. Oliveira and A.T. Patera, Reduced-basis techniques for rapid reliable optimization of systems described by affinely parametrized coercive elliptic partial differential equations. *Optim. Eng.* **8** (2007) 43–65.
- [50] E. Qian, M. Grepl, K. Veroy and K. Willcox, A certified trust region reduced basis approach to PDE-constrained optimization. *SIAM J. Sci. Comput.* **39** (2017) S434–S460.
- [51] A. Quarteroni, A. Manzoni and F. Negri, Reduced Basis Methods for Partial Differential Equations, 1st edition. La Matematica per il 3+2. Springer International Publishing (2016).
- [52] R. Rannacher, On the adaptive discretization of PDE-based optimization problems. In: PDE Constrained Optimization. Springer (2006).
- [53] S. Rogg, S. Trenz and S. Volkwein, Trust-region POD using a-posteriori error estimation for semilinear parabolic optimal control problems. Konstanzer Schriften in Mathematik No. 359 (2017). <http://nbn-resolving.de/urn:nbn:de:bsz:352-0-401106>.
- [54] A. Rösch and D. Wachsmuth, A-posteriori error estimates for optimal control problems with state and control constraints. *Numer. Math.* **120** (2012) 733–762.
- [55] 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 Eng.* **15** (2007) 1.
- [56] E.W. Sachs and M. Schu, A-priori error estimates for reduced order models in finance. *ESAIM: M2AN* **47** (2013) 449–469.
- [57] Y. Yue and K. Meerbergen, Accelerating optimization of parametric linear systems by model order reduction. *SIAM J. Optim.* **23** (2013) 1344–1370.
- [58] M.J. Zahr and C. Farhat, Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *Int. J. Numer. Methods Eng.* **102** (2015) 1111–1135.