A POSTERIORI ERROR ESTIMATION FOR REDUCED ORDER SOLUTIONS OF PARAMETRIZED PARABOLIC OPTIMAL CONTROL PROBLEMS [*]

MARK KÄRCHER[1] AND MARTIN A. GREPL[2]

Abstract. We consider the efficient and reliable solution of linear-quadratic optimal control problems governed by parametrized parabolic partial differential equations. To this end, we employ the reduced basis method as a low-dimensional surrogate model to solve the optimal control problem and develop a posteriori error estimation procedures that provide rigorous bounds for the error in the optimal control and the associated cost functional. We show that our approach can be applied to problems involving control constraints and that, even in the presence of control constraints, the reduced order optimal control problem and the proposed bounds can be efficiently evaluated in an offline-online computational procedure. We also propose two greedy sampling procedures to construct the reduced basis space. Numerical results are presented to confirm the validity of our approach.

October 1, 2013.

INTRODUCTION

Many problems in science and engineering can be modeled in terms of optimal control problems governed by parametrized partial differential equations (PDEs), see e.g. [15, 25, 26] for theoretical results and applications. While the PDE describes the underlying system or component behavior, the parameters often serve to identify a particular configuration of the component — such as boundary and initial conditions, material properties, and geometry. The solution of these problems using classical discretization techniques such as finite elements or finite volumes is sometimes computationally expensive and time-consuming. One way to decrease the computational burden is the surrogate model approach, where the original high-dimensional model is replaced by a reduced order approximation. These ideas have received a lot of attention in the past and various model order reduction techniques have been used in this context: proper orthogonal decomposition (POD) e.g. in [3, 23, 24, 34, 35], reduction based on inertial manifolds in [18], and reduced basis methods in [6, 7, 19, 20, 28, 34]; for a review of various model order reduction techniques we also refer to [2, 5]. However, the solution of the reduced order optimal control problem is generally suboptimal and reliable error estimation is thus crucial.

[1] Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Schinkelstraße 2, 52062 Aachen, Germany (`kaercher@aices.rwth-aachen.de`).

[2] Numerical Mathematics, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany (`grepl@igpm.rwth-aachen.de`).

In this paper we employ the reduced basis method [30, 32] as a surrogate model for the solution of optimal control problems. We extend our previous work in [10, 22] in the following two directions: First, we consider optimal control problems governed by *time-dependent (parabolic)* PDEs. To this end, we allow for multiple controls which are scalar functions of time. Second, we consider problems involving box *constraints on the controls*, i.e., upper and lower bounds. We develop rigorous *a posteriori* error bounds for the optimal control *and* the associated cost functional and show that the reduced order optimal control problem *and* error bounds can be efficiently evaluated in an offline-online computational procedure. We note that the efficient, real-time solution of optimal control problems is essential in model predictive control of PDEs, see e.g. [1, 17].

A posteriori error bounds for reduced order solutions of optimal control problems have been proposed for proper orthogonal decomposition (POD) and reduced basis surrogate models in [35] and [6, 7, 28], respectively. In [35], the authors estimate the distance between the computed suboptimal control and the unknown optimal control using a perturbation argument proposed in [14, 27]. The approach allows to use the POD approximation to efficiently solve the optimal control problem. The evaluation of the *a posteriori* error bounds, however, requires a forward-backward solution of the underlying high-dimensional state and adjoint equations and, as pointed out in [35], is thus computationally expensive. In [6, 7], reduced basis approximations and associated *a posteriori* error estimation procedures have been derived to estimate the error in the optimal value of the cost functional. However, although the estimator is efficient to evaluate, it is not a rigorous upper bound for the error. Recently, a reduced basis approach to optimal control problems based on a saddle-point formulation has been considered in [28]. The resulting *a posteriori* error bound follows directly from previous work on reduced basis methods for saddle point problems [33]. However, the saddle point theory only provides a combined bound for the error in the state, adjoint, and control variable. Furthermore, the approach is only applicable to optimal control problems without control constraints involving stationary (time-independent) PDEs. The former restriction is due to the fact that the results from [33] do not apply to variational inequalities, the latter since calculating the stability constant of the space-time saddle point problem, i.e., (a lower bound to) the Babuška inf-sup constant, is computationally prohibitive.

This paper is organized as follows. We introduce the optimal control problem in Section 1: we start with the general problem statement, state the first order optimality conditions, and illustrate how the reduced basis approximation can be used as a surrogate model. In Section 2 we turn to the *a posteriori* error estimation and develop bounds for the optimal control and the associated cost functional. Finally, we present numerical results for a model problem in Section 3 and offer concluding remarks in Section 4.

1. Optimal Control Problem

1.1. **Preliminaries**

Let Y_{e} with $\mathrm{H}_0^1(\Omega) \subset Y_{\mathrm{e}} \subset \mathrm{H}^1(\Omega)$ be a Hilbert space over the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d, d = 1, 2, 3$, with boundary Γ.[1] The inner product and induced norm associated with Y_{e} are given by $(\cdot, \cdot)_{Y_{\mathrm{e}}}$ and $\|\cdot\|_{Y_{\mathrm{e}}} = \sqrt{(\cdot, \cdot)_{Y_{\mathrm{e}}}}$, respectively. We assume that the norm $\|\cdot\|_{Y_{\mathrm{e}}}$ is equivalent to the $\mathrm{H}^1(\Omega)$-norm and denote the dual space of Y_{e} by Y_{e}'. We also recall the Hilbert space $W(0, T) = \{v \in L^2(0, T; Y_{\mathrm{e}}) : v_t \in L^2(0, T; Y_{\mathrm{e}}')\}$ for a fixed final time T with its standard inner product, see for example [31]. We also introduce the control space $\mathcal{U}_{\mathrm{e}} = L^2(0, T; \mathbb{R}^m), m \in \mathbb{N}$, together with its inner product $(w, v)_{\mathcal{U}_{\mathrm{e}}} = \int_0^T (w(t), v(t))_{\mathbb{R}^m} \mathrm{d}t$, induced norm $\|\cdot\|_{\mathcal{U}_{\mathrm{e}}} = \sqrt{(\cdot, \cdot)_{\mathcal{U}_{\mathrm{e}}}}$, and associated dual space $\mathcal{U}_{\mathrm{e}}'$. Furthermore, let $\mathcal{D} \subset \mathbb{R}^P$ be a prescribed P-dimensional compact parameter set in which our P-tuple (input) parameter $\mu = (\mu_1, \ldots, \mu_P)$ resides.

We next introduce the (for the sake of simplicity) parameter-independent bilinear form $m(w, v) = (w, v)_{L^2(\Omega)}$, $\forall w, v \in L^2(\Omega)$, and the parameter-dependent bilinear form $a(\cdot, \cdot; \mu) : Y_{\mathrm{e}} \times Y_{\mathrm{e}} \to \mathbb{R}$. We shall assume that $a(\cdot, \cdot; \mu)$ is continuous,

$$\gamma_{\mathrm{e}}(\mu) = \sup_{w \in Y_{\mathrm{e}} \backslash \{0\}} \sup_{v \in Y_{\mathrm{e}} \backslash \{0\}} \frac{a(w, v; \mu)}{\|w\|_{Y_{\mathrm{e}}} \|v\|_{Y_{\mathrm{e}}}} \le \gamma_0 < \infty, \quad \forall \mu \in \mathcal{D}, \tag{1.1}$$

---

[1]The subscript e denotes "exact".

coercive,

$$\alpha_{\mathrm{e}}(\mu) = \inf_{v \in Y_{\mathrm{e}} \backslash \{0\}} \frac{a(v, v; \mu)}{\|v\|_{Y_{\mathrm{e}}}^2} \geq \alpha_0 > 0, \quad \forall \mu \in \mathcal{D}, \tag{1.2}$$

and affinely parameter-dependent,

$$a(w, v; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(w, v), \quad \forall w, v \in Y_{\mathrm{e}}, \quad \forall \mu \in \mathcal{D}, \tag{1.3}$$

for some (preferably) small integer Q_a. Here, the coefficient functions $\Theta_a^q : \mathcal{D} \to \mathbb{R}$ are continuous and depend on μ, but the continuous bilinear forms a^q do *not* depend on μ. We also introduce the continuous and linear operator $\mathcal{B}_{\mathrm{e}} : \mathcal{U}_{\mathrm{e}} \to L^2(0, T; Y_{\mathrm{e}}')$, given by

$$\langle (\mathcal{B}_{\mathrm{e}} u_{\mathrm{e}})(t), \cdot \rangle_{Y_{\mathrm{e}}', Y_{\mathrm{e}}} = \sum_{i=1}^{m} b_i(\cdot) u_{\mathrm{e},i}(t), \tag{1.4}$$

where $\langle \cdot, \cdot \rangle_{Y_{\mathrm{e}}', Y_{\mathrm{e}}}$ denotes the dual pairing between Y_{e}' and Y_{e}, $b_1, \ldots, b_m$ are given bounded linear functionals on $L^2(\Omega)$ and $u_{\mathrm{e}} \in \mathcal{U}_{\mathrm{e}}$ is the control with time-dependent control components $u_{\mathrm{e},i} \in L^2(0, T)$, $1 \leq i \leq m$. For simplicity, we assume that the functionals $b_1, \ldots, b_m$ do not depend on the parameter; however, (affine) parameter dependence of the b_i and thus of the operator $\mathcal{B}_{\mathrm{e}}$ itself is readily admitted [11]. Finally, we require that all linear and bilinear forms are independent of time – the system is thus linear time-invariant (LTI).

1.2. General Problem Statement

We consider the parametrized optimal control problem

$$\min J_{\mathrm{e}}(y_{\mathrm{e}}, u_{\mathrm{e}}; \mu) \quad \text{s.t.} \quad (y_{\mathrm{e}}, u_{\mathrm{e}}) \in W(0, T) \times \mathcal{U}_{\mathrm{e,ad}} \quad \text{solves}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} m(y_{\mathrm{e}}(t), v) + a(y_{\mathrm{e}}(t), v; \mu) = \langle (\mathcal{B}_{\mathrm{e}} u_{\mathrm{e}})(t), v \rangle_{Y_{\mathrm{e}}', Y_{\mathrm{e}}}, \quad \forall v \in Y_{\mathrm{e}}, \quad \text{f.a.a. } t \in (0, T], \tag{P_{e}}$$

with initial condition $y_{\mathrm{e}}(0) = y_0 \equiv 0$; the quadratic cost functional, $J_{\mathrm{e}}(\cdot, \cdot; \mu) : W(0, T) \times \mathcal{U}_{\mathrm{e}} \to \mathbb{R}$, is given by

$$J_{\mathrm{e}}(y_{\mathrm{e}}, u_{\mathrm{e}}; \mu) = \frac{\sigma_1}{2} \int_0^T \|y_{\mathrm{e}} - y_{d,\mathrm{e}}(\mu)\|_{L^2(D)}^2 \mathrm{d}t + \frac{\sigma_2}{2} \|y_{\mathrm{e}}(T) - y_{d,\mathrm{e}}(T; \mu)\|_{L^2(D)}^2 + \frac{\lambda}{2} \|u_{\mathrm{e}} - u_{d,\mathrm{e}}\|_{\mathcal{U}_{\mathrm{e}}}^2. \tag{1.5}$$

Here, $D \subset \Omega$ is a measurable set; $y_{d,\mathrm{e}}(\mu) \in L^2(0, T; L^2(D))$ and $u_{d,\mathrm{e}} \in \mathcal{U}_{\mathrm{e}}$ are the desired state and control, respectively; and $\lambda > 0$ and $\sigma_1, \sigma_2 \geq 0$ are given regularization parameters governing the trade-off between the cost associated with the deviation from the desired control and the desired state (over the whole time trajectory and/or at the final time), respectively. We assume that the parameter-dependent desired state $y_{d,\mathrm{e}}(\mu)$ admits the affine representation

$$y_{d,\mathrm{e}}(x, t; \mu) = \sum_{q=1}^{Q_{yd}} \Theta_{yd}^q(t; \mu) y_{d,\mathrm{e}}^q(x), \tag{1.6}$$

with parameter-dependent and time-dependent coefficient functions $\Theta_{yd}^q : [0, T] \times \mathcal{D} \to \mathbb{R}$ and parameter-independent functions $y_{d,\mathrm{e}}^q \in L^2(D)$. For simplicity, we assume that the desired control $u_{d,\mathrm{e}}$ is parameter-independent; however, (affine) parameter dependence is readily admitted. We also introduce the non-empty convex subset of admissible controls

$$\mathcal{U}_{\mathrm{e,ad}} = \{ u_{\mathrm{e}} \in \mathcal{U}_{\mathrm{e}} : u_{a,\mathrm{e}}(t) \leq u_{\mathrm{e}}(t) \leq u_{b,\mathrm{e}}(t) \} \subset \mathcal{U}_{\mathrm{e}}, \tag{1.7}$$

where $u_{a,e}$, $u_{b,e} \in L^2(0, T; \mathbb{R}^m)$, with $u_{a,e}(t) \leq u_{b,e}(t)$, f.a.a. $t \in [0, T]$, are given lower and upper bounds for the control components and the inequalities are interpreted component-wise in $\mathbb{R}^m$. It follows from our assumptions that there exists a unique optimal solution (y_e^*, u_e^*) to (P$_e$) [26].

Employing a Lagrangian approach we obtain the first-order optimality system consisting of the state equation, the adjoint equation, and the optimality condition: Given $\mu \in \mathcal{D}$, the optimal solution $(y_e^*, p_e^*, u_e^*) \in W(0, T) \times W(0, T) \times \mathcal{U}_{e,\mathrm{ad}}$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} m(y_e^*(t), \phi) + a(y_e^*(t), \phi; \mu) = \langle (\mathcal{B}_e u_e^*)(t), \phi \rangle_{Y_e', Y_e}, \qquad \forall \phi \in Y_e, \quad \text{f.a.a. } t \in (0, T], \qquad (1.8a)$$

$$y_e^*(0) = y_0 \equiv 0 \qquad (1.8b)$$

$$-\frac{\mathrm{d}}{\mathrm{d}t} m(\varphi, p_e^*(t)) + a(\varphi, p_e^*(t); \mu) = \sigma_1(y_{d,e}(t; \mu) - y_e^*(t), \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y_e, \quad \text{f.a.a. } t \in [0, T), \qquad (1.8c)$$

$$m(\varphi, p_e^*(T)) = \sigma_2(y_{d,e}(T; \mu) - y_e^*(T), \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y_e, \qquad (1.8d)$$

$$\left(\lambda(u_e^* - u_{d,e}) - \mathcal{B}_e^\top p_e^*, \psi - u_e^* \right)_{\mathcal{U}_e} \geq 0, \qquad \forall \psi \in \mathcal{U}_{e,\mathrm{ad}}. \qquad (1.8e)$$

Here, p_e is the adjoint variable and the superscript * denotes optimality. Furthermore, the linear and bounded dual operator of $\mathcal{B}_e$ in (1.8e) is given by $\mathcal{B}_e^\top : L^2(0, T; Y_e) \to \mathcal{U}_e$, where we identify $(L^2(0, T; Y_e'))'$ with $L^2(0, T; Y_e)$ and $\mathcal{U}_e'$ with $\mathcal{U}_e$. From the relationship

$$\langle \mathcal{B}_e u, \phi \rangle_{L^2(0, T; Y)', L^2(0, T; Y)} = \int_0^T \sum_{i=1}^m b_i(\phi(t)) u_i(t) \mathrm{d}t = \int_0^T \left(u(t), (\mathcal{B}_e^\top \phi)(t) \right)_{\mathbb{R}^m} \mathrm{d}t = (u, \mathcal{B}_e^\top \phi)_{\mathcal{U}_e} \qquad (1.9)$$

it follows that, for given $\phi \in L^2(0, T; Y_e)$, the dual operator $\mathcal{B}_e^\top \phi$ can be expressed as

$$(\mathcal{B}_e^\top \phi)_i(t) = b_i(\phi(t)), \quad 1 \leq i \leq m, \quad t \in [0, T]. \qquad (1.10)$$

We note that for the linear-quadratic optimal control problem (P$_e$) the first-order conditions (1.8) are necessary and sufficient for the optimality of (y_e^*, u_e^*) [26].

Note that we consider here the classical initial value problem formulation to extend our earlier elliptic work [10,22] to the parabolic case. Another approach worth pursuing is a space-time formulation of the optimal control problem [12]; also see [36] for a reduced basis space-time formulation for parabolic problems. However, in the context of the current paper — spatially coercive operators — there would be little quantitative improvement in the results (cf. the very good performance of the energy-norm bounds in Table 1).

In practice, the regularization parameters often serve as design parameters which are tuned to achieve a desired performance of the optimal controller. From a reduced basis point of view, however, the regularization parameters may simply be considered input parameters of the parametrized optimal control problem. This allows us to vary (say) λ online and thus to efficiently design the optimal controller with the approach presented here. For our model problem in Section 3 we will in fact consider λ as an additional (reduced basis) input parameter.

1.3. **Truth Approximation**

In general, we of course cannot expect to find an analytic solution to (1.8). We thus consider a temporal and spatial "truth" discretization: we divide the time interval $[0, T]$ into K subintervals of equal length $\tau = \frac{T}{K}$ and define $t^k = k\tau, 0 \leq k \leq K$, and $\mathbb{K} = \{1, \ldots, K\}$; we also introduce a finite element approximation space $Y \subset Y_e$ of typically very large dimension $\mathcal{N}$. Note that Y shall inherit the inner product and norm from Y_e: $(\cdot, \cdot)_Y = (\cdot, \cdot)_{Y_e}$ and $\|\cdot\|_Y = \|\cdot\|_{Y_e}$. Clearly, the continuity and coercivity properties of the bilinear form a are

inherited by the truth approximation, i.e.,

$$\gamma(\mu) = \sup_{w \in Y \setminus \{0\}} \sup_{v \in Y \setminus \{0\}} \frac{a(w,v;\mu)}{\|w\|_Y \|v\|_Y} \le \gamma_e(\mu) \le \gamma_0 < \infty, \quad \forall \mu \in \mathcal{D}, \tag{1.11}$$

and

$$\alpha(\mu) = \inf_{v \in Y \setminus \{0\}} \frac{a(v,v;\mu)}{\|v\|_Y^2} \ge \alpha_e(\mu) \ge \alpha_0 > 0, \quad \forall \mu \in \mathcal{D}. \tag{1.12}$$

We also define the operator $\mathcal{B} : \mathcal{U} = (\mathbb{R}^m)^K \to (Y')^K$ by

$$\langle (\mathcal{B}u)_k, \cdot \rangle_{Y',Y} = \sum_{i=1}^m b_i(\cdot) u_i^k, \quad k \in \mathbb{K}, \tag{1.13}$$

where $\mathcal{U} = (\mathbb{R}^m)^K$ is the discretized control space with inner product $(u,v)_{\mathcal{U}} = \tau \sum_{k=1}^K (u^k, v^k)_{\mathbb{R}^m}$ and induced norm $\|\cdot\|_{\mathcal{U}} = \sqrt{(\cdot,\cdot)_{\mathcal{U}}}$. Here, a control is denoted by $u = (u^1, \ldots, u^K), u^k \in \mathbb{R}^m$, such that $u_i^k \in \mathbb{R}$ corresponds to the i-th control input at time t^k, i.e., $u_i^k = u_i(t^k)$.

We thus obtain the corresponding truth optimal control problem[2]

$$\min J(y,u;\mu) \quad \text{s.t.} \quad (y,u) \in Y^K \times \mathcal{U}_{\mathrm{ad}} \quad \text{solves}$$
$$m(y^k, v) + \tau\, a(y^k, v; \mu) = m(y^{k-1}, v) + \tau \langle (\mathcal{B}u)_k, v \rangle_{Y',Y}, \quad \forall v \in Y, \quad \forall k \in \mathbb{K}, \tag{P}$$

with initial condition $y^0 = y_0 = 0$. Here, y^k denotes the truth solution at time t^k and the discretized cost functional $J(\cdot, \cdot; \mu) : Y^K \times \mathcal{U} \to \mathbb{R}$ is given by

$$J(y,u;\mu) = \frac{\sigma_1}{2}\tau \sum_{k=1}^K \|y^k - y_d^k(\mu)\|_{L^2(D)}^2 + \frac{\sigma_2}{2}\|y^K - y_d^K(\mu)\|_{L^2(D)}^2 + \frac{\lambda}{2}\tau \sum_{k=1}^K \|u^k - u_d^k\|_{\mathbb{R}^m}^2, \tag{1.14}$$

$u_d^k = u_{d,e}(t^k) \in \mathbb{R}^m$, $y_d^k(\mu) \in Y$ is the L^2–projection of $y_{d,e}(t^k; \mu)$, and the discretized admissible control set is $\mathcal{U}_{\mathrm{ad}} = \{u \in \mathcal{U} : u_a^k \le u^k \le u_b^k, \ k \in \mathbb{K}\}$, where $u_a^k = u_{a,e}(t^k) \in \mathbb{R}^m$ and $u_b^k = u_{b,e}(t^k) \in \mathbb{R}^m$.

The associated first-order optimality system reads: Given $\mu \in \mathcal{D}$, the optimal solution $(y^*, p^*, u^*) \in Y^K \times Y^K \times \mathcal{U}_{\mathrm{ad}}$ satisfies

$$m(y^{*,k} - y^{*,k-1}, \phi) + \tau\, a(y^{*,k}, \phi; \mu) = \tau \langle (\mathcal{B}u^*)_k, \phi \rangle_{Y',Y}, \qquad \forall \phi \in Y, \quad \forall k \in \mathbb{K}, \tag{1.15a}$$

$$y^{*,0} = y_0 \equiv 0, \tag{1.15b}$$

$$m(\varphi, p^{*,k} - p^{*,k+1}) + \tau\, a(\varphi, p^{*,k}; \mu) = \tau \sigma_1(y_d^k(\mu) - y^{*,k}, \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y, \quad \forall k \in \mathbb{K}, \tag{1.15c}$$

$$m(\varphi, p^{*,K+1}) = \sigma_2(y_d^K(\mu) - y^{*,K}, \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y, \tag{1.15d}$$

$$\left(\lambda(u^* - u_d) - \mathcal{B}^\top p^*, \psi - u^*\right)_{\mathcal{U}} \ge 0, \qquad \forall \psi \in \mathcal{U}_{\mathrm{ad}}, \tag{1.15e}$$

where the dual operator $\mathcal{B}^\top : Y^K \to \mathcal{U}$ is given for $\phi \in Y^K$ by $(\mathcal{B}^\top \phi)_i^k = b_i(\phi^k)$, which is the discrete counterpart to the dual operator $\mathcal{B}_e^\top$ defined in (1.10). We further note that the Ansatz and test spaces are identical for the state and adjoint equations. This ensures that the solution of the optimality system (1.15) is indeed also an optimal solution of the truth optimal control problem (P).

The optimality system (1.15) constitutes a coupled set of equations (and variational inequalities) of dimension $2K\mathcal{N} + Km$ and is thus expensive to solve, especially if one is interested in various values of $\mu \in \mathcal{D}$. Our goal

---

[2]We shall employ the backward Euler method for the time integration; we note, however, that the treatment of higher-order schemes such as Crank-Nicolson (or a general Θ-scheme) is also possible.

is therefore to significantly speed up the solution of (1.15) by employing the reduced basis approximation as a surrogate model for the PDE constraint in (P).

1.4. Reduced Basis Approximation

We first assume that we are given the reduced basis spaces $Y_N = \text{span}\{\zeta_n, \ 1 \leq n \leq N\}, \ 1 \leq N \leq N_{\max}$, where the $\zeta_n, \ 1 \leq n \leq N$, are mutually $(\cdot, \cdot)_Y$-orthogonal basis functions and $N, N_{\max}$ are even. We comment on the POD/Greedy sampling procedure to construct the spaces Y_N in Section 2.4.

We next replace the truth approximation of the PDE constraint in (P) with its reduced basis approximation. The reduced basis optimal control problem is thus given by

$$\begin{aligned} \min J(y_N, u_N; \mu) \quad \text{s.t.} \quad (y_N, u_N) &\in Y_N^K \times \mathcal{U}_{\text{ad}} \quad \text{solves} \\ m(y_N^k, v) + \tau \, a(y_N^k, v; \mu) &= m(y_N^{k-1}, v) + \tau \langle (\mathcal{B}u_N)_k, v \rangle_{Y', Y}, \quad \forall v \in Y_N, \quad \forall k \in \mathbb{K}, \end{aligned} \tag{P_N}$$

with initial condition $y_N^0 = y_0 = 0$. Notice that the already low-dimensional truth control space $\mathcal{U}$ is not reduced, although this is possible for problems with high-dimensional control spaces, e.g., problems with distributed controls over parts of the domain or its boundary [21, 28].

We can also directly state the associated first-order optimality system: Given $\mu \in \mathcal{D}$, find $(y_N^*, p_N^*, u_N^*) \in Y_N^K \times Y_N^K \times \mathcal{U}_{\text{ad}}$ such that

$$m(y_N^{*,k} - y_N^{*,k-1}, \phi) + \tau \, a(y_N^{*,k}, \phi; \mu) = \tau \, \langle (\mathcal{B}u_N^*)_k, \phi \rangle_{Y', Y}, \qquad \forall \phi \in Y_N, \quad \forall k \in \mathbb{K}, \tag{1.16a}$$

$$y_N^{*,0} = y_0 \equiv 0, \tag{1.16b}$$

$$m(\varphi, p_N^{*,k} - p_N^{*,k+1}) + \tau \, a(\varphi, p_N^{*,k}; \mu) = \tau \, \sigma_1 (y_d^k(\mu) - y_N^{*,k}, \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y_N, \quad \forall k \in \mathbb{K}, \tag{1.16c}$$

$$m(\varphi, p_N^{*,K+1}) = \sigma_2 (y_d^K(\mu) - y_N^{*,K}, \varphi)_{L^2(D)}, \qquad \forall \varphi \in Y_N, \tag{1.16d}$$

$$\left(\lambda (u_N^* - u_d) - \mathcal{B}^\top p_N^*, \psi - u_N^* \right)_{\mathcal{U}} \geq 0, \qquad \forall \psi \in \mathcal{U}_{\text{ad}}. \tag{1.16e}$$

The reduced basis optimality system is only of dimension $2KN + Km$ and can be evaluated efficiently using an offline-online computational decomposition.

We note that we use a single "integrated" reduced basis Ansatz and test space for the state and adjoint equations. The reason is twofold: first, the reduced basis optimality system (1.16) reflects the reduced basis optimal control problem (P_N) only if the spaces of the state and adjoint equations are identical; and second, using different spaces may result in an unstable system (1.16). This issue is closely related to the stability of reduced basis formulations for saddle point problems, see [9] for details. If we use the same space Y_N for the state *and* the adjoint equation, on the other hand, the system (1.16) is provably stable. Finally, since the state and adjoint solutions need to be well-approximated using the single space Y_N, we choose "integrated" spaces, i.e., we integrate both snapshots of the state and adjoint equations into the reduced basis space Y_N.

1.5. Computational Procedure

We now turn to the computational details of the reduced basis approximation of the optimality system. To this end, we express the reduced basis state and adjoint solutions as $y_N^k(\mu) = \sum_{i=1}^N y_{Ni}^k(\mu)\zeta_i$ and $p_N^k(\mu) = \sum_{i=1}^N p_{Ni}^k(\mu)\zeta_i$ and denote the coefficient vectors by $\underline{y}_N^k(\mu) = [y_{N1}^k(\mu), \ldots, y_{NN}^k(\mu)]^T \in \mathbb{R}^N$ and $\underline{p}_N^k(\mu) = [p_{N1}^k(\mu), \ldots, p_{NN}^k(\mu)]^T \in \mathbb{R}^N$, respectively. If we choose as test functions $\phi = \zeta_i, \ 1 \leq i \leq N$, and $\varphi =$

ζ_i, $1 \le i \le N$, in (1.16), the reduced basis optimality system can be written as

$$(M_N + \tau A_N(\mu))\, \underline{y}_N^k(\mu) = M_N\, \underline{y}_N^{k-1}(\mu) + \tau B_N u_N^k, \qquad\qquad\qquad k \in \mathbb{K}, \qquad\qquad (1.17a)$$

$$\underline{y}_N^0(\mu) = 0, \qquad\qquad\qquad\qquad (1.17b)$$

$$(M_N + \tau A_N(\mu))\, \underline{p}_N^k(\mu) = M_N\, \underline{p}_N^{k+1}(\mu) + \tau\sigma_1\,(Y_{d,N}^k(\mu) - D_N\, \underline{y}_N^k(\mu)), \qquad k \in \mathbb{K}, \qquad\qquad (1.17c)$$

$$M_N\, \underline{p}_N^{K+1}(\mu) = \sigma_2\,(Y_{d,N}^K(\mu) - D_N\, \underline{y}_N^K(\mu)), \qquad\qquad\qquad (1.17d)$$

$$(\lambda(u_N^k - u_d^k) - B_N^T \underline{p}_N^k(\mu), \psi^k - u_N^k)_{\mathcal{U}} \ge 0, \qquad\qquad\qquad \forall \psi^k \in \mathcal{U}_{\mathrm{ad}}^k, \quad k \in \mathbb{K}. \qquad (1.17e)$$

Here, $M_N \in \mathbb{R}^{N \times N}$, $B_N \in \mathbb{R}^{N \times m}$, and $D_N \in \mathbb{R}^{N \times N}$ are matrices with entries $(M_N)_{ij} = m(\zeta_j, \zeta_i)$, $1 \le i, j \le N$, $(B_N)_{ij} = b_j(\zeta_i)$, $1 \le i \le N$, $1 \le j \le m$, and $(D_N)_{ij} = (\zeta_j, \zeta_i)_{L^2(D)}$, $1 \le i, j \le N$ respectively. Invoking the affine parameter dependence (1.3) yields the expansion $A_N(\mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu) A_N^q$, where the parameter-independent matrices $A_N^q \in \mathbb{R}^{N \times N}$ are given by $(A_N^q)_{ij} = a^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N$, $1 \le q \le Q_a$. Similarly, from (1.6) we obtain the affine representation of $Y_{d,N}^k(\mu) \in \mathbb{R}^N$, $k \in \mathbb{K}$, as $Y_{d,N}^k(\mu) = \sum_{q=1}^{Q_{yd}} \Theta_{yd}^q(t^k; \mu) Y_{d,N}^q$, where the parameter-independent vectors $Y_{d,N}^q \in \mathbb{R}^N$ are given by $(Y_{d,N}^q)_i = (y_d^q(x), \zeta_i)_{L^2(D)}$, $1 \le i \le N$, $1 \le q \le Q_{yd}$. Finally, to allow an efficient evaluation of the cost functional in the online stage, we also compute and store the matrix $\tilde{Y}_d \in \mathbb{R}^{Q_{yd} \times Q_{yd}}$, given by $(\tilde{Y}_d)_{p,q} = (y_d^p, y_d^q)_{L^2(D)}$.

The offline-online decomposition is now clear. In the offline stage — performed only *once* — we first construct the reduced basis space Y_N. We then assemble the parameter-independent quantities A_N^q, $1 \le q \le Q_a$, M_N, D_N, B_N, $Y_{d,N}^q$, $1 \le q \le Q_{yd}$, and $\tilde{Y}_d$. The computational cost clearly depends on the truth finite element dimension $\mathcal{N}$. In the online stage — for each new parameter value μ — we first assemble the parameter-dependent quantities $A_N(\mu)$ and $Y_{d,N}^k(\mu), k \in \mathbb{K}$, in $\mathcal{O}(Q_a N^2)$ and $\mathcal{O}(Q_{yd} K N)$ operations, respectively. We then solve the reduced basis optimality system (1.17) iteratively with a BFGS Quasi-Newton method. The cost for one BFGS-iteration on the "reduced" cost functional $j_N(u_N; \mu) := J(y_N(u_N), u_N; \mu)$ is to leading order $\mathcal{O}(N^3 + KN^2 + KNm + (Km)^2)$. In the control constrained case we use the primal dual active set method (PDAS) resulting in an outer loop around the BFGS iteration. In our numerical tests we needed at most 4 PDAS iterations.

Given a reduced basis approximation, the cost functional can be evaluated efficiently from

$$J(y_N, u_N; \mu) = \frac{\sigma_1}{2}\tau \sum_{k=1}^{K} \left((\underline{y}_N^k)^T D_N \underline{y}_N^k - 2(Y_{d,N}^k(\mu))^T \underline{y}_N^k + \sum_{p,q=1}^{Q_{yd}} \Theta_{yd}^p(t^k; \mu)\Theta_{yd}^q(t^k; \mu)(\tilde{Y}_d)_{p,q} \right)$$

$$+ \frac{\sigma_2}{2}\left((\underline{y}_N^K)^T D_N \underline{y}_N^K - 2(Y_{d,N}^K(\mu))^T \underline{y}_N^K + \sum_{p,q=1}^{Q_{yd}} \Theta_{yd}^p(t^K; \mu)\Theta_{yd}^q(t^K; \mu)(\tilde{Y}_d)_{p,q} \right) + \frac{\lambda}{2}\tau \sum_{k=1}^{K} \|u_N^k - u_d^k\|_{\mathbb{R}^m}^2 \quad (1.18)$$

in (to leading order) $\mathcal{O}(KN^2 + KQ_{yd}N + Q_{yd}^2 + Km)$ operations.

Hence, the computational cost for the online stage is independent of $\mathcal{N}$, the dimension of the underlying "truth" finite element approximation space. Since $N \ll \mathcal{N}$, we expect significant computational savings in the online stage relative to the solution of (1.15). However, we need to rigorously and efficiently assess the error introduced.

2. A Posteriori Error Estimation

We will now develop *a posteriori* error bounds for the error in the optimal control and the error in the associated cost functional. We discuss the control and cost functional bounds in Sections 2.1 and Section 2.2, respectively. We summarize the computational procedure in Section 2.3 and discuss the greedy algorithm to generate Y_N in Section 2.4.

2.1. **Error Bound for the Optimal Control**

The point of departure for our bound is a result from [35], where the authors estimate the distance between the computed POD suboptimal control and the unknown truth optimal control using a perturbation argument proposed in [14, 27]. The idea is to introduce a perturbation function $\zeta \in \mathcal{U}$ such that the RB optimal control u_N^*, i.e., the perturbed control, satisfies the optimality condition

$$(\lambda(u_N^* - u_d) - \mathcal{B}^\top \tilde{p} + \zeta, \psi - u_N^*)_\mathcal{U} \geq 0, \quad \forall \psi \in \mathcal{U}_{\mathrm{ad}}, \tag{2.1}$$

of a perturbed optimal control problem. Here, $\tilde{p} = p(y(u_N^*))$ is the solution of the (truth) adjoint equation (1.15c) with $y(u_N^*)$ instead of $y^*(u^*)$ on the right-hand side, and $\tilde{y} = y(u_N^*)$ is the solution of the (truth) state equation (1.15a) with control u_N^* instead of u^*. It is then possible to explicitly construct ζ in terms of u_N^*, $\tilde{y}$, and $\tilde{p}$ such that (2.1) holds and to bound the error in the optimal control, $u^* - u_N^*$, in terms of ζ. More specifically, for notational convenience we define

$$\xi = \lambda(u_N^* - u_d) - \mathcal{B}^\top \tilde{p} \tag{2.2}$$

with components ξ_i^k, $i = 1, \ldots, m$, $k \in \mathbb{K}$. We then distinguish for every component ζ_i^k, $i = 1, \ldots, m$, $k \in \mathbb{K}$, the following three cases to construct ζ (see [35] for a more detailed discussion):

(1) If $u_{N,i}^{*,k} = u_{a,i}^k$, then $\psi_i^k - u_{N,i}^{*,k} \geq 0$ for all $\psi_i^k \in \mathcal{U}_{\mathrm{ad},i}^k$ and hence $\xi_i^k + \zeta_i^k \geq 0$ has to hold: we set $\zeta_i^k = \left[\xi_i^k\right]_-$.

(2) If $u_{N,i}^{*,k} = u_{b,i}^k$, then $\psi_i^k - u_{N,i}^{*,k} \leq 0$ for all $\psi_i^k \in \mathcal{U}_{\mathrm{ad},i}^k$ and hence $\xi_i^k + \zeta_i^k \leq 0$ has to hold: we set $\zeta_i^k = -\left[\xi_i^k\right]_+$.

(3) If $u_{N,i}^{*,k} \in \left(u_{a,i}^k, u_{b,i}^k\right)$, then $\psi_i^k - u_{N,i}^{*,k}$ can attain positive and negative values for $\psi_i^k \in \mathcal{U}_{\mathrm{ad},i}^k$ and hence $\xi_i^k + \zeta_i^k = 0$ has to hold: we set $\zeta_i^k = -\xi_i^k$.

Here, $[\cdot]_+$ and $[\cdot]_-$ denote the positive and negative part functions, respectively. To summarize, $\zeta \in \mathcal{U}$ is given component-wise by

$$\zeta_i^k = \begin{cases} \left[\xi_i^k\right]_- & \text{if } u_{N,i}^{*,k} = u_{a,i}^k; \\ -\left[\xi_i^k\right]_+ & \text{if } u_{N,i}^{*,k} = u_{b,i}^k; \\ -\xi_i^k & \text{if } u_{N,i}^{*,k} \in \left(u_{a,i}^k, u_{b,i}^k\right). \end{cases} \tag{2.3}$$

We can now state the main result (see Theorem 4.11 in [35]).

Theorem 2.1. *Let u^* and u_N^* be the optimal solutions of the truth and reduced basis optimal control problems* (P) *and* (P$_N$), *respectively. The error in the optimal control then satisfies*

$$\|u^* - u_N^*\|_\mathcal{U} \leq \frac{1}{\lambda} \|\zeta\|_\mathcal{U}, \quad \forall \mu \in \mathcal{D}. \tag{2.4}$$

We note that evaluation of ζ and thus the bound in (2.4) is computationally expensive since it requires a forward-backward solve of the truth state and adjoint equations. Our goal is to develop a bound which can be computed using an offline-online decomposition such that the computational cost for the online-evaluation is independent of $\mathcal{N}$. The main idea is to replace the truth approximation $p(y(u_N^*))$ in (2.4) with the reduced basis approximation $p_N^*(y_N^*(u_N^*))$ and to bound the error term $p(y(u_N^*)) - p_N^*(y_N^*(u_N^*))$.

Before we continue, let us make some notational remarks. Following the notation and terminology in [6], we refer to $\tilde{e}^{y,k} = y^k(u_N^*) - y_N^{*,k}(u_N^*)$ as the state predictability error and to $\tilde{e}^{p,k} = p^k(y(u_N^*)) - p_N^{*,k}(y_N^*(u_N^*))$ as the adjoint predictability error. They reflect the ability of the corresponding reduced basis solutions to approximate the truth state and adjoint solutions for a prescribed control. In contrast, we define the state, adjoint, and control optimality errors as $e^{y,*,k} = y^{*,k}(u^*) - y_N^{*,k}(u_N^*)$, $e^{p,*,k} = p^{*,k}(y^*(u^*)) - p_N^{*,k}(y_N^*(u_N^*))$, and $e^{u,*} = u^* - u_N^*$, respectively. We start with the following definition.

Definition 2.2. The residuals for the state equation, the adjoint equation, and the optimality condition are defined by

$$r^{y,k}(\phi;\mu) = \langle (\mathcal{B}u_N^*)_k, \phi \rangle_{Y',Y} - a(y_N^{*,k}, \phi; \mu) - \frac{1}{\tau}m(y_N^{*,k} - y_N^{*,k-1}, \phi), \quad \forall \phi \in Y, \quad \forall k \in \mathbb{K}, \tag{2.5}$$

$$r^{p,k}(\varphi;\mu) = \sigma_1(y_d^k(\mu) - y_N^{*,k}, \varphi)_{L^2(D)} - a(\varphi, p_N^{*,k}; \mu) - \frac{1}{\tau}m(\varphi, p_N^{*,k} - p_N^{*,k+1}), \quad \forall \varphi \in Y, \quad k \in \mathbb{K} \setminus \{K\}, \tag{2.6}$$

$$r^{p,K}(\varphi;\mu) = (\sigma_1 + \frac{\sigma_2}{\tau})(y_d^K(\mu) - y_N^{*,K}, \varphi)_{L^2(D)} - a(\varphi, p_N^{*,K}; \mu) - \frac{1}{\tau}m(\varphi, p_N^{*,K}), \quad \forall \varphi \in Y, \tag{2.7}$$

$$r_i^{u,k}(\psi;\mu) = (\lambda(u_{N,i}^{*,k} - u_{d,i}^k) - b_i(p_N^{*,k}))\psi_i^k, \quad \forall \psi_i^k \in \mathcal{U}_{\mathrm{ad},i}^k, \quad \forall k \in \mathbb{K}, \quad i = 1, \ldots, m. \tag{2.8}$$

Before turning to the bound for the optimal control we require two intermediate results for the state and adjoint predictability errors. We specify the inner product $(v,w)_Y = \frac{1}{2}\big(a(v,w;\mu^{\mathrm{ref}}) + a(w,v;\mu^{\mathrm{ref}})\big)$, where $\mu^{\mathrm{ref}} \in \mathcal{D}$ is a reference parameter value, and assume that we are given a positive lower bound $\alpha_{\mathrm{LB}}(\mu) : \mathcal{D} \to \mathbb{R}_+$ for the coercivity constant $\alpha(\mu)$ such that

$$\alpha(\mu) \geq \alpha_{\mathrm{LB}}(\mu) \geq \alpha_0 > 0, \quad \forall \mu \in \mathcal{D}; \tag{2.9}$$

various recipes exist to construct this lower bound [16, 30, 37]. We can now state

Lemma 2.3. Let $\tilde{e}^{y,k} = y^k(u_N^*) - y_N^{*,k}(u_N^*)$ be the state predictability error and define

$$|||v^k|||^y \equiv \left(m(v^k, v^k) + \sum_{k'=1}^{k} \tau\, a(v^{k'}, v^{k'}; \mu) \right)^{\frac{1}{2}}, \quad \forall k \in \mathbb{K}. \tag{2.10}$$

The state predictability error satisfies

$$|||\tilde{e}^{y,k}|||^y \leq \tilde{\Delta}_N^{y,k}(\mu), \quad \forall \mu \in \mathcal{D}, \ \forall k \in \mathbb{K}, \tag{2.11}$$

where the error bound $\tilde{\Delta}_N^{y,k}(\mu)$ *is defined as*

$$\tilde{\Delta}_N^{y,k}(\mu) \equiv \left(\frac{\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=1}^{k} \|r^{y,k'}(\cdot;\mu)\|_{Y'}^2 \right)^{\frac{1}{2}}. \tag{2.12}$$

This is the standard *a posteriori* error bound for parabolic PDEs; for a proof see [11]. We state the corresponding result for the adjoint in the following lemma; see Appendix A.1 for the proof.

Lemma 2.4. Let $\tilde{e}^{p,k} = \tilde{p}^k - p_N^{*,k}$ be the adjoint predictability error and define

$$|||v^k|||^p \equiv \left(m(v^k, v^k) + \sum_{k'=k}^{K} \tau\, a(v^{k'}, v^{k'}; \mu) \right)^{\frac{1}{2}}. \tag{2.13}$$

and $C_D \equiv \sup_{v \in Y \setminus \{0\}} \frac{\|v\|_{L^2(D)}}{\|v\|_Y}$. *The adjoint predictability error satisfies*

$$|||\tilde{e}^{p,k}|||^p \leq \tilde{\Delta}_N^{p,k}(\mu), \quad \forall \mu \in \mathcal{D}, \ \forall k \in \mathbb{K}, \tag{2.14}$$

where the error bound $\tilde{\Delta}_N^{p,k}(\mu)$ *is defined as*

$$\tilde{\Delta}_N^{p,k}(\mu) \equiv \left(\frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} \|r^{p,k'}(\cdot;\mu)\|_{Y'}^2 + \left(\frac{2C_D^4 \sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)^2} + \frac{\sigma_2^2}{2} \right)(\tilde{\Delta}_N^{y,K}(\mu))^2 \right)^{\frac{1}{2}}. \tag{2.15}$$

For simplicity of exposition we first derive the control error bound for the unconstrained case in the following section and then turn to the constrained case in Section 2.1.2.

2.1.1. *Problems without Control Constraints*

We obtain the following result for the error in the optimal control.

Proposition 2.5. *Let u^* and u_N^* be the optimal solutions of the truth and reduced basis optimal control problems* (P) *and* (P$_N$), *respectively. Given $\tilde{\Delta}_N^{p,k}(\mu)$ defined in* (2.15), *the error in the optimal control satisfies*

$$\|u^* - u_N^*\|_{\mathcal{U}} \leq \Delta_N^{u,*}(\mu) \equiv \frac{1}{\lambda\sqrt{\alpha_{\mathrm{LB}}(\mu)}} \left(\sum_{i=1}^m \|b_i\|_{Y'}^2 \right)^{\frac{1}{2}} \tilde{\Delta}_N^{p,1}(\mu), \quad \forall \mu \in \mathcal{D}. \tag{2.16}$$

Proof. In the unconstrained case $\zeta \in \mathcal{U}$ is given by

$$\zeta = - \left(\lambda(u_N^* - u_d) - \mathcal{B}^\top \tilde{p} \right). \tag{2.17}$$

We add $\pm \mathcal{B}^\top p_N^*$ and note that $\lambda(u_N^* - u_d) - \mathcal{B}^\top p_N^* = 0$ to obtain

$$\zeta = \mathcal{B}^\top (\tilde{p} - p_N^*), \tag{2.18}$$

or component-wise

$$\zeta_i^k = b_i(\tilde{p}^k - p_N^{*,k}), \quad \forall i = 1, \ldots, m, \forall k \in \mathbb{K}. \tag{2.19}$$

We can then bound

$$|\zeta_i^k| \leq \|b_i\|_{Y'} \|\tilde{p}^k - p_N^{*,k}\|_Y, \quad \forall i = 1, \ldots, m, \forall k \in \mathbb{K}, \tag{2.20}$$

and thus arrive at

$$\|u^* - u_N^*\|_{\mathcal{U}} \leq \frac{1}{\lambda}\|\zeta\|_{\mathcal{U}} \leq \frac{1}{\lambda} \left(\tau \sum_{k=1}^K \sum_{i=1}^m \|b_i\|_{Y'}^2 \|\tilde{p}^k - p_N^{*,k}\|_Y^2 \right)^{\frac{1}{2}} = \frac{1}{\lambda} \left(\sum_{i=1}^m \|b_i\|_{Y'}^2 \right)^{\frac{1}{2}} \left(\tau \sum_{k=1}^K \|\tilde{p}^k - p_N^{*,k}\|_Y^2 \right)^{\frac{1}{2}}. \tag{2.21}$$

Finally, it follows from Lemma 2.4 that

$$\left(\tau \sum_{k=1}^K \|\tilde{p}^k - p_N^{*,k}\|_Y^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} \left(\tau \sum_{k=1}^K a(\tilde{e}^{p,k}, \tilde{e}^{p,k}; \mu) \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} \left(m(\tilde{e}^{p,1}, \tilde{e}^{p,1}) + \tau \sum_{k=1}^K a(\tilde{e}^{p,k}, \tilde{e}^{p,k}; \mu) \right)^{\frac{1}{2}}$$

$$\leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} \tilde{\Delta}^{p,1}(\mu). \tag{2.22}$$

The desired result directly follows from (2.21) and (2.22). $\qquad\square$

2.1.2. *Problems with Control Constraints*

The construction of ζ from (2.3) in the control constrained case requires evaluation of the positive and negative part functions of ξ defined in (2.2). Again, we want to avoid an explicit evaluation of ξ since this would require a forward-backward truth solve. The idea here is to construct an efficiently evaluable upper and lower bound for ξ and to conservatively replace the positivity resp. negativity condition by the lower resp. upper bound in order to obtain an approximation of ζ. To this end, we first define

$$\xi_N = \lambda(u_N^* - u_d) - \mathcal{B}^\top p_N^* \tag{2.23}$$

and realize that

$$\xi_i^k = \xi_{N,i}^k - b_i(\tilde{p}^k - p_N^{*,k}), \quad \forall i = 1, \ldots, m. \quad \forall k \in \mathbb{K}. \tag{2.24}$$

We can then prove the following lemma.

Lemma 2.6. *The function ξ defined in (2.2) satisfies*

$$\xi_{N,i,\mathrm{LB}}^k(\mu) \le \xi_i^k \le \xi_{N,i,\mathrm{UB}}^k(\mu), \quad \forall i = 1, \ldots, m. \quad \forall k \in \mathbb{K}, \tag{2.25}$$

where the upper and lower bound are given by

$$\xi_{N,i,\mathrm{LB}}^k(\mu) \equiv \xi_{N,i}^k - \Delta_{N,i}^{\xi,k}(\mu), \qquad \xi_{N,i,\mathrm{UB}}^k(\mu) \equiv \xi_{N,i}^k + \Delta_{N,i}^{\xi,k}(\mu), \quad \forall i = 1, \ldots, m. \quad \forall k \in \mathbb{K}, \tag{2.26}$$

the error bound, $\Delta_{N,i}^{\xi,k}(\mu)$, is defined as

$$\Delta_{N,i}^{\xi,k}(\mu) \equiv \frac{1}{\sqrt{2}} \|\tilde{b}_i\|_{L^2(\Omega)} \tilde{\Delta}_N^{p,k}(\mu), \quad \forall i = 1, \ldots, m. \quad \forall k \in \mathbb{K}, \tag{2.27}$$

and $\tilde{b}_i$ is the $L^2(\Omega)$ Riesz representation of $b_i \in Y'$.

Proof. We note that

$$|b_i(\tilde{p}^k - p_N^{*,k})| \le \|\tilde{b}_i\|_{L^2(\Omega)} \|\tilde{p}^k - p_N^{*,k}\|_{L^2(\Omega)} \le \|\tilde{b}_i\|_{L^2(\Omega)} \frac{1}{\sqrt{2}} \tilde{\Delta}_N^{p,k}(\mu) = \Delta_{N,i}^{\xi,k}(\mu), \tag{2.28}$$

where we used the L^2-norm bound $\|\tilde{e}^{p,k}\|_{L^2(\Omega)} \le \tilde{\Delta}_N^{p,k}(\mu)/\sqrt{2}$, $\forall k \in \mathbb{K}$, (see [13]) and Lemma 2.4 for the last inequality. The desired result directly follows from (2.24) $\qquad\square$

To clarify the next steps we sketch the various quantities involved in constructing the control error bound in Figure 1. For simplicity, we only consider a single control component. On top we show the (unknown) truth optimal control, $u^*(t)$, and reduced basis optimal control, $u_N^*(t)$, over time. The function $\xi(t)$ and its approximation $\xi_N(t)$ as well as the upper and lower bounds $\xi_{N,\mathrm{UB}}(t)$ $\xi_{N,\mathrm{LB}}(t)$ are shown in the middle. We note that $\xi_N(t)$ is identically zero if the constraints for $u_N^*(t)$ are inactive and that $\xi(t)$ is bounded from below and above by $\xi_{N,\mathrm{LB}}(t)$ and $\xi_{N,\mathrm{UB}}(t)$, respectively. Given $u_N^*(t)$ and $\xi(t)$ we may now evaluate ζ from (2.3); we plot the absolute value of ζ on the bottom of Figure 1.

To construct an upper bound for $|\zeta(t)|$, we first note from Lemma 2.6 that

$$|\xi_i^k| \le |\xi_{N,i}^k| + \Delta_{N,i}^{\xi,k}(\mu) = \max\left(|\xi_{N,i,\mathrm{LB}}^k(\mu)|, |\xi_{N,i,\mathrm{LB}}^k(\mu)|\right), \ \forall i = 1, \ldots, m, \ \forall k \in \mathbb{K}. \tag{2.29}$$

Furthermore, we can conservatively replace the positive and negative part function of $\xi(t)$ in (2.3) by the lower and upper bound of $\xi(t)$: For example, assume that for some index (i, k) we have $u_{N,i}^{*,k} = u_{a,i}^k$. If additionally $\xi_{N,i,\mathrm{LB}}^k(\mu) > 0$ we know that $\xi_i^k > 0$ and thus $\left[\xi_i^k\right]_- = 0$, i.e., there is no contribution to the error bound. This corresponds to the index set I_1 defined next. To distinguish the relevant different cases we define the index sets

$$\begin{aligned}
I_1 &= \{(i, k) \in \{1, \ldots, m\} \times \mathbb{K} : u_{N,i}^{*,k} = u_{a,i}^k \text{ and } \xi_{\mathrm{LB}}^{i,k}(\mu) \ge 0\} \\
&\quad \cup \{(i, k) \in \{1, \ldots, m\} \times \mathbb{K} : u_{N,i}^{*,k} = u_{b,i}^k \text{ and } \xi_{\mathrm{UB}}^{i,k}(\mu) \le 0\}; \\
I_2 &= \{(i, k) \in \{1, \ldots, m\} \times \mathbb{K} : u_{N,i}^{*,k} = u_{a,i}^k \text{ and } \xi_{\mathrm{LB}}^{i,k}(\mu) < 0\} \\
&\quad \cup \{(i, k) \in \{1, \ldots, m\} \times \mathbb{K} : u_{N,i}^{*,k} = u_{b,i}^k \text{ and } \xi_{\mathrm{UB}}^{i,k}(\mu) > 0\}; \\
I_3 &= \{(i, k) \in \{1, \ldots, m\} \times \mathbb{K} : u_{N,i}^{*,k} \in ((u_a)_i, (u_b)_i)\}.
\end{aligned} \tag{2.30}$$

We note that I_1 corresponds to the indices where the constraint is active *and* we can guarantee the positivity or negativity of ξ_i^k, I_2 corresponds to the indices where the constraint is active *but we do not know* if ξ_i^k is positive or negative, and I_3 corresponds to the indices where the constraint is inactive. Note that we can evaluate the index sets $I_{1,2,3}$ efficiently online. We may thus define

$$\zeta_{N,i}^k = \begin{cases} 0 & \text{if } (i,k) \in I_1 \\ |\xi_{N,i}^k| + \Delta_{N,i}^{\xi,k}(\mu) & \text{if } (i,k) \in I_2 \cup I_3. \end{cases} \tag{2.31}$$

It follows by construction that $|\zeta_i^k| \leq |\zeta_{N,i}^k|$, $\forall i = 1, \ldots, m$, $\forall k \in K$. We sketch $|\zeta_{N,i}^k|$ and mark the index sets $I_{1,2,3}$ in the bottom coordinate system of Figure 1. To summarize, we obtain the following result for the error in the optimal control.
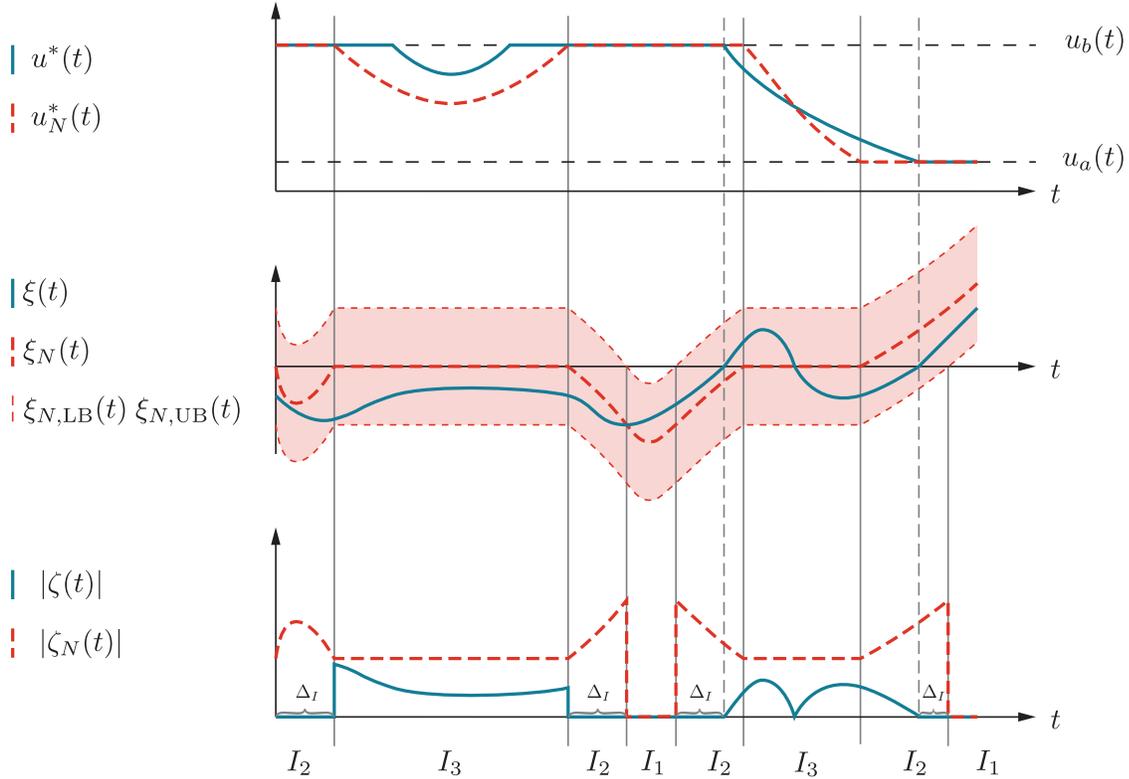

FIGURE 1. Sketch of the error bound construction for control constraints.

Proposition 2.7. *Let u^* and u_N^* be the optimal solutions of the truth and reduced basis optimal control problems* (P) *and* (P$_N$), *respectively. The error in the optimal control satisfies*

$$\|u^* - u_N^*\|_{\mathcal{U}} \leq \Delta_N^{u,*}(\mu) \equiv \frac{1}{\lambda} \|\zeta_N\|_{\mathcal{U}}, \quad \forall \mu \in \mathcal{D}, \tag{2.32}$$

where ζ_N is defined in (2.31).

Proof. The result directly follows from Theorem 2.1, (2.3), (2.29), and the definition of the index sets (2.30). □

We again note that we overestimate the error bound (2.4) for two reasons: first, we replace $|\xi(t)|$ by the upper bound (2.29) which we can efficiently evaluate using the standard offline-online decomposition; and second, we incur additional contributions to the bound by including the time-intervals where we cannot guarantee the required sign of ξ. In Figure 1 we denote these additional intervals by Δ_I. However, since the error bound

$\Delta_{N,i}^{\xi,k}(\mu)$ defined in (2.27) converges to zero as N increases, we can expect Δ_I to decrease. We will confirm this numerically in Section 3.

Finally, we stress that Δ_I is not identical to the index set I_2: consider some index (i,k) where $u_{N,i}^{*,k} = u_{a,i}^k$ and $\xi_i^k < 0$, then $\left[\xi_i^k\right]_-$ in (2.3) is nonzero and we need to rightfully account for its contribution to the error bound (see the rightmost interval I_2 in Figure 1). We do this by replacing ξ_i^k with the upper bound (2.29) and by including all contributions in I_2 in (2.31). Also note that we cannot determine Δ_I online, since this would require the evaluation of ξ_i^k. In fact, we introduce Δ_I solely for the presentation of the numerical results, i.e., to show the convergence of Δ_I to zero as N increases.

2.2. Error Bound for the Cost Functional

Given the error bound $\Delta_N^{u,*}(\mu)$ for the optimal control we may readily derive a bound for the error in the cost functional. We again distinguish between the control constrained and control unconstrained case in Sections 2.2.1 and 2.2.2, respectively. However, we first require the following two preparatory lemmata stating the *a posteriori* error bounds for the state and adjoint optimality errors. We note that the proofs of these lemmata are similar to the proof of Lemma 2.4, i.e., the error in the optimal control — or, more precisely, the error bound of the optimal control — propagates and appears as an additional term in the state and adjoint optimality error bound.

Lemma 2.8. *The state optimality error $e^{y,*,k} = y^{*,k}(u) - y_N^{*,k}(u_N^*)$ satisfies*

$$|||e^{y,*,k}|||^y \leq \Delta_N^{y,*,k}(\mu), \qquad \forall \mu \in \mathcal{D}, \ \forall k \in \mathbb{K}, \tag{2.33}$$

where the error bound $\Delta_N^{y,*,k}(\mu)$ is defined as

$$\Delta_N^{y,*,k}(\mu) \equiv \left(\frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=1}^{k} \|r^{y,k'}(\cdot;\mu)\|_{Y'}^2 + \frac{2}{\alpha_{\mathrm{LB}}(\mu)} \left(\sum_{i=1}^{m} \|b_i\|_{Y'}^2 \right) \left(\Delta_N^{u,*}(\mu) \right)^2 \right)^{\frac{1}{2}}. \tag{2.34}$$

We refer the reader to Appendix A.2 for the proof of Lemma 2.8.

Lemma 2.9. *The adjoint optimality error $e^{p,*,k} = p^{*,k} - p_N^{*,k}$ satisfies*

$$|||e^{p,*,k}|||^p \leq \Delta_N^{p,*,k}(\mu), \qquad \forall \mu \in \mathcal{D}, \ \forall k \in \mathbb{K}, \tag{2.35}$$

where the error bound $\Delta_N^{p,*,k}(\mu)$ is defined as

$$\Delta_N^{p,*,k}(\mu) \equiv \left(\frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} \|r^{p,k'}(\cdot;\mu)\|_{Y'}^2 + \left(\frac{2C_D^4\sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)^2} + \frac{\sigma_2^2}{2} \right) \left(\Delta_N^{y,*,K}(\mu) \right)^2 \right)^{\frac{1}{2}}. \tag{2.36}$$

Proof. The desired result follows directly from the proof of Lemma 2.4 by replacing $\tilde{e}^{y,k}$ with $e^{y,*,k}$, $\tilde{e}^{p,k}$ with $e^{p,*,k}$, and finally invoking Lemma 2.8. We therefore omit the detailed proof. $\square$

2.2.1. *Problems without Control Constraints*

We can now state

Proposition 2.10. *Let $J^* = J(y^*, u^*; \mu)$ and $J_N^* = J(y_N^*, u_N^*; \mu)$ be the optimal values of the cost functionals of the truth and reduced basis optimal control problems, respectively. The error then satisfies*

$$|J^* - J_N^*| \leq \Delta_N^{J,*}(\mu) \equiv \frac{1}{2} \left(\Delta_N^{y,K}(\mu) \, \Delta_N^{p,*,1}(\mu) + \Delta_N^{p,1}(\mu) \, \Delta_N^{y,*,K}(\mu) \right), \quad \forall \mu \in \mathcal{D}, \tag{2.37}$$

where the standard "spatio-temporal" energy norm error bounds $\Delta_N^{y,k}(\mu)$ and $\Delta_N^{p,k}(\mu)$ for the state and adjoint equations are given by

$$\Delta_N^{y,k}(\mu) \equiv \left(\frac{\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=1}^{k} \|r^{y,k'}(\cdot;\mu)\|_{Y'}^2 \right)^{\frac{1}{2}} \tag{2.38}$$

and

$$\Delta_N^{p,k}(\mu) \equiv \left(\frac{\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} \|r^{p,k'}(\cdot;\mu)\|_{Y'}^2 \right)^{\frac{1}{2}}, \tag{2.39}$$

respectively.

Proof. We use the standard result from [4] to estimate the error in the cost functional by

$$J^* - J_N^* = \frac{1}{2}\tau \sum_{k=1}^{K} r^{y,k}(e^{p,*,k};\mu) + \frac{1}{2}\tau \sum_{k=1}^{K} r^{p,k}(e^{y,*,k};\mu), \quad \forall \mu \in \mathcal{D}. \tag{2.40}$$

From the Cauchy-Schwarz inequality we then obtain

$$|J^* - J_N^*| \leq \frac{1}{2} \left(\tau \sum_{k=1}^{K} \|r^{y,k}(\cdot;\mu)\|_{Y'}^2 \right)^{\frac{1}{2}} \left(\tau \sum_{k=1}^{K} \|e^{p,*,k}\|_Y^2 \right)^{\frac{1}{2}} + \frac{1}{2} \left(\tau \sum_{k=1}^{K} \|r^{p,k}(\cdot;\mu)\|_{Y'}^2 \right)^{\frac{1}{2}} \left(\tau \sum_{k=1}^{K} \|e^{y,*,k}\|_Y^2 \right)^{\frac{1}{2}}$$

$$\leq \frac{1}{2} \left(\Delta_N^{y,K}(\mu) \, |||e^{p,*,1}|||^p + \Delta_N^{p,1}(\mu) \, |||e^{y,*,K}|||^y \right), \tag{2.41}$$

where we used the fact that (see (2.22))

$$\left(\tau \sum_{k=1}^{K} \|e^{y,*,k}\|_Y^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} |||e^{y,*,K}|||^y \qquad \text{and} \qquad \left(\tau \sum_{k=1}^{K} \|e^{p,*,k}\|_Y^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} |||e^{p,*,1}|||^p. \tag{2.42}$$

The result then follows from Lemma 2.8 and 2.9. $\qquad\square$

2.2.2. *Problems with Control Constraints*

There are two main differences between the cost functional error bound for the constrained case compared to the unconstrained case. First, we cannot bound the absolute value of the error, $J^* - J_N^*$, and thus only obtain an upper bound for the value of J^*; and second, we need to account for the non-vanishing residual of the optimality condition in the formulation of the bound. We thus obtain the following result.

Proposition 2.11. *Let $J^* = J(y^*, u^*; \mu)$ and $J_N^* = J(y_N^*, u_N^*; \mu)$ be the optimal values of the cost functionals of the truth and reduced basis optimal control problems, respectively. The error then satisfies*

$$J^* - J_N^* \leq \Delta_N^{J,*}(\mu) \equiv \frac{1}{2} \left(\Delta_N^{y,K}(\mu) \, \Delta_N^{p,1}(\mu) + \Delta_N^{p,1}(\mu) \, \Delta_N^{y,*,K}(\mu) + \left(\tau \sum_{i=1}^{m} \sum_{k=1}^{K} |\tilde{r}_i^{u,k}|^2 \right)^{\frac{1}{2}} \Delta_N^{u,*}(\mu) \right), \quad \forall \mu \in \mathcal{D}, \tag{2.43}$$

where $\tilde{r}_i^{u,k} = \lambda(u_{N,i}^{*,k} - u_{d,i}^k) - b_i(p_N^{*,k}).$

Proof. We again use the standard result from [4] to estimate the error in the cost functional by

$$J^* - J_N^* \leq \frac{1}{2}\tau \sum_{k=1}^{K} r^{y,k}(e^{p,*,k};\mu) + \frac{1}{2}\tau \sum_{k=1}^{K} r^{p,k}(e^{y,*,k};\mu) + \frac{1}{2}\tau \sum_{i=1}^{m} \sum_{k=1}^{K} r_i^{u,k}(e_i^{u,*,k};\mu), \quad \forall \mu \in \mathcal{D}, \tag{2.44}$$

where $e_i^{u,*,k} = (u_{N,i}^{*,k} - u_i^{*,k})$, $i = 1, \ldots, m$, $k \in \mathbb{K}$. We can bound the first two terms in (2.44) following the proof of Proposition 2.10. For the last term we obtain

$$\frac{1}{2}\tau \sum_{i=1}^{m} \sum_{k=1}^{K} r_i^{u,k}(e_i^{u,*,k}; \mu) \leq \frac{1}{2}\tau \sum_{i=1}^{m} \sum_{k=1}^{K} |\tilde{r}_i^{u,k}||e_i^{u,*,k}| \leq \frac{1}{2}\left(\tau \sum_{i=1}^{m} \sum_{k=1}^{K} |\tilde{r}_i^{u,k}|^2\right)^{\frac{1}{2}} \left(\tau \sum_{i=1}^{m} \sum_{k=1}^{K} |e_i^{u,*,k}|^2\right)^{\frac{1}{2}}$$

$$= \frac{1}{2}\left(\tau \sum_{i=1}^{m} \sum_{k=1}^{K} |\tilde{r}_i^{u,k}|^2\right)^{\frac{1}{2}} \|e^{u,*}\|_{\mathcal{U}}. \tag{2.45}$$

The result then follows directly by invoking Proposition 2.7. $\qquad\square$

We remark that the cost functional bound for the unconstrained case defined in (2.37) converges superlinearly with respect to the state and adjoint optimality errors. The bound for the constrained case defined in (2.43) on the other hand loses this property if the optimal control has active sets, i.e., one of the control constraints is active. The reason is that the last term in (2.43) contains the sum of the residuals $|\tilde{r}_i^{u,k}|$ which does not converge to zero as N increases but tends to a (positive) constant if the optimal control has active sets. We may thus expect the bound (2.43) to perform considerably worse than the bound (2.37); we will confirm this behavior when discussing numerical results in Section 3.

2.3. Computational Procedure

For the evaluation of the control and cost functional error bounds described in Section 2.1 and 2.2 the following quantities need to be computed: the dual norms of the state and adjoint equation residuals $\|r^{y,k}(\cdot; \mu)\|_{Y'}$ and $\|r^{p,k}(\cdot; \mu)\|_{Y'}$, respectively; the constant C_D; the dual norms of the linear functionals $\|b_i\|_{Y'}$, $1 \leq i \leq m$; and the coercivity lower bound $\alpha_{\mathrm{LB}}(\mu)$. Since all of these quantities can be evaluated using the standard offline-online decomposition [11], we only summarize the computational cost in the online stage. Given a new parameter $\mu \in \mathcal{D}$ and associated optimal solution (y_N^*, p_N^*, u_N^*), evaluation of $\Delta_N^{u,*}(\mu)$ and $\Delta_N^{J,*}(\mu)$ requires (to leading order) $\mathcal{O}(K((Q_a + Q_{yd})N + m)^2)$ operations, and is thus *independent* of $\mathcal{N}$. We also note that the online computational cost to evaluate the rigorous error bounds introduced here is equivalent to the online cost to evaluate the non-rigorous bounds introduced in [6,7].

2.4. Greedy Algorithm

We propose two alternatives for generating the reduced basis space Y_N. The optimal sampling approach discussed in Section 2.4.1 is more expensive during the offline stage — it requires the solution of the truth optimal control problem (P) at selected parameter values — but is more efficient online due to a smaller reduced basis dimension N for a given desired error tolerance. The impulse sampling proposed in Section 2.4.2, on the other hand, is more efficient during the offline stage — we generate a reduced basis space for the impulse response and *do not* solve the truth optimal control problem — but results in a higher reduced basis dimension N and thus higher online cost. The two alternatives thus present a trade-off between offline and online efficiency. However, we can in fact combine both ideas following the idea presented in [8]; we briefly comment on this option at the end of this section.

2.4.1. *Optimal Sampling*

The optimal sampling approach is a straightforward extension of the POD/Greedy sampling procedure introduced in [13]. The method is briefly summarized in Algorithm 1. Here, $\Xi_{\mathrm{train}} \subset \mathcal{D}$ is a finite but suitably large parameter train sample; $\mu^1 \in \Xi_{\mathrm{train}}$ is the initial parameter value; and $\epsilon_{\mathrm{tol,min}} > 0$ is a prescribed desired error tolerance. Furthermore, for a given time history $v_k \in Y$, $k \in \mathbb{K}$, the operator $\mathrm{POD}_Y(\{v_k : k \in \mathbb{K}\})$ returns the largest POD mode with respect to the $(\cdot, \cdot)_Y$ inner product (and normalized with respect to the Y-norm), and $v_{\mathrm{proj},N}^k(\mu)$ denotes the Y-orthogonal projection of $v^k(\mu)$ onto the reduced basis space Y_N. We apply the POD in steps 6 and 7 of Algorithm 1 to the time history of the optimal state and adjoint projection errors,

i.e., $\{e_{\text{proj},N}^{y,k}(\mu) = y^{*,k}(\mu) - y_{\text{proj},N}^{*,k}(\mu) : k \in \mathbb{K}\}$ and $\{e_{\text{proj},N}^{p,k}(\mu) = p^{*,k}(\mu) - p_{\text{proj},N}^{*,k}(\mu) : k \in \mathbb{K}\}$, and not to the optimal solutions $\{y^{*,k}(\mu) : k \in \mathbb{K}\}$ and $\{p^{*,k}(\mu) : k \in \mathbb{K}\}$ themselves.

Note that we expand the reduced basis space with the largest POD mode of the state and the adjoint equation, i.e., we use "integrated" spaces as discussed previously. Also, we may use different metrics $\Delta_N(\mu)$, i.e., the control or cost functional error bound, during the greedy parameter search. This choice could for example be based on (i) the metric of primary interest, i.e., the error in the control or cost; (ii) the rigor of the bound; and/or (iii) the sharpness of the two bounds. Assuming that the error in the cost is of primary interest, we propose to use the (relative) cost functional error bound, $\Delta_N^{J,*}(\mu)/J_N^*(\mu)$, in the sampling procedure in the unconstrained case, and to use the (relative) control error bound, $\Delta_N^{u,*}(\mu)/\|u_N^*(\mu)\|_{\mathcal{U}}$, in the control constrained case. The reason is that in the presence of control constraints the cost functional error estimator is not a provable upper bound for the error in the cost functional as discussed in Section 2.2.2. This choice is also used in Section 3 for the numerical results.

Algorithm 1 Optimal POD/Greedy Sampling Procedure

1: Choose $\Xi_{\text{train}} \subset \mathcal{D}$, $\mu^1 \in \Xi_{\text{train}}$ (arbitrary), and $\epsilon_{\text{tol,min}} > 0$

2: Set $N \leftarrow 1$, $\zeta_1 = \text{POD}_Y(\{y^{*,k}(\mu^1) : k \in \mathbb{K}\})$, $Y_1 \leftarrow \text{span}\{\zeta_1\}$

3: Set $N \leftarrow 2$, $\zeta_2 = \text{POD}_Y(\{e_{\text{proj},N-1}^{p,k}(\mu^1) : k \in \mathbb{K}\})$, $Y_2 \leftarrow Y_1 \oplus \text{span}\{\zeta_2\}$

4: $\mu^* \leftarrow \underset{\mu \in \Xi_{\text{train}}}{\arg\max} \; \Delta_N(\mu)$

5: **while** $\Delta_N(\mu^*) > \epsilon_{\text{tol,min}}$ **do**

6: $\quad N \leftarrow N + 1$, $\zeta_N = \text{POD}_Y(\{e_{\text{proj},N-1}^{y,k}(\mu^*) : k \in \mathbb{K}\})$, $Y_N \leftarrow Y_{N-1} \oplus \text{span}\{\zeta_N\}$

7: $\quad N \leftarrow N + 1$, $\zeta_N = \text{POD}_Y(\{e_{\text{proj},N-1}^{p,k}(\mu^*) : k \in \mathbb{K}\})$, $Y_N \leftarrow Y_{N-1} \oplus \text{span}\{\zeta_N\}$

8: $\quad \mu^* \leftarrow \underset{\mu \in \Xi_{\text{train}}}{\arg\max} \; \Delta_N(\mu)$

9: **end while**

10: $N_{\max} \leftarrow N$

2.4.2. *Impulse Sampling*

Algorithm 1 requires the solution of $N/2$ truth optimal control problems (P). Although this operation is performed during the offline stage, the cost can be considerable. We therefore propose an alternative sampling procedure which is purely based on the impulse response of the state and adjoint equations. To this end, we follow the procedure described in [11] to generate an integrated reduced basis space. More precisely, we apply the standard POD/Greedy sampling procedure but alternate between the following equations after each POD/Greedy step:

(1) the state equation (1.15a) with an impulse consecutively applied at each input;
(2) the adjoint equation (1.15c) with $\tau(y_d^q, \varphi)_{L^2(D)}$ on the right-hand side (sequentially for each $1 \le q \le Q_{yd}$);
(3) the adjoint equation (1.15c) with $\tau(\hat{y}_m^k, \varphi)_{L^2(D)}$ on the right-hand side, where $\hat{y}_m^k$ is the step response for the m-th input (note that we again apply these different right-hand sides consecutively).

Note that we sample on the standard energy-norm error bound and consecutively expand the basis until the desired error tolerance is reached. We are aware that the linear time-invariance (LTI) property — which justifies the impulse approach — does not hold for the adjoint equation (1.15c). However, we still expect the reduced basis to well-approximate the adjoint solution by applying the step response $\hat{y}_m^k$ on the right-hand side of the adjoint in the POD/Greedy procedure. Furthermore, we can always confirm the fidelity of the optimal control solution online thanks to our *a posteriori* error bounds. We will present numerical results in in Section 3.

The appealing feature of the impulse sampling is that we can provide an *a posteriori* error bound for the optimal control problem *without ever* solving the truth optimal control problem. The disadvantage is that we end up with a higher-dimensional basis than with the optimal sampling approach. This is due to the fact that the impulse basis can well-approximate all possible control inputs, whereas the basis using the optimal sampling is tailored towards — and can thus only well-approximate — the optimal control.

Finally, we note that we can combine both sampling procedures following the approach presented in [8]. The idea is as follows: We first generate a reduced basis space Y_N using the impulse sampling. Then, given Y_N, we run the optimal sampling in Algorithm 1 but instead of solving the truth optimal control problem (P) to generate the snapshots we solve the reduced optimal control problem (P_N) with the given Y_N. The second step allows us to condense the basis Y_N and obtain an optimal "derived" basis Y_M of smaller dimension $M \leq N$. This approach combines the advantages of both sampling procedures: a small reduced basis dimension without having to solve the truth optimal control problem.

3. Numerical Results

We consider a linear-quadratic optimal control problem governed by unsteady heat conduction in a two-dimensional domain [29]. The spatial domain, a typical point of which is $x = (x_1, x_2)^T$, is given by $\Omega = (0,7) \times (0,3)$ and is subdivided into the five subdomains $\Omega_1 = (1,2) \times (0,1)$, $\Omega_2 = (3,4) \times (0,1)$, $\Omega_3 = (5,6) \times (0,1)$, $\Omega_4 = (1,6) \times (2.5,3)$, and $\Omega_5 = \Omega \setminus \{\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4\}$. A sketch of the domain is shown in Figure 2. We impose zero Dirichlet conditions on the left and right boundaries and zero Neumann conditions on the bottom and top boundaries. The amount of heat supply in the heater domains Ω_1, Ω_2, and Ω_3 is regulated by the first, second, and third component of the (time–dependent) control $\{u^k = (u_1^k, u_2^k, u_3^k)^T\}_{k \in \mathbb{K}} \in \mathcal{U} \equiv (\mathbb{R}^3)^K$, respectively. The (reference) conductivity in the subdomain $\Omega_1 \cup \Omega_2 \cup \Omega_3$ is set to unity. We consider the normalized conductivity κ in the subdomains Ω_5 and Ω_4 as our first and second parameter $\mu_1, \mu_2 \in [0.5, 5]$, respectively.
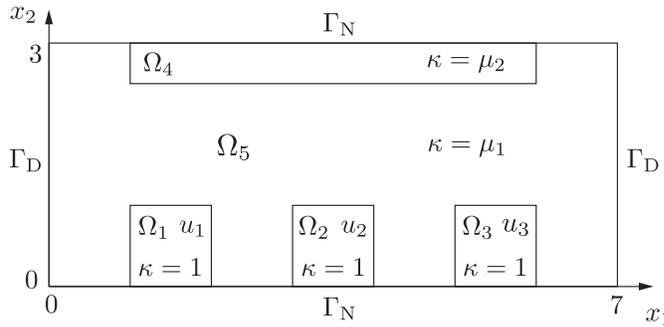

FIGURE 2. Domain Ω for the model problem.

The underlying partial differential equation is the heat (diffusion) equation and we shall directly consider the truth approximation of the problem. The temperature $y^k(\mu) \in Y$ thus satisfies the governing equation in (P), where $Y \subset Y^e \equiv \{v \in \mathrm{H}^1(\Omega) : v|_{\Gamma_D} = 0\}$ is a linear finite element truth approximation subspace of dimension $\mathcal{N} = 9097$ over a triangulation of Ω. We shall consider the time interval $[0, T] = [0, 6]$ and a time step $\tau = 0.06$; we thus have $K = 100$. The initial condition is set to zero. The bilinear and linear forms are given by $a(w, v; \mu) = \mu_1 \int_{\Omega_5} \nabla w \nabla v \, dx + \mu_2 \int_{\Omega_4} \nabla w \nabla v \, dx + \int_{\Omega_1 \cup \Omega_2 \cup \Omega_3} \nabla w \nabla v \, dx$ and $b_i(v) = \int_{\Omega_i} v \, dx$ for $i = 1, 2, 3$, respectively. The bilinear form $a(\cdot, \cdot; \mu)$ admits the affine representation (1.3) with $\Theta_a^1(\mu) = \mu_1, \Theta_a^2(\mu) = \mu_2, \Theta_a^3(\mu) = 1$ and $Q_a = 3$. We also define the inner product $(w, v)_Y = \mu_1^{\mathrm{ref}} \int_{\Omega_5} \nabla w \nabla v \, dx + \mu_2^{\mathrm{ref}} \int_{\Omega_4} \nabla w \nabla v \, dx + \int_{\Omega_1 \cup \Omega_2 \cup \Omega_3} \nabla w \nabla v \, dx$ for $(\mu_1^{\mathrm{ref}}, \mu_2^{\mathrm{ref}})^T = (\sqrt{2.5}, \sqrt{2.5})^T$; we may hence choose $\alpha_{\mathrm{LB}}(\mu) = \min(\mu_1/\mu_1^{\mathrm{ref}}, \mu_2/\mu_2^{\mathrm{ref}}, 1)$ in (2.9).

We consider the quadratic cost functional $J(y, u; \mu) = \frac{1}{2} \sum_{k=1}^{K} \|y^k - y_d^k(\mu)\|_{L^2(D)}^2 + \frac{\lambda}{2} \|u - u_d\|_{\mathcal{U}}^2$, corresponding to (1.14) with fixed regularization parameters $\sigma_1 = 1, \sigma_2 = 0$. We further choose $u_d \equiv 0$ and $D = \Omega_4$; the parametrized desired state $y_d(\mu) \in Y^K$ is given as the discretization of $y_{d,e}(x, t; \mu) = \sin(\mu_3 \pi t) \frac{2}{5}(x_1 - 3.5)\chi_D(x) + 0.5\chi_D(x)$, where $\chi_D(x)$ is the characteristic function on D. Thus the desired state admits the affine representation (1.6) with $\Theta_{yd}^1(t; \mu) = \sin(\mu_3 \pi t)$, $\Theta_{yd}^2(t; \mu) = 0.5$ and $y_{d,e}^1 = \frac{2}{5}(x_1 - 3.5)\chi_D(x)$, $y_{d,e}^2 = \chi_D(x)$. In the presence of control constraints we choose $u_{a,1}^k = u_{a,2}^k = u_{a,3}^k = 0$ and $u_{b,1}^k = u_{b,2}^k = u_{b,3}^k = 0.5, k \in \mathbb{K}$. Finally, we assume that the regularization parameter λ is allowed to vary in the range from 0.1 to 1. The full parametrization of our problem is thus given by $\mu = (\mu_1, \mu_2, \mu_3, \mu_4 = \lambda) \in \mathcal{D} \equiv [0.5, 5] \times [0.5, 5] \times [0.5, 1] \times [0.1, 1]$; we have $P = 4$ parameters.

We first present representative results for the solution of the truth optimal control problem (P). In Figure 3 we plot the optimal control $u^*(t)$ for two different parameter combinations for the unconstrained and constrained case on the top and bottom, respectively. We note that the parameters clearly influence the solution of the optimal control problem and that the lower and upper bound on the control become active depending on the parameter.
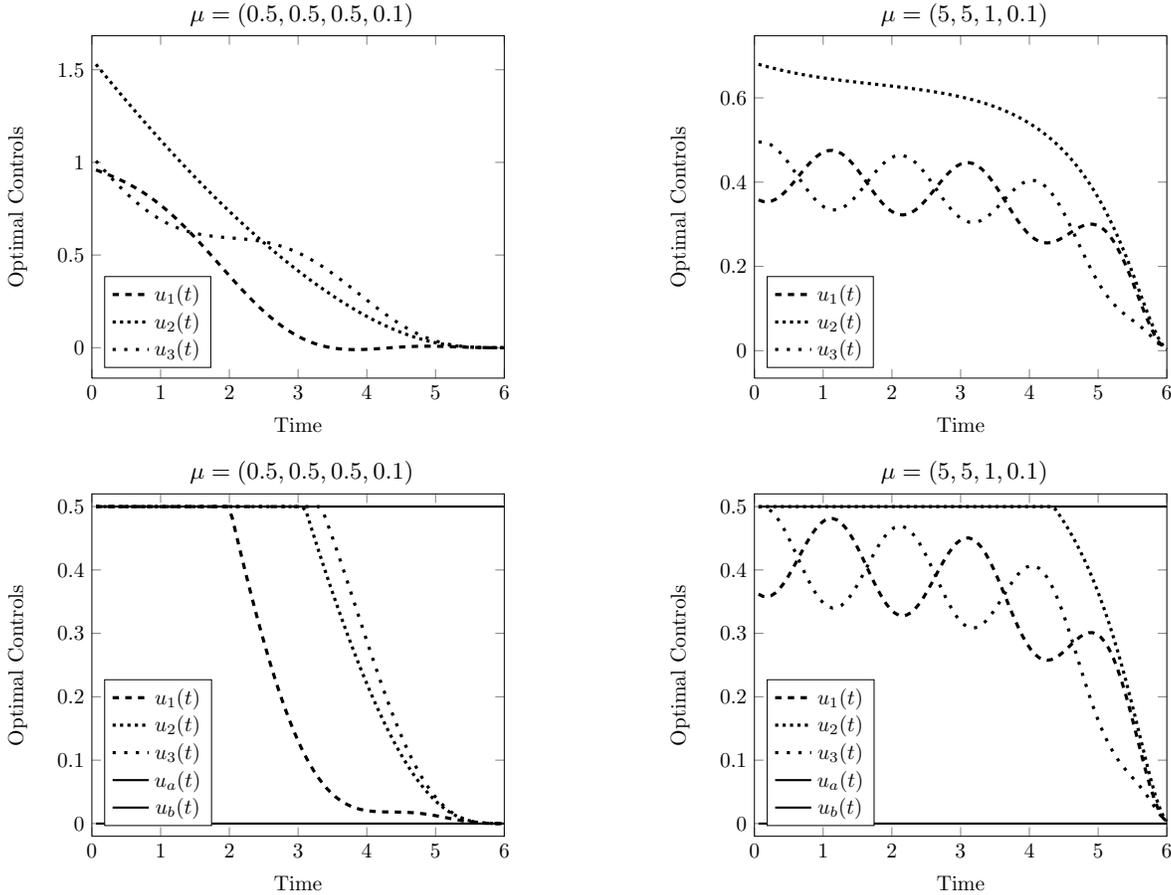
FIGURE 3. Optimal control $u^*(\mu)$ for different representative parameter values

We construct the reduced basis space Y_N according to both sampling procedures described in Section 2.4. To this end, we employ the train sample $\Xi_{\text{train}} \subset \mathcal{D}$ consisting of $n_{\text{train}} = 720$ parameter points over $\mathcal{D}$ and choose $\mu^1 = (0.5, 0.5, 0.5, 0.1)$ as the initial parameter value. We also introduce a parameter test sample Ξ_{test} of size

$n_{\text{test}} = 40$ with a log-random distribution in μ_1, μ_2, λ and a uniform-random distribution in μ_3. We first employ the optimal sampling described in Section 2.4.1 and present results for the unconstrained and constrained case in Sections 3.1 and 3.2, respectively. In the unconstrained case we sample on $\Delta_N(\mu) = \Delta_N^{J,*}(\mu)/J_N^*(\mu)$; in the constrained case on $\Delta_N^{u,*}(\mu)/\|u_N\|_{\mathcal{U}}$. In Section 3.3 we present results for the impulse sampling described in Section 2.4.2.

Throughout this section, we present maximum relative errors and bounds as well as average effectivities. We normalize the state (resp. adjoint) predictability error and bound by $|||y^K(u_N^*)|||^y$ (resp. $|||p^1(y(u_N^*))|||^p$), the state (resp. adjoint) optimality error and bound by $|||y^{*,K}|||^y$ (resp. $|||p^{*,1}|||^p$), the control error and bound by $\|u^*\|_{\mathcal{U}}$, and the cost functional error and bound by J^*. The effectivities are given by the ratios of the error bound to the actual error. The maxima and average values are taken over Ξ_{test}. More precisely, we define the maximum relative state predictability error and bound as

$$\epsilon_{N,\text{max,rel}}^y \equiv \max_{\mu \in \Xi_{\text{test}}} \frac{|||\tilde{e}^{y,K}(\mu)|||^y}{|||y^K(u_N^*(\mu))|||^y}, \qquad \tilde{\Delta}_{N,\text{max,rel}}^y \equiv \max_{\mu \in \Xi_{\text{test}}} \frac{\tilde{\Delta}_N^{y,K}(\mu)}{|||y^K(u_N^*(\mu))|||^y} \qquad (3.1)$$

and the average effectivity, $\bar{\eta}_N^y$, as the average over Ξ_{test} of $\tilde{\Delta}_N^{y,K}(\mu)/|||\tilde{e}^{y,K}(\mu)|||^y$. The remaining quantities are defined analogously and the detailed definitions thus omitted for the sake of brevity.

3.1. Optimal sampling without control constraints

In Table 1 we present, as a function of N, the maximum relative predictability errors $\epsilon_{N,\text{max,rel}}^y$ and $\epsilon_{N,\text{max,rel}}^p$, the maximum relative error bounds $\tilde{\Delta}_{N,\text{max,rel}}^y$ and $\tilde{\Delta}_{N,\text{max,rel}}^p$, and the average effectivities $\bar{\eta}_N^y$ and $\bar{\eta}_N^p$ for the state and adjoint equation, respectively. We also present the maximum relative control error $\epsilon_{N,\text{max,rel}}^{u,*}$, error bound $\Delta_{N,\text{max,rel}}^{u,*}$, and average effectivity $\bar{\eta}_N^{u,*}$. We observe that the state and adjoint predictability errors and bounds are decreasing rapidly with increasing dimension of the reduced basis space and that the error bounds are very sharp for all values of N. The slightly larger effectivities of the adjoint error bound are due to the error contribution of the state predictability error bound in (2.15). We also observe that the error in the optimal control exhibits a superlinear convergence with respect to the state and adjoint predictability errors. Since the control error bound (2.16) is proportional to the adjoint predictability error bound it cannot quite capture the error decay. The mean effectivities thus deteriorate slightly as N increases, but are still acceptable for all values of N.

	State			Adjoint			Control		
N	$\epsilon_{N,\text{max,rel}}^y$	$\tilde{\Delta}_{N,\text{max,rel}}^y$	$\bar{\eta}_N^y$	$\epsilon_{N,\text{max,rel}}^p$	$\tilde{\Delta}_{N,\text{max,rel}}^p$	$\bar{\eta}_N^p$	$\epsilon_{N,\text{max,rel}}^{u,*}$	$\Delta_{N,\text{max,rel}}^{u,*}$	$\bar{\eta}_N^{u,*}$
8	$1.40E{-}1$	$2.13E{-}1$	$1.41E{+}0$	$1.48E{-}1$	$6.79E{-}1$	$2.77E{+}0$	$2.02E{-}1$	$5.15E{+}0$	$2.22E{+}1$
16	$3.83E{-}2$	$6.23E{-}2$	$1.37E{+}0$	$3.72E{-}2$	$1.57E{-}1$	$2.64E{+}0$	$1.33E{-}2$	$1.17E{+}0$	$5.17E{+}1$
32	$1.46E{-}2$	$2.37E{-}2$	$1.27E{+}0$	$7.75E{-}3$	$2.66E{-}2$	$3.52E{+}0$	$2.87E{-}3$	$2.06E{-}1$	$5.79E{+}1$
48	$4.20E{-}3$	$6.41E{-}3$	$1.28E{+}0$	$2.32E{-}3$	$7.58E{-}3$	$3.53E{+}0$	$4.74E{-}4$	$5.72E{-}2$	$1.27E{+}2$
64	$7.04E{-}4$	$1.08E{-}3$	$1.31E{+}0$	$1.11E{-}3$	$2.66E{-}3$	$2.73E{+}0$	$8.91E{-}5$	$2.10E{-}2$	$5.78E{+}2$
80	$3.00E{-}4$	$4.96E{-}4$	$1.34E{+}0$	$2.95E{-}4$	$7.97E{-}4$	$3.06E{+}0$	$1.15E{-}5$	$5.62E{-}3$	$2.20E{+}3$
96	$1.02E{-}4$	$1.34E{-}4$	$1.34E{+}0$	$1.38E{-}4$	$2.95E{-}4$	$2.68E{+}0$	$7.18E{-}6$	$2.74E{-}3$	$3.94E{+}3$

TABLE 1. Unconstrained case with optimal sampling: state predictability, adjoint predictability, and control errors, error bounds, and effectivities as a function of N.

We next turn to the optimality errors and the cost functional. In Table 2 we present, as a function of N, the maximum relative optimality errors $\epsilon_{N,\text{max,rel}}^{y,*}$ and $\epsilon_{N,\text{max,rel}}^{p,*}$, the maximum relative error bounds $\Delta_{N,\text{max,rel}}^{y,*}$ and $\Delta_{N,\text{max,rel}}^{p,*}$, and the average effectivities $\bar{\eta}_N^{y,*}$ and $\bar{\eta}_N^{p,*}$ for the state and adjoint variable, respectively. We also present the maximum relative cost functional error $\epsilon_{N,\text{max,rel}}^{J,*}$, the maximum relative error bound $\Delta_{N,\text{max,rel}}^{J,*}$, and the average effectivities $\bar{\eta}_N^{J,*}$. Again, we observe a rapid decay of both the state and adjoint optimality errors

with increasing N. However, the effectivities of the optimality error bounds are larger than the predictability ones because of the contribution of the control error bound in (2.34) and subsequently in (2.36). We note that — as opposed to the control — the error in the cost functional *and* the error bound converge superlinearly with respect to the predictability and optimality errors and the effectivities thus do not deteriorate with increasing N. Unfortunately, the effectivities are $\mathcal{O}(10^3)$ for all values of N, i.e., we consistently overestimate the error in the cost functional because of the propagation/contribution of the control error bound. However, given the superlinear convergence, the additional online computational cost due to the overestimation is still acceptable.

We finally consider the online computational cost for solving the reduced basis optimal control problem compared to the truth optimal control problem. For $N = 32$ — corresponding to a relative error bound for the cost functional of approximately 1% — we obtain an average (over Ξ_{test}) speed-up of approximately 85 for the solution of the optimal control problem; furthermore, we obtain an average speed-up of approximately 65 for the solution of the optimal control problem *and* evaluation of the error bound for the control and cost functional. Note that the online evaluation of the error bound is approximately 3 times faster than the online solution of the *reduced basis optimal control problem.*

	State			Adjoint			Cost Functional		
N	$\epsilon_{N,\max,rel}^{y,*}$	$\Delta_{N,\max,rel}^{y,*}$	$\bar{\eta}_N^{y,*}$	$\epsilon_{N,\max,rel}^{p,*}$	$\Delta_{N,\max,rel}^{p,*}$	$\bar{\eta}_N^{p,*}$	$\epsilon_{N,\max,rel}^{J,*}$	$\Delta_{N,\max,rel}^{J,*}$	$\bar{\eta}_N^{J,*}$
8	$2.05E{-}1$	$2.11E{+}1$	$4.84E{+}1$	$1.48E{-}1$	$9.20E{+}1$	$9.44E{+}1$	$3.80E{-}3$	$5.23E{+}0$	$6.97E{+}2$
16	$4.23E{-}2$	$4.87E{+}0$	$5.02E{+}1$	$3.72E{-}2$	$2.12E{+}1$	$9.05E{+}1$	$2.34E{-}4$	$2.75E{-}1$	$2.13E{+}3$
32	$1.53E{-}2$	$8.17E{-}1$	$3.89E{+}1$	$7.75E{-}3$	$3.56E{+}0$	$1.27E{+}2$	$1.43E{-}5$	$8.19E{-}3$	$2.80E{+}3$
48	$4.24E{-}3$	$2.36E{-}1$	$4.03E{+}1$	$2.32E{-}3$	$1.03E{+}0$	$1.30E{+}2$	$6.18E{-}7$	$4.81E{-}4$	$1.55E{+}3$
64	$7.05E{-}4$	$8.27E{-}2$	$5.74E{+}1$	$1.11E{-}3$	$3.60E{-}1$	$9.74E{+}1$	$7.97E{-}8$	$6.12E{-}5$	$2.98E{+}3$
80	$3.00E{-}4$	$2.39E{-}2$	$4.88E{+}1$	$2.95E{-}4$	$1.04E{-}1$	$1.17E{+}2$	$1.11E{-}8$	$4.90E{-}6$	$1.76E{+}3$
96	$1.02E{-}4$	$9.18E{-}3$	$6.13E{+}1$	$1.38E{-}4$	$3.99E{-}2$	$1.05E{+}2$	$6.47E{-}10$	$7.20E{-}7$	$2.12E{+}3$

TABLE 2. Unconstrained case with optimal sampling: state optimality, adjoint optimality, and cost functional errors, error bounds, and effectivities as a function of N.

3.2. **Optimal sampling with control constraints**

We next turn to the problem involving constraints on the control. We again present analogous results as in the previous section. The predictability and control errors, bounds, and effectivities are shown in Table 3, the optimality and cost functional errors, bounds, and effectivities are presented in Table 4.

We observe that the results for the state and adjoint predictability and optimality errors, bounds, and hence effectivities very much resemble the unconstrained case. The same holds true for the error bound of the optimal control, i.e., we obtain a certain desired accuracy for the same value of N in the constrained and unconstrained case. However, the error in the optimal control converges even faster than in the unconstrained case and the effectivities thus deteriorate faster. We also observe that the convergence of the cost functional error bound is close to linear with respect to the state and adjoint optimality errors and thus much slower than before. The reason — as discussed in Section 2.2.2 — is the additional term in the cost functional error bound (2.43). As a result, the effectivities deteriorate considerably. The computational times closely resemble the unconstrained case: for $N = 32$ we obtain an average speed-up of approximately 80 for solving the optimal control problem and of 70 for solving the optimal control problem *and* evaluating the control and cost functional error bounds.

Finally, we recall our discussion concerning the control error bound at the end of Section 2.1.2. We pointed out that one reason for overestimating the error bound (2.4) are the additional contributions from time steps where we cannot guarantee the sign of ξ; these intervals were marked with Δ_I in Figure 1. In Table 5 we present, as a function of N, the maximum number of time steps contained in Δ_I for the first ($\Delta_{I,1}$), second ($\Delta_{I,2}$), and third ($\Delta_{I,3}$) control component. Here, the maximum is taken over Ξ_{test}. We note that $\max \Delta_{I,i}$, $i = 1, 2, 3$ — and hence the corresponding contributions to the error bound (2.32) — tend to zero very fast.

N	State			Adjoint			Control		
	$\epsilon^y_{N,\mathrm{max,rel}}$	$\tilde{\Delta}^y_{N,\mathrm{max,rel}}$	$\bar{\eta}^y_N$	$\epsilon^p_{N,\mathrm{max,rel}}$	$\tilde{\Delta}^p_{N,\mathrm{max,rel}}$	$\bar{\eta}^p_N$	$\epsilon^{u,*}_{N,\mathrm{max,rel}}$	$\Delta^{u,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{u,*}_N$
8	$1.98E{-}1$	$2.90E{-}1$	$1.32E{+}0$	$1.36E{-}1$	$6.19E{-}1$	$3.03E{+}0$	$1.38E{-}1$	$1.12E{+}1$	$7.75E{+}1$
16	$9.25E{-}2$	$1.15E{-}1$	$1.29E{+}0$	$4.19E{-}2$	$2.45E{-}1$	$3.91E{+}0$	$9.82E{-}3$	$4.27E{+}0$	$2.89E{+}2$
32	$1.25E{-}2$	$1.82E{-}2$	$1.33E{+}0$	$6.16E{-}3$	$2.29E{-}2$	$3.68E{+}0$	$2.99E{-}4$	$4.02E{-}1$	$1.47E{+}3$
48	$5.39E{-}3$	$6.26E{-}3$	$1.29E{+}0$	$2.15E{-}3$	$9.37E{-}3$	$4.24E{+}0$	$5.65E{-}5$	$1.36E{-}1$	$3.90E{+}3$
64	$1.16E{-}3$	$1.63E{-}3$	$1.32E{+}0$	$6.52E{-}4$	$1.74E{-}3$	$3.25E{+}0$	$6.78E{-}6$	$2.80E{-}2$	$1.02E{+}4$
80	$4.33E{-}4$	$6.20E{-}4$	$1.33E{+}0$	$2.41E{-}4$	$8.50E{-}4$	$3.18E{+}0$	$9.50E{-}7$	$9.89E{-}3$	$3.59E{+}4$
96	$1.63E{-}4$	$2.03E{-}4$	$1.36E{+}0$	$1.01E{-}4$	$3.06E{-}4$	$3.11E{+}0$	$1.17E{-}7$	$3.89E{-}3$	$7.95E{+}4$

TABLE 3. Constrained case with optimal sampling: state predictability, adjoint predictability, and control errors, error bounds, and effectivities as a function of N.

N	State			Adjoint			Cost Functional		
	$\epsilon^{y,*}_{N,\mathrm{max,rel}}$	$\Delta^{y,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{y,*}_N$	$\epsilon^{p,*}_{N,\mathrm{max,rel}}$	$\Delta^{p,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{p,*}_N$	$\epsilon^{J,*}_{N,\mathrm{max,rel}}$	$\Delta^{J,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{J,*}_N$
8	$1.96E{-}1$	$4.25E{+}1$	$7.85E{+}1$	$1.36E{-}1$	$1.38E{+}2$	$1.77E{+}2$	$5.51E{-}3$	$6.76E{+}0$	$4.62E{+}3$
16	$9.30E{-}2$	$1.80E{+}1$	$7.48E{+}1$	$4.19E{-}2$	$5.84E{+}1$	$2.60E{+}2$	$3.36E{-}4$	$1.10E{+}0$	$3.01E{+}3$
32	$1.25E{-}2$	$1.62E{+}0$	$6.95E{+}1$	$6.16E{-}3$	$5.25E{+}0$	$2.14E{+}2$	$7.44E{-}6$	$2.50E{-}2$	$5.52E{+}3$
48	$5.39E{-}3$	$5.73E{-}1$	$6.21E{+}1$	$2.15E{-}3$	$1.86E{+}0$	$2.45E{+}2$	$1.19E{-}6$	$5.02E{-}3$	$5.64E{+}4$
64	$1.16E{-}3$	$8.88E{-}2$	$6.78E{+}1$	$6.52E{-}4$	$2.70E{-}1$	$1.69E{+}2$	$5.51E{-}8$	$1.06E{-}3$	$2.72E{+}4$
80	$4.33E{-}4$	$4.03E{-}2$	$6.89E{+}1$	$2.41E{-}4$	$1.31E{-}1$	$1.70E{+}2$	$7.14E{-}9$	$3.31E{-}4$	$8.79E{+}4$
96	$1.63E{-}4$	$1.42E{-}2$	$6.96E{+}1$	$1.01E{-}4$	$4.61E{-}2$	$1.55E{+}2$	$8.42E{-}10$	$8.78E{-}5$	$8.22E{+}4$

TABLE 4. Constrained case with optimal sampling: state optimality, adjoint optimality, and cost functional errors, error bounds, and effectivities as a function of N.

N	8	16	32	64	96	128	160
$\max \Delta_{I,1}$	41	52	35	16	5	2	1
$\max \Delta_{I,2}$	74	73	12	2	1	1	0
$\max \Delta_{I,3}$	52	50	37	25	10	6	4

TABLE 5. Maximal size of the interval Δ_I for the first, second, and third control component.

3.3. Impulse sampling without control constraints

Finally, we consider the impulse sampling approach discussed in Section 2.4.2. We focus on the unconstrained case and again present analogous results as in the previous two sections. The predictability and control errors, bounds, and effectivities are shown in Table 6, the optimality and cost functional errors, bounds, and effectivities are presented in Table 7.

We immediately note that the effectivities of the state and adjoint predictability and optimality error bounds are very similar to the ones in Section 3.1. However, the rate of convergence of the error and bound itself is much smaller. To achieve a certain desired accuracy we require almost twice as many basis functions with the impulse sampling compared to the optimal sampling. In contrast, the convergence of the error in the optimal control is similar to before. Again, since the control error bound is proportional to the adjoint predictability error bound, this means an even larger increase in the effectivities of the control error bound. The results for the cost functional are also very similar to the ones in Section 3.1. The convergence of the error and bound is slower — we need to choose N almost twice as large as before to obtain a certain desired accuracy — but the effectivities are the same as the ones we obtained using the optimal sampling approach. Following our discussion in Section 2.4.2, however, we expect that one can recover the "optimal" convergence by generating a derived reduced basis following the ideas in [8].

	State			Adjoint			Control		
N	$\epsilon^{y}_{N,\mathrm{max,rel}}$	$\tilde{\Delta}^{y}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{y}_{N}$	$\epsilon^{p}_{N,\mathrm{max,rel}}$	$\tilde{\Delta}^{p}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{p}_{N}$	$\epsilon^{u,*}_{N,\mathrm{max,rel}}$	$\Delta^{u,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{u,*}_{N}$
8	$7.72E{-}1$	$1.27E{+}0$	$1.39E{+}0$	$2.86E{-}1$	$4.79E{+}0$	$6.31E{+}0$	$2.25E{-}1$	$3.60E{+}1$	$6.30E{+}1$
16	$3.24E{-}1$	$4.56E{-}1$	$1.31E{+}0$	$1.17E{-}1$	$1.16E{+}0$	$5.87E{+}0$	$4.77E{-}2$	$9.28E{+}0$	$1.85E{+}2$
32	$4.73E{-}2$	$6.27E{-}2$	$1.25E{+}0$	$8.99E{-}3$	$2.26E{-}1$	$5.42E{+}0$	$8.07E{-}4$	$1.65E{+}0$	$6.45E{+}2$
64	$4.38E{-}3$	$6.05E{-}3$	$1.34E{+}0$	$1.95E{-}3$	$1.58E{-}2$	$4.39E{+}0$	$1.64E{-}5$	$1.56E{-}1$	$4.98E{+}3$
96	$9.73E{-}4$	$1.34E{-}3$	$1.31E{+}0$	$2.40E{-}4$	$5.39E{-}3$	$6.27E{+}0$	$7.47E{-}7$	$3.93E{-}2$	$3.08E{+}4$
128	$2.46E{-}4$	$3.36E{-}4$	$1.29E{+}0$	$4.04E{-}5$	$1.14E{-}3$	$9.35E{+}0$	$3.68E{-}8$	$8.32E{-}3$	$1.76E{+}5$

TABLE 6. Unconstrained case with impulse sampling: state predictability, adjoint predictability, and control errors, error bounds, and effectivities as a function of N.

	State			Adjoint			Cost Functional		
N	$\epsilon^{y,*}_{N,\mathrm{max,rel}}$	$\Delta^{y,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{y,*}_{N}$	$\epsilon^{p,*}_{N,\mathrm{max,rel}}$	$\Delta^{p,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{p,*}_{N}$	$\epsilon^{J,*}_{N,\mathrm{max,rel}}$	$\Delta^{J,*}_{N,\mathrm{max,rel}}$	$\bar{\eta}^{J,*}_{N}$
8	$7.71E{-}1$	$1.53E{+}2$	$3.72E{+}1$	$2.86E{-}1$	$6.66E{+}2$	$2.61E{+}2$	$4.13E{-}2$	$1.98E{+}2$	$2.80E{+}3$
16	$3.23E{-}1$	$3.62E{+}1$	$3.62E{+}1$	$1.17E{-}1$	$1.58E{+}2$	$2.49E{+}2$	$1.02E{-}2$	$1.57E{+}1$	$4.10E{+}3$
32	$4.73E{-}2$	$7.02E{+}0$	$3.29E{+}1$	$8.99E{-}3$	$3.05E{+}1$	$2.32E{+}2$	$1.08E{-}4$	$4.02E{-}1$	$2.18E{+}3$
64	$4.38E{-}3$	$4.89E{-}1$	$3.96E{+}1$	$1.95E{-}3$	$1.75E{+}0$	$1.75E{+}2$	$4.52E{-}7$	$4.38E{-}3$	$9.21E{+}3$
96	$9.73E{-}4$	$1.68E{-}1$	$3.48E{+}1$	$2.40E{-}4$	$7.29E{-}1$	$2.71E{+}2$	$2.72E{-}8$	$2.28E{-}4$	$2.98E{+}3$
128	$2.46E{-}4$	$3.54E{-}2$	$3.38E{+}1$	$4.04E{-}5$	$1.54E{-}1$	$3.80E{+}2$	$3.25E{-}10$	$1.11E{-}5$	$3.63E{+}4$

TABLE 7. Unconstrained case with impulse sampling: state optimality, adjoint optimality, and cost functional errors, error bounds, and effectivities as a function of N.

4. CONCLUSIONS

The solution of optimal control problems governed by (parametrized) parabolic partial differential equations is a challenging and often time-consuming task. The surrogate model approach, where the original high-dimensional PDE approximation is replaced by a reduced order model, allows to decrease the computational burden: the (online) solution of the reduced order optimal control problem is significantly faster than the solution of the high-dimensional problem and at the same time provides a very good accuracy with small errors (compared to the high-dimensional problem). However, an efficient *and* rigorous error control for the surrogate model approach was lacking so far.

In this paper, we considered parametrized optimal control problems with and without control constraints. We developed rigorous *a posteriori* error bounds for the optimal control and the associated cost functional which can be efficiently evaluated using a standard offline-online decomposition. To generate the reduced basis spaces we proposed two greedy algorithms. Finally, we presented numerical results for a model problem that corroborate the theory. In summary, our approach enables the certified *and* real-time solution of parametrized optimal control problems and can be gainfully employed, e.g., in model predictive control.

APPENDIX A. PROOFS

A.1. **Proof of Lemma 2.4**

Proof. We immediately derive from (1.15c), (1.15d) and (2.6), (2.7) that $\tilde{e}^{p,k} = p^k(y(u_N^*)) - p_N^{*,k}(y_N^*(u_N^*))$ satisfies

$$m(\varphi, \tilde{e}^{p,k} - \tilde{e}^{p,k+1}) + \tau a(\varphi, \tilde{e}^{p,k}; \mu) = \tau r^{p,k}(\varphi; \mu) + \tau \sigma_1(y_N^{*,k} - \tilde{y}^k, \varphi)_{L^2(D)}, \quad \forall \varphi \in Y, \quad K - 1 \geq k \geq 1, \quad \text{(A.1)}$$

with final condition

$$m(\varphi, \tilde{e}^{p,K}) + \tau a(\varphi, \tilde{e}^{p,K}; \mu) = \tau r^{p,K}(\varphi; \mu) + \tau \sigma_1(y_N^{*,K} - \tilde{y}^K, \varphi)_{L^2(D)} + \sigma_2(y_N^{*,K} - \tilde{y}^K, \varphi)_{L^2(D)}, \quad \forall \varphi \in Y. \quad \text{(A.2)}$$

We now choose $\varphi = \tilde{e}^{p,k}$ in (A.1) and $\varphi = \tilde{e}^{p,K}$ in (A.2) to obtain

$$m(\tilde{e}^{p,k}, \tilde{e}^{p,k}) + \tau a(\tilde{e}^{p,k}, \tilde{e}^{p,k}; \mu) = m(\tilde{e}^{p,k}, \tilde{e}^{p,k+1}) + \tau r^{p,k}(\tilde{e}^{p,k}; \mu) + \tau \sigma_1 (y_N^{*,k} - \tilde{y}^k, \tilde{e}^{p,k})_{L^2(D)}, \quad K-1 \geq k \geq 1, \quad \text{(A.3)}$$

and

$$m(\tilde{e}^{p,K}, \tilde{e}^{p,K}) + \tau a(\tilde{e}^{p,K}, \tilde{e}^{p,K}; \mu) = \tau r^{p,K}(\tilde{e}^{p,K}; \mu) + \tau \sigma_1 (y_N^{*,K} - \tilde{y}^K, \tilde{e}^{p,K})_{L^2(D)}$$
$$+ \sigma_2 (y_N^{*,K} - \tilde{y}^K, \tilde{e}^{p,K})_{L^2(D)}, \quad \forall \varphi \in Y. \quad \text{(A.4)}$$

We first consider (A.3) and note that we can bound the first and third term on the right-hand side using Cauchy-Schwarz and Young's inequality by

$$2m(\tilde{e}^{p,k}, \tilde{e}^{p,k+1}) \leq m(\tilde{e}^{p,k}, \tilde{e}^{p,k}) + m(\tilde{e}^{p,k+1}, \tilde{e}^{p,k+1}), \quad \text{(A.5)}$$

and

$$2\tau \sigma_1 (y_N^{*,k} - \tilde{y}^k, \tilde{e}^{p,k})_{L^2(D)} \leq 2\tau \sigma_1 C_D^2 \|y_N^{*,k} - \tilde{y}^k\|_Y \|\tilde{e}^{p,k}\|_Y \leq C_D^4 \frac{2\tau \sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)} \|y_N^{*,k} - \tilde{y}^k\|_Y^2 + \frac{\tau \alpha_{\mathrm{LB}}(\mu)}{2} \|\tilde{e}^{p,k}\|_Y^2, \quad \text{(A.6)}$$

where we also used the definition of C_D. Furthermore, using the definition of the dual norm and Young's inequality the second term on the right-hand side of (A.3) can be bounded by

$$2\tau r^{p,k}(\tilde{e}^{p,k}; \mu) \leq 2\tau \|r^{p,k}(\cdot; \mu)\|_{Y'} \|\tilde{e}^{p,k}\|_Y \leq \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \|r^{p,k}(\cdot; \mu)\|_{Y'}^2 + \frac{\tau \alpha_{\mathrm{LB}}(\mu)}{2} \|\tilde{e}^{p,k}\|_Y^2. \quad \text{(A.7)}$$

It thus follows from (A.3), (A.5), (A.6), (A.7), and invoking (1.12) and (2.9) that

$$m(\tilde{e}^{p,k}, \tilde{e}^{p,k}) - m(\tilde{e}^{p,k+1}, \tilde{e}^{p,k+1}) + \tau a(\tilde{e}^{p,k}, \tilde{e}^{p,k}; \mu) \leq \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \|r^{p,k}(\cdot; \mu)\|_{Y'}^2 + C_D^4 \frac{2\tau \sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)} \|y_N^{*,k} - \tilde{y}^k\|_Y^2. \quad \text{(A.8)}$$

We next consider the final condition (A.4) where we bound the last term on the right-hand side using Cauchy-Schwarz and Young's inequality by

$$2\sigma_2 (y_N^{*,K} - \tilde{y}^K, \tilde{e}^{p,K})_{L^2(D)} \leq 2\sigma_2 \|y_N^{*,K} - \tilde{y}^K\|_{L^2(D)} \|\tilde{e}^{p,K}\|_{L^2(D)}$$
$$\leq \sigma_2^2 \|y_N^{*,K} - \tilde{y}^K\|_{L^2(D)}^2 + \|\tilde{e}^{p,K}\|_{L^2(D)}^2. \quad \text{(A.9)}$$

It thus follows from (A.4), (A.6) and (A.7) for $k = K$, (A.9), and invoking (1.12) and (2.9) that

$$m(\tilde{e}^{p,K}, \tilde{e}^{p,K}) + \tau a(\tilde{e}^{p,K}, \tilde{e}^{p,K}; \mu) \leq \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \|r^{p,K}(\cdot; \mu)\|_{Y'}^2$$
$$+ C_D^4 \frac{2\tau \sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)} \|y_N^{*,K} - \tilde{y}^K\|_Y^2 + \sigma_2^2 \|y_N^{*,K} - \tilde{y}^K\|_{L^2(D)}^2, \quad \text{(A.10)}$$

where we also used the fact that $\|\cdot\|_{L^2(D)} \leq \|\cdot\|_{L^2(\Omega)} = m(\cdot, \cdot)$. We now perform the sum from $k' = k$ to $K-1$ of (A.8) and add (A.10), leading to

$$m(\tilde{e}^{p,k}, \tilde{e}^{p,k}) + \tau \sum_{k'=k}^{K} a(\tilde{e}^{p,k}, \tilde{e}^{p,k}; \mu) \leq \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} \|r^{p,k'}(\cdot; \mu)\|_{Y'}^2$$
$$+ C_D^4 \frac{2\tau \sigma_1^2}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} \|y_N^{*,k'} - \tilde{y}^{k'}\|_Y^2 + \sigma_2^2 \|y_N^{*,K} - \tilde{y}^K\|_{L^2(D)}^2. \quad \text{(A.11)}$$

Finally, we note that we can bound the sum of the error $\tilde{e}^{y,k} = y_N^{*,k} - \tilde{y}^k$ by

$$\tau \sum_{k'=k}^{K} \|\tilde{e}^{y,k'}\|_Y^2 \leq \frac{\tau}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=k}^{K} a(\tilde{e}^{y,k'}, \tilde{e}^{y,k'}; \mu) \leq \frac{1}{\alpha_{\mathrm{LB}}(\mu)} \left(m(\tilde{e}^{y,K}, \tilde{e}^{y,K}) + \tau \sum_{k'=1}^{K} a(\tilde{e}^{y,k'}, \tilde{e}^{y,k'}; \mu) \right)$$

$$\leq \frac{1}{\alpha_{\mathrm{LB}}(\mu)} \left(\tilde{\Delta}_N^{y,K}(\mu) \right)^2 \tag{A.12}$$

and from the L^2-norm bound [13] that

$$\|y_N^{*,K} - \tilde{y}^K\|_{L^2(D)}^2 \leq \frac{1}{2} \left(\tilde{\Delta}_N^{y,K}(\mu) \right)^2. \tag{A.13}$$

The desired result then follows from (A.11), (A.12), and (A.13). □

A.2. Proof of Lemma 2.8

Proof. Since the proof follows along the same lines as the proof of Lemma 2.4 we only sketch the main steps. We immediately derive from (1.15a) and (2.5) that $e^{y,*,k} = y^{*,k}(u^*) - y_N^{*,k}(u_N^*)$ satisfies

$$m(e^{y,*,k} - e^{y,*,k-1}, \phi) + \tau a(e^{y,*,k}, \phi; \mu) = \tau r^{y,k}(\phi; \mu) + \tau \sum_{i=1}^{m} b_i(\phi)(u_i^{*,k} - u_{N,i}^{*,k}), \quad \forall \phi \in Y, \ 1 \leq k \leq K, \tag{A.14}$$

with initial condition $e^{y,*,0} = 0$ since $y^{*,0} = y_N^{*,0} = 0$ by assumption. We now choose $\phi = e^{y,*,k}$, again apply Cauchy-Schwarz, Young's inequality, and the definition of the dual norm — analogous to (A.5) and (A.7) — and invoke (1.12) and (2.9) to obtain

$$m(e^{y,*,k}, e^{y,*,k}) - m(e^{y,*,k-1}, e^{y,*,k-1}) + \tau a(e^{y,*,k}, e^{y,*,k}; \mu)$$

$$\leq \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \|r^{y,k}(\cdot; \mu)\|_{Y'}^2 + \frac{2\tau}{\alpha_{\mathrm{LB}}(\mu)} \left(\sum_{i=1}^{m} \|b_i\|_{Y'}(u_i^{*,k} - u_{N,i}^{*,k}) \right)^2 \tag{A.15}$$

The desired result follows by performing the sum from $k' = 1$ to k, recalling that $e^{y,*,0} = 0$, and invoking Proposition 2.5 resp. 2.7. □

References

[1] N. Altmüller and L. Grüne. Distributed and boundary model predictive control for the heat equation. *GAMM Mitteilungen*, 35(2):131–145, 2012.

[2] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Advances in Design and Control. SIAM, 2005.

[3] J. A. Atwell and B. B. King. Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations. *Math. Comput. Model.*, 33(1-3):1–19, 2001.

[4] R. Becker, H. Kapp, and R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: Basic concept. *SIAM J. Control Optim.*, 39:113–132, 2000.

[5] P. Benner, V. Mehrmann, and D. Sorensen, editors. *Dimension reduction of large-scale systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, Berlin, 2005. Springer.

[6] L. Dedè. Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. *SIAM J. Sci. Comput.*, 32(2):997–1019, 2010.

[7] L. Dedè. Reduced basis method and error estimation for parametrized optimal control problems with control constraints. *J. Sci. Comput.*, 50(2):287–305, 2012.

[8] J. Eftang, D. Huynh, D. Knezevic, and A. Patera. A two-step certified reduced basis method. *J. Sci. Comput.*, 51:28–58, 2012.

[9] A.-L. Gerner and K. Veroy. Certified reduced basis methods for parametrized saddle point problems. *SIAM J. Sci. Comput.*, 34(5):A2812–A2836, 2012.

[10] M. A. Grepl and M. Kärcher. Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems. *C. R. Math.*, 349(15-16):873–877, 2011.

[11] M. A. Grepl and A. T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *ESAIM: Math. Model. Num.*, 39(1):157–181, 2005.

[12] M. Gunzburger and A. Kunoth. Space-time adaptive wavelet methods for optimal control problems constrained by parabolic evolution equations. *SIAM J. Control and Optimization*, pages 1150–1170, 2011.

[13] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM: Math. Model. Num.*, 42(02):277–302, 2008.

[14] W. Hager. Multiplier methods for nonlinear optimal control. *SIAM J. Numer. Anal.*, 27(4):1061–1080, 1990.

[15] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, 2009.

[16] 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(8):473–478, 2007.

[17] K. Ito and K. Kunisch. Receding horizon optimal control for infinite dimensional systems. *ESAIM: Contr. Optim. Ca.*, 8:741–760, 2002.

[18] K. Ito and K. Kunisch. Reduced-order optimal control based on approximate inertial manifolds for nonlinear dynamical systems. *SIAM J. Numer. Anal.*, 46(6):2867–2891, 2008.

[19] K. Ito and S. S. Ravindran. A reduced-order method for simulation and control of fluid flows. *J. Comput. Phys.*, 143(2):403 – 425, 1998.

[20] K. Ito and S. S. Ravindran. A reduced basis method for optimal control of unsteady viscous flows. *Int. J. Comput. Fluid D.*, 15:97–113, 2001.

[21] M. Kärcher. The reduced-basis method for parametrized linear-quadratic elliptic optimal control problems. Master's thesis, Technische Universität München, April 2011.

[22] M. Kärcher and M. A. Grepl. A certified reduced basis method for parametrized elliptic optimal control problems. *ESAIM: Contr. Optim. Ca.*, 2013. accepted for publication.

[23] K. Kunisch and S. Volkwein. Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition. *J. Optimiz. Theory App.*, 102:345–371, 1999.

[24] K. Kunisch, S. Volkwein, and L. Xie. HJB-POD based feedback design for the optimal control of evolution problems. *SIAM J. Appl. Dyn. Syst.*, 3:701–722, 2004.

[25] G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich, and S. Ulbrich, editors. *Constrained Optimization and Optimal Control for Partial Differential Equations*. (International Series of Numerical Mathematics). Birkhäuser Basel, 2012.

[26] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer, 1971.

[27] K. Malanowski, C. Buskens, and H. Maurer. *Convergence of approximations to nonlinear optimal control problems*, volume 195, pages 253–284. CRC Press, 1997.

[28] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Reduced basis method for parametrized elliptic optimal control problems. *SIAM J. Sci. Comput.*, 2013. accepted for publication.

[29] I. B. Oliveira. *A "HUM" Conjugate Gradient Algorithm for Constrained Nonlinear Optimal Control: Terminal and Regulator Problems*. PhD thesis, Massachusetts Institute of Technology, 2002.

[30] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *J. Fluid. Eng.*, 124(1):70–80, 2002.

[31] A. M. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer, 2008.

[32] 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. Method. E.*, 15(3):229–275, 2008.

[33] G. Rozza and K. Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *Comput. Methods Appl. Mech. Engrg.*, 196(7):1244 – 1260, 2007.

[34] T. Tonn, K. Urban, and S. Volkwein. Comparison of the reduced-basis and pod a posteriori error estimators for an elliptic linear-quadratic optimal control problem. *Math. Comput. Model. Dyn.*, 17(4):355–369, 2011.

[35] F. Tröltzsch and S. Volkwein. POD a-posteriori error estimates for linear-quadratic optimal control problems. *Comput. Optim. Appl.*, 44:83–115, 2009.

[36] K. Urban and A. T. Patera. A new error bound for reduced basis approximation of parabolic partial differential equations. *C. R. Math.*, 350(3–4):203–207, 2012.

[37] K. Veroy, D. V. Rovas, and A. T. Patera. A posteriori error estimation for reduced-basis approximation of parametrized elliptic coercive partial differential equations: "convex inverse" bound conditioners. *ESAIM: Contr. Optim. Ca.*, 8:1007–1028, 2002. Special Volume: A tribute to J.L. Lions.