Hamburger Beiträge
zur Angewandten Mathematik

POD basis updates for nonlinear PDE control

Carmen Gräßle, Martin Gubisch, Simone Metzdorf,
Sabrina Rogg, Stefan Volkwein

Nr. 2016-18
August 2016

C. Gräßle, M. Gubisch, S. Metzdorf, S. Rogg, and S. Volkwein

POD basis updates for nonlinear PDE control

Abstract: In the present paper a semilinear boundary control problem is considered. For its numerical solution proper orthogonal decomposition (POD) is applied. POD is based on a Galerkin type discretization with basis elements created from the evolution problem itself. In the context of optimal control this approach may suffer from the fact that the basis elements are computed from a reference trajectory containing features which are quite different from those of the optimally controlled trajectory. Therefore, different POD basis update strategies which avoids this problem of unmodelled dynamics are compared numerically.

1 Introduction

In this paper we consider an optimal control problem governed by a semilinear parabolic equation together with control constraints. For the numerical solution we apply a Galerkin approximation, which is based on proper orthogonal decomposition (POD), a method for deriving reduced-order models of dynamical systems; cf. Holmes et al. (2012). However, to obtain the state data underlying the POD model, it is necessary to solve once the full state system using a reference control. Consequently, the POD approximations depend on the chosen reference control, so that the choice of a reference control turned out to be essential for the computation of a POD basis for the optimal control problem. To overcome this problem we investigate two different basis update strategies for improving the POD basis: optimality-system POD (cf. Kunisch, Volkwein (2008)) and trust-region POD (cf. Arian et al. (2000)). In both update strategies the POD basis is changed in the optimization method in order to ensure convergence and a certain accuracy for the obtained controls. Let us also refer to the papers Yue, Meerbergen (2013) and Qian et al. (2016), where the trust-region optimization is efficiently combined with the reduced-basis method.

C. Gräßle: Universität Hamburg, Fachbereich Mathematik, Bundesstraße 55, 20146 Hamburg, Germany, carmen.graessle@uni-hamburg.de
M. Gubisch, S. Metzdorf, S. Rogg, S. Volkwein: Universität Konstanz, Fachbereich Mathematik und Statistik, Universitätsstraße 10, 78457 Konstanz, Germany, {martin.gubisch,simone.metzdorf,sabrina.rogg,stefan.volkwein}@uni-konstanz.de

The paper is organized as follows: In Section 2 we introduce our semilinear control problem and review optimality conditions. Our applied numerical methods are explained in Section 3. Section 4 is devoted to explain the application of the POD method to our optimal control problem and in Section 5 we present our numerical experiments.

2 Optimal control problem

Let $\Omega \subset \mathbb{R}^{\mathsf{n}}$, $\mathsf{n} \in \{1,2,3\}$, be a bounded, open domain with Lipschitz-continuous boundary $\Gamma = \partial\Omega$. The boundary is divided into m disjunct segments Γ_i with $\Gamma = \cup_{i=1}^m \Gamma_i$. The corresponding characteristic functions are denoted by χ_{Γ_i}, $i = 1, \dots, m$. We consider the time interval $[0,T]$ with $T > 0$ and define the sets $Q = (0,T) \times \Omega$, $\Sigma = (0,T) \times \Gamma$. The quadratic objective has the form

$$J(y,u) = \frac{1}{2} \int_\Omega \big| y(T,\boldsymbol{x}) - y_d(\boldsymbol{x}) \big|^2 \,\mathrm{d}\boldsymbol{x} + \frac{\sigma_u}{2} \int_0^T \int_\Gamma \Big| \sum_{i=1}^m u_i(t) \chi_{\Gamma_i}(\boldsymbol{x}) \Big|^2 \,\mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{1}$$

In (1) the desired state $y_d : \Omega \to \mathbb{R}$ is assumed to be bounded and $\sigma_u > 0$ holds. The admissible controls belong to the closed, convex set

$$U_{\mathsf{ad}} = \big\{ u : [0,T] \to \mathbb{R}^m \,\big|\, u_a \le u \le u_b \text{ in } [0,T] \big\}, \tag{2}$$

where u_a, $u_b : [0,T] \to \mathbb{R}^m$ are bounded functions and "$\le$" is understood componentwise in $\mathbb{R}^m$. For a given control $u \in U_{\mathsf{ad}}$ the state $y = y(t,\boldsymbol{x})$ satisfies the semilinear parabolic differential equation

$$\begin{aligned} c y_t(t,\boldsymbol{x}) - \Delta y(t,\boldsymbol{x}) + y(t,\boldsymbol{x})^3 &= 0, && (t,\boldsymbol{x}) \in Q,\\ \frac{\partial y}{\partial \boldsymbol{n}}(t,\boldsymbol{x}) + q y(t,\boldsymbol{x}) &= \sum_{i=1}^m u_i(t) \chi_{\Gamma_i}(\boldsymbol{x}), && (t,\boldsymbol{x}) \in \Sigma,\\ y(0,\boldsymbol{x}) &= y_\circ(\boldsymbol{x}), && \boldsymbol{x} \in \Omega. \end{aligned} \tag{3}$$

We suppose that the initial condition $y_\circ : \Omega \to \mathbb{R}$ is bounded and c, q are positive constants. Throughout the paper we focus on the specific nonlinearity y^3, which is utilized in our numerical tests. Of course, other nonlinearities – satisfying certain boundedness and monotonicity conditions – are also possible; see (Tröltzsch, 2010, Chapter 5.1). In (3) the vector $\boldsymbol{n} = \boldsymbol{n}(\boldsymbol{x})$ stands for the normal vector defined on Γ and $\frac{\partial y}{\partial \boldsymbol{n}}$ is the normal derivative. Recall that $H = L^2(\Omega)$ and $V = H^1(\Omega)$ are defined as

$$H = \Big\{ \varphi : \Omega \to \mathbb{R} \,\Big|\, \|\varphi\|_H = \Big(\int_\Omega |\varphi(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} \Big)^{1/2} < \infty \Big\},$$

$$V = \Big\{ \varphi \in H \,\Big|\, \|\varphi\|_V = \Big(\|\varphi\|_H^2 + \sum_{i=1}^{\mathsf{n}} \|\varphi_{x_i}\|_H^2 \Big)^{1/2} < \infty \Big\},$$

respectively, where φ_{x_i} denotes the (weak) partial derivative of φ with respect to i-th component x_i of $\boldsymbol{x} \in \Omega$. A solution to (3) is understood in the following weak sense: $y(t) = y(t,\cdot) \in V$ satisfies $y(0) = y_\circ$ in Ω and

$$\begin{aligned}\int_\Omega c y_t(t)\varphi + \nabla y(t) \cdot \nabla\varphi + y(t)^3\varphi \,\mathrm{d}\boldsymbol{x}&\\ + \int_\Gamma q y(t)\varphi \,\mathrm{d}\boldsymbol{x} = \sum_{i=1}^m u_i(t) \int_{\Gamma_i} \varphi \,\mathrm{d}\boldsymbol{x}&\end{aligned} \tag{4}$$

for all $\varphi \in V$, where we set $|\Gamma_i| = \int_{\Gamma_i} 1 \mathrm{d}\boldsymbol{x}$. Let us define

$$\begin{aligned}\langle e(y,u), (p, p_\circ)\rangle &= \int_\Omega \big(y(0) - y_\circ\big) p_\circ \,\mathrm{d}\boldsymbol{x}\\ &+ \int_\Omega c y_t(t) p(t) + \nabla y(t) \cdot \nabla p(t) + y(t)^3 p(t) \,\mathrm{d}\boldsymbol{x}\\ &+ \int_\Gamma q y(t) p(t) \,\mathrm{d}\boldsymbol{x} - \sum_{i=1}^m u_i(t) \int_{\Gamma_i} p(t) \,\mathrm{d}\boldsymbol{x}\end{aligned}$$

for all (test) functions $p : Q \to \mathbb{R}$ with $\int_0^T \|p(t)\|_V^2 \,\mathrm{d}t < \infty$ and $p_\circ \in H$. Then, the operator equation $e(y,u) = 0$ is equivalent to the fact that $y = y(u)$ is a weak solution to (3) for given $u \in U_{\mathsf{ad}}$. Now the optimal control problem is expressed as

$$\min J(y,u) \quad \text{s.t.} \quad e(y,u) = 0 \text{ and } u \in U_{\mathsf{ad}}, \tag{$\mathbf{P}$}$$

where "s.t." stands for "subject to". Under appropriate assumptions on $y_\circ$ we can ensure that (3) has a unique weak solution $y(u)$ for any admissible control $u \in U_{\mathsf{ad}}$; i.e., we have $e(y(u),u) = 0$. This motivates the introduction of the reduced objective $\hat{J}(u) = J(y(u),u)$. We consider – instead of ($\mathbf{P}$) – the reduced problem

$$\min \hat{J}(u) \quad \text{s.t.} \quad u \in U_{\mathsf{ad}}. \tag{$\hat{\mathbf{P}}$}$$

The notion "reduced" expresses the fact that $\hat{J}$ depends on the control only, whereas J is dependent on the state and control. It follows from (Tröltzsch, 2010, Chapter 5.3) that ($\hat{\mathbf{P}}$) admits an optimal solution $\bar{u} \in U_{\mathsf{ad}}$. If $\bar{y} = y(\bar{u})$ is the unique weak solution to (3), then $\bar{x} = (\bar{y}, \bar{u})$ is an optimal solution to ($\mathbf{P}$). To compute a locally optimal solution we make use of the following first-order necessary optimality condition for a locally optimal solution (Tröltzsch, 2010, Chapter 2.8):

$$\hat{J}'(\bar{u})(u - \bar{u}) \geq 0 \quad \text{for all } u \in U_{\mathsf{ad}}, \tag{5}$$

where $\hat{J}'(\bar{u})$ denotes the (Gateaux) derivative of $\hat{J}$ at $\bar{u}$. Following the general approach in Chapter 5.5 in Tröltzsch (2010) or the one for our present optimal control problem in Gräßle (2014); Rogg (2014) we derive that (5) is equivalent to the variational inequality

$$\int_0^T \sum_{i=1}^m \Big(\sigma_u \bar{u}_i(t) - \int_{\Gamma_i} \bar{p}(t) \,\mathrm{d}\boldsymbol{x} \Big) (u_i(t) - \bar{u}_i(t)) \,\mathrm{d}t \geq 0 \tag{6}$$

for all $u \in U_{\mathsf{ad}}$, where $\bar{p}(t) = \bar{p}(t,\cdot) \in V$ is the (unique) weak solution to the dual or adjoint equation

$$\begin{aligned} &- c\bar{p}_t(t,\boldsymbol{x}) - \Delta\bar{p}(t,\boldsymbol{x}) = -3\bar{y}(t,\boldsymbol{x})^2 p(t,\boldsymbol{x}), && (t,\boldsymbol{x}) \in Q,\\ &\frac{\partial \bar{p}}{\partial \boldsymbol{n}}(t,\boldsymbol{x}) + q\bar{p}(t,\boldsymbol{x}) = 0, && (t,\boldsymbol{x}) \in \Sigma,\\ &\bar{p}(T,\boldsymbol{x}) = y_d(\boldsymbol{x}) - \bar{y}(T,\boldsymbol{x}), && \boldsymbol{x} \in \Omega \end{aligned} \tag{7}$$

and $\bar{y} = y(\bar{u})$ is solves (4) for the optimal control $u = \bar{u}$. The associated dual variable $\bar{p}_\circ$ is determined by $\bar{p}_\circ = c\bar{p}(0,\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega$. Note that ($\hat{\mathbf{P}}$) and also ($\mathbf{P}$) are nonlinear (and therefore non convex) problems. To ensure that the solution $\bar{u}$ to (5) is a locally optimal solution to ($\hat{\mathbf{P}}$) we have to investigate second-order sufficient optimality conditions; see, e.g., Chapter 5.7 in Tröltzsch (2010) and Section 2.3 in Gräßle (2014).

3 Optimization methods

In this section we briefly review the numerical algorithms utilized to numerically solve ($\mathbf{P}$) and ($\hat{\mathbf{P}}$), respectively. For more details we refer to Gräßle (2014); Metzdorf (2015); Rogg (2014).

3.1 Sequential quadratic programming

We introduce the Lagrangian associated with ($\mathbf{P}$) as

$$\mathcal{L}(y,u,p,p_\circ) = J(y,u) + \langle e(y,u), (p,p_\circ)\rangle,$$

where the mapping $e(y,u)$ has already been defined in Section 2. It can be shown that the Lagrangian is twice continuously (Fréchet) differentiable. The principal idea of the sequential quadratic programming (SQP) method is to solve ($\mathbf{P}$) in an iterative procedure, whereby in each SQP iteration k a linear-quadratic SQP subproblem has to be solved. This SQP subproblem is obtained by a quadratic approximation of the Lagrangian $\mathcal{L}$ and a linearization of the equality constraint at the current iterates $x^k = (y^k, u^k)$ and $z^k = (p^k, p_\circ^k)$. For brevity, we set $\mathcal{L}^k = \mathcal{L}(x^k, z^k)$ and $e^k = e(x^k)$. The linear-quadratic SQP subproblem is given as

$$\begin{aligned} \min_{x_\delta^k=(y_\delta^k,u_\delta^k)} \; & \mathcal{L}^k + \mathcal{L}_x^k x_\delta^k + \frac{1}{2}\mathcal{L}_{xx}^k(x_\delta^k, x_\delta^k) \\ \text{s.t. } & e^k + e_x^k x_\delta^k = 0 \text{ and } u^k + u_\delta^k \in U_{\mathsf{ad}}. \end{aligned} \tag{8}$$

By x we denote the partial (Fréchet) derivative with respect to the argument $x = (y, u)$. Problem (8) is well-defined if second-order sufficient optimality conditions hold at the point (x^k, z^k). If x_δ is determined, the new SQP iterate is $x^{k+1} = x^k + x_\delta^k$. For the Lagrange multiplier z^k the update is given in Algorithm 1. It can be shown that the SQP method is locally equivalent to the Newton method, which consists in finding a stationary point of the Lagrangian. Hence, the rate of convergence is locally quadratic, provided that the second-order sufficient optimality condition holds; cf. Hinze et al. (2009).

3.2 Primal-dual active set strategy

Since (8) involves inequality constraints, we utilize a primal-dual active set strategy (PDASS); cf. Bergounioux et al. (1997). Let k and j denote the current SQP and PDASS iteration level, respectively. The PDASS method works on the primal variable u_δ^k, which has to satisfy the inequality constraint $u^k + u_\delta^k \in U_{\mathsf{ad}}$, and on the dual variable μ^k, which corresponds to the control constraint. For a given iterate $(u_\delta^{k,j}, \mu^{k,j})$ of the PDASS algorithm we define the active and inactive sets

$$\begin{aligned} \mathcal{A}_a^{k,j} &= \big\{t \in [0,T] : u^{k,j}(t) + u_\delta^{k,j}(t) + \mu^{k,j}(t) < u_a(t)\big\}, \\ \mathcal{A}_b^{k,j} &= \big\{t \in [0,T] : u^{k,j}(t) + u_\delta^{k,j}(t) + \mu^{k,j}(t) > u_b(t)\big\}, \\ \mathcal{A}^{k,j} &= \mathcal{A}_a^{k,j} \cup \mathcal{A}_b^{k,j}, \quad \mathcal{I}^{k,j} = [0,T] \setminus \mathcal{A}^{k,j}, \end{aligned}$$

where the inequalities are understood componentwise. Then, the solution of (8) is computed by solving its associated KKT system:

$$\begin{aligned} &\begin{pmatrix} \mathcal{L}_{yy}^k & 0 & e_y^{k,\star} \\ 0 & \mathcal{L}_{uu}^k & e_u^{k,\star} \\ e_y^k & e_u^k & 0 \end{pmatrix} \begin{pmatrix} y_\delta \\ u_\delta \\ z_\delta \end{pmatrix} + \begin{pmatrix} 0 \\ \mu^{k,j+1}|_{\mathcal{A}^{k,j}} \\ 0 \end{pmatrix} = \begin{pmatrix} -\mathcal{L}_y^k \\ -\mathcal{L}_u^k \\ -e^k \end{pmatrix} \\ &u_\delta|_{\mathcal{A}^{k,j}} = \chi_{\mathcal{A}_a^{k,j}}(u_a - u^k) + \chi_{\mathcal{A}_b^{k,j}}(u_b - u^k), \end{aligned} \tag{9}$$

where we set $x_\delta^{k,j+1} = (y_\delta, u_\delta)$ and $z_\delta^{k,j+1} = z_\delta$. By $e_y^{k,\star}$ and $e_u^{k,\star}$ we denote the dual operators of e_y^k and e_u^k, respectively. The SQP method with PDASS is summarized in Algorithm 1. In general, Algorithm 1 is only locally convergent. In order to ensure convergence for any starting point (x^0, z^0), a globalization strategy is utilized, which is made from two ingredients: a modification of the second-order term $\mathcal{L}_{xx}^k$ to ensure coercivity and the inclusion of an Armijo backtracking line search for an ℓ_1-merit function. For more details we refer to Hintermüller (2001) for the general theory and to Gräßle (2014) for our specific optimal control problem.

Algorithm 1: SQP method with PDASS

Require: Initial value (x^0, z^0), set $k = 0$;
repeat
  Initialize $(u_\delta^{k,0}, \mu^{k,0})$ and set $j = 0$;
  repeat
    Set $j = j + 1$;
    Determine $\mathcal{A}_a^{k,j}$, $\mathcal{A}_b^{k,j}$, $\mathcal{A}^{k,j}$ and $\mathcal{I}^{k,j}$;
    Solve (9) for $(y_\delta, u_\delta, z_\delta)$ and $\mu^{k,j}|_{\mathcal{A}^{k,j}}$;
    Set $j = j + 1$, $x_\delta^{k,j} = (y_\delta, u_\delta)$, $z_\delta^{k,j} = z_\delta$, and $\mu^{k,j}|_{\mathcal{I}^{k,j-1}} = 0$;
  until $j \geq 1$, $\mathcal{A}_a^{k,j} = \mathcal{A}_a^{k,j-1}$, $\mathcal{A}_b^{k,j} = \mathcal{A}_b^{k,j-1}$;
  Set $x^{k+1} = x^k + x_\delta^{k,j}$ and $z^{k+1} = z^k + z_\delta^{k,j}$;
until SQP stopping criterium is fulfilled;

3.3 Trust region method

In this paper we compare Algorithm 1 with the trust-region (TR) method for the solution of $(\hat{\mathbf{P}})$; cf. Conn et al. (2000). The key idea of TR methods is as follows: At a current iterate the objective $\hat{J}$ is replaced by a "simpler" (often quadratic) model function which is then approximately minimized in a TR to find the next iterate. Assume we are in iteration k with current iterate u^k. Then, we build a model function m^k to approximate the objective $\hat{J}$ at u^k. In Algorithm 2, we restrict ourselves to the quadratic model

$$m^k(d) = \hat{J}(u^k) + g^k d + \frac{1}{2}\mathcal{H}^k(d,d), \tag{10}$$

where d is a mapping from $[0,T]$ to $\mathbb{R}^m$, $g^k = \hat{J}'(u^k)$ holds and $\mathcal{H}^k$ denotes the Hessian $\hat{J}''(u^k)$ or a self-adjoint approximation to it. In line 2 a trial step d^k is computed by approximately solving the TR subproblem. In lines 4 to 13 it is then decided whether or not to accept the trial point $u^k + d^k$ as next iterate and the TR radius Δ_k gets adjusted. This decision is based upon the quotient ρ_k in line 3 which compares the reduction pred_k of the model function to the reduction ared_k of the cost functional and

Algorithm 2: (Trust region method)

Require: Initial TR radius $\Delta_0 > 0$, initial point u^0, constants η_1, η_2, γ_1, γ_2, γ_3 satisfying $0 < \eta_1 \leq \eta_2 < 1$, $0 < \gamma_1 \leq \gamma_2 < 1 \leq \gamma_3$; set $k = 0$;

1: Build up the model function m^k;
2: Determine an approximate solution d^k to
$$\min_d m^k(d) \quad \text{s.t.} \quad u^k + d \in U_{\mathsf{ad}} \text{ and } \|d\| \leq \Delta_k;$$
3: Compute the quotient
$$\rho_k = \frac{\mathsf{ared}_k}{\mathsf{pred}_k} = \frac{\hat{J}(u^k) - \hat{J}(u^k + d^k)}{m^k(0) - m^k(d^k)};$$
4: **if** $\rho_k \geq \eta_2$ **then**
5: Set $u^{k+1} = u^k + d^k$ and $\Delta_{k+1} \in [\Delta_k, \gamma_3 \Delta_k]$;
6: Set $k = k + 1$ and go to line 1;
7: **else if** $\eta_1 \leq \rho_k < \eta_2$ **then**
8: Set $u^{k+1} = u^k + d^k$ and $\Delta_{k+1} \in [\gamma_2 \Delta^{(k)}, \Delta_k]$;
9: Set $k = k + 1$ and go to line 1;
10: **else if** $\rho_k < \eta_1$ **then**
11: Set $u^{k+1} = u^k$ and $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$;
12: Set $k = k + 1$ and go to line 2;
13: **end if**

which provides a measure for the approximation quality of the model function. In case $\rho_k \geq \eta_1$ (e.g. $\eta_1 = 0.2$), the trial point is accepted. If even $\rho_k \geq \eta_2$ (e.g. $\eta_2 = 0.8$), the approximation quality of the model function is as good that the TR radius can be increased for the next iteration. Otherwise the radius should be kept or decreased. In case $\rho_k < \eta_1$, the quadratic model is a poor approximation to the cost functional on the TR, the trial point is rejected and the TR subproblem gets again solved for a decreased TR radius. Disregarding standard assumptions, Algorithm 2 is globally convergent in the sense $\lim_{k \to \infty} \|g^k\| = 0$ if the trial steps d^k lead to a model decrease pred_k at least as good as that of the so-called Cauchy step. If the model function possesses inexact gradient information $g^k \neq \hat{J}'(u^k)$, global convergence still holds if the Carter condition

$$\|\hat{J}'(u^k) - g^k\| \leq \zeta \|g^k\|, \quad \zeta \in (0, 1 - \eta_2) \tag{11}$$

is satisfied in every iteration; see Carter (1991). We compute truncated Newton steps using the Steihaug-CG algorithm, see Nocedal, Wright (2006), together with active set projection according to Kelley (1999). Hence, we fulfill this condition. Note that the computation of $\hat{J}''(u^k)d$ requires the solutions to a linearized state and linearized adjoint equation; cf. Hinze et al. (2009).

4 POD reduced-order modelling

In this section we recall the basics of the POD method for optimal control problems. For more details we refer, e.g., to Benner et al. (2014); Sachs, Volkwein (2010).

4.1 POD method

Suppose that we are given trajectories $z^\nu(t) = z^\nu(t, \cdot) \in H$, $t \in [0, T]$, $1 \leq \nu \leq \wp$ and $\wp \in \mathbb{N}$. We introduce the snapshot subspace as

$$\mathcal{V} = \operatorname{span}\left\{ z^\nu(t) \,\middle|\, t \in [0, T] \text{ and } \nu = 1, \ldots, \wp \right\}$$

with $\mathsf{d} = \dim \mathcal{V} \leq \infty$. The inner product in H is given as

$$\langle \varphi, \phi \rangle_H = \int_\Omega \varphi(\boldsymbol{x}) \phi(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \quad \text{for all } \varphi, \phi \in H.$$

For every ℓ with $1 \leq \ell \leq \mathsf{d}$, a POD basis of rank ℓ is defined as a solution to the minimization problem

$$\begin{gathered} \min \sum_{\nu=1}^{\wp} \int_0^T \Big\| z^\nu(t) - \sum_{i=1}^{\ell} \langle z^\nu(t), \psi_i \rangle_H \psi_i \Big\|_H^2 \,\mathrm{d}t \\ \text{s.t } \{\psi_i\}_{i=1}^{\ell} \subset H, \ \langle \psi_i, \psi_j \rangle_H = \delta_{ij}, \ 1 \leq i, j \leq \ell. \end{gathered} \tag{12}$$

A solution to (12) is given by the eigenvalue problem

$$\mathcal{R}\psi_i = \lambda_i \psi_i \quad \text{for } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell \geq \lambda_{\mathsf{d}} > 0 \tag{13}$$

with the linear, bounded integral operator

$$\mathcal{R} : H \to \mathcal{V} \subset H, \quad \mathcal{R}\psi = \sum_{\nu=1}^{\wp} \int_0^T \langle z^\nu(t), \psi \rangle_H z^\nu(t) \,\mathrm{d}t.$$

In our numerical experiments we consider $\wp = 2$. We include by z^1 and z^2 information of the (linearized) state and the dual equation, respectively.

4.2 Reduced-order modelling for the control problem

Suppose that a POD basis $\{\psi_i\}_{i=1}^{\ell}$ is computed. We introduce the subspace $V^\ell = \operatorname{span}\{\psi_1, \ldots, \psi_\ell\}$ and define the orthogonal projection $\mathcal{P}^\ell \varphi = \sum_{i=1}^{\ell} \langle \varphi, \psi_i \rangle_H \psi_i$ for $\varphi \in H$. The POD scheme for (4) is as follows: $y^\ell(t) \in V^\ell$ satisfies $y^\ell(0) = \mathcal{P}^\ell y_\circ$ in H and

$$\begin{aligned} &\int_\Omega c y_t^\ell(t) \psi + \nabla y^\ell(t) \cdot \nabla \psi + y(t)^3 \psi \,\mathrm{d}\boldsymbol{x} \\ &\quad + \int_\Gamma q y^\ell(t) \psi \,\mathrm{d}\boldsymbol{x} = \sum_{i=1}^{m} u_i(t) \int_{\Gamma_i} \psi \,\mathrm{d}\boldsymbol{x} \end{aligned} \tag{14}$$

for all $\psi \in V^\ell$ and $t \in (0, T]$. Then, we replace $(\mathbf{P})$ by the following POD approximation

$$\min J(y^\ell, u) \text{ s.t. } (y^\ell, u) \text{ satisfies (14) and } u \in U_{\mathsf{ad}}. \quad (\mathbf{P}^\ell)$$

We assume that (14) has a unique solution $y^\ell(u)$ for any $u \in U_{\mathsf{ad}}$. Then, setting $\hat{J}_\ell(u) = J(y^\ell(u), u)$ for any $u \in U_{\mathsf{ad}}$ we obtain the POD approximation of $(\hat{\mathbf{P}})$

$$\min \hat{J}_\ell(u) \quad \text{s.t.} \quad u \in U_{\mathsf{ad}}. \qquad (\hat{\mathbf{P}}^\ell)$$

Suppose that $\bar{u}^\ell$ is a locally solution to $(\hat{\mathbf{P}}^\ell)$. We interpret $\bar{u}^\ell$ as a suboptimal solution to $(\hat{\mathbf{P}})$. First-order necessary optimality conditions for $(\hat{\mathbf{P}}^\ell)$ are given by (compare (6))

$$\int_0^T \sum_{i=1}^m \Big(\sigma_u \bar{u}_i^\ell(t) - \int_{\Gamma_i} \bar{p}^\ell(t)\,\mathrm{d}\boldsymbol{x}\Big)(u_i(t) - \bar{u}_i^\ell(t))\,\mathrm{d}t \geq 0$$

for all $u \in U_{\mathsf{ad}}$, where $\bar{p}^\ell(t) = \bar{p}(t, \cdot) \in V^\ell$ is the weak solution to a POD Galerkin scheme for (7) utilizing the optimal POD state $\bar{y}^\ell = y^\ell(\bar{u})$. More details are presented in Gräßle (2014); Metzdorf (2015); Rogg (2014).

4.3 POD basis update strategies

The choice of a reference control $u^{\mathsf{ref}} \in U_{\mathsf{ad}}$ turned out to be essential for the computation of a POD basis of rank ℓ. When using an arbitrary control, the obtained accuracy was not at all satisfying even when using a huge number of basis functions. On the other hand, an optimal POD basis (computed from the FE optimally controlled state) led to far better results; Hinze, Volkwein (2008). To overcome this problem different techniques for improving the POD basis have been proposed.

4.3.1 Optimality system POD

Let us just roughly recall optimality system POD (OS-POD) introduced in Kunisch, Volkwein (2008). The idea of OS-POD is to include the equations determining the POD basis in the optimization process

$$\begin{aligned} &\min \hat{J}_\ell(u) \\ &\text{s.t. } u \in U_{\mathsf{ad}} \text{ and } y(u) \text{ is the solution to (4),} \\ &\quad \{\psi_i\}_{i=1}^\ell \text{ solves (13) for } \wp = 1, z^1 = y(u), \\ &\quad y^\ell(u) \text{ satisfies (14).} \end{aligned} \qquad (\mathbf{P}_{\mathsf{os}}^\ell)$$

A thereby obtained basis is optimal for the considered problem. Note that $(\mathbf{P}_{\mathsf{os}}^\ell)$ is more complex than the original problem $(\mathbf{P})$. Therefore, we do not solve $(\mathbf{P}_{\mathsf{os}}^\ell)$. Starting with an initial control $u^0 \in U_{\mathsf{ad}}$, we compute a few projected gradient steps for $(\mathbf{P}_{\mathsf{os}}^\ell)$ in order to find a POD basis, which is appropriate for the underlying optimal control problem. Then, this POD basis is kept fix and the POD approximation $(\mathbf{P}^\ell)$ of $(\mathbf{P})$ is derived. We discuss two possibilities to use OS-POD:

1) First, starting with a reference control $u^{\mathsf{ref}} \in U_{\mathsf{ad}}$ we apply a few projected gradient steps to solve $(\mathbf{P}_{\mathsf{os}}^\ell)$ numerically. Then, the obtained POD basis is kept fix and the POD approximation $(\mathbf{P}^\ell)$ is solved numerically by the SQP method. This approach is used in Grimm et al. (2014); Metzdorf (2015) for linear parabolic equations.
2) In the second approach we utilize OS-POD within the SQP framework. In each SQP iteration a linear-quadratic optimal control problem has to be solved up to a certain accuracy. For each linear-quadratic SQP subproblem we formulate the OS-POD optimization problem as in $(\mathbf{P}_{\mathsf{os}}^\ell)$. Again, we apply a few (projected) gradient steps, but now to the linear-quadratic problem in order to find a POD basis which is appropriate for the current SQP subproblem. Then, we apply the PDASS algorithm to solve the linear-quadratic SQP subproblem using the POD basis found by the OS-POD gradient steps.

4.3.2 Trust region POD

Within the trust-region POD (TR-POD) algorithm we guarantee that the reduced-order models are sufficiently accurate by ensuring gradient accuracy according to the Carter condition (11); see Schu (2010). This is realized by successively adapting the reduced-order models in the course of optimization just by increasing the number ℓ of POD basis functions or, if this does not lead to a sufficient accuracy, by computing a new POD basis from the current control iterate (and then adapting the number ℓ). Recall the approximative model gradient g^k and the approximative Hessian $\mathcal{H}^k$ from (10). In the TR-POD approach g^k and $\mathcal{H}^k$ are computed from the respective POD approximations of $\hat{J}$:

$$m_\ell^k(d) = \hat{J}(u^k) + g_\ell^k d + \frac{1}{2}\,\mathcal{H}_\ell^k(d, d).$$

The TR-POD method is summarized in Algorithm 3. The lower part is completely identical to Algorithm 2. But here, at the beginning of each iteration, the approximation quality of the current POD basis is tested and if necessary recomputed.

Algorithm 3: TR-POD method

Require: Initial TR radius $\Delta_0 > 0$, initial point u^0, a maximal number $\ell_{\max}$ of POD basis functions and constants η_1, η_2, γ_1, γ_2, γ_3 satisfying $0 < \eta_1 \le \eta_2 < 1$, $0 < \gamma_1 \le \gamma_2 < 1 \le \gamma_3$; set $k = 0$;

Compute a reduced-order model for some $\ell \le \ell_{\max}$;

Guarantee (11) by computing new POD basis and/or expanding $\ell \le \ell_{\max}$;

Build up the model function $m_\ell^k(d)$;

Determine an approximate solution d^k to

$$\min_d m_\ell^k(d) \quad \text{s.t.} \quad u^k + d \in U_{\mathsf{ad}} \text{ and } \|d\| \le \Delta_k;$$

Compute the quotient

$$\rho_k = \frac{\mathsf{ared}_k}{\mathsf{pred}_k} = \frac{\hat{J}(u^k) - \hat{J}(u^k + d^k)}{m_\ell^k(0) - m_\ell^k(d^k)};$$

Lines 4 to 13 of Algorithm 2.

4.3.3 POD based inexact SQP method

For large-scale systems it is costly to solve each SQP subproblem (9) exactly. Thus, an inexact version of the SQP method is used, where the inexactness is caused by replacing (9) by its POD surrogate model; cf. Kahlbacher, Volkwein (2011). An a-posteriori error bound for linear-quadratic optimal control problems (see Tröltzsch, Volkwein (2009)) is used to control the accuracy of the POD model and hence guarantee convergence. For more details we refer to Grimm et al. (2014); Sachs, Volkwein (2010).

5 Numerical Results

Problem setting. We choose $\Omega = (0,1) \times (0,1) \subset \mathbb{R}^2$ and $T = 0$. The parameters are $c = 10$ and $q = 0.01$. We divide the boundary Γ into $m = 4$ parts so that the characteristic functions $\{\chi_{\Gamma_i}\}_{i=1}^4$ represent each of the four boundary parts. The cost parameter is $\sigma_u = 5 \cdot 10^{-3}$. For the spatial discretization we use piecewise linear finite elements (FE) with diameter $h_{\max} = 0.06$ leading to 498 degrees of freedom. As time integration scheme we utilize the implicit Euler method with $N_t = 250$ equidistant time steps. The reference control for snapshot generation in the context of Section 4.1 is $u^{\mathsf{ref}} = 0$. The snapshot set for POD basis computation differs depending on the chosen optimization strategy. In case of TR-POD, state and adjoint state snapshots are taken, whereas for SQP linearized state and linearized adjoint state snapshots are utilized. Moreover, we define the absolute and relative

method	ε^u_{abs}	ε^u_{rel}
SQP-POD, no basis update	$2.21 \cdot 10^0$	$2.78 \cdot 10^{-1}$
OS-POD approach 1	$3.15 \cdot 10^{-1}$	$3.96 \cdot 10^{-2}$
OS-POD approach 2	$2.43 \cdot 10^{-1}$	$3.06 \cdot 10^{-2}$
SQP-POD, optimal basis	$1.07 \cdot 10^{-1}$	$1.35 \cdot 10^{-2}$
TR-POD	$1.89 \cdot 10^{-1}$	$2.45 \cdot 10^{-2}$

Table 1. Run 1: errors between POD and FE optimal controls.

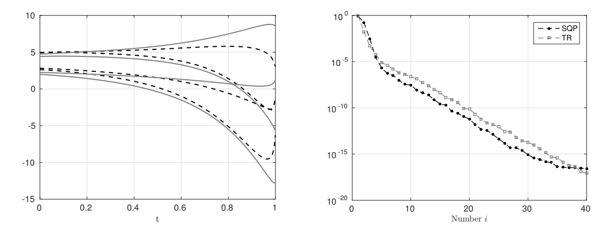

Fig. 1. Run 1, left: SQP-POD controls $\{u_i(t)\}_{i=1}^4$ for a fixed POD basis with $\ell = 10$ (black line) and SQP-FE solution (black dotted). Right: Decay of the eigenvalues for SQP-POD and TR-POD.

errors $\varepsilon^u_{\text{abs}} = \|\bar{u}^h - u\|$ and $\varepsilon^u_{\text{rel}} = \varepsilon^u_{\text{abs}}/\|\bar{u}^h\|$ with

$$\|u\|^2 := \sum_{i=1}^m \Delta t \Big(\frac{(u_i^{(0)})^2}{2} + \sum_{j=1}^{N_t} (u_i^{(j)})^2 + \frac{(u_i^{(N_t+1)})^2}{2} \Big)$$

to measure the difference between any computed control u and the optimal FE control $\bar{u}^h$.

Run 1 (unconstrained case). At first we consider $(\mathbf{P})$ without control constraints. We choose the desired state $y_d(\boldsymbol{x}) = \sin(2\pi(x_1 - 0.25)) + 4x_1x_2$ and the initial condition $y_\circ(\boldsymbol{x}) = \sin(2\pi(x_1 - 0.25))$ for $\boldsymbol{x} = (x_1, x_2) \in \Omega$. In the left plot of Fig. 1 we present the SQP-FE optimal controls (dotted lines) compared to the SQP-POD optimal controls obtained with a fixed POD basis of rank $\ell = 10$. As can be seen, the accuracy is not satisfying; see also Table 1, row 1. This confirms the necessity of a POD basis update strategy. The POD eigenvalues computed for SQP-POD and TR-POD are shown in the right plot of Figure 1. As first POD basis update strategy, we use the OS-POD approach 1) of Section 4.3.1. One gradient step is applied to $(\mathbf{P}^\ell_{\mathsf{os}})$. Then, the POD basis of rank $\ell = 10$ is fixed and $(\mathbf{P}^\ell)$ is solved by the SQP method. This update strategy leads to an improvement of the error in the control variable of one order (see Table 1, row 2). The next POD basis update strategy consists in a combination of the POD based inexact SQP method with OS-POD approach 2) of Section 4.3.1. Two SQP iterations are needed for convergence. The first SQP iteration is cheap, since $\ell = 5$ POD bases suffice. In the second SQP iteration, it is detected that a higher accuracy of the POD surro-

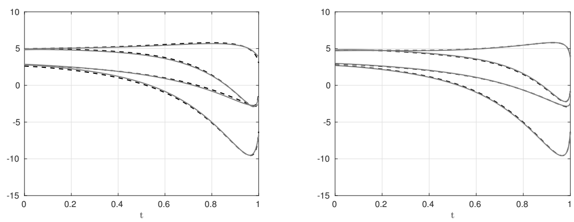

Fig. 2. Run 1: POD controls $\{u_i(t)\}_{i=1}^4$ (black line) and FE solution (black dotted), OS-POD basis update (left) and TR-POD (right).

gate model for the linear-quadratic SQP subproblem (8) is needed. Thus, we apply one gradient step to solve the OS-POD optimization problem for the linear-quadratic SQP subproblem and the number of utilized POD basis functions is increased to $\ell = 10$. The POD solution $\bar{u}^\ell$ for the time dependent control intensities coincides convincingly with the FE solution (see Fig. 2, left and Table 1, row 3). Since the POD basis is adapted to the desired accuracy within the solution process, this strategy leads to better results than the first POD basis update strategy. For comparison, the error between the FE solution $\bar{u}^h$ and the SQP-POD solution $\bar{u}^\ell$ with optimal POD basis of length $\ell = 10$ is listed in Table 1, row 4. Now, we compare these results to those obtained by TR-POD. Recall that the TR-POD algorithm guarantees convergence to a stationary point. This is why the deviation of the TR-POD optimal control from the FE optimal control is within discretization accuracy; see the last row in Table 1 and Fig. 2, right plot. The iterations of the TR-POD method are very similar to those of the POD based inexact SQP method with OS-POD approach 2. In the first iteration $\ell = 4$ POD basis functions suffice to fulfill the Carter condition (11). In the second iteration a new POD basis gets computed and $\ell = 10$ POD modes are required. We visualize the adaptation of the POD basis during optimization by means of the shape of the first three basis functions. In Fig. 3 the adaptation corresponding the POD based inexact SQP method with OS-POD approach 2 is shown. For comparison, in Fig. 4 we present the POD basis adaptation obtained from the TR-POD algorithm.

Run 2 (constrained case). Now we impose the box constraints $u_a = 0$ and $u_b = 7$. We choose $y_d(\boldsymbol{x}) = 2+2|2x_1 - x_2|$ and $y_\circ(\boldsymbol{x}) = 3-4(x_2-0.5)^2$ for $\boldsymbol{x} = (x_1, x_2) \in \Omega$. The approximation results for the POD solution $\bar{u}^\ell(t)$ with snapshots computed from the reference controls $u^{\mathsf{ref}} = 0$ and $u^{\mathsf{ref}} = \bar{u}^h(t)$ are listed in Table 2, row 1 and 4, respectively. The application of the POD basis update strategy by OS-POD approach 1) leads to a reduction of the error in the control by a factor of two in comparison to the case

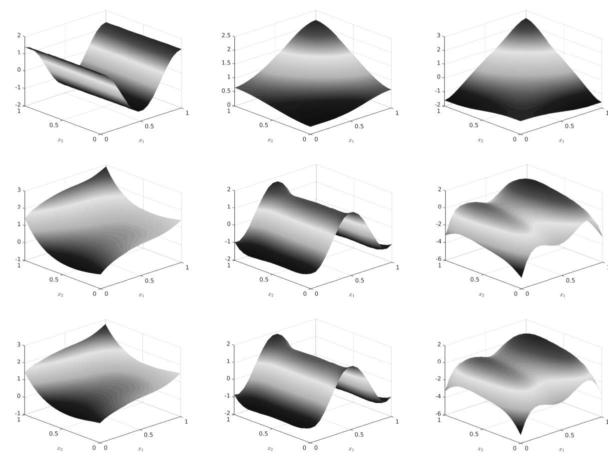

Fig. 3. Run 1, OS-POD approach 2): POD bases ψ_1, ψ_2, ψ_3 computed in the first (top) and second iteration (middle) of SQP-POD method; optimal basis functions computed by reference control $u^{\mathsf{ref}} = \bar{u}^h$ (bottom).

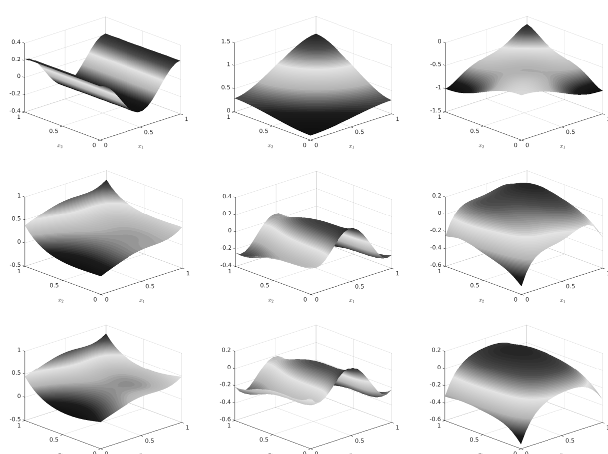

Fig. 4. Run 1, TR-POD: POD bases ψ_1, ψ_2, ψ_3 computed in the first (top) and second iteration (middle); optimal basis functions computed by the reference control $u^{\mathsf{ref}} = \bar{u}^h$ (bottom).

method	ε^u_{abs}	ε^u_{rel}
SQP-POD, no basis update	$7.92 \cdot 10^{-1}$	$8.81 \cdot 10^{-2}$
OS-POD approach 1	$3.73 \cdot 10^{-1}$	$4.15 \cdot 10^{-2}$
OS-POD approach 2	$1.12 \cdot 10^{-1}$	$1.25 \cdot 10^{-2}$
SQP-POD, optimal basis	$1.30 \cdot 10^{-1}$	$1.45 \cdot 10^{-2}$
TR-POD	$8.60 \cdot 10^{-2}$	$9.60 \cdot 10^{-3}$

Table 2. Run 2: errors between POD and FE optimal controls.

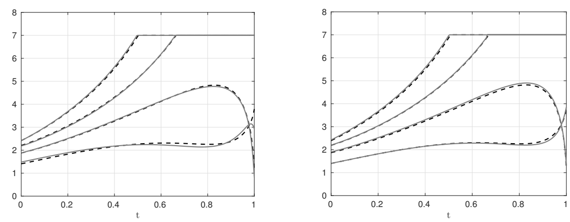

Fig. 5. Run 2: controls $\{u_i(t)\}_{i=1}^4$, black dotted: FE solution, black line: POD solution. Right: TR-POD.

with no POD basis update. Finally, we utilize OS-POD approach 2) in combination with the inexact SQP-POD method. In this case, four SQP iterations are needed for convergence. In the first two iterations, $\ell = 3$ POD modes suffice. In the third iteration, one gradient step is performed to compute approximatively a solution for the OS-POD optimization problem for the linear-quadratic SQP subproblem. Moreover, the number of utilized POD basis functions is increased to $\ell = 7$. In the last iteration, the number of POD modes is increased to the maximum number $\ell = 10$. The resulting POD controls are compared to the FE controls in the left plot of Fig. 5 as well as in row 3 of Table 2. From the computational point of view, OS-POD approach 1) is the fastest. The computational time for snapshot generation and OS-POD basis computation is 2.6 sec. The SQP-POD run then takes 7.3 sec. In comparison, the SQP-FE run takes 85.8 sec. The TR-POD Algorithm needs five iterations for optimization. In the first three iterations the initial POD basis suffices to guarantee the Carter condition with only 3, 3 and 5 POD basis functions. In the fourth iteration a new POD basis is computed and $\ell = 7$ POD modes are needed. This basis of rank seven can then be kept for the last step.

Acknowledgment: C. Gräßle and S. Metzdorf have been supported partially by the project *A-posteriori error estimation for TR-POD methods in Design and Control* financed by the University of Konstanz.

References

Arian, E., Fahl M., Sachs, E.W. "Trust-region proper orthogonal decomposition for flow control." Technical Report 2000-25, ICASE, 2000.

Bergounioux, M., Ito, K., Kunisch, K. "Primal-dual strategy for constrained optimal control problems." *SIAM J. Control Optim.,* **35** *1524-1543* (1997)

Benner, P., Sachs, E.W., Volkwein, S. "Model order reduction for PDE constrained optimization." *Internat. Ser. Numer. Math.,* **165** *303-326*, (2014)

Carter, R.G. "On the global convergence of trust region algorithms using inexact gradient information." *SIAM J. Numer. Anal.,* **28** *251-265* (1991)

Conn, A.R., Gould, N.I.M., Toint, P.L. "Trust-region methods." *SIAM* (2000)

Gräßle, C. "POD based inexact SQP methods for optimal control problems governed by a semilinear heat equation." *Diploma Thesis, University of Konstanz* (2014), https://kops.uni-konstanz.de/handle/123456789/29390

Grimm, E., Gubisch, M., Volkwein, S. "A-Posteriori Error Analysis and Optimality-System POD for Constrained Optimal Control." *Comp. Sci. Eng.,* **105**, *297-317* (2014)

Holmes, P., Lumley, J.L., Berkooz, G., Romley, C.W. "Turbulence, Coherent Structures, Dynamical Systems and Symmetry." *Cambridge Monographs on Mechanics*, Cambridge University Press (2012)

Hintermüller, M. "On a globalized augmented Lagrangian-SQP algorithm for nonlinear optimal control problems with box constraints." *Internat. Ser. Numer. Math.,* **138** *139-153* (2001)

Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, R. "Optimization with ODE constraints." *Math. Mod.,* **23**, *Springer Series* (2009)

Hinze, M., Volkwein, S. "Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition." *Comput. Optim. Appl.,* **39** *319-345* (2008)

Kahlbacher, M., Volkwein, S. "POD a-posteriori error based inexact SQP method for bilinear elliptic optimal control problems." *ESAIM: M2AN,* **46** *491-511* (2011)

Kelley, C.T. "Iterative Methods for Optimization." *Frontiers in Applied Mathematics, SIAM, Philadelphia, PA* (1999)

Kunisch, K., Volkwein, S. "Proper orthogonal decomposition for optimality systems." *ESAIM: M2AN* **42**, *1-23* (2008)

Metzdorf, S. "Optimality system POD for time-variant, linear-quadratic control problems." *Diploma Thesis, University of Konstanz* (2015), http://nbn-resolving.de/urn:nbn:de:bsz:352-0-329322

Nocedal, J., Wright, S.J. "Numerical Optimization." *2nd ed., Springer, New York* (2006)

Qian, E., Grepl, M., Veroy, K., Willcox, K. "A certified trust region reduced basis approach to PDE-constrained optimization." ACDL Technical Report TR16-3, 2016

Rogg, S. "Trust region POD for optimal boundary control of a semilinear heat equation." *Diploma Thesis, University of Konstanz* (2014), https://kops.uni-konstanz.de/handle/123456789/29194

Sachs, E.W., Volkwein, S. "POD Galerkin approximations in PDE-constrained optimization." *GAMM-Mitt.* **33**, *194-208* (2010)

Schu, M. "Adaptive trust-region POD methods and their applications in finance." *Ph.D thesis, University of Trier* (2012)

Tröltzsch, F. "Optimal Control of Partial Differential Equations: Theory, Methods and Applications." *AMS American Mathematical Society, 2nd ed.* (2010)

Tröltzsch, F., Volkwein, S. "POD a-posteriori error estimates for linear-quadratic optimal control problems." *Comput. Optim. Appl.,* **44** *83-115* (2009)

Yue, Y., Meerbergen, K. "Accelerating PDE-constrained optimization by model order reduction with error control." *SIAM J. Optim.* **23**, *1344-1370* (2013)