Parameter identification for nonlinear elliptic-parabolic systems with application in lithium-ion battery modeling

Oliver Lass · Stefan Volkwein

Abstract In this paper the authors consider a parameter estimation problem for a nonlinear systems, which consists of one parabolic equation for the concentration and two elliptic equations for the potentials. The measurements are given as boundary values for one of the potentials. For its numerical solution the Gauss Newton method is applied. To speed up the solution process, a reduced-order approach based on proper orthogonal decomposition (POD) is utilized, where the accuracy is controlled by error estimators. Parameters, which can not be identified from the measurements, are identified by the subset selection method with QR pivoting. Numerical examples show the efficiency of the proposed approach.

The authors gratefully acknowledges support by the German Science Funds *Numerical and Analytical Methods for Elliptic-Parabolic Systems Appearing in the Modeling of Lithium-Ion Batteries* (Excellence Initiative) and DFG grant VO 1658/2-1 *A-posteriori-POD Error Estimators for Nonlinear Optimal Control Problems governed by Partial Differential Equations.*

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

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

1 Introduction

We consider an elliptic-parabolic partial differential equation (PDE) system consisting of two elliptic and one parabolic equation. Coupled multi component systems of this type can be viewed as generalizations of mathematical models for lithium ion batteries; see, e.g., [11,32,33]. The elliptic equations in the nonlinear system of PDEs model the potentials in the liquid and the solid phase and the parabolic equation for the concentration of lithium ions. The three equations are coupled by a strong nonlinear term involving the hyperbolic sine, the square root and the logarithmic function and a nonlinear diffusion coefficient.

The discretization of the nonlinear system of PDEs using, e.g., finite element techniques, lead to very large systems that are expensive to solve. The goal is to develop a reduced order model for the nonlinear system of PDEs that is cheap to evaluate. This is motivated by applications like parameter estimations, optimal control and design, where repeated evaluations of the nonlinear systems are required. Therefore, the spatial approximation is realized by the Galerkin scheme using proper orthogonal decomposition (POD); see, e.g., [18,21,29]. POD is used to generate a basis of a subspace that expresses the characteristics of the expected solution. This is in contrast to more general approximation methods, such as the finite element method, that do not correlate to the dynamics of the underlying system. We refer the reader to [5], where the authors apply a POD reduced-order modeling for the lithium ion battery model presented in [11].

After obtaining an efficient reduced order model we want to utilize it in a parameter estimation problem. The nonlinear systems arising from modeling of lithium ion batteries contain a variety of parameters. Those parameters have to be identified in order to calibrate the model. Not all parameters are equally sensitive, hence a strategy has to be applied in order to identify the parameters best suited for the identification process. For this the sensitivities are computed and with the help of a subset selection method the sensitive parameters are extracted [3,4,10].

When using a reduced order model in the optimization process an error is introduced. Therefore, an a posteriori error estimator has to be developed in order to quantify the quality of the obtained solution. We here use recent results from [8,19]. Further, it is important to understand that the obtained reduced oreder model by the POD is only a local approximation of the nonlinear system. Hence, it is necessary to garantee that the approximation is good throughout the optimization process. For this we make use of a simplified error indicator known from the reduced basis strategies [14]. Using this error indicator an adaptive reduced order model strategy is proposed to solve the optimization problem in an efficient way.

The paper is organized in the following manner: In Section 2 the nonlinear elliptic-parabolic system is formulated and the parameter identification problem is introduced. Further, we briefly describe the subset selection method. Section 3 is devoted to the discretization of the nonlinear system. Both the

finite element discretization and the POD Galerkin approximation method are introduced. Moreover, we review the a posteriori error estimates for the optimization using the reduced order model. The numerical strategy used to solve the nonlinear parameter estimation is outlined in Section 4 and numerical examples are presented to demonstrate the efficiency of the proposed approach. Finally, a conclusion is drawn in the last section.

2 Problem description

In this section the model under investigation is introduced. Further, the parameter estimation problem formulated as a nonlinear least squares problem is stated. Lastly, the subset selection method is outlined briefly.

2.1 The Model

In this section we formulate the nonlinear elliptic-parabolic system. Suppose that $\Omega = (a, b) \subset \mathbb{R}$, $a < b$, is the spatial domain with boundary $\Gamma = \{a, b\}$. We set $H = L^2(\Omega)$, $V = H^1(\Omega)$ and

$$V_a = \{\varphi \in H^1(\Omega) \,|\, \varphi(a) = 0\}.$$

For the terminal time $T > 0$ let $Q = (0, T) \times \Omega$ and $\Sigma = (0, T) \times \Gamma$. The space $L^2(0, T; V)$ stands for the space of (equivalence classes) of measurable abstract functions $\varphi : [0, T] \to V$, which are square integrable, i.e.,

$$\int_0^T \|\varphi(t)\|_V^2 \, \mathrm{d}t < \infty.$$

When t is fixed, the expression $\varphi(t)$ stands for the function $\varphi(t, \cdot)$ considered as a function in Ω only. Recall that

$$W(0, T; V) = \{\varphi \in L^2(0, T; V) \,|\, \varphi_t \in L^2(0, T; V)\}$$

is a Hilbert space supplied with its corresponding inner product; see, e.g., [1, 2, 6, 9].

After introducing the required functional settings let us state the nonlinear elliptic-parabolic system under investigation. First we introduce the admissible set $\mathcal{M}_{ad}$ for the parameter μ by

$$\mathcal{M}_{ad} = \left\{(\mu_1, \mu_2, \mu_3, \mu_4) \in \mathbb{R}^4 \,\middle|\, 1 < \mu_1 \leq 3/2,\ \mu_2 < 0, \mu_3 < 0 \text{ and } \underline{\mu} \leq \mu_4 \leq \overline{\mu}\right\}$$

with $\underline{\mu}, \overline{\mu} \in \mathbb{R}$ and $\underline{\mu} \leq \overline{\mu}$. For a given parameter $\mu = (\mu_1, \mu_2, \mu_3, \mu_4) \in \mathcal{M}_{ad}$ the triple $(y, p, q) : \overline{Q} \to \mathbb{R}$ satisfies the nonlinear system

$$y_t(t, \mathbf{x}) - \big(c_1(\mathbf{x}) y_{\mathbf{x}}(t, \mathbf{x})\big)_{\mathbf{x}} + \mathcal{N}(\mathbf{x}, y(t, \mathbf{x}), p(t, \mathbf{x}), q(t, \mathbf{x}); \mu) = 0, \tag{2.1a}$$

$$-\big(c_2(y(t, \mathbf{x}); \mu) p_{\mathbf{x}}(t, \mathbf{x})\big)_{\mathbf{x}} + \mathcal{N}(\mathbf{x}, y(t, \mathbf{x}), p(t, \mathbf{x}), q(t, \mathbf{x}); \mu) = 0, \tag{2.1b}$$

$$-\big(c_3(\mathbf{x}) q_{\mathbf{x}}(t, \mathbf{x})\big)_{\mathbf{x}} - \mathcal{N}(\mathbf{x}, y(t, \mathbf{x}), p(t, \mathbf{x}), q(t, \mathbf{x}); \mu) = 0 \tag{2.1c}$$

for almost all (f.a.a.) $(t, \mathbf{x}) \in Q$ together with the homogeneous Neumann boundary conditions for y and p

$$y_{\mathbf{x}}(t, a) = y_{\mathbf{x}}(t, b) = p_{\mathbf{x}}(t, a) = p_{\mathbf{x}}(t, b) \quad \text{f.a.a. } t \in (0, T), \tag{2.1d}$$

Dirichlet-Neumann condition for the variable q

$$q(t, a) = 0 \quad \text{and} \quad q_{\mathbf{x}}(t, b) = \mathcal{I}(t) \quad \text{f.a.a. } t \in (0, T) \tag{2.1e}$$

and the initial condition

$$y(0, \mathbf{x}) = y_\circ(\mathbf{x}) \quad \text{f.a.a. } \mathbf{x} \in \Omega \tag{2.1f}$$

with $y_\circ : \Omega \to \mathbb{R}$ positive and essentially bounded. The diffusion coefficients c_1 and c_3 are supposed to be piecewise constant and positive. For the nonlinear diffusion coefficient we choose a cubic polynomial in y and μ_4 of the form

$$c_2(y; \mu) = (1 + \mu_4 y)^3 - 1. \tag{2.2}$$

For the bounds in the admissible set we choose $\underline{\mu} = 0$ and $\overline{\mu} = 3$. Let us introduce the variable $z = (y, p, q) \in \mathcal{Z}_{ad}$ and the corresponding admissible set

$$\mathcal{Z}_{ad} = \big\{(y, p, q) \in \mathbb{R}^3 \,\big|\, y \geq y_{\min}\big\}.$$

The nonlinear coupling term $\mathcal{N} : \Omega \times \mathcal{Z}_{ad} \times \mathcal{M}_{ad} \to \mathbb{R}$ introduced in the model (2.1) is given by

$$\mathcal{N}(\mathbf{x}, z; \mu) = \chi(\mathbf{x}; \mu)\sqrt{y} \sinh\big(\mu_1(q - p) - \ln y\big), \tag{2.3}$$

where χ is an indicator function of the form

$$\chi(\mathbf{x}; \mu) = \begin{cases} \mu_2 \text{ for } \mathbf{x} \in [a, s_1] \\ 0 \quad \text{for } \mathbf{x} \in (s_1, s_2) \\ \mu_3 \text{ for } \mathbf{x} \in [s_2, b] \end{cases} \tag{2.4}$$

with $a < s_1 \leq s_2 < b$. With these choices (2.1) can be seen as a generalization of a mathematical model for lithium ion batteries; see, e.g., [11,32,33].

Remark 2.1 1) The positivity of the y component is needed to evaluate the terms $\sqrt{y}$ and $\ln y$ in the nonlinearity.

2) Notice that $\mathcal{N}(\mathbf{x}, \cdot; \mu) : \mathcal{Z}_{ad} \to \mathbb{R}$ is continuously differentiable for any parameter $\mu \in \mathcal{M}_{ad}$. Furthermore, we get

$$\frac{\partial \mathcal{N}}{\partial p}(\mathbf{x}, z; \mu) = -\mu_1 \chi(\mathbf{x}, \mu)\sqrt{y} \cosh(\mu_1(q - p) - \ln y) = -\frac{\partial \mathcal{N}}{\partial q}(\mathbf{x}, z; \mu) > 0.$$

$\Diamond$

Let us next look at the weak formulation of (2.1). For this we multiply (2.1) by testfunctions $\varphi \in V$ and $\varphi \in V_a$, respectively. Then integrating by parts the weak formulation is given by

$$\int_\Omega y_t(t)\varphi + c_1 y_{\mathbf{x}}(t)\varphi' + \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\varphi \,\mathrm{d}\mathbf{x} = 0 \qquad \forall \varphi \in V, \quad (2.5a)$$

$$\int_\Omega c_2(y(t);\mu)p_{\mathbf{x}}(t)\varphi' + \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\varphi \,\mathrm{d}\mathbf{x} = 0 \qquad \forall \varphi \in V, \quad (2.5b)$$

$$\int_\Omega c_3 q_{\mathbf{x}}(t)\varphi' - \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\varphi \,\mathrm{d}\mathbf{x} = \int_\Gamma \mathcal{I}(t)\varphi \,\mathrm{d}S \quad \forall \varphi \in V_a. \quad (2.5c)$$

The triple (y, p, q) solving (2.5) is called a *weak solution* to (2.1) if $z = (y, p, q) \in Z$, $y(0) = y_\circ$ in H with

$$Z = \big(W(0,T;V) \times L^2(0,T;V) \times L^2(0,T;V_a)\big) \cap L^\infty(Q)^3.$$

In the following theorem we state the analytical results about existence and uniqueness of weak solutions to the nonlinear system (2.1).

Theorem 2.1 *Let $\mu \in \mathcal{M}_{ad}$ and assume there exist $k_1 \le 3/2 + \mu_1$ and $k_2 > 0$ such that*

$$y^{k_1} \le c_2(y;\mu) \le k_2 y^{3/2+\mu_1}, \quad y \le 1.$$

Then there exists a unique global weak solution to the system given by (2.1) *for all $T > 0$. Furthermore, there exists a constant $C > 0$ such that*

$$\|p(t)\|_{L^\infty(\Omega)} \le Ce^{Ct}, \quad \|q(t)\|_{L^\infty(\Omega)} \le Ce^{Ct} \quad \textit{and} \quad y_{\min} = \frac{1}{C} \le y(t,\mathbf{x}) \le C$$

f.a.a. $(t, \mathbf{x}) \in Q$ hold.

Proof The proof to this theorem was presented in [33] for the case of all homegenous Neumann boundary condition for the variable q, i.e., $q_x(t,a) = q_x(t,b) = 0$ together with

$$\int_\Omega q(t,\mathbf{x}) \,\mathrm{d}\mathbf{x} = 0.$$

The introduced constraints on the parameters μ_1, μ_2 and μ_3 in the admissible set $\mathcal{M}_{ad}$ are essential for the proof. By using the results from [28] the proof can be extended to the case of the boundary conditions presented in (2.1e).□

Note that the choice for the nonlinear diffusion coefficient (2.2) fulfills the conditions required in Theorem 2.1 for the existence and uniqueness of the solution. Using this result we get that the nonlinear solution operator $\mathcal{S} : \mathcal{M}_{ad} \to Z$ is well-defined, where $z = \mathcal{S}(\mu)$ is the weak solution to (2.1) for the parameter value $\mu \in \mathcal{M}_{ad}$.

2.2 Parameter estimation

In this work we consider a parameter estimation problem for (2.5). The goal is to identify a parameter μ in the set $\mathcal{M}_{ad}$ of admissible parameter values from given boundary measurements for the solution component q. Therefore, we consider

$$\min_{\mu\in\mathcal{M}_{ad}} J(\mu) = \frac{1}{2}\int_0^T \left|\eta(t,\mu) - q^d(t)\right|^2 \mathrm{d}t, \tag{\textbf{P}}$$

where $(y,p,q) = \mathcal{S}(\mu)$ and $\eta(t,\mu) = q(t,b)$ hold f.a.a. $t \in (0,T)$. To solve (**P**) we apply the so-called sensitivity approach to evaluate the gradient of J; see, e.g., [17, Section 1.6.1]. The derivatives of the cost J are given by [22]

$$\frac{\partial J(\mu)}{\partial \mu_i} = \int_0^T \big(\eta(t,\mu) - q^d(t)\big)\frac{\partial \eta}{\partial \mu_i}(t,\mu)\,\mathrm{d}t, \quad i = 1,\dots,4,$$

and

$$\frac{\partial^2 J(\mu)}{\partial \mu_i \partial \mu_j} = \int_0^T \frac{\partial \eta}{\partial \mu_i}(t,\mu)\frac{\partial \eta}{\partial \mu_j}(t,\mu)\,\mathrm{d}t + \int_0^T \big(\eta(t,\mu) - q^d(t)\big)\frac{\partial^2 \eta}{\partial \mu_i \partial \mu_j}(t,\mu)\,\mathrm{d}t.$$

Let us introduce the first order sensitivities for a given parameter $\mu \in \mathcal{M}_{ad}$: We set $s^{y,i} = \frac{\partial y}{\partial \mu_i}$ for $i = 1,\dots,4$. Analogously, $s^{p,i}$ and $s^{q,i}$ are defined. Then, $\frac{\partial \eta}{\partial \mu_i}(t,\mu)$ is equal to $s^{q,i}$ evaluated at (t,b). The sensitivity equations for $s^{y,i}$, $s^{p,i}$, and $s^{q,i}$ are given as

$$\begin{aligned}
&s_t^{y,i}(t) - \big(c_1 s_{\mathbf{x}}^{y,i}(t)\big)_{\mathbf{x}} \\
&\qquad + \mathcal{N}_y(\cdot\,, y(t), p(t), q(t); \mu)\big(s^{y,i}(t) + s^{p,i}(t) + s^{q,i}(t)\big) = \mathcal{G}_y^i(t), \\
&\quad -\big(c_2(y(t);\mu)s_{\mathbf{x}}^{p,i}(t)\big)_{\mathbf{x}} + c_y(y(t);\mu)p_{\mathbf{x}}(t)s_{\mathbf{x}}^{y,i}(t) \\
&\qquad + \mathcal{N}_p(\cdot\,, y(t), p(t), q(t); \mu)\big(s^{y,i}(t) + s^{p,i}(t) + s^{q,i}(t)\big) = \mathcal{G}_p^i(t), \\
&-\big(c_3 s_{\mathbf{x}}^{q,i}(t)\big)_{\mathbf{x}} - \mathcal{N}_q(\cdot\,, y(t), p(t), q(t); \mu)\big((s^{y,i}(t) + s^{p,i}(t) + s^{q,i}(t)\big) = \mathcal{G}_q^i(t)
\end{aligned}$$

with

$$\begin{aligned}
&\mathcal{G}_y^i(t) = -\mathcal{N}_{\mu_i}(\cdot\,, y(t), p(t), q(t); \mu) \text{ for } i = 1,2,3, \quad \mathcal{G}_y^4(t) = 0 \\
&\mathcal{G}_p^i(t,\mathbf{x}) = \mathcal{G}_y^i(t) \text{ for } i = 1,2,3, \quad \mathcal{G}_p^4(t) = -\big((c_2(y(t);\mu))_{\mu_4} y_{\mathbf{x}}(t)\big)_{\mathbf{x}} \\
&\mathcal{G}_q^i(t) = -\mathcal{G}_y^i(t) \text{ for } i = 1,\dots,4
\end{aligned}$$

together with the boundary conditions

$$s_{\mathbf{x}}^{y,i}(t,a) = s_{\mathbf{x}}^{y,i}(t,b) = s_{\mathbf{x}}^{p,i}(t,a) = s_{\mathbf{x}}^{p,i}(t,b) = 0 \quad \text{f.a.a. } t \in (0,T)$$

and

$$s^{q,i}(t,a) = 0 \quad \text{and} \quad s_x^{q,i}(t,b) = 0 \quad \text{f.a.a. } t \in (0,T)$$

and the initial condition

$$s^{y,i}(0,\mathbf{x}) = 0 \quad \text{f.a.a. } \mathbf{x} \in \Omega.$$

for $i = 1, \dots, 4$. To obtain the hessian $\nabla^2 J(\mu) \in \mathbb{R}^{4\times 4}$ we derive the adjoint equations by utilizing the Lagrangian for (**P**):

$$\begin{aligned}\mathcal{L}(z,\mu,\lambda) = \frac{1}{2}\int_0^T |q(t,b) - q^d(t)|^2 \,\mathrm{d}t \\ + \int_0^T \int_\Omega y_t(t)\lambda^y(t) + c_1 y_{\mathbf{x}}(t)\lambda^y_{\mathbf{x}}(t) + \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\lambda^y(t)\,\mathrm{d}\mathbf{x}\mathrm{d}t \\ + \int_0^T \int_\Omega c_2(y(t);\mu)p_{\mathbf{x}}(t)\lambda^p_{\mathbf{x}}(t) + \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\lambda^p(t)\,\mathrm{d}\mathbf{x}\mathrm{d}t \\ + \int_0^T \int_\Omega c_3 q_{\mathbf{x}}(t)\lambda^q_{\mathbf{x}}(t) - \mathcal{N}(\cdot\,, y(t), p(t), q(t); \mu)\lambda^q(t)\,\mathrm{d}\mathbf{x}\mathrm{d}t \\ - \int_0^T \int_\Gamma \frac{\partial q}{\partial n}(t)\lambda^q(t)\,\mathrm{d}S\mathrm{d}t + \int_\Omega (y(0) - y_\circ)\lambda^y(0)\,\mathrm{d}\mathbf{x}\end{aligned}$$

for $z = (y, p, q) \in Z$, $\mu \in \mathcal{M}_{ad}$ and $\lambda = (\lambda^y, \lambda^p, \lambda^q) \in L^2(0,T;V)\times L^2(0,T;V)\times L^2(0,T;V_a)$. We derive the following adjoint system [22,30]

$$\begin{aligned}-\lambda^y_t(t) - \big(c_1\lambda^y_{\mathbf{x}}(t)\big)_{\mathbf{x}} + c_y(y(t);\mu)p_{\mathbf{x}}(t)\lambda^p_{\mathbf{x}}(t) \\ +\mathcal{N}_y(\cdot\,, y(t), p(t), q(t); \mu)\big(\lambda^y(t) + \lambda^p(t) - \lambda^q(t)\big) = 0, \\ -\big(c_2(y(t);\mu)\lambda^p_{\mathbf{x}}(t)\big)_{\mathbf{x}} + \mathcal{N}_p(\cdot\,, y(t), p(t), q(t); \mu)\big(\lambda^y(t) + \lambda^p(t) - \lambda^q(t)\big) = 0, \\ -\big(c_3(\mathbf{x})\lambda^q_{\mathbf{x}}(t,x)\big)_{\mathbf{x}} + \mathcal{N}_q(\cdot\,, y(t), p(t), q(t); \mu)\big(\lambda^y(t) + \lambda^p(t) - \lambda^q(t)\big) = 0\end{aligned}$$

f.a.a. $(t, \mathbf{x}) \in Q$ together with the boundary conditions

$$\lambda^y_{\mathbf{x}}(t,a) = \lambda^y_{\mathbf{x}}(t,b) = \lambda^p_{\mathbf{x}}(t,a) = \lambda^p_{\mathbf{x}}(t,b) = 0 \quad \text{f.a.a. } t \in (0,T)$$

and

$$\lambda^q(t,a) = 0 \quad \text{and} \quad \lambda^q_{\mathbf{x}}(t,b) = q^d(t) - q(t,b) \quad \text{f.a.a. } t \in (0,T)$$

and the initial condition

$$\lambda(T,\mathbf{x}) = 0 \quad \text{f.a.a. } \mathbf{x} \in \Omega.$$

Combining the sensitivities and the adjoint the hessian is obtained; see, e.g., [17, Section 1.6.5].

2.3 Subset selection

To address the question of identifiability of our four parameters we apply the QR factorization with column pivoting proposed by Golub in [12]. This method turns out to be computationally very efficient. The algorithm consists in using a column pivoting strategy to compute a permutation matrix Π so that the hessian satisfies $H\Pi = QR$. While the QR algorithm with column pivoting works well for our numerical examples, there are examples, where the method fails [15].

When looking at (**P**) the question arises whether all parameters are identifiable. Following the notion of sensitivity identifiability the matrix

$$\big(\big(H(\mu)\big)\big)_{ij} = \int_0^T \frac{\partial \eta}{\partial \mu_i}(t,\mu)\frac{\partial \eta}{\partial \mu_j}(t,\mu)\,\mathrm{d}t, \quad 1 \le i,j \le 4, \tag{2.6}$$

has to have rank four, or $\det(H) \neq 0$ [27]. This matrix is also known as the Fisher information matrix. When utilizing a Gauss-Newton method the matrix $H(\mu)$ is used as the approximation of the hessian of J. Note that H is a symmetric and positive semidefinite matrix. To eliminate non-identifiable parameters we apply the subset selection method; see, e.g., [3,4,10]:

1) Compute the nonnegative eigenvalues $\{\lambda_i\}_{i=1}^4$ of $H = V\Lambda V^\top$ and truncate at $\varepsilon > 0$, i.e.,

$$\lambda_1 \ge \ldots \ge \lambda_k \ge \varepsilon \quad \text{with } k \in \{1,2,3,4\};$$

2) Use the QR method with column-pivoting to reorder μ, i.e.,

$$H\Pi = QR \quad \text{and} \quad \bar{\mu} = \Pi^\top \mu;$$

3) Select the first k parameters of $\bar{\mu}$ as new parameter and set the others to academic/reasonable values.

By applying the subset selection method it is guranteed that the matrix H is invertible. In the Gauss-Newton context this means that the search direction can be computed.

3 Discretization

In this section we introduce our finite element approximation for (2.1) and the associated reduced-order approach based on the proper orthogonal decomposition (POD) method.

3.1 Finite element method

Let us next introduce the Galerkin scheme for (2.1). To discretize the nonlinear system (2.1) in space we use a Galerkin scheme. For this we introduce the finite dimensional subspaces

$$V^{g,N_x} = \mathrm{span}\{\varphi_,^g \ldots, \varphi_{dof}^g\} \subset V$$

with $g = \{y,p\}$, where the φ_i's denote the dof basis functions that are used to approximate the space V. Analogously we proceed with $V_a^{q,N_x} \subset V_a$. At this point we do not make a restriction of what type of basis functions are being used. With the superindex g we emphasize the possibility of using different types of basis functions for each variable. In this section we will use the same

basis functions for each variable and hence omit the superindex g from now on. We proceed by introducing the standard Galerkin ansatz of the form

$$y^N(t,\mathbf{x}) = \sum_{i=1}^{dof} \mathrm{y}_i^N(t)\varphi_i(\mathbf{x}), \quad p^N(t,\mathbf{x}) = \sum_{i=1}^{dof} \mathrm{p}_i^N(t)\varphi_i(\mathbf{x}), \quad \varphi \in V^{N_x}, \tag{3.1a}$$

and

$$q^N(t,\mathbf{x}) = \sum_{i=1}^{dof} \mathrm{q}_i^N(t)\varphi_i(\mathbf{x}) \quad \varphi \in V_a^N. \tag{3.1b}$$

Here y_i^N, p_i^N and q_i^N denote the time-dependent Galerkin coefficients corresponding to the variables y, p and q. Inserting the Galerkin ansatz into the weak form (2.5) we get

$$\int_\Omega y_t^N(t)\varphi + c_1 y_{\mathbf{x}}^N(t)\varphi' + \mathcal{N}(\cdot\,, y^N(t), p^N(t), q^N(t);\mu)\varphi \,\mathrm{d}\mathbf{x} = 0 \ \forall \varphi \in V^{N_x}, \tag{3.2a}$$

$$\int_\Omega c_2(y^N(t);\mu) p_{\mathbf{x}}^N(t)\varphi' + \mathcal{N}(\cdot\,, y^N(t), p^N(t), q^N(t);\mu)\varphi \,\mathrm{d}\mathbf{x} = 0 \ \forall \varphi \in V^{N_x}, \tag{3.2b}$$

$$\int_\Omega c_3 q_{\mathbf{x}}^N(t)\varphi' - \mathcal{N}(\cdot\,, y^N(t), p^N(t), q^N(t);\mu)\varphi \,\mathrm{d}\mathbf{x} = \int_\Gamma \mathcal{I}(t)\varphi \,\mathrm{d}S \ \forall \varphi \in V_a^{N_x}, \tag{3.2c}$$

and $1 \le i \le dof$. Note that here also the test functions φ_i are chosen from the same space as the ansatz functions. There are also techniques to choose them from different spaces which then results in a Petrov-Galerkin ansatz. Utilizing the structure of the Galerkin ansatz for y^N, p^N and q^N the obtained system can be written in matrix-vector form. For this we introduce

$$\left(\left(M_f^N\right)\right)_{ij} = \int_\Omega f(\mathbf{x})\varphi_j\varphi_i \,\mathrm{d}\mathbf{x}, \qquad \text{(mass matrix)} \tag{3.3a}$$

$$\left(\left(S_f^N\right)\right)_{ij} = \int_\Omega f(\mathbf{x})\varphi_j'\varphi_i' \,\mathrm{d}\mathbf{x}, \qquad \text{(stiffness matrix)} \tag{3.3b}$$

$$\left(\mathcal{N}^N(\mathrm{y}^N(t), \mathrm{p}^N(t), \mathrm{q}^N(t);\mu)\right)_i = \int_\Omega \mathcal{N}(\mathbf{x}, y^N(t), p^N(t), q^N(t);\mu)\varphi_i \,\mathrm{d}\mathbf{x}, \tag{3.3c}$$

$$\left(\mathcal{I}^N(t)\right)_i = \int_\Gamma \mathcal{I}(t)\varphi_i \,\mathrm{d}S, \tag{3.3d}$$

$$\left(\mathcal{Y}_0^N\right)_i = \int_\Omega y_\circ(\mathbf{x})\varphi_i \,\mathrm{d}S, \tag{3.3e}$$

for $1 \le i,j \le dof$, $t \in [0,T]$ and

$$\mathrm{y}^N(t) = \left(\mathrm{y}_i^N(t)\right)_{1\le i\le dof}, \quad \mathrm{p}^N(t) = \left(\mathrm{p}_i^N(t)\right)_{1\le i\le dof}, \quad \mathrm{q}^N(t) = \left(\mathrm{q}_i^N(t)\right)_{1\le i\le dof}.$$

Hence the weak form (2.5) can be written in semi-discrete form using matrix vector notation together with the initial condition as

$$M_1^N \mathrm{y}_t^N(t) + S_{c_1}^N \mathrm{y}^N(t) + \mathcal{N}^N(\mathrm{y}^N(t), \mathrm{p}^N(t), \mathrm{q}^N(t); \mu) = 0, \, t \in (0,T), \quad (3.4a)$$

$$M_1^N \mathrm{y}^N(0) = \mathcal{Y}_0^N, \quad (3.4b)$$

$$S^N_{c_2(\mathrm{y}^N(t);\mu)} \mathrm{p}^N(t) + \mathcal{N}^N(\mathrm{y}^N(t), \mathrm{p}^N(t), \mathrm{q}^N(t); \mu) = 0, \qquad t \in (0,T), \quad (3.4c)$$

$$S_{c_3}^N \mathrm{q}^N(t) - \mathcal{N}^N(\mathrm{y}^N(t), \mathrm{p}^N(t), \mathrm{q}^N(t); \mu) = \mathcal{I}^N(t), \qquad t \in (0,T). \quad (3.4d)$$

Note that (3.4) is a nonlinear semi-discrete system since only the space domain has been discretized. Furthermore, the dimension of the matrices M^N and S^N is $dof \times dof$. Consequently the semi-discrete nonlinear system (3.4) is of dimension $3 \times dof$. Throughout the paper we assume that for any $\mu \in \mathcal{M}_{ad}$ the system (3.4) admits a unique solution satisfying

$$y^N(t) = \sum_{i=1}^{dof} \mathrm{y}_i^N(t) \varphi_i(\mathbf{x}) \geq y_{\min} \quad \text{f.a.a. } (t, \mathbf{x}) \in Q.$$

Further, there exists a constant $C > 0$ independent of h such that the following estimates hold:

$$\|p^N(t)\|_{L^\infty(\Omega)} \leq Ce^{Ct}, \ \|q^N(t)\|_{L^\infty(\Omega)} \leq Ce^{Ct} \text{ and } y_{\min} = \frac{1}{C} \leq y^N(t, \mathbf{x}) \leq C$$

f.a.a. $(t, \mathbf{x}) \in Q$.

To solve the obtained nonlinear system an implicit method is used for the time discretization. The obtained discrete nonlinear system is then solved by applying the Newton method with a damping strategy [7].

3.2 The POD method

We explain the POD method for the solution component $y^N(t)$, $t \in [0,T]$. The components p^N and q^N are treated analogously. In practical computations, we do not have the whole trajectory $y^N(t)$ at hand, but rather snapshots $\mathbb{R}^{dof} \ni \mathrm{y}^{N,k} \approx \mathrm{y}^N(t_j)$, $0 \leq j \leq N_t$, at discrete times

$$0 = t_0 < t_1 < \ldots < t_{N_t} = T.$$

This setting coincides with the results obtained by the finite element discretization introduced in Section 3.1. We set $\mathcal{V}^{y,N_x} = \operatorname{span}\{\mathrm{y}^{N,0}, \ldots, \mathrm{y}^{N,N_t}\} \subset \mathbb{R}^{dof}$ with $\mathrm{d}_y = \dim \mathcal{V}^{y,N_x} \leq \min(dof, N_t)$. Then, a POD basis $\{\Psi_i^y\}_{i=1}^{\ell_y}$ is given by the solution to

$$\begin{cases} \min\limits_{\Psi_1^y, \ldots, \Psi_{\ell_y}^y \in \mathbb{R}^{dof}} \sum\limits_{k=0}^{N_t} \alpha_k \left| \mathrm{y}^{N,k} - \sum\limits_{i=1}^{\ell_y} \langle \mathrm{y}^{N,k}, \Psi_i^y \rangle_{\mathrm{W}} \Psi_i^y \right|_{\mathrm{W}}^2 \mathrm{d}t \\ \quad \text{s.t. } \langle \Psi_i^y, \Psi_j^y \rangle_{\mathrm{W}} = \delta_{ij} \text{ for } 1 \leq i, j \leq \ell_y, \end{cases} \quad (3.5)$$

where the α_k are trapezoidal weights for the integration over the discretized time interval, i.e.,

$$\alpha_0 = \frac{\mathrm{t}_1 - \mathrm{t}_0}{2},\ \alpha_k = \frac{\mathrm{t}_{k+1} - \mathrm{t}_{k-1}}{2} \text{ for } k = 1, \dots, N_t - 1, \alpha_{N_t} = \frac{\mathrm{t}_{N_t} - \mathrm{t}_{N_t - 1}}{2}.$$

Moreover, $\langle \cdot\,, \cdot \rangle_{\mathrm{W}}$ stands for the weighted inner product in $\mathbb{R}^{dof}$ with a positive definite, symmetric matrix $\mathrm{W} \in \mathbb{R}^{dof \times dof}$. Let us introduce the matrices

$$\mathrm{D} = \operatorname{diag}\left(\alpha_1, \dots, \alpha_{N_t}\right) \in \mathbb{R}^{N_t \times N_t} \quad \text{and} \quad \mathrm{Y}^N = \left(\mathrm{y}^{N,1} \,\middle|\, \dots \,\middle|\, \mathrm{y}^{N,N_t}\right) \in \mathbb{R}^{dof \times N_t}.$$

Then we can write the operator $\mathcal{R}^y$ arising from the optimization problem (3.5) as

$$\mathcal{R}^y \Psi^y = \sum_{k=0}^{N_t} \alpha_k \langle \mathrm{y}^{N,k}, \Psi^y \rangle_{\mathrm{W}} \mathrm{y}^{N,k} = Y^N D \big(Y^N\big)^\top \mathrm{W} \Psi^y \quad \text{for } \Psi^y \in \mathbb{R}^{dof}$$

which leads to the unsymmetric eigenvalue problem

$$Y^N D \big(Y^N\big)^\top \mathrm{W} \Psi_i^y = \lambda^y \Psi_i^y \quad \text{for } i = 1, \dots, \mathrm{d}_y.$$

This eigenvalue problem has to be solved in order to obtain the solution to the optimal control problem (3.5). In order to obtain a symmetric eigenvalue problem we introduce the matrix

$$\bar{\mathrm{Y}^N} = \mathrm{W}^{1/2} \mathrm{Y}^N \mathrm{D}^{1/2} \in \mathbb{R}^{dof \times N_t}.$$

Recall that W is supposed to be symmetric and positive definite. Thus, W possesses an eigenvalue decomposition of the form $\mathrm{W} = \mathrm{P} \Lambda \mathrm{P}^\top$, where Λ is a diagonal matrix containing the positive eigenvalues of W and $\mathrm{P} \in \mathbb{R}^{N_x \times N_x}$ is an orthogonal matrix. Hence we can define the square root of W by setting $\mathrm{W}^{1/2} = \mathrm{P} \Lambda^{1/2} \mathrm{P}^\top$, where $\Lambda^{1/2}$ is a diagonal matrix containing the square roots of the eigenvalues. Note that $\mathrm{D}^{1/2} = \operatorname{diag}\left(\alpha_1^{1/2}, \dots, \alpha_{N_t}^{1/2}\right)$ holds, but usually the computation of the square root $\mathrm{W}^{1/2}$ is more involved. Hence the solution $\{\Psi_i^y\}_{i=1}^{\ell_y}$ to (3.5) can be determined by three different ways (see, e.g., [13]):

1) Solve the symmetric $dof \times dof$ eigenvalue problem

$$\bar{\mathrm{Y}^N} \big(\bar{\mathrm{Y}^N}\big)^\top \mathrm{u}_i = \lambda_i^y \mathrm{u}_i, \quad 1 \le i \le \ell_y, \tag{3.6}$$

   and set $\Psi_i^y = \mathrm{W}^{-1/2} \mathrm{u}_i$ for $1 \le i \le \ell_y$.
2) Determine the solution to the symmetric $N_t \times N_t$ eigenvalue problem

$$\big(\bar{\mathrm{Y}^N}\big)^\top \bar{\mathrm{Y}^N} v_i = \lambda_i^y v_i, \quad 1 \le i \le \ell_y, \tag{3.7}$$

   and set $\Psi_i^y = \mathrm{Y} \mathrm{D}^{1/2} v_i / \sqrt{\lambda_i^y}$ for $1 \le i \le \ell_y$.

3) Compute the singular value decomposition for $\bar{\mathrm{Y}}$, i.e.,

$$\mathrm{U}^\top \bar{\mathrm{Y}} \mathrm{V} = \left(\begin{array}{ccc|c} \sigma_1^y & & & \\ & \ddots & & 0 \\ & & \sigma_{\mathrm{d}_y}^y & \\ \hline & 0 & & 0 \end{array}\right)$$

with orthogonal matrices $\mathrm{U} = (\mathrm{u}_1^y \,|\, \ldots \,|\, \mathrm{u}_{dof}^y) \in \mathbb{R}^{dof \times dof}$, $\mathrm{V} \in \mathbb{R}^{N_t \times N_t}$ and singular values $\sigma_1^y \geq \sigma_2^y \geq \ldots \geq \sigma_{\mathrm{d}_y}^y > 0$. Then, set $\Psi_i^y = \mathrm{W}^{-1/2}\mathrm{u}_i$ for $1 \leq i \leq \ell_y$ as in 1) above.

Notice that $\Psi_i^{y,N} = \mathrm{W}^{-1/2}\mathrm{u}_i$ is realized by solving the linear system $\mathrm{W}^{1/2}\Psi_i^y = \mathrm{u}_i$. Moreover, $(\sigma_i^y)^2 = \lambda_i^y$ holds true for $1 \leq i \leq \mathrm{d}_y$. In the case $N_t \ll dof$ the approach 2) turns out to be the fastest strategy.

Suppose that we have determined three POD bases $\{\Psi_i^\wp\}_{i=1}^{\ell_\wp}$ for $\wp \in \{y, p, q\}$. Let us define the matrices

$$\Psi^\wp = \left(\Psi_1^\wp \,\middle|\, \ldots \,\middle|\, \Psi_{\ell_\wp}^\wp\right) \in \mathbb{R}^{dof \times \ell_\wp} \quad \text{for } \wp \in \{y, p, q\}.$$

The vectors $\Psi_i^\wp$, $\wp \in \{y, p, q\}$ can be interpreted as finite element coefficients of $\mathbf{x}$-dependent POD basis functions:

$$\psi_i^\wp(\mathbf{x}) = \sum_{j=1}^{dof} \left(\Psi_i^\wp\right)_j \varphi_j(\mathbf{x}) \quad \text{for } i = 1, \ldots, \ell_\wp \text{ with } \wp \in \{y, p, q\}.$$

which implies that

$$y^\ell(t) = \sum_{i=1}^{\ell_y} \mathrm{y}_i^\ell(t)\psi_i^y = \sum_{i=1}^{\ell_y} \mathrm{y}_i^\ell(t) \sum_{j=1}^{dof} \left(\Psi_i^y\right)_j \varphi_j = \sum_{j=1}^{dof} \left(\Psi^y \mathrm{y}^\ell(t)\right)_j \varphi_j \tag{3.8}$$

for $i = 1, \ldots, \ell^y$. Thus, the vector $\Psi^y \mathrm{y}^\ell(t)$ contains the finite element nodal coefficients of the function y^ℓ at time $t \in [0, T]$. An analogous result follows for the p- and q-variable. Moreover, the matrices and vectors for the reduced order model are then given by

$$\begin{aligned} M_f^{\ell_\wp} &= \left(\Psi^\wp\right)^\top M_f^N \Psi^\wp, \quad \mathrm{S}_f^{\ell_\wp} = \left(\Psi^\wp\right)^\top S_f^N \Psi^\wp, \\ \mathcal{I}^{\ell_q}(t) &= \left(\Psi^q\right)^\top \mathcal{I}^N(t), \quad \mathcal{Y}_0^{\ell_y} = \left(\Psi^y\right)^\top \mathcal{Y}_0^N, \\ \mathcal{N}^{\ell_\wp}(\mathrm{y}^\ell(t), \mathrm{p}^\ell(t), \mathrm{q}^\ell(t); \mu) &= \left(\Psi^\wp\right)^\top \mathcal{N}^N\left(\Psi^y \mathrm{y}^\ell(t), \Psi^p \mathrm{p}^\ell(t), \Psi^q \mathrm{q}^\ell(t); \mu\right), \end{aligned} \tag{3.9}$$

for $\wp \in \{y, p, q\}$, where the finite element matrices and vectors are given by (3.3). The reduced order model then reads as

$$M_1^{\ell_y} \mathrm{y}_t^{\ell,N}(t) + S_{c_1}^{\ell_y} \mathrm{y}^\ell(t) + \mathcal{N}^{\ell_y}(\mathrm{y}^\ell(t), \mathrm{p}^\ell(t), \mathrm{q}^\ell(t); \mu) = 0, \tag{3.10a}$$

$$M_1^{\ell_y} \mathrm{y}^{\ell,N}(0) = \mathcal{Y}_0^{\ell_y}, \tag{3.10b}$$

$$S_{c_2(\mathrm{y}^\ell(t);\mu)}^{\ell_p} \mathrm{p}^\ell(t) + \mathcal{N}^{\ell_p}(\mathrm{y}^\ell(t), \mathrm{p}^\ell(t), \mathrm{q}^\ell(t); \mu) = 0, \tag{3.10c}$$

$$S_{c_3}^{\ell_q} \mathrm{q}^\ell(t) - \mathcal{N}^{\ell_q}(\mathrm{y}^\ell(t), \mathrm{p}^\ell(t), \mathrm{q}^\ell(t); \mu) = \mathcal{I}^{\ell_q}(t), \tag{3.10d}$$

f.a.a. $t \in [0,T]$. This obtained reduced order model is based on the finite element method introduced in Section 3.1 but is of much smaller dimension. The model is generated by projection of the finite element matrices and vectors onto the subspace spanned by the POD basis. Throughout the paper we assume that for any $\mu \in \mathcal{M}_{ad}$ the system (3.10) admits a unique solution satisfying

$$y^\ell(t,\mathbf{x}) = \sum_{i=1}^{\ell_y} \mathrm{y}_i^\ell(t)\psi_i^y(\mathbf{x}) \geq y_{\min} \quad \text{f.a.a. } (t,\mathbf{x}) \in Q.$$

Further, there exists a constant $C > 0$ independent of ℓ_y, ℓ_p and ℓ_q such that the following estimates hold:

$$\|p^\ell(t)\|_{L^\infty(\Omega)} \leq Ce^{Ct}, \quad \|q^\ell(t)\|_{L^\infty(\Omega)} \leq Ce^{Ct} \quad \text{and} \quad y_{\min} = \frac{1}{C} \leq y^\ell(t,\mathbf{x}) \leq C$$

f.a.a. $(t,\mathbf{x}) \in Q$.

In (3.9) the evaluation of the nonlinear terms is hidden. Note that for the evaluation of the nonliear term $\mathcal{N}^{\ell_\wp}$, $\wp \in \{y,p,q\}$ it is required to evaluate the finite element counter part $\mathcal{N}^N$. This is computationally very expensive. By applying the *empirical interpolation method* (EIM) [24] this can be overcome. The method has been applyed succesfully to models of this type in [23] and reduced order models indepentend ot the finite element dimension are obtained.

To solve the obtained nonlinear system the same numerical methods are used as in the case of the finite element discretization.

3.3 A posteriori error estimator

By discretizing the nonlinear least squares problem (**P**) using the finite element method we obtain J^N and the corresponding optimal solution $\mu^{N,*}$. Analogously, when using the reduced order model obtained by the POD Galerkin scheme we get J^ℓ and $\mu^{\ell,*}$. In the optimization process the goal is to replace the finite element model by the reduced order model. The obtained solution is suboptimal, i.e. $\mu^{\ell,*}$ is not an optimal solution to the reference finite element problem J^N. The aim is to find bounds for the error $\|\mu^{N,*} - \mu^{\ell,*}\|_2$. Thereby the optimal solution $\mu^{\ell,*}$ obtained by the reduced order model can be certified. For this we reformulate the results from [8,19] for the settings introduced in this work.

The result – obtained in [19] by a second-order analysis – are in a more general setting including bound constraints for μ which are omitted in the numerical part of this work. Let us restate the results for our specific settings [19, Theorem 3.4].

Theorem 3.1 *Let $\mu^{N,*} \in \mathcal{M}_{ad}$ be an inactive, local minimizer of* (**P**) *utilizing the finite element discretization, i.e., $\nabla J^N(\mu^{N,*}) \equiv 0$ holds true. Further, there exists a $\delta > 0$ such that the second-order sufficient optimality condition*

$$\nabla^2 J^N(\mu^{N,*})(v,v) > \delta \|v\|2_2 \quad \forall v \in \mathbb{R}^4$$

is satisfied. If $\mu^{\ell,*} \in \mathcal{M}_{ad}$ is given such that $\|\mu^{N,*} - \mu^{\ell,*}\|_2$ is sufficiently small. Then the error estimator

$$\|\mu^{N,*} - \mu^{\ell,*}\|_2 \leq \frac{2}{\delta} \|\nabla J^N(\mu^{\ell,*})\|_2$$

holds.

This gives an upper bound for the error when utilizing the reduced order model with respect to the solution obtained when using the finite element model. The open question with this estimator is the value for δ. In [19] the constant δ was chosen as the smallest eigenvalue of the hessian $\nabla^2 J^N(\mu^{\ell,*})$ which turned out to work very well. The error estimator then reads as

$$\|\mu^{N,*} - \mu^{\ell,*}\|_2 \leq 2 \|\nabla^2 J^N(\mu^{\ell,*})^{-1}\|_2 \|\nabla J^N(\mu^{\ell,*})\|_2. \tag{3.11}$$

Note that in order to compute the error estimator the hessian is needed. This can become computationally expensive since it involves finite element solves. In the context of reduced basis a similar result, which is based on the Newton-Kantorovich theory, was obtained. For our setting the result reads as follows [8, Proposition 4]:

Proposition 3.1 *Let $\mu^{\ell,*} \in \mathcal{M}_{ad}$ be the optimal solution found by solving the least squares probelm* (**P**) *utilizing the reduced order model* (3.10) *satisfying the stopping criterion $\|\nabla J^\ell(\mu^{\ell,*})\|_2 \leq \varepsilon_{opt}$. Further, let us introduce*

$\bar{B}(\mu^{\ell,*}, \alpha)$ *a closed ball around $\mu^{\ell,*}$ with radius α,*
the hessian $\nabla^2 J^N(\mu^{\ell,*})$ is regular,
$\gamma := \|\nabla^2 J^N(\mu^{\ell,*})^{-1}\|_2$,
$\varepsilon := \Delta_{\nabla J(\mu^{\ell,*})} + \varepsilon_{opt}$ *with* $\Delta_{\nabla J(\mu)} = \nabla J^N(\mu) - \nabla J^\ell(\mu)$,
$L(\alpha) := \sup_{\mu \in \bar{B}(\mu^{\ell,*}, \alpha)} \|\nabla^2 J^N(\mu^{\ell,*}) - \nabla^2 J^N(\mu)\|_2$.

If the condition

$$2\gamma L(2\gamma\epsilon) \leq 1$$

holds, then there exists a unique solution $\mu^{N,*}$ to the optimization problem (**P**) *utilizing the finite element model* (3.4) *with $\mu^{N,*} \in \bar{B}(\mu_\ell^*, 2\gamma\varepsilon)$ and the rigorous error bound*

$$\|\mu_N^* - \mu^{\ell,*}\|_2 \leq \|\nabla^2 J^N(\mu^{\ell,*})^{-1}\|_2 (\Delta_{\nabla J(\mu^{\ell,*})} + \varepsilon_{opt}) \tag{3.12}$$

holds.

When comparing the results it can be seen that they are very similar although the derivations are very different. This also verifies the choice $\delta = \|\nabla^2 J^N(\mu^{\ell,*})^{-1}\|_2$ in (3.11). Note that not all conditions can be checked and hence some assumptions have to be made.

4 Numerical methods and results

In this section we describe the numerical procedure to solve the parameter estimation problem (**P**) introduced in Section 2.2. The usage of the subset selection method and the reduced order model will be outlined. The section is concluded with numerical results that illustrate the advantages of the proposed approach.

4.1 Numerical method

There are different numerical strategies to solve parameter estimation problems formulated in nonlinear least squares form. As a numerical method to solve the introduced parameter estimation problem (**P**) we apply the Gauss-Newton method [20,25,31]. In the Gauss-Newton method the Hessian matrix $\nabla^2 J(\mu)$ is approximated by the matrix $H(\mu) \in \mathbb{R}^{4\times 4}$ given by (2.6). Note that for the evaluation of the approximated Hessian only the first order sensitivities are needed which are used to evaluate the gradient $\nabla J(\mu)$. Further, recall that H corresponds to the Fisher information matrix which is utilized in the subset selection method. By applying the subset selection method it is ensured that the inverse of H exists which is essential in computing the search direction in the proposed optimization strategy. In every iteration k the Gauss-Newton search direction d^k is computed as the solution of the linear system

$$H(\mu^k)d^k = -\nabla J(\mu^k).$$

In order to speed up the optimization process we will replace the finite element model (3.4) by the reduced order model (3.10). This will allow us to solve the parameter estimation problem at less computational costs.

When looking at the construction of the reduced order model it can be seen that the POD basis is computed with respect to a given parameter μ. This implies that when varying the parameter μ we do not have a guarantee that the reduced order model is of good quality. Hence it can be necessary to recompute the POD basis. For this we need a measure for the error introduced by the reduced order model. In this work we utilize the residual as an error indicator. In particular we look at the error obtained by inserting the reduced order solution into the finite element model. The equations read as follows

$$\begin{aligned}
\mathrm{res}_y(t,\mu) &= M_1^N \Psi^y \mathrm{y}_t^\ell(t) + S_{c_1}^N \Psi^y \mathrm{y}^\ell(t) + \mathcal{N}^N(\Psi^y \mathrm{y}^\ell(t), \Psi^p \mathrm{p}^\ell(t), \Psi^q \mathrm{q}^\ell(t); \mu),\\
\mathrm{res}_p(t,\mu) &= S_{c_2(\Psi^y \mathrm{y}^\ell(t);\mu)}^N \Psi^p \mathrm{p}^\ell(t) + \mathcal{N}^N(\Psi^y \mathrm{y}^\ell(t), \Psi^p \mathrm{p}^\ell(t), \Psi^q \mathrm{q}^\ell(t); \mu),\\
\mathrm{res}_q(t,\mu) &= S_{c_3}^N \Psi^q \mathrm{q}^\ell(t) - \mathcal{N}^N(\Psi^y \mathrm{y}^\ell(t), \Psi^p \mathrm{p}^\ell(t), \Psi^q \mathrm{q}^\ell(t); \mu) - \mathcal{I}^N(t).
\end{aligned}$$

Then the error indicator is chosen as

$$\rho^2(\Psi,\mu) = \int_0^T \|\mathrm{res}_y(t,\mu)\|_{\mathrm{W}^{-1}} + \|\mathrm{res}_p(t,\mu)\|_{\mathrm{W}^{-1}} + \|\mathrm{res}_q(t,\mu)\|_{\mathrm{W}^{-1}}\,\mathrm{d}t, \quad (4.1)$$

with $\Psi = [\Psi^y, \Psi^p, \Psi^q]$ and $\mathrm{W} = M^N$. It is required to compute the dual norms of the residual which corresponds to solving linear systems. This type of error indicator is motivated by error estimators which are used in the reduced basis context [14,16,26]. The indicator measures the error of the given reduced order solution with respect to the finite element discretization for a given parameter μ. This error indicator is used to ensure that the reduced order model is sufficiently accurate during the optimization process.

Let us next describe how the different modules are used in the optimization process. For a given initial value $\tilde{\mu}$ the finite element model (3.4) is solved and a POD basis is computed for each variable using (3.7). By utilizing the obtained POD basis the reduced order model (3.10) is generated. Additionally, reduced order models are generated for the sensitivity equations. The subset selection is performed to determine the parameters that are identifiable. Non identifiable parameters are set to some reasonable or academic value and hence the parameter vector μ^0 is obtained. Starting with this parameter the optimization strategy is started. During the optimization process the reduced order model is utilized to solve the nonlinear system (2.1). After each solution process of (3.10) the error indicator (4.1) is evaluated in order to quantify the accuracy of the reduced order model. If the error indicator is too large (i.e. $\rho(\Psi, \mu) > \varepsilon_{res}$), the finite element model (3.4) is solved, new POD bases are determined and the associated reduced order models are generated. An Armijo backtracking strategy is applied for the step size control during the optimization process. The optimization method is stopped when a predefined tolerance ε_{opt} is reached (i.e. $\|\nabla J(\mu)\| < \varepsilon_{opt}$). Lastly, the error estimator (3.11) or (3.12) can be computed to verify that the obtained optimal solution is of good quality. The described strategy is summarized in Algorithm 1 in detail.

The introduced algorithm ensured that the solution to the finite element model and the reduced order model do not deviate too much from each other. Possible extension to the described strategy are to introduce additional POD bases for the sensitivity equations. In the presented examples this is not required in order to obtain accurate results. Note that in the presented algorithm the sensitivity equations are always solved through a reduced order model. Furthermore, the subset selection method could be applied in every iteration of the optimization problem. This particular scenario is not investigated in this paper. The focus here is to apply the subset selection method once to obtain a reasonable identifiable system.

The Gauss-Newton method can be replaced by any other optimization strategy. Especially, in order to incorporate constraints on the parameter μ a SQP method can be utilized, where the search direction d is solved by a quadratic programming problem in every iteration [25].

Algorithm 1 (Adaptive optimization strategy)

Require: Initial guess $\tilde{\mu}$, residual tolerance ε_{res}, stopping criteria ε_{opt}, maximum number of iteration $k_{\max}$ and $j_{\max}$

$[Y^N, P^N, Q^N] \leftarrow$Solve FE model (3.4) for $\tilde{\mu}$
$\Psi = [\Psi^y, \Psi^p, \Psi^q] \leftarrow$Compute POD basis by solving (3.7) for Y^N, P^N, and Q^N
Solve the sensitivity equation using POD for $\tilde{\mu}$ and Ψ
$\mu^0 \leftarrow$Perform the subset selection for $\tilde{\mu}$
$[Y^\ell, P^\ell, Q^\ell] \leftarrow$Solve ROM model (3.10) for μ^0 and Ψ
$\rho(\Psi, \mu^0) \leftarrow$ Evaluate error indicator (4.1) for Ψ and μ^0
if $\rho(\Psi; \mu^0) > \varepsilon_{res}$ **then**
  $[Y^N, P^N, Q^N] \leftarrow$Solve FE model (3.4) for μ^0
  $\Psi = [\Psi^y, \Psi^p, \Psi^q] \leftarrow$Compute POD basis by (3.7) for Y^N, P^N, and Q^N
end if
Solve the sensitivity equation using the reduced order model for μ^0 and Ψ
Set $k \leftarrow 0$ and evaluate cost $J^k \leftarrow J(\mu^k)$ and $\nabla J^k \leftarrow \nabla J(\mu^k)$
while $\nabla J^k > \varepsilon_{opt}$ and $k \leq k_{\max}$ **do**
  Evaluate Gauss-Newton matrix H^k
  Solve search direction d^k given by $H^k d^k = -\nabla J^k$
  Set $\alpha \leftarrow 1$ and $\bar{\mu} \leftarrow \mu^k + \alpha d^k$
  $[Y^\ell, P^\ell, Q^\ell] \leftarrow$Solve ROM model (3.10) for $\bar{\mu}$ and Ψ
  $\rho(\Psi; \bar{\mu}) \leftarrow$ Evaluate error indicator (4.1) for Ψ and $\bar{\mu}$
  if $\rho(\Psi, \bar{\mu}) > \varepsilon_{res}$ **then**
    $[Y^N, P^N, Q^N] \leftarrow$Solve FE model (3.4) for $\bar{\mu}$
    $\Psi = [\Psi^y, \Psi^p, \Psi^q] \leftarrow$Compute POD basis by (3.7) for Y^N, P^N, and Q^N
  end if
  Solve the sensitivity equation using the reduced order model for $\bar{\mu}$ and Ψ
  Evaluate cost $\bar{J} \leftarrow J(\bar{\mu})$ and $\nabla \bar{J} \leftarrow \nabla J(\bar{\mu})$
  while $\bar{J} \geq J^k + \sigma\alpha\langle \nabla J^k, d^K \rangle_2$ and $j \leq j_{\max}$ **do**
    Set $\alpha \leftarrow \frac{1}{2}\alpha$ and $\bar{\mu} \leftarrow \mu^k + \alpha d^k$
    Solve ROM model (3.10) for $\bar{\mu}$ and Ψ
    $\rho(\Psi, \bar{\mu}) \leftarrow$ Evaluate error indicator (4.1) for Ψ and $\bar{\mu}$
    if $\rho(\Psi, \bar{\mu}) > \varepsilon_{res}$ **then**
      $[Y^N, P^N, Q^N] \leftarrow$Solve FE model (3.4) for $\bar{\mu}$
      $\Psi = [\Psi^y, \Psi^p, \Psi^q] \leftarrow$Compute POD basis by (3.7) for Y^N, P^N, and Q^N
    end if
    Solve the sensitivity equation using POD for $\bar{\mu}$ and Ψ
    Evaluate cost $\bar{J} \leftarrow J(\bar{\mu})$ and $\nabla \bar{J} \leftarrow \nabla J(\bar{\mu})$
    Set $j \leftarrow j + 1$
  end while
  Set $\mu^{k+1} \leftarrow \bar{\mu}$, $J^{k+1} \leftarrow \bar{J}$ and $\nabla J^{k+1} \leftarrow \nabla \bar{J}$
  Set $k \leftarrow k + 1$;
end while
return Optimal solution $\mu^* \leftarrow \mu^k$ ($\varepsilon \leftarrow$Error estimator (3.11) or (3.12) (Optional))

4.2 Numerical results

To show the efficiency of the described algorithm we will present some numerical examples. We will consider two scenarios. First we will have a look at the case that the initial guess is close to the optimal solution. In the second case we choose an initial guess that is further away from the optimal solution. In both scenarios the obtained solution of Algorithm 1 is compared to a strategy using only the finite element model. This strategy is obtained by replacing the

reduced order solvers by finite element solvers in Algorithm 1. Additionally, after the optimization the error estimator is evaluated.

Let us introduce the settings used in the numerical experiments. We start by introducing the spatial domain $\Omega = [0, 5]$. Further, for the partitioning of the domain we set $s_1 = 2$ and $s_2 = 3$. Hence we get $\Omega = \Omega_l \cup \Omega_c \cup \Omega_r$ with $\Omega_l = [0, 2]$, $\Omega_c = (2, 3)$ and $\Omega_r = [3, 5]$. This partitioning is motivated by the structure found in lithium ion batteries; see, e.g., [11,32,33].

Next let us introduce the missing quantities for system (2.1). We start by introducing the diffusion coefficients. The piecewise constant diffusion coefficients are given as

$$c_1(x) = \begin{cases} 3 \text{ for } \mathbf{x} \in \Omega_l, \\ 4 \text{ for } \mathbf{x} \in \Omega_c, \\ 2 \text{ for } \mathbf{x} \in \Omega_r, \end{cases} \quad \text{and} \quad c_3(x) = \begin{cases} 1 & \text{for } \mathbf{x} \in \Omega_l, \\ 10^{-3} & \text{for } \mathbf{x} \in \Omega_c, \\ 5 & \text{for } \mathbf{x} \in \Omega_r. \end{cases} \tag{4.2a}$$

The nonlinear term c_2 and $\mathcal{N}$ are chosen as introduced in (2.2) and (2.3) together with the indicator function (2.4). Further, the boundary value for (2.1e) are given by

$$\mathcal{I}(t) = \frac{t}{2} \sin(2\pi t). \tag{4.2b}$$

Lastly, we set the initial condition (2.1f) for the variable y as

$$y_\circ(x) = 1. \tag{4.2c}$$

In the numerical experiments we set $T = 1$. For the space discretization second order finite elements are used. We choose 1000 elements and hence we end up with 1999 degrees of freedom for each variable. The time discretization is performed by an implicit Euler method with 100 time steps. The obtained nonlinear system is then solved by a Newton method to an accuracy of 10^{-10}. To guarantee that the Newton method converges a globalization strategy is applied [7]. For the reduced order model the number of basis functions for the three variables are set to $\ell_y = 16$, $\ell_p = 17$ and $\ell_q = 15$. These numbers are obtained by solving (3.7) under consideration of the accuracy of the eigenvalue solver. The nonlinear terms are evaluated with the empirical interpolation method as shown in [23]. For the nonlinearity $\mathcal{N}$ we utilize 22 basis functions and for c_3 we use 23. This corresponds to a maximum interpolation error of 10^{-11} during the construction.

Next let us introduce the settings for the optimization procedure. As the tolerance for the stopping criteria we chose $\varepsilon_{opt} = 10^{-6}$ and set the maximum number of iterations to $k_{\max} = 100$. Further, we set the tolerance for the error indicator to $\varepsilon_{res} = 10^{-4}$. The settings for the Armijo condition are $\sigma = 0.1$ and $j_{\max} = 10$. The target $q^d(t)$ is generated by solving the nonlinear system (2.1) for $\mu^\star = (1.1, -0.7, -0.1, 0.4)$ which then should be recovered by the parameter estimation.

In the subset selection procedure parameters associated to eigenvalues larger then 10^{-5} are selected for the optimization. This is a reasonable choice

and coincides with the accuracy of the finite element discretization and eigenvalue solver. The parameters excluded from the optimization are set to the corresponding exact values of $\mu^\star$. This is done in order to show the performance of the optimization algorithm and omit problems arising from the ill posedness of the problem.

All tests are performed on the same computer and software. To compare the performance we state the CPU times for the computation and the number of required finite element solves for the nonlinear system and for the sensitivity equations in parentheses.

4.2.1 Run 1

For the first numerical experiment we consider the case that the deviation of the initial parameter $\tilde{\mu}$ is in the neighborhood of the exact solution $\mu^\star$. For this we choose

$$\tilde{\mu} = (1.3 \cdot \mu_1^\star, 1.5 \cdot \mu_2^\star, 1.5 \cdot \mu_3^\star, 1.5 \cdot \mu_4^\star) = (1.43, -1.05, -0.15, 0.60).$$

This corresponds to a deviation of 30% in the first parameter and 50% in the other parameters. By applying the subset selection the parameter μ_2 is marked as non identifiable and therefore set to $\mu_2^0 = 0.7$. This is also indicated in the numerical results in Table 4.1 by a dagger. The obtained results of the optimization are presented in Table 4.1. Further, the computational expenses for the two approaches are outlined in the last two columns in form of CPU time and finite element solves. In Figure 4.1 (left plot) the target function $q^d(t)$ together with the output corresponding to the initial guess μ^0, i.e. after the subset selection, are shown. To compare the output corresponding to the optimal solution obtained by the two methods the absolute difference is plotted in Figure 4.1 (right plot). Lastly, in Table 4.1 the error estimator is presented.

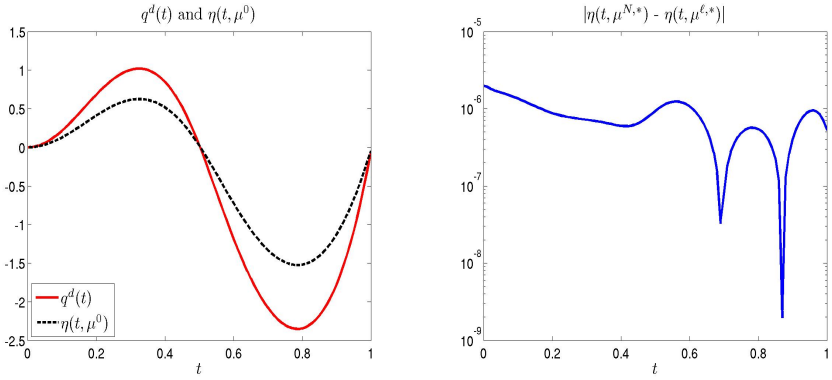

Fig. 4.1 Target function q^d and output function $\eta(t,\mu^0)$ (right plot) and difference between the output $\eta(t,\mu)$ for the optimal solutions μ^* obtained by using finite elements and adaptive reduced order models (right plot).

When looking at the results it can be seen that the optimization using adaptive reduced order models is very efficient. Only one finite element solution is required. This is due to the fact that we started close to the exact solution $\mu^\star$. The optimization strategy terminates after five iterations. Hence a speed up of more then seven is a very good result. Further, let us note that computing a single finite element solution takes approximately 27 seconds. This implies that the optimization using the reduced order model takes nine seconds which corresponds to a speedup of factor 26. The accuracy of the optimization using the reduced order model is very good when looking at the right plot presented in Figure 4.1. The obtained optimal parameter $\mu^{N,*}$ and $\mu^{\ell,*}$ by both strategies deliver almost the same values. Comparing it to the exact $\mu^\star$ it can be seen that the optimization strategies could recover the parameter almost exact in both cases.

Method	$\mu^{N,*}$ and $\mu^{\ell,*}$	FE solves	CPU time
FE	$(1.100000, -0.700000^\dagger, -0.100000, 0.400000)$	8(21)	260
ROM	$(1.100003, -0.700000^\dagger, -0.100000, 0.399995)$	1(0)	36

Table 4.1 Optimal solution obtained by the finite element (FE) approach and the adaptive algorithm using reduced order models (ROM). The † indicates the parameter fixed by the subset selection. In the last two columns the total expenses of the two approaches is compared.

Lastly, we look at the error estimator. We want to avoid computing the optimization using only finite element solvers. Further, the exact solution $\mu^\star$ is in general not known. Hence we need to estimate the error in the parameter $\mu^{\ell,*}$ obtained by the adaptive strategy. For this we evaluate the error estimator introduced in Section 3.3. The evaluation of this estimator is expensive. The computational time is approximately 39 seconds and involved the solution of the nonlinear system, the sensitivity equations and the adjoint equation using the finite element solvers. The obtained solution by the error estimator is very satisfying since it indicates that the error is below 1%. Note that the estimator is an upper bound and the actual error is lower. The efficiency of the error estimator is comparable to the results presented in [8,19].

$\|\mu^{N,*} - \mu^{\ell,*}\|_2$	(3.11)	(3.12)
0.000005	0.005945	0.008495

Table 4.1 Error estimator for the obtained solution $\mu^{\ell,*}$ by the adaptive reduced order model optimization.

4.2.2 Run 2

In the second numerical experiment we will use an initial guess that is further from the optimal solution. For this we set

$$\tilde{\mu} = (1.3 \cdot \mu_1^\star, 2 \cdot \mu_2^\star, 2 \cdot \mu_3^\star, 2 \cdot \mu_4^\star) = (1.43, -1.40, -0.20, 0.80).$$

Note that the first parameter $\tilde{\mu}_1$ is again deviated by 30%. The reason for this is that we need to stay in the admissible set $\mathcal{M}_{ad}$ in order to satisfy the conditions for Theorem 2.1. The other parameters are deviated by 100%. The subset selection method again selects the parameter μ_2 which is then set to the exact value 0.7 in the initial value μ^0. Compared to the previous experiment we expect that now the adaptivity of the algorithm will be active.

The results are outlined in Table 4.1. It can be seen that the adaptive strategy is active since now the nonlinear system has to be solved two times using finite elements. This leads to the fact that now the adaptive strategy became more expensive compared to the first scenario. The POD bases are recomputed in the second iteration. Still a speed up factor of approximately five is obtained. The optimization procedure again requires five iteration. This time the Armijo condition is active more often, hence the higher computational cost in the finite element case. The optimization using the adaptive reduced order models takes approximately ten seconds when excluding the time required by the finite element solves (27 seconds each). This corresponds to a speedup factor of 31. Comparing this to the previous run it can be seen that the cost for the optimization stayed the same only more finite element evaluations are needed to update the POD bases. In the case that the POD bases have to be updated more frequently, the optimization using adaptive reduced order models will have roughly the same computational cost as the optimization using finite elements. When introducing additional POD bases for the sensitivity equations and setting $\varepsilon_{res} = 0$ the adaptive strategy is equivalent to using only finite element solvers.

Method	$\mu^{N,*}$ and $\mu^{\ell,*}$	FE solves	CPU time
FE	$(1.100000, -0.700000^\dagger, -0.100000, 0.400000)$	10(21)	315
ROM	$(1.099997, -0.700000^\dagger, -0.100001, 0.399998)$	2(0)	64

Table 4.1 Optimal solution obtained by the finite element (FE) approach and the adaptive algorithm using reduced order models (ROM). The † indicates the parameter fixed by the subset selection. In the last two columns the total expenses of the two approaches is compared.

Lastly we again look at the error estimator. As in the previous case the estimators give an upper bound for the error in the optimal parameter $\mu^{\ell,*}$. The results are again very good and of equal quality and computational expenses.

$\|\mu^{N,*} - \mu^{\ell,*}\|_2$	(3.11)	(3.12)
0.000004	0.006375	0.002471

Table 4.1 Error estimator for the obtained solution $\mu^{\ell,*}$ by the adaptive reduced order model optimization.

Conclusion

In this paper an adaptive optimization strategy using reduced order models is presented. The reduced order models are obtains by the POD method. An application to a nonlinear system of partial differential equations with application in lithium ion battery modeling is outlined. In the numerical results the efficiency of the proposed method is shown.

References

1. R.A. Adams. *Sobolev Spaces*. Pure and Applied Mathematics, Vol. 65. Academic Press, New York-London, 1975.
2. H.W. Alt. *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Springer, Berlin Heidelberg, 5th edition, 2006.
3. J.J. Batzel, S. Fürtinger, M. Bachar, M. Fink and F. Kappel. Sensitivity identifiability of a baroreflex control system model. *Technical report*, Graz University of Technology, 2009.
4. M. Burth, G.C. Verghese and M. Vélez-Reyes. Subset selection for Improved parameter estimation in on-line identifiaction of a synchronous generators. *IEEE Transactions on Power Systems*, 14(1):218–225, 1999.
5. L. Cai and R. E. White. Reduction of model order based on proper orthogonal decomposition for lithium-ion battery simulations. *Journal of The Electrochemical Society*, 156(3):A154–A161, 2009.
6. R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I*. Series Computational Mathematics 35. Addison-Wesley, Springer-Verlag, 1992.
7. P. Deuflhard. *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*. Series Computational Mathematics 35. Springer, Berlin, 2004.
8. M. Dihlmann and B. Haasdonk. Certified PDE-constrained parameter optimization using reduced basis surrogate models for evolution problems. *Submitted*, 2013.
9. L.C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 2002.
10. M. Fink, A. Attarian and H. Tran. Subset selection for parameter estiamtion in an HIV model. *Proc. Appl. Math. Mech.*, 7:1121501–1121502, 2007.
11. T. Fuller, M. Doyle and J. Newman. Modeling of galvanostatic charge and discharge of the lithium ion battery/polymer/insertion cell. *Journal of the Chemical Society*, 140(6):1526–1533, 1993.
12. G. Golub. Numerical methods for solving least squares problems. *Numerische Mathematik*, 7:206-216, 1965.
13. M. Gubisch and S. Volkwein. *Proper Orthogonal Decomposition for Linear-Quadratic Optimal Control*. Submitted, 2013.
`www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/scripts.php`
14. M. Grepl. Certified reduced basis method for nonaffine linear time-varying and nonlinear parabolic partial differential equations. *WSPC: Mathematical Models and Methods in Applied Sciences*, 22(3):40, 2012.

15. M. Gu and S.C. Eisenstat. Efficient algorithms for computing a strong rank-revealing QR factorization. *SIAM Journal on Scientific Computing*, 17:848-869, 1996.
16. B. Haasdonk. Convergence rates of the pod-greedy method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2012.
17. M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer, Netherlands, 2009.
18. P. Holmes, J.L. Lumley, G. Berkooz, and C.W. Rowley. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge Monographs on Mechanics. Cambridge University Press, Cambridge, 2nd ed. edition, 2012.
19. E. Kammann, F. Tröltzsch and S. Volkwein. A method of a-posteriori error estimation with application to proper orthogonal decomposition. *ESAIM: M2AN*, 47:933934 218-225, 2013.
20. C.T. Kelley. *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, 1999.
21. K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, 90:117–148, 2001.
22. O. Lass. *Reduced-order modeling and parameter identi?cation for coupled nonlinear PDE systems*. Ph.D thesis, in preparation.
23. O. Lass and S. Volkwein. POD Galerkin schemes for nonlinear elliptic-parabolic systems. *SIAM Journal on Scientific Computing*, 35(3):A1271–A1298, 2013.
24. Y. Maday, N. C. Nguyen, M. Barrault and A. T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004.
25. J. Nocedal and S.J. Wright. *Numerical optimization*. Springer-Verlag, New York, 2006.
26. A.T. Patera and G. Rozza. *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*. MIT Pappalardo Graduate Monographs in Mechanical Engineering, MIT, version 1.0 edition, 2006.
27. J.G. Reid. Structural identifiability in linear time-invariant systems. *IEEE Trans Automat Control*, 22:242–246, 1977.
28. T. Seger. *Elliptic-Parabolic Systems with Application to Lithium-Ion Battery Models*. PhD thesis, University of Konstanz, 2013.
29. L. Sirovich. Turbulence and the dynamics of coherent structures. parts i-ii. *Quarterly of Applied Mathematics*, XVL:561–590, 1987.
30. F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*. American Math. Society, Providence, volume 112, 2010.
31. C.R. Vogel. *Computational Methods for Inverse Problems*. Frotiers in Applied Mathematics, SIAM, Philadelphia,
32. Y. Vutov, S. Mergenov, P. Popov and O. Iliev. Finite volume discretization of equations describing nonlinear diffusion in li-ion batteries. *Numerical Methods and Applications, Lecture Notes in Computer Science*, 6046:338–346, 2011.
33. J. Xu, J. Wu and H. Zou. On the well posedness of mathematical model for lithium-ion battery systems. *Methods and Applications of Analysis*, 13:275–298, 2006.