Parameter Estimation for High-Dimensional PDE Models Using a Reduced Approach

Robert Kircheis and Stefan Körkel

Abstract Partial differential equations (PDE) are indispensable to describe complex processes. PDE constrained parameter estimation is still a prevailing topic of research. The increase in computation time with increasing complexity of the problem is one of the main problems.

With the application of multiple shooting, the number of required derivatives for the generalized Gauss–Newton method rises rapidly. We introduce a method to overcome this challenge. By using directional derivatives the computational effort can be reduced to the minimal number. We demonstrate our methods with help of the heat equation.

1 Introduction

Validated models are essential for process optimization and optimal control in chemistry, engineering etc. Usually these models depend on parameters that are not known initially but have to be identified from measurement data. Derivative based methods, such as the generalized Gauss–Newton method for direct multiple shooting by Bock [5], have shown good results for parameter estimation problems with ordinary differential equations (ODEs).

If spatially distributed processes are taken into account, we have to consider constraints given by partial differential equations (PDEs). If PDEs are discretized by means of a method of lines, we end up with a high-dimensional system of differential algebraic equations (DAEs). In general, we formulate the parameter estimation problem as a nonlinear least squares problem and apply the generalized Gauss–Newton method. In the context of multiple shooting, the effort for the computation of the Jacobians in each iteration of the generalized Gauss–Newton method is tremendous. We present a method which couples the evaluation of the Jacobians and the subsequent block-Gaussian elimination. Thus, the number of required derivatives is reduced to the minimal number. The *reduced approach* was

R. Kircheis (✉) • S. Körkel
IWR Heidelberg, Heidelberg University, Heidelberg, Germany
e-mail: robert.kircheis@iwr.uni-heidelberg.de; stefan.koerkel@iwr.uni-heidelberg.de

T. Carraro et al. (eds.), *Multiple Shooting and Time Domain Decomposition Methods*, Contributions in Mathematical and Computational Sciences 9,
DOI 10.1007/978-3-319-23321-5_5

introduced by Schlöder [13] for parameter estimation problems constrained by high-dimensional systems of ODEs in 1987. A first extension to DAE constrained problem was presented by Bauer [2]. In this paper, we develop a different formulation of the reduced approach for DAE constraints which we consider as an approximation of the solution of a partial differential equation. The first application of the reduced approach to PDE constrained parameter estimation problems was presented by Dieses [1]. Dieses considered only ODEs to approximate the solution of a PDE.

We first introduce the general formulation of a parameter estimation problem and the generalized Gauss–Newton method. Afterwards, we present the reduced approach. In the end, an application example is investigated to show the advantages of our method compared to the conventional approach. In the end, some conclusions are drawn.

2 Problem Formulation

We consider a dynamic system defined by partial differential equations,

$$0 = \mathscr{F}\left(t, x, u, \frac{\partial u_i}{\partial t}, \frac{\partial u_i}{\partial x_j}, \frac{\partial u_i}{\partial^2 x_j \partial x_k}, \ldots, p\right), \tag{1}$$

on the bounded domain $\Omega \subset \mathbb{R}^d$. $\mathscr{F}$ is an arbitrary function with time $t = [t_0, t_{\text{end}}]$, independent variables $x \in \mathbb{R}^d$, dependent variables $u \in \mathbb{R}^{n_u}$ and parameters $p \in \mathbb{R}^{n_p}$. We examine transient problems. Thus, let initial values of the form

$$u(t_0) = u_0(p) \tag{2}$$

be given that may depend on the parameters. Additionally, let boundary conditions be defined by

$$au + b\frac{\partial u}{\partial n} = c \quad \text{on } \partial\Omega \tag{3}$$

for some constants a and b that can be zero but not at the same time and a given function c on the boundary of the domain.

Assume that n_{ex} experiments have been executed which have provided a set of measurements η_k^j, $k = 1, \ldots, m^j$, $j = 1, \ldots, n_{\text{ex}}$. By $h_k^j(t_k^j, u^{*,j}(t_k^j), p^*)$, $k = 1, \ldots, m^j$, $j = 1, \ldots, n_{\text{ex}}$, we denote the corresponding model response evaluated for the true, but unknown parameters p^* and states $u^{*,j}(t_k^j)$ computed by Eq. (1) for p^*. We assume the measurement errors

$$\varepsilon_k^j := \eta_k^j - h_k^j(t_k^j, u^{j,*}(t_k^j), p^*), \quad k = 1, \ldots, m^j, \ j = 1, \ldots, n_{\text{ex}},$$

to be normally distributed

$$\varepsilon_k^j \sim \mathcal{N}\left(0, \left(\sigma_k^i\right)^2\right), \quad k = 1, \dots, m^j,\ j = 1, \dots, n_{\mathrm{ex}},$$

with mean 0 and standard deviation σ_k^j. The difference between measurement values and model response can be evaluated for other values of p and u, too, and we obtain the residuals

$$\eta_k^j - h_k^j(t_k^j, u^j(t_k^j), p), \quad k = 1, \dots, m^j\ j = 1, \dots, n_{\mathrm{ex}}. \tag{4}$$

The parameter estimation problem is then to minimize the weighted sum of the residuals

$$\frac{1}{2}\sum_{j=1}^{n_{\mathrm{ex}}}\sum_{k=1}^{m^j}\left(\frac{\eta_k^j - h_k^j(t_k^j, u^j(t_k), p)}{\sigma_k^j}\right)^2 \tag{5}$$

with the variances σ_k^j as weights. Equation (5) can also be interpreted as a log-likelihood estimator for the parameters, see Seber [14].

Often, we have to deal with additional interior point and boundary constraints as well:

$$0 = \sum_{l=1}^{n_r^j} r_l^j(u^j(t_l^j), p) \tag{6}$$

with $r_l : \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to \mathbb{R}^r$.

Considering the model equations (1)–(3) and the interior point constraints (6) as additional constraints, we state the PDE constrained parameter estimation problem

$$\min_{y,p} \quad \frac{1}{2}\sum_{j=1}^{n_{\mathrm{ex}}}\sum_{k=1}^{m^j}\left(\frac{\eta_k^j - h_k^j(t_k^j, u^j(t_k^j), p)}{\sigma_k^j}\right)^2 \tag{7a}$$

$$\text{s.t.} \qquad 0 = \mathscr{F}^j\left(t, x, u^j, \frac{\partial u_i^j}{\partial t}, \frac{\partial u_i^j}{\partial x_j}, \frac{\partial u_i^j}{\partial^2 x_j \partial x_k}, \dots, p\right) \quad j = 1, \dots, n_{\mathrm{ex}}. \tag{7b}$$

$$u^j(t_0) = u_0^j(p), \tag{7c}$$

$$c^j = a^j u^j + b^j \frac{\partial u^j}{\partial n} \quad \text{on } \partial\Omega, \tag{7d}$$

$$0 = \sum_{l=1}^{n_r^j} r_l^j(u^j(t_l^j), p). \tag{7e}$$

3 Discretization in Space and in Time

In this section, we present concepts for the discretization of the model equations (1) in space and the concept of multiple shooting. If not declared otherwise, the following methods are presented for the first experiment only. For the remaining experiments, the steps have to be repeated. We neglect the superscript 1.

The first step in the parametrization of the parameter estimation problem (7) consists of the discretization of the PDE constraints (1) by a method of lines, cf. Schiesser [12] , e.g., by a finite difference methods (FDMs) or a finite element methods (FEMs).

The approach leads to a high-dimensional system of differential algebraic equations

$$A(y(t), z(t), p)\dot{y} = f(t, y(t), z(t), p), \tag{8a}$$

$$0 = g(t, y(t), z(t)), \tag{8b}$$

$$y(t_0) = y_0(p), \tag{8c}$$

where $A(y(t), z(t), p)$ denotes the mass matrix that may depend on the spatially discretized states $y(t) \in \mathbb{R}^{n_y}$ and $z(t) \in \mathbb{R}^{n_z}$ and the parameters p. We consider only DAEs with differentiation index 1, i.e. the matrix $\frac{\partial g}{\partial z}$ is regular.

To solve System (8) in time, we apply direct multiple shooting, see Bock [6]. Thus, the parameter estimation problem is transformed into a problem with finite dimensional constraints.

We define the *shooting grid*, i.e., a partition of the time interval $[t_0, t_{\text{end}}]$,

$$\tau_0 = t_0 < \tau_1 < \ldots < \tau_{n_{\text{ms}}} < \tau_{n_{\text{ms}}+1} = t_{end},$$

and the shooting intervals

$$\mathscr{I}_i = [\tau_i, \tau_{i+1}), \quad i = 0, \ldots, n_{\text{ms}}.$$

We introduce artificial initial values $s_i = (s_i^{y^T}, s_i^{z^T})^T$, $i = 0, \ldots, n_{\text{ms}}$, with

$$s_0^y = y_0(p) \tag{9}$$

for the differential and algebraic states $y(t)$ and $z(t)$, respectively, and examine the relaxed DAE system

$$A(y(t), z(t), p)\dot{y}(t) = f(t, y(t), z(t), p), \qquad t \in \mathscr{I}_i \tag{10a}$$

$$y(\tau_i) = s_i^y, \tag{10b}$$

$$0 = g(t, y(t), z(t), p) - \beta(t) g(\tau_i, s_i, p), \quad i = 0, \ldots, n_{\text{ms}}, \tag{10c}$$

$$\beta(t) \in [0, 1], \quad \beta(\tau_i) = 1, \tag{10d}$$

on each of the $n_{\mathrm{ms}} + 1$ subintervals. Here, $\beta(t)$ is a continuous, monotonically decreasing function with

$$\lim_{t \to \tau_{i+1}} \beta(t) = 0.$$

The evaluation of the DAE system leads to a step-by-step formulation of the trajectory

$$\begin{pmatrix} y(t) \\ z(t) \end{pmatrix} = \psi(t; s^i, p), \quad t \in \mathscr{I}_i,\ i = 0, \dots, n_{\mathrm{ms}}. \tag{11}$$

We refer to $\psi_i(t; s^i, p),\ t \in \mathscr{I}_i,\ i = 0, \dots, n_{\mathrm{ms}}$ as the nominal trajectory

By Eq. (11), we obtain a piecewise continuous, finite dimensional parametrization of the nominal trajectories of (10). To assure continuity of the trajectory for the solution $\hat{p}$ of the parameter estimation problem for the whole time interval and consistency for the algebraic equations, we impose continuity constraints

$$c(\tau_i, s_i, s_{i-1}, p) := \psi^y(\tau_i; s^y_{i-1}, p) - s^y_i = 0, \quad i = 1, \dots, n_{\mathrm{ms}} \tag{12}$$

and consistency constraints

$$g(\tau_i, s_i, p) = 0, \quad i = 0, \dots, n_{\mathrm{ms}}. \tag{13}$$

The variables $s_i,\ i = 0, \dots, n_{\mathrm{ms}}$ are additional degrees of freedom of the parameter estimation problem.

Before we formulate the finite dimensional constrained parameter estimation problem, we have to adjust the interior point constraints (6) to the shooting discretization. With the initial conditions (9) added to the set of constraints, we introduce a new vector of variables s^r that is locally uniquely determined by the interior point constraints, i.e. the matrix $\frac{\partial r}{\partial s^r}$, has full rank. Here, we use the definition

$$r := \sum_{l=1}^{n_r} r_l(\psi(t_l), p, s^r), \quad j = 1, \dots, n_{\mathrm{ex}},$$

where we neglect the dependencies of the nominal trajectories ψ of s and p, respectively.

When we present the generalized Gauss–Newton method in Sect. 5, we will clarify the necessity of the variables s^r in more detail. For later considerations, we define the vector

$$s := (s_0, \dots, s_{\mathrm{ms}}, s^r)$$

and the function

$$d(s,p) := \begin{pmatrix} y_0(s^r,p) - s_0^y, \\ r \end{pmatrix}. \tag{14}$$

Note, that the initial conditions may depend on the variables s^r, too.

Summarized, this results in the following finite dimensional nonlinear least squares problem for n_{ex} experiments:

$$\min_{s,p} \quad \frac{1}{2}\sum_{j=1}^{n_{\text{ex}}}\sum_{k=1}^{m^j}\left(\frac{\eta_k^j - h_k^j(t_k,\psi(t_k),p)}{\sigma_k^j}\right)^2 \tag{15a}$$

$$\text{s.t.} \quad 0 = c^j(\tau_i, s_i, s_{i-1}, p) - s_i^y, \qquad i = 1,\dots,n_{\text{ms}}^j,\ j = 1,\dots,n_{\text{ex}}, \tag{15b}$$

$$0 = g^j(\tau_i, s_i, p), \qquad i = 0,\dots,n_{\text{ms}}^j, \tag{15c}$$

$$0 = d^j(s,p). \tag{15d}$$

4 The Generalized Gauss–Newton Method

For readability, we introduce a shorter notation of Problem (15)

$$\min_{s,p} \quad \|F_1(s,p)\|_2^2 \tag{16a}$$

$$\text{s.t. } 0 = F_2(s,p). \tag{16b}$$

with $F_1(s,p) \in \mathbb{R}^{n_1}$, $F_2(s,p) \in \mathbb{R}^{n_2}$ and $(s,p) \in \mathbb{R}^n$. We use the definitions

$$F_1 := \left(\frac{\eta_k^j - h_k^j(t_k,\psi(t_k),p)}{\sigma_k^j}\right)_{\substack{k=1,\dots,m^j \\ j=1,\dots,n_{\text{ex}}}} \tag{17a}$$

$$F_2 := \begin{pmatrix} \left(c^j(\tau_i, s_i, s_{i-1}, p) - s_i^y\right)_{\substack{i=1,\dots,n_{ms}^j \\ j=1,\dots,n_{\text{ex}}}} \\ \left(g^j(\tau, s_i, p)\right)_{\substack{i=0,\dots,n_{ms}^j \\ j=1,\dots,n_{\text{ex}}}} \\ \left(d^j(s,p)\right)_{j=1,\dots,n_{\text{ex}}} \end{pmatrix} \tag{17b}$$

Bock [5] suggested to apply the generalized Gauss–Newton method to solve nonlinear least squares problems with ODE constraints. The first application to DAE constrained problems was presented by Bock et al. [7]. For a detailed description we refer to Körkel [10]. Problem (16) is solved iteratively by examining linearized equations

$$\min_{\Delta s, \Delta p} \quad \left\| F_1 + J_1 \begin{pmatrix} \Delta s \\ \Delta p \end{pmatrix} \right\|_2^2 \tag{18a}$$

$$\text{s.t. } 0 = F_2 + J_2 \begin{pmatrix} \Delta s \\ \Delta p \end{pmatrix}. \tag{18b}$$

Therefore, we need to compute the Jacobians

$$J_1 := \frac{dF_1}{d(s,p)}, \tag{19a}$$

$$J_2 := \frac{dF_2}{d(s,p)}. \tag{19b}$$

For problem (15) and $n_{\text{ex}} = 1$, the Jacobian has the following structure:

$$J = \begin{pmatrix} J_1 \\ \hline J_2 \end{pmatrix} = \left(\begin{array}{cccccc} D_0^1 & D_1^1 & \cdots & D_{n_{ms}}^1 & D_{s^r}^1 & D_p^1 \\ \hline G_0 & (-I,0) & & & G_0^{s^r} & G_0^p \\ & \ddots & \ddots & & \vdots & \vdots \\ & & G_{n_{\text{ms}}-1} & (-I,0) & G_{n_{\text{ms}}-1}^{s^r} & G_{n_{\text{ms}}-1}^p \\ H_0 & & & & H_0^{s^r} & H_0^p \\ & & \ddots & & \vdots & \vdots \\ & & & H_{n_m s} & H_{n_{ms}}^{s^r} & H_{n_{ms}}^p \\ D_0^2 & D_1^2 & \cdots & D_{n_{ms}}^2 & D_{s^r}^2 & D_p^2 \end{array} \right), \tag{20}$$

with the derivatives of

- of the residual of the measurements

$$D_i^1 := \frac{\partial F_1}{\partial s_i},\ i = 0, \ldots, n_{\text{ms}}, \quad D_{\hat{v}}^1 := \frac{\partial F_1}{\partial \hat{v}},$$

- of the continuity constraints

$$G_i := \frac{\partial \psi^y}{\partial s_i}(\tau_{i+1}; s_i, p),\ i = 0, \ldots, n_{\text{ms}} - 1, \quad G_i^{\hat{v}} := \frac{\partial \psi^y}{\partial \hat{v}}(\tau_{i+1}, s_i, p),$$

- of the consistency conditions

$$H_i := \frac{\partial g}{\partial s_i}(\tau_i, s_i, p),\ i = 0, \ldots, n_{\text{ms}}, \quad H_i^{\hat{v}} := \frac{\partial g}{\partial \hat{v}}(\tau_i, s_i, p),$$

- and the initial conditions and the interior point constraints

$$D_i^2 := \frac{\partial d}{\partial s_i},\ i = 0, \dots, n_{\text{ms}}, \quad D_{\hat{v}}^2 := \frac{\partial d}{\partial \hat{v}}.$$

with $\hat{v} = (p, s^r)$.

To assure uniqueness of the solution of Problem (16), we assume that the following two conditions hold for all values of $(s,\ p)$, where we have to evaluate F and J :

- Constraint Qualification (CQ)

$$\operatorname{rank} J_2(s,p) = n_2, \tag{21}$$

- Positive Definiteness (PD)

$$\operatorname{rank} J(s,p) = n. \tag{22}$$

We recall from Sect. 3 the dimension of F_2. The continuity conditions (12) sum up to $n_y \cdot n_{\text{ms}}$ constraints, the consistency constraints (13) provide $n_z \cdot (n_{\text{ms}} + 1)$ additional constraints and the initial conditions and the interior point and boundary constraints (14) results in $n_y + n_r$ equations. We end up with

$$n_2 = (n_y + n_z) \cdot (n_{\text{ms}} + 1) + n_r.$$

Since we have already defined

$$(n_y + n_z) \cdot (n_{\text{ms}} + 1)$$

shooting variables, we need to introduce n_r additional variables s^r to guarantee that (CQ) is fulfilled.

The linearization of (16) leads to a comparatively large, but sparse Jacobian. Bock [5] introduced the condensing algorithm for ODE constrained parameter estimation problems that exploits the sparse structure of (20) and eliminates the shooting variables $s_i,\ i = 1, \dots, n_{ms}$ by a block-Gaussian elimination. The condensed system depends only on $\Delta s_0^y,\ \Delta s^r$ and Δp

$$\min_{\Delta s_0, \Delta p} \quad \left\| u_1 + E_1 \Delta s_0^y + E_1^r \Delta s^r + E_1^p \Delta p \right\|_2^2 \tag{23a}$$

$$\text{s.t.} \quad 0 = u_2 + E_2 \Delta s_0^y + E_2^r \Delta s^r + E_2^p \Delta p \tag{23b}$$

By projecting on the algebraic variables $s_i^z,\ i = 0, \dots, n_{\text{ms}}$, the method can be applied to DAE constrained parameter estimation problems too, cf. Leineweber [11].

The procedure of evaluating the Jacobians first and applying the condensing method afterwards is referred to as the general approach. The general approach is implemented in the software package for parameter estimation *PARFIT* which is based on the methods presented in this section and in Bock [5] and Körkel [10].

5 The Reduced Approach

Especially for DAE constraints, that result from parametrized PDEs, the effort to compute (20) is tremendous due to the high dimension of s_i. The computation of submatrices G_i and H_i scales with the number of (discretized) states. For $(n_y + n_z) \gg n_p$, the effort to evaluate the blocks G_i and H_i dominates the evaluation of the Jacobian. The computation of these blocks requires the evaluation of $(n_y + n_z)$ variational differential equations. That is why the common approach is not suitable to solve PDE constrained parameter estimation problems in the context of multiple shooting.

Schlöder [13] developed an approach for high-dimensional ODE systems that couples the evaluation of the Jacobians and the subsequent condensing by using directional derivatives. Thereby, the effort of the computation of the Jacobian (20) reduces to the one of single shooting, i.e., the smallest possible number. Bauer [3] extended this method to parameter estimation problems with differential algebraic constraints.

We developed a different formulation of the reduced approach which fully eliminates the algebraic constraints and, thus, leads to a reduced condensed system of equal size as Problem (23). The approach of Bauer leads to redundant constraints which may cause numerical problems.

As in Sect. 3, we present the following method only for the first experiment and neglect the superscript 1. For the next steps, we assume that the interior point constraints and the residuals of the measurements can be written in the following form

$$0 = \sum_{i=0}^{n_{\text{ms}}} \sum_{t_k \in \mathscr{I}_i} \hat{h}_k(\psi(t_k), p) = \sum_{i=0}^{n_{\text{ms}}} R_i^1, \qquad \hat{h}_k = \frac{\eta_k - h_k}{\sigma_k}, \tag{24a}$$

$$0 = \sum_{i=0}^{n_{\text{ms}}} \sum_{t_l \in \mathscr{I}_i} r_l(\psi^j(t_l), p, s^r) = \sum_{i=0}^{n_{\text{ms}}} R_i^2. \tag{24b}$$

We refer to Eqs. (24) as separability conditions. We define the derivatives of (6) and (4) with respect to $\hat{v} = (p, s^r)$ according to

$$D_{\hat{v}}^1 = \frac{d}{d\hat{v}} \sum_{i=0}^{n_{\text{ms}}} \sum_{t_k \in \mathscr{I}_i} \hat{h}_k(\psi(t_k), p) = \sum_{i=0}^{n_{\text{ms}}} \frac{d}{d\hat{v}} \sum_{t_k \in \mathscr{I}_i} \hat{h}_k(\psi(t_k), p) = \sum_{i=0}^{n_{\text{ms}}} D_{\hat{v},i}^1.$$

$$D_{\hat{v}}^2 = \frac{d}{d\hat{v}} \sum_{i=0}^{n_{\text{ms}}} \sum_{t_l \in \mathscr{I}_i} r_l(\psi(t_l), p) = \sum_{i=0}^{n_{\text{ms}}} \frac{d}{d\hat{v}} \sum_{t_l \in \mathscr{I}_i} r_l(\psi(t_l), p) = \sum_{i=0}^{n_{\text{ms}}} D_{\hat{v},i}^2.$$

cf. Eq. (20). To eliminate the shooting variables at $t = \tau_0$, we define

$$d_0(s,p) := y_0(s^r,p) - s_0^y. \tag{25}$$

and we examine the rows of the Jacobian (20) that belong to the initial conditions (25) and to the consistency constraints at $t = \tau_0$

$$\begin{aligned} D_0^{s_0^y}\Delta s_0^y + D_0^{s_0^z}\Delta s_0^z + D_0^{s^r}\Delta s^r + D_0^p\Delta p + d_0(s,p) &= 0, \\ H_0^{s_0^y}\Delta s_0^y + H_0^{s_0^z}\Delta s_0^z + H_0^{s^r}\Delta s^r + H_0^p\Delta p + g(\tau_0,s_0,p) &= 0. \end{aligned} \tag{26}$$

Obviously, it holds

$$D_0^{s_0^y} = -I_{n_y}, \quad D_0^{s_0^z} = 0.$$

Since we consider only DAEs with differentiation index 1, the matrix

$$M := \begin{pmatrix} -I_{n_y} & 0 \\ H_0^{s_0^y} & H_0^{s_0^z} \end{pmatrix} \tag{27}$$

is regular. We eliminate $n_y + n_z$ variables Δs_0 formally from (26) and obtain

$$\begin{aligned} \begin{pmatrix} \Delta s_0^y \\ \Delta s_0^z \end{pmatrix} &= -\begin{pmatrix} -I_{n_y} & 0 \\ H_0^{s_0^y} & H_0^{s_0^z} \end{pmatrix}^{-1} \left[\begin{pmatrix} D_0^{s^r} \\ H_0^{s^r} \end{pmatrix} \Delta s^r + \begin{pmatrix} D_0^p \\ H_0^p \end{pmatrix} \Delta p + \begin{pmatrix} d_0 \\ g(\tau_0,s_0,p) \end{pmatrix} \right] \\ &= M_s^0 \Delta s^r + M_p^0 \Delta p + M_r^0 \end{aligned} \tag{28}$$

with

$$M_r^0 := -\begin{pmatrix} -I_{n_y} & 0 \\ H_0^{s_0^y} & H_0^{s_0^z} \end{pmatrix}^{-1} \begin{pmatrix} d_0 \\ g(\tau_0,s_0,p) \end{pmatrix}, \tag{29a}$$

$$M_s^0 := -\begin{pmatrix} -I_{n_y} & 0 \\ H_0^{s_0^y} & H_0^{s_0^z} \end{pmatrix}^{-1} \begin{pmatrix} D_0^{s^r} \\ H_0^{s^r} \end{pmatrix}, \tag{29b}$$

$$M_p^0 := -\begin{pmatrix} -I_{n_y} & 0 \\ H_0^{s_0^y} & H_0^{s_0^z} \end{pmatrix}^{-1} \begin{pmatrix} G_0^p \\ H_0^p \end{pmatrix}. \tag{29c}$$

We refer to (29) as *seed* matrices. The following steps are closely related to the ODE formulation of the reduced approach presented by Schlöder [13] with an extension to the consistency constraints which are solved locally for $s_j^z,\ j = 0, n_{\text{ms}}$.

The idea is to apply the explicit representation for the increments Δs_0 given by Eq. (28) to the evaluation of the remaining constraints. We define

$$\hat{E}_l^{p,0} := D_0^l \cdot M_p^0 + D_{p,0}^l, \tag{30a}$$

$$\hat{E}_l^{s,0} := D_0^l \cdot M_s^0 + D_{s^r,0}^l, \quad l = 1,2, \tag{30b}$$

$$\hat{u}_l^0 := D_0^l \cdot M_r^0 + R_0^l. \tag{30c}$$

Since we use the matrices $\hat{E}_l^{p,0}$, $\hat{E}_l^{s,0}$ and the vectors $\hat{u}_l^0$, $l = 1,2$, for the computation of the reduced condensed system, the notations in Eqs. (30) correspond to (23). Then, we compute recursively

$$\hat{E}_l^{p,i} = \hat{E}_l^{p,i-1} + D_i^l \cdot M_p^i + D_{p,i}^l, \tag{31a}$$

$$\hat{E}_l^{s,i} = \hat{E}_l^{s,i-1} + D_i^l \cdot M_s^i + D_{s^r,i}^l, \quad i = 1,\dots,n_{\text{ms}} \quad l = 1,2, \tag{31b}$$

$$\hat{u}_l^i = \hat{u}_l^{i-1} + D_i^l \cdot M_r^i + R_i^l. \tag{31c}$$

We set

$$\tilde{E}_l^p := \hat{E}_l^{p,n_{\text{ms}}}, \tilde{E}_l^s := \hat{E}_l^{s,n_{\text{ms}}}, \tilde{u}_l := \hat{u}_i^{n_{\text{ms}}}, \quad l = 1,2,$$

and obtain the reduced condensed system

$$\min_{\Delta p, \Delta s^r} \quad \frac{1}{2} \left\| \tilde{u}_1 + \tilde{E}_1^p \Delta p + \tilde{E}_1^s \Delta s^r \right\|_2^2 \tag{32a}$$

$$\text{s.t.} \quad 0 = \tilde{u}_2 + \tilde{E}_2^p \Delta p + \tilde{E}_2^s \Delta s^r. \tag{32b}$$

The seed matrices are updated iteratively by

$$M_p^i := \begin{pmatrix} G_i \\ H_{i+1} \end{pmatrix} \cdot M_p^{i-1} + \begin{pmatrix} G_i^p \\ H_{i+1}^p \end{pmatrix}, \tag{33a}$$

$$M_s^i := \begin{pmatrix} G_i \\ H_{i+1} \end{pmatrix} \cdot M_s^{i-1} + \begin{pmatrix} G_i^{s^r} \\ H_{i+1}^{s^r} \end{pmatrix}, \qquad k = 1,\dots,n_{\text{ms}}, \tag{33b}$$

$$M_r^i := \begin{pmatrix} G_i \\ H_{i+1} \end{pmatrix} \cdot M_r^{i-1} + \begin{pmatrix} c(\tau_{i+1}, s_{i+1}, s_i, p) \\ g(\tau_{i+1}, s_{i+1}, p) \end{pmatrix}. \tag{33c}$$

In Eqs. (30), (31) and (33), the expressions given by " $\cdot$ " are evaluated by directional derivatives and do not denote matrix products. In difference to the common approach, described in Sect. 4, we have to evaluate $n_p + n_r + 1$ directions instead of $n_y + n_z + n_p$.

If the initial conditions (14) are given in the form

$$d_0(s,p) = y_0(p) - s_0^y,$$

i.e., Eq. (14) does not depend explicitly on the variables s^r, the number of required directions is independent of the number of states.

After we have solved Problem (32) for Δp and Δs^r, the increments for the shooting variables $(s_1, \ldots, s_{\mathrm{ms}})$ are determined by

$$\Delta s_i^j = M_r^{i,j} + M_s^{i,j} \Delta s^r + M_p^{i,j} \Delta p, \quad i = 0, \ldots, n_{\mathrm{ms}}^j,\ j = 1, \ldots, n_{\mathrm{ex}}. \tag{34}$$

The methods, which have been presented in this section, have been implemented in a software package for parameter estimation called *PAREMERA*. *PAREMERA* is a new implementation in Fortran90 that is suited for the treatment of multiple experiments.

6 Example

In the following, we examine the 1D heat equation, e.g., see Evans [9]. which describes the distribution of heat in a given region over time. The system is defined by

$$0 = \frac{\partial u}{\partial t} - p_1 \nabla^2 u \tag{35a}$$

$$0 = u(t, 0) = u(t, 1), \tag{35b}$$

$$u(0, x) = -4 \cdot x \cdot (x - 1) \tag{35c}$$

with homogeneous Dirichlet conditions on the domain $\Omega \times T = [0, 1] \times [0, 1]$ (Fig. 1).

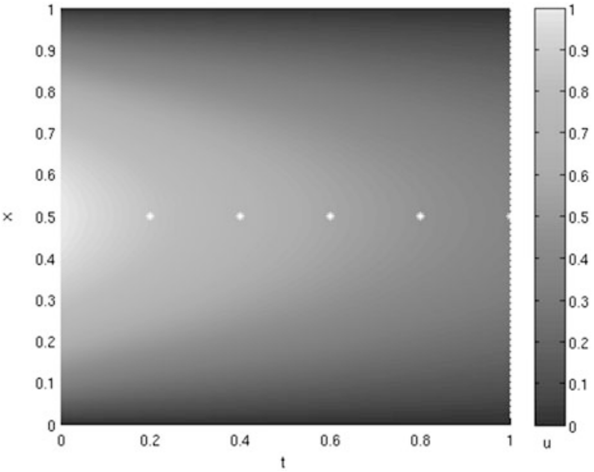

Fig. 1 Heat distribution over the domain $\Omega \times T$. The *white asterisks* mark the measurement points

Here, u is an arbitrary function, usually referred to as temperature. The parameter p_1 is the thermal diffusivity. We discretize (35) with second order central finite differences for three different mesh sizes $\Delta x = \{0.01, 0.002, 0.001\}$ to obtain an ODE system of $101,\ 501$ and 1001 states, respectively.

To compare the results between the reduced approach and the common one, the time interval is decomposed into four subintervals $[\tau_i, \tau_{i+1})$ with $\tau_i = 0.25 \cdot i,\ i = 0, \dots, 4$. We measure the peak at $x = 0.5$ at $t \in \{0.2,\ 0.4,\ 0.6,\ 0.8,\ 1.0\}$. The measurement data is determined by integrating the ODE system applying the software package *DAESOL* by Bauer [4] with known true parameter

$$p_1^* = 0.1$$

and adding Gaussian noise. For the calculations, p_1^* is scaled to one.

We compare the results of the two previously mentioned parameter estimation tools *PAREMERA* and *PARFIT*. Here, we use a version of *PARFIT* which is eligible for the exploitation of multiple experiment structures, see von Schwerin [15]. Both tools are embedded in software toolbox *VPLAN* by Körkel et al. [10].

We apply $p_1^0 = 2$ as starting parameter for all six settings (three mesh sizes and two algorithms). All computations are executed on a 64bit computer with 4 GB memory and an Intel® Core2Duo with 2 × 2.8 GHz. The results are listed in Table 1.

Both algorithms converge for all mesh sizes to approximately the same solution ($\hat{p}_1 \approx 1.00806$), but there is a significant difference in the time per iteration. For 101 states, PAREMERA is around four times faster then PARFIT. With increasing number of states, the difference in computational time increases. PAREMERA is 13 and 26 times faster for 501 and 1001 states, respectively.

Another fact that differs drastically, is the number of iterations. For all discretizations, PAREMERA achieves convergence after 5 iterations while PARFIT finds the solution after 10 iterations. This can be explained by the different types of globalization strategies. In PAREMERA, the restricted monotonicity test (RMT) is implemented as it is presented in Bock et al. [8]. In Parfit, only a first order Taylor series expansion is used to compute the curvature information.

Note, that the computed increments $(\Delta s, \Delta p)$ for the first iteration of both algorithms are exactly the same since both algorithms solve the same system.

Table 1 Survey of the results

# states		*PAREMERA*	*PARFIT*
101	# iterations	5	10
	Time per iteration	0.268 s	1.01 s
501	# iterations	5	10
	Time per iteration	3.44 s	45.62 s
1001	# iterations	5	10
	Time per iteration	15.06 s	392.5 s

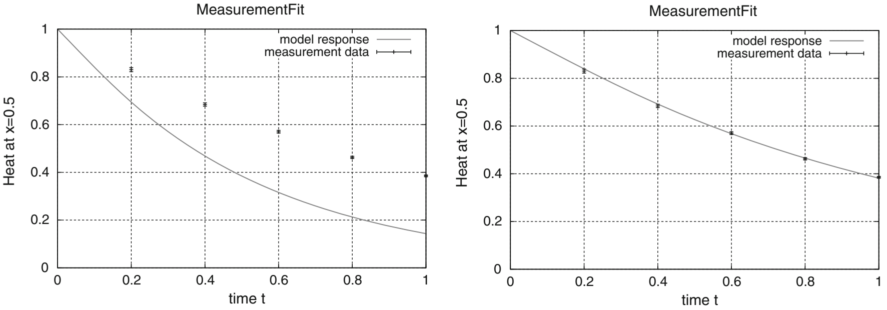

Fig. 2 Data fits for the starting parameter $p_1^0 = 2$ and the estimated parameter $\hat{p}_1 = 1.00806$

In Fig. 2, a comparison is shown between the two fitting curves with $p_1^0 = 2$ on the left hand side and the resulting parameter $\hat{p}_1 = 1.00806$ computed with *PAREMERA* on the right hand side, respectively.

Even for this rather small example we could show the advantages of the reduced approach. For more complex problems we expect even more significant savings in computation time. This is important to solve higher dimensional PDE problems or to do online parameter estimation. Thus, the reduced approach should be favored to solve this kind of problems.

Acknowledgements Financial support by BASF SE and HGS MathComp is gratefully acknowledged.

References

1. Altmann-Dieses, A., Schlöder, J.P., Bock, H.G., Richter, O.: Optimal experimental design for parameter estimation in column outflow experiments. Water Resour. Res. **38**, 1186ff (2002)
2. Bauer, I.: Numerische Verfahren zur Lösung von Anfangswertaufgaben und zur Generierung von ersten und zweiten Ableitungen mit Anwendungen bei Optimierungsaufgaben in Chemie und Verfahrenstechnik. Ph.D. thesis, Universität Heidelberg (1999)
3. Bauer, I.: Numerische Verfahren zur Lösung von Anfangswertaufgaben und zur Generierung von ersten und zweiten Ableitungen mit Anwendungen in Chemie und Verfahrenstechnik. Preprint, SFB 359, Universität Heidelberg (2001)
4. Bauer, I., Bock, H.G., Schlöder, J.P.: DAESOL—a BDF-code for the numerical solution of differential algebraic equations. Internal report, IWR, SFB 359, Universität Heidelberg (1999)
5. Bock, H.G.: Randwertproblemmethoden zur parameteridentifizierung in systemen nichtlinearer differentialgleichungen. Bonner Mathematische Schriften 183 (1987)
6. Bock, H.G., Plitt, K.J.: A Multiple Shooting algorithm for direct solution of optimal control problems. In: Proceedings of the 9th IFAC World Congress, pp. 242–247. Pergamon Press, Budapest (1984). Available at http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1984.pdf
7. Bock, H.G., Schlöder, J.P., Schulz, V.H.: Numerik großer Differentiell-Algebraischer Gleichungen—Simulation und Optimierung. In: Schuler, H. (ed.), Prozeß-Simulation, Chap. 2, pp. 35–80. VCH, Germany (1995)

8. Bock, H.G., Kostina, E.A., Schlöder, J.P.: On the role of natural level functions to achieve global convergence for damped Newton methods. In: Powell, M.J.D., Scholtes, S. (eds.) System Modelling and Optimization: Methods, Theory and Applications, pp. 51–74. Kluwer, Dodrecht (2000)
9. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society, Providence, RI (2010)
10. Körkel, S.: Numerische Methoden für optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen. Ph.D. thesis, Universität Heidelberg, Heidelberg (2002)
11. Leineweber, D.B.: Efficient reduced SQP methods for the optimization of chemical processes described by large sparse DAE models. Fortschritt-Berichte VDI Reihe 3, Verfahrenstechnik, vol. 613. VDI Verlag, Düsseldorf (1999)
12. Schiesser, W.E.: The Numerical Method of Lines: Integration of Partial Differential Equations, vol. 17. Academic, San Diego (1991)
13. Schlöder, J.P.: Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung. Dissertation, Hohe Mathematisch-Naturwissenschaftliche Fakultät der Rheinischen Friedrich-Wilhelms-Universität zu Bonn (1987)
14. Seber, G.A.F., Wild, C.J.: Nonlinear Regression. Wiley, New York (1989)
15. von Schwerin, R.: Numerical methods, algorithms, and software for higher index nonlinear differential-algebraic equations in multibody system simulation. Ph.D. thesis, Universität Heidelberg (1997)