IDENTIFICATION OF TRANSPORT COEFFICIENT MODELS IN CONVECTION-DIFFUSION EQUATIONS

MAKA KARALASHVILI[*], SVEN GROß[†], WOLFGANG MARQUARDT[*], ADEL MHAMDI[*], AND ARNOLD REUSKEN[‡]

Abstract. A rigorous method is presented for the systematic identification of the structure and the parameters of transport coefficient models in three-dimensional, transient convection-diffusion systems using high resolution measurement data. The transport is represented by a convection term with known convective velocity and a diffusion term with an unknown, generally state-dependent, transport coefficient. The identification of a transport coefficient model constitutes an ill-posed, highly nonlinear inverse problem. In our previous work [29], we presented a novel *incremental identification method*, which decomposes this inverse problem into easier to handle inverse subproblems. This way, the incremental identification method not only allows for the identification of the structure and the parameters of the model, but also supports the rigorous decision making on the best suited transport model structure. Due to the decomposition approach, the identified transport model structure and parameters are subject to errors. To cope with the error propagation inherent to the incremental method, the present work suggests a model correction procedure as a *supplement* to the incremental identification method [29], which results in a transport model of higher precision. The correction refers to both, the model structure and parameters. No a priori knowledge on the unknown transport model structure is necessary. The identification approach is numerically illustrated for a three-dimensional, transient convection-diffusion equation which has its origin in the modeling and simulation of energy transport in a laminar wavy film flow.

Key words. Modeling, identification, transport, convection-diffusion equation, inverse problem, regularization, parameter estimation, model selection, inexact Newton-type method.

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. This work is motivated by inverse problems which occur in the modeling of laminar wavy film flows. The goal of this paper is the identification of the structure *and* the parameters of transport models describing transport phenomena in laminar wavy film flows.

Falling film flows are characterized by the complex dynamics of developing surface waves (cf. e. g. , [33]). The direct numerical simulation of the detailed flow model in three space dimensions is complex and computationally demanding. Hence, such models can not be employed for model-based design of technical systems (e. g. , absorbers, evaporators, cooling towers, etc.). Therefore, simplified flow models are desirable. To this end, the real computational domain with its dynamical free boundary, i.e., the liquid phase, is approximated by a simplified computational domain with a *known* stationary (cf. e. g. , [44]) or slowly varying boundary. At the same time, *effective* transport coefficients are introduced to capture the enhanced, wave-induced transport in the reduced flow geometry. To support the design of technical systems, suitable transport coefficient models have to be developed, which accurately depict the characteristics of the film flow localised in time and space. In this paper, we consider the following transport model to approximate the behavior of wavy film flows.

Let $\Omega \subset \mathbb{R}^3$ be a computational domain corresponding to a reduced flow geometry. Let boundary parts $\Gamma_D \cup \Gamma_N = \partial\Omega := \Gamma$ denote the Dirichlet (index D) and Neumann

(index N) parts of the boundary, respectively, and let $[t_0, t_f]$ be the time interval of interest. Consider the balance equation

$$\rho(\frac{\partial u}{\partial t} + \mathbf{w} \cdot \nabla u) - \nabla \cdot ((a_{\mathrm{mol}} + f_w(u, \mathbf{x}, t, \theta))\nabla u) = 0 \,, \tag{1.1a}$$

$$(\mathbf{x}, t) \in \Omega \times (t_0, t_f]\,,$$

with initial and boundary conditions

$$\begin{aligned}
u(\mathbf{x}, t_0) &= u_0(\mathbf{x}) \,, \quad \mathbf{x} \in \Omega \,, \\
u(\mathbf{x}, t) &= g_D(\mathbf{x}, t) \,, \quad (\mathbf{x}, t) \in \Gamma_D \times [t_0, t_f] \,, \\
\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}, t) &= g_N(\mathbf{x}, t) \,, \quad (\mathbf{x}, t) \in \Gamma_N \times [t_0, t_f] \,.
\end{aligned} \tag{1.1b}$$

The scalar state variable $u(\mathbf{x}, t)$ represents, e. g., specific enthalpy in case of energy transport or mass density in case of mass transport. The vector field $\mathbf{w}(\mathbf{x}, t) \in \mathbb{R}^3$ represents the mass-averaged convection velocity, which is assumed to be known throughout this paper. $\rho(\mathbf{x}, t)$ stands for the density of the fluid and the vector $\mathbf{n}$ denotes the outer normal on the boundary Γ.

The *effective diffusive transport* distinguishes between different transport mechanisms, namely, - the molecular transport and the enhanced transport, which is induced by waves [9, 11, 44]. The transport coefficient $a(\mathbf{x}, t)$ is represented as a sum of two contributions (cf. (1.1a)):

$$a(\mathbf{x}, t) = a_{\mathrm{mol}} + f_w(u, \mathbf{x}, t, \theta) \,, \quad (\mathbf{x}, t) \in \Omega \times [t_0, t_f] \,. \tag{1.1c}$$

The *molecular transport* (e.g., the heat conduction in case of energy transport) is expressed by the *constant* molecular transport coefficient a_{mol}, the value of which corresponds to the thermal properties of the fluid and is thus *known*. The *unknown, wavy transport*, captures all remaining transport enhancing effects. It is expressed by a scalar model function $f_w(u, \mathbf{x}, t, \theta)$ of the state variable u, space $\mathbf{x}$, time t and model parameters $\theta \in \mathbb{R}^p$. We assume that the diffusive transport is constant near the boundary Γ, such that the boundary condition given at Γ_N (cf. (1.1b)) is well-defined (cf. e. g., [27]). The problem (1.1) is called the *direct* problem.

Our goal is the identification of a wavy transport model $f_w^\star(\cdot, \theta)$ in eq. (1.1) from transient, distributed measurement data $u_m(\mathbf{x}_j, t_k)$ describing energy or mass transport at a finite set of sampling points $(\mathbf{x}_j, t_k) \in \Omega \times [t_0, t_f]$. No specific structure of such a model is assumed to be available. Hence, a set of candidate model structures derived from different perspectives have to be considered. The "best" transport model has then to be selected from this set of candidates based on a meaningful measure of model quality. We call this model $f_w^\star(\cdot, \theta)$ the *"best-performing"* model. *With the term "best-performing" we refer to a transport model determined from a given measurement data set, which, among the proposed candidates, exhibits highest precision with respect to its structure $f_w^\star(\cdot, \theta)$ and its parameters θ in terms of the model quality measure employed.* As for the latter, we use Akaike's *minimum information theoretic criterion AIC* [1] (cf. Section 3.2), which is well-established and widely used in model identification, though there are alternative criteria for model structure discrimination [13]. Obviously, despite our choice of the term "best-performing model", we cannot claim that this model is reflecting reality close to perfection for the following two reasons. First, the set of candidate structures has to contain a model structure which matches reality well. Second, the identification of such a model constitutes a highly

nonlinear *inverse* problem, which is ill-posed in nature (cf. e.g., [17]). Since an evaluation of different candidate model structures is necessary, a naive treatment of this problem would require the solution of a series of complex, nonlinear inverse problems to estimate the parameters $\theta_i \in \mathbb{R}^{p_i}$ for *each* transport model $f_{w,i}(\,\cdot\,, \theta_i)$ from the list of candidates $i \in I := \{1, 2, \ldots, m\}, m \in \mathbb{N}$, in eq. (1.1).

Though there are many papers on parameter estimation in partial differential equations (cf. e.g., [3, 7, 8, 25, 35, 45] and references therein), the efficient solution of these problems is still difficult, since the computational effort grows strongly with the number of unknown model parameters (cf. [18]). Moreover, the numerical case studies treated are often restricted to one or two space dimensions. Furthermore, these studies do *not* aim at the identification of the most favorable transport model structure from a set of candidates.

In the above stated identification problem the following difficulties occur. First, the existence of a solution of the direct problem (1.1) for some transport model $f_w(\,\cdot\,, \theta)$ is not guaranteed [27]; second, in case of measurement errors an incorrect candidate model structure, biased or poor estimates are typically to be expected [43]; third, even in case the correct model structure is available, the lack of good initial guesses for parameter estimation may lead to estimates of poor quality (cf. e.g., [17]) or to a failure of the algorithm. These issues render the systematic identification of the most suitable model structure very challenging.

In contrast to the direct identification of a transport model in eq. (1.1) from a given set of candidates, we rely on the *incremental identification method*. The incremental identification method comprises a general framework, cf. [6, 10, 29, 31, 32, 34], and is not restricted to the particular class of inverse problems addressed in this paper. The main idea of incremental identification is the *decomposition* of the nonlinear identification problem in a series of subproblems which are easier to handle. This way, not only more reliable parameter estimates can be obtained for a given model structure, but the identification of the best model structure can be dealt with much more efficiently. On the other hand, due to the solution of a sequence of inverse subproblems, error propagation is inevitable (cf. [5, 29]). This paper will show however that such errors can be reduced to a large extent.

Recently, we applied and validated the incremental identification method for the identification of the transport model (1.1) [29]. In that paper, however, we have assumed *a known model structure* for the transport model. In the present paper, we relax this assumption and present a method, which aims at the identification of the transport model structure $f_w^\star(\,\cdot\,, \theta)$ and its parameters θ from a given candidate set, the so-called "best-performing" transport model. This method comprises a procedure for the efficient *reduction* of the errors resulting from the sequence of steps in incremental identification and is thus a *supplement* to the method reported previously [29]. It will be shown that no a priori knowledge on the transport mechanism is needed.

The paper is organized as follows. We first introduce the incremental identification method in Section 2. In Section 3, we outline the procedure for the reduction of the error induced by incremental identification. To this end, we aim at a correction in both, the structure and the parameters. This requires the solution of a transient inverse coefficient problem (cf., e.g., [2]), which is solved by means of an inexact Newton method. Regularization of this inverse problem is discussed in detail. A numerical case study is presented in Section 4 for an illustration of the suggested approach. The numerical example is the same as the one treated in [29]. In Section 5, conclusions are given and suggestions for future work are formulated.
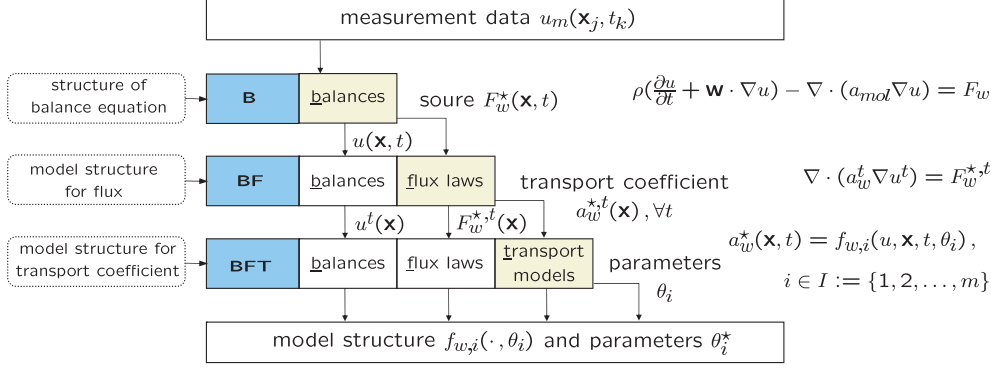
FIG. 2.1. *Incremental modeling and identification of transport model structure and parameters.*

2. Incremental modeling and identification.

Incremental identification reflects the steps of *incremental modeling* (cf. Fig. 2.1). Incremental modeling refines the model structure in a structured step-by-step procedure [31, 32] to allow for a transparent and systematic modeling process which exploits the structure naturally occurring in any model. In particular, the constitutive equations are introduced on different modeling levels to (hopefully) reduce the structural uncertainty (to a minimum). Incremental identification follows the steps of incremental modeling and decouples the identification process in a sequence of easier to handle subproblems. This strategy facilitates an efficient identification of a suitable model structure and its associated parameters. We present a brief description of incremental identification method here, further details are given in [29].

Incremental identification assumes that transient, distributed measurement data $u_m(\mathbf{x}_j, t_k)$ is available at sufficiently high resolution at spatial positions $\mathbf{x}_j \in \Omega$ and at time instants $t_k \in [t_0, t_f]$ (cf. Fig. 2.1).

The balance equations constitute the first modeling level – level B – of the incremental modeling procedure (cf. Fig. 2.1). On this level, the source $F_w(\mathbf{x}, t)$ which models the *unknown*, wavy diffusive transport is introduced without further specifications. Consequently, in the first identification step, this source F_w is reconstructed as a function of space $\mathbf{x}$ and time t using the measurement data $u_m(\mathbf{x}_j, t_k)$. Hence, the inverse source problem [2]

$$B: \qquad \rho(\frac{\partial u}{\partial t} + \mathbf{w} \cdot \nabla u) - \nabla \cdot (a_{\mathrm{mol}} \nabla u) = F_w(\mathbf{x}, t), \ (\mathbf{x}, t) \in \Omega \times (t_0, t_f]. \qquad (2.1)$$

has to be solved. The initial and boundary conditions are the same as in eq. (1.1b). This inverse problem is affine-linear, such that solution techniques for linear inverse problems can be employed (cf., e. g., [19, 24]). We denote the solution of this inverse problem as $F_w^\star(\mathbf{x}, t)$.

In the second modeling level – level BF – a suitable constitutive relation for the wavy source F_w is chosen (cf. Fig. 2.1). Here, this source is described by a diffusive transport law, where the flux is assumed to be proportional to ∇u. Accordingly, the *wavy transport coefficient* $a_w(\mathbf{x}, t)$ is introduced to depict the enhanced wavy transport contribution. This defines the second identification step, where the previously reconstructed source $F_w^\star(\mathbf{x}, t)$ is used as inferential measurement data to reconstruct the unknown, wavy transport coefficient a_w as a function of space $\mathbf{x}$ and time t (cf. Fig. 2.1). Since diffusive transport is typically time-invariant, we introduce a decoupling of

4

the space- and time-dependent function $\xi(\mathbf{x}, t)$ as $\xi^t(\mathbf{x}) := \xi(\mathbf{x}, t)$, $(\mathbf{x}, t) \in \Omega \times [t_0, t_f]$. The wavy transport coefficients $a_w^t(\mathbf{x})$ are then reconstructed for each time $t \in [t_0, t_f]$ in the $t-$parametric diffusion equation

$$BF: \qquad \nabla \cdot (a_w^t(\mathbf{x}) \nabla u^t) = F_w^{\star, t}(\mathbf{x}), \ (\mathbf{x}, t) \in \Omega \times [t_0, t_f]. \qquad (2.2)$$

The boundary conditions carry over from eq. (1.1b) for each t. These inverse problems are referred to as coefficient inverse problems [2]. They are stationary and nonlinear. The solutions of these problems are denoted by $a_w^{\star, t}(\mathbf{x})$.

On the last modeling level – level BFT – a set of different transport models $f_{w,i}(\cdot, \theta_i)$, $i \in I$, are proposed as closure constitutive relations (cf. Fig. 2.1) to correlate a_w with the state u and the model parameters $\theta_i \in \mathbb{R}^{p_i}$. Hence, the reconstructed wavy transport coefficients $a_w^{\star, t}(\mathbf{x})$ are assembled for $t \in [t_0, t_f]$ to result in the function

$$a_w^\star(\mathbf{x}, t) = a_w^{\star, t}(\mathbf{x}), \ (\mathbf{x}, t) \in \Omega \times [t_0, t_f]. \qquad (2.3)$$

This "model-free" wavy transport coefficient is first employed for the generation of candidate transport model structures $f_{w,i}(u, \mathbf{x}, t, \theta_i)$, $i \in I$, and next for the estimation of their associated parameters θ_i in each of them. In this step, nonlinear regression problems are stated and solved. These inverse problems are of algebraic nature only. The resulting parameter estimates are denoted by $\theta_i^\star$, $i \in I$.

Note, that the inverse problems arising in the incremental steps can be solved by applying *different, problem-adapted* solution techniques. In our previous paper [29], we have formulated the inverse problems for source estimation in the first step B (cf. eq. (2.1)) and for coefficient estimation in the second step BF (cf. eq. (2.2)) as optimization problems and solved them using suitable iterative solution techniques for linear and nonlinear inverse problems. In the last step BFT (cf. Fig. 2.1), we have employed standard least-squares solution techniques (cf. [36]) for parameter estimation.

It is clear, however, that the decoupling into a sequence of inverse problems and their numerical solution results in an inevitable error propagation. Nevertheless, a subset $I_s \subseteq I = \{1, 2, \ldots, m\}$ of *reasonable* transport model structures can be chosen by eliminating models showing a poor quality measure (such as the *AIC* criterion used in this paper). To identify a transport model $f_w^\star(\cdot, \theta)$ of and its parameters θ, it suffices to restrict the list of all candidate models to the chosen subset I_s in the model correction procedure, i. e., only the model structures

$$f_{w,i}(\cdot, \theta_i), \ i \in I_s \subseteq I. \qquad (2.4)$$

will be considered.

We remark that after carrying out all three steps of the incremental identification method *without* the correction step, a *single* transport model $f_{w,i}^\star(\cdot, \theta_i)$ showing the best quality measure could be chosen. However, in general, this model is *different* from the "best-performing" transport model $f_w^\star(\cdot, \theta)$ determined *with* the correction step, due to uncertainty in both, structure $f_w(\cdot, \theta)$ and parameters θ_i.

3. Model correction. In this section, we propose a correction approach in which we address the error propagation through the steps of the incremental identification method. To facilitate efficient error correction *independently* of the employed transport model structure, we treat the error caused in the first two steps of the incremental identification method (cf. Fig. 2.1) separately from the error in the last step.
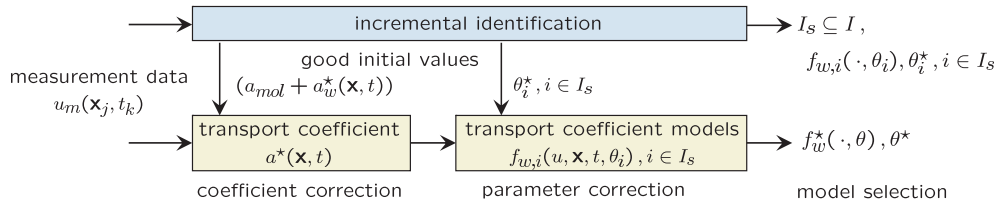
FIG. 3.1. *Correction of the transport model structure and parameters.*

Hence, we again decouple parameter estimation and model selection in the nonlinear problem (1.1) by carrying out *two* subsequent steps (cf. Fig. 3.1).

In the *coefficient correction* step, we first correct the transport coefficient $a(\mathbf{x}, t)$ (cf. eq. (1.1c)) by assembling the first two steps of the incremental identification method into *one*. Thus, the constitutive equation (2.2) and the balance equation (2.1) are combined into *one* equation. Consequently, the transport coefficient a is reconstructed as a function of space $\mathbf{x}$ and time t in the convection-diffusion equation

$$\rho(\frac{\partial u}{\partial t} + \mathbf{w} \cdot \nabla u) - \nabla \cdot ((a_{\mathrm{mol}} + a_w)\nabla u) = 0\,, \ (\mathbf{x}, t) \in \Omega \times (t_0, t_f]\,, \qquad (3.1)$$

with initial and boundary conditions as in eq. (1.1b). This, again, corresponds to a coefficient inverse problem [2]. However, compared to the coefficient inverse problems in eq. (2.2), this nonlinear inverse problem has higher complexity due to the time-dependence and the presence of a convective transport term. An efficient treatment of this inverse problem is achieved by using the wavy transport coefficient $a_w^\star$ (cf. eq. (2.3)) obtained from solving the inverse problem involving eq. (2.2) as a *good* initial value, i.e.,

$$a^{\mathrm{init}}(\mathbf{x}, t) = a_{\mathrm{mol}} + a_w^\star(\mathbf{x}, t)\,. \qquad (3.2)$$

As a result, the diffusion transport coefficient $a^\star(\mathbf{x}, t)$ is obtained (cf. Fig. 3.1).

In the subsequent *parameter correction* step of the correction procedure, the model structures (2.4) are considered. For the estimation of parameters of each considered candidate model, the wavy transport coefficient $(a^\star(\mathbf{x}, t) - a_{\mathrm{mol}})$ from the coefficient correction step is used as inferential measurement data. The optimal parameter values $\theta_i^\star, i \in I_s$, available from the third step BFT of the incremental identification procedure are employed as initial guesses for efficient parameter correction (cf. Fig. 3.1). Since the estimation problems in this step are of algebraic nature, the consideration of a larger number of candidate model structures $f_{w,i}(\,\cdot\,, \theta_i)\,, i \notin I_s$ does not constitute a significant effort.

Finally, the model selection step is carried out to select the "best-performing" transport model structure $f_w^\star(\,\cdot\,, \theta)$ from the set of regressed algebraic candidate models. Here, the quality measures (such as the AIC criterion used in this work) of the considered candidate models (cf. eq. (2.4)) are recomputed and the best performing candidate is selected (cf. Fig. 3.1).

3.1. Correction of the transport coefficient. In this section, we consider the inverse coefficient problem involving eq. (3.1) for the reconstruction of a corrected transient transport coefficient $a^\star(\mathbf{x}, t)$ from the transient, distributed measurement data $u_m(\mathbf{x}_j, t_k)\,, (\mathbf{x}_j, t_k) \in \Omega \times [t_0, t_f]$ to implement the coefficient correction step depicted in Fig. 3.1. Considerable literature is available on the solution of inverse

coefficient problems (e. g., [3, 14, 26, 35, 40]). However, only a few of them address the problem in three spatial dimensions (e. g., [4, 20, 22, 23]). Moreover, such problems are typically considered without a convective transport term, which is difficult to treat numerically.

3.1.1. Optimization-based formulation. We formulate the coefficient inverse problem as an optimization problem and solve it using an inexact Newton method which is considered to be very suitable for the solution of highly nonlinear ill-posed problems [17]. We use the so-called *weighted minimum-norm* solution [17] (cf. Section 3.1.4). We assume the unknown transport coefficient $a(\mathbf{x}, t)$ in $\mathcal{X} \subset \mathcal{L}^2(t_0, t_f; \mathcal{H}^1(\Omega))$. The transient measurement data $u_m(\mathbf{x}_j, t_k)$ can be represented in the finite element space $\mathcal{Y}_h \subset \mathcal{Y} \subset \mathcal{L}^2(\Omega \times [t_0, t_f])$. This representation is denoted by $u_m \in \mathcal{Y}$. The corresponding norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$ are

$$\|\cdot\|_{\mathcal{X}}^2 = \int_{t_0}^{t_f} \int_{\Omega} (\cdot)^2 d\mathbf{x} dt + \int_{t_0}^{t_f} \int_{\Omega} \nabla(\cdot)^2 d\mathbf{x} dt \,, \tag{3.3a}$$

$$\|\cdot\|_{\mathcal{Y}}^2 = \int_{t_0}^{t_f} \int_{\Omega} (\cdot)^2 d\mathbf{x} dt \,. \tag{3.3b}$$

Note, that the use of the space $\mathcal{X}$ with norm (3.3a) for the unknown quantity implies an a priori smoothness assumption on the solution of the inverse coefficient problem.

The optimization-based formulation of the coefficient inverse problem allows for a model-free, *function* estimation. The transport coefficient $a^\star(\mathbf{x}, t) \in \mathcal{X}$ minimizes the quadratic objective function

$$J_1(a) = \frac{1}{2} \|u(\mathbf{x}, t; a) - u_m\|_{\mathcal{Y}}^2 \tag{3.4}$$

with transient measurement data u_m as defined before. Here, $u(\mathbf{x}, t; a)$ is the solution of the direct problem (3.1) with initial and boundary conditions (1.1b) for some transport coefficient a. The initial and boundary data u_0, g_D and g_N, respectively, are assumed to be known [2].

3.1.2. Iterative optimization strategy. For the minimization of the objective function (3.4) subject to the constraints (3.1) and (1.1b), an adapted Newton type method is employed. In this method, an outer (Newton) iteration is combined with an inner conjugate gradient (CG) iteration applied to the normal equation (NE), which is referred to as CGNE method in the literature [21]. The stated nonlinear inverse problem is first *linearized* to benefit from the regularization properties of the inner (CGNE) iteration. A special stopping rule of the CGNE iteration results in a *regularized* update in the Newton iteration. For a (convergence) analysis of this method we refer to [23]. In case of nonlinear inverse problems, an additional strong condition has to be imposed on the nonlinear remainder of the linearization of the problem to achieve convergence (cf. Chapter 11 in [17], [23]). This condition restricts essentially the solution space of the nonlinear inverse problem and requires initial values sufficiently close to the solution.

The *truncated Newton-CGNE* [23] algorithm starts with an initial guess $a^0 \in \mathcal{X}$ for the transport coefficient and updates it in the Newton iterations according to

$$a^{n+1} = a^n + \alpha^n \hat{x}^n \,, \ n \in \mathbb{N} \,. \tag{3.5}$$

$\hat{x}^n$ stands for the Newton update, which is controlled by introducing a step-length parameter α^n. Instead of the *exact* Newton step $\alpha^n = 1 \,, \forall n$, we implement a *damped*

Newton step by choosing the step-length using the (first-order) standard Armijo-rule [36]. This damping strategy is different from [23]; since the studied coefficient inverse problem is highly nonlinear the Newton unit step-length needs to be appropriately shortened.

The Newton update $\hat{x}^n$ results from the inner CGNE iteration which solves the linearization of the stated inverse problem. Starting from $x^0 = 0$, this value is updated by moving along a conjugate descent direction d^k with an optimal step-length μ^k at each CGNE iteration k. CGNE iterations also require updates for (linear) residuals r^k. The update formulas are given by (cf. [21])

$$x^{k+1} = x^k + \mu^k d^k\,, \tag{3.6a}$$

$$r^{k+1} = r^k - \mu^k S(a^n, d^k)\,. \tag{3.6b}$$

Here, each new descent direction d^k in eq. (3.6a) is calculated from

$$d^{k+1} = G(a^n, r^k) + \gamma^k d^k\,, \tag{3.7}$$

with the gradient of the objective functional $G(a^n, r^k) := \frac{dJ_1}{da}(a^n)r^k$ and the conjugate coefficients γ^k given by

$$\gamma^0 = 0\,, \ \gamma^k = \frac{\left\|G(a^n, r^{k+1})\right\|_{\mathcal{X}}^2}{\left\|G(a^n, r^k)\right\|_{\mathcal{X}}^2}\,, k \geq 1\,. \tag{3.8}$$

Using variational calculus, it can be shown that $G := G(a^n, r^k)$ is the solution of the boundary value problem

$$
\begin{aligned}
-\Delta G + G &= -\nabla u(a^n)\nabla\psi(a^n, r^k)\,, \ \mathbf{x} \in \Omega\,, \\
G &= 0\,, \ \mathbf{x} \in \Gamma_D\,, \\
\frac{\partial G}{\partial\mathbf{n}} &= g_N\psi(a^n, r^k)\,, \ \mathbf{x} \in \Gamma_N\,.
\end{aligned}
\tag{3.9}
$$

for each $t \in [t_0, t_f]$. $u(a^n) = u(\mathbf{x}, t)$ denotes the solution of the *direct* problem (3.1), (1.1b) with the insertion $a(\mathbf{x}, t) = a^n(\mathbf{x}, t)$ for the transport coefficient from the n-th Newton iteration. $\psi(a^n, r^k) = \psi(\mathbf{x}, t)$ represents the solution of the *adjoint* problem

$$\rho(-\frac{\partial\psi}{\partial t} - \mathbf{w} \cdot \nabla\psi) - \nabla \cdot (a^n\nabla\psi) = r^k\,, \ (\mathbf{x}, t) \in \Omega \times [t_0, t_f)\,, \tag{3.10a}$$

with *terminal* and boundary conditions

$$
\begin{aligned}
\psi(\mathbf{x}, t_f) &= 0\,, \quad \mathbf{x} \in \Omega\,, \\
\psi(\mathbf{x}, t) &= 0\,, \quad (\mathbf{x}, t) \in \Gamma_D \times [t_0, t_f]\,, \\
\frac{\partial\psi}{\partial\mathbf{n}}(\mathbf{x}, t) &= 0\,, \quad (\mathbf{x}, t) \in \Gamma_N \times [t_0, t_f]\,.
\end{aligned}
\tag{3.10b}
$$

The derivation of the adjoint problem is based on standard procedures (cf. e.g. [2]). By introducing a new time variable $t_a = (t_0 + t_f - t)$ we obtain an equation with the same structure as the direct problem (3.1), (1.1b), however, with different boundary conditions.

The CGNE step-length μ^k in eq. (3.6) is given by

$$\mu^k = \frac{\left\|G(a^n, r^k)\right\|_{\mathcal{X}}^2}{\left\|S(a^n, d^k)\right\|_{\mathcal{Y}}^2}\,. \tag{3.11}$$

Here and in the update formula for residuals (3.6b), $S(a^n, d^k)$ is the solution of the *sensitivity* problem

$$\rho(\frac{\partial S}{\partial t} + \mathbf{w} \cdot \nabla S) - \nabla \cdot (a^n \nabla S) = \nabla \cdot (d^k \nabla u(a^n)) , \ (\mathbf{x}, t) \in \Omega \times (t_0, t_f] , \quad (3.12\text{a})$$

$$\begin{aligned} S(\mathbf{x}, t_0) &= 0 , & \mathbf{x} \in \Omega , \\ S(\mathbf{x}, t) &= 0 , & (\mathbf{x}, t) \in \Gamma_D \times [t_0, t_f] , \\ \frac{\partial S}{\partial \mathbf{n}}(\mathbf{x}, t) &= 0 , & (\mathbf{x}, t) \in \Gamma_N \times [t_0, t_f] . \end{aligned} \quad (3.12\text{b})$$

Here, again, $u(a^n) = u(\mathbf{x}, t)$ represents the solution of the direct problem (3.1), (1.1b) for the Newton update a^n of the transport coefficient and the conjugate descent direction d^k. For the derivation of the sensitivity problem we refer to, e. g. , [2].

In summary, the truncated Newton-CGNE method solves the direct problem (3.1) in each Newton iteration. In each CGNE iteration one has to solve an adjoint problem (3.10) and additionally eq. (3.9) for each $t \in [t_0, t_f]$ for the determination of the descent direction. Furthermore, in each iteration a sensitivity problem (3.12) for the computation of the step length has to be solved. It is very important to terminate the CGNE iterations early enough to handle the ill-posedness of the linearization of the stated inverse problem [21, 23]. One possible *stopping rule* for the CGNE iterations [23] is given by

$$\left\| r^k \right\|_{\mathcal{Y}} < \eta \left\| u_m - u(a^n) \right\|_{\mathcal{Y}} , \ 0 < \eta < 1 . \quad (3.13)$$

Here, again, r^k represents the (linear) residual from the k-th CGNE iteration (cf. (3.6b)) and $u(a^n) = u(\mathbf{x}, t)$ is the solution of the direct problem (3.1), (1.1b) with $a(\mathbf{x}, t) = a^n(\mathbf{x}, t)$ of the n-th Newton iteration. The tolerance parameter η has to be chosen appropriately (cf. Section 3.1.4).

3.1.3. Solution of the PDE problems. For the numerical realization of the iterative optimization strategy, the direct problem (3.1), (1.1b), the adjoint problem (3.10) and the sensitivity problem (3.12) have to be solved. Also, eq. (3.9) has to be solved for each $t \in [t_0, t_f]$ for the gradient calculation. All problems are of elliptic type, hence, similar numerical techniques can be employed for their solution.

Approximation of the solutions of all three-dimensional problems are calculated by means of the software package DROPS [16], which is based on multilevel nested grids and conforming finite-element discretization methods. For time discretization, a standard one-step θ-method is used. For space discretization, piecewise linear finite-elements on a tetrahedral grid are employed. The resulting discrete systems of linear equations are solved by suitable preconditioned Krylov subspace methods. The calculation of the norms (cf. eq. (3.11)) and computation of the conjugate coefficients $\gamma_k , k \geq 1$ (cf. eq. (3.8)) in the numerical optimization procedure are efficiently realized in DROPS.

In this paper, we do not study the performance of the solvers employed for the solution of the partial differential equations that occur in the truncated Newton-CGNE method. We use a fixed (quasi-uniform) mesh for discretization and prescribe a tolerance to which the resulting linear systems are solved.

3.1.4. Regularization. We outline the principal advantages concerning the regularization of the inverse problems in an appropriate norm and give the particular regularization properties of the truncated Newton-CGNE method.

To enforce the uniqueness of the solution of the inverse problem, usually, the *minimum-norm* solution is introduced which comprises the best approximation of the solution with minimum norm [17]. In many applications, however, and particularly in case of nonlinear inverse problems, the weighted minimum-norm solution is introduced by putting additional conditions on the minimum-norm solution in terms of the chosen norm space. For example, the norm $\|\cdot\|_{\mathcal{X}}$ in eq. (3.3a) introduces additional smoothness into the solution. This type of regularization also contributes to convergence (cf. Chapter 8.1 in [17]). The availability of good initial guesses for the optimization is essential for nonlinear inverse problems (cf. Chapter 11.1 in [17], [23, 38]).

The regularization of the employed solution method is introduced via the fixed spatial and temporal discretization and proper stopping criteria for the Newton (outer) and the CGNE (inner) iterations as follows. As mentioned, before entering the optimization strategy, the coefficient inverse problem is linearized to benefit from the regularization properties of the CGNE method. Therefore, the Newton iteration can only make further progress if the Taylor remainder of the linearization is *not* dominant. Hence, the CGNE iterations should run as long as the linearized equation still provides the additional information. In particular, the stopping rule of the CGNE iteration (3.13) with $0 < \eta < 1$ prevents the unwanted dominance of the Taylor remainder. Besides, the linearized problem needs to be regularized. It is well known that the regularizing effect of the CGNE method comes from the early termination of iterations [21]. Hence, depending on the value of η, the CGNE iteration should result in a *regularized* Newton update $\hat{x}^n$ (cf. eq. (3.5)). The larger $\eta \uparrow 1$, the less CGNE iterations are allowed, and the resulting Newton updates tend to be *overregularized*; the smaller the parameter $\eta \downarrow 0$, the more CGNE iterations have to be carried out. However, one should be careful not to allow a too large number of CGNE iterations which would result in *underregularized* Newton updates. Consequently, the stopping rule (3.13) should control the Taylor remainder *and* the ill-posedness of the linearized problem simultaneously: in the vicinity of the solution (e. g., in case of sufficiently good initial values) relatively small values of η are advantageous, since in this case more benefit from the quadratic Newton approximation can be gained; however, during the Newton iterations it might be necessary to increase the value of η to handle the ill-posedness of the linearized problem by early termination of the CGNE iteration (cf. [38, 39] for more details on the choice of η).

We apply regularization by iteration to determine the optimal number of Newton iterations. The discrepancy principle [21] is employed as parameter choice rule, i. e. the iterations are stopped once the noise level in the data has been reached. Details on the stopping criterion for the Newton iteration and on the choice of the parameter η are given for the case study in Section 4.

3.2. Correction of parameters and model selection. In this section, we turn to the parameter correction step of the overall correction procedure (cf. Fig. 3.1). We formulate this problem as a nonlinear least-squares minimization problem with the objective function

$$J_2(f_{w,i}(u, \mathbf{x}, t, \theta_i)) = \frac{1}{2} \left\| (a^\star(\mathbf{x}, t) - a_{\mathrm{mol}}) - f_{w,i}(u, \mathbf{x}, t, \theta_i) \right\|_{\mathcal{Y}}^2 , \quad i \in I_s , \qquad (3.14)$$

with I_s being the subset of preselected transport model structures (cf. (2.4)).

To measure the model quality and to select a "best-performing" transport model $f_w^\star(\,\cdot\,,\theta)$, we use Akaike's minimum information theoretic criterion AIC [1]

$$AIC(f_{w,i}(\,\cdot\,,\theta_i^\star)) = 2p_i + M[\ln(2\pi RSS(\theta_i^\star)/M) + 1]\,,\ i \in I_s\,. \qquad (3.15)$$

It incorporates the residual sum of squares $RSS(\theta_i^\star)$ of each estimated candidate model $f_{w,i}(\,\cdot\,,\theta_i^\star)\,, i \in I_s$, at the corresponding optimum $\theta_i^\star$, the number of candidate model parameters p_i and the number M of measurements employed for the solution of the estimation problems. The model with *minimum AIC* is selected. Consequently, this criterion chooses models with the best fit of the data and hence high precision in the parameters and at the same time penalizes the number of model parameters.

Model identification for transport coefficients. We summarize the methodology of the identification of a "best-performing" transport model in pseudo-code notation as follows:

Incremental model identification: Decompose the problem of identifying the transport model structure $f_w(\,\cdot\,,\theta)\,,\theta \in \mathbb{R}^p$, in balance equation (1.1) into several easier to handle subproblems (cf. Fig. 2.1):

(B) Reconstruct the source term $F_w(\mathbf{x},t)$ using the transient, distributed measurement data u_m (cf. eq. (2.1));
  $\rightsquigarrow$ **return** $F_w^\star(\mathbf{x},t)$;

(BF) Reconstruct the wavy transport coefficient $a_w(\mathbf{x},t)$ using the measurement data u_m by employing the source estimate $F_w^\star(\mathbf{x},t)$ as inferential measurement data (cf. eqs. (2.2), (2.3));
  $\rightsquigarrow$ **return** $a_w^\star(\mathbf{x},t)$;

(BFT) Use $a_w^\star(\mathbf{x},t)$ to
  (a) propose a list $I = \{1, 2, \ldots, m\}$ of candidate transport model structures $f_{w,i}(\,\cdot\,,\theta_i)\,, i \in I$;
  (b) estimate the parameters of each proposed transport model structure.
  Compute the quality measures of each candidate model and eliminate candidate models of poor quality resulting in a subset $I_s \subseteq I$ of reasonable candidates (cf. eq. (3.15));
  $\rightsquigarrow$ **return** $I_s \subseteq I\,; f_{w,i}(\,\cdot\,,\theta_i)$ and $\theta_i^\star\,, i \in I_s$.

Model correction: Decompose the correction of the model structure in eq. (1.1) (cf. Fig. 3.1):

(1) Correct the transport coefficient $a(\mathbf{x},t)$ (cf. eq. (1.1c)) by combining the steps B and BF into one (cf. eq. (3.4) subject to eqs. (3.1), (1.1b)). Use $a_w^\star(\mathbf{x},t)$ (cf. eq. (3.2)) as good initial value for the solution of this inverse problem.
  $\rightsquigarrow$ **return** $a^\star(\mathbf{x},t)$;

(2) Use $(a^\star(\mathbf{x},t) - a_{\mathrm{mol}})$ to
  − estimate the parameters in the transport coefficient model structures $f_{w,i}(\,\cdot\,,\theta_i)\,,\ i \in I_s$, using $\theta_i^\star, i \in I_s$ as good initial values;
  − optionally propose new candidate transport coefficient model structures.
  Recompute the quality measures of the candidate models considered and select the best candidate (cf. eq. (3.15));
  $\rightsquigarrow$ **return** $f_w^\star(\,\cdot\,,\theta)$ and $\theta^\star$.

11

REMARK 1. *For a known velocity* $\mathbf{w}(\mathbf{x}, t)$ *and molecular transport coefficient* a_{mol}

(i) *it suffices to reconstruct the source term* $F_w^\star(\mathbf{x}, t)$ *in B and the wavy transport coefficient* $a_w^\star(\mathbf{x}, t)$ *in BF only once in the incremental model identification;*

(ii) *the reconstruction of* $a^\star(\mathbf{x}, t)$ *in the model correction step has to be carried out only once;*

(iii) *the (repeated) estimation and the selection of suitable candidate models for the transport coefficient is of* algebraic nature *in both, the incremental model identification and model correction steps.*

4. Numerical case study. In this section, we consider an illustrative case study describing energy transport in laminar wavy film flows, without incorporating a priori knowledge on the unknown transport. The simulation experiment considered corresponds to the so-called *flat-film* model problem [44], which has been introduced for the design of technical systems [9, 44].

The flat-film model reduces the complexity of the real film by mapping the three-dimensional wavy, time-varying domain $\Omega_W(t)$ corresponding to the liquid phase of the film flow to a three-dimensional time-invariant *waveless* domain $\Omega := (0, L_x) \times (0, L_y) \times (0, L_z) \subset \mathbb{R}^3$. This reduction is compensated by the introduction of space- and time-dependent *effective* transport coefficients [9, 15, 44], which capture wave-induced transport enhancing effects in this flat film geometry.

In our previous work [29], we applied the incremental identification method (cf. Fig. 2.1) to the flat film model problem to identify the energy transport model for the effective energy transport coefficient, the *effective thermal diffusivity* $a_{\text{eff}}(\mathbf{x}, t)$ assuming an a priori known model structure. In this section, we reconsider the same test problem but relax this assumption. We now eliminate model structures of poor quality measure from a list of proposed candidate models (cf. (2.4)) to identify a subset of reasonable model structures and apply the proposed correction procedure (cf. Fig. 3.1). We briefly describe the flat film example model problem of the simulation experiment.

We consider the convection-diffusion system describing energy transport in a single component fluid of density ρ on the flat domain $\Omega = (0, 1)^3 [\text{mm}^3]$, with boundary Γ consisting of the inflow $\Gamma_{in} = \{x = 0\}$, the wall $\Gamma_{wall} = \{y = 0\}$, the outflow $\Gamma_{out} = \{x = 1\}$ and the remaining $\Gamma_r = \{y = 1\} \cup \{z = 0\} \cup \{z = 1\}$ boundaries. Here, x corresponds to the flow direction of the falling film, y is the direction in the film thickness and z the direction along the film width (cf. Fig. 4.1).

In this case, the state variable in eq. (1.1) is $u(\mathbf{x}, t) = cT(\mathbf{x}, t)$ with temperature $T(\mathbf{x}, t)$ and heat capacity c of the fluid. We assume ρ and c to be constants. The transport model $f_w(\cdot, \theta)$, $\theta \in \mathbb{R}^p$, models the enhanced wavy part of the effective thermal diffusivity $a(\mathbf{x}, t) = a_{\text{eff}}(\mathbf{x}, t)$ in eq. (1.1). The velocity $\mathbf{w}(\mathbf{x}, t)$ is given by
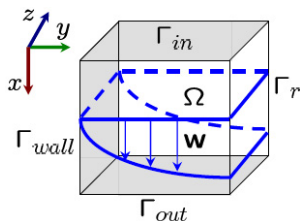

FIG. 4.1. *The flat film model problem.* (Γ_r *consists of the non-shaded faces.*)

the one-dimensional Nusselt profile $\mathbf{w}(\mathbf{x},t) = 4.2857(2y - y^2)$ [37][1]. The initial and boundary conditions are

$$T(\mathbf{x}, 0) = 15\,[^\circ C],\ \mathbf{x} \in \Omega\,, \tag{4.1a}$$

$$T_{in}(\mathbf{x}, t) = (-30yt + 15)\,[^\circ C],\ (\mathbf{x}, t) \in \Gamma_{in} \times [t_0, t_f]\,, \tag{4.1b}$$

$$T_{wall}(\mathbf{x}, t) = \left(100\left(1 - \cos(\pi\tfrac{x}{2})\right)t + 15\right)\,[^\circ C],\ (\mathbf{x}, t) \in \Gamma_{wall} \times [t_0, t_f]. \tag{4.1c}$$

At the Neumann boundaries Γ_{out} and Γ_r, a zero flux condition is used, i.e.,

$$\frac{\partial T}{\partial \mathbf{n}}(\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in (\Gamma_{out} \cup \Gamma_r) \times [t_0, t_f]. \tag{4.1d}$$

The effective transport coefficient a_{eff} (cf. eq. (1.1c)) in this simulation experiment comprises the sum of a constant *molecular thermal diffusivity* $a_{\text{mol}} = 0.35[\frac{\text{mm}^2}{\text{s}}]$ and the exact wavy transport model

$$f_w^{ex}(\mathbf{x}, t, \theta) = 5(\vartheta_1 + \vartheta_2 y \sin(\vartheta_3 x + \vartheta_4 t) + \vartheta_5 xy + \vartheta_6 xyz)\,, \tag{4.2a}$$
$$(x, y, z, t) \in \Omega \times [t_0, t_f]\,,$$

with the exact parameter values $\theta^{ex} \in \mathbb{R}^6$

$$\theta^{ex} = (\vartheta_1^{ex}, \ldots, \vartheta_6^{ex})^T = (1.1, 1.0, \pi, 0.02, 0.2, 0.02)^T\,, \tag{4.2b}$$

which corresponds to the wavy transport coefficient $a_w(\mathbf{x}, t)$ (cf. eq. (2.2), Fig. 2.1). Motivated by physical considerations, we have chosen a sinusoidal pattern over the space coordinate in the flow direction of the falling film (i.e., the x-direction). The time-dependency is introduced such that the waves travel along the x-direction and propagate along the y- and z-directions, with a larger gradient in the y-direction (film thickness) and with a relatively small gradient in the z-direction (film width).

High-quality temperature simulation data are generated by solving the linear problem (1.1) with the exact transport model (4.2) on a uniform fine grid with the spatial discretization consisting of $48 \times 48 \times 38$ intervals in x, y, and z directions, respectively. This yields a space discretization with $89,856$ unknowns and $525,312$ tetrahedra. As measurement data, we use the temperature data T_m on the coarser grid with $24 \times 24 \times 19$ intervals in x, y, and z directions, respectively, to avoid the so-called inverse crime [28]. For the time discretization we use the implicit Euler scheme with time step $\tau = 0.01$ s and apply 50 time steps starting from the initial time $t_0 = 0$ s (i.e., $t_f = 0.5$ s). This results in the number of measurements $M = 637,500$.

Furthermore, we analyze the entire methodology both for noise-free and noisy measurements by artificially perturbing the noise-free temperature T_m with measurement error ω the values of which are generated from a zero mean normal distribution with variance one. Hence, we compute the perturbed temperature

$$\tilde{T}_m = T_m + \sigma\omega\,, \tag{4.3}$$

with the standard deviation $\sigma = 0.1$ of the measurement error.

We now reconstruct the wavy thermal diffusivity (cf. eq. (2.3)) by the incremental identification method. Fig. 4.2 shows the wavy thermal diffusivity resulting from the second step BF at time instance $t = 0.01\,s$ and constant $z = 0.5\,mm$ reconstructed

---

[1] Alternatively, a more precise solution to the film Navier-Stokes equations can be used.

i	$f_{w,i}(\mathbf{x},t,\theta), \theta \in \mathbb{R}^p, i \in I := \{1,\ldots,6\}$	$AIC/10^6$ noise-free	$AIC/10^6$ noisy
1	$5(\vartheta_1 + \vartheta_2 y \sin(\vartheta_3 x + \vartheta_4 t) + \vartheta_5 xy + \vartheta_6 xyz)$	-0.194	0.4272
2	$5(\vartheta_1 + \vartheta_2 y \sin(\vartheta_3 x + \vartheta_4 t))$	-0.112	0.6467
3	$5(\vartheta_1 + \vartheta_2 y \sin(\vartheta_3 x + \vartheta_4 t) + \vartheta_5 xy)$	-0.184	0.4289
4	$5(\vartheta_1 + \vartheta_3 x^2 + \vartheta_4 t + \vartheta_5 xy)$	1.785	1.9362
5	$5(\vartheta_1 + \sin(\vartheta_3 x + \vartheta_4 t))$	2.210	2.2432
6	$5(\vartheta_1 + \cos(\vartheta_3 x + \vartheta_4 t))$	2.334	2.3892

TABLE 4.1

List of candidate transport models $f_{w,i}(\cdot,\theta), i \in I$ with the corresponding AIC values both for noise-free and noisy measurement data.

in [29] (no reconstruction has been carried out at the time instance $t = 0$ s, since $T(\mathbf{x},0) = \text{const.}$ (cf. eq. (4.1)). As it can be seen, the chosen constant initial guess $a_w^{\text{init},0.01\,s} = 5.5[\frac{\text{mm}^2}{\text{s}}]$ is very different from the true solution at $t = 0.01\,s$. Since the reconstruction of the wavy transport coefficient in eq. (2.2) is decoupled in time, the obtained optimal solution at time instance $t = 0.01\,s$ serves as a good initial value for the efficient optimization at later times $t > 0.01\,s$. Hence, the incremental identification method *generates* good initial values for optimization, such that no a priori knowledge is necessary for the reconstruction of the wavy thermal diffusivity $a_w^{\star,t}(\mathbf{x})$ at *all* times $t \in [t_0, t_f]$ [29]. By assembling these estimates, $a_w^\star(\mathbf{x},t)$ is obtained as

$$a_w^\star(\mathbf{x},t) = a_w^{\star,t}(\mathbf{x}), \forall t \in [0, 0.5\,s]. \tag{4.4}$$

We explore the shape of this reconstructed wavy transport coefficient $a_w^\star(\mathbf{x},t)$ (cf. Fig. 4.2) to *manually* develop a list I of model structures $f_{w,i}(\cdot,\theta_i), i \in I$. The results in Fig. 4.2 suggest that a reasonable model structure should incorporate a trigonometric function in the flow direction (i.e., the x-direction) with a period changing in time. The side and the surface views of this function (not shown) reveal
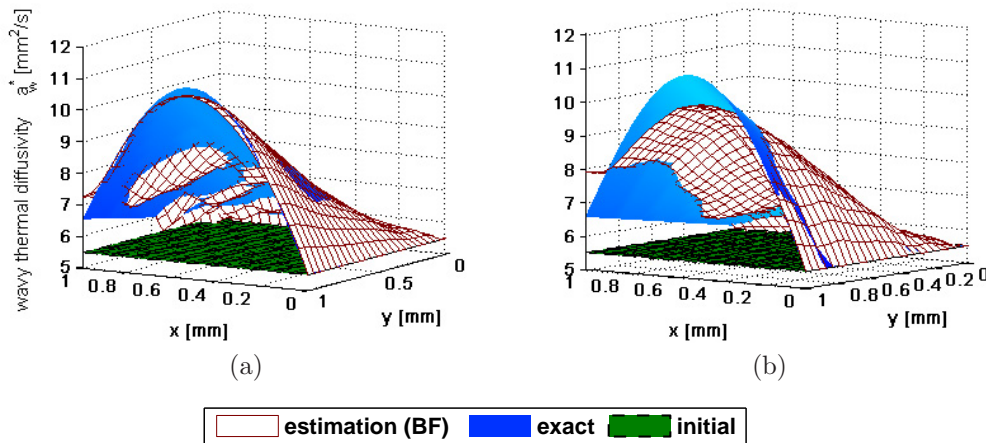

(a)   (b)

estimation (BF)   exact   initial

FIG. 4.2. *Estimated wavy thermal diffusivity $a_w^{\star,t}$ at $t = 0.01$ s for constant $z = 0.5$ mm for (a) noise-free and (b) noisy ($\sigma = 0.1$) measurements. Adopted from [29].*

14

a slight nonlinear shape change. Based on these observations, we propose a set of six candidate models $I := \{1, \ldots, 6\}$ as listed in Tab. 4.1. Note that the structure of the candidate models $f_{w,i}$, $i \in \{4, 5, 6\}$ does not match the correct structure. Obviously, the choice of model candidates requires intuition and physical insight. However, this choice can be efficiently guided by the results of the transport coefficient estimation step of the incremental identification method.

In what follows, we first enter the third step BFT of the incremental identification method (cf. Fig. 2.1) to estimate model parameters for each candidate model by using $a_w^\star(\mathbf{x}, t)$ as inferential measurement data. After choosing the corresponding subset $I_s \subseteq I$ of reasonable transport models (cf. (2.4)), we turn to the model correction procedure.

Noise-free measurements will be considered first. Subsequently, the influence of artificially perturbed measurements will be studied.

4.1. Identification with noise-free measurements. We fit each of the candidate models from Tab. 4.1 to the function $a_w^\star(\mathbf{x}, t)$ (cf. eq. (4.4), Fig. 4.2). In particular, we choose 500 different random initial values for each candidate model. From the set of the resulting *local* solutions the one showing the lowest residual sum of squares is selected. Note, that global optimization (e.g. , BARON, [42]) could replace this multi-start strategy for solving the non-convex optimization problem. The resulting AIC values (cf. eq. (3.15)) for each candidate model at the corresponding optimum are given in Tab. 4.1. The candidate models $f_{w,i}$, $i \in \{4, 5, 6\}$, with wrong model structure show much larger AIC values than the other candidates and are thus discarded. The remaining candidate models show comparable quality. Note, that candidate models $f_{w,i}$, $i \in \{2, 3\}$ constitute a simplification of the exact model structure (4.2a), whereas $f_{w,1}$ matches the exact model structure. These observations reveal that the estimate $a_w^\star(\mathbf{x}, t)$ is of sufficient quality to distinguish between candidate model structures. Obviously, in a real situation, the true model structure would not be among the candidates. The case study convincingly illustrates however, that the AIC is able to adequately rank model candidates.

The subset $I_s \subseteq I$ of reasonable model structures (cf. eq. (2.4)) in this case is

$$I_s = \{1, 2, 3\}. \tag{4.5}$$

Note, that if the model selection step would have been carried out directly after completing all three steps of the incremental identification method, the model $f_{w,1}^\star(\cdot, \theta_1) := f_{w,1}(\cdot, \theta_1)$ would have been chosen as the solution, since it exhibits the smallest AIC value among the candidates $f_{w,i}$, $i \in I_s$ (cf. Tab. 4.1). Consequently, in the scenario chosen, the solution would comprise the *exact* model (4.2a) with the optimal parameter vector

$$\theta_1^\star = (\vartheta_1^\star, \ldots, \vartheta_6^\star)^T = (1.121, 0.917, 3.417, -0.016, 0.512, 0.051)^T. \tag{4.6}$$

The comparison with the exact parameter vector (4.2b) reveals reasonable estimation quality, but shows deviations. We now show that the error can be significantly reduced if the correction step is carried out.

In the first step of the correction procedure (cf. Fig. 3.1), we employ (cf. eq. (3.2))

$$a_{\text{eff}}^{\text{init}} = a_{\text{mol}} + a_w^\star(\mathbf{x}, t), \quad a_{\text{mol}} = 0.35[\frac{\text{mm}^2}{\text{s}}], \tag{4.7}$$

15

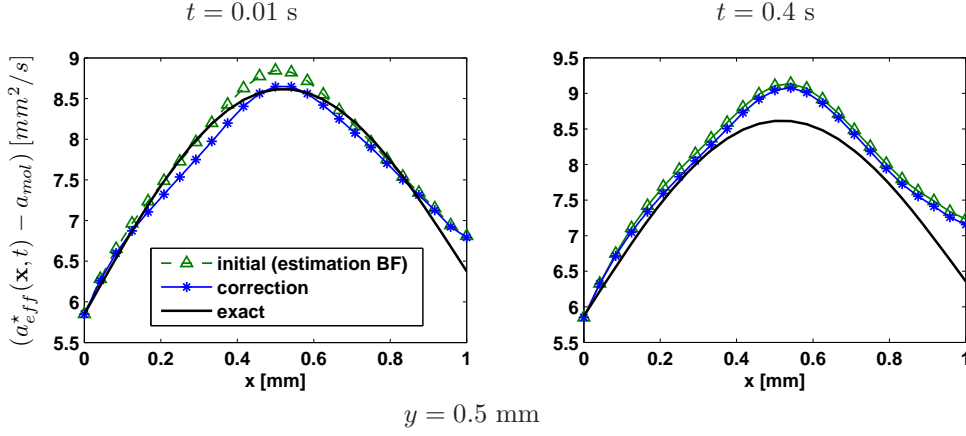FIG. 4.3. *Initial $(a_{eff}^{init} - a_{mol})$ (cf. (4.7)), exact and corrected wavy thermal diffusivity $(a_{eff}^{\star}(\mathbf{x}, t) - a_{mol})$ at different times and constant $z = 0.5$ mm.*

with $a_w^{\star}(\mathbf{x}, t)$ from eq. (4.4) as initial guess for the solution of the transient, nonlinear inverse coefficient problem (3.4) subject to eqs. (3.1), (1.1b) to reconstruct the effective thermal diffusivity $a_{\text{eff}}^{\star}(\mathbf{x}, t)$. Some results are presented in Fig. 4.3 for selected times $t \in \{0.01, 0.4\}$ s. The truncated Newton-CGNE method is stopped after 5 Newton iterations when the (Euclidean) norm of the Newton update $\hat{x}^n$ (cf. eq. (3.5)) becomes smaller than the chosen tolerance $\epsilon = 10^{-5}$. For the stopping rule (3.13) in the CGNE iteration we use $\eta = 0.5$ at the beginning of the Newton iteration and increase it to $\eta = 0.6$ in the last iteration. Our experiments reveal that this stopping rule needs to be tuned to efficiently handle the nonlinear Taylor remainder and the ill-posedness of the inverse problem (cf. discussion in section 3.1.4). One can see (cf. Fig. 4.3) that the initial values have been slightly changed. This already suffices for a significant correction in the transport model as will be shown below.

REMARK 2. *Recall, that in the solution of the stationary coefficient inverse problems in the second step of the incremental identification procedure (cf. eq. (2.2)), we choose a constant initial guess $a_w^{init,0.01\,s} = 5.5[\frac{mm^2}{s}]$ at time instance $t = 0.01\,s$ (cf. Fig. 4.2). By applying a similar initial guess, i. e., $a_{eff}^{init} = (a_{mol} + 5.5)[\frac{mm^2}{s}], a_{mol} = 0.35[\frac{mm^2}{s}]$ for the solution of the transient inverse coefficient problem (cf. eq. (3.4) subject to eqs. (3.1), (1.1b)), we were not able to find a value of the parameter $0 < \eta < 1$ in the stopping rule of the CGNE iteration (cf. eq. (3.13)) for the algorithm to converge. Thus, as theoretically predicted, good initial values are a necessity to successfully solve this nonlinear inverse problem (cf. Sections 3.1.2 and 3.1.4).*

To obtain the "best-performing" transport coefficient model structure, we solve the nonlinear least-squares problem (3.14) with the corrected wavy thermal diffusivity $(a_{\text{eff}}^{\star}(\mathbf{x}, t) - a_{\text{mol}})$ (cf. Fig. 4.3) and candidate models $f_{w,i}(\mathbf{x}, t, \theta_i), i \in I_s = \{1, 2, 3\}$, (cf. Tab. 4.1 and eq. (4.5)) in the second step of the correction procedure. In addition to the optimal values available from the BFT step of the incremental identification method, we again employ a multi-start strategy using 500 random initial values and select the solution corresponding to the lowest residual sum of squares. The resulting
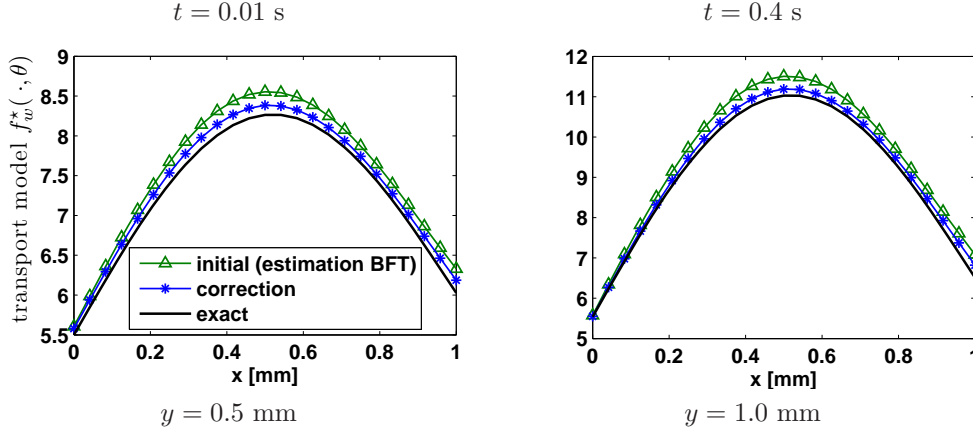
16

FIG. 4.4. *Initial, exact and "best-performing" transport model* $f_w^\star(\mathbf{x}, t, \theta) = f_{w,1}(\mathbf{x}, t, \theta_1)$ *at different times and constant* $z = 0.5\,mm$.

AIC values for $f_{w,i}$, $i \in I_s$ at their corrected optima are

$$\frac{AIC(f_{w,i}(\mathbf{x}, t, \theta_i^\star))}{10^4} = \{-2.6340, 0.3915, -2.0248\}, i \in I_s = \{1, 2, 3\}. \qquad (4.8)$$

As the "best-performing" model, we select the one with the smallest AIC value among the structures in I_s in eq. (4.8). This results in the *same* correct transport model $f_w^\star(\cdot, \theta) := f_{w,1}(\cdot, \theta_1)$ in this scenario (cf. Tab. 4.1). Fig. 4.4 shows the results of the correction procedure. The corrected model structure is identified as

$$f_w^\star(\mathbf{x}, t, \theta) = f_{w,1}(\mathbf{x}, t, \theta_1) = 5(\vartheta_1 + \vartheta_2 y \sin(\vartheta_3 x + \vartheta_4 t) + \vartheta_5 xy + \vartheta_6 xyz), \qquad (4.9a)$$

with estimated parameters

$$\theta_1^\star = (\vartheta_1^\star, \ldots, \vartheta_6^\star)^T = (1.116, 0.923, 3.315, -0.015, 0.378, 0.045)^T. \qquad (4.9b)$$

The comparison of these parameters with the exact parameters (4.2b) shows that the estimation quality in *all* parameters has been improved over those in (4.6). However, the corrected estimate (cf. Fig. 4.3) still shows deviations especially in the parameters $\vartheta_4^\star, \vartheta_5^\star$ and $\vartheta_6^\star$. One reason is due to the properties of the simulated transport process. In the considered unit cube geometry, the first two summands (i.e., the constant term and the trigonometric part of the model structure) of the transport coefficient model (4.2) depict the dominant part, whereas the remaining summands have comparably low influence and are thus hard to identify from any type of experiment. A second reason is due to the experiment. Its short duration $(t_f - t_0) = 0.5\,s$ results only in a very slight shift in time that can not be correctly recovered. Still, the reduction of the propagation errors resulting from the incremental decomposition can be clearly identified, e.g., by comparing the AIC values of the model alternatives (cf. eq. (4.8)) with each other. Namely, now the AIC value of the candidate model $f_{w,2}$ is much larger than that of model $f_{w,3}$ because of its simpler structure.

The results of the parameter correction step are summarized for all models in I_s (cf. (4.5)) before and after correction in Tab. 4.2. One can see that the corrected parameter values have moved towards the exact solution (4.2b), though some deviations are still present.

17

incremental identification method	correction procedure
noise-free data	
$\theta_1^\star = (1.121, 0.917, 3.417, -0.016, 0.512, 0.051)^T$	$\theta_1^\star = (1.116, 0.923, 3.315, -0.015, 0.378, 0.045)^T$
$\theta_2^\star = (1.125, 1.173, 2.941, 0.045)^T$	$\theta_2^\star = (1.117, 1.118, 2.951, 0.049)^T$
$\theta_3^\star = (1.121, 0.918, 3.416, -0.016, 0.538)^T$	$\theta_3^\star = (1.116, 0.924, 3.315, -0.015, 0.401)^T$
noisy data	
$\theta_1^\star = (1.140, 0.803, 4.077, -0.112, 0.989, 0.0336)^T$	$\theta_1^\star = (1.104, 0.723, 4.069, -0.149, 0.826, 0.186)^T$
$\theta_2^\star = (1.150, 1.203, 2.975, 0.123)^T$	$\theta_2^\star = (1.108, 1.098, 2.926, 0.176)^T$
$\theta_3^\star = (1.140, 0.804, 4.077, -0.112, 1.006)^T$	$\theta_3^\star = (1.104, 0.723, 4.069, -0.149, 0.919)^T$

TABLE 4.2

Parameter values after the incremental identification and after the correction.

To show that the proposed correction procedure is essential for handling error propagation through the steps of the incremental identification method, we solve *parameter estimation* problem directly using balance equation (1.1) and assuming that the candidate model structure $f_{w,1}(\,\cdot\,,\theta_1)$ resulting from the incremental identification method were the correct one. Moreover, we use the optimal parameter values (4.6) resulting from the incremental identification as initial values. As solution strategy we employ the so-called Landweber-iteration method [17], a prototype of the steepest-descent method for well-posed problems (cf. [36]), because this method is known not to be as sensitive with respect to the initial parameter values as the Newton type methods. The estimation result is given in Tab. 4.3 (a), left.

The comparison of this estimate with the one obtained after the correction procedure in Tab. 4.2 reveals that the latter is more favorable in terms of error propagation. One reason for this is the strong nonlinearity and the *nonconvexity* of the underlying direct parameter estimation problem. More precisely, the step-lengths in the Landweber iteration are so small that only tiny parameters updates result. A second reason is due to the *uncertainty* with respect to the model structure $f_{w,1}(\,\cdot\,,\theta_1)$ analyzed in the preceding paragraph. Using the parameter estimate $\theta_1^\star$ from the Tab. 4.3 (a), left, a sensitivity analysis of this model (cf. [12, 30]) has been carried out. It revealed that all six parameters are *identifiable*. However, the parameters $\vartheta_3, \dots, \vartheta_6$ are difficult to estimate because of (i) the very small sensitivity of the model with respect to these parameters (in particular, with respect to ϑ_6) and (ii) the relatively high correlation of the parameters ϑ_3, ϑ_4 and ϑ_5 with ϑ_2, as well as of ϑ_5 with ϑ_6. Exemplarily, if we fix the values of parameters $\vartheta_3, \dots, \vartheta_6$ to their corresponding *exact* values (cf. (4.2b)) we obtain the estimate $\theta_1^\star = (\vartheta_1, \vartheta_2)^T$ as given in the Tab. 4.3 (b), left. These values are closer to their exact counterparts (cf. (4.2b)). Also note, that this estimate compares well with the corresponding estimate for the first two parameters obtained after the correction procedure (cf. Tab. 4.2). We remark that the convergence can

noise-free data	noisy data
$\theta_1^\star = (1.123, 0.898, 3.408, -0.034, 0.512, 0.051)^T$	$\theta_1^\star = (1.209, 0.748, 4.048, -0.139, 1.009, 0.043)^T$
$\theta_1^\star = (1.099, 0.982)^T$	$\theta_1^\star = (1.047, 0.839)^T$

TABLE 4.3

Parameter values after direct parameter estimation using eq. (1.1) the model structure $f_{w,1}(\,\cdot\,,\theta_1)$ and the initial parameter estimate (4.6) both resulting from the incremental identification method upper row for the full parameter set and lower row for the reduced parameter set with remaining parameters fixed at their exact values (cf. (4.2b))

18

be achieved for the reduced parameter vector $\theta_1^\star = (\vartheta_1, \vartheta_2)^T$ even for initial value, far from the exact solution.

We emphasize that the solution of the direct parameter estimation problem with balance equation (1.1) and model structure $f_{w,1}(\,\cdot\,,\theta_1)$ using the same initial values for the parameters as the ones we employed in the third step of the incremental identification method (recall, these were randomly chosen vectors), *failed* to converge.

Hence, at least for the simulation example studied in this paper we can conclude that the incremental identification approach is able to deliver valuable results, even when the level of detail of the investigated model structure is high and when poor initial estimates are available, and thus represents a sound, rigorous alternative to the so-called *simultaneous* (i. e. direct) identification approach. The disadvantageous error propagation in the incremental identification approach can be effectively treated with the correction procedure proposed in this paper. Because of the strong correlation present, some of the parameter values could not be corrected well. However, the correction procedure is still able to provide a significant correction in the significant parameters.

4.2. Identification with noisy measurements. In this section, we carry out the estimation procedure with perturbed temperature data as in eq. (4.3) with $\sigma = 0.1$.

The AIC values of the candidate models resulting from the multi-start strategy are listed in the last column of Tab. 4.1. In the presence of noise, the AIC values are significantly larger for all candidate models. Similar to the noise-free case, the candidate models $f_{w,i}$, $i \in \{4,5,6\}$, with an incorrect model structure show poor quality, whereas all the remaining candidates are of comparable quality. Hence, a subset $I_s \subseteq I$ of reasonable model structures results again in

$$I_s = \{1, 2, 3\}. \tag{4.10}$$

The model of best quality obtained directly from the incremental identification method is again the *correct* model $f_{w,1}^\star(\,\cdot\,,\theta_1) := f_{w,1}(\,\cdot\,,\theta_1)$ (cf. AIC values in Tab. 4.1). The corresponding optimal parameter vector is

$$\theta_1^\star = (\vartheta_1^\star, \ldots, \vartheta_6^\star)^T = (1.140, 0.803, 4.077, -0.112, 0.989, 0.0336)^T. \tag{4.11}$$

The comparison with the result obtained using noise-free data (cf. eq. (4.6)) and with the exact parameter vector (cf. eq. (4.2b)) shows that the deviation in most of the parameters is now more significant. The reason for this is the larger error in the wavy transport coefficient estimate $a_w^\star(\mathbf{x},t)$ (cf. Fig. 4.2 (b)) as compared to the one without the presence of the noise in the data (cf. Fig. 4.2 (a)). However, despite the measurement noise, the same model structure as in the noise-free case can be recovered. This result shows in fact, how difficult the solution of such ill-posed identification problems is if (inevitable) noise is present in the measurements. Though in the considered case the choice of the best model structure is not sensitive to noise, the quality of the estimated parameters deteriorates significantly despite the favorable situation that the model structure is correct.

In case of noisy data, the solution of the inverse coefficient problem (3.4) subject to eqs. (3.1), (1.1b) has to be appropriately regularized to prevent unwanted resolution of measurement noise. We employ again the optimal, *regularized* wavy thermal diffusivity resulting from the incremental identification method (cf. eq. (3.2) and Fig. 4.2) as initial value. To obtain the regularized Newton updates $\hat{x}^n$ (cf. eq. (3.5)),
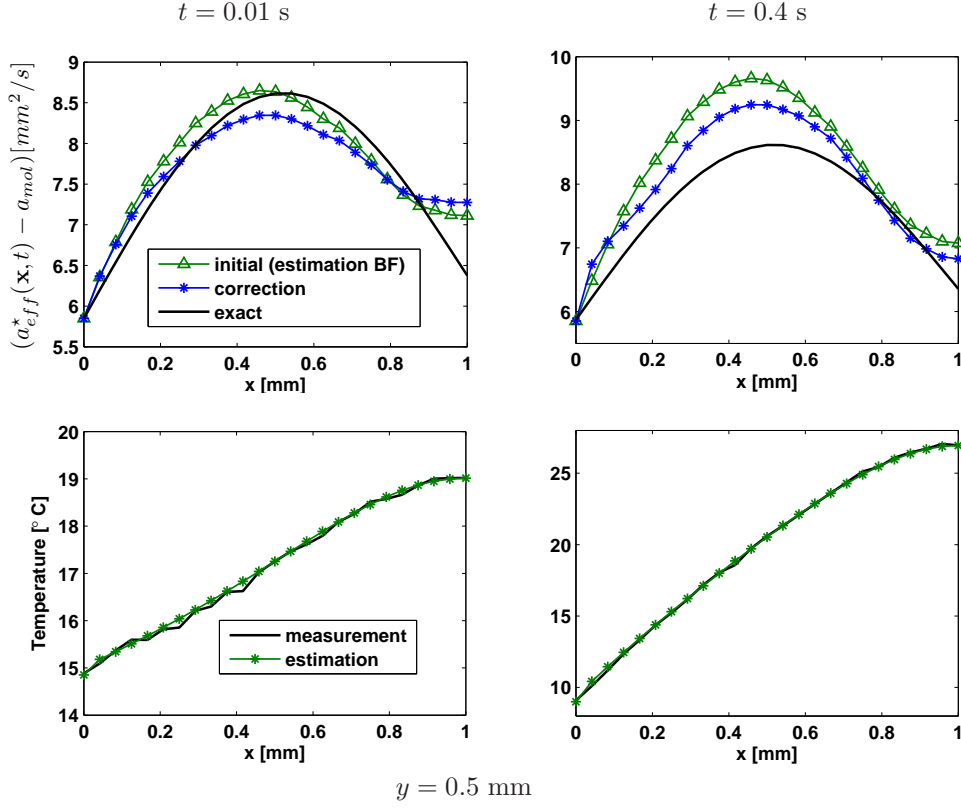
FIG. 4.5. *Regularized initial* $(a_{eff}^{init} - a_{mol})$ *(cf. (4.7)), exact and corrected wavy thermal diffu-sivity* $(a_{eff}^{\star}(\mathbf{x}, t) - a_{mol})$ *(upper row) and the agreement in the temperatures (lower row) at different times and constant* $z = 0.5$ *mm for noisy data* $(\sigma = 0.1)$.

we choose $\eta = 0.4$ in the stopping rule (3.13) for all Newton iterations, and constrain the maximal number of CGNE iterations to $k_{\max} = 50$. Since good initial values are available here, smaller values of η are more appropriate. However, now the number of CGNE iterations has to be constrained to handle the ill-posedness of the problem. This last statement applies especially to the noisy case (cf. discussion in section 3.1.4).

For the determination of the number of Newton iterations, we employ the discrepancy principle [21], which suggests to stop the iterations once the condition

$$J_1(a^n) \leq \kappa \sigma^2 , \ \kappa > 1 , \tag{4.12}$$

is satisfied for the residual norm (3.4) with parameter κ and noise level σ. The parameter κ has to be chosen close to 1, to allow for a small residual norm. On the other hand, the tolerance η (cf. eq. (3.13)) for the inexact Newton step should be suffciently small to benefit from the quadratic Newton approximation. Therefore, these two parameters can not be chosen independently [23]. Theory suggests the quite restrictive condition (for details cf. [23])

$$\kappa > 2/\eta^2 > 2 . \tag{4.13}$$

However, in case of good initial values, and under the imposed additional smoothness assumption on the inverse problem solution $a_{\text{eff}} \in \mathcal{L}^2(t_0, t_f; \mathcal{H}^1(\Omega))$, this condition

can be weakened [23, 41] such that smaller values of κ are allowed. In general, an appropriate value of κ is highly problem-dependent.

For the studied example, the value of $\kappa = 4.6$ has shown to be the best choice. This value is much smaller than the value suggested by eq. (4.13) for $\eta = 0.4$. For this choice of κ the discrepancy principle (4.12) results in 4 Newton iterations. The corresponding regularized solution is presented in Fig. 4.5 (upper row). The agreement in the measured and estimated temperatures is presented, too (lower row). It can be seen, that because of the time dependence of the temperature (note the different scales in the shown temperatures for the different time instances $t \in \{0.01, 0.4\}\,s$) the *same* noise level in the data influences the quality of the solution remarkably: The higher the slope in the temperature (i. e., the lower the presence of the noise in the data) the better the correction in the solution ($t = 0.4\,s$). This is explained by the extreme ill-posedness of the inverse coefficient problem (cf., e. g., [17]).

In the second step of the correction procedure, we consider three candidate models $f_{w,i}$, $i \in I_s = \{1, 2, 3\}$, (cf. eq. (4.10) and Tab. 4.1) for the corrected wavy thermal diffusivity $(a_{\mathrm{eff}}^\star(\mathbf{x}, t) - a_{\mathrm{mol}})$ (cf. Fig. 4.5) and employ it as inferential measurement data to solve the least-squares problems (3.14) for each of these candidate models. Again, besides the corresponding optimal values of parameters $\theta_i^\star$ available from the incremental identification procedure, other 500 randomly chosen initial values are used. The resulting AIC values for each of these candidates at their corrected optima result in

$$\frac{AIC(f_{w,i}(\mathbf{x}, t, \theta^\star))}{10^6} = \{0.635, 0.782, 0.671\}\,, i \in I_s = \{1, 2, 3\}\,. \qquad (4.14)$$

The selection of the "best-performing" model with the smallest AIC value from $f_{w,i}$, $i \in I_s$, from eq. (4.14) results again in the model $f_w^\star(\mathbf{x}, t, \theta) := f_{w,1}(\mathbf{x}, t, \theta_1)$ (cf. (4.2a)). Fig. 4.6 depicts the estimation result in comparison to the exact transport coefficient. The corresponding corrected optimal parameter vector results now in

$$\theta_1^\star = (\vartheta_1^\star, \ldots, \vartheta_6^\star)^T = (1.104, 0.723, 4.069, -0.149, 0.826, 0.186)^T\,. \qquad (4.15)$$

The comparison with the parameter estimates that follow directly after the incremental identification method (cf. (4.11)) and with the exact parameters in eq. (4.2b)
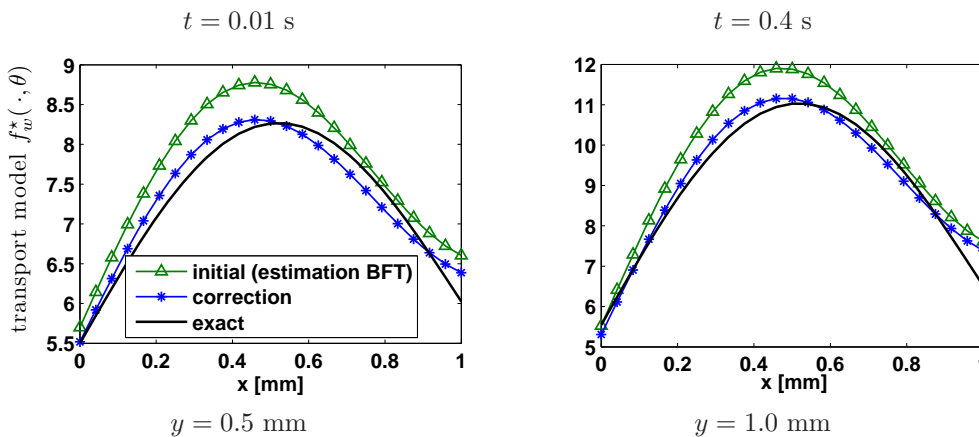

$$t = 0.01 \text{ s} \qquad\qquad t = 0.4 \text{ s}$$

$$y = 0.5 \text{ mm} \qquad\qquad y = 1.0 \text{ mm}$$

FIG. 4.6. *Initial, exact and "best-performing" transport model $f_w^\star(\mathbf{x}, t, \theta) = f_{w,1}(\mathbf{x}, t, \theta_1)$ (upper row) at different times and constant $z = 0.5\,mm$ for noisy data ($\sigma = 0.1$).*

21

reveals that most of the parameter estimates are moved towards the exact parameter values. Note, that $\vartheta_4^\star$ showing large deviations from the correct value governs the time-dependency in the model structure. Because of the short duration of the experiment, it can not be correctly recovered as in the noise-free case before (cf. (4.9b)). The even larger deviation is due to the error in the measurement data that, as already mentioned above, is more severe for small (cf. Fig. 4.5 lower row) rather than for larger values of time t.

The results of the correction for all candidate models from I_s (cf. (4.5)) in the case of noisy measurements arealso shown in Tab. 4.2. Again, most of the corrected estimates have moved towards the exact parameter values (cf. (4.2b)), though deviations are still present.

The results obtained from a solution of the direct parameter estimation problem (1.1) with the model $f_{w,1}(\cdot,\theta_1)$ using the parameter estimate (4.11) as initial value is given in Tab. 4.3 (a), right. Similar to the noise-free case, this estimate is not as good as the one obtained after the correction procedure. In addition to the reasons explained in the previous section, the error in the measurement data contributes to the deficiencies in the estimation. A sensitivity analysis of the model using the parameter estimate $\theta_1^\star$ Tab. 4.3 (a), right, lead to the very same insights as reported for the noise-free case. The values of the correlated parameters $\vartheta_3,\ldots,\vartheta_6$ are fixed to their exact values (cf. (4.2b)), the results in given in Tab. 4.3 (b), right are obtained. As in the noise-free case, these values are much closer to their exact counterparts. Moreover, the corresponding estimates obtained after the correction procedure (cf. Tab. 4.2) compare with these values well.

Also in the case of noisy measurements, convergence to the solution could not be achieved for the direct parameter estimation problem using the same initial values employed in the third step of the incremental identification method. Consequently, continuing the conclusion of the previous section, also in the more realistic case of noisy measurement data, the incremental identification method supplied with the proposed correction procedure represents an attractive strategy to handle nonlinear, ill-posed, transient, distributed (three-dimensional) parameter systems with structural model uncertainty as considered in this paper.

5. Conclusions. A method for the identification of a "best-performing" model for transport coefficients in a transient, distributed system of convection-diffusion type given a set of measurements is presented. No a priori knowledge on the transport model is necessary. The principal ingredient is the method of incremental identification, which decomposes the problem (1.1) into three sequential steps. In this way, a systematic identification of an appropriate model structure and associated parameters becomes possible. Nevertheless, the error propagation through the sequence of inverse subproblems and the associated numerical methods has to be treated to obtain "best-performing" transport model with highest precision in the parameters among the set of proposed candidates. Therefore, a two-step correction procedure is proposed as a supplement to the incremental identification. The computed optimal solutions available from the incremental identification are used as initial values in both steps of the correction procedure facilitating convergence to a high quality correction.

The identification method is exemplarily applied to the identification of a simulated transport model for the effective thermal diffusivity in a three-dimensional convection-diffusion model problem. The assumption of a known model structure for the transport coefficient made in [29] is relaxed. Rather, a list of candidate models is introduced by inspection of the results of the first two steps of the incremental

identification method. After parameter estimation a subset of reasonable candidate models is chosen from this list to apply the correction procedure in order to obtain the "best-performing" model.

Instead of re-solving problem (1.1) for each candidate model from the chosen subset, the decoupled, two-step correction procedure is introduced for the reasons of computational efficiency and estimation quality. The transient coefficient inverse problem occurring in the coefficient correction step is solved using the truncated Newton-CGNE method [23], which is suitable for the solution of such nonlinear inverse problems [17]. For reliable solutions, the problem has to be regularized in an appropriate norm. Moreover, initial values, sufficiently close to the solution are necessary. While these are usually not available in practice, they become accessible due to the incremental decomposition introduced in our approach. Yet, regularization properties of the truncated Newton-CGNE method should be carefully applied to obtain solutions of good quality.

After employing the corrected transport coefficient in the parameter correction step of the correction procedure, the model selection step is carried out. As a result, for a "best-performing" transport model a significant correction can be observed, both for noise-free and noisy data.

For comparison, the direct parameter estimation problem (1.1) assuming the *correct* candidate model structure $f_{w,1}(\cdot, \theta_1)$ has been addressed. The parameter estimates $\theta_1^\star \in \mathbb{R}^6$ resulting from the incremental identification method (cf. (4.6) and (4.11) for the noise-free and the noisy cases, respectively) were used as initial values. The estimates resulting for noise-free and noisy measurements are worse than those obtained after the correction procedure. This is attributed to the strong nonlinearity and the nonconvexity of the parameter estimation problem and the uncertainty in the transport model structure employed. This conjecture is farther supported by a sensitivity analysis of the model, which uncovered the poor identifiability of some of the parameters. In particular, though all parameters are identifiable, only after fixing the poorly identifiable parameters to their exact values, it was possible to obtain good estimates for the remaining parameters. These estimates are comparable to those obtained after the correction procedure in the incremental approach. Hence, incremental identification supplied with an appropriate correction procedure to address inavitable error propagation can provide valuable estimates for *all* parameters even when the level of detail of the exact transport model is high. Moreover, we emphasize that we were not able to find a solution of the direct parameter estimation problem if the initial values were not very close to the exact solution. This shows too that the incremental identification method constitutes an attractive and sound strategy for handling transient, distributed systems of convection-diffusion type in a rigorous manner.

In our future work, the method will be applied to real data in a laminar film flow experiment. Furthermore, we will aim to improve the identifiability of the transport coefficients a (cf. eq. (1.1)) from experimental data by means of an optimal design of experiments [43]. We will investigate to which extent different experimental conditions (e.g., measurement resolution and/or boundary conditions) result in experimental data with sufficient information content for an even more reliable selection of the model structure from a set of candidates and for an improvement of the precision of the parameter estimates.

23

REFERENCES

[1] H. Akaike, *A new look at the statistical model identification*, IEEE Trans. Automat. Contr., 19 (1974), pp. 716–723.

[2] O. M. Alifanov, E.A. Artyukhin, and S.V. Rumyantsev, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begell House, Inc., U.S.A., 1995.

[3] O. M. Alifanov, Y. Jarny, P.V. Prosuntsov, and G.A. Ivanov, *Complex identification of thermophysical properties of anisotropic composite material*, in Proceedings of the 5th international Conference on Inverse Problems in Engineering: Theory and Practice, Cambridge, U.K., 2005.

[4] U. M. Ascher and E. Haber, *Grid refinement and scaling for distributed parameter estimation problems*, Inverse Probl., 17 (2001), pp. 571–590.

[5] A. Bardow and W. Marquardt, *Incremental and simultaneous identification of reaction kinetics: methods and comparison*, Chem. Eng. Sci., 59 (2003), pp. 2673–2684.

[6] ———, *Identification of diffusive transport by means of an incremental approach*, Comp. Chem. Eng., 28 (2004), pp. 585–595.

[7] J. V. Beck, B. Blackwell, and A. Haji-Sheikh, *Comparison of some inverse heat conduction methods using experimental data*, Int. J. Heat Mass Tran., 39 (1996), pp. 3649–3657.

[8] J. V. Beck and K. A. Woodbury, *Inverse problems and parameter estimation: integration of measurements and analysis*, Meas. Sci. Technol., 9 (1998), pp. 839–847.

[9] H. Brauer, *Strömung und Wärmeübergang bei Rieselfilmen*, in VDI-Forsch.-Heft 457, VDI-Verlag, Düsseldorf, 1956.

[10] M. Brendel, D. Bonvin, and W. Marquardt, *Incremental identification of kinetic models for homogeneous reaction systems*, Chem. Eng. Sci., 61 (2006), pp. 5404–5420.

[11] R. S. Brodkey and H. C. Hershey, *Transport Phenomena — A Unified Approach*, McGraw-Hill Book Company, U.S.A, 1976.

[12] R. Brun and P. Reichert, *Practical identifiability analysis of large environmental simulation models*, Water Resour. Res., 37 (2001), pp. 1015–1030.

[13] K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference*, Springer Verlag, New York, 2002.

[14] M. J. Colaco, G. S. Dulikravich, H. R. B. Orlande, and F. A. Rodrigues, *A comparison of two solution techniques for the inverse problem of simultaneously estimating the spatial variations of diffusion coefficient and source terms*, in Proceedings of IMECE'03, 2003 ASME International Mechanical Engineering Congress and Exposition, Washington, D.C., U.S.A., 2003.

[15] G. Dietze, V.V. Lel, and R. Kneer, *Modelling of heat transfer in stable wavy film flow based on effective thermal diffusivity*, in Proceedings of IHTC-13 2006, Sidney, 2006.

[16] *DROPS package.* http://www.igpm.rwth-aachen.de/DROPS/.

[17] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[18] S. Finsterle, *Multiphase inverse modeling: Review and iTOUGH2 applications*, Vadose Zone Journal, 3 (2004), pp. 747–762.

[19] S. Gross, M. Soemers, A. Mhamdi, F. Al Sibai, A. Reusken, W. Marquardt, and U. Renz, *Identification of boundary heat fluxes in a falling film experiment using high resolution temperature measurements*, Int. J. Heat Mass Tran., 48 (2005), pp. 5549–5562.

[20] E. Haber, U. M. Ascher, and D. Oldenburg, *On optimization techniques for solving nonlinear inverse problems*, Inverse Probl., 16 (2000), pp. 1263–1280.

[21] M. Hanke, *Conjugate Gradient Type Methods for Ill-Posed Problems*, Longman Scientific & Technical, 1995.

[22] ———, *A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Probl., 13 (1997), pp. 79–95.

[23] ———, *Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse prob-

lems, Numer. Funct. Anal. Optim., 18 (1997), pp. 971–993.

[24] Y. Heng, A. Mhamdi, S. Gross, A. Reusken, M. Buchholz, H. Auracher, and W. Marquardt, *Reconstruction of local heat fluxes in pool boiling experiments along the entire boiling curve from high resolution transient temperature measurements*, Int. J. Heat Mass Tran., 51 (2008), pp. 5072–5087.

[25] C. H. Huang and W.-C. Chen, *A three-dimensional inverse forced convection problem in estimating surface heat flux by conjugate gradient method.*, Int. J. Heat Mass Tran., 43 (2000), pp. 3171–3181.

[26] Cheng-Hung Huang, Chun-Ying Yeh, and H. R. B. Orlande, *A nonlinear inverse problem in simultaneously estimating the heat and mass production rates for a chemically reacting fluid*, Chem. Eng. Sci., 58 (2003), pp. 3741–3752.

[27] V. Isakov, *Inverse Problems for Partial Differential Equations*, Springer-Verlag New York, Inc., 1998.

[28] J. P. Kaipio and E. Somersalo, *Computational and Statistical Methods for Inverse Problems*, Springer Verlag, New York, 2004.

[29] M. Karalashvili, S. Gross, A. Mhamdi, A. Reusken, and W. Marquardt, *Incremental identification of transport coefficients in convection-diffusion systems*, SIAM J. Sci. Comput., 30 (2008), pp. 3249–3269.

[30] M. Karalashvili, A. Mhamdi, G. F. Dietze, H. M. Bücker, A. Vehreschild, R. Kneer, C. H. Bischof, and W. Marquardt, *Sensitivity analysis and identification of an effective heat transport model in wavy liquid films*, in Progress in Computational Heat and Mass Transfers, ICCHMT'09, Guangzhou, China, 2009, pp. 644–651.

[31] W. Marquardt, *Towards a process modeling methodology*, vol. 293 of R. Berber: Methods of Model-Based Control, NATO-ASI Ser. E, Applied Sciences, Kluwer Academic Pub., Dordrecht, Kluwer Academic Pub, 1995, pp. 3–41.

[32] ———, *Model-based experimental analysis of kinetic phenomena in multi-phase reactive systems*, Chem. Eng. Res. Des., 83 (2005), pp. 561–573.

[33] C. E. Meza and V. Balakotaiah, *Modeling and experimental studies of large amplitude waves on vertically falling films*, Chem. Eng. Sci., 63 (2008), pp. 4704–4734.

[34] C. Michalik, M. Brendel, and W. Marquardt, *Incremental identification of fluid multi-phase reaction systems*, AlChE J., 55 (2009), pp. 1009–1022.

[35] T. K. Nilssen, K. H. Karlsen, T. Mannseth, and X. C. Tai, *Identification of diffusion parameters in a nonlinear convection-diffusion equation using the augmented lagrangian method*, Computational Geosciences, 13 (2009), pp. 317–329.

[36] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, Springer, Berlin, Heidelberg, New York, 1999.

[37] L. Prandtl, *Über Flüssigkeitsbewegung bei sehr kleiner Reibung*, in Gesammelte Abhandlungen zur angewandten Mechanik, Hydro- und Aerodynamik, W. Tolmien, H. Schlichting, and H. Görtler, eds., vol. 2, Springer-Verlag, Berlin, 1961.

[38] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Probl., 15 (1999), pp. 309–327.

[39] ———, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal., 43 (2005), pp. 604–622.

[40] J. Su and F. H. Geoffrey, *Inverse heat conduction problem of estimating time-varying heat transfer coefficient*, Numer. Heat Transf., 45 (2004), pp. 777–789.

[41] U. Tautenhahn, *On a general regularization scheme for nonlinear ill-posed problems: II. Regularization in Hilbert scales*, Inverse Probl., 14 (1998), pp. 1607–1616.

[42] M. Tawarmalani and N. V. Sahinidis, *A polyhedral branch-and-cut approach to global optimization*, Mathematical Programming, 103 (2005), pp. 225–249.

[43] E. Walter and L. Pronzato, *Identification of Parametric Models from Experimental Data*, Springer, Berlin, 1997.

[44] W. Wilke, *Wärmeübergang an Rieselfilmen*, in VDI-Forsch.-Heft 490, VDI-Verlag, Düsseldorf, 1962.

[45] L. Zhang, K. Kulkarni, R. M. Somayaji, M. Xenos, and A. A. Linninger, *Discovery of transport and reaction properties in distributed systems*, AIChE J., 53 (2007), pp. 381–396.