Frequency domain sample Maximum Likelihood Estimation for spatially dependent parameter estimation in PDEs [★]

Matthijs van Berkel [a,b,c] Gerd Vandersteen [b] Egon Geerardyn [b] Rik Pintelon [b]
Hans Zwart [a] Marco de Baar [a,c]

[a] *Eindhoven University of Technology, Dept. of Mechanical Engineering, Control Systems Technology group, PO Box 513, 5600 MB Eindhoven, The Netherlands*

[b] *Vrije Universiteit Brussel, Department of Fundamental Electricity and Instrumentation, Pleinlaan 2, 1050 Brussels, Belgium*

[c] *FOM institute DIFFER, Dutch Institute for Fundamental Energy Research, Association EURATOM- FOM, Trilateral Euregio Cluster, PO Box 1207, 3430 BE Nieuwegein, The Netherlands*

Key words: Partial Differential Equations, Maximum likelihood estimators, Systems transfer functions, Heat flows

1 Introduction

Many processes in physics are described by partial differential equations (PDEs). Typically, the parameters describing these PDEs depend on the spatial coordinate. The estimation of these spatially dependent parameters is important for understanding the physical behavior of these PDEs.

The identification of PDEs is treated in the class of infinite dimensional systems or Distributed Parameter Systems (DPS). In the DPS literature the problem of estimating parameters is extensively treated in the monograph by Banks et al. [3], and references therein. In addition, different methodologies exist to identify spatially varying parameters in PDEs [18], of which a number focus specifically on parabolic PDEs [4,19,22]. Their emphasis is on the regularization of the least-squares cost-function used to estimate the parameters, to assure well-posedness in Hadamard's sense, i.e., to guarantee the existence of its solution, the uniqueness of this solution and its stability with respect to the measurement data [38]. In addition, regularization is often used to make the optimization problems convex and to constrain the

solution in some sense [16]. Moreover, different methods for reducing the infinite dimensional PDEs to finite dimension exist and are discussed in the above mentioned references. Generally, they are using some discretization of the spatial coordinate [14,32] or basis functions [2,6].

A different approach to estimate the parameters is to solve the problem in the frequency domain via the Laplace Transform [8]. This reduces the PDE to an parameterized Ordinary Differential Equation (ODE). These ODEs can be approximated or sometimes even be solved analytically, thereby avoiding approximation errors. The solutions are generally of non-rational or fractional form [7], of which the parameters can be identified using frequency domain identification techniques [10,17,24,37]. The disadvantage of this approach is that the non-rational form of the model can complicate the identification significantly. On the other hand, the use of periodic excitations enables the removal of the unexcited noisy frequency lines from the measured data, hence high Signal-to-Noise ratios (SNR) can be obtained at the excited frequency lines by averaging the signals over consecutive periods. In addition, it is no longer required to use rational approximations in the frequency domain making the identification easier.

In this paper, the following second order parabolic PDE with spatially varying parameters on a one-dimensional domain is considered

$$\frac{\partial z}{\partial t} = f_D\left(x,\theta\right)\frac{\partial^2 z}{\partial x^2} + f_V\left(x,\theta\right)\frac{\partial z}{\partial x} + f_K\left(x,\theta\right)z + Q\left(x,t\right) \tag{1}$$

where the spatial coordinate is denoted by x. We assume that the functions $f_D\left(x,\theta\right)$, $f_V\left(x,\theta\right)$, and $f_K\left(x,\theta\right)$ de-

pend on the unknown parameters θ, which we want to identify.

A specific process described by (1) is the radial heat-transport inside the tokamak plasma [11], a toroidal nuclear fusion device. For the optimization of this radial heat-transport, which determines the efficiency, the different profiles need to be identified, i.e. f_D, and possibly f_V and f_K.

In the tokamak the initial conditions are unknown, hence only measurements are considered for "$t \gg$ the dominant time constant" such that the transients are negligible compared to the forced response. The source (microwave heating) term periodically excites the plasma and is defined as

$$Q(x,t) = \begin{cases} q(x) \sum_{k=1}^{K} A_k \cos(\Omega_k t + \phi_k) & x_1 \leq x \leq x_2 \\ 0 & elsewhere \end{cases}$$
(2)

where A_k, ϕ_k, and Ω_k are the amplitude, phase, and frequency respectively. The spatial dependence of the source $q(x)$ is unknown, but $q(x) \neq 0$. The plasma temperatures, z, can be measured locally by microwave radiation, which is prone to additive normally distributed noise. Moreover, boundary conditions exist such that (1) has a unique solution. However, these boundary conditions are subject to debate and are assumed to be unknown. This means that for the parameter estimation only measurements are available.

The estimation of parameters from noisy measurements only is known as an Errors-in-Variables problem (EIV). In the EIV literature it is well known that the least-squares estimator is not consistent [33], i.e. the parameter estimates will be biased. Moreover, this bias depends on the SNR, which decreases with the distance to the source $Q(x,t)$. This is caused by the low-pass characteristic of (1), which also causes the higher frequencies to be more noisier than low frequencies. This problem can be partly overcome using a low-pass filter to suppress the noisy "high-frequency" components in the measurements [34]. However, the optimal cut-off frequency of this filter depends on the unknown system parameters.

In this paper a frequency domain sample Maximum Likelihood Estimator (SMLE) is used for this EIV problem. The SMLE is based on the Probability Density Function of the noise, allowing for a consistent estimate under weak assumptions [23]. Moreover, it naturally weights the different frequency components avoiding the necessity of a low-pass filter. In the SMLE framework the confidence bounds on the estimated parameters can be calculated and model validation tests exist. However, the SMLE requires knowledge about noise properties of the measurements. A disadvantage of the SMLE is that its

optimization problem is more complex and is no longer convex in contrast to linear least-squares estimators.

Here we choose to estimate a set of small sub-domains, instead of modeling the entire domain. On each sub-domain a complex valued non-rational analytic model description is used, hence avoiding discretization errors [5]. The advantage of this approach is that the decoupling in sub-domains assures that errors do not propagate from one sub-domain to the others. In addition, the SMLE optimization problem remains solvable, because only a few parameters need to be estimated.

A sub-domain needs to consist of at least three measurements. The reason is that two measurements act as the boundary conditions (inputs) as such defining the solution of the second order PDE. Then, at least one measurement point between the boundaries (output) is necessary to compare it with the solution of the PDE at that location, which is determined by the parameters, the model structure, and the boundary conditions. In principle the number of measurements can be extended, but then a Multiple-Input Multiple-Output system needs to be identified. Therefore, in this paper three adjacent measurements are used to define a sub-domain. Note, that we only consider sub-domains that are outside the interval $[x_1, x_2]$, i.e. domains that do not contain an excitation source $q(x)$ (see (1) and (2)). The reason is that the exact dependency of $q(x)$ on x is unknown.

A general framework for different spatial dependencies and geometries is discussed, but results are only shown for sub-domains on which parameters can be modeled as constants. Consequently, only piecewise smooth profiles are identifiable, which is generally true for the considered application.

An identification scheme consists of the following three components: a model of the system, derived in Section 2; a cost function minimization scheme based on a realistic noise model, explained in Section 3; and simulations which are generated by means of a finite difference model such that the result can be validated. The latter is discussed in Section 4. Finally, a number of conclusions are summarized and discussed.

2 Modeling

This section derives the transfer functions based on a smart choice of the boundary conditions. The most important concept is the replacement of boundary conditions by measurements. In addition, derivations are done without specifying any spatial dependency.

2.1 Considered Partial Differential Equation

The one-dimensional, second-order linear PDE introduced in (1) is excited by the local source term $Q(x,t)$.

The PDE is modeled by the homogeneous form of (1) outside this source domain. It is also possible to solve the inhomogeneous PDE, allowing the use of a domain with a source. However, the inhomogeneous PDE results in a more complex relationships and therefore will not be discussed.

The Laplace transform of (1) simplifies the PDE towards a complex valued Ordinary Differential Equation (ODE) of the form

$$0 = f_D(x,\theta) \frac{d^2 Z}{dx^2} + f_V(x,\theta) \frac{dZ}{dx} + (f_K(x,\theta) - s) Z \quad (3)$$

with $Z = \mathcal{L}\{z\}$ and s the Laplace variable. The functions $f_D(x,\theta)$, $f_V(x,\theta)$, and $f_K(x,\theta)$ depend on the spatial coordinate x and the time-invariant unknown parameters θ. In many cases this ODE can be solved analytically, with the general solution given by

$$Z(x,s) = C_1(s)\,\xi(x,\theta,s) + C_2(s)\,\zeta(x,\theta,s), \quad (4)$$

where $C_1(s)$ and $C_2(s)$ denote the free variables set by the boundary conditions. The choice of the spatially dependent functions $f_D(x,\theta)$, $f_V(x,\theta)$, and $f_K(x,\theta)$ in (3) determine the solutions $\xi(x,\theta,s)$ and $\zeta(x,\theta,s)$, which are the complex eigenfunctions of (3). Solutions for many different choices of $f_D(x,\theta)$, $f_V(x,\theta)$ and $f_K(x,\theta)$ exist, such as constant, and linear functions. An extended list of these analytic solutions can be found in [27]. The next step is to derive the local transfer functions.

2.2 Boundary Condition description

This paper considers a set of small sub-domains, which prevents the errors from propagating. Moreover, the boundary conditions are no longer approximated but measured, in contrast to other strategies such as the infinite domain boundary condition reported in [7]. Every domain uses three adjacent frequency spectra of the measurements z to estimate the local parameters. The outer two measurements act as the boundary conditions i.e. $Z(x_{i-1}, s)$ and $Z(x_{i+1}, s)$. These two boundary conditions allow for the calculation of $C_1(s)$ and $C_2(s)$. Rearranging and defining $Z(x_i, s)$ as the output measurement results in the following Multiple-Input Single-Output system (θ and s are omitted):

$$Z(x_i) = \left(\frac{\xi(x_{i+1})\zeta(x_i) - \zeta(x_{i+1})\xi(x_i)}{\zeta(x_{i-1})\xi(x_{i+1}) - \zeta(x_{i+1})\xi(x_{i-1})} \right) Z(x_{i-1})$$

$$- \left(\frac{\xi(x_{i-1})\zeta(x_i) - \zeta(x_{i-1})\xi(x_i)}{\zeta(x_{i-1})\xi(x_{i+1}) - \zeta(x_{i+1})\xi(x_{i-1})} \right) Z(x_{i+1}), \quad (5)$$

with $i = 2, ..., m-1$, where m denotes the number of sensors. The inputs and outputs are defined as $U_1(s) =$

$Z(x_{i-1}, s)$, $U_2(s) = Z(x_{i+1}, s)$, and $Y(s) = Z(x_i, s)$ on the interval x_{i-1}, x_{i+1}:

$$Y(s) = G_1(\theta, s)\, U_1(s) - G_2(\theta, s)\, U_2(s). \quad (6)$$

In (5), the choices of $\xi(x,\theta,s)$ and $\zeta(x,\theta,s)$ are deliberately undefined, as different dependencies of f_D, f_V, and f_K can be used. Here, we only consider constant parameters.

2.3 Constant Parameters

This section derive the transfer functions in the case of constant parameters over a small domain. Consequently, $\xi(x,\theta,s)$ and $\zeta(x,\theta,s)$ in (5) are constant over the domain. The local PDE description is given by

$$\frac{\partial z}{\partial t} = D\,\frac{\partial^2 z}{\partial x^2} + V\,\frac{\partial z}{\partial x} + K\,z, \quad (7)$$

where D denotes the diffusivity, V the convective velocity, and K the damping. The eigenfunctions are fixed by the choice of local constant spatial dependencies i.e.

$$\xi(x,\theta,s) = \exp(\lambda_1 x) \text{ and } \zeta(x,\theta,s) = \exp(\lambda_2 x), \quad (8)$$

where $\lambda_{1,2}$ are the eigenvalues of the underlying ODE:

$$\lambda_{1,2} = \frac{-V \mp \sqrt{V^2 + 4D(s-K)}}{2D}. \quad (9)$$

As the combination of (5) and (9) result in a non-rational Multiple-Input Single-Output system with complex exponents and square roots, identifying the parameters is not straightforward. Therefore, a change of variables is introduced to facilitate the parameter estimation.

2.4 Change of Variables

The parameters will be estimated by minimizing a cost function. The computation time and possibly avoidance of local minima can be improved by simplifying the equations to be evaluated. Therefore, a substitution is introduced such that the eigenvalues are simplified, avoiding parameter divisions, i.e.

$$\lambda_{1,2} = -a \mp \sqrt{a^2 + b + c\,s}, \quad (10)$$

with $a = \dfrac{V}{2D}$, $b = -\dfrac{K}{D}$, and $c = \dfrac{1}{D}$.

This parameter set will be denoted as $\theta = \begin{bmatrix} a & b & c \end{bmatrix}^T$ and the estimated set by $\hat{\theta}$. Although not discussed here, similar simplifications are possible for other choices of the spatial dependent functions.

3

In the next section, not only the parameters are estimated, but also their covariance. Consequently, the covariance matrix of the estimated parameters can be recalculated using

$$\mathrm{Cov}\left(\hat{D}, \hat{V}, \hat{K}\right) = J_D \, \mathrm{Cov}\left(\hat{\theta}\right) J_D^T$$

$$\text{with } J_D = \frac{1}{\hat{c}^2} \begin{bmatrix} 0 & 0 & -1 \\ 2\hat{c} & 0 & -2\hat{a} \\ 0 & -\hat{c} & \hat{b} \end{bmatrix}. \tag{11}$$

3 Sample Maximum Likelihood Estimator

In this section, the Errors-in-Variables (EIV) problem is discussed, for which the sample Maximum Likelihood Estimator (SMLE) offers a solution. It is also discussed how to minimize the sample Maximum Likelihood cost function, how to construct confidence bounds on the estimated parameters, and how to validate the estimated models.

Maximum Likelihood Estimation is a method for estimating parameters given some statistical properties, where the MLE maximizes a known likelihood function. The likelihood function can be interpreted as a Probability Density Function (PDF), but with respect to the measured data. For example, if the difference between the output and a predicted output based on the model and input is studied, the remainder in the absence of model errors is fully characterized by the likelihood function of the noise. The sample likelihood function differs from the likelihood function in the sense that the real noise (co-)variances are unknown. Hence, they need to be replaced by sample variances, which can be determined using a pre-processing step.

3.1 Error Model: Errors-in-Variables

The transfer functions introduced in (6) are based on local spatial measurements. This means that an EIV approach is necessary to handle the noise on the measurements. In this paper, the EIV problem is solved via the sample Maximum Likelihood Estimator (SMLE) in the frequency domain. It is based on additive circular complex normally distributed noise in the frequency domain [23], which is the result of Gaussian noise in the time domain [12].

In the SMLE, the true unknown noise (co-)variances are replaced by sample estimates obtained from the periodic signal. This is achieved by calculating the average over the different periods and variances per frequency line resulting in the deterministic spectra and the estimated noise spectra. A minimum number of 4 periods is necessary to make a parameter estimate, however, if

at least 7 periods are used other desirable properties of the SMLE are also retained (see [31] for the details). In principle, also measurements containing transients can be used to obtain the (co-)variances using the local polynomial method, but at the cost of a more complex pre-processing step (see [26]).

There also exist different approaches to handle an EIV problem often relying on multiple experiments and specific assumptions on the noise. An overview of the different methods can be found in [33]. In contrast, to the frequency domain (sample) MLE also MLEs in time domain exist [1,9], which handle the spectral factorization and possible transients differently [33]. Moreover, MLEs can also be constructed for non-Gaussian noise distributions [13]. However, in contrast to the non-Gaussian MLE, the MLE for Gaussian noise are extensively studied in e.g. [23,35].

3.2 Maximum Likelihood Cost

The sample MLE cost function is derived on the basis of the system model and the error model. The sample log likelihood cost function V_{SML} is used [23]. It has the same global minimum as the SMLE for complex normally distributed noise, but is computationally less intensive. It is defined as

$$V_{SML} = \frac{1}{F} \sum_{k=1}^{F} \frac{|e\left(\Omega_k, \theta\right)|^2}{\sigma_e^2\left(\Omega_k, \theta\right)}, \tag{12}$$

with Ω_k the excited frequencies and F the number of frequencies used. The error $e\left(\Omega_k, \theta\right)$ is defined as

$$e\left(\Omega_k, \theta\right) = Y\left(\Omega_k\right) - \left(G_1\left(\theta\right) U_1\left(\Omega_k\right) - G_2\left(\theta\right) U_2\left(\Omega_k\right)\right), \tag{13}$$

where the transfer functions G_1 and G_2 are evaluated at Ω_k. The variability, which takes the different noise contributions into account, is given by (dependency on Ω_k and θ are omitted)

$$\sigma_e^2\left(\Omega_k, \theta\right) = \sigma_Y^2 + \sigma_{U_1}^2 |G_1|^2 + \sigma_{U_2}^2 |G_2|^2$$

$$-2\,\mathrm{Re}\left(G_1 \sigma_{U_1 U_2}^2 \overline{G_2} + \sigma_{Y U_1}^2 \overline{G_1} - \sigma_{Y U_2}^2 \overline{G_2}\right), \tag{14}$$

where the variances and covariances are estimated for every Ω_k using M realizations (periods). The complex conjugate of G is denoted as $\overline{G}$. The parameters are estimated by minimizing V_{SML}

$$\hat{\theta} = \arg\min_{\theta} V_{SML}\left(\Omega_k, \theta\right). \tag{15}$$

Minimizing this cost function gives the estimated parameters. For filtered white noise disturbances the minimizer of (12), based on the non-rational transfer models, has exactly the same asymptotic ($F \to \infty$) properties as

the SMLE for rational transfer function models. For example, it is consistent and asymptotically normally distributed (see [23] for the details). The cost function (12) can be naturally interpreted as a weighting of the error with the uncertainty of the measurements. Measurements with small noise variances have a higher weighting and vice versa. If σ_e is constant for all Ω_k, then (12) reduces to a non-linear least-squares (NLS) estimator (output error).

3.3 Optimization and confidence bounds

The minimization of (12) is in principle non-convex. On the other hand, only a few parameters need to be estimated. Therefore, the entire relevant parameter space can easily be searched for the cost function's minimum. In addition, if the noise levels are reasonable and the parameters are optimized in terms of θ, gradient based algorithms converge to the global minimum for many initialization values. These gradient methods are computationally cheap, especially if an analytic Jacobian is used.

The analytical Jacobian is also used to estimate the confidence bounds on the parameters. Section 3.2 explains that minimizing the cost function (12) provides the best parameter set. The uncertainty of the parameters is determined using the Jacobian from a first-order Taylor series expansion

$$J_\theta = \frac{\partial}{\partial\theta}\left(\frac{e\left(\Omega_k,\theta\right)}{\sigma_e\left(\Omega_k,\theta\right)}\right). \qquad (16)$$

The resulting covariance matrix on the different transformed parameters is given by

$$\text{Cov}\left(\hat{\theta}\right) \approx \left(\frac{M-1}{M-3}\right)\left[\text{Re}\left(2J_\theta^H J_\theta\right)\right]^{-1}, \qquad (17)$$

with M the number of realizations (periods) [25]. This correction is necessary, because estimated sample variances are used instead of the real (but unknown) variances. The uncertainty on the real valued parameters can be calculated by (11).

3.4 Model validation

A cost function analysis is used to detect model errors. If the noise is indeed normally distributed, no model errors are present, and if a number of weak assumptions are fulfilled [23], then the expected value of the sample Maximum Likelihood cost function equals the number of frequency lines F minus the number of free real-valued parameters n_θ divided by two i.e.

$$V_{noise} = \left(F - \frac{n_\theta}{2}\right). \qquad (18)$$

In addition, the theoretical variance of the SMLE cost equals V_{noise}. However, as variances are estimated using M repeated experiments, a correction is necessary to take this extra uncertainty into account i.e. $\mathbb{E}\left\{V_{SML}\right\} = (M-1)/(M-2)\,V_{noise}$ and $\text{var}\left\{V_{SML}\right\} = (M-1)^3/\left((M-3)(M-2)^2\right)V_{noise}$. The estimated variance is used to construct confidence bounds. Model errors generally lead to a higher value of the cost function at the global minimum.

3.5 Input design and choice of domain

The SMLE developed in this paper always considers a domain of three measurement points, of which the locations can be non-uniformly distributed. In principle, the larger the domain the better the estimate of the parameters. The reason is that the suppression of the amplitude increases with the distance due to the source. Hence, the amplitude difference between the inputs and output is also larger. However, the larger the domain, the more stringent the assumption of constant parameters becomes. This means there is a trade-off between the sensor distance and the assumption of constant parameters. These considerations are more important than the overall number of sensors used, because every domain is treated separately in this approach. The validation test and confidence bounds can be helpful in making a choice for the sensor locations. If the model (of constant parameters) is rejected by the validation test, the model assumption is incorrect. In that case the sensor locations should be closer to each other. On the other hand, if the confidence bounds are large then increasing the distance between the sensors could improve the accuracy.

The other important aspect is the source defined in (2). The theoretical minimum number of sinusoidal components in the source should be $F > \frac{n_\theta}{2}$. However, every extra sinusoidal component increases the accuracy of the estimated parameters. The domain to be identified is preferably close to the source, but should not contain the source. In case the domain has some distance to the source, the diffusive process acts as a low-pass filter, which generally reduces the optimal frequencies Ω_k with respect to the parameter accuracy when compared with a similar domain closer to the source. In addition, the (unknown) parameters on the domain, and between the source and the domain, directly influence the choice of the optimal excitation frequency. This directly follows from (9) if $s = i\omega, \omega \to \infty$ then V and K become negligible. On the other hand, K can be estimated using sine components at a very low-frequency, which follows from $s \to 0$, and V is best identified somewhere in the intermediate region.

It is clear, that selecting the optimized identification accuracy is complex and depends on many factors. Moreover, it depends on which parameter needs to be identified accurately and which domains need to be identi-

fied accurately. The accuracy can be evaluated using the Fisher information matrix. As such this matrix plays an important role in designing optimal excitation signals and sensor placement. However, the exact design of the optimal excitation signal and sensor placement is complicated, and will not be considered in this paper. The reader is referred to [28,36] for a treatment of sensor placement and excitation design in DPS. On the other hand, the design of optimal excitation signals for transfer functions is treated in, e.g. [29,30]. Note, that also the crest factor and randomness of the excitation signal should be considered to reduce non-linear effects [23].

4 Results

This section shows a number of analysis steps to validate the estimation procedure presented in the previous sections through a Monte Carlo Analysis. Then, a finite difference simulation presents the estimation results for varying profiles of D and V. The models estimated using a finite difference simulation are validated by means of a cost function analysis.

4.1 Estimator and confidence bound validation

The possibility to construct confidence bounds on the estimated parameters is an important advantage of the SMLE. However, the implementation of any estimation algorithm can be error-prone. Therefore, it is important to validate the implementation. This especially holds for schemes which use an analytic Jacobian. A number of trivial tests, such as Finite Difference Model comparison, convergence to the true minimum, and comparison of the Fourier transforms were done.

Next to validating the model, the estimated confidence bounds on the estimator need to be verified. The SMLE has a number of advantages when compared to the least-squares estimators. The estimates remain consistent if the additive complex circular normal noise is filtered or the different noise sources are correlated (covariance), because the noise models are also estimated based on the realizations of the periodic signals. Under the conditions of filtered or colored noise, the estimator should find a consistent estimate and should approximate the confidence bounds accordingly. This can be tested by means of Monte Carlo simulations. The noise sources have been filtered and have been correlated. The original model is used and different noise realizations (10000) with the same variance and mean are simulated. The confidence bounds were validated under these conditions. It turned out that the parameter $b = -K/D$ is extremely sensitive to noise. This is caused by the necessity of excitation signals with very low Ω_k and the presence of V and D. Therefore, the damping K is fixed at zero in the rest of this paper.

4.2 Finite difference simulations

Finite difference simulations are used to validate the methodology. Only the estimated diffusivity D and convective velocity V are considered as the damping K is very uncertain for reasonable excitation frequencies. Both the diffusivity profile and convective profile contain a step (simulated by an erf function), but at different locations. The reason for this choice of profiles is to show that the domains can be identified independently from the profile outside the domain. The problem of identifying the parameters with discontinuities is not treated here, but can be found in e.g. [20]. In the specific example of tokamak plasmas the steps have experimentally been observed and can be interpreted as transport barriers [15].

The excitation source $Q(x, t)$ is placed at $x = 0.1$ and hence influences the estimates in the gray area. The sensors are positioned at $x = 0.1, 0.15, ..., 0.95$, only $x = 0.75$ is not present. The source is a multi-sine with equal amplitude exciting the harmonics from 20 to 400 Hz. In total 40 periods are observed (2 s). Noise is added, which is normally distributed and is for 20% uncorrelated and 80% correlated (over different x_i), which is a realistic situation. Although, gradient based methods work well for lower noise levels, at higher noise levels it sometimes converges to a local minimum. Therefore, a simple grid search algorithm is applied. The resulting local estimates are presented in Fig. 1.

The SMLE is compared to the non-linear least-squares (NLS) estimator (12) with $\sigma_e = 1$, which is comparable to an output error approach in the time-domain [21] (Parseval's theorem). At low noise levels their performance is very similar. If the measurements become more noisy, due to a larger distance to the source, the estimates diverge from the true values. However, the confidence bounds for the SMLE still give correct confidence bounds. On the other hand, at higher noise levels the NLS estimates diverge significantly from the true values and perform worse than the SMLE. Moreover, the estimated D seems relatively insensitive for a step in $f_V(x)$. In addition, the estimates of the non-symmetric domains using the measurements at $x = [0.65, 0.7, 0.8]$ and $x = [0.7, 0.8, 0.85]$, which are plotted at $x = 0.7$ and $x = 0.8$ respectively, perform as expected. Although, the performance at the steps seems good, it needs to be decided on the basis of the validation test.

4.3 Model validation of finite difference simulation

In Section 3.4 a model validation test is introduced, which is used here to validate the estimates of the finite difference simulation. The value of the cost function at the global minimum should equal the degrees of freedom (18) within some confidence region. This is verified in Fig. 2, where not only the values of the cost function

Fig. 2. (a) Contour plot of the cost function values as function of the transformed parameters a and c at a specific spatial location. Note, that the original parameters are defined in terms of θ so D is plotted from 1 to 200 and a is a more complicated combination of D and V. (b) The values of the cost function at the global minimum ($\times$), V_{SML} for the different spatial estimates ($\times$). The dashed lines are the 95% confidence bounds based on the expected variance of V_{SML}.

Fig. 1. Comparison between the SMLE and the non-linear least-squares estimator (NLS) ($\sigma_e = 1$). The true profiles (solid lines) end estimates of D and V are presented including the 95% confidence bounds (SMLE only). Some NLS estimates of V become very large, are not included). The estimate is always plotted at the central measurement x_i of the domain. The measurements are generated with a finite difference model of 4000 grid points, time step of 0.1 ms, where the boundary conditions are $\frac{\partial z(x=0)}{\partial x} = 0$ and $z(x=1) = 0$. The amplitude of the ground frequency $|Z(x_i, \Omega_1)|$ and the amplitude of the highest excited frequency $|Z(x_i, \Omega_{20})|$ are shown. The other amplitudes of the excited frequencies are situated in between. In addition, the standard deviation of the noise in the frequency domain $|\sigma_{noise}(x_i, \Omega_k)| = \sigma_y$ is presented, which is constant for all frequencies. In comparison also the SNR in the time domain is presented, which is defined as $SNR_{time}(x_i) = 20\log_{10}\left(\frac{z_{rms}(x_i)}{noise_{rms}(x_i)}\right)$. The crosses represent the measurement locations.

are plotted but also the cost function as function of the parameters at one spatial location.

Fig. 2(a) shows the cost function, due to the parameter transformation introduced in Section 2.4 many initialization values will allow for a gradient based method to converge to the global minimum or its neighborhood. Fig. 2(b) shows that most models describe the data well. However, at lower noise levels than the noise level chosen here, the estimated models at the step in $f_D(x)$ and to a lesser extend at the step in $f_V(x)$ are rejected by the validation.

5   Conclusions and Discussion

This paper presents a new methodology to identify the spatial dependent parameters. The estimation is performed in the frequency domain, allowing analytic models to be used for simple dependencies. A transfer function is derived, based on three measurement points such that unknown boundary conditions can be handled. Moreover, a sample Maximum Likelihood Estimator is used to estimate the parameters. It takes the noise into account, which is present at the inputs and the output. This allows for consistent estimates of the parameter values and their uncertainties.

The SMLE performs better than the non-linear least squares estimator at high noise levels. A validation is applied showing that most estimates are accepted. The local estimation with three parameters ensures a simple and robust minimization of the cost function as only a few parameters need to be estimated. The local estimates also overlap, resulting in some redundancy, which can further improve the estimation procedure by using the overlap information.

Extending the methodology to higher dimensions is not straightforward. The main problem is foreseen in terms of the boundary conditions of a domain. The out and inflow of energy between two measurement locations needs to be described, which probably would require some interpolation. In addition, the required number of measurements increases to at least four in 2D. This means

that the SMLE needs to be extended accordingly.

References

[1] K. J. Åström. Maximum likelihood and prediction error methods. *Automatica*, 16(5):551 – 574, 1980.

[2] H. T. Banks, J. Crowley, and K. Kunisch. Cubic spline approximation techniques for parameter estimation in distributed systems. *IEEE Trans. Autom. Control*, 28(7):773–786, 1983.

[3] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhauser Boston, 1989.

[4] H. T. Banks and P. K. Lamm. Estimation of variable coefficients in parabolic distributed systems. *IEEE Trans. Autom. Control*, 30(4):386–398, 1985.

[5] B. Bhikkaji and T. Söderström. Reduced order models for diffusion systems. *Int. J. Contr.*, 74(15):1543–1557, 2001.

[6] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods in fluid dynamics*. Springer-Verlag Berlin Heidelberg, 1988.

[7] R. F. Curtain and K. Morris. Transfer functions of distributed parameter systems: A tutorial. *Automatica*, 45(5):1101–1116, 2009.

[8] R. F. Curtain and H. J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*, volume 21. Springer-Verlag New York, 1995.

[9] R. Diversi, R. Guidorzi, and U. Soverini. Maximum likelihood identification of noisy input-output models. *Automatica*, 43(3):464 – 472, 2007.

[10] J.-D. Gabano and T. Poinot. Fractional modelling and identification of thermal systems. *Signal Process.*, 91(3):531 – 541, 2011.

[11] K. W. Gentle. Dependence of heat pulse propagation on transport mechanisms: Consequences of nonconstant transport coefficients. *Physics of Fluids*, 31:1105, 1988.

[12] N. R. Goodman. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 34(1):152–177, 1963.

[13] G. C. Goodwin and R. L. Payne. *Dynamic system identification: experiment design and data analysis*. Academic press New York, 1977.

[14] M.T. Heath. *Scientific computing*. McGraw-Hill New York, 1997.

[15] G. M. D. Hogeweij, N. J. Lopes Cardozo, M. R. de Baar, and A. M. R. Schilham. A model for electron transport barriers in tokamaks, tested against experimental data from RTP. *Nuclear Fusion*, 38(12):1881, 1998.

[16] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15. SIAM Philadelphia (PA), 2008.

[17] A. Jalloul, K. Jelassi, P. Melchior, and J.-C. Trigeassou. Comparison of identification techniques for fractional models. In *IEEE Multi-Conference on Systems, Signals and Devices (SSD), 2011 8th International*, pages 1–6, 2011.

[18] C. Kravaris and J. H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM J. Control Optim.*, 23(2):217–241, 1985.

[19] K. Kunisch and G. Peichl. Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented lagrangian technique. *Numerische Mathematik*, 59(1):473–509, 1991.

[20] P. K. Lamm. Estimation of discontinuous coefficients in parabolic systems: applications to reservoir simulation. *SIAM J. Control Optim.*, 25(1):18–37, January 1987.

[21] L. Ljung. Some results on identifying linear systems using frequency domain data. In *In Proc. 32nd IEEE Conference on Decision and Control*, pages 3534–3538, 1993.

[22] M. Mochi, G. Pacelli, M. C. Recchioni, and F. Zirilli. Inverse problem for a class of two-dimensional diffusion equations with piecewise constant coefficients. *Journal of optimization theory and applications*, 100(1):29–57, 1999.

[23] R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*. Wiley-IEEE Press, Hoboken (NJ), 2012.

[24] R. Pintelon, J. Schoukens, L. Pauwels, and E. Van Gheem. Diffusion systems: Stability, modeling, and identification. *IEEE Trans. Instrum. Meas.*, 54(5):2061–2067, 2005.

[25] R. Pintelon, J. Schoukens, and Y. Rolain. Uncertainty of transfer function modelling using prior estimated noise models. *Automatica*, 39(10):1721–1733, 2003.

[26] R. Pintelon, J. Schoukens, G. Vandersteen, and K Barbé. Estimation of nonparametric noise and FRF models for multivariable systems-Part I: Theory. *Mechanical Systems and Signal processing*, 24(3):573–595, 2010.

[27] A. D. Polyanin and V. F. Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*, volume 2. Chapman & Hall/CRC, London, 2003.

[28] E. Rafajlowicz. Optimal experiment design for identification of linear distributed-parameter systems: Frequency domain approach. *IEEE Trans. Autom. Control*, 28(7):806–808, Jul 1983.

[29] C. R. Rojas, J. S. Welsh, G. C. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43(6):993–1008, 2007.

[30] J. Schoukens and R. Pintelon. *Identification of linear systems: a practical guideline to accurate modeling*. Pergamon press New York, 1991.

[31] J. Schoukens, R. Pintelon, G. Vandersteen, and P. Guillaume. Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33(6):1073–1086, 1997.

[32] G. D. Smith. *Numerical solution of partial differential equations: finite difference methods*. Oxford University Press, 1985.

[33] T. Söderström. Errors-in-variables methods in system identification. *Automatica*, 43(6):939 – 958, 2007.

[34] T. Söderström and S. Halvarsson. Parameter estimation for diffusion models. Technical report, Dept. of Information Technology, Uppsala University, Sweden, 1999. http://www.it.uu.se/research/reports/1999-003/1999-003.pdf.

[35] T. Söderström, M. Hong, J. Schoukens, and R. Pintelon. Accuracy analysis of time domain maximum likelihood method and sample maximum likelihood method for errors-in-variables and output error identification. *Automatica*, 46(4):721–727, 2010.

[36] D. Ucinski. *Optimal measurement methods for distributed parameter system identification*. CRC Press Boca Raton (FL), 2004.

[37] D. Valério and J. S. da Costa. Identification of fractional models from frequency data. In *Advances in Fractional Calculus*, pages 229–242. Springer-Verlag, 2007.

[38] C. R. Vogel. *Computational methods for inverse problems*. SIAM Philadelphia (PA), 2002.