2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

A novel frequency domain maximum likelihood approach for estimating transport coefficients in cylindrical geometry for nuclear fusion devices

Matthijs van Berkel[1,2], Gerardus W. Oosterwegel[1,3], Martijn Anthonissen[4], Hans J. Zwart[5,6], Gerd Vandersteen[7]

Abstract— **This paper introduces a novel maximum likelihood approach to determine the local thermal transport coefficients belonging to diffusion and convection from excitation (perturbative) transport experiments. It extends previous work developed for linear (slab) geometry to cylindrical (toroidal) geometry for fusion reactors. The previous linear geometry approach is based on analytic solutions of the partial differential equation. However, for cylindrical geometries with convection the analytic solutions are confluent hypergeometric functions (CHFs) with complex valued arguments. Most numerical libraries do not support CHFs evaluation with complex valued arguments. Hence, this paper proposes the use of an ultra-fast transfer function evaluation based on sparse numerical solutions for the discretized partial differential equation. This solution is implemented in MATLAB© and incorporated in the frequency domain Maximum Likelihood Estimation framework. Consequently, transport coefficients can be estimated consistently when measurements are perturbed by coloured and spatially correlated noise.**

I. INTRODUCTION

Transport of heat and particles plays an important role in many applications in physics and particularly nuclear fusion. Although different descriptions exist for thermal transport and mass transport, in general the processes of diffusion, convection, and sometimes damping determine the global transport.

In nuclear fusion, the (electron) confinement time is, besides the temperature and density, the most important characterizing performance measure which needs to be optimized for performance [1]. The confinement time has a direct relationship to the transport coefficients and especially to the diffusion coefficient. In addition, to predicting the transport coefficients based on physical modeling [2], it is important to assess these transport coefficients experimentally. A very effective and unambiguous method is by exciting or perturbing the plasma around an operating point which allows to distinguish between the different transport processes. In this paper, we extend a methodology previously introduced for linear (slab) geometry [3] to cylindrical geometry such that it can be used to determine the transport coefficients in terms of the dimensionless minor radius for nuclear fusion reactors. This is highly necessary for regions close to the core where slab-approximations give erroneous results. The method is developed in the frequency domain and uses the concept of maximum likelihood estimation to handle the uncertainties in the measurements optimally.

As the identification and estimation of parabolic partial differential equations in cylindrical geometry extends over many fields such as mathematics (inverse problems) [4], control theory and system identification [5]–[7], and nuclear fusion [8], it is not possible to give an extensive overview here. Hence, we limit ourselves to frequency domain methods and specifically for nuclear fusion. The reason is that most heat sources (electron/ion cyclotron resonance heating, neutral beam injection, etc.) used for excitation experiments are block-wave modulated and as such the temperature perturbations are periodic and consist only of approximately 1-3 harmonic components with significant enough signal-to-noise ratios to be used for estimation purposes. Moreover, assuming that the PDE is linear, it can transformed to the frequency domain reducing to an ordinary differential equation (ODE) which significantly simplifies the estimation.

Other methods exist to estimate transport coefficients, such as, in nuclear fusion there is a focus on methods using only one harmonic/frequency component as this allows for a one-to-one inversion from amplitude and phase information to diffusion coefficient and convective velocity [8]–[10]. On the other hand, in the system identification literature the diffusion results in a non-rational transfer function (fractional form) which as such is taken into account [5], [11]–[13]. However, the two methods described above have in common that they do not take spatial cross-correlation of the noise on these measurements into account. Moreover, in nuclear fusion fitting procedures to continuate amplitude and phase between sensors is common practice. This both results in biased estimates (systematic errors) of the diffusion and convective velocity as noise is present on all temperature measurements. This is known in the literature as an errors-in-variables (EIV) problem [14]. Here, we further develop

[1]DIFFER-Dutch Institute for Fundamental Energy Research, Energy Systems and Control group, PO Box 6336, 5600HH Eindhoven, The Netherlands. E-mail: M.vanBerkel@differ.nl

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

[3]Eindhoven University of Technology, Science and Technology of Nuclear Fusion, Dept. of Applied Physics, Flux, PO Box 513, 5600MB, Eindhoven, The Netherlands

[4]Eindhoven University of Technology, Dept. of Mathematics and Computer Science, Postbus 513, MetaForum, 5600 MB, Eindhoven, The Netherlands

[5]Eindhoven University of Technology, Dept. of Mechanical Engineering, Dynamics and Control Group, PO Box 513, 5600MB Eindhoven, The Netherlands

[6]University of Twente, Dept. of Applied Mathematics, PO Box 217, 7500AE, Enschede, The Netherlands

[7]Vrije Universiteit Brussel (VUB), Dept. of Fundamental Electricity and Instrumentation, Pleinlaan 2, 1050 Brussels, Belgium

the maximum likelihood estimation (MLE) procedure such that it does not result in biased estimates [3]. Moreover, the MLE estimates are local, therefore, avoiding the spreading of errors over a large spatial region. The MLE approach can take multiple harmonics into account resulting in estimates with minimum variance. The work extends previous work developed for linear (slab) geometry [3], [15] to cylindrical (toroidal) geometry for fusion reactors. The linear geometry approach is based on analytic solutions of the partial differential equation. However, cylindrical geometry analytic solutions (confluent hypergeometric functions) are impractical to evaluate in the complex plane as it is not supported by most numerical libraries or are slow in its evaluation. Hence, in this paper the transfer functions and Jacobians that are analytic expressions in [3] are replaced by sparse matrix inversions. This is ultra-fast when only a few harmonic components need to be considered [15].

Procedural overview of our algorithm: 1) perform excitation experiment (here simulation) with a sufficiently small perturbation and (nearly) constant density; 2) estimate mean and co-variances from data; 3) choose/calculate starting values and transform in new parameter set; 4) optimize the MLE cost function using Levenberg-Marquardt; 5) retransform parameters to original set and apply propagation of uncertainty to calculate the confidence bounds.

This paper is structured as follows. Sec. II gives a brief overview of thermal transport in fusion reactors and introduces the partial differential equation under consideration. Then, Sec. III describes our new discretization approach. Sec. IV describes the maximum likelihood approach and how the noise is taken into account in the estimation. Sec. V gives some simulation results showing the merit of our approach. Finally, Sec. VI summarizes the approach and gives an outlook on further research.

II. Transport modelling in a fusion reactor

Thermal transport inside a fusion reactor is determined by turbulent transport. This is often characterized by axisymmetric transport due to the magnetic confined plasma topology with a diffusion coefficient which depends on various gradients such as the electron temperature gradient of ion gradient [16]–[18]. To study the turbulent transport a sufficiently small perturbation is applied such that the non-linear partial differential equation can be linearized around this operating point to characterize the transport. At this moment, this technique is mainly used to validate gyro-kinetic (theory which describes transport in tokamaks) predictions and to built empirical models for transport which but can be used for prediction and control. The linearized partial differential equation considered is

$$\frac{\partial}{\partial t}\left(nT\right) = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(n\rho\chi\left(\rho\right)\frac{\partial T}{\partial \rho} + n\rho V\left(\rho\right)T\right) - n\tau_{inv}\left(\rho\right)T + P_{ecrh}, \quad (1)$$

where χ is the diffusivity, V the convective velocity, τ_{inv} the (inverse) damping ($\tau_{inv} = 1/\tau$), T denotes the electron temperature, n the density, ρ the dimensionless radius, and P_{ecrh} a perturbative heat source which is spatially localized. For the estimation of χ, V, and τ_{inv}, the partial differential equation in (1) needs to be simplified.

The following assumptions are used to simplify (1):

- The latter density is actively controlled to remain constant over time and the temperature perturbation is sufficiently small to assume linearity of the experiment. Hence, within the local domains based on three sensors, we assume the transport coefficients and the density to be constant, i.e. $\frac{d\chi}{d\rho} = \frac{dV}{d\rho} = \frac{dn}{d\rho} = 0$.
- For the estimation only local domains are considered where there is no direct perturbed heating present, i.e. $P_{ecrh} = 0$. Hence, the perturbation due to the source acts from the outside on the three sensor domains. Note, that if static sources ($\partial/\partial t = 0$) are present they do not affect the estimation as long as the operating point stays the same.
- We consider (periodic) frequency domain measurements, i.e., spectra, instead of time domain data. Consequently, we assume that the transients due to the initial condition can be neglected, i.e., the system is in a regime with initial condition zero.

Under these assumptions (1) can be transformed into the Laplace domain yielding on our local domain

$$s\Theta\left(\rho,s\right) = \frac{1}{\rho}\frac{d}{d\rho}\left(\rho\chi\frac{\partial\Theta\left(\rho,s\right)}{\partial\rho} + \rho V\Theta\left(\rho,s\right)\right) - \tau_{inv}\Theta\left(\rho,s\right), \quad (2)$$

where $\Theta\left(\rho,s\right) = \mathfrak{L}\{T\left(\rho,t\right)\}$ with $\mathfrak{L}$ denoting the Laplace transform and s the Laplace variable. This complex valued ordinary differential equation (ODE) can be solved analytically as is shown in the next section.

A. Analytic solutions in frequency for cylindrical domains

In this section, we describe the analytic solutions for the transformed transport equation (2) and discuss its complexity when used in practice.

The analytic solution for the complex valued ODE in (2) can be derived, see [19], and is given by

$$\begin{aligned}\Theta\left(\rho,s\right) &= e^{\lambda_1\rho}D_1\Psi\left(\frac{\lambda_2}{\lambda_2-\lambda_1},1,\left(\lambda_2-\lambda_1\right)\rho\right) \\ &+ e^{\lambda_1\rho}D_2\Phi\left(\frac{\lambda_2}{\lambda_2-\lambda_1},1,\left(\lambda_2-\lambda_1\right)\rho\right), \\ \text{with}\qquad \lambda_{1,2} &= -\frac{V}{2\chi} \mp \sqrt{\left(\frac{V}{2\chi}\right)^2 + \frac{s+\tau_{inv}}{\chi}}, \end{aligned} \quad (3)$$

where $D_1\left(s\right)$ and $D_2\left(s\right)$ denote the boundary constants. The functions $\Phi\left(\rho,s\right)$ and $\Psi\left(\rho,s\right)$ denote the confluent hypergeometric functions of the first and the second kind, respectively [20]. For the special case that the convective velocity $V = 0$ these simplify to the Bessel functions [19].

Both the confluent hypergeometric functions and the Bessel functions are transcendental functions. This means

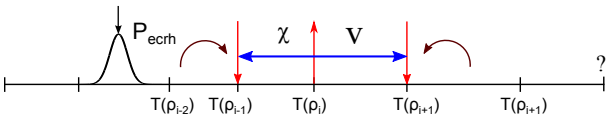

Fig. 1. Schematic overview of the three sensor domain. The circular arrows represent that the outside perturbations due to the source (P_{ecrh}) interact with the domain through the temperature measured at the edges of the domain. The direction of the red arrows represent the inputs (up-arrow) and output (down-arrow) related to the sensor location. The Question mark represents the unknown physical boundary condition.

that they are numerically approximated using special algorithms such as power series, asymptotic expansions, recursive algorithms, etc. In case of the Bessel functions, this is well developed [21]. However, for the confluent hypergeometric functions in (3) many programs only include standard power series approximations which are slow and often do not support evaluation of complex valued arguments. To make the method implementable for a large set of programs, in the next section a numerical approach to evaluate the solutions in case $V \neq 0$ is introduced.

B. Local domain based on three measurements

This section discusses the idea of using local domains briefly and is largely based on [3]. In this paper both an analytic approach and numerical approach is developed.

The solution of a partial differential equation is determined by its boundary conditions. However, in practice such boundary conditions are often not known precisely. Hence, a different approach was developed in [3] where instead of using physical boundaries, temperature measurements at sensor points are assumed to be the boundaries of a local domain. This allows to split the domain into smaller sub-domains for which the transport coefficients can be estimated. This has a number of advantages compared to domains consisting of more than 3 sensors:

1) For a consistent estimate with noisy measurements we require the MLE. This results in a non-convex optimization problem. Hence, splitting the optimization problem into small domains for which only three parameters need to be estimated makes the optimization problem tractable.
2) Signal-to-noise ratios tend to change depending on spatial locations, e.g., depending on the distance to the perturbation. Hence, errors do not propagate over space, but are contained locally.

The disadvantage is that within the local domain the transport coefficients are assumed to be constant. Hence, errors are introduced when the parameters vary significantly within this domain.

Using the temperature measurements at sensor locations as boundary conditions results in two Dirichlet boundary conditions, i.e., $T(\rho_{i-1})$ and $T(\rho_{i+1})$. In the frequency domain this is expressed as $\Theta(\rho_{i-1}, s)$ and $\Theta(\rho_{i+1}, s)$. This is graphically represented in Fig. 1.

These virtual boundary conditions can be used to analytically solve the temperature $\Theta(\rho_i, s)$ at any minor dimensionless radius within the local domain. The smallest local domains that can be constructed consist of three adjacent sensors as two sensors act as the two boundary conditions (second order PDE) and one sensor gives the resulting temperature used to estimate the transport coefficients. In other words, with m sensors, we have $i = 2, \ldots, m-1$ (excluding those with a source), the local domain is described by $[\rho_{i-1}, \rho_{i+1}]$. This is done by solving $D_1(s)$ and $D_2(s)$ in (3) using the Dirichlet boundary conditions. This results in the following multi-input single-output description for the temperature $\Theta(\rho_i, s)$ at the central sensor in the domain

$$\Theta(\rho_i) = \frac{\xi(\rho_{i+1})\zeta(\rho_i) - \zeta(\rho_{i+1})\xi(\rho_i)}{\zeta(\rho_{i-1})\xi(\rho_{i+1}) - \zeta(\rho_{i+1})\xi(\rho_{i-1})}\Theta(\rho_{i-1}) - \frac{\xi(\rho_{i-1})\zeta(\rho_i) - \zeta(\rho_{i-1})\xi(\rho_i)}{\zeta(\rho_{i-1})\xi(\rho_{i+1}) - \zeta(\rho_{i+1})\xi(\rho_{i-1})}\Theta(\rho_{i+1}), \quad (4)$$

where dependencies on the parameters and s have been omitted and ξ and ζ are defined as

$$\xi(\rho, s) = e^{\lambda_1 \rho}\Psi\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}, 1, (\lambda_2 - \lambda_1)\rho\right) \quad (5)$$

$$\zeta(\rho, s) = e^{\lambda_1 \rho}\Phi\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}, 1, (\lambda_2 - \lambda_1)\rho\right). \quad (6)$$

If we redefine (4) in terms of inputs and outputs, then we see that it consists of two inputs $\Theta(\rho_{i-1}, s)$ and $\Theta(\rho_{i+1}, s)$, and one output $\Theta(\rho_i, s)$ on the interval $[\rho_{i-1}, \rho_{i+1}]$, i.e.,

$$\Theta(\rho_i, s) = G_1(\theta, s)\Theta(\rho_{i-1}, s) - G_2(\theta, s)\Theta(\rho_{i+1}, s), \quad (7)$$

which are connected through the transfer functions G_1 and G_2 depending on the transport coefficients $\theta = [\chi, V, \tau_{inv}]$. As explained, analytic functions for ζ and ξ are unsuitable to use (see Sec. II-A) for the transport coefficients to be estimated. The next section, therefore describes a numerical computationally efficient implementation for G_1 and G_2.

III. Finite difference discretization approach

This section describes the approach to acquire the transfer functions G_1 and G_2 using a finite difference approximation of the complex valued ODE in (2).

A. Diffusion and damping

Here a finite difference model is derived for a local domain. The index of the nodes of the discretization grid is denoted by n such that ρ_{i-1} corresponds to $n = 1$ and ρ_{i+1} corresponds to $n = N$. For clarity, the index i denotes the sensor positions on the global domain. This means that T_1 will be the measured temperature at $T(\rho_{i-1})$ and T_N will be the measured temperature at $T(\rho_{i+1})$.

The explicit central difference scheme is applied such that the diffusive part of the ODE in (2) has the following

discretization

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\chi\rho\frac{\partial\Theta(\rho,s)}{\partial\rho}\right)\approx \chi\frac{\left(\rho_n-\frac{1}{2}\Delta\rho\right)\Theta_{n-1}-2\rho_n\Theta_n+\left(\rho_n+\frac{1}{2}\Delta\rho\right)\Theta_{n+1}}{\rho_n\,\Delta\rho^2} \quad (8)$$

evaluated for $n = 2,\ldots,N-1$. In principle, also the frequency needs to be discretized. However, as (2) is linear for every frequency defined as $s = i\Omega_k$, (8) can be evaluated independently. Consequently, the discretization over frequency with nodes k is omitted. The damping is a simple evaluation

$$-\tau_{inv}\Theta \approx -\tau_{inv}\Theta_n. \quad (9)$$

The next step is to express (8) in matrix form which results in a tri-diagonal matrix known as the Laplacian.

B. Convection

The spatial derivative with respect to the convective velocity part of (2) is discretized as follows for $V \geq 0$

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho V\Theta)\approx V\left(\frac{\Theta_n-\Theta_{n-1}}{\Delta\rho}+\frac{1}{2}\frac{\Theta_n+\Theta_{n-1}}{\rho_n-\frac{1}{2}\Delta\rho}\right), \quad (10)$$

whereas when $V < 0$

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho V\Theta)\approx V\left(\frac{\Theta_{n+1}-\Theta_n}{\Delta\rho}+\frac{1}{2}\frac{\Theta_{n+1}+\Theta_n}{\rho_n+\frac{1}{2}\Delta\rho}\right), \quad (11)$$

evaluated for $n = 2,\ldots,N-1$. The second terms in (10) and (11) are chosen such that the average uses the same information as the two-point difference instead of using only the temperature at one grid point which is also a valid choice. Moreover, two different implementations are used depending on the flow direction. These discretizations are combined to calculate the transfer functions.

C. State-space representation and boundary conditions

The spatial discretizations can be cast into a state-space representation with respect to the spatial coordinate ρ using boundary inputs and the middle sensor as output. This results in the following state-space representation where A, B, and C are independent of s

$$\begin{aligned} s\overrightarrow{\Theta} &= A\overrightarrow{\Theta} + B\begin{bmatrix}\Theta_1(\rho_{i-1})\\ \Theta_N(\rho_{i+1})\end{bmatrix} \\ \Theta(\rho_i) &= C\overrightarrow{\Theta}\end{aligned} \quad (12)$$

with $\overrightarrow{\Theta} = \mathrm{col}(\Theta_2,\ldots,\Theta_{N-1})$. The state matrix A with dimensions $(N-2)\times(N-2)$ can be written as a sum of the individual finite difference schemes presented in (8), (10) or (11), and (9), i.e.,

$$A = \chi A_\chi + VA_V + \tau_{inv}A_\tau. \quad (13)$$

The input matrix B corresponds to the boundary nodes due to the choice of Dirichlet boundary conditions at the bounds, i.e., $\Theta(\rho_{i-1}) = \Theta_1$ and $\Theta(\rho_{i+1}) = \Theta_N$. Hence, the B matrix has dimensions $(N-2)\times 2$ parametrized as $B = \left(B_l^{\pm},\, B_r^{\pm}\right)$. As the boundary nodes depend on the sign of the convective velocity, B changes when the sign of V changes. Hence, for $V < 0$

$$B_l^- = \mathrm{col}\left(\chi\,\frac{\rho_2-\frac{1}{2}\Delta\rho}{\rho_2\,\Delta\rho^2},0,\ldots,0\right), \quad (14)$$

$$B_r^- = \mathrm{col}\Bigg(0,\ldots,0,\chi\,\frac{\rho_{N-1}+\frac{1}{2}\Delta\rho}{\rho_{N-1}\,\Delta\rho^2} + V\left(\frac{1}{2\rho_{N-1}+\Delta\rho}+\frac{1}{\Delta\rho}\right)\Bigg). \quad (15)$$

and for $V \geq 0$

$$B_l^+ = \mathrm{col}\Bigg(\chi\frac{\rho_2-\frac{1}{2}\Delta\rho}{\rho_2\,\Delta\rho^2} + V\left(\frac{1}{2\rho_2-\Delta\rho}-\frac{1}{\Delta\rho}\right),0,\ldots,0\Bigg). \quad (16)$$

$$B_r^+ = \mathrm{col}\left(0,\ldots,0,\chi\,\frac{\rho_{N-1}+\frac{1}{2}\Delta\rho}{\rho_{N-1}\,\Delta\rho^2}\right), \quad (17)$$

which is a clear reflection of the single boundary condition necessary for the convective part. The sensor matrix C has dimension $1\times(N-2)$ and is sparse with only one non-zero entry equal to 1 where $\rho_n = \rho_i$. This state-space model can be efficiently solved to find G_1 and G_2 as is discussed in the next section.

D. Frequency domain sparse inversion and their Jacobians with respect to the transport coefficients

The approximated transfer functions are derived here based on the state-space representation in (12). In the case of experimental data from perturbative experiments only a few harmonic components are excited. Hence, the computation should only be performed for these harmonics. We only compute $G_1(\theta,\Omega_k)$ and $G_2(\theta,\Omega_k)$ at the excited harmonics Ω_k, i.e., $s = i\Omega_k$. This results in the well-known relationship

$$\begin{bmatrix}G_1(\theta,\Omega_k)\\ G_2(\theta,\Omega_k)\end{bmatrix} = C\left(\Omega_k I - A(\theta)\right)^{-1}B(\theta), \quad (18)$$

with I the identity matrix of size $(N-2)\times(N-2)$ and θ the unknown thermal transport coefficients to be estimated. These numerical transfer functions can be calculated efficiently as only a matrix inversion based on sparse matrices needs to be performed, i.e, A is tri-diagonal, B has only two non-zero elements, and C has only one non-zero element. Moreover, the presented algorithm is implemented in MATLAB© efficient in solving sparse linear matrices.

An efficient evaluation of the transfer functions and hence the cost function is crucial for estimation of the thermal transport coefficients. In addition, an analytic Jacobian also significantly decreases the computational time of the optimization procedure discussed in the next section. The chosen state-space representation simplifies the derivation of the derivatives of (18), i.e.,

$$\begin{bmatrix}\frac{\partial G_1(\theta,\Omega_k)}{\partial\theta}\\ \frac{\partial G_2(\theta,\Omega_k)}{\partial\theta}\end{bmatrix} = C\left(\Omega_k I - A(\theta)\right)^{-1}\frac{\partial B(\theta)}{\partial\theta} + C\left(\Omega_k I - A(\theta)\right)^{-1}\frac{\partial A(\theta)}{\partial\theta}\left(\Omega_k I - A(\theta)\right)^{-1}B(\theta) \quad (19)$$

where the derivatives of A are given by

$$\frac{\partial A}{\partial \chi} = A_\chi,\ \frac{\partial A}{\partial V} = A_V,\ \frac{\partial A}{\partial \tau_{inv}} = A_\tau; \tag{20}$$

and that of B by

$$\frac{\partial B}{\partial \chi} = \left[B_l^- (\chi = 1), B_r^+ (\chi = 1)\right],\ \frac{\partial B}{\partial \tau_{inv}} = 0 \tag{21}$$

and

$$\frac{\partial B^-}{\partial V} = \left[0_{(N-2)\times 1},\ \begin{matrix} 0_{(N-3)\times 1} \\ \frac{1}{2\rho_N\ _1 + \Delta\rho} + \frac{1}{\Delta\rho} \end{matrix}\right], \tag{22}$$

or

$$\frac{\partial B^+}{\partial V} = \left[\begin{matrix} \frac{1}{2\rho_1 - \Delta\rho} - \frac{1}{\Delta\rho} \\ 0_{(N-3)\times 1} \end{matrix}, 0_{(N-2)\times 1}\right]. \tag{23}$$

As also the derivatives are sparse only one additional inversion is calculated corresponding to the combination $C(\Omega_k I - A(\theta))^{-1}$. Both the numerical transfer functions and its Jacobians with respect to the parameters are used in the estimator introduced in the next section.

IV. MAXIMUM LIKELIHOOD ESTIMATION

This section gives a brief overview of the maximum likelihood approach for three sensors as presented in [3] and [22]. The main theory behind maximum likelihood estimation is extensively described in [23]. The maximum likelihood estimation method is used to estimate the diffusivity, convective velocity, and the damping based on both the analytical implementation ($V = 0$) and the previously introduced numerical approximation of the transfer functions.

The maximum likelihood estimation requires, in addition to the input parameters, also information about the noise in terms of its co-variances. Moreover, the optimization and confidence bound calculation are briefly discussed.

A. Maximum Likelihood Estimation

Maximum likelihood estimation is a method for estimating parameters given some statistical model, where the maximum likelihood estimator maximizes a known likelihood function [24]. This function can be interpreted as a probability density function (PDF) but with respect to the measured data.

In the maximum likelihood estimator the parameters are adapted such that the residue will resemble the likelihood function as good as possible. This results in the parameter estimates with the highest possible accuracy given the underlying assumptions and a specific measurement set [23].

For the PDF of the noise, it is assumed that every harmonic has a zero mean additive circular complex normal distribution (CCND), which is a two-dimensional Gaussian distribution. The distribution of the Fourier coefficients naturally converges to this CCND for increasing number of time samples due to the central limit theorem. This assumption is verified for ECE-measurements in [25].

Maximizing the likelihood function is rather complicated due to the exponent of the CCND (see [23]). Instead, of maximizing the PDF the negative log likelihood function V_{MLE} (taking the logarithm of the PDF and adding a minus sign) is minimized [23]. As the log-function is a monotonic function, it does not change the minima of the likelihood function. Consequently, the parameters $\theta = [\chi, V, \tau_{inv}]$ are estimated by minimizing V_{MLE}

$$\hat{\theta} = \arg\min_\theta V_{MLE}(\Omega_k, \theta), \tag{24}$$

resulting in the parameter estimates $\hat{\theta}$, where the hat denotes the estimated parameters. The V_{MLE} cost function can also be written as

$$V_{MLE} = \frac{1}{F}\sum_{k=1}^{F} \frac{\left|\hat{\Theta}(\rho_i, \Omega_k) - \Theta(\rho_i, \Omega_k)\right|^2}{\sigma_e^2(\theta, \Omega_k)}, \tag{25}$$

with

$$\Theta(\rho_i, \Omega_k) = G_1(\theta, \Omega_k)\hat{\Theta}_k(\rho_{i-1}) - G_2(\theta, \Omega_k)\hat{\Theta}_k(\rho_{i+1}), \tag{26}$$

where $\hat{\Theta}$ denotes the measured Fourier coefficients and F the total number of harmonics used. The variability is given by (see [3])

$$\begin{aligned}\sigma_e^2(\omega_k, \theta) = {} & \sigma_i^2 + \sigma_{i-1}^2 |G_1|^2 + \sigma_{i+1}^2 |G_2|^2 \\ & - 2\,\mathrm{Re}\left(G_1 \sigma_{i-1,i+1}^2 \overline{G_2} + \sigma_{i,i-1}^2 \overline{G_1} - \sigma_{i,i+1}^2 \overline{G_2}\right),\end{aligned} \tag{27}$$

where σ_j^2 and $\sigma_{j,i}^2$ are the (co-)variances of the measured Fourier coefficients [25]. The $\overline{G_1}$ denotes the complex conjugate of G_1. Only considering (26) without (27) results in the non-linear least-squares estimator. The maximum likelihood estimator modifies the least-squares estimator via the variability σ_e. The variability σ_e acts as a natural weighting of the harmonics. In other words, harmonics with a small variance have more weight than harmonics with a high variance. Consequently this results in a non-convex minimization problem. The weighting is fully automated and consequently there is no need for additional tuning. The resulting parameter estimates are consistent, i.e., if the number of frequencies $k \to \infty$, then $\hat{\theta} \to \theta$ [3], [23].

The variability σ_e takes into account information on the (co-)variances to assure the consistency of the estimator. The necessary (co-)variances for the MLE can be estimated from measurements which is explained next.

B. Optimization

The cost function in (25) needs to be minimized with respect to the parameters θ. In principle, the minimization problem is non-convex due to the variability in the denominator of V_{MLE}, i.e., $\sigma_e(\theta, \Omega_k)$, which depends on the parameters. As such non-linear optimization techniques are necessary, but for reasonable noise levels, a gradient-based method suffices to find the global minimum. The region of convergence can also be even further extended using a change of parameters as shown in [3]. If we combine this parameter change with the analytical Jacobian based on the sparse matrix inversions of the state-space model in Sec. III-C, the region of convergence, computational efficiency, and stability of the optimization scheme are significantly increased. The gradient based method used here is the Levenberg-Marquardt method, which is a modified Newton-Gauss gradient method [26]. However, as the optimization

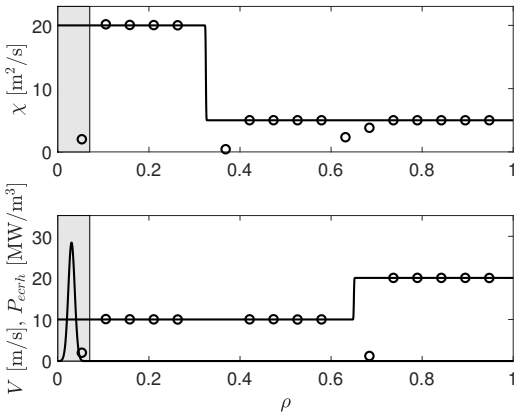

Fig. 2. The true profiles (solid lines) end estimates of χ and V (circles). The estimate is always plotted at the central measurement ρ_i of the local domain and are generated with a finite difference model. The simulation boundary conditions are $\partial\Theta/\partial\rho(\rho = 0) = 0$ and $\Theta(\rho = 1) = 0$. Note that some estimates are outside the plotted domain where there are gradients in the diffusive and or convective profile.

per local domain only requires the estimation of maximally three coefficients any search algorithm suffices to find the global minimum.

V. Simulation results

In this section, we present some results based on the introduced estimation procedure.

A. Validation of numerical and analytic implementation

First, we show that both the diffusion coefficient and convective velocity can correctly be estimated when the parameters are simultaneously locally constant. This is shown in Fig. 2. This example is deliberately chosen similarly to the example for linear geometry in [3]. It shows that indeed for the constant domains χ and V can be estimated correctly. However, when the gradients in the thermal coefficients become strong, the estimates fail unlike in the linear geometry example in [3], where still the diffusion could be estimated correctly. At this moment it is unclear why there is such a strong dependency on cylindrical geometry. This is not disastrous as validation tools exist to detect these errors [22].

When the gradient in the spatial varying coefficients become smoother, then it is also possible to estimate spatially varying transport coefficients.

As the Bessel functions in Sec. II-A can only be used for $V = 0$, we compare the new Bessel function approach to the new finite difference approximation in a diffusion only simulation. The comparison is shown for a spatially varying diffusion profile in Fig 3. We observe that both the finite difference approximation and Bessel functions result in the same estimates. As constant domains are assumed errors increase significantly when estimating V and are devastating for the τ_{inv} estimates as was already observed in Fig. 2 and [3]. Here, we have deliberately chosen not to add noise to study the new implementation of the algorithm. If noise is added, then very similar results can be observed as in [3].

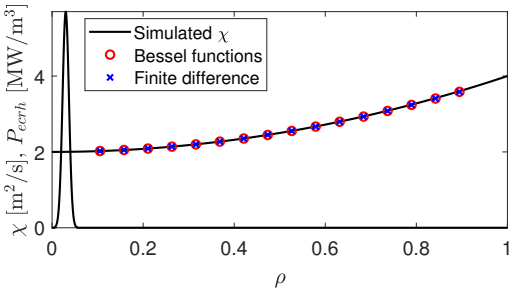

Fig. 3. Comparison of the estimates between (new) Bessel function implementation and the new finite difference implementation tested on a spatially varying diffusion profile in cylindrical geometry. As initial guess χ = 5 [m^2/s] is used. The perturbative source (P_{ecrh}) is localized and is modulated with a block-wave of 25 Hz and 75% duty-cycle.

Instead, we decided to analyze noise in more detail through a Monte-Carlo analysis.

B. Monte-Carlo verification

The primary advantage of the maximum likelihood estimator (MLE) over the least-squares estimator (LSE) is that its estimates are consistent (see Sec. IV-A) and that the MLE gives an estimate of the confidence bounds (CB) as is described in [3], [23]. This can be analyzed in detail using a Monte-Carlo analysis. To test and visualize this, one local-domain ($\chi = 5$, $V = \tau_{inv} = 0$, $[\rho_{i-1} = 0.1, \rho_i = 0.2, \rho_{i+1} = 0.3]$) is simulated and noise is added using a full co-variance matrix with cross-correlation. We use 10^4 realizations of the noise such that we can make 10^4 estimates using a parameter transformation $\hat{c} = 1/\hat{\chi}$ and the corresponding estimate of the CB. One estimate takes here approximately 5 ms but is strongly dependent on its initialization value. These can then be compared to the statistical properties of the Monte-Carlo estimates. This is shown in Fig. 4.

Both the MLE and LSE estimation was initialized at $\chi = 53$ and hence we see that the MLE finds a value close to the correct χ. The simulated value of χ = 5 where the estimated value with the MLE is $\hat{\chi}_{MLE}$ = 4.98 and that of the least-squares estimator $\hat{\chi}_{LSE}$ = 4.15. This means that the MLE and the LSE for this example have a relative error of 0.4% and 17%, respectively. Hence, the MLE significantly outperforms the LSE especially for increased noise levels.

We also observe a significant difference between the predicted variance of the estimates and the actual variance based on the MLE (in terms of CB). The reason is that for the propagation of uncertainty a first order Taylor expansion is used for the calculation of the MLE variances. However, in reality the transformation of variance is (strongly) non-linear and results in non-Gaussian distribution functions. This is to some extend true for the distribution of $1/\chi$. This is even more apparent when the histogram of χ is plotted. This results in a distribution function which is clearly non-Gaussian and requires non-symmetric CB.

In case the noise levels decrease both the histograms of $1/\chi$ and χ become more Gaussian-like and hence also

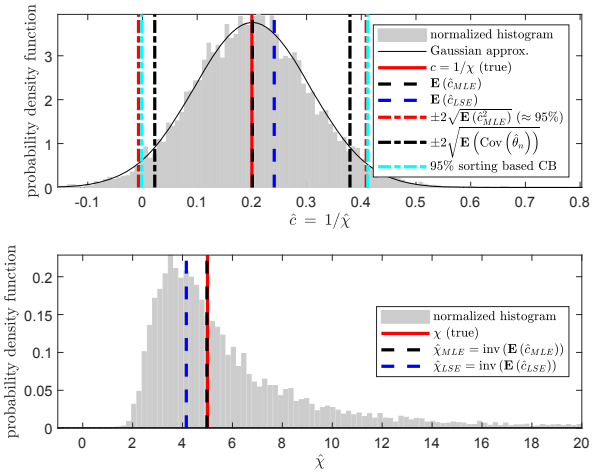

Fig. 4. The grey bell-curve is the histogram for the MLE estimates. The expected value $\mathbf{E}$ is calculated for the MLE and compared to that of the LSE (blue) and compared against the correct χ in black ($\mathbf{E}$ of the histogram). The CB are estimated based on 1) the average of the estimated covariances of the MLE (black); 2) using the variance of the histogram (red); 3) based on sorting the 10^4 realizations (cyan). The CB for the LSE are not shown as they will be strongly biased due to the EIV-problem.

the variance predictions of the MLE converge to the true variances. Nevertheless, we see that the MLE estimates converge to the true value of χ.

VI. CONCLUSION & OUTLOOK

In this paper, we presented a novel approach to estimate the transport coefficients in cylindrical geometry which is an extension of the linear geometry case presented in previous work. We derived an analytic approach based on Bessel functions and a numerical approach based on finite difference approximations of the transfer functions. Both are integrated in the maximum likelihood framework showing significant better performance compared to least-squares estimation for high noise levels. Although the estimation method is derived for local domains that assume spatially constant transport coefficients, it can be used to estimate varying profiles but this results in some errors. Especially, errors in the convective velocity requires further investigation and quantification. Moreover, for high noise levels the confidence bounds predicted by the MLE are erroneous and hence either a higher order Taylor expansion with respect to confidence bounds needs to implemented or alternatively, given the ultra-fast MLE estimates, the confidence bounds can be constructed through a Monte-Carlo analysis. At this moment, this algorithm is foreseen to be used offline only. Finally, as the perturbative source can be uncertain simultaneous estimation of the source domain and diffusion coefficient is currently being developed.

REFERENCES

[1] J. P. Freidberg, *Plasma Physics and Fusion Energy*. Cambridge University Press, Cambridge, 2007.

[2] J. Citrin, C. Bourdelle, *et al.*, "Multi-channel integrated modelling with the quasilinear gyrokinetic transport model qualikiz," *under revision in Plasma Phys. Control. Fusion*, 2017.

[3] M. van Berkel, G. Vandersteen, E. Geerardyn, R. Pintelon, H. J. Zwart, and M. R. de Baar, "Frequency domain sample maximum likelihood estimation for spatially dependent parameter estimation in PDEs," *Automatica*, vol. 50, no. 8, pp. 2113 – 2119, 2014.

[4] H. T. Banks and K. Kunisch, "Estimation techniques for distributed parameter system," *Birkhaäuser, Boston*, 1989.

[5] R. Pintelon, J. Schoukens, L. Pauwels, and E. Van Gheem, "Diffusion systems: Stability, modeling, and identification," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 5, pp. 2061–2067, 2005.

[6] S. Mechhoud, E. Witrant, L. Dugard, and D. Moreau, "Estimation of heat source term and thermal diffusion in tokamak plasmas using a kalman filtering method in the early lumping approach," *IEEE T. Contr. Syst. T.*, vol. 23, no. 2, pp. 449–463, 2015.

[7] B. Mavkov, E. Witrant, and C. Prieur, "Distributed control of coupled inhomogeneous diffusion in tokamak plasmas," *IEEE Transactions on Control Systems Technology*, vol. 27, pp. 443–450, Jan 2019.

[8] A. Jacchia, P. Mantica, F. De Luca, and G. Gorini, "Determination of diffusive and nondiffusive transport in modulation experiments in plasmas," *Phys. Fluids B-Plasma*, vol. 3, no. 11, pp. 3033–3040, 1991.

[9] D. F. Escande and F. Sattin, "Calculation of transport coefficient profiles in modulation experiments as an inverse problem," *Phys. Rev. Lett.*, vol. 108, no. 12, p. 125007, 2012.

[10] F. Ryter, R. Dux, P. Mantica, and T. Tala, "Perturbative studies of transport phenomena in fusion devices," *Plasma Phys. Control. Fusion*, vol. 52, p. 124043, 2010.

[11] J.-D. Gabano and T. Poinot, "Fractional modelling and identification of thermal systems," *Signal Process.*, vol. 91, no. 3, pp. 531 – 541, 2011.

[12] 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*, pp. 1–6, 2011.

[13] D. Valério and J. S. da Costa, "Identification of fractional models from frequency data," in *Advances in Fractional Calculus*, pp. 229–242, Springer-Verlag, Berlin-Heidelberg, 2007.

[14] T. Söderström, "Errors-in-variables methods in system identification," *Automatica*, vol. 43, no. 6, pp. 939 – 958, 2007.

[15] A. Das, S. Weiland, and M. Van Berkel, "Frequency domain estimation of spatially varying parameters in heat and mass transport," in *2019 American Control Conference (ACC)*, pp. 600–605, July 2019.

[16] M. van Berkel, H. J. Zwart, N. Tamura, G. M. D. Hogeweij, S. Inagaki, M. R. de Baar, and K. Ida, "Explicit approximations to estimate the perturbative diffusivity in the presence of convectivity and damping I Semi-infinite slab approximations," *Phys. Plasmas*, vol. 21, p. 112507, 2014.

[17] N. J. Lopes Cardozo, "Perturbative transport studies in fusion plasmas," *Plasma Phys. Control. Fusion*, vol. 37, p. 799, 1995.

[18] S. P. Eury, E. Harauchamps, X. Zou, and G. Giruzzi, "Exact solutions of the diffusion-convection equation in cylindrical geometry," *Phys. Plasmas*, vol. 12, p. 102511, 2005.

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

[20] L. Slater, *Confluent hypergeometric functions*. Cambridge University Press, Cambridge, 1960.

[21] D. Amos, "Algorithm 644: A portable package for bessel functions of a complex argument and nonnegative order," *Acm. T. Math. Software*, vol. 12, no. 3, pp. 265–273, 1986.

[22] M. van Berkel, G. Vandersteen, H. J. Zwart, G. M. D. Hogeweij, and M. R. de Baar, "Maximum likelihood estimation of diffusion and convection in tokamaks using infinite domains," in *IEEE Multi-conference on Systems and Control*, pp. 1230–1234, 2013.

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

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

[25] M. van Berkel, H. J. Zwart, G. M. D. Hogeweij, G. Vandersteen, H. van den Brand, M. R. de Baar, and the ASDEX Upgrade Team, "Estimation of the thermal diffusion coefficient in fusion plasmas taking frequency measurement uncertainties into account," *Plasma Phys. Control. Fusion*, vol. 56, p. 105004., 2014.

[26] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM Journal on Applied Mathematics*, pp. 431–441, 1963.