A systematic approach to optimize excitations for perturbative transport experiments

M. van Berkel[1,2], A. de Cock[3], T. Ravensbergen[1,2], G.M.D. Hogeweij[1], H.J. Zwart[4,5], G. Vandersteen[3]

[1]*DIFFER-Dutch Institute for Fundamental Energy Research,*
PO Box 6336, 5600 HH Eindhoven, The Netherlands
[2]*Eindhoven University of Technology, Dept. of Mechanical Engineering,*
Control Systems Technology, PO Box 513, 5600MB Eindhoven, The Netherlands
[3]*Vrije Universiteit Brussel (VUB), Dept. of Fundamental Electricity and Instrumentation, Pleinlaan 2, 1050 Brussels, Belgium*
[4]*University of Twente, Dept. of Applied Mathematics,*
PO Box 217, 7500AE, Enschede, The Netherlands and
[5]*Eindhoven University of Technology, Dept. of Mechanical Engineering,*
Dynamics and Control, PO Box 513, 5600MB Eindhoven, The Netherlands

ABSTRACT

In this paper, techniques for optimal input design are used to optimize the waveforms of perturbative experiments in modern fusion devices. The main focus of the paper is to find the modulation frequency for which the accuracy of the estimated diffusion coefficient is maximal. Mathematically this problem can be formulated as an optimization problem in which the Fisher information matrix is maximized. First this optimization problem is solved for a simplified diffusion model, while assuming a slab geometry and a semi-infinite domain. Later, the optimization is repeated under more general conditions such as a cylindrical geometry, finite domain, and simultaneous estimation of multiple transport coefficients. Based on the results of these optimizations, guidelines are offered to select the modulation frequency and to determine the optimality of the corresponding experiment.

I. INTRODUCTION

Since many years perturbative experiments have been used to study the transport of heat and particles in tokamaks and stellarators [1–3]. These experiments allow to measure quantities that in steady-state cannot be identified separately [2]. Some of these important quantities that can be identified through such experiments are the transport coefficients, which can be calculated based on the Fourier spectra of the measured time-traces [4, 5]. However, how well the transport coefficients can be determined strongly depends on experimental conditions [6]. Some of these conditions are fixed by the physical properties of the measurement setup, e.g., the thermal noise level, the unknown transport coefficient and the non-linear distortions, while others can be influenced by the expermentalist, e.g., the total available power,

the deposition profile, waveform of the heat source, and sometimes the location of the sensors.

Performing (perturbative) experiments in the field of nuclear fusion is often very costly, due to restricted machine availability to perform experiments. As a result, a lot of effort and time is put into determining the experimental conditions which optimize the use of the measurement resources. In the plasma fusion community determining the optimal experimental conditions is often done based on first order principles, heuristics, and experience of experimentalists. In contrast, the field of system identification provides a vast library of methods to design the optimal experiment in a more systematic way for linear [7–9] and non-linear systems [10–12]. These methods are based on model knowledge, which is rarely used explicitly to optimize the perturbations.

The key idea of optimal experiment design in the system identification community is to minimize estimation error on the measured quantities with respect to the experimental design choices. In general, the estimation error is determined by the bias (systematic error) and covariance (stochastic error) of estimated parameters [13, 14]. However, in the field of optimal input design it is common to assume that the estimator is asymptotically unbiased and efficient [15]. Hence, minimization of the estimation error reduces to the minimization of the covariance matrix. Moreover, the assumption of an asymptotic unbiased and efficient estimator allows us to approximate the covariance matrix only with the inverse of the Fisher information matrix. This matrix can be computed prior to the experiment based on the model, noise distribution, and prior guess of the model parameters [14, 16].

In theory the presented optimal input design can be extended to biased estimators by replacing the Fisher information matrix with an analytic expression of the mean square error matrix (which depends on the expression for the bias and covariance matrix). However, if an analytic expression for the bias is available, one can simply use this information to correct the estimation, which reduces the problem back to an unbiased problem (be it with an altered expression for the Fisher information matrix). The two cases where this is not possible is when the bias depends on the measured quantities or no

analytical expression is available. Unfortunately, these conditions also prohibit the computation of the optimal input prior to the experiment.

In principle the above methodology can be applied to optimize any controllable aspect of a perturbative experiment such that the uncertainty of the measured quantities is minimized. However, in this paper the main focus lies on computing the modulation frequency of the heat source that is used to generate the perturbation of the experiment. Often the range for modulation frequency is set by physical restrictions (see Sec. II for more details). However, to the best of the authors knowledge the relation between the modulation frequency and the uncertainty of the estimation has not yet been explored.

Initially it will be assumed that only the diffusion coefficient is estimated during the experiment. This means that the Fisher information matrix becomes a simple scalar. By reducing the dimension of the optimization problem, it will be possible to derive an analytical expression for the optimal modulation frequency given some simplifying assumption. Later, it will be explained how the method can be extended for simultaneous estimation of multiple transport coefficients.

The remainder of this paper is organized as follows. First, in Sec. II an overview is given from the physical consideration that are made when choosing the waveform of a perturbative experiment. Next, in Sec. III an introductory example is presented, and the concept of the Fisher information matrix is introduced. Additionally, it is explained how the Fisher information matrix can be used to optimize an experiment. Then, in Sec. IV, the optimal modulation frequency is computed for various conditions and modulation waveforms. Finally, in Sec. VI it is explained how the method can be extended when multiple transport coefficients are estimated, as well as how the spectra of complex waveforms can be optimized.

II. PHYSICAL RESTRICTIONS ON THE MODULATION FREQUENCY

In most perturbation experiments a periodic blockwave is used to modulate the heat source. Since these excitations are periodic, i.e., they have a fixed modulation frequency (f_{mod}), the relevant perturbation can be extracted from the experimental data with correlation methods, e.g., using the fast Fourier transform (FFT). This improves the signal-to-noise ratio and allows one to minimize the magnitude of the perturbation. Due to the well-defined frequency f_{mod}, the FFT exhibits a frequency spectrum with narrow peaks exactly at f_{mod}, and possibly higher harmonics, depending on the modulation scheme. At each radial position of the measurement, one gets FFT's amplitude and phase at f_{mod} and its higher harmonics. The ensemble of measurement points provides radial profiles of amplitude and phase whose shape is determined by the sources and by the propagation of the perturbation.

One of the strengths of perturbative experiments is that they can provide a separation between diffusive and non-diffusive (convective) contributions [1], which cannot be separated in a steady-state (local power balance) analysis. An example of such separation is the detection of a heat pinch in various experiments [17, 18]. As the effect of non-diffusive (convective) terms on the pulse propagation diminishes with increasing f_{mod}, it is very useful to be able to analyze higher harmonics of the ground frequency. The choice of a non-standard (i.e. far away from 50%) duty cycle helps to enhance the signal of higher harmonics. The relative strength of higher harmonics can be further enhanced by choosing complicated waveforms. If one has modulated sources at different modulation frequencies, say $f_{mod,i}, i = 1, 2$, the interacting modulations may produce perturbations at some beat frequencies $m f_{mod,1} + n f_{mod,2}$, which can provide additional transport information. This can be done using natural interaction between sawtooth and modulated electron cyclotron heating (ECH) [19] or using directly two modulated ECH sources [20, 21]. Analyzing higher harmonics can specifically be used to test the linearity of the experiment, which is a necessary condition when comparing to linearized physics models.

When designing the waveform of the modulation, it is also important to consider that f_{mod} and the amplitude of the waveform determine the modulation depth and linearity of the experiments. One wants to avoid non-linearities as for the estimation of transport coefficients linearized models are used. Non-linear effects are usually present due to a large perturbation in the non-linear heat flux relation or by modulating a non-linear boundary condition. As such, we want to avoid perturbations that lead us too far from the equilibrium. Additionally, the excitation should be small enough to minimize perturbation of quantities other than the one to be studied.

On the other hand, one wants to be able to analyze the perturbation over a significant radial range of the plasma. For pure diffusive propagation in a region free of the modulated source, the amplitude profile decreases exponentially with a decay length, $\lambda \sim \sqrt{\chi/f_{mod}}$ (where χ is the diffusion coefficient of the perturbed quantity). The amplitude profile should be large enough to allow for a measurable amplitude in the region of interest, but remain smaller than the typical plasma size to avoid the influence of the plasma boundary which might prevent interpreting the transport results locally. Moreover, when the perturbation is settled (or almost settled) to an equilibrium, decreasing the modulation frequency no longer increases the amount of information in the measurement. As a rule of thumb this lower bound on f_{mod} is similar to the inverse of the confinement time $1/\tau_E$. However, as this settling down time relates to many aspects it is not entirely clear how accurate it is. In conclusion under given plasma conditions, these requirements determine the best range for f_{mod}.

A last well-known phenomenon in heat plasma's that should be taken into consideration, is the existence of

critical gradients. Above a certain threshold of the inverse scale length $(\nabla X/X)_{crit}$ of quantity X (e.g. electron temperature T_e) some type of turbulence is triggered, which makes it very hard to further enhance the inverse scale length. This phenomenon is known as stiff profiles. Perturbative experiments are very suited to study this type of phenomenon. Below the threshold generally $\chi^{pert} = \chi$. Above threshold χ^{pert} is (much) larger than χ^{LPB}, and χ^{pert} is a measure for stiffness [1, 22, 23]. The original critical gradient models have been build based on linearization of transport around different operating points [24, 25] for which we will optimize the modulation frequency in this paper. At exactly the transition point, the so-called knee, the linearization may not exist due to the discontinuous behaviour of the critical gradient. If during the perturbative experiment this discontinuity lies in the domain of the perturbation. Then, new higher harmonic components will appear showing that the linearization does not hold and the estimates of χ^{pert} are unreliable. The appearence of higher harmonics due to non-linear behaviour is discussed in more detail in Sec. VI B.

III. HOW TO OPTIMIZE THE PERTURBATION?

In this section, the basic concepts to optimize the modulation signal are explained. First, an introductory example is given to get a better understanding of what determines the optimal modulation frequency. Next, the concept of the Fisher information matrix is introduced, and it is explained how this matrix can be used to determine the optimal modulation signal. To conclude this section, an expression for the Fisher information matrix is derived for the specific case were we want to optimize the modulation frequency in order to reduce the uncertainty on the estimate diffusion coefficient.

A. Introductory example

Assume that we want to optimize the modulation frequency of the perturbation for a linearized transport model depicted in Fig. 1 , which is described by

$$n_e \frac{\partial T_e(t,x)}{\partial t} = n_e \chi \frac{\partial^2 T_e(t,x)}{\partial x^2} + p_{ech}(t,x), \quad (1)$$

where

$$p_{ech}(t,x) = p(t)\frac{1}{a\sqrt{\pi}}\exp\left(-\frac{(x-x_{dep})^2}{a^2}\right), \quad (2)$$

with boundary conditions $\partial T_e/\partial x\,(x=0) = 0$ and at the boundary x_{end}, $T(x_{end}=\infty) = 0$.

Moreover, a localized Gaussian deposition profile is considered with a chosen to be small. The other quantities in this model are T_e the temperature, x the spatial

coordinate, constant density n_e, χ the diffusion coefficient, and the source $p_{ech}(t,x)$ with the center of deposition x_{dep} and dispersion a. Note, that the possible static terms in (1) do not need to be taken into account in an perturbative analysis [2].

Assuming that the perturbative source is localized, the diffusion coefficient can be estimated outside the source domain, i.e., where, see (2), contributions become negligible. The estimation of the transport coefficients such as χ is then generally determined on a local domain between two temperature measurement locations (x_1 and x_2) [4]. The solution of (1) can be calculated analytically in the frequency domain assuming constant transport coefficients and is given by [5]

$$G(\omega,\chi) = \frac{\Theta(\omega,x_2)}{\Theta(\omega,x_1)} = \exp\left(-\sqrt{\frac{i\omega}{\chi}}\Delta x\right), \quad (3)$$

where $\Theta(\omega,x) = \mathcal{F}(T_e(t,x))$ in which $\mathcal{F}$ denotes the Fourier transform, Δx represents the distance between x_1 and x_2, and $G(\omega,\chi)$ is referred to as the transfer function which models the relation between the input $\Theta(\omega,x_1)$ and the output $\Theta(\omega,x_2)$. It is chosen to represent the solution in transfer function form [5] instead of the more common form in terms of logarithmic spatial derivatives A'/A and ϕ' [3–5] because it significantly simplifies the calculation of the optimized perturbation signals later.

It is important to note that $G(\omega,\chi)$ does not depend on the source, but only describes how $\Theta(\omega,x_2)$ changes as result of $\Theta(\omega,x_1)$ because $x_1 \leqslant x$, $p_{ech}(x) \approx 0$. In Fig. 2 the dependence of $G(\omega,\chi)$ on the frequency is presented for two different values of the diffusion coefficient $\chi = 1$ and $\chi = 10$. Notice, that aside from the frequency and the transport coefficient χ, the transfer functions in (3) also depends on the sensor locations x_1 and x_2. However, these are considered to be known prior to the experiment.

The goal of the experiment is to estimate the transport coefficient χ based on temperature measurements $T(x_1)$ and $T(x_2)$. It is assumed that the temperature measurements are corrupted by additive Gaussian noise, which results in complex circular normal noise (CCND) in the frequency domain [26] with variance σ_ω^2. As a result the exact transfer function cannot be obtained during the experiment. Instead a noisy version of the transfer function is obtained, as represented by the point clouds in Fig. 2.

The optimal modulation frequency corresponds to the frequency for which it is the easiest to discriminate between different values of χ. Qualitative assessment of this optimal frequency can be obtained based on Fig. 2. For low-frequency, i.e., $\omega \to 0$, it is observed that both transfer functions converge to the same value. In other words, at low-frequency different values for χ are indistinguishable. This is well known in the literature [2, 4]. On the other hand, at high frequencies, the noise contribution starts to dominate the measurement. This results in an overlap between the measurement points
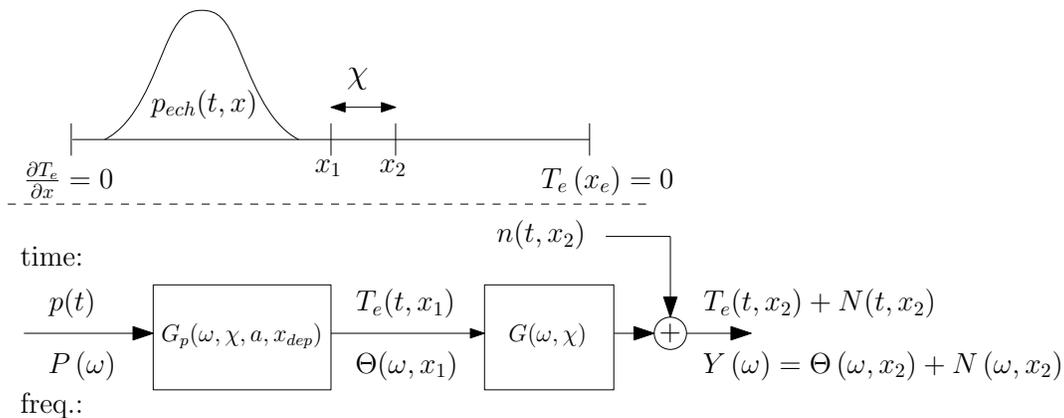
Figure 1. Graphical overview of transfer functions in relationship to 1D domain. Note that depending what problem is analyzed the boundary condition can be different.
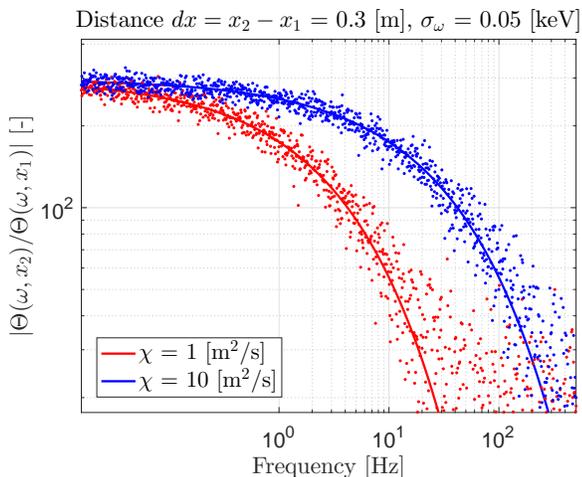

Figure 2. Amplitude ratio of (3) between two spatial locations as function of frequency for two different values of χ where on the output $\Theta(\omega, x_2)$ additive stationary Gaussian distributed white noise. The solid lines represent the amplitude of the noiseless transfer functions $G(\omega, \chi)$.

of the two transfer functions, which again implies that it is difficult to distinguish between different values of χ. Hence, somewhere in between these regions lies the optimal modulation frequency, which we will derive in this paper.

B. Defining the information of an experiment

In the previous example it is was intuitively shown that certain choices of modulation frequency facilitate the distinction between different diffusion coefficients. Two important properties that influence this choice are the noise distribution and the sensitivity of the model output with respect to the diffusion coefficient. Mathematically these factors can be taken into account

through the use of the Fisher information matrix F_i [27]

$$F_i(\theta_0) = \mathbb{E}\left\{ \left(\frac{\partial \ln(f_z)}{\partial \theta} \right)^T \left(\frac{\partial \ln(f_z)}{\partial \theta} \right) \right\} \Bigg|_{\theta=\theta_0}, \quad (4)$$

where f_z stands for the probability distribution of the measurement vector z which contains all the measurement data (e.g. the measured temperatures), θ represents the parameters (e.g. diffusion coefficient), θ_0 represents the true value of the parameters, and where the expected value $\mathbb{E}\{\cdot\}$ is taken with respect to measurement vector z. Notice that the expected value ensures that the Fisher information matrix is independent of the actual measured samples and thus can be computed prior to the experiment.

The importance of the Fisher information matrix follows from the Cramer-Rao lower bound [16]. In the absence of a systematic error, this bound states that the *inverse* of the Fisher information matrix is the lower bound on the uncertainty of the estimated parameters. If the estimation of the parameters is also minimum variance (which is often the case for maximum likelihood estimations) then the inverse of the Fisher information matrix asymptotically approximates the estimation uncertainty [16]. Under these assumptions, the Fisher information matrix allows us to asses the quality of the estimation.

The goal of optimal experiment design is to choose the experiment setting, e.g. the modulation frequency, for which the Fisher information matrix is maximized since this minimizes the uncertainty on the estimated parameters. Finding the largest positive definite matrix is not a straightforward task, since a matrix is a higher dimensional object. To resolve this issue, a scalar function of the Fisher information matrix is optimized instead. Examples of such functions are the determinant, trace, or smallest eigenvalue of the Fisher information matrix. Each choice for the scalar function corresponds to a different information criterion. The value of this criterion is called the information of the experiment. A

more detailed discussion of the different information criteria is given in Sec. VI.

The computation of the Fisher information matrix and thus the information criterion often requires the knowledge of the true system parameters. This implies that in order to estimate the parameters in the most optimal way, the parameters themselves need to be known a priori. This chicken and egg problem is a well-known problem in the field of optimal input design. Different strategies have been followed to circumvent this problem in practice.

- Nominal input design: During nominal input design, good initial estimates of the parameters are used instead of the true parameters to evaluate the Fisher information matrix. This approach only works well if the initial parameter values are already close to the true values [7, 8].

- Robust input design: Robust input design tries to circumvent the shortcomings of nominal design through the use of a robust version of the information criterion, for example the expected value of the information criterion over the distribution of possible parameter values is used [28–30].

- Iterative input design: An iterative or sequential input design consists of an alternation between an estimation step and a design step. In each design step the current best estimation of the parameters is used [7, 30, 31].

It is important to realize that both the robust input design and the iterative design are based on nominal designs. Therefore, a nominal design is always the first step when exploring new optimal input design problems. In this paper, we will assume that the true system parameters are known.

C. Evaluating the Fisher information matrix

In order to use the Fisher information matrix to asses the quality of the estimation, equation (4) needs to be expanded in more detail. This is done by explicitly filling in the distribution of the measurement vector. To simplify the computation of the Fisher information matrix, it is assumed that only the output measurement (y) is corrupted by noise which is Gaussian distributed with known covariance C_y. This implies that the measurement vector contains only the samples of the output signal. The Gaussian distribution of these samples is given by

$$f_z \equiv f_y(y|u,\theta_0,C_y) = \frac{1}{\sqrt{(2\pi)^n \det(C_y)}}$$
$$\exp\left(-\left(y - y_p\left(\theta\right)\right)^T C_y^{-1}\left(y - y_p\left(\theta\right)\right)\right), \quad (5)$$

in which C_y is the covariance matrix of the noise, y_p is a vector with the deterministic part of the output

determined by the plasma transport properties such as the transport coefficients θ, y is the vector containing the measured output samples and n is the number of collected samples. Inserting the distribution in (4) and taking the expected value leads to the following expression for the Fisher information matrix

$$F_i(\theta_0) = \left\{\left(\frac{\partial y_p}{\partial \theta}\right)^T C_y^{-1}\left(\frac{\partial y_p}{\partial \theta}\right)\right\}\bigg|_{\theta=\theta_0}, \quad (6)$$

where $\partial y_p/\partial \theta$ is a vector containing the partial derivatives of y_p with respect to the parameters. These derivatives can be computed based on (1) which describes the physics of the system. Notice that equation (6) is in accordance with the intuition obtained in Sec. III A. The derivatives represent the sensitivity of the output with respect to the parameters, while the covariance expresses the uncertainty introduced by the noise.

Alternatively, we can also work in the frequency domain, the measured time domain samples are transformed to complex spectra using the discrete Fourier transform. This transform can be represented through a linear mapping.

$$Y = A_{DFT}\, y, \qquad y = A_{DFT}^H Y\ , \quad (7)$$

with Y the complex spectrum of the measured output, and A_{DFT} is the discrete Fourier transform matrix. Substituting this expression in (6) allows us to reformulate the equation of the Fisher information matrix as a function of the complex quantities. This results in the following alternative expression [16],

$$F_i(\theta_0) = \left\{\left(\frac{\partial Y_p}{\partial \theta}\right)^H C_Y^{-1}\left(\frac{\partial Y_p}{\partial \theta}\right)\right\}\bigg|_{\theta=\theta_0}, \quad (8)$$

where Y_p contains the complex spectra of the true output, and C_Y is the complex covariance matrix of Y. Notice that the Fisher information matrix still remains a real valued positive definite matrix.

Since we consider perturbative experiments induced by forced perturbations such as ECH, we opt to continue the computations of the Fisher information matrix in the frequency domain. Under the assumption that the system is linear, the true output of the system can be described by using the transfer function

$$Y_p(\omega_k) = G(\omega_k,\theta_0)U_p(\omega_k), \quad (9)$$

in which U_p is the true input spectrum, Y_P is the true output spectrum, and G is the transfer function dictated by the ODE or PDE describing the relation between input and output.

To better illustrate this equation consider again the example of Sec. III A, in that case $Y_p\left(\omega_k\right) = \Theta\left(\omega_k, x_2\right)$, is the output, $U_p\left(\omega_k\right) = \Theta\left(\omega_k, x_1\right)$, is the input, and $G\left(\omega,\theta\right) = \exp\left(-\sqrt{i\omega/\chi}\left(x_2 - x_1\right)\right)$ is the transfer

function. Using this insight and our previous assumptions to compute the Fisher information matrix leads to

$$F_i(\theta_0) = U_p^H \left(\frac{\partial G}{\partial \theta}\right)^H C_Y^{-1} \left(\frac{\partial G}{\partial \theta}\right) U_p. \qquad (10)$$

The $\partial G/\partial \theta$ is a matrix containing the partial derivatives of the transfer function with respect to each of estimated parameters, U contains the complex spectrum of the input which is considered noise free, and where C_Y is the covariance matrix of the measured output spectrum.

In the case only one parameter needs to be optimized and assuming that C_Y is a diagonal matrix with $\sigma_Y^2(\omega_k)$ $k = 1, \ldots, F$ on the diagonal, where the variance can change with frequency, as diagonal elements, the above equation reduces to

$$F_i(\theta_0) = \sum_{k=1}^{F} \frac{1}{\sigma_Y^2(\omega_k)} \left| \frac{\partial G(\omega_k, \theta)}{\partial \theta} \right|_{\theta=\theta_0}^2 \left|^2 |U_p(\omega_k)|^2 \right., \qquad (11)$$

where F corresponds to the number of frequency components in the input signal. This expression for the Fisher information matrix will be used in the remainder of the paper to derive the optimal frequency to estimate the transport coefficients.

IV. OPTIMIZING THE MODULATION FREQUENCY FOR SLAB DIFFUSION ESTIMATES

This section shows how to optimize the modulation frequency such that the diffusion coefficient can be estimated with minimum uncertainty. First, the analytical calculation is given based on an ordinary differential equation (ODE), which is basically the upper bound on the modulation frequency. Then, the full partial differential equation (PDE) solution is given showing that for PDEs the optimal excitation frequency becomes significantly lower.

A. Optimizing the boundary input

Reconsider the example in (3). The transfer function can be split in its amplitude and phase contribution

$$G(\omega, \chi) = \exp\left(-\sqrt{\frac{\omega}{2\chi}}\Delta x\right) \exp\left(-\sqrt{\frac{\omega}{2\chi}}\Delta x \cdot i\right). \qquad (12)$$

By using the general expression in (11), the Fisher information can be calculated with respect to the diffusion

coefficient, i.e., $\theta = \chi$. This results in

$$F_i(\chi) =$$
$$\sum_{k=1}^{F} \frac{\Delta x^2}{2\chi^3} \frac{\omega_k}{\sigma_1^2(\omega_k)} \exp\left(-\sqrt{\frac{2\omega_k}{\chi}}\Delta x\right) |\Theta(\omega_k, x_1)|^2, \qquad (13)$$

for an arbitrary modulation. For simplicity, a single frequency is used (sinusoidal, $F = 1$) instead of the typical block waves. This will still give an accurate estimate for symmetric block waves as most of their energy is contained in the first harmonic anyway. Considering, only one frequency component means that (13) simplifies to

$$F_i(\chi) = \frac{\Delta x^2}{2\sigma^2\chi^3}\omega \exp\left(-\sqrt{\frac{2\omega}{\chi}}\Delta x\right) |\Theta(\omega, x_1)|^2. \qquad (14)$$

Ignoring for a moment the dependence of $\Theta(\omega, x_1)$ on ω, the maximum of (14) is found by taking its derivative and setting it to zero. This results in an optimal excitation frequency ω_{opt} as function of χ and the sensor distance Δx

$$\omega_{opt} = \frac{2\chi}{(\Delta x)^2}. \qquad (15)$$

As we will show later, this is the absolute upper bound on the modulation frequency.

In Fig. 3, the frequency dependence of $F_i^{-1}(\chi)$, which corresponds to the variance of the estimate, is plotted for three different diffusion coefficients. The minima of these curves, which are marked by a black asterix, correspond to the (sinusoidal) optimal excitation frequency $\omega_{opt} = 2\pi f_{opt}$. The decrease of f_{opt} with decreasing χ is in accordance with our intuition, since with decreasing χ the transport is suppressed which holds also for the input signal. However, what is not so intuitive is that the uncertainty increases significantly for higher than optimal frequencies, whereas for lower than optimal frequencies the increase in uncertainty is more modest. In other words, based on this model it is better to choose a low modulation frequency when the diffusion coefficient is unknown.

B. Validation of approach

The Fisher information $F_i(\chi)$ predicts the confidence of the estimate of the diffusion coefficient χ. Therefore, to validate the Fisher information matrix approach a Monte-Carlo analysis is used to validate the statistical outcome. Therefore, the diffusion coefficient is 10000 times estimated in the case there is (dominant) measurement noise at the output temperature. The result is shown in Fig. 4, where the only difference is the modulation frequency, optimal versus a standard frequency. The results show that the confidence of the estimate of χ for the optimal modulation frequency is significantly smaller as is predicted (here more than a factor 4). This
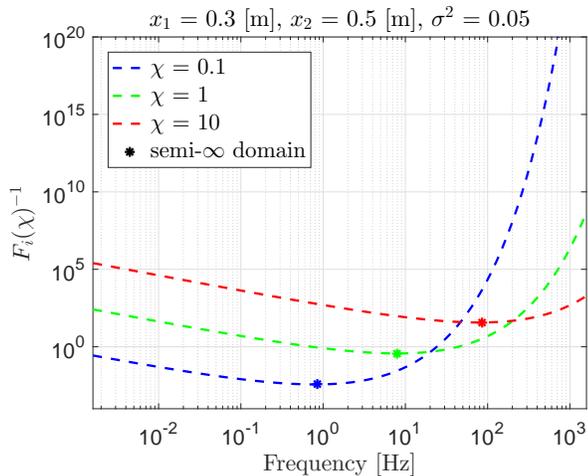
Figure 3. Inverse Fisher information matrix based on (3) assuming a sinusoidal boundary input as a function of frequency for three different values of χ for the semi-infinite domain. The $*$ give the minima of F_i with respect to frequency.
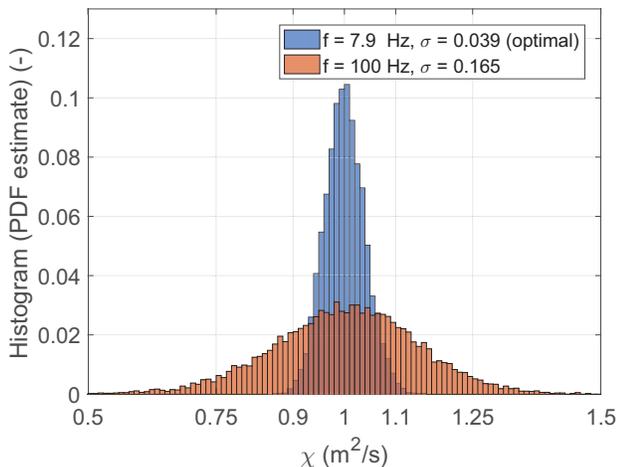

Figure 4. Histograms of 10000 estimations of the diffusion coefficient based on (3) where the optimal modulation frequency is used and a non-optimal frequency is used. Input of the simulation is $\chi = 1$ and a noise level of $\sigma = 0.01$ for $\Theta(\omega, x_x)$ used to avoid estimates of χ around 0. The optimal modulation frequency is invariant for a change in σ in the case of only input or output noise.

validates the Fisher information matrix approach, which has also been validated in much more complicated cases [7, 32, 33].

C. Optimizing the source perturbation

In the previous section it was assumed that $\Theta(\omega, x_1)$ can be directly controlled both in amplitude and frequency. However, in a real transport experiments $\Theta(\omega, x_1)$ cannot be controlled directly, but depends on

the transport over the entire domain and the source. This dependence should be included in the optimization of the optimal modulation frequency.

Hence in this section, the whole PDE in (1) is taken into account including the boundary conditions, which also reintroduces the source term including key parameters such as x_{dep} and a. Its transfer function can also be calculated analytically (or numerically) for constant parameters and is given in the appendix. Here the transfer function from the source to x_1 is shortened to

$$\Theta(\omega, x_1) = G_p(\omega, \chi, x_1, x_{dep}, a) P(\omega), \qquad (16)$$

with $P(\omega) = \mathcal{F}(p(t))$ in (2). This model is graphically depicted in Fig. 1, where it is shown that G_p describes the model over the entire domain till the location x_1 and G describes a local domain between x_1 and x_2 in which we are interested to estimate the diffusion coefficient. Another modification is the introduction of a more realistic boundary condition. Instead of assuming a semi-infinite domain, the following boundary condition is used $T_e(x_e = 2.2) = 0$. This is because an semi-infinite domain has an unrealistic impact on the modulation frequency, which will be explained later.

Next, there are two approaches which can be chosen:

A) The semi-infinite domain approach in which an approximation of the transfer function is used for the domain on which the transport coefficients need to be determined in combination with the transfer function G_p between the input power and the temperature at the spatial location x_1 (semi-infinite + source).

B) The (numerical) transfer function approach which calculates the actual transfer function between x_1 and x_2 in combination with the transfer function G_p between the input power and the temperature at the spatial location x_1 (full numerical solution).

In approach A) only the input power $|\Theta(\omega, x_1)|^2$ becomes ω dependent through G_p, i.e.,

$$F_i(\chi) = \sum_{k=1}^{F} \frac{\Delta x^2}{2\sigma_1^2(\omega_k)\chi^3}$$

$$\omega_k \exp\left(-\sqrt{\frac{2\omega_k}{\chi}}\Delta x\right) |G_p(x_1, \omega_k)|^2 |P(\omega_k)|^2, \quad (17)$$

which needs to be optimized. In approach B) the whole transfer function (i.e. combined effect of G_p and G) and its derivatives are numerically approximated using finite difference scheme. Note that in both cases the standard deviation of the estimated parameter scales reciprocally with the amount of modulation power. Hence, increasing the modulation power or reducing noise is the most straightforward approach to increase SNR.

Again considering a sinusoidal input, the resulting optimal modulation frequency is shown in Fig. 5. In case of approach A) this yields an optimal excitation frequency
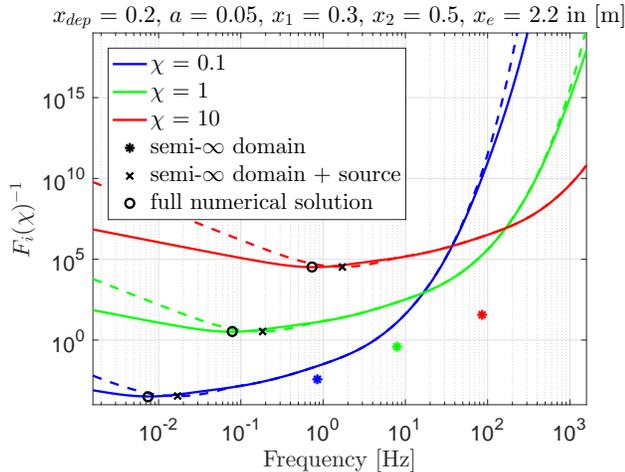
Figure 5. Inverse Fisher information matrix based on (3) assuming a sinusoidal boundary input, a sinusoidal source input, as function of frequency for three different values of χ. The $*$ give the minima of F_i with respect to frequency as were shown in Fig. 3. The dashed lines with minima $\times$ are the results for F_i^{-1} using approach A) as defined in (3). The full lines with minima $\circ$ give the results for F_i^{-1} using approach B).

Figure 6. Time evolution at $T_e(t, x_1)$ for a symmetric block-wave with frequency 0.74 Hz and 12.04 Hz for the full numerical case B).

which is slightly higher than the solution found in case of approach B). Both approaches find an optimal modulation frequency (the crosses and circles) which is significantly lower compared to our previous solution (black asterix) where we do not consider the entire domain. This can be explained by the fact that the amplitude of $\Theta(\omega, x_1)$ decreases with increasing ω.

Based on the difference between the full and dashed lines, it becomes apparent that the semi-infinite domain approximation (A) diverges from the numerical solution (B) for low frequencies. This should be taken into account when using this approximation during the estimation of the transport coefficients. The deviation for higher frequencies can be explained by the fact that a small part of the heating was applied inside the domain to illustrate the effect of the source on the model. In other words, the existence of a source term on the domain results in differences between both methods at high frequencies.

D. Qualitative explanation

To validate our explanation for the decrease in optimal modulation frequency, the perturbation at $|\Theta(\omega, x_1)|$ was simulated for asymmetric block-waves with different modulation frequencies. In Fig. 6 the results of this simulation is shown. This plot clearly shows that the amplitude of the perturbation at $\Theta(\omega, x_1)$ is significantly larger at frequency $f_{opt} = 0.74$ Hz compared to $f_{opt} = 12.04$ Hz. This difference between the modulation amplitudes immediately explains why this low-frequent modulation is more optimal.

To understand why the modulation frequency should not be reduced indefinitely, the amplitude of the transfer function G_p is plotted for different locations of the bounding conditions (x_e) in Fig. 6. From these plots it can be observed that the amplitude G_p flattens for the lower frequencies. This means that lowering the frequency below a certain value no longer leads to an increased amplitude of the modulation. However, the sensitivity of the transfer function G with respect to the diffusion coefficient still decreases for lower frequencies. This explains why frequencies below the optimal modulation frequency are less optimal.

When comparing the shape of G_p for different values of x_e it becomes apparent that the frequency where the flattening of G_p starts, becomes increasingly smaller for larger values of x_e and the gain is reduced due to the boundary condition. If one would decrease χ, the length scale decreases, as such this can also be seen as x_e increasing and as such also the flattening is reduced and the gain increases (not shown). This explains why using the semi-infinite domain approximation $(x_e \to \infty)$ is not accurate as it results in unrealistically low optimal modulation signal.

Note that when optimizing the modulation frequency in a distributed context the influence of Δx on the optimal frequency is strongly diminished. This stands in contrast with (15). The reason for this is that G_p has a significant impact on the modulation frequency, but does not depend on Δx. This, in combination with the fact that f_{opt} is reduced significantly, explains why Δx has little influence (if $\omega \to 0$, then the impact of Δx through G becomes zero). On the other hand, G_p does depend on χ and as such its impact of the diffusion coefficient remains strong.
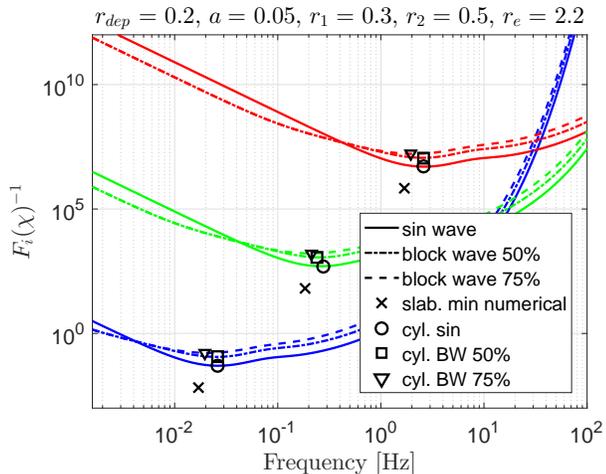
Figure 7. Comparison inverse Fisher information matrices for sinusoidal modulation waveforms and block waveforms in a cylindrical domain calculated numerically using a finite difference approximation.

V. OPTIMIZING THE MODULATION FREQUENCY FOR CYLINDRICAL ESTIMATES

In the previous sections, we have calculated the Fisher information matrix based on sinusoidal perturbations and slab geometry. However, generally block-wave modulations are used in cylindrical like geometry. Hence, the Fisher information matrix is calculated numerically using a finite difference approximation of

$$\frac{\partial}{\partial t}\left(n_e T_e\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(n_e \rho \chi \frac{\partial T_e}{\partial r}\right) + p_{ech}\left(t, r\right), \quad (18)$$

using both (11) and (16) in terms of cylindrical geometry.

A. From ideal sinusoidal slab geometry approximations to cylindrical block wave solutions

The resulting Fisher information matrix is shown in Fig. 7 for a block-wave modulation in cylindrical geometry with a duty-cycle of 50% and 75% in terms of the fundamental frequency of the waveform. The cylindrical geometry has not such a large impact as it slightly increases the optimal modulation frequency due to the enhanced suppression towards small radii. This also means that as r_1 or Δr are becoming small this effect is enhanced.

Fig. 7 shows there is a quantitative difference between the optimal frequency for block-waves and sinusoidal waves. However, the evolution of the inverse Fisher information matrix has qualitatively the same behavior for both the sinusoidal and block wave types. Therefore, it can be concluded that it suffices to use the optimization for sinusoidal waves, as it is sufficiently close to the optimal fundamental frequency of the block wave

modulation. The reason is that even in the case of a block-wave most energy is contained in the first few harmonic components. Moreover, higher harmonic components are suppressed by transport, which reduces their amount of information. Hence, for the identification of the diffusion coefficient a block-wave is not so beneficial. Of course, if one wants to compare harmonic components for validation, then extra harmonic components are desirable.

B. Cylindrical block wave solutions with a broad deposition profiles

In the previous section, we have shown the behavior of the Fisher information matrix for block-wave modulation in cylindrical geometry. As this approach is based on numerical evaluation of the partial differential equations using finite difference, extensions in which the deposition profiles encompass the estimation domain or an off-axis deposition can also be simulated. This is shown in Fig. 8 for two broad deposition profiles. The figures show that in this case the optimal modulation frequency does not change significantly. However, as phase and amplitude differences between spatial locations decrease when there is a source on the domain the confidence of the estimates will go down. The example where the modulation source extends over the estimation domain shows that this method can be applied to other transport channels such as neutral beam injection where source free domains do not exist unlike ECRH.

C. Impact of other transport contributions

In real experiments it is possible that transport contributions other than diffusive contributions are also relevant. The presence of these contributions will of course alter the optimal modulation frequency. To investigate the effect of different transport contributions on the optimal modulation frequency the following PDE is considered

$$\frac{\partial}{\partial t}\left(n_e T_e\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(n_e \rho \chi \frac{\partial T_e}{\partial r} + n_e r V T_e\right)$$
$$- n_e \tau_{inv} T_e + p_{ech}\left(t, r\right), \quad (19)$$

which is the result of linearizing the coupled PDE of mass and electron thermal transport [34]. In (19), V is the convective velocity of the heat pinch, τ_{inv} is the damping ($\tau_{inv} = 1/\tau$). The optimal modulation frequency is computed for different values of V and τ_{inv} with the same method as before.

The evolution of the inverse Fisher information matrix with the modulation frequency is plotted in Fig. 9. Initially each of the three transport coefficients is estimated independently while the other coefficients are considered to be known. In Fig. 9 the black curves correspond to an estimation of the diffusion coefficient, the
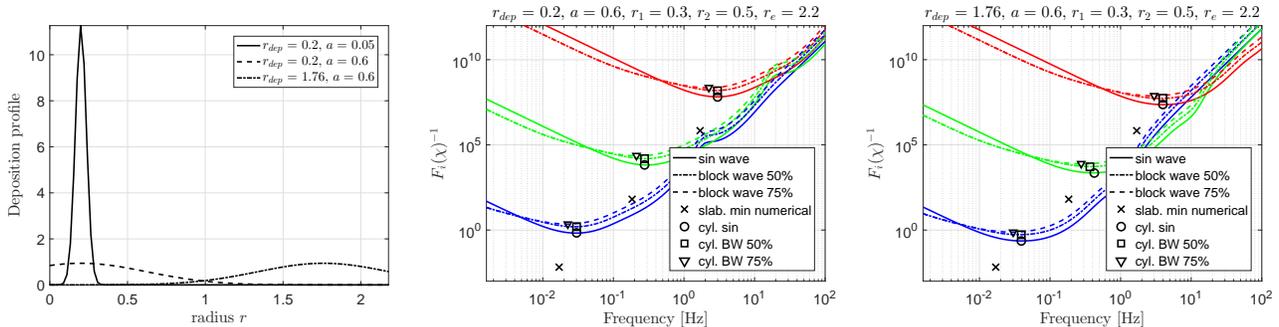
Figure 8. Deposition profiles (right), inverse of the Fisher information matrix for a deposition profile encompassing the estimation domain (middle) and an off-axis modulation (right). Note that only the deposition profile has changes also with respect to Fig. 7.

magenta curves correspond to an estimation of V and the cyan curves correspond to an estimation of τ_{inv}.

For the estimation of χ the presence of a convective velocity and damping term result in a slightly lower optimal modulation frequency. However, if we try to estimate instead V or τ_{inv}, then the optimal modulation frequency drops significantly. Moreover, as the gain factor is different, i.e., convergence of the ratio is no-longer one ($\omega \to 0, |G(\omega, \chi, V)| \neq 1$), extreme low frequencies also perform well to identify V. However, extreme low frequencies give rise to large amplitudes which may lead to non-linearities. Therefore, it is undesirable to go as low as suggested by the calculations.

D. Overview plot of the optimal frequency

Fig. 10 shows the relation between the optimal modulation frequency and diffusion coefficient for various domains (i.e. different values of x_1). Based on this overview graph, it becomes clear that x_1 seems to have little influence on the optimal frequency, with exception of small radii (red curve) since there the impact of the cylindricity is strong.

It is clear from this figure that mainly the diffusion coefficient determines the optimal modulation frequency. As explained before Δx has little influence as such the graphs only show the change in diffusion coefficient χ.

As explained this optimization does not take non-linearities into account. The calculated optimal modulation frequency leads to large amplitudes (see Fig. 6) and as such is more prone to exciting non-linearities. Consequently, this calculated optimal modulation frequency should be seen as a lower bound on the modulation frequency when the regime has non-linear dependencies. The upper bound is shown by the dashed lines and follows from the slab optimization under ideal conditions given by (15).

VI. FURTHER EXTENSIONS

In the previous sections, we have introduced specific descriptions for the Fisher information matrix for transport models commonly used in the fusion community when optimizing the modulation frequencies in perturbation experiments. In this section two possible extensions of the method are further discussed. First, finding the optimal modulation frequency in case multiple transport coefficients are estimated and secondly how to handle non-linearities.

A. Simultaneous estimation of multiple transport coefficients

Uptill now we have always optimized the modulation frequency for the estimation of one transport coefficient. However, in many experiments there are multiple transport coefficients which need to be estimated. Consequently, the Fisher information matrix will be a (positive definite) matrix instead of a scalar, which is not always comparable on a matrix level [35]. Hence, the Fisher information matrix needs to be reduced to a scalar information criterium again allowing the selection of an optimal modulation frequency or an alternative quantity which needs to be optimized in the experiment.

Deciding which information criterion to use, is strongly related to the envisioned purpose of the model [32]. Since the estimated parameters have a physical interpretation, it is sensible to use an information criterion that is related to uncertainty in the estimated parameters. Three common information criteria used for accurate parameter estimates are [33]:

- A-optimality: An A-optimal input minimizes the trace of the inverse of the Fisher information matrix. Geometrically this corresponds to minimizing the sum of edges of the bounding box surrounding the uncertainty region of the estimated parameters [36]. Note that scaling of the units transport coefficients influences the optimality.
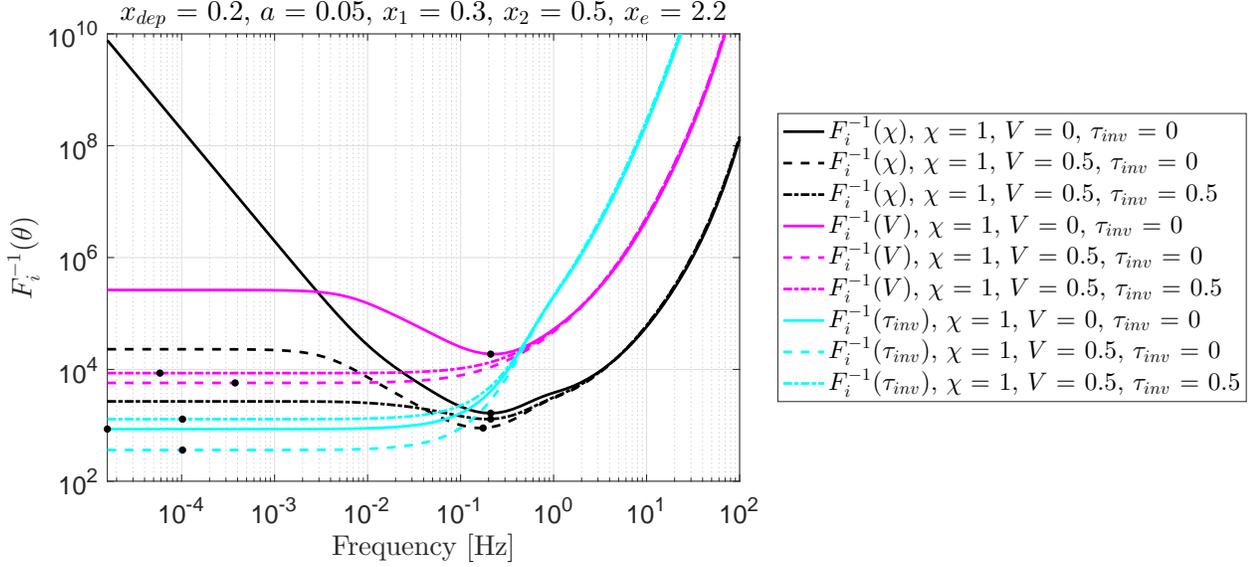
Figure 9. Inverse Fisher information matrix for various values of f_{mod} where only one transport coefficient is varied. The colors correspond to which parameter is varied and the line types to a different combination of the other coefficients.
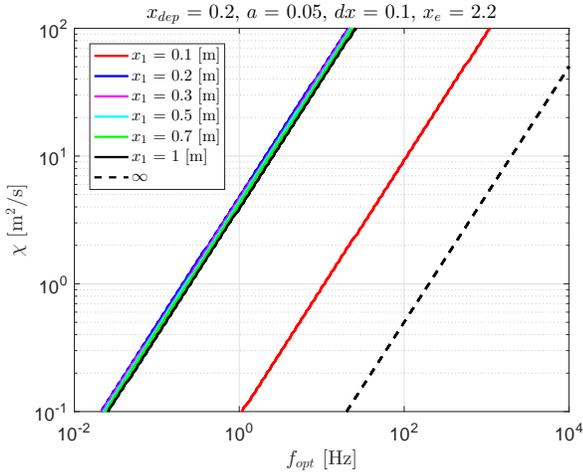
Figure 10. Optimal modulation frequency versus diffusion coefficient for various values of x_1. In addition, the slab geometry approximation is shown by the dashed line, which is independent of x_1.

- D-optimality: A D-optimal input maximizes the determinant of the Fisher information matrix. Geometrically this corresponds to minimizing the uncertainty volume of the estimated parameters [37].

- E-optimality: An E-optimal input maximizes the smallest eigenvalue of the Fisher information matrix. Geometrically this corresponds to minimizing the largest axis of the uncertainty ellipse [38]. Note that scaling of the units transport coefficients influences the optimality.

For a more in-depth study of the difference between these criteria the reader is referred to [39].

The A-optimality and D-optimality criteria are applied to the problem in (19) where both the diffusion coefficient χ and the convective velocity V need to be estimated. The resulting contour plots of the optimal modulation frequency for various values of χ and V are shown in Fig. 11.

Both show that when V is small, the optimal modulation frequency is quite similar for both criteria. However, if V becomes negative the two criteria start to diverge significantly. In the top left corner, the convective term dominates over the diffusion and as such is complicated to estimate, this is reflected by the optimal modulation frequency changing quickly here.

B. Finding the minimum frequency modulation and avoiding non-linearities

In the previous sections, it is shown that in a purely linear experiment the modulation frequency is small compared to what is expected in real experiments. The reason is that in real experiments the transport depends non-linearly on the perturbation. On the other hand, the transport coefficients such as the diffusion coefficient are based on the linearized transport models. Hence, we want to estimate the transport coefficients generally in the linear regime. Therefore, the perturbative experiment has another constraint and that is that the perturbation should be sufficiently small such that the transport coefficients can be estimated. However, this constraint is not in the linear model and as such in the optimization of the excitation signal. Moreover, the reason why small frequencies were optimal in the linear case is due to the perturbation becoming very large as heat is accumulated in the system (see Sec. IV D and specif-
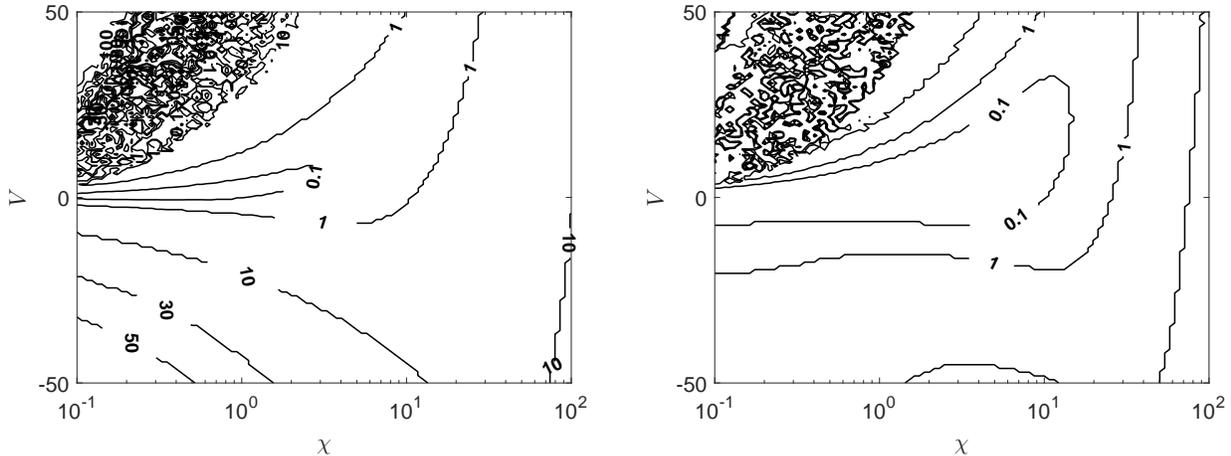
Figure 11. Contour plots of (left) $\max\left(\det\left(F_i\right)\right)$ and (right) $\max\left(\operatorname{trace}\left(F_i\right)\right)$ where the contour lines show the corresponding optimal modulation frequencies in [Hz] as function of constant diffusion coefficient χ and constant convective velocity V in cylindrical geometry. The scaling in the case of the trace (A-optimality) is done in SI units.

ically Fig. 6). Hence, in reality the optimal modulation frequency should be higher such that the perturbation is sufficiently small for the temperature perturbation to stay in the linear regime. Hence, there are two methods to take this extra constrained into account, i.e., to avoid non-linearities in the optimization

- Optimize the Fisher information matrix for the underlying non-linear model.

- Verify in the experiment (or in full simulation) if non-linearities occur at a certain amplitude in combination with modulation frequency.

The first option is still a field of research and is beyond the scope of this paper. Moreover, it is unclear what non-linear model should be used. The reader interested in the optimal input designs for non-linear models is referred to [10–12, 40–42]. Alternatively, we can verify in simulation (or experimentally) when we enter the non-linear regime. This can be done in the frequency domain by analyzing the nonlinear components (i.e. higher harmonics) of the output spectrum as is shown in Fig. 12. For more details we refer to [16, Chapter 3].

The amplitude of the signals show that in the case of non-linear response new harmonic components appear at multiples of the ground harmonic components and inter-modulation harmonic components [43]. When these extra harmonics are at the same level as the noise, then we consider the experiment as linear. Therefore, the optimized modulation frequency must simultaneously fulfill this condition. This test has also been experimentally applied and can be found in [21]. There, it is shown that indeed if the perturbation is chosen relatively small even for a small perturbation one can observe non-linear components. The experiment being non-linear can be caused by various non-linear dependencies such as $\chi\left(T, \nabla T\right)$ or we can have a non-linear boundary condition. In the latter case, we want to avoid

that we are modulating the boundary too strongly. As a simple approximation, we could say for on-axis modulation that the perturbation should be optimal for the interval till the wall. In that case in (15), Δx must be replaced by the minor radius a such that

$$\omega_{min} = \frac{2\chi}{a^2} \longleftrightarrow f_{mod} \sim \frac{\chi}{\lambda^2}. \tag{20}$$

Then, we see that the result is closely linked to f_{mod} in Sec. II, the standard measure of choosing the modulation frequency. This clearly shows the link between classic interpretation and the systematic optimization performed in this paper.

C. Optimization for non-linear dependencies

Non-linear dependencies are rarely estimated directly in perturbative fusion experiments, but are mapped out through the use of a number of linearizations for a set of operating points. These are coupled together to acquire a non-linear model description. Hence, for every operating point the optimization of the modulation signal is exactly done as described in this paper. In case of a critical gradient model at the threshold the diffusion coefficient can change significantly. Hence, depending on the operating point the diffusion coefficient is different and also the optimal modulation frequency. As Fig. 4 shows, using a too low modulation frequency gives a significantly better result than using a too high modulation frequency when remaining in the linear regime. Hence, if it is unclear in which transport regime the experiment is performed one should choose an optimal frequency which is towards the transport coefficient belonging to the regime with the lower optimal frequency.
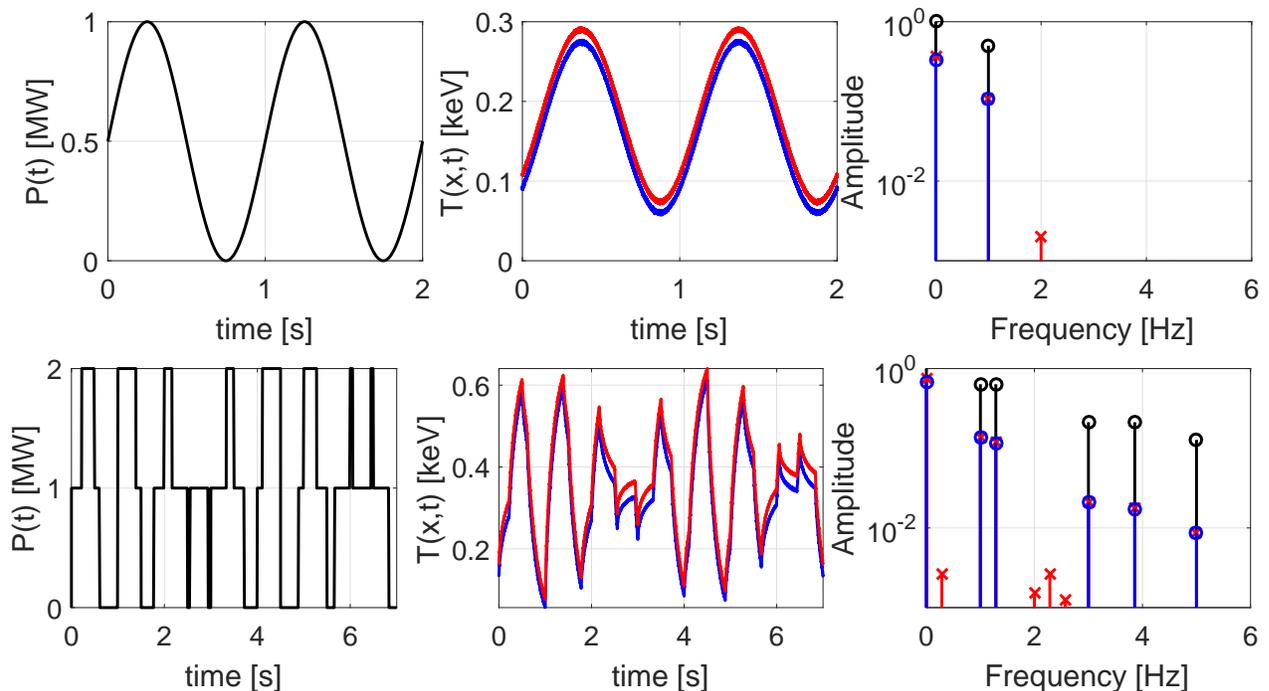
Figure 12. Graphical representation of two perturbation: 1) signals sinusoidal with a frequency of 1 Hz (top) and 2) sum of two block-wave modulations with frequencies 1 Hz and 9/7 Hz (bottom). In blue the linear response is shown and in red the non-linear response. On the right are the corresponding amplitudes of the Fourier transformed time signals with corresponding colors.

D. Perturbative transport problems in general

The method to optimize perturbation experiments described in this paper can be applied, in principle, to every linear perturbation experiment. The extensions proposed in this manuscript, with regard to standard linear experimental optimization theory [44], are on the distributed character of transport experiments. This means that for modulated neutral beam injection and modulated ion cyclotron resonance heating the same technique can be applied. After all, the partial differential equation models are very similar for all these problems. In case of, for instance gas-puffing, the optimization can be further simplified because the perturbation is applied on the boundary of the domain. This is described in Sec. IV A where the frequency dependent perturbation amplitude does not play a role. This is the mathematical part of the optimization. Of course, in practice, gas-puffing is much more complicated due to its lack of symmetry in its propagation and the interaction with other gasses. Nevertheless, the method presented in this paper is applicable to the optimization of any perturbation problem as long as the response to the perturbation remains linear.

VII. CONCLUSION AND SUMMARY

In this paper, we discuss how the design of an optimal modulation experiment based on the concept of the Fisher information matrix. First, this method was used to determine analytical expression for the optimal modulation frequency under simplifying assumptions. It turned out this solution forms an upper bound for the optimal modulation frequency. Later, we showed how more realistic conditions can be incorporate into the optimization, which lead to a decrease of the optimal frequency. To conclude it was explained how the computation of the optimal modulation frequency could be extended in the case of simultaneous estimation of multiple transport coefficients and waveforms with tunable power spectrum.

Special attention went to understanding the qualitative reasoning behind a low optimal modulation frequency. We showed that the optimal frequency of the source depends both on the amplitude and the modulation frequency of the perturbation at the boundary x_1. Decreasing the modulation frequency increases the amplitude and as such has a favorable impact on the signal-to-noise ratio. This leads to very small optimal modulation frequencies when assuming a linear model.

In reality perturbative experiments with large amplitudes are more likely to induce non-linear effects. Hence, an experiment with a very low modulation frequency may violate the linear assumptions that were

made during the design. To resolve this issue it is theoretically possible to directly optimize the non-linear experiment, assuming non-linear models are available. However, such optimization scheme's are significantly more involved than the ones discussed here. Alternatively the validity of the linear conditions could be evaluated experimentally or through simulation of a non-linear transport model. Hence, the final conlusion of the paper is

- Absolute upper bound for the modulation frequency is given by the diffusion coefficient divided by π times the distance squared between the measurement points one wants to estimate.

- Absolute lower bound on modulation frequency is given by the combination of non-linearity avoidance and perturbation size, which both implicitly depends on the frequency.

As a final remark optimization of experiments can be considered a so-called chicken-and-egg problem. If one knows the transport coefficient exactly, then there is no need to do an experiment (except for conformation perhaps). On the other hand, if one has absolutely no idea of the transport coefficient, then any modulation frequency could be optimal. The methodology gives insight if one has some idea of the transport coefficient. The range of the modulation frequency to be used depends on the range of the transport coefficient a priori known. Hence, as a first experiment one applies a wide band modulation signal, with a number of frequency components in the region of interest. The transport coefficient can be identified with some accuracy reducing the range of possible transport coefficients significantly. Redesigning the band or making an improved choice of the modulation frequency reduces the uncertainty on the transport coefficients even further until a desired accuracy is achieved or the absolute highest limit of accuracy theoretical possible is reached (Cramer-Rao lower bound). Consequently, the methods described in this paper give the best result if used in a recursive fashion.

VIII. ACKNOWLEDGMENTS

APPENDIX: ANALYTIC SOLUTION FOR A SLAB GEOMETRY WITH GAUSSIAN SOURCE

This appendix presents the analytical solutions for the transfer function G_p in case of slab geometry and a Gaussian deposition profile used for the analytical optimization of the modulations frequency. Consider again the slab geometry solution with constant density and constant diffusion coefficient, i.e.,

$$n_e \frac{\partial T_e}{\partial t} = n_e \chi \frac{\partial^2 T_e}{\partial x^2} + p_{ech}\left(t, x\right), \qquad (21)$$

with a source of the form

$$p_{ech}\left(t, x\right) = p\left(t\right) \frac{1}{a\sqrt{\pi}} \exp\left(-\frac{(x - x_{dep})^2}{a^2}\right) \qquad (22)$$

and boundary conditions: $\partial T_e / \partial x \left(x = 0\right) = 0$ and $T_e\left(x_{end}\right) = 0$. This is transformed to the Fourier domain, which results in

$$n_e i\omega \Theta = n_e \chi \frac{\partial^2 \Theta}{\partial x^2} + P_{ech}\left(\omega, x\right), \qquad (23)$$

with $\Theta\left(\omega, x\right) = \mathcal{F}\left(T_e\left(t, x\right)\right)$ and $P_{ech}\left(\omega, x\right) = \mathcal{F}\left(p\left(t, x\right)\right)$. This can be solved analytically

$$\Theta\left(\omega, x\right) = c_1 e^{x\sqrt{\frac{i\omega}{\chi}}} + c_2 e^{-x\sqrt{\frac{i\omega}{\chi}}} + G_p\left(\omega, x\right) P\left(\omega\right), \quad (24)$$

where $P\left(\omega\right) = \mathcal{F}\left(p\left(t\right)\right)$

$$G_p\left(\omega, x\right) = \alpha e^{(x - x_{dep})\sqrt{\frac{i\omega}{\chi}}} \text{erf}\left(\frac{a}{2}\sqrt{\frac{i\omega}{\chi}} + \frac{x - x_{dep}}{a}\right)$$
$$+ \alpha e^{(x_{dep} - x)\sqrt{\frac{i\omega}{\chi}}} \text{erf}\left(\frac{a}{2}\sqrt{\frac{i\omega}{\chi}} + \frac{x_{dep} - x}{a}\right) \quad (25)$$

with $\alpha = \frac{i\sqrt{i}\pi}{4n\sqrt{\chi\omega}} \exp\left(\frac{ia^2\omega}{4\chi}\right)$. Boundary conditions $\partial T_e\left(x = 0\right) / \partial x = 0$ and $T_e\left(x = x_{end}\right) = 0$, which translates to $\partial \Theta\left(x = 0\right) / \partial x = 0$ and $\Theta\left(x = x_{end}\right) = 0$ and results in

$$c_1 = \frac{e^{x_{end}\sqrt{\frac{i\omega}{\chi}}} G_p\left(\omega, x_{end}\right) + \sqrt{\frac{\chi}{i\omega}}\frac{\partial G_p\left(\omega, x=0\right)}{\partial x}}{1 + e^{2x_{end}\sqrt{\frac{i\omega}{\chi}}}} P\left(\omega\right)$$
$$(26)$$

and

$$c_2 = P\left(\omega\right) \cdot$$
$$\left(\frac{e^{x_{end}\sqrt{\frac{i\omega}{\chi}}} G_p\left(\omega, x_{end}\right) + \sqrt{\frac{\chi}{i\omega}}\frac{\partial G_p\left(\omega, 0\right)}{\partial x}}{1 + e^{2x_{end}\sqrt{\frac{i\omega}{\chi}}}} - \sqrt{\frac{\chi}{i\omega}}\frac{\partial G_p\left(\omega, 0\right)}{\partial x}\right).$$
$$(27)$$

[1] P. Mantica and F. Ryter, "Perturbative studies of turbulent transport in fusion plasmas," *C. R. Phys.*, vol. 7, no. 6, pp. 634–649, 2006.

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

[3] 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.

[4] 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.

[5] 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.

[6] M. Chilenski, M. Greenwald, Y. Marzouk, J. Rice, and A. White, "Efficient design and verification of diagnostics for impurity transport experiments," *Review of Scientific Instruments*, vol. 89, no. 1, p. 013504, 2018.

[7] V.V.Federov, *Theory of Optimal Experiments*, Z. Birn-Baum and E. Lukacs, Eds. Academic Press, 1972.

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

[9] M. Gevers, "Identification for control: From the early achievements to the revival of experiment design," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, Dec 2005, pp. 12–12.

[10] H. Hjalmarsson, J. Martensson, and B. Ninness, "Optimal input design for identification of non-linear systems: Learning from the linear case," in *American Control Conference, 2007. ACC '07*, July 2007, pp. 1572–1576.

[11] P. Valenzuela, C. Rojas, and H. Hjalmarsson, "A graph theoretical approach to input design for identification of nonlinear dynamical models," *Automatica*, vol. Vol 51, pp. pp 233–242, January 2015.

[12] K. Mahata, J. Schoukens, and A. D. Cock, "Information matrix and d-optimal design with Gaussian inputs for Wiener model identification," *Automatica*, vol. 69, pp. 65 – 77, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109816300589

[13] S. M. Kay, "Fundamentals of statistical signal processing. vol 1, estimation theory," 1993.

[14] A. van den Bos, *Parameter Estimation for Scientists and Engineers.* Hoboken: John Wiley & Sons., 2007.

[15] L. Ljung, *System identification: theory for the user*, 2nd ed. Prentice Hall, Upper Saddle River (NJ), 1999.

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

[17] P. Mantica, F. Ryter, C. Capuano, H. U. Fahrbach, F. Leuterer, W. Suttrop, J. Weiland, and ASDEX Upgrade Team, "Investigation of electron heat pinch in AS-DEX Upgrade," *Plasma Phys. Control. Fusion*, vol. 48, no. 3, p. 385, 2006.

[18] P. Mantica, G. Gorini, G. M. D. Hogeweij, N. J. Lopes Cardozo, and A. M. R. Schilham, "Heat convection and transport barriers in low-magnetic-shear rijnhuizen tokamak project plasmas," *Phys. Rev. Lett.*, vol. 85, pp. 4534–4537, Nov 2000.

[19] A. Jacchia, F. De Luca, G. Hogeweij, G. Gorini, J. Konings, N. L. Cardozo, P. Mantica, and M. Peters, "Simultaneous analysis of ech modulation and sawtooth activity in the plasma core of the rtp tokamak," *Nucl. fusion*, vol. 34, no. 12, p. 1629, 1994.

[20] H. J. Hartfuss, L. Giannone, U. Stroth, V. Erckmann, U. Gasparino, H. Maassberg, W7-AS-team, and ECRH-team, "Heat wave studies on W7-AS stellarator," in *Proceedings of the workshop on 'Local Transport Studies in Fusion', Varenna 1993*, 1994, pp. 119 – 125.

[21] M. van Berkel, H. Igami, G. Vandersteen, G. Hogeweij, K. Tanaka, N. Tamura, M. R. de Baar, H. J. Zwart, S. Kubo, S. Ito, H. Tsuchiya, and the LHD Experiment Group, "New evidence and impact of electron transport non-linearities based on new perturbative inter-modulation analysis," 2017.

[22] A. Jacchia, F. De Luca, S. Cirant, C. Sozzi, G. Bracco, A. Bruschi, P. Buratti, S. Podda, and O. Tudisco, "Gradient length driven electron heat transport study in modulated electron cyclotron heating ftu tokamak," *Nucl. fusion*, vol. 42, no. 9, p. 1116, 2002.

[23] G. Tardini, A. Peeters, G. Pereverzev, F. Ryter, J. Stober, and A. U. Team, "Comparison of theory based transport models with asdex upgrade data," *Nucl. Fusion*, vol. 42, no. 3, p. 258, 2002.

[24] F. Imbeaux, F. Ryter, and X. Garbet, "Modelling of ech modulation experiments in asdex upgrade with an empirical critical temperature gradient length transport model," *Plasma Phys. Control. Fusion*, vol. 43, no. 11, p. 1503, 2001.

[25] F. Ryter, C. Angioni, A. G. Peeters, F. Leuterer, H.-U. Fahrbach, and W. Suttrop, "Experimental study of trapped-electron-mode properties in tokamaks: Threshold and stabilization by collisions," *Phys. Rev. Lett.*, vol. 95, p. 085001, 2005.

[26] 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.

[27] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[28] L. Pronzato and E. Walter, "Robust experiment design via maximin optimization," *Mathematical Biosciences*, vol. 89, no. 2, pp. 161 – 176, 1988. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0025556488900971

[29] J. Mï¿œrtensson and H. Hjalmarsson, "Robust input design using sum of squares constraints," *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 1352 – 1357, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1474667015354550

[30] C. R. Rojas, J. S. Welsh, G. C. Goodwin, and A. Feuer, "Robust optimal experiment design for system identification," *Automatica*, vol. 43, no. 6, pp. 993–1008, 2007.

[31] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, vol. 58, no. 5, pp. 527–535, 09 1952. [Online]. Available: http://projecteuclid.org/euclid.bams/1183517370

[32] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, no. 5, pp. 543 – 554, 1986.

[Online]. Available: http://www.sciencedirect.com/science/article/pii/0005109886900646

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

[34] K. W. Gentle, "Dependence of heat pulse propagation on transport mechanisms: Consequences of nonconstant transport coefficients," *Phys. Fluids B-Plasma*, vol. 31, p. 1105, 1988.

[35] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York, NY, USA: Cambridge University Press, 2004.

[36] H. Chernoff, "Locally optimal designs for estimating parameters," *The Annals of Mathematical Statistics*, vol. 24, no. 4, pp. 586–602, 1953. [Online]. Available: http://www.jstor.org/stable/2236782

[37] Z. Galil and J. Kiefer, "D-optimum weighing designs," *The Annals of Statistics*, vol. 8, no. 6, pp. 1293–1306, 1980. [Online]. Available: http://www.jstor.org/stable/2240943

[38] S. Ehrenfeld, "On the efficiency of experimental designs," *The Annals of Mathematical Statistics*, vol. 26, no. 2, pp. 247–255, 1955. [Online]. Available: http://www.jstor.org/stable/2236879

[39] W. Wong, "Comparing robust properties of A, D, E and G-optimal designs," *Computational Statistics and Data Analysis*, vol. 18, no. 4, pp. 441 – 448, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0167947394901619

[40] C. Larsson, H. Hjalmarsson, and C. Rojas, "On optimal input design for nonlinear FIR-type systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec 2010, pp. 7220–7225.

[41] D. Telen, D. Vercammen, F. Logist, and J. V. Impe, "Robust optimal experiment design for nonlinear dynamic systems," in *22nd Mediterranean Conference on Control and Automation*, June 2014, pp. 930–935.

[42] A. De Cock, M. Gevers, and J. Schoukens, "D-optimal input design for nonlinear fir-type systems: A dispersion-based approach," vol. 73, pp. 88–100, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109816301844

[43] P. Wambacq and W. Sansen, *Distortion analysis of analog integrated circuits.* Springer Science & Business Media, 2013, vol. 451.

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