

Estimation of Heat Source Term and Thermal Diffusion in Tokamak Plasmas Using a Kalman Filtering Method in the Early Lumping Approach*

Sarah Mechhoud¹, Emmanuel Witrant¹, Luc Dugard¹ and Didier Moreau²

Abstract— In this paper, early lumping estimation of space-time varying diffusion coefficient and source term for a non-homogeneous linear parabolic partial differential equation (PDE) describing Tokamak plasma heat transport is considered. The analysis of this PDE is achieved in a finite dimensional framework using the cubic b-splines finite element method with the Galerkin formulation. This leads to a finite dimensional linear time-varying state-space model with unknown parameters and inputs. The Extended Kalman Filter with Unknown Inputs Without Direct Feed-through (EKF-UI-WDF) is applied to estimate simultaneously the unknown parameters and inputs and an adaptive fading memory coefficient is introduced in the EKF-UI-WDF, to deal with time varying parameters. Conditions under which the direct problem is well posed and the reduced order model converges to the initial one are established. *In silico* and real data simulations are provided to evaluate the performances of the proposed technique.

I. INTRODUCTION

Distributed parameter systems (DPS) widely exist in industrial processes. These physical and chemical systems are governed by partial differential equations (PDE) and complex spatio-temporal nonlinear dynamics. In many situations, it is difficult to get an accurate nominal PDE description due to incomplete physical or chemical knowledge (unknown system parameters, unknown disturbances...). These uncertainties make the modelling problem tedious. Three different problems in DPS are of prime interest: (i) model reduction when the objective is to reduce the process high-order to limit computation loads, (ii) system identification (“black-box modelling”) for which the structure of the PDE is unknown and the problem is to capture the dominant dynamics, and (iii) parameter estimation (“grey-box modelling”), where the PDE structure is available and only some parameters need to be known. This last problem constitutes our field of interest. Thermonuclear fusion is a very complex physical process where several DPS phenomena (magnetohydrodynamics) occur. It has been proposed as a promising alternative to fossil fuels and as a sustainable energy source since the 40s. A nuclear fusion reaction between light atoms such as

the hydrogen isotopes deuterium and tritium yields to large amounts of energy: from 100 *mg* of deuterium plus 150 *mg* of tritium, it is possible to extract the same energy as the one produced by 7 barrels of oil. To extract energy from a fusion reaction, different confinement concepts exist: mainly inertial confinement (high-energy beams of laser light, electrons or ions) and magnetic confinement (in stellarators and tokamaks). Tokamak plasmas are considered in this work. A major drawback of nuclear fusion is that the fuel burns at a temperature of hundreds of million Kelvin, rendering precise physical modelling and feedback design particularly difficult to achieve.

To obtain and maintain plasma conditions that are optimized for energy generation and to guarantee safe fusion operations (especially in advanced scenarios), the control of heat and particles transport in tokamaks is mandatory. Understanding heat transport mechanisms helps in explaining plasma instability and energy losses and this may lead to reliable predictions of the tokamak performances [1]. In this work, electron heat transport is described by a one-dimensional diffusion equation in a cylinder, where electrons and ions heat diffusivities are distributed (space-time varying). Several empirical models for the diffusion coefficient in hot plasmas exist ([2], [3], [4] and references therein). They depend on several conditions (tokamak dimensions, discharge parameters and temperature profiles to name a few), but all of them assert the diffusivity dependence on the temperature gradient and the magnetic and velocity shears. As a consequence, the heat model becomes nonlinear, complex and coupled with other variables. Another unknown quantity is the heating energy absorbed by the particles, called *the source term*. It depends on the power deposition profiles and on the efficiencies of the various heating systems (radio-frequency waves and high-energy neutral beams), and is sometimes difficult to model because of parasitic phenomena and anomalous energy losses. To derive an experimentally-based model, the diffusion coefficient formula is assumed to be of an unknown form and we aim to reconstruct this coefficient and the source term using parameter identification tools. Note that, in all previous studies which dealt with the heat diffusivity estimation in tokamak plasmas (see [5] and references therein), the source term was supposed to be a known quantity. In these works, only heating by Electron Cyclotron Resonance Heating (ECRH) was considered: experimental results have proved that using this mean of heating, plasma’s electrons absorb almost $\simeq 100\%$ of the heating power [6], [5]. When

This project received partial funding provided by the Fdration de Recherche Fusion par Confinement Magnétique (FR-FCM) and by the Bonus Quality Research of Grenoble-INP.

¹S. Mechhoud, E. Witrant and L. Dugard are with UJF-Grenoble 1/CNRS, Grenoble Image Parole Signal Automatique (GIPSA-lab), UMR 5216, B.P. 46, F-38402 St Martin d’Hères, France {sarah.mechhoud, emmanuel.witrant, luc.dugard}@gipsa-lab

²D. Moreau is with CEA, IRFM, F-13108 Saint Paul-Lez-Durance, France didier.moreau@cea.fr

the heating mode is different (Low Hybrid , Neutral Beam Injection and Ion Cyclotron Resonance Heating) the source term is an ambiguity (see [7]).

The main contribution of this paper is that it considers the estimation problem of both the diffusion coefficient and the source term with no *a priori* assumptions and free of the Tokamak's operational conditions.

The estimation of these parameters is needed to optimize current and pressure profiles which allows the reactor to be run close to the stability limits of magnetic confinement in a controlled manner (see [1]).

In general, estimating a partial differential equation (PDE) is a challenging task. This work is dedicated to propose some answers to this problem in the finite dimensional framework, called also early lumping approach. It consists in performing a spatial discretization of the PDE to generate a set of ordinary differential equations (ODE) that constitutes an approximation of the original PDE model. For this reduced model, an estimation design is developed in the framework of the classical estimation/identification theory of lumped parameter systems (LPS). It must be noted that through early lumping, the finite dimensional model may be of high order and thus difficult to implement.

In this paper, we first convert the PDE model for heat transport into a finite dimensional system, i.e. a set of ODEs by means of the finite element method (FEM) in the Galerkin formulation using cubic b-splines functions [8] (FEM-Galerkin cubic b-splines method). Since diffusion and source term are functional parameters, in the early lumping approach the distributed PDE equation is converted into a linear time-varying system (LTV) with unknown parameters and inputs using the projection/interpolation methods. In order to reconstruct them, an estimation procedure has to be developed.

In the literature, model-based state and input estimation have received considerable attention since the 70's. In [9], the unknown input was considered as a part of the state with a Gaussian distribution. This assumption is not valid for every input variable: some are deterministic whereas many others are not Gaussian. In [10], an unbiased minimum variance filter was proposed to estimate the state independently of the unknown inputs. Since [10], several papers dealt with this problem in the MVU (minimum variance unbiased) framework [11], [12], [13], [14], where optimality conditions in the scope of unbiased minimum variance estimation were generated. Other methods, especially for observer-based design using matrix manipulations [15], [16], sliding modes [17] or linear matrix inequalities [18] were developed. Most of these methods address the problem of joint state and input estimation and cannot be easily extended to include the parameter estimation. They typically do not guarantee optimality in the least square sense. To cope with this situation, the Kalman Filter for Unknown Inputs Without Direct Feed-through (KF-UI-WDF) is a good approach [19]. This filter guarantees the optimality both in the MVU and the least-squares sense.

By minimizing a weighted least squares objective function

with respect to an extended variable including the actual states and all unknown inputs from the initial measurement time t_0 to the current one t_k , a recursive least-squares estimation approach is performed and the KF-UI-WDF is derived. Throughout this paper, we assume that distributed sensing and measurements are available. The question of input estimation is not only related to heat transport but arises, to cite few, in fault detection, machine tool and manipulator applications, chaotic systems and general inverse problems.

This paper is organized as follows. In Section II, the electron heat model and its well-posedness's conditions are presented. In Section III, the early lumping estimation approach using the projection method and the KF-UI-WDF are introduced after proving the convergence of the finite dimensional estimation problem. The case of time varying parameters using an adaptive fading memory is considered in Section IV. Computer simulations using both simulated and real data are provided to demonstrate the efficiency of the proposed methodology.

II. ELECTRON HEAT TRANSPORT IN A TOKAMAK

A. Model description

Assuming poloidal and toroidal axisymmetry, the tokamak is considered as an infinite cylinder where space variations occur along only the small plasma radius a . This hypothesis typically referred to as the "cylindrical approximation" and implies the symmetry of the variables with respect to the plasma center. Applying the energy conservation principle, the electron heat transport model is given by the following parabolic partial differential equation [20]:

$$\frac{3}{2} \frac{\partial(nT)}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} (rn\bar{\chi}_e(r,t) \frac{\partial T}{\partial r}) - \frac{3}{2\tau} nT(r,t) + P_e(r,t) \quad (1)$$

where t is the time, r is the radial variable along a , $\bar{\chi}_e$ is the electron diffusivity, $\tau (< \infty)$ is a damping time modelling the energy losses, T is the electron temperature, n is the electron density and P_e is the power density absorbed by the particles from an external heating system. The spatio-temporal variations of the electron density are assumed to be negligible with respect to the temperature variations during the heating process, and using the normalized variable $z = \frac{r}{a}$, (1) can be written as:

$$\begin{cases} \frac{\partial T}{\partial t} = \text{div} \left(\chi_e(z,t) \frac{\partial T}{\partial z} \right) - \frac{1}{\tau} T(z,t) + S(z,t), & z \in \Omega, t \in [0, t_f] \\ \frac{\partial T}{\partial z}(z=0,t) = 0, \quad T(z=1,t) = 0, & t \in [0, t_f], \\ T(z,t=0) = 0, & z \in \Omega, \end{cases} \quad (2)$$

where Ω is the interval $]0, 1[$, t_f is the final time and div is the divergence operator in the cylindrical axisymmetry configuration: $\text{div}(\cdot) := \frac{1}{z} \frac{\partial}{\partial z} (z \cdot)$. The normalized source term $S(z,t)$ and diffusivity coefficient $\chi_e(z,t)$ are respectively

given by:

$$\begin{cases} S(z, t) = \frac{2}{3n} P_e(z, t) \\ \chi_e(z, t) = \frac{2}{3a^2} \overline{\chi_e}(r, t) \end{cases} \quad (3)$$

In system (2), the second and third equations represent initial and boundary conditions, chosen to guaranty the plasma symmetry (no gradient at $z = 0$) and a negligible temperature at the plasma edge ($T(1, t) = 0$) in comparison with the central temperature. The initial temperature is set to zero by considering that all the energy is brought by the external sources $S(z, t)$.

Even if it has been shown (e.g see [21]) that the diffusion coefficient χ_e depends on the temperature gradient, in order to take advantage of the linear aspect of our PDE, in this paper it is assumed that there is no *a priori* assumption on χ_e .

B. Well-posedness of the direct problem

A well-posed problem is a problem which satisfies the Hadamard well-posedness conditions: (i) existence of at least one solution to the problem, (ii) uniqueness of this solution and (iii) stability of this solution with respect to data [22]. In this work, the variational formulation is used to find under which conditions a weak solution of (2) is guaranteed to satisfy Hadamard conditions.

We first introduce the separable Hilbert spaces of interest:

$$* L^2(\Omega) = \left\{ f : \int_{\Omega} f^2 \partial\Omega < \infty \right\} \text{ and its usual norm } \|\cdot\|_0$$

$$\text{defined by: } \|f\|_0 = \left(\int_{\Omega} f^2 \partial\Omega \right)^{1/2}.$$

* $H_{0,\{1\}}^1(\Omega) = \{f \in L^2(\Omega) : f|_1 = 0, \nabla f \in L^2(\Omega)\}$ endowed with the inner product (\cdot, \cdot) and the norm $\|\cdot\|_1$, defined respectively as follows:

$$\begin{cases} \forall f, g \in (H_{0,\{1\}}^1(\Omega))^2 : (f, g) = \int_{\Omega} f g \partial\Omega. \\ \forall f \in H_{0,\{1\}}^1(\Omega) : \|f\|_1 = \|f\|_0 + \|\nabla f\|_0. \end{cases} \quad (4)$$

The spaces $H_{0,\{1\}}^1(\Omega)$, $L_2(\Omega)$ and $H^{-1}(\Omega)$ form a "Gelfand triple" (see [23], chapter 5), where $H^{-1}(\Omega)$ is the dual space of $H_{0,\{1\}}^1(\Omega)$.

Let X_e be a special Banach space defined as: $\forall x \in \Omega, \forall t \in [0, t_f]$

$$X_e = \left\{ f \in L^2(0, t_f; L^2(\Omega)) : \exists c_1, c_2 \in \mathbb{R}_+^* : c_2 > f(x, t) > c_1 \right\}$$

To get the variational formulation of this problem, we first multiply equation (2) by a test function $v(z) \in H_{0,\{1\}}^1(\Omega)$ and then integrate on Ω :

$$\begin{aligned} \int_0^1 \frac{\partial T}{\partial t} v(z) dz &= \int_0^1 \text{div} \left(\chi_e(z, t) \frac{\partial T}{\partial z} \right) v(z) dz \\ &- \frac{1}{\tau} \int_0^1 T(z, t) v(z) dz + \int_0^1 S(z, t) v(z) dz. \end{aligned} \quad (5)$$

Using the Gauss' divergence formula:

$$\int_0^1 \text{div} \left(\chi_e(z, t) \frac{\partial T}{\partial z} v(z) \right) dz = \chi_e(z, t) \frac{\partial T}{\partial z} v(z) \Big|_0^1$$

and:

$$\text{div} \left(\chi_e \frac{\partial T}{\partial z} v(z) \right) = v(z) \text{div} \left(\chi_e \frac{\partial T}{\partial z} \right) + \chi_e \frac{\partial T}{\partial z} \frac{dv}{dz},$$

we obtain the following integration by parts for the divergence term:

$$\int_0^1 \text{div} \left(\chi_e \frac{\partial T}{\partial z} \right) v(z) dz = \chi_e \frac{\partial T}{\partial z} v(z) \Big|_0^1 - \int_0^1 \chi_e \frac{\partial T}{\partial z} \frac{dv}{dz}.$$

Taking into account the boundary conditions, equation (5) becomes:

$$\begin{aligned} \int_0^1 \frac{\partial T}{\partial t} v(z) dz &= - \int_0^1 \chi_e(z, t) \frac{\partial T}{\partial z} \frac{dv}{dz} dz \\ &- \frac{1}{\tau} \int_0^1 T(z, t) v(z) dz + \int_0^1 S(z, t) v(z) dz. \end{aligned} \quad (6)$$

Since we are dealing with evolution equations, it is convenient to adopt the following viewpoint. Assume that for every $t \in [0, t_f]$ (or at least almost everywhere in $[0, t_f]$) the function $u(z, t)$ belongs to the Hilbert space $H_{0,\{1\}}^1(\Omega)$. Then the function $u(t)$ is considered as a function of the real variable t with values in $H_{0,\{1\}}^1(\Omega)$:

$$u : [0, t_f] \longrightarrow H_{0,\{1\}}^1(\Omega). \quad (7)$$

We can then write $T(t)$ and $\dot{T}(t)$ instead of $T(z, t)$ and $\frac{\partial T}{\partial t}(z, t)$. Accordingly $\chi_e(z, t) := \chi_e(t)$, $S(z, t) := S(t)$.

From the variational formulation (6), the bilinear form $b(T, v; \chi_e)$ defined on $(H_{0,\{1\}}^1(\Omega))^2$ which gets its values in \mathbb{R} is given by:

$$b(T, v; \chi_e) = \int_0^1 \chi_e(t) \frac{\partial T}{\partial z} \frac{dv}{dz} dz + \frac{1}{\tau} \int_0^1 T(t) v(z) dz, \quad (8)$$

and $L(v; S)$ is a linear form defined on $L^2(0, t_f; H^{-1}(\Omega))$ such that:

$$\forall v \in H_{0,\{1\}}^1(\Omega), L(v; S) = \int_0^1 S(t) v(z) dz. \quad (9)$$

Equation (6) can then be written as:

$$\begin{cases} (\dot{T}, v) + b(T, v; \chi_e) = L(v; S); \quad \forall v \in H_{0,\{1\}}^1(\Omega), \\ T(0) = 0, \end{cases} \quad (10)$$

which is the weak formulation of problem (2).

Theorem 2.1: For all χ_e in X_e , τ in \mathbb{R}_+^* and S in $L^2(0, t_f; L^2(\Omega))$, the system (10) admits a unique solution in $L^2(0, t_f; H_{0,\{1\}}^1(\Omega))$ which is stable with respect to the data (χ_e, τ, S) .

Proof See Appendix I.

To get the classical solution of system (2) such that $T(z, t) \in C^1(0, t_f; C^2(\Omega))$, *i.e.* continuously differentiable once over the time range $[0, t_f]$ and twice over the space range Ω the following regularity conditions are required:

$$\begin{cases} \chi_e \in C^0(0, t_f; C^0(\Omega)) \cap X_e, \\ S \in C^0(0, t_f; C^1(\Omega)) \cap L^2(0, t_f; L^2(\Omega)). \end{cases}$$

Note that these regularity conditions are not restrictive and are consistent with the physical properties of χ_e and S .

Finding the well-posedness conditions of the heat model (2) ensures the existence, uniqueness and stability of its solution and allows us to derive an approximate solution to this direct problem. In the following we will investigate the well-posedness conditions of the inverse problem: given the measurements of T , is it possible to find a unique $\chi_e \in Q_{\chi_e} \subset X_e$ and $S \in Q_S \subset L^2(0, t_f; L^2(\Omega))$ such that the PDE (2) is satisfied.

The answers to the direct and inverse problems guarantee the bijection of the mapping relating T to its parameters χ_e and S . The bijection proof will be brought in the finite dimensional framework as we design an early lumping approach for parameters' estimation. Note that it is well known that even if we are able to prove the existence and uniqueness of the inverse problem solution, its stability is not guaranteed, due to measurement noises. This is an inherent difficulty for inverse problems.

III. EARLY LUMPING APPROACH FOR THE JOINT DIFFUSION AND INPUT ESTIMATION

In early lumping approaches, the PDE is first converted into a finite dimensional system and then an estimation method is used to recover the unknown variables.

In this section, we first prove the convergence of the approximate finite dimensional estimation problem to the infinite dimensional initial one. Then, the b-spline Galerkin formulation and the projection (or parameter interpolation) method are combined to transform the PDE into an approximate set of ordinary differential equations (ODEs). This set provides a reduced state-space model with unknown time-varying inputs and parameters. Finally, an estimation method based on the Kalman filter approach is used to identify the unknown parameters.

A. Convergence of the finite dimensional estimation problem

First, we assume that the forms b and L defined in the previous section are continuous and $H_{0,\{1\}}^1(\Omega)$ -coercive uniformly in $\chi_e \in Q_{\chi_e}$ and $S \in Q_S$. Let $q = (\chi_e, S) \in Q := Q_{\chi_e} \times Q_S$ denote the unknown parameter vector.

The identification problem can be formulated as a standard output least squares optimization problem: finding the parameters which give the best fit of the parameter-dependent solution of the partial differential equation to the system response measurements. Thus, given the experimental temperature profile $\{T_{exp}(t_i)\}_{i=1}^{N_t}$ corresponding to the PDE

(2) with sampled observations at each time t_i , we aim at minimizing the least square output functional over $q \in Q$:

$$\min_{q \in Q} J(q, T_{exp}) = \sum_{i=0}^{N_t} \int_0^1 (T_{exp}(t_i) - T(t_i; q))^2 dz, \quad (11)$$

and satisfying the model constraints (2) and where $T(t_i; \chi_e; S)$ is the parameter-dependent solution of (2). Note that distributed sensing of the system state T is assumed to be available.

The minimization of (11) involves an infinite dimensional state space $H_{0,\{1\}}^1(\Omega)$ and an infinite dimensional admissible parameter set Q . To construct a method which can be implemented without loosing the parameters distribution property, we consider Galerkin type approximations in the context of the variational formulation (10). Thus, let H^N be a sequence of finite dimensional subspaces of $H_{0,\{1\}}^1(\Omega)$ and Q^{n_p} a sequence of finite dimensional subsets of Q . The approximate solution of T denoted T^N is obtained using the orthogonal projection P^N of $H_{0,\{1\}}^1(\Omega)$ onto H^N as follows:

$$\begin{aligned} P^N : H_{0,\{1\}}^1(\Omega) &\longrightarrow H^N \\ T &\longmapsto T^N \end{aligned}$$

and T^N is the solution of the finite dimensional approximation of (10) given by:

$$\begin{cases} (T^N, v) + b(T^N, v; \chi_e^r) = L(v; S^d), \quad \forall v \in H^N(\Omega), \\ T(0) = 0. \end{cases} \quad (12)$$

Note $q^{n_p} := [\chi_e^r, S^d]$ the vector of the approximate parameters with $n_p = r + d$. Thereby, the minimization problem of (11) is converted to a family of approximation estimation problems with finite dimensional states and parameters, and the finite dimensional least squares criterion minimization is given by:

$$\min_{q^{n_p} \in Q^{n_p}} J^N(q^{n_p}, T_{exp}) = \sum_{i=0}^{N_t} \int_0^1 (T_{exp}(t_i) - T^N(t_i; q^{n_p}))^2 dz, \quad (13)$$

under constraints (12).

In order to prove the convergence of the approximate problem (12) and (13) toward the original one involving (10) and (11), we have first to attest under which conditions $T^N(t, q^{n_p})$, solution of the problem (12), converges to $T(t; q)$, solution of the variational problem (10). Then, it is necessary to show that for some assumptions on the parameter space Q (and Q^{n_p}), the sequence $\{q^{n_p}\}$ (or a subsequence of $\{q^{n_p}\}$) solution of (13) converges to the solution of (11). In what follows, $H_{0,\{1\}}^1$ is denoted H , and since we are dealing with functional parameters, the parameter space Q is considered as a Banach space endowed with a norm denoted $|\cdot|_Q$.

Theorem 3.1: Consider the state spaces H (and H^N) previously defined and assume that they satisfy:

$$\forall \psi \in H : \|\psi - P^N \psi\|_1 \rightarrow 0 \quad \text{for } N \rightarrow \infty. \quad (14)$$

Suppose that the form b (and L) defined in (8)-(9) in addition to being continuous, bilinear (linear) symmetric and H -coercive, satisfy the following inequalities:

$$\forall T, v \in H : |b(T, v; \chi_e) - b(T, v; \chi_e^*)| \leq \gamma_1 |\chi_e - \chi_e^*|_Q \|T\|_1 \|v\|_1 \quad (15)$$

$$\forall v \in H : |L(v; S) - L(v; S^*)| \leq \gamma_2 |S - S^*|_Q \|v\|_1. \quad (16)$$

Let q^{n_p} be an arbitrary sequence in the parameter space Q such that: $\lim_{n_p \rightarrow \infty} q^{n_p} = q$, ($q \in Q$).

Then the following result holds:

$$\lim_{\substack{N \rightarrow \infty \\ n_p \rightarrow \infty}} T^N(t; q^{n_p}) \rightarrow T(t; q)$$

Proof See Appendix II.

Once the convergence conditions of problem (12) to (10) are established, conditions under which the solution of the approximate estimation problem (12) and (13) converges to the solution of the initial infinite dimensional one, given in (10) and (11), have to be found. The following proposition answers to this question.

Proposition 3.2: Let Q and Q^{n_p} be two compact spaces endowed with a norm denoted $|\cdot|_Q$. Let i^{n_p} be a "mapping" from Q into Q^{n_p} such that:

$$\begin{aligned} i^{n_p} : Q &\longrightarrow Q^{n_p} \\ q &\longmapsto i^{n_p}(q), \end{aligned} \quad (17)$$

and suppose that $i^{n_p}(q) \rightarrow q$ when $n_p \rightarrow \infty$ uniformly in Q .

If $\{q^r\}_{r \in \mathbb{N}^*}$ is a sequence solution of (12) and (13), $q^r \rightarrow \bar{q}$ ($r \rightarrow \infty$) and $T^N(t; q^r) \rightarrow T(t; q)$, \bar{q} is a solution of (10) and (11).

Proof See Appendix III.

The key of this proof comes from the fact that parameter spaces Q and Q^{n_p} are Banach compact spaces. For more details on the choice of approximate spaces H^N and Q^{n_p} , see [24].

The compactness property plays a critical theoretical and computational role in proving the convergence of the approximate finite dimensional solution of the problem (13) to the solution of the initial problem (11). From a computational viewpoint, the compactness criterion is reduced to uniform constraints on the admissible parameters and their derivatives. It can be added explicitly by imposing these constraints in the estimation algorithm as it was done in [25] where it was shown that the compactness ensures the convergence of the resulting algorithms, or implicitly using the *Tikhonov regularisation* as it was discussed for example by [26] and [27].

Remark:

In the early lumping approach, the choice of the basis generating the approximate state-space H^N and the parameter space Q^{n_p} is of paramount importance. It is not only a question of the approximate spaces dimensions but even the properties of the basis and thus the method of discretization

itself. In [28] there is an exhaustive bibliography on Grey-box modelling in finite dimensional framework, where it is shown that many discretization methods were developed and implemented in various applications. It illustrates that an a priori knowledge of the studied system is needed in order to guarantee the effectiveness of the chosen discretization technique. This means that some experience has to be collected to be able to choose the discretization technique.

B. Building the approximate parameter estimation problem

The Galerkin formulation belongs to the weighted residual methods (WRM) which is one of the most often used and most efficient lumping methods [28]. It is an extension of the eigenfunction method where the basis functions are chosen such that the residual of the PDE approximation is made orthogonal to each basis function. Thus, in the Galerkin method the approximate solutions T^N , v^N of the variational form established in (12) are written as:

$$\begin{cases} T^N(z, t) = \sum_{k=0}^N x_k(t) \omega_k(z), \\ v^N(z) = \sum_{k=0}^N \kappa_k \omega_k(z), \end{cases} \quad (18)$$

where $\{\omega_i\}_{i=0}^N$ is the sequence of basis functions spanning H^N and $\{x_k(t)\}_{k=0}^N$, $\{\kappa_k\}_{k=0}^N$ is the sequence of the corresponding weighting functions. The key is to select appropriate basis functions satisfying (14) and to construct the finite-order (reduced-order) temporal model. Several choices are available [28].

By replacing (18) in (12), we get the following expression for $i = 0, \dots, N$:

$$\sum_{i=0}^{i=N} \left(\int_0^1 w_i(z) w_j(z) dz \right) \dot{x}_i(t) = \int_0^1 S(z, t) w_j(z) dz \quad (19)$$

$$- \sum_{i=0}^{i=N} \left[\int_0^1 \chi_e(z, t) \frac{dw_i}{dz} \frac{dw_j}{dz} dz + \frac{1}{\tau} \int_0^1 w_i(z) w_j(z) dz \right] x_i(t).$$

By introducing the state vector $X(t) = [x_0(t), x_1(t), \dots, x_N(t)]^T$ ($X \in \mathbb{R}^{N+1}$), the stiffness matrix $M \in \mathbb{R}^{N+1 \times N+1}$, non-singular (by definition), symmetric and diagonally dominant, the transition matrix $A(t) \in \mathbb{R}^{(N+1) \times (N+1)}$ and the input vector $B(t) \in \mathbb{R}^{N+1}$ given by:

$$M = \begin{pmatrix} \int_0^1 w_0^2(z) dz & \dots & \int_0^1 w_0(z) w_N(z) dz \\ \vdots & \ddots & \vdots \\ \int_0^1 w_0(z) w_N(z) dz & \dots & \int_0^1 w_N^2(z) dz \end{pmatrix},$$

$$A(t) = \underbrace{\begin{pmatrix} \int_0^1 \chi_e(z, t) w_0'^2(z) dz & \dots & \int_0^1 \chi_e(z, t) w_0'(z) w_N'(z) dz \\ \vdots & \dots & \vdots \\ \int_0^1 \chi_e(z, t) w_0'(z) w_N'(z) dz & \dots & \int_0^1 \chi_e(z, t) w_N'^2(z) dz \end{pmatrix}}_{A_1(t)} + \frac{1}{\tau} M,$$

$$B(t) = \begin{pmatrix} \int_0^1 S(z, t) w_0(z) dz \\ \vdots \\ \int_0^1 S(z, t) w_N(z) dz \end{pmatrix},$$

where the symbol $(\cdot)'$ denotes the derivative operator with respect to the spatial variable z , the equation (19) is transformed to a state-space continuous model given by:

$$\begin{cases} \dot{X}(t) = -M^{-1} \left(A_1(t) + \frac{1}{\tau} M \right) X(t) + M^{-1} B(t), \\ X(0) = 0. \end{cases} \quad (20)$$

Since the bilinear form b is coercive, the transition matrix $A(t)$ is symmetric, definite and positive. Note that this model provides the approximate solution to the direct problem. It is based on the knowledge of χ_e and S .

To construct efficiently the approximation spaces H^N and Q^{n_p} , the cubic b-splines functions are used to span them. They naturally fulfil the requirements of (14), (15) and (16) (see [24], chapter 4).

For Q^{n_p} , standard cubic b-splines [29] are utilized whereas modified ones which satisfy the boundary conditions of (2) have to be formulated to generate H^N .

The cubic b-splines finite element method is one of the widely used methods that provide accurate approximate solutions of class $C^2(\Omega)$. In the following, this technique is briefly summarized (details can be found in [8], [30], [31]). Specifically, let $\{z_i\}_{i=0}^N$ be a uniform mesh of Ω , $\Delta z = 1/N$ the mesh size and $\{\pi_i\}_{i=0}^N$ the sequence of standard piecewise cubic b-splines functions that vanish outside $]z_{i-2}, z_{i+2}[$, has value 4 and slope 0 at z_i , value 1 and slope $-3/\Delta z$ at z_{i+1} , and value 1 and slope $3/\Delta z$ at z_{i-1} for all π_i , $i = 0, \dots, N$.

This reasoning is applied to the functional parameters χ_e and S where the dimensions of their approximate spaces may differ. For χ_e we choose a basis of dimension r while for S the basis dimension is equal to d . The values of r and d are not free since we have to impose that $r + d \leq N + 1$ to ensure the structural identifiability of the estimation problem [32]. This leads to the following approximate parameter expressions:

$$\chi_{e_r}(z, t) = \sum_{k=1}^r \lambda_k(t) \pi_k(z) = \theta^T(t) P(z), \quad (21)$$

where:

$$\begin{cases} \theta^T(t) = [\lambda_1(t), \lambda_2(t), \dots, \lambda_r(t)] \quad , \quad \theta \in \mathbb{R}^r \\ P(z) = [\pi_1(z), \pi_2(z), \dots, \pi_r(z)]^T \end{cases}$$

The source term writes as:

$$S_d(z, t) = \sum_{k=1}^d \zeta_k(t) \pi_k(z) = \mathcal{B}(z) \beta(t). \quad (22)$$

where :

$$\begin{cases} \beta(t) = [\zeta_1(t), \zeta_2(t), \dots, \zeta_d(t)]^T \quad , \quad \beta \in \mathbb{R}^d \\ \mathcal{B}(z) = [\pi_1(z), \pi_2(z), \dots, \pi_d(z)] \end{cases}$$

To span the space H^N as it was noted above some modifications on the basis functions has to be done. In fact, the modified cubic b-splines basis elements $\{\omega_i(z)\}_{i=0}^N$ have the following expression [30]:

$$\omega_i(z) = \begin{cases} \pi_0(z), & \text{if } i = 0, \\ \pi_{-1}(z) + \pi_1(z), & \text{if } i = 1, \\ \pi_i(z), & \text{for } i = 2, \dots, N-2, \\ \pi_{N-1}(z) - \pi_{N+1}(z), & \text{if } i = N-1, \\ \pi_N(z) - 4\pi_{N+1}(z), & \text{if } i = N. \end{cases} \quad (23)$$

Replacing (21) and (22) in (20)'s matrices, the PDE given in (2) is converted to a set of ordinary differential equations (ODE) defining a linear time-varying state-space time-continuous model (using the combined Galerkin and projection method). This leads to:

$$\begin{cases} \dot{X}(t) = \left(M^{-1} A(\theta^T(t)) + \frac{1}{\tau} I_{N+1} \right) X(t) + M^{-1} \mathcal{D} \beta(t) \\ X(0) = 0 \\ y(t) = X(t) \end{cases} \quad (24)$$

where the new form of the transition (or damping) matrix $A(t) \in \mathbb{R}^{(N+1) \times (N+1)}$ writes as:

$$A(\theta^T(t)) = -\theta^T(t) \otimes \begin{pmatrix} \int_0^1 P(z) \omega_0'^2(z) dz \cdots \int_0^1 P(z) \omega_0'(z) \omega_N'(z) dz \\ \vdots \\ \int_0^1 P(z) \omega_0'(z) \omega_N'(z) dz \cdots \int_0^1 P(z) \omega_N'^2(z) dz \end{pmatrix}$$

and the input vector B is converted into a known input matrix $\mathcal{D} \in \mathbb{R}^{(N+1) \times d}$ multiplying the unknown input vector $\beta(t)$. The matrix \mathcal{D} has the following form:

$$\mathcal{D} = \begin{pmatrix} \int_0^1 \omega_1(z) \pi_1(z) dz & \cdots & \int_0^1 \omega_1(z) \pi_d(z) dz \\ \vdots & \ddots & \vdots \\ \int_0^1 \omega_N(z) \pi_1(z) dz & \cdots & \int_0^1 \omega_N(z) \pi_d(z) dz \end{pmatrix}$$

Remark:

The FEM-Galerkin formulation is an efficient method for linear stable parabolic equations, and the B-splines are becoming (if not they already are) standard approximate functions that provide a unique C^2 solution and are widely used in many domains, such as interpolation and curve fitting. It is important to point out that once the discretization method is chosen it is, at the best knowledge of the authors, impossible to change it when the estimation procedure starts.

Remark:

The discretization method has an important impact on the approximate state-space model and the solution of the approximate inverse problem characteristics. It has to ensure that the model structure is structurally and numerically identifiable in order to guarantee the well-posedness of the

estimation solution. As it is a discretization-based estimation method, in order to end with a good approximation close to the description of the original infinite dimensional phenomenon, the dimensions of the approximate space H^N have to be chosen such that the major system dynamics are captured. This may lead to a high order model and which will complicate the estimation problem. Conversely, a reduced order H^N implies a limited basis dimension for Q^{n_p} . Hence, to get a structurally identifiable estimation problem, the approximation of the functional parameters χ_e and S is constrained by the H^N dimension. The number of unknowns $r + d$ has to be less than $N + 1$ the number of the equations. A trade-off between the parameters approximation accuracy and the number of spatial sensors has to be made. Nevertheless, the early lumping approach is widely used and developed in the literature [28] as it is reasonably easy to implement. Many softwares integrating different FEM techniques are available.

IV. JOINT ESTIMATION OF UNKNOWN PARAMETERS AND INPUTS USING THE EXTENDED KALMAN FILTER FOR UNKNOWN INPUTS WITHOUT DIRECT FEED-THROUGH (EKF-UI-WDF)

Once the approximate finite dimensional estimation problem is built and its convergence to the original one is established, an estimation method can be applied to retrieve the unknown parameter vector θ and the input vector $\beta(t)$ of the reduced model (24). Note that this problem is different from what is usually considered in unknown input observers [33], [11], [12]. Our objective is not the estimation of θ in the presence of some unknown inputs β (the free-unknown inputs estimation) but to infer the values of both θ and β from the measurements.

In [19], a Kalman filter based estimator which guarantees, under the model observability condition, the global optimality for both state and unknown inputs estimation in a least square sense (LS), was proposed. This filter was developed for a discrete state space model with unknown inputs and without direct feed-through. Hence it was named the Kalman Filter for Unknown Inputs Without Direct Feedthrough (KF-UI-WDF). Unlike the minimum-variance unbiased (MVU) estimators ([12], [34], [35]) where only the MVU optimality is provided, the KF-UI-WDF is a natural extension of the Kalman filter (KF) to the unknown inputs estimation problem without any constraints (a pre-determined input gain matrix) and conserves all the KF performances and practical knowledge [19]. By minimizing a weighted least squares objective function with respect to an extended variable including the actual states and all unknown inputs from the initial measurement time t_0 to the actual one t_k , a recursive least squares estimation (LSE) approach is performed and the KF-UI-WDF is derived.

In this section the KF-UI-WDF is extended to estimate also the parameter θ using the same philosophy as the EKF [36]. To this end, we first extend the state vector in (24) to include the unknown parameter vector θ and then discretize the system dynamics.

Define:

$$f(X, \theta, \beta, k) = \begin{pmatrix} (I + dtM^{-1}A(\theta^T(k)) - \frac{1}{\tau}I_{N+1})X(k) + dtM^{-1}\mathcal{D}\beta(k) \\ \theta(k) \end{pmatrix}$$

The discrete extended model is given by:

$$\begin{cases} x_{ext}(k+1) = f(X, \theta, \beta, k) + w(k) \\ y(k) = C x_{ext}(k) + v(k) \end{cases} \quad (25)$$

where $x_{ext}(k+1) = \begin{pmatrix} X(k+1) \\ \theta(k+1) \end{pmatrix}$ is the extended state, $C = [I_{N+1} \ 0]$ is the observation matrix, dt is the time step, $w(k) \in \mathbb{R}^{N+1+r}$ and $v(k) \in \mathbb{R}^{N+1}$ are respectively the model uncertainty and the measurements noise vectors, assumed to be independent, white and Gaussian. They are characterized by:

$$\begin{cases} \mathbb{E}(w) = \mathbb{E}(v) = 0, \\ \mathbb{E}(w(i) w^T(j)) = W\delta_{ij}, \\ \mathbb{E}(v(i) v^T(j)) = V\delta_{ij}, \\ \mathbb{E}(v(i) w^T(j)) = 0. \end{cases} \quad (26)$$

Based on the above representation (25)-(26), the EKF-UI-WDF approach can be used to estimate the extended state $\hat{x}_{ext}(k|k)$ and the inputs $\hat{\beta}(k-1|k)$, given all the available observations (prior and including time k). The EKF-UI-WDF algorithm is given by Table I, where: \mathbb{E} is the expectation operator, K is the extended state gain matrix, U is the input gain matrix, $\hat{x}_{ext}^-(k+1)$ is the a priori estimate of $x_{ext}(k+1)$, $\hat{\beta}^-(k+1)$ is the a priori estimate of $\beta(k+1)$ and $\hat{x}_{ext}^+(k+1)$ is the a posteriori estimate of $x_{ext}(k+1)$. The a priori and a posteriori estimates of a random variable α at time k are defined as:

$$\begin{cases} \hat{\alpha}^-(k) = \mathbb{E}[\alpha(k) | y(1), \dots, y(k-1)] \\ \hat{\alpha}^+(k) = \mathbb{E}[\alpha(k) | y(1), \dots, y(k)] \end{cases}$$

The EKF asymptotic convergence for observable systems is proved in [38]. In [19], the optimality conditions are analysed. The only restriction of this filter is to impose that the dimension of the outputs has to be larger than that of the inputs ($N+1 > d$), to ensure the uniqueness of the estimated variables. For the extended case, $N+1$ has to be larger than or equal to $r + d$, where r is the length of the parameters vector.

Unfortunately, like the KF, the limitations of this filter are the hypotheses on the model and measurements noises (only additive noises) and the need of a perfect knowledge of the covariance matrices W and V . Nevertheless, the EKF has proved its performances in practice even with some missing knowledge, provided that all the implementation steps listed in [39] are respected.

EKF-UI-WDF Algorithm
Initialize:
$\hat{x}_{ext}^+(0) = \mathbb{E}(x_{ext}(0)), W = \mathbb{E}[w w^T], R = \mathbb{E}[v v^T],$
$P^+(0) = \mathbb{E}[(x_{ext}(0) - \hat{x}_{ext}^+(0) (x_{ext}(0) - \hat{x}_{ext}^+(0))^T],$
for $k = 0$ to $N - 1$:
Time update equations:
$\bar{x}_{ext}(k+1) = f(\hat{x}_{ext}(k), dt) _{\beta=0}$
$P^-(k+1) = A_k P^+(k) A_k^T + W$
$K(k+1) = P^-(k+1) C^T [C P^-(k+1) C^T + V]^{-1}$
$U(k+1) = [(D)^T C^T R^{-1} (I - C K(k+1)) C (D)]^{-1}$
$\hat{\beta}^-(k+1) = U(k+1) \mathcal{D}^T C^T R^{-1} (I - C K(k+1))$
$[y_{exp}(k+1) - C \bar{x}_{ext}(k+1)]$
$\hat{x}_{ext}^-(k+1) = f(\hat{x}_{ext}(k), \hat{\beta}^-(k), dt)$
Measurement update equations:
$\hat{x}_{ext}^+(k+1) = \hat{x}_{ext}^-(k+1) + K(k+1)$
$[y_{exp}(k+1) - f(\hat{x}_{ext}^-(k+1), \mathcal{D} \hat{\beta}^-(k+1))]$
$P^+(k+1) = (I - K_x(k+1)C)$
$[P^-(k+1) + \mathcal{D}U(k+1)\mathcal{D}^T(I - K(k+1)C)^T]$
end

TABLE I: EKF-UI-WDF's algorithm [37]

In this paper, since the estimation is considered in a deterministic framework, the covariance matrices W and V are used as tuning parameters.

For the EKF, it is well known that we can, at most, have asymptotic convergence properties for well-behaved systems (small size, well-scaled, well-conditioned,...) and unfortunately divergence in the opposite case.

In subsection III-A, it has been shown that the compactness of the approximate parameter space is primordial even from a numerical point of view to prove the convergence of the estimation method. In the EKF-UI-WDF, this can be incorporated using ideas developed in [39]. In MATLAB, this can be implemented using the Optimization Toolbox with the *fmincon* solver.

V. CASE OF TIME VARYING PARAMETERS: THE ADAPTIVE EKF-UI-WDF

The EKF-UI-WDF as it is formulated can not take into account time-variations of the parameters. As mentioned in [40], adding a fading memory parameter may recover this problem. In [41], based on Taylor expansion, a technique to estimate time varying parameters in the case of RLS estimation was proposed. The generalization of this method to Kalman filtering is not an easy task. In the other hand, expanding a parameter vector multiplies its dimension which may complicate its estimation.

In the literature, several formulations of the fading memory parameter are available and a summary can be found in [42]. In this section, the adaptive tuning law of the fading memory parameter presented in [43] is chosen. In this method, the fading memory parameter is tuned only by the parameter time-variations contribution independently of the measurement noise effect. Therefore, in the EKF-UI-WDF algorithm presented in Table I, the *a priori* estimation error covariance matrix is modified as follows:

$$P^-(k) = \lambda(k) (A_e(k-1) P^+(k-1) A_e(k-1)^T) + W_e(k),$$

where $\lambda(k)$ is the adaptive fading memory parameter computed using [43]'s technique (see [37] for more details on how to implement this technique in a Kalman filtering approach). Throughout this paper, the EKF-UI-WDF using [43]'s method is called the adaptive EKF-UI-WDF.

Remark:

For the EKF-UI-WDF, as it was discussed, since it is a Kalman-based filter it inherits all its advantages and disadvantages. If only state and input simultaneous estimation in a linear state-space model with perfect knowledge on the noise and model uncertainties covariances is considered, then optimality in MVU and LSE is guaranteed even in the presence of a coloured noise [39]. In the nonlinear case (the extended state to parameters), only asymptotic convergence is assumed [38]. However, for mild nonlinear problems such as the simultaneous state and parameters estimation, the EKF has proved in practice to be an efficient algorithm.

VI. SIMULATION AND EXPERIMENTAL RESULTS

A. Using simulated data

Simulated data is generated based on the *a priori* qualitative knowledge on χ_e and S experimental profiles. It was reported in plasma physics that the diffusion coefficient has an increasing profile from the center to the edge in the validity domain and that S has a Gaussian form (see [21], [44] and [45]). In the following simulations, χ_e and S are described by:

$$\begin{cases} \chi_e(z, t) = (1 + 9z - 36z^2 + 32z^3) g(t), \\ S(z, t) = \frac{10^6}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(z-\mu)^2}{2\sigma^2}\right) f(t), \\ z \in [0, 1], t \in [0, 25], dz = 0.05, dt = 0.01, \tau = 0.1, \end{cases} \quad (27)$$

where the reaction coefficient τ is assumed to be constant, $g(t)$ and $f(t)$ model respectively the time variations of χ_e and S .

To prove the EKF-UI-WDF tracking abilities using the adaptive fading memory technique developed in [43], a space-time varying χ_e is assumed in simulation even if in plasma physics, only its space variations are considered ([5], [44], [2]).

Two cases are investigated in this section;

In **Example 1**, χ_e has a slowly time variation to which the temperature profile T is insensitive.

In **Example 2**, χ_e has a time sinusoidal form.

In both cases, the time variations of the source term S are described by a rectangular periodic signal given by:

$$f(t) = \begin{cases} 2 & \text{if } t \leq \alpha t_{period}, \\ 1 & \text{if } t \in [\alpha t_{period}, t_{period}] \\ f(t + t_{period}) = f(t). \end{cases} \quad (28)$$

where α is the signal duty cycle (equal to 0.4 in our simulations) and t_{period} is the signal $f(t)$ period. The formula of $g(t)$ will be presented for each example.

The initial variables of the adaptive EKF-UI-WDF algorithm are selected as: $\mu = 0.9$, $\hat{P}(0|0) = 10^4 I_{r+N+1}$, $W = 10^{-2} I_{N+r+1}$, $V = 10^{-4} I_{N+1}$, $\hat{\theta}(0)$ and $\hat{\beta}(0)$ are arbitrary. However, it should be noted that there are no constraints on the fading memory variable λ and various trial and error tests were necessary to tune μ , W and V .

1) *Example 1: $\chi_e(z, t)$ is slowly time varying:* First, we start with the profiles presented in Fig. 1, generated using equations (27) and (28) with:

$$g(t) = 0.1 + \frac{1}{1 + e^{-20*(t-5)}}, \quad t \in [0, 25]. \quad (29)$$

Thus, $g(t)$ is a sigmoid function modelling a smooth jump of about 10% at $t = 5s$. From Fig. 1, it is clear that the time variations of T are mainly due to the source term S , whereas the time variation's impact of χ_e on T profile is invisible. This choice of χ_e is a challenging task to the adopted algorithm as it tests its capacities to estimate such a small variation. From the estimation results given in Fig. 2, the adaptive EKF-UI-WDF performs well and is able to estimate both χ_e and S . After each input jump, the filter estimates converge rapidly to the real parameters.

2) *Example 2: $\chi_e(z, t)$ and $S(z, t)$ are time varying:* The purpose of this example is to prove that even when χ_e is a time varying coefficient (and the influence of this variation is visible on the T profile), the adaptive KF-UI-WDF is able to estimate efficiently the unknown parameters (χ_e and S) without modifying the previous initial variables.

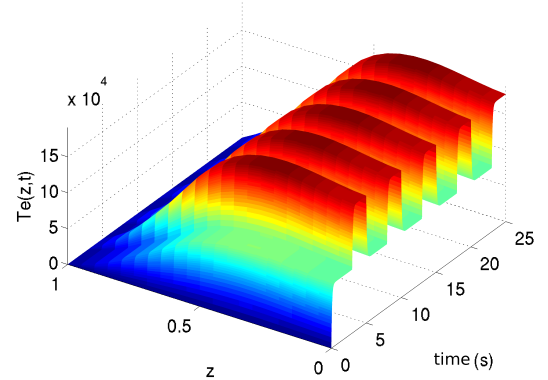
In this example, the χ_e time profile $g(t)$ is given by:

$$g(t) = \sin(2\pi t(2/t_f)) + 2, \quad (30)$$

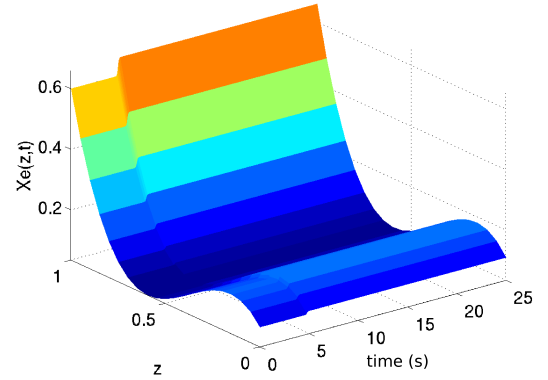
where t_f is the final time. The profiles of T , χ_e and S are illustrated in Fig. 3. Comparing the results presented Fig. 4 to those in Fig. 2 shows that the adaptive EKF-UI-WDF conserves qualitatively its properties even if the relative estimation errors are more important in this example, but still acceptable ($\leq 1\%$ between two input jumps).

B. Using experimental data

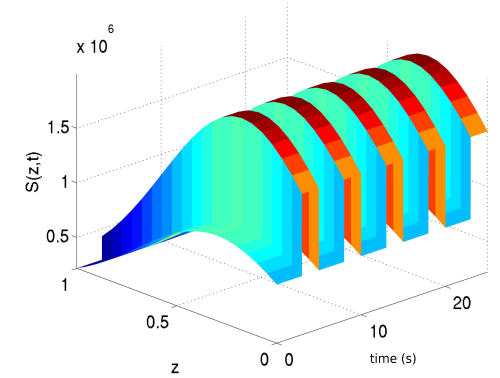
The chosen estimation strategy is implemented on real data provided by the Tore Supra tokamak. It is a large tokamak with supraconducting toroidal magnets of minor radius $a = 0.72m$, major radius $R = 2.4m$, toroidal magnetic field $B_T \leq 4T$, circular cross-section and which often runs discharges in the range of 10 s to 30 s. It has



(a) Spatio-temporal profile of $T(z, t)$



(b) $\chi_e(z, t)$ profile



(c) $S(z, t)$ profile

Fig. 1: Example 1: Profiles of $T(z, t)$, $S(z, t)$ and $\chi_e(z, t)$

even obtained discharges of 6 min [46]. The availability of multi-megawatt radio frequency heating and the possibility to vary the injected power during the shot make Tore Supra a unique machine to study plasma transport properties and their dynamics.

On Tore Supra tokamak, electron temperature is measured using Electron Cyclotron Emission (ECE) diagnostics. Based on Rayleigh-Jeans approximation of Planck's black body radiation laws, ECE can be used for spatially and tempo-

rally resolvent measurement of electron temperature T (see [1], [47] and [48] for more details on ECE physics). The standard system employed for T profile measurement is a 32-channel heterodyne ECE radiometer with 1 GHz spacing and 500 MHz bandwidth. Fast acquisition for the profile radiometer is done with a video bandwidth of 42 kHz at a sampling rate of 84 kHz [46]. More relevant details on this detection system, its electronics and data-analysis may be found in [46]. To reduce the thermal noise and improve physical information obtained from these measurements, advanced data-analysis techniques such as correlation ECE are employed. The main features of this technique are discussed in [46], where the ECE diagnostics setup, noise filtering using correlation analysis (a sequence of passband filters with shifted central frequencies followed by a detection to estimate the power spectrum of the measured signals) and estimation of measurement error are also discussed. This diagnostic provides temperature fluctuation measurements with a radial resolution of 1 cm , which supports the hypothesis that the full temperature state is available for estimation purpose.

In this section, we consider the discharge TS-33632 in Fig. 5, where the heating is mainly due to the radio-frequency power at the Ion Cyclotron Resonant Heating (ICRH). For the projection method, as with simulated-data the temperature projection basis is $N = 20$ and for χ_e and S , the orders of the bases are $r = d = 10$.

Fig. 7 presents the estimated profiles of χ_e and S in the spatial validity interval ($z \leq 0.8$). Both are positive without enforcing this constraint in the Kalman filter criterion. From Fig. 6, the EKF-UI-WDF performs well since the relative estimation error of T is strictly less than 1%.

Using the estimated source term, the absorbed power is computed by the following formula:

$$Pe^*(r, t) = \frac{3}{2} n_e(r, t) S(r, t),$$

where n_e is the electron density. The absorbed power P_e is given by:

$$P_e(t) = \iiint_{volume} Pe^*(r, t) dvolume. \quad (31)$$

The tore volume writes as:

$$volume = 2\pi^2 R r^2,$$

where $r = az$. By considering the tokamak axisymmetry, equation (31) becomes:

$$P_e(t) = 6\pi^2 a^2 e \int_0^1 n_e(z, t) S(z, t) dz, \quad (32)$$

where $e = 1.6 \times 10^{-19} J$ is the electron charge ensuring the conversion of eV to J ($1eV = |e| J$).

In Fig. 8, the absorbed power Pe_{estim} computed using (32) is compared to the total power Pe_{input} . This figure shows that the time variations of the absorbed power correspond to those of the input one and gives an idea on the amplitude of the energy losses.

VII. CONCLUSION

In this paper, the problem of space-time parameter estimation in a distributed linear parabolic PDE equation modelling the heat transport in tokamak plasma was considered in the early lumping framework. To get a good approximate reduced order model (finite dimensional) that respects the distributed behaviour of the studied phenomenon, the Galerkin finite element method and the parameter projection/interpolation technique were associated to discretize the PDE and to define a linear time-varying state-space model with unknown parameters and inputs. The EKF-UI-WDF was then chosen to estimate these unknown variables and an adaptive tuning fading memory coefficient was introduced to take into account the time variations of the parameters.

Simulation results on both simulated and experimental data attest the efficiency of the chosen methodology to handle parameter and input estimation for heat transport in tokamak plasmas. However, if the basis dimensions constraint can be covered using late lumping methods (as we have done in [49] and [50]), the problem of the noise measurement effect was not addressed in this paper. This is beyond the scope of our present work since we dealt with filtered measurements using correlation electron cyclotron emission diagnostics (see [46] and references therein). The question of how to develop such a filter or to include the filtering in the tuning of the measurement noise covariance matrix is not an easy task and will may be considered in the future.

APPENDIX I

Proof of Theorem 2.1 First note that the bilinear form b is upper bounded by the following expression:

$$\forall \chi_e \in Xe, \tau \in \mathbb{R}_+^* :$$

$$\begin{aligned} |b(T, v; t)| &\leq \left(\|\chi_e(z, \cdot)\|_0 + \frac{1}{\tau} \right) \|v\|_1 \|T\|_1, \\ &\leq \left(c_2 + \frac{1}{\tau} \right) \|v\|_1 \|T\|_1 \end{aligned} \quad (33)$$

from the Cauchy-Schwarz inequality and norm's $\|\cdot\|_1$ definition and where $|\cdot|$ is the absolute value operator. Second:

$$\forall v \in H_{0, \{1\}}^1(\Omega), \forall \chi_e \in Xe : |b(v, v)| \geq \frac{c_1}{\sqrt{(1 + C(\Omega))}} \|v\|_1$$

where $C(\Omega)$ is the Poincaré constant. Hence, the bilinear form b is continuous and $H_{0, \{1\}}^1$ -coercive.

For the linear form L , we have from (9) that:

$$\forall S \in L^2(0, t_f; L^2(\Omega)) : |L(t)(v)| \leq \|S(x, t)\|_0 \|v(x)\|_1$$

and thus L is continuous. Using the Lax-Milgram theorem ([51], chapter 07) the results of theorem are guaranteed, which conclude the proof.

APPENDIX II

Proof of Theorem 3.1 The proof is in the spirit of the one given in [52]. The difference is in the type of PDE problem considered. In this work, we consider a parabolic PDE while

in [52] it was a hyperbolic PDE.

For all $T \in H$ and $T^N \in H^N$ we have:

$$\begin{aligned} \|T^N(t; q^{n_p}) - T(t; q)\|_1 &\leq \|T^N(t; q^{n_p}) - P^N T(t; q)\|_1 \\ &\quad + \|P^N T(t; q) - T(t; q)\|_1. \end{aligned} \quad (34)$$

and from (14): $\|P^N T(t; q) - T(t; q)\|_1 \xrightarrow{N \rightarrow \infty} 0$ thus if:

$$\|T^N(t; q^{n_p}) - P^N T(t; q)\|_1 \xrightarrow{N \rightarrow \infty} 0 \quad (35)$$

then:

$$\|T^N(t; q^{n_p}) - T(t; q)\|_1 \xrightarrow{N \rightarrow \infty} 0$$

Thus it suffices to show (35) to get the convergence result.

Denote:

$$\begin{aligned} T^N &:= T^N(t; q^{n_p}), \quad T := T(t; q), \\ \Delta^N &:= T^N - P^N T, \quad \dot{\Delta}^N := \dot{T}^N - \frac{d}{dt} P^N T = \dot{T}^N - P^N \dot{T}. \end{aligned}$$

Considering (10) and (12):

$$\begin{aligned} (\dot{\Delta}^N, \psi) &= \left(\dot{T}^N - \dot{T} + \dot{T} - P^N \dot{T}, \psi \right) \\ &= \left(\dot{T} - P^N \dot{T}, \psi \right) + L(\psi; S^d) \\ &\quad - b(T^N, \psi; \chi_e^r) - L(\psi; S) + b(T, \psi; \chi_e). \end{aligned} \quad (36)$$

In the other hand, $b(T^N, \psi; \chi_e^r)$ can be written as:

$$b(T^N, \psi; \chi_e^r) = b(\Delta^N, \psi; \chi_e^r) + b(T, \psi; \chi_e^r) - b(T - P^N T, \psi; \chi_e^r), \quad (37)$$

so (36) becomes:

$$\begin{aligned} (\dot{\Delta}^N, \psi) + b(\Delta^N, \psi; \chi_e^r) &= \left(\dot{T} - P^N \dot{T}, \psi \right) + L(\psi; S^d) \\ &\quad + b(T, \psi; \chi_e) - b(T, \psi; \chi_e^r) \\ &\quad - L(\psi; S) + b(T - P^N T, \psi; \chi_e^r). \end{aligned} \quad (38)$$

Equality (38) is satisfied for any test function ψ in H (or H^N) and thus it is true for $\psi = \Delta^N$, and using the fact that

$$\left(\dot{\Delta}^N, \Delta^N \right) = \frac{1}{2} \frac{d}{dt} \|\Delta^N\|_1^2, \quad (38) \text{ is equivalent to:}$$

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\Delta^N\|_1^2 + b(\Delta^N, \Delta^N; \chi_e^r) &= \left(\dot{T} - P^N \dot{T}, \Delta^N \right) \\ &\quad + L(\Delta^N; S^d - S) \\ &\quad + b(T, \Delta^N; \chi_e - \chi_e^r) \\ &\quad + b(T - P^N T, \Delta^N; \chi_e^r). \end{aligned} \quad (39)$$

The bilinear form b is continuous and H -coercive; so from the definition of space X_e , we have:

$$b(\Delta^N, \Delta^N; \chi_e^r) \geq c_1 \|\Delta^N\|_1^2$$

$$b(T - P^N T, \Delta^N; \chi_e^r) \leq c_2 \|T - P^N T\|_1 \|\Delta^N\|_1$$

Using (15) et (16) and the Young's inequality:

$$|b(T, \Delta^N; \chi_e - \chi_e^r)| \leq \frac{\gamma^2}{4\epsilon} |\chi_e - \chi_e^r|_Q^2 \|T\|_1^2 + \epsilon \|\Delta^N\|_1^2,$$

we get:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\Delta^N\|_1^2 + c_2 \|\Delta^N\|_1^2 &\leq \frac{1}{2} \frac{d}{dt} \|\Delta^N\|_1^2 + b(\Delta^N, \Delta^N; \chi_e^r) \\ &\leq \frac{\gamma^2}{4\epsilon} |\chi_e - \chi_e^r|_Q^2 \|T\|_1^2 + \epsilon \|\Delta^N\|_1^2 \\ &\quad + \left(\dot{T} - P^N \dot{T}, \Delta^N \right) \\ &\quad + c_1 \|T - P^N T\|_1 \|\Delta^N\|_1 \end{aligned} \quad (40)$$

Note that the constants c_1 and c_2 (defining the upper and lower bounds imposed on χ_e with respect to the Q -norm) are uniform (independent of time t).

When: $N \rightarrow \infty$: $T - P^N T \rightarrow 0$, $n_p \rightarrow \infty$ ($r \rightarrow \infty$ and $d \rightarrow \infty$): $S^d \rightarrow S$ and $\chi_e^r \rightarrow \chi_e$ thus for n_p , $N \rightarrow \infty$ the right term in inequality (40) converges to 0 and:

$$0 \leq \frac{1}{2} \frac{d}{dt} \|\Delta^N\|_1^2 + (c_2 - \epsilon) \|\Delta^N\|_1^2 \leq 0 \quad \text{for } N, n_p \rightarrow \infty$$

and hence: $\Delta^N \rightarrow 0$ in $L^2(0, t_f; H)$ for $N, n_p \rightarrow \infty$ and $T^N(t; q^{n_p}) \rightarrow T(t; q)$ in $L^2(0, t_f; H)$. This concludes the proof.

APPENDIX III

Proof of Lemma 3.2 Let $\{q^r\}_{r \in \mathbb{N}^*}$ be a sequence solution of the optimization problem (12) and (13), and let $\{\hat{q}^k\}_{k \in \mathbb{N}^*}$ be an arbitrary sequence in Q such that $i^{n_p}(\hat{q}^k) = q$.

Since Q et Q^{n_p} are two compact spaces, we can always extract a subsequence of $\{q^r\}$ and $\{\hat{q}^k\}$ such that: $\lim_{r \rightarrow \infty} q^r = \bar{q}$ and $\lim_{r, k \rightarrow \infty} \hat{q}^k = q$.

Since $\{q^r\}$ is the solution of (12) and (13), $\{q^r\}$ is an optimal solution and thus:

$$J^N(q^r) \leq J^N(i^{n_p}(\hat{q}^k))$$

when: $N \rightarrow \infty$ and $r \rightarrow \infty$: $T^N(t; q^r) \rightarrow T(t; q)$ (from theorem 3.1) and:

$$J^N(q^r) \rightarrow J(\bar{q})$$

$$J^N(i^{n_p}(\hat{q}^k)) \rightarrow J(q),$$

Consequently:

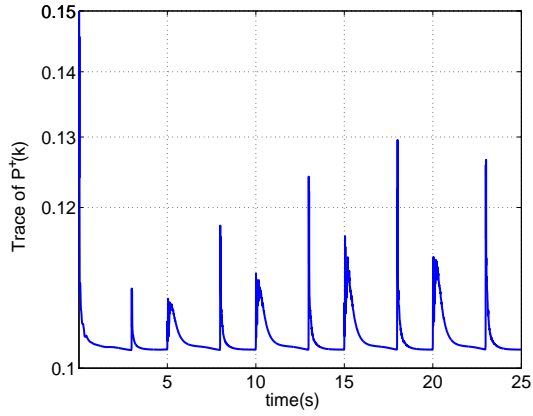
$$J(\bar{q}) \leq J(q)$$

We conclude that \bar{q} is the solution of (10) and (11).

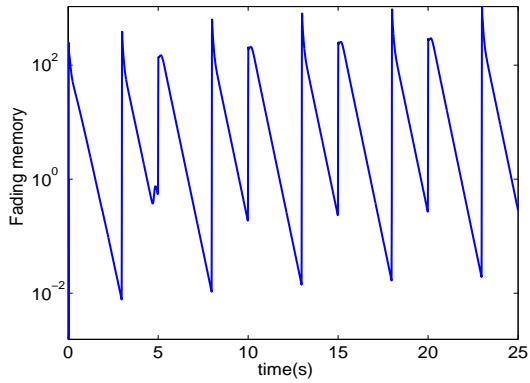
REFERENCES

- [1] G. W. Spakman, "Two-dimensional heat transport in tokamak plasma," Ph.D. dissertation, Technische Universiteit Eindhoven, 2011.
- [2] D. R. Mikkelsen, S. Scott, and W. Dorland, "Testing the ρ_* scaling of thermal transport models: Predicted and measured temperatures in the tokamak fusion test reactor dimensionless scaling experiments," *Physics of Plasmas*, vol. 4, no. 5, pp. 1362–1370, 1997.
- [3] G. Hoang, B. Saoutic, L. Guizou, V. Basiuk, A. Bécoulet, R. Budny, F. Clairet, L. Colas, P. Devynck, C. Gil, E. Joffrin, X. Litaudon, J.-L. Ségui, I. Voitsekhovitch, and X. Zou, "Magnetic shear effects on confinement and electron heat transport in Tore Supra discharges with electron heating," *Nuclear Fusion*, vol. 38, no. 1, pp. 117–132, 1998.

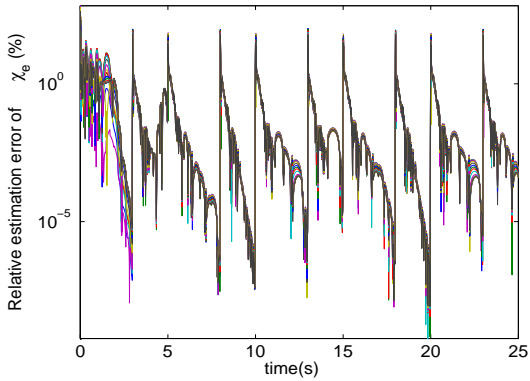
- [4] F. Ryter, Y. Camenen, J. C. DeBoo, F. Imbeaux, P. Mantica, G. Regnoli, C. Sozzi, U. Stroth, A. Upgrade, DIII-D, FTU, J.-E. contributors, T. S. TCV, and W.-A. Teams, "Electron heat transport studies," *Plasma Physics and Controlled Fusion*, vol. 48, pp. B453–B463, 2006.
- [5] F. Ryter, R. Dux, P. Mantica, and T. Tala, "Perturbative studies of transport phenomena in fusion devices," *Plasma Physics and Controlled Fusion*, vol. 52, no. 124043, 2010.
- [6] X. L. Zou, G. Giruzzi, J. F. Artaud, F. Bouquay, A. Cléménçon, C. Darbos, R. J. Dumont, C. Guivarch, M. Lennholm, R. Magne, and J. L. Ségui, "Electron heat transport and ECRH modulation experiments in Tore Supra tokamak," *Nuclear Fusion*, vol. 43, pp. 1411–1420, 2003.
- [7] V. D. Pustovitov, "Nonlocal effects in energy in an equilibrium plasma during its fast heating/cooling in tokamaks and stellarators," *Plasma Physics and Controlled Fusion*, vol. 54, no. 124036, 2012.
- [8] C. de Boor, *A practical guide to splines*, N. Y. Springer, Ed. Applied ; mathematics series, 1978, vol. 27.
- [9] J. D. Glover, "The linear estimation of completely unknown signals," *IEEE Transactions on Automatic Control*, pp. 766–767, 1969.
- [10] P. K. Kitanidis, "Unbiased minimum-variance linear state estimation," *Automatica*, vol. 23, no. 6, pp. 775–778, 1987.
- [11] M. Darouach and M. Zasadzinski, "Unbiased minimum-variance estimation for systems with unknown exogenous inputs," *Automatica*, vol. 33, no. 4, pp. 717–719, 1997.
- [12] M. Darouach, M. Zasadzinski, and M. Boutayeb, "Extension of minimum-variance estimation for systems with unknown inputs," *Automatica*, vol. 39, pp. 867–876, 2003.
- [13] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation in discrete-time systems," *Automatica*, vol. 43, pp. 111–116, 2007.
- [14] H. Fang, Y. Shi, and J. Yi3, "On stable simultaneous input and state estimation for discrete-time linear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, pp. 671–686, 2011.
- [15] R. W. F. Yang, "Observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 33, p. 677681, 1988.
- [16] M. Darouach, M. Zasadzinski, and S. Xu, "Full-order observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 39, p. 606609., 1994.
- [17] T. Floquet, C. Edwards, and S. K. Spurgeon, "On sliding mode observers for systems with unknown inputs," *International Journal of Adaptive Control and Signal Processing*, vol. 21, no. 8-9, pp. 638–656, 2007.
- [18] Q. P. Ha and H. Trinh, "State and input simultaneous estimation for a class of nonlinear systems," *Automatica*, vol. 40, no. 10, p. 17791785, 2004.
- [19] S. Pan, H. Su, H. Wang, J. Chu, and R. Lu, "Input and state estimation for linear systems: A least squares estimation approach," *Proceedings of the 7th Asian Control Conference*, pp. 378–383, August 27-29 2009.
- [20] F. Hinton and R. Hazeltine, "Theory of plasma transport in toroidal confinement systems," *Reviews of Modern Physics*, vol. 48, no. 2, pp. 239–308, 1976.
- [21] A. Taroni, M. Erba, E. Springmann, and F. Tibone, "Global and local energy confinement properties of simple transport coefficients of the bohm type," *Plasma Physics and Controlled Fusion*, vol. 36, pp. 1629–1646, 1994.
- [22] J. Hadamard, "Sur les problèmes aux dérivées partielles et leur signification physique," *Princeton University Bulletin*, vol. 13, pp. 49–52, 1902.
- [23] H. T. Banks, *A functional analysis framework for modelling, estimation and control in science and engineering*, C. Press, Ed. Taylor & Francis Group, 2012.
- [24] H. T. Banks, R. C. Smith, and Y. Wang, *Smart Materials Structures: Modeling, Estimation and Control*, M. Wiley, Ed., Paris/Chichester, 1996.
- [25] H. T. Banks and D. W. Iles, "On compactness of admissible parameter sets: convergence and stability in inverse problems," *Lecture notes in control and information sciences*, vol. 97, pp. 130–142, 1987.
- [26] C. Kravis and J. H. Seinfeld, "Identification of parameters in distributed parameter systems by regularization," *SIAM J. Control and Optimization*, vol. 23, pp. 217–241, 1985.
- [27] W. Yu and J. H. Seinfeld, "Identification of parabolic distributed parameter systems by regularization with differential operators," *Journal of mathematical analysis and applications*, vol. 132, pp. 365–387, 1988.
- [28] H. X. Li and C. Qi, "Modeling of distributed parameter systems for application- a synthesised review from time-space separation," *Journal of Process Control*, vol. 20, pp. 891–901, 2010.
- [29] C. De Boor, *A practical guide to splines*, Springer-Verlag, Ed. Applied Mathematical Sciences, 2001.
- [30] P. M. Prenter, *Splines and variational methods*, Pure and applied mathematics, Eds. Jhon Wiley & Sons, 1989.
- [31] K. Höllig, *Finite Element Methods with B-Splines*, F. in applied mathematics, Ed. Siam, 2003.
- [32] S. Mechhoud, E. Witrant, L. Dugard, and D. Moreau, "Combined distributed parameters and source estimation in tokamak plasma heat transport," *European Control Conference*, July 2013.
- [33] Y. Guan and M. Saif, "A novel approach to the design of unknown input observers," *IEEE Transactions on Automatic Control*, vol. 36, no. 5, pp. 632–635, 1991.
- [34] S. Gillijns and B. D. Moor, "Unbiased minimum-variance estimation for systems with unknown inputs," *Automatica*, vol. 43, pp. 934–937, 2007.
- [35] H. Palanthandalam-Madapusi and D. S. Bernstein, "Unbiased minimum-variance filtering for input reconstruction," *Proceedings of the 2007 American Control Conference*, pp. 11–13, July 2007.
- [36] A. H. Jazwinski, *Stochastic process and filtering theory*, A. Press, Ed., New York, 1970.
- [37] S. Pan, "System identification and damage detection of structures with unknown excitations," Ph.D. dissertation, Dept. Civil and Environmental Engineering, University of California, Irvine, USA, 2006.
- [38] L. Ljung, "Asymptotic behaviour of extended Kalman filter as a parameter estimator for linear systems," *IEEE Trans. on Auto. Control*, vol. AC-24, no. 1, pp. 36–50, 1979.
- [39] D. Simon, *Optimal state estimation: Kalman, H_∞ and nonlinear approaches*. Jhon Wiley & Sons, 2006.
- [40] L. Ljung and T. Söderström, *Theory and practice of recursive identification*. The MIT Press, Cambridge Massachusetts, London England, 1983.
- [41] Y. Zhu and P. Pagillia, "Adaptive estimation of time varying parameters in linearly parametrized systems," *Journal of Dynamic Systems Measurements and Control*, vol. 128, pp. 691–695, 2005.
- [42] I. D. Landau, *Identification des systèmes*, ser. Système automatisé, Information Commande Communication. Paris: Hermès, 1998.
- [43] J. Yang and S. Lin, "Identification of parametric variations of structures based on least squares estimation and adaptive tracking technique," *Journal of engineering mechanics*, 2005.
- [44] M. Erba, T. Aniel, V. Basiuc, A. Becoulet, and X. Litaudon, "Validation of a new mixed bohm/gyro-bohm model for electron and heat transport against the iter, tore supra and start database discharges," *Nuclear Fusion*, vol. 38, no. 7, pp. 1013–1028, 1998.
- [45] A. Cléménçon, C. Guivarch, S. Eury, X. Zou, and G. Giruzzi, "Analytical solution of the diffusion equation in a cylindrical medium with step-like diffusivity," *Physics of Plasmas*, vol. 11, no. 11, pp. 4998–5009, 2004.
- [46] V. Udintsev, M. Goniche, J. L. Segui, G. Giruzzi, D. Molina, F. Turco, G. T. A. Huysmans, P. Maget, T. S. team, and A. Krämer-Flecken, "First results of correlation electron cyclotron emission on tore supra," *Fusion Science and Technology*, vol. 50, pp. 508–520, 2006.
- [47] B. Baiocchi, "Understanding and predicting ion heat transport in tokamaks," Ph.D. dissertation, Università Degli Studi di Milano, Facoltà di Scienze Matematiche, Fisiche e Naturali, 2010.
- [48] I. K. et al., "Electron cyclotron emission spectroscopy on the tokamak á configuration variable," *Review of Scientific Measurements*, vol. 76, no. 9, 2005.
- [49] S. Mechhoud, E. Witrant, L. Dugard, and D. Moreau, "Adaptive distributed parameter and input estimation in plasma tokamak heat transport," *1st IFAC Workshop on systems governed by Partial Differential Equations, Paris, France*, 2013.
- [50] —, "Adaptive space-time distributed parameter and input estimation in heat transport with unknown bounds," *IEEE International Conference on Systems and Control, Algiers, Algeria*, 2013.
- [51] L. C. Evans, *Partial differential equations: Graduate studies in Mathematics*, 2nd ed. American Mathematical Society, 2010, vol. 19.
- [52] H. T. Banks and M. Pedersen, "Well-posedness of inverse problems for systems with time dependent parameters," *Arabian Journal for Science and Engineering. Section B: Engineering*, vol. 34, pp. 39–58., 2009.



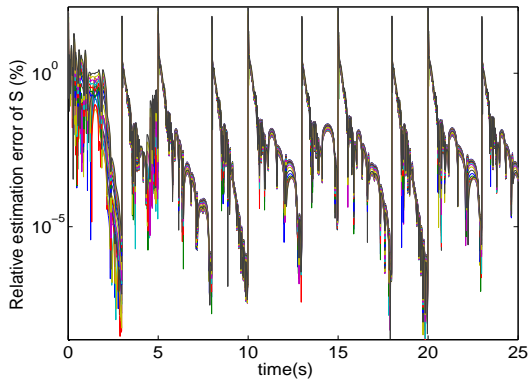
(a) Trace of $P^+(k)$



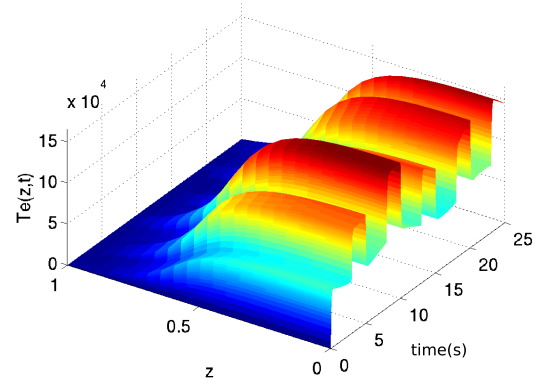
(b) Time evolution of the fading memory $\lambda(k)$



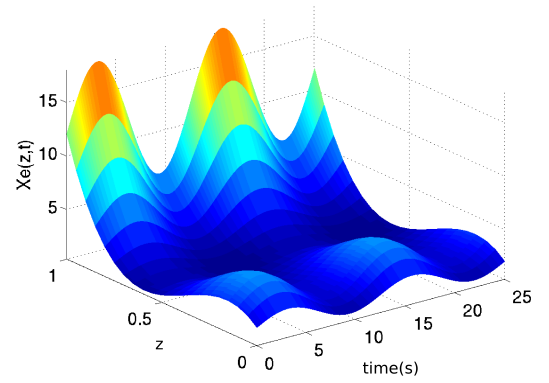
(c) Relative estimation error of $\chi_e(z, t)$



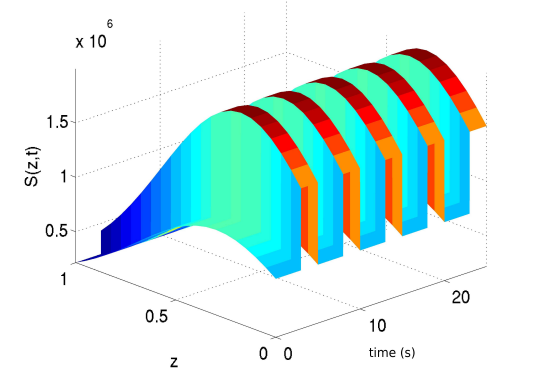
(d) Relative estimation error of $S(z, t)$



(a) Spatio-temporal profile of $T(z, t)$



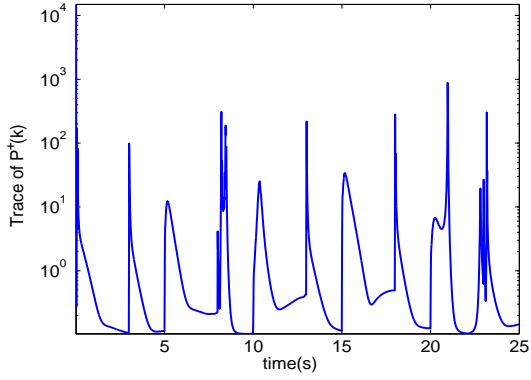
(b) $\chi_e(z, t)$ profile



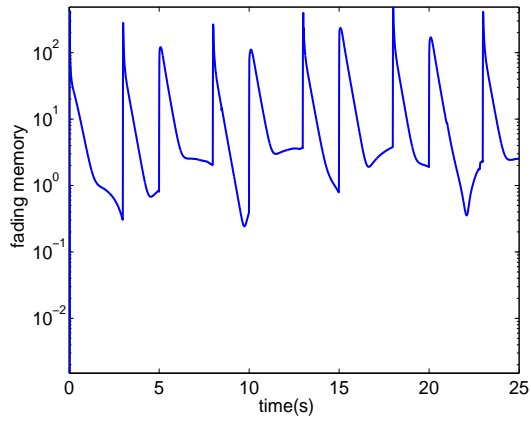
(c) $S(z, t)$ profile

Fig. 3: Example 2: Profiles of $T(z, t)$, $S(z, t)$ and $\chi_e(z, t)$

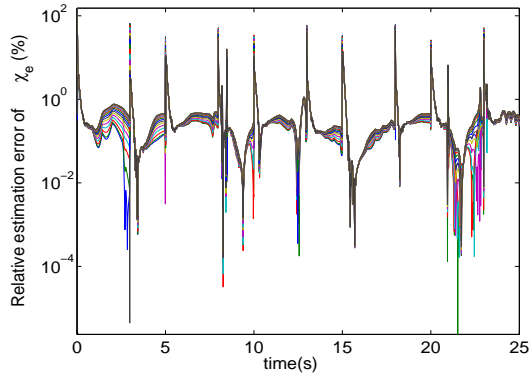
Fig. 2: Simulation results using the adaptive EKF-UI-WDF for the reconstruction of S and χ_e in Example 1



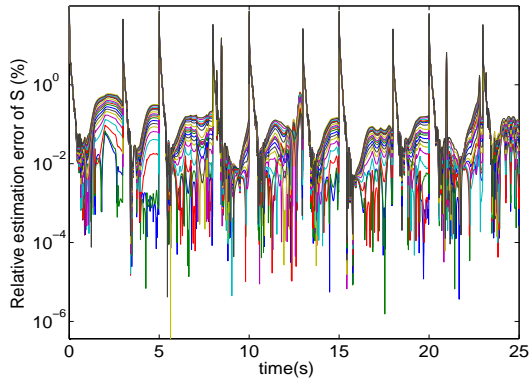
(a) Trace of $P(k|k)$



(b) Time evolution of the fading memory $\lambda(k)$



(c) Relative estimation error of $\chi_e(z, t)$



(d) Relative estimation error of $S(z, t)$

Fig. 4: Simulation results using the adaptive EKF-UI-WDF for the reconstruction of S and χ_e in Example 2

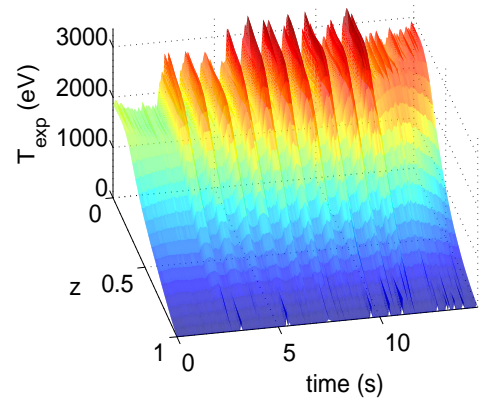
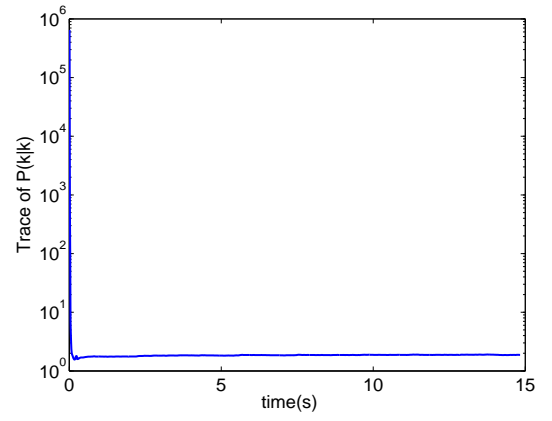
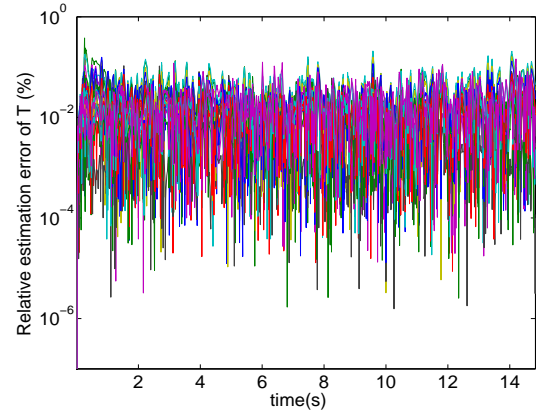


Fig. 5: Profile of T_{exp}

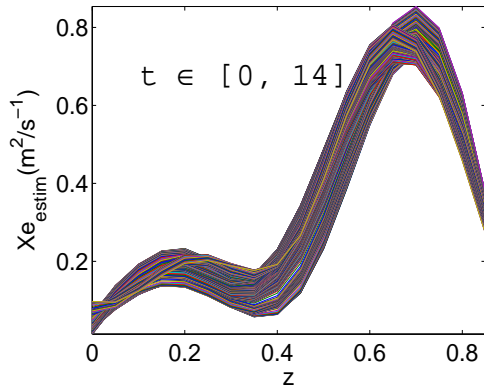


(a) Trace of $P^+(k)$

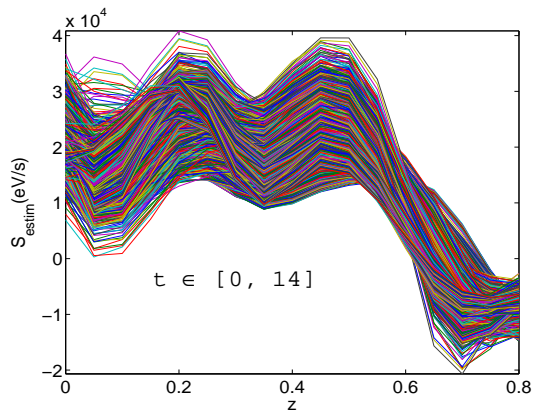


(b) Relative estimation error of T_{exp}

Fig. 6: Performances of the EKF-UI-WDF for the estimation of T_{exp} for the experimental data of Tore Supra shot $TS - 33632$.



(a) Estimated profile of $\chi_e(z, t)$



(b) Estimated profile of $S(z, t)$

Fig. 7: Estimated experimental profiles of S and χ_e using the EKF-UI-WDF

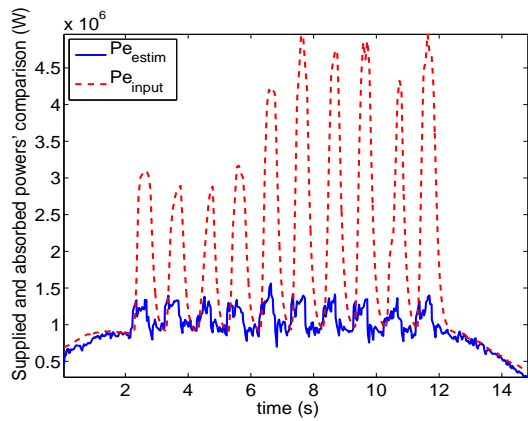


Fig. 8: Comparison between the absorbed Pe_{estim} and the total Pe_{input} powers