Parameter estimation for a nonlinear control-oriented tokamak profile evolution model

P Geelen¹, F Felici¹, A Merle² and O Sauter²

¹ Department of Mechanical Engineering, Eindhoven University of Technology, Control Systems Technology Group, PO Box 513, 5600MB Eindhoven, The Netherlands
² Centre de Recherches en Physique des Plasmas, École Polytechnique Fédérale de Lausanne (EPFL), Association EURATOM-Suisse, 1015 Lausanne, Switzerland

E-mail: f.felici@tue.nl

Received 28 May 2015, revised 15 August 2015
Accepted for publication 18 August 2015
Published 16 October 2015

Abstract

A control-oriented tokamak profile evolution model is crucial for the development and testing of control schemes for a fusion plasma. The RAPTOR (RApid Plasma Transport simulatOR) code was developed with this aim in mind (Felici 2011 Nucl. Fusion 51 083052). The performance of the control system strongly depends on the quality of the control-oriented model predictions. In RAPTOR a semi-empirical transport model is used, instead of a first-principles physics model, to describe the electron heat diffusivity \( \chi_e \) in view of computational speed. The structure of the empirical model is given by the physics knowledge, and only some unknown physics of \( \chi_e \), which is more complicated and less well understood, is captured in its model parameters. Additionally, time-averaged sawtooth behavior is modeled by an ad hoc addition to the neoclassical conductivity \( \sigma_n \) and electron heat diffusivity. As a result, RAPTOR contains parameters that need to be estimated for a tokamak plasma to make reliable predictions. In this paper a generic parameter estimation method, based on the nonlinear least-squares theory, was developed to estimate these model parameters. For the TCV tokamak, interpretative transport simulations that used measured \( T_e \) profiles were performed and it was shown that the developed method is capable of finding the model parameters such that RAPTOR’s predictions agree within ten percent with the simulated \( q \) profile and twenty percent with the measured \( T_e \) profile. The newly developed model-parameter estimation procedure now results in a better description of a fusion plasma and allows for a less ad hoc and more automated method to implement RAPTOR on a variety of tokamaks.

Keywords: tokamaks, transport modeling, parameter estimation

(Some figures may appear in colour only in the online journal)
capable RAPTOR (RApid Plasma Transport simulatOR) code [14, 5]. It solves the coupled poloidal flux diffusion equation and the electron temperature transport under the assumption of a fixed magnetic equilibrium and assuming all thermal transport is diffusive. The energy confinement time, which is in the order of 1 s for TCV and 3.6 s for ITER, determines the required speed for real-time execution. In order to meet this requirement the RAPTOR code uses, at present, an empirical, parameterized model for the heat conductivity. While these empirical models are constructed to satisfy general dependencies of the transport coefficient on local plasma quantities such as the safety factor and magnetic shear, they have some free parameters which must be tuned based either on experiments or more complete physical models. This is presently done by manual tuning of the appropriate coefficients to match experimental or simulated data. While this is feasible in practice due to the short time required for a RAPTOR simulation (typically 10 s for a full discharge), it would be appropriate to have an automated method of tuning these parameters. This should yield a more optimal parameter choice, treatment of larger experimental databases, and an overall faster tuning procedure. In addition, the resultant parameters can yield new information for comparison with theory.

Extensive work has also been done on the use of linear models for plasma profile control [2, 15, 16]. In these models, the plasma profile dynamics around an operating point is represented by a set of linear ordinary differential equations which include the different timescales of resistive and thermal transport. The coefficients of this linear model are automatically identified based on experiments or simulated data around that operating point. Since the equations are linear, the parameter estimation problem is also linear, which is easily tractable numerically. However, the obtained linear model is valid only in a region close to the chosen operating point and can not be guaranteed to work in other regimes.

In this paper, we present an automated parameter estimation method for the nonlinear transport simulation equations used in RAPTOR based on data. We thereby extend the application of system identification from the linear models, as treated in [15] to the nonlinear model used in RAPTOR.

The data used to tune the model can be obtained either from experiments or from more complete and time-intensive simulations using a more comprehensive interpretative transport calculation. The latter can be either predictive simulations (full simulation of plasma evolution) or interpretative simulations (using some measured data as input for the simulation and solving only for the unknown quantities). For simplicity, we refer to this data as ‘experimental’ data, irrespective of whether it was obtained from a true tokamak experiment or from a numerical (simulation) experiment. In this paper we will use interpretative simulations of TCV tokamak experiments to tune the transport model.

The problem of finding the unknown model parameters is formulated as a nonlinear least-squares optimization problem, with a cost function reflecting the discrepancy between simulated and known (i.e. measured) profile evolutions. Since the model is nonlinear, also the parameter optimization problem is inherently nonlinear, which makes its solution prone to local minima and other numerical issues.

An important step that greatly facilitates solving the nonlinear optimization problem is to construct an analytical expression for the gradient of this cost function, giving the steepest descent direction of the cost function in parameter space. Once a local minimum has been found, local properties of the solution can be analyzed to give information about confidence intervals or non-identifiability of some combinations of parameters.

One may legitimately wonder why one does not resort to more complete physics-based transport models, such as TGFL [17], QuaLiKiz [18] or MMM8.1 [19] to determine the required parameters. This is due to several reasons. Firstly, despite significant progress, such models are not yet able to fully predict tokamak transport in all regimes and locations over a full discharge timescale. Moreover, they are not yet computationally fast enough to be evaluated practically for a very large region of transport parameters, though very recent work showed how this can be circumvented using a neural-network fit of a large database of QuaLiKiz results [20]. Finally, by ensuring that the simulations match (measured) experimental profile evolution, any systematic mismatch between the model and reality can be partly absorbed into the estimated parameters as well.

The rest of this paper is structured as follows: in section 2 an introduction to RAPTOR is given, followed by the section which describes the parameters of the empirical models $\chi_c$ and $\sigma_c$ in RAPTOR that need to be identified. In section 4 the nonlinear parameter estimation method for RAPTOR is presented. To demonstrate the developed method, known model parameters from simulated data are first recovered in section 5.1. As the main result of this paper, the parameters of $\chi_c$ and $\sigma_c$ were identified for a set of discharges from the TCV tokamak in section 5.2. The conclusions are presented in section 7.

2. Overview of the RAPTOR code

RAPTOR is a control-oriented, physics-based code for simulating the 1D plasma transport equations, that has been designed for real-time control applications [5, 9]. Using a prescribed evolution of the equilibrium flux surface layout, density profile evolution, and actuator inputs ($I_a$ and power to auxiliary heating and current drive systems), it simulates the time evolution of the poloidal flux and temperature profiles. In the present work, we assume a fixed flux surface geometry, a prescribed density profile and a fixed ratio of ion temperature to electron temperature, though RAPTOR presently includes the possibility to vary these in time.

The physics model solved is that of 1D (flux surface averaged) diffusion of poloidal flux ($\psi$) and electron temperature ($T_e$) profiles as a function of the radial variable $\rho$, the normalized toroidal flux coordinate defined as $\rho = \sqrt{\psi}/\psi_0$. The equation reads, following [10],

$$\frac{\partial \psi}{\partial t} = \frac{F^2}{16\pi^2\mu_0\Phi_{fe}^2} \frac{\partial}{\partial \rho} \left[ g_2 g_3 \frac{\partial \psi}{\partial \rho} \right] - \frac{B_0}{2\Phi_{fe}} V(j_{cd} + j_{bs}) \tag{1}$$

where $F = RB_{e\rho}$, $g_2 = \left(\frac{\partial \psi}{\partial \rho}\right)^2$, $g_3 = \left(\frac{1}{\rho}\right)$ and $V = \frac{\partial \psi}{\partial \rho}$. $\Phi_{fe}$ are functions of $\rho$ that depend on the 2D MHD equilibrium.
[21]. Together with $\Phi_b$, the (scalar) toroidal magnetic flux at the plasma boundary, these quantities are computed for one representative equilibrium and kept constant thereafter. $\sigma_f$ and $j_{bn}$ are, respectively, the neoclassical electrical conductivity and bootstrap current density, using the Sauter–Angioni formulas [22, 23]. The driven current density $j_{\text{act}}$ is modeled using a gaussian deposition shape for ECH/ECCD, which is the only heating/current drive actuator considered in this paper.

Assuming $V'$ does not vary in time [14], the equation for the spatial evolution of the electron temperature $T_e$ in time is

$$V' \frac{\partial}{\partial t} (n_e T_e) = \frac{\partial}{\partial \rho} \left( \frac{q_0}{V' n_e \chi_e} \frac{\partial T_e}{\partial \rho} \right) + V' P_e. \tag{2}$$

Here $n_e$ is the electron density, $q_0 = \left( \frac{\partial V}{\partial \rho} \right)^2 \left( \nabla \psi \right)^2$ is another geometric profile quantity and $P_e$ is the total power to the electrons. This includes ohmic and auxiliary power sources and losses via equipartition with ions, line radiation and bremsstrahlung. For the electron heat diffusivity, $\chi_e$, a nonlinear analytical function of the local safety factor $q$ and the magnetic shear $s = \frac{\partial q_b}{\partial \rho}$ is used, which will be described in detail in section 3.1.

Density profiles are prescribed (but can be time-varying) and the ion temperature is assumed to scale proportionally to the electron temperature, with proportionality factor $\Pi$ that does not vary in time [14], the equation for $T_i$ becomes discrete profiles

$$T_i(\rho, t) = C_i(\rho) \chi_k,$$ \hspace{1cm} \(7\)

$$\iota(\rho, t) = C(\rho) \chi_k,$$ \hspace{1cm} \(8\)

$$U_{\text{pl}}(\rho, t) = C_{\text{pl}} \chi_k.$$ \hspace{1cm} \(9\)

For further details, see [5].

3. Model parameters to be estimated in RAPTOR

As discussed in the previous section, RAPTOR uses an empirical model for the electron heat diffusivity $\chi_e$. This model captures some main physical dependencies of tokamak thermal transport, but includes unknown parameters that have to be identified for each tokamak. This model was extended, with respect to the original version in [14, 5], by the inclusion of the time-averaged effect of the sawteeth and an increase of diffusivity at higher temperatures, introduced to mimic thermal confinement degradation at higher input power (as appears in confinement scaling laws). Furthermore, an ad hoc addition of the time-averaged effect of the sawtooth behavior to the model for the electrical conductivity, $\sigma_f$ was made. Both models contain parameters that cannot be easily computed from physical laws and therefore need to be identified. The structure of the assumed empirical transport model is described here.

3.1. The electron heat diffusivity model and inclusion of sawtooth effects

The expression for $\chi_e$ in RAPTOR includes the anomalous diffusion and the dependence of the confinement on the $q$ profile [5]. In this paper the electron heat diffusivity model was extended to include the effect of sawtooth crashes and confinement deterioration with increasing temperature and temperature gradient. The expression for the local heat diffusivity (to be used in equation (2)) reads:

$$\chi_e = \chi_{\text{neo}} + c_{\text{ano}} G(q) + c_{\text{ano}} q F(s) (T_e [\text{keV}])^2 s.$$ \hspace{1cm} \(10\)

In this equation $\chi_{\text{neo}}$ is intended to represent a minimum level of transport that is due to neoclassical effects. This is chosen here as a small constant, though it could be calculated based on neoclassical transport physics [21]. The much larger anomalous diffusion is captured by $c_{\text{ano}}$ and the presence of $q$ in the anomalous diffusion term accounts for increased confinement observed at higher plasma currents $I_p$. The term $F(s)$ is a shear-dependent function to include the effect of improved confinement (ic) at low and negative magnetic shear $s$. This is used to model the improved confinement observed in some
advanced scenarios. The expression for the function \( F(s) \), plotted in figure 1, is

\[
F(s) = \frac{d_w}{1 + \exp[w_d(d_w - s)]} + (1 - a_w). \tag{11}
\]

The term \( a_w \) accounts for the amount of reduction of transport, \( d_w \) is the level of the shear at which the transition takes place. The sharpness of the transition is governed by the term \( w_d \).

A scaling of the heat transport at increasing temperature is modeled by the \( c_T \) exponent, where \( T_e(\rho) \) is the electron temperature at the center of the plasma. This term is included to capture the effect of reduced thermal confinement at higher input powers. The identified coefficient \( c_T \) can be related to the power degradation exponent in thermal confinement scaling laws\(^3\). Other choices of the \( \chi_e \) dependency are possible.

Sawtooth crashes occur when the safety factor \( q \) is below unity and cause periodic expulsion of plasma current and energy from the core region. Considering the time-averaged effect, this amounts to a decreased heat confinement in this region. This effect is modeled by the second term in the \( \chi_e \) equation. The function \( G(q) \), similar to figure 1, is used to model the effect as a function of \( q \) using a sigmoidal function centered at \( q = 0.95 \) with smoothness of the transition governed by the parameter \( w_{\chi e} \).

The amount of conductivity increase is governed by the parameter \( c_{\chi e} \).

\(^3\)The parameter \( c_T \) in (10) can be related to the power degradation exponent in thermal confinement scaling laws as follows. The (electron) heat flux in stationary conditions can written as \( Q_e = n_e \chi_e \frac{\partial T_e}{\partial \rho} = P \) where \( P \) is the total (volume-integrated) input power. The scaling law can then be written in the form \( \tau_e \sim P^{\nu \gamma} \). Assuming constant density this gives \( T_e \sim P^{1+\gamma} \). Filling this into the power balance, and introducing from (10) \( \chi_e \sim \frac{T_e^{\gamma}}{\nu} \) yields \( \sim n_e \chi_{e} T_e^{\gamma+1} \sim (P^{1+\gamma} n_e)^{\gamma+1} \). Therefore \( (1 + \alpha_p)(1 + c_T) = 1 \) or \( c_T = \frac{1}{1 + \alpha_p} - 1 \). This allows one to directly compare the identified value of \( c_T \) to the scaling law exponent. For example, typical values of \( \alpha_p = [-0.5, -0.7] \) correspond to \( c_T = [1, 2.3] \).
sawtoothing plasmas. Note that we included W(q) to obtain realistic q0 values and avoid too peaked current density profiles in sawtoothing plasmas.

W(q) for typical parameters is plotted in figure 2.

4. Nonlinear parameter estimation method

4.1. Theory of the nonlinear least-squares method

The nonlinear least-squares (NLS) method [24] is an estimation technique to determine the parameter set p for a nonlinear model f(t, p) such that it best describes the given measurements yk at time points tk of the output of the system. The general approach to the solution of the parameter estimation problem is to minimize a cost function, which is a function of the squares of the model residuals (errors) at the time step k: r_k(p) = y_k - f(t_k, p). The cost function J(p) is given as

\[ J(p) = \sum_{k=1}^{n_k} (y_k - f(t_k, p))^2, \]

where \( n_k \) is the number of data points. The estimate of the parameter set \( p \) is the minimizer of the cost function

\[ p^{\text{NLS}} = \min_p J(p). \]

Due to the nature of the nonlinear model \( f(t, p) \) only a local minimum can be guaranteed, not the global.

4.1.1. Application. In this paper the model parameters \( p \) are determined by minimizing the difference between RAPTOR’s predicted radial \( T_{e,k}^{\text{pred}}(\rho) \) and \( T_{e,k}^{\exp}(\rho) \) profiles and the experimental profiles \( T_{e,k}^{\exp}(\rho) \) for all time points k. Also the loop voltage \( U_{pl}(\rho) \) was included in the cost function, but using only the profile at the final time step \( n_k \) as this profile does not vary much in time.

Before formulating the cost function, we introduce a parameter scaling \( \Gamma = \Gamma^{-1}p \), with \( \Gamma = \text{diag}\{p_1, \ldots, p_{n_k}\} \) containing the initial guess of the model parameters. This ensures that the parameters are of similar order of magnitude which aids convergence of the nonlinear optimization [24].

The cost function to be minimized is now written as:

\[ J(p) = \nu_T \sum_{k=1}^{n_k} \| T_{e,k}^{\text{pred}}(\rho, p) - T_{e,k}^{\exp}(\rho) \|^2 + \nu_U \sum_{k=1}^{n_k} \| U_{pl,k}^{\text{pred}}(\rho, p) - U_{pl,k}^{\exp}(\rho) \|^2, \]

which is written more compactly as

\[ J(\varphi) = \nu_T \| \varphi_T \|_2 + \nu_U \| \varphi_U \|_2, \]

where the norm \( \| \cdot \|_2 \) of the radial profiles are approximated by the sum-of-squares over several points \( \rho_i \) chosen on the radial \( \rho \) grid, for example:

\[ \| \varphi_T \|_2 = \sum_{i=1}^{n_{\rho}} (T^{\text{pred}}(\rho_i, \varphi_T) - T^{\exp}(\rho_i))^2, \]

Weighting scalars \( \nu_T, \nu_U, \nu_{Upl} \) were chosen such that the contribution of \( U_{pl}^2 [\text{V}^2] (\mathcal{O}(1)), T_e^2 [\text{eV}^2] (\mathcal{O}(1 \times 10^6)) \) and \( \psi^2 [\text{V}] \) (\( \mathcal{O}(1) \)) are of the same order.

By minimizing the cost function in equation (19) the model parameters \( p \) are found such that RAPTOR describes the temporal evolution of the radial \( U_{pl}, T_e \) and \( \psi \) profiles more accurately than when using the initial guess of the model parameters. The cost function can also be extended to estimate the model parameters for \( n_d \) datasets at once. This feature can be used to identify one set of model parameters which accurately describe the data for various plasma shots done in different conditions. In this case the cost function definition becomes:

\[ J(\varphi) = \sum_{d=1}^{n_d} \nu_T \| \varphi_{T,d} \|_2 + \nu_U \| \varphi_{U,d} \|_2 + \nu_{Upl} \| \varphi_{Upl,d} \|_2. \]

4.2. Outline of the parameter estimation algorithm

The parameter estimation algorithm has to identify a (sub)set of eight \( \chi_e \) model parameters that appear in equation (10):

\[ \lambda_{\text{neo}}, c_{\text{saw}}, w_{\text{saw}}, c_{\text{ano}}, a_{\text{ic}}, w_{\text{ic}}, d_{\text{ek}} \text{ and } c_{\text{Te}}, \]

and/or two \( \sigma \) model parameters in equation (15)

\[ c_{\text{saw}} \text{ and } w_{\text{saw}}. \]

Figure 3 shows the flow diagram of the developed parameter estimation algorithm. To start the parameter estimation routine, an initial guess of the model parameters \( p \) must be specified. RAPTOR uses this initial guess to solve the nonlinear poloidal flux diffusion equation and nonlinear electron energy transport equation to obtain the state vector \( x(t) \), defined by equation (4). This state vector is used to compute the profiles.
The cost function $J$ was calculated for a nominal value of the model parameter $p$ and in case a small disturbance $\delta p$ ($O(10^{-6})$) was added to this value $p + \delta p$. In the table below, the relative difference between $\delta J_{\text{analyt}}$ and $\delta J_{\text{num}}$ [\%] are shown for the 10 model parameters.

| Parameter | $\chi_{\text{neo}}$ | $c_{\text{neo}}$ | $d_{\text{ic}}$ | $w_{\text{ic}}$ | $d_{\text{ic}}$ |
|-----------|---------------------|------------------|-----------------|----------------|----------------|
| Value     | $2.24 \times 10^{-5}$ | $7.97 \times 10^{-4}$ | $3.97 \times 10^{-4}$ | $2.59 \times 10^{-3}$ | $2.36 \times 10^{-2}$ |
| $c_{\text{raw}}$ | $6.86 \times 10^{-4}$ | $2.76 \times 10^{-5}$ | $1.40 \times 10^{-4}$ | $0.16$ | $9.38 \times 10^{-4}$ |

Note: The difference in the value of the cost function yields the finite difference number $\delta J_{\text{num}}$. The value $\delta J_{\text{analyt}}$ was obtained using the analytical expression for the cost function gradient described in this section (equation (22)): $\delta J_{\text{analyt}} = (\partial J/\partial p)_{\text{analyt}} \delta p$. The gradient $(\partial J/\partial p)_{\text{analyt}}$ was computed using the nominal model parameter $p$. The relative difference between the two various approaches to calculate $\delta J$ are small (maximum 0.16\% for $\sigma_{\text{c saw}}$).

Figure 4. Time traces of key quantities for the simulated experiment. The plasma current $I_p$, electron cyclotron power deposition, electron density $n_{eo}$ and the effective nuclear charge were imposed. Note the rapid evolution of the electron temperature, caused by fast changes in the applied EC power and the short energy confinement times in TCV L-mode plasmas (few ms).

$T_e(p)$, $\iota(p)$ and $U_{el}(p)$ at all times as defined in equations (7)–(9). The model predictions are compared to experimental data by computing the cost function $J$. The cost function is minimized with an optimization routine. The sequential quadratic programming (SQP) algorithm [25], which is readily available in MATLAB®, was used for this purpose. The SQP algorithm finds the (local) minimum of the cost function by changing the model parameters $p$ in an iterative way. As long as the cost function is not in a local minimum, the algorithm attempts to compute a new set of parameters $p$ for which the cost function is lower. The SQP algorithm can also deal with constraints, which is useful to impose realistic bounds on the parameters to be estimated. If the algorithm behaves correctly, the parameters obtained at the (local) minimum yield $\iota$, $U_{el}$ and $T_e$ profiles in RAPTOR that match the experimental data better than the initial guess of the model parameters. Due to the nonlinearity of the parameter estimation problem, no strict guarantees can be given on whether a local or global minimum has been found.

One can merely rely on practical techniques, such as shooting methods, to reduce the chance of finding local minima.

### 4.3. Computation of the cost function gradient

The performance of the SQP algorithm critically depends on the derivative of the cost function to the model parameters $dJ/\partial p$. This cost function gradient gives information on the direction in parameter space in which the cost function decreases. In RAPTOR this gradient $dJ/\partial p$ is computed analytically. This has two major advantages, the first one being that the SQP algorithm converges faster to the (local) minimum because no finite difference approximation is needed to approximate the derivative of the cost function to the model parameters. The second advantage is that the analytical gradient is exact in contrast with the numerical approximation, which enhances the performance of the SQP algorithm. The analytical expression of $dJ/\partial p$ is computed as follows.
The term \(\frac{\partial}{\partial t} J(x)\) can be obtained from direct differentiation of the cost function. The term \(\frac{\partial}{\partial x} p\) is obtained from the forward sensitivity equation, which is discussed next.

Recall that to obtain the state vector \(x\), RAPTOR solves equation (6) (which contains the flux diffusion equation and electron energy transport) at discrete-time points \(k\):

\[
f_k = f(x_{k+1}, x_k, u_k) = 0 \quad \forall k.
\]

Differentiating this equation with respect to a vector of model parameters \(p\) results in the forward sensitivity equation

\[
0 = \frac{df_k}{dp} = \frac{\partial f_k}{\partial x_{k+1}} \frac{\partial x_{k+1}}{\partial p} + \frac{\partial f_k}{\partial x_k} \frac{\partial x_k}{\partial p} + \frac{\partial f_k}{\partial u_k} \frac{\partial u_k}{\partial p} + \frac{\partial f_k}{\partial p}.
\]

In this equation

- The first two terms on the right hand side contain the Jacobians \(\frac{\partial f_k}{\partial x_{k+1}}\) and \(\frac{\partial f_k}{\partial x_k}\). These are already computed at each time step in RAPTOR since they are necessary for solving the Newton problem used to obtain the next time step of the simulation;
- The third term is nonzero only if the inputs are prescribed in terms of parameters, as was the case in [5], but not in this work;
- The last term is nonzero for model parameters that appear in the equations (10) and (15).
Now every term is known, the ODE sensitivity equation (23) is recursively solved for each parameter in the parameter vector \( p \) starting from the initial condition \( \frac{\partial \chi_0}{\partial p} \) (which is typically 0, since the initial condition does not depend on any parameters). The recursive solution of this finite difference equation yields \( \frac{\partial \chi_k}{\partial p} \), for \( k \in k_{\text{n1}}, \ldots, k_{\text{n}} \). This set of matrices is referred to as the state sensitivities with respect to the model parameters \( p \). Once the state sensitivities are known, the cost function gradient with respect to the model parameters can be computed using (22).

We now verify that the cost function gradient is correctly computed. As shown in table 1 the analytical calculation of the cost function gradient and its finite difference estimate are in good agreement (maximum relative difference between the two approaches is smaller than 0.16%), which confirms the correct implementation.

5. Results

In this section we present the application of the parameter estimation algorithm to the TCV tokamak. We first present a numerical experiment to validate the parameter estimation algorithm by finding known parameters from simulated data (section 5.1). Then we will show the application of the procedure to Astra interpretative transport simulations of TCV tokamak plasmas (section 5.2).

5.1. Demonstration: finding known model parameters from simulated data

5.1.1. Noise-free case. A numerical experiment was performed to demonstrate that the developed estimation algorithm is able to find the model parameters in \( \lambda_e \) and \( \sigma_0 \). In this experiment the experimental data contains a set of \( T_e \), \( \iota \) and \( U_{pl} \) profiles generated by a forward simulation of RAPTOR with known parameters and without any artificial noise. A plasma with TCV parameters exhibiting reversed shear and sawtooth behavior was simulated by applying fast changes to the plasma current and EC power. A time step of 1 ms was used in a 0.4 s simulation and 11 \( \rho \) gridpoints were equally distributed from 0 to 1. In figure 4 the time traces of key quantities of the plasma during this numerical experiment are shown and in figure 5 various plasma profiles are depicted at three distinct time points.

The true model parameters \( p \) of \( \lambda_e \) and \( \sigma_0 \) that were used to generate the synthetic data are given in table 2. The developed estimation method was used to recover the model parameters.
parameters which were set to generate the data. To start the parameter estimation procedure an arbitrary initial guess of the model parameters was used (table 2). In figure 6 the cost function is plotted versus the iteration steps. The cost function reduces as the number of iterations increase, showing that the estimated model parameters at each increasing iteration step allow for a better description of the synthetic data. At the final iteration step, the relative error between the estimated model parameters (pest) and true model parameters (pref) is

\[ \| \text{pest} - \text{pref} \|_2/\| \text{pref} \|_2 = 3.5004 \cdot 10^{-10}, \]

meaning that the estimation routine almost perfectly recovers the model parameters which were used to generate the synthetic data.

In figure 7(a) the relative error of \( e \) is shown versus time for a variety of different iteration steps. In figure 7(b), the relative error of \( T_e \) versus time and a variety of different iteration steps is plotted, and in figure 8 the relative error of the loop voltage \( U_{pl} \) at the final time point \((0.4 \text{ s})\) is shown versus iteration steps. The relative errors reduce with increasing number of iterations as the estimated model parameters converge to the true model parameters which were used to generate the data within machine precision.

This numerical experiment demonstrates that the developed parameter estimation algorithm is able to recover the original model parameters from simulated data.

5.1.2. Noisy case. In this section the reliability and precision of the estimated model parameters in \( n_e \) and \( \sigma_t \) are investigated.

A parameter is ill-determined if its estimated value is affected with the difference that the synthetic data of the

\[ T_e, t \text{ and } U_{pl} \text{ profiles are always contaminated with noise.} \]

\( \hat{\rho} \) denotes the estimated parameter vector. To investigate the reliability and precision of the estimated model parameters one needs to compute the covariance matrix for the parameters. According to the asymptotic theory, this matrix can be approximated as [26]

\[ C = \text{cov}(\hat{\rho}) \approx \sigma^2(H(\hat{\rho}))^{-1}, \]

in which \( H(\hat{\rho}) = \partial^2 J/\partial \rho^2 \) denotes the Hessian matrix of the cost function evaluated at \( \hat{\rho} \), the local minimum of the cost function. The standard deviations \( \sigma \) of the noise can be estimated from the residuals, assuming that the errors are independent and normally distributed with expected value 0 [23]

\[ \sigma^2 = 1/n_k - n_p \]

with \( n_p \) the number of model parameters and \( n_k \) the number of measurements.

The diagonal elements of the covariance matrix contain the variance of the parameters \( \hat{\rho} \)

\[ \text{var}(\hat{\rho}_i) = C_{ii}. \]

The off-diagonal elements give information about the covariance of the parameters. The correlation matrix \( R \), with elements \( R_{ij} \), can be computed from the covariance matrix \( C \)

\[ R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \]

The relative 100(1 - \( \alpha \)% confidence interval for the \( i \)th parameter is given by [23]:

\[ 1 \pm t_{n_k-n_p}^{\alpha/2}\sqrt{\text{var}(\hat{\rho}_i)/\hat{\rho}_i}, \]

in which \( t_{n_k-n_p}^{\alpha/2} \) denotes a quantile of the \( t \)-distribution with \( n_k - n_p \) degrees of freedom. Note that the statistical analysis only provides local properties for the parameter estimation problem instead of global ones, as multiple local minima with different properties could exist due to the nonlinear nature of the equations in RAPTOR. In table 3 the estimated model parameters \( \hat{\rho} \) and the relative deviations from the true model parameters are presented for the numerical experiment with noise when using the tabled initial guess of the model parameters. The table also contains the 95% confidence intervals.

The true model parameters are all within the 95% confidence intervals. The model parameters \( \chi_{neo} \) has a large

| Table 3. Results of the parameter estimation for the numerical experiment contaminated with white noise. |
|---|---|---|---|---|---|---|---|---|
| \( \chi_{neo} \) | \( c_{neo} \) | \( a_{ic} \) | \( w_{ic} \) | \( c_{savw} \) | \( w_{savw} \) | \( c_{e} \) | \( c_{e,pl} \) | \( w_{e,pl} \) |
| Initial guess \( p \) | 0.7 | 4 | 0.45 | 2.8 | 8 | 8 | 0.7 | 0.3 | 9.5 |
| True \( p \) | 0.5 | 7 | 0.5 | 3 | 10 | 10 | 0.8 | 0.5 | 10 |
| Estimated \( \hat{\rho} \) | 0.4965 | 7.0181 | 0.4924 | 3.1156 | 9.9469 | 9.9815 | 0.7960 | 0.4930 | 10.3615 |
| Relative error [\%] | 0.6981 | 0.2587 | 1.5220 | 3.8529 | 0.5310 | 0.1846 | 0.5048 | 1.3931 | 3.6154 |
| ±95% conf. int. \( \hat{\rho} \) [\%] | 21.5129 | 1.8492 | 4.2277 | 11.3485 | 2.3436 | 5.1799 | 1.6734 | 3.1060 | 8.3973 |
degree of uncertainty. Typically, the value of the electron heat diffusivity $\chi_e$ profile is more than a factor 10 greater than $\chi_{neo}$. The relative contribution of $\chi_{neo}$ to the $\chi_e$ profile is therefore very small. Perturbing the value of $\chi_{neo}$ will have little influence on $\chi_e$ and thus on the cost function. It can be concluded that the model parameter $\chi_{neo}$ is poorly observable from the data and could be kept outside the set of identified parameters when performing the parameter estimation for experimental data.

For the nonlinear parameter estimation problem, the 95% confidence intervals commonly underestimate the actual uncertainty of the estimated parameters. This is because the off-diagonal elements of the covariance matrix are ignored for the computation of the confidence intervals. As a consequence, parameter correlation is ignored. For example, from the correlation matrix $R$ it follows that there is a strong negative correlation ($-0.8767$) between $\chi_{neo}$ and $c_{ano}$. By increasing $\chi_{neo}$ and simultaneously decreasing $c_{ano}$, a solution can be obtained that is very nearly as good as the one with the estimated model parameters. The 95% confidence intervals in table 3 should therefore be interpreted as lower limits on the parameter uncertainties.

### 5.2. Parameter estimation for TCV

In this section, the parameter estimation procedure is used to identify the unknown model parameters in RAPTOR, so as to give reliable predictions of the $T_e$, $\iota$ and $U_{pl}$ profile evolution in TCV experiments.

#### 5.2.1. Experimental setup

Since direct experimental measurements of the $q$ profile are not available at TCV, interpretative transport simulations were run with the more complete ASTRA code [10] using the measured loop voltage $U_{pl}$, plasma current $I_p$, plasma boundary, electron temperature profile $T_e$ and electron density profile $n_e$. The following procedure was followed:

- TCV experiments were performed, taking measurements of the electron temperature (from Thomson Scattering), line-averaged density (interferometer), loop voltage and plasma current;
- No experimental measurements of $T_i$ were available, so we used instead the following formula for the ion pressure $P_i = P_e \cdot 2I_{Zeff} \cdot 0.07 \cdot (\pi/10)^{0.8} \cdot (\rho_{e0})^{0.6}$ and $T_i = p_i/(e \cdot n_i)$. ($\delta$ is the line-averaged density in electrons per cubic meter, $e$ is the electric charge). This formula was obtained based on a series of simulations and on the observation that, on TCV, the ion temperature does not change much when EC power is used;
- The deposition location of heat by the auxiliary EC heating system was obtained using the TORAY-GA code [27];
- The ‘experimental’ thermal conductivity was calculated filling in the spline-fitted experimental profiles of $n_e$ and $T_e$ into (2), also using information about the plasma shape obtained from equilibrium reconstruction;
- When appropriate, this thermal conductivity is modified to include an effective reduction of core confinement due to sawtooth activity;
- The thermal conductivity was used as input in a new predictive ASTRA simulation, evolving $T_e$ and $\psi$ as a result of the applied ECH/ECCD actuator powers. This was done to get a high-resolution representation of the evolution of all the profiles.

This procedure was carried out for a selection of 12 relevant TCV shots, which mutually featured a wide range of confinement regimes and settings for the EC heating and current drive system. The individual discharges also contained a large variation in time of the applied auxiliary power. Time evolutions of the key parameters of these discharges can be seen in figures 10–13.

#### 5.2.2. Parameter estimation procedure

A three-way data split approach [28] was used in order to identify the model parameters in RAPTOR for TCV discharges. The twelve TCV shots were divided into three disjoint sets: a training, validation and test set. The training set was used to estimate the model parameters with the developed parameter estimation procedure that yield reliable predictions of the $T_e$, $\iota$ and $U_{pl}$ profiles. The aim is not that RAPTOR perfectly fits the training data, but gives reliable predictions for shots it has not been trained on. Therefore a validation set was used to investigate if during the learning phase overfitting [29] occurred. This would imply that RAPTOR not only captures the underlying phenomenon of the training data but also the noise associated with the particular data set. Lastly, the test set was used to assess the real performance of RAPTOR on unseen data when using the estimated model parameters. The training set consists of shots 46 712, 46 713, 4 7500, 48 041, 48 042 and 49 293, the validation set of shots 46 715, 47 501 and 48 042 and 49 293, the test set which was utilized to assess the performance of RAPTOR’s predictions. The ramp-up and ramp-down phases of these discharges were disregarded to avoid issues related to the initial conditions and strongly time-varying geometric profiles.

The following model parameters in the model for $\sigma_b$ and $\chi_e$ were estimated:

| $\chi_{neo}$ | $c_{ano}$ | $c_{saww}$ | $w_{saww}$ | $c_T$ | $c_{fsww}$ | $w_{fsww}$ |
|--------------|-----------|-------------|------------|------|------------|------------|
| Initial guess $p^*$ | 0.1 | 4 | 25 | 10 | 0.1 | 0.4 | 20 |
| Estimated $p'$ | 0.1785 | 4.0948 | 46.4064 | 15.1183 | 0.7907 | 0.2174 | 50.0000 |
| ±95% conf. int. $p'^{\%}$ | 3.2251 | 2.1540 | 8.5822 | 3.8327 | 3.0530 | 4.1603 | 3.0827 |
The model parameters that describe sawteeth were estimated, because sawtooth crashes are present in the TCV shots and the $q$ profile drops below 1. However, the model parameters $a_{ic}$, $w_{ic}$ and $d_{ic}$ were not estimated, because the shots considered do not represent advanced plasma scenarios characterized by a very low or negative magnetic shear $s$. Therefore the data does not contain information about $a_{ic}$, $w_{ic}$ and $d_{ic}$, and these parameters are therefore unidentifiable.

Conversely, since they do not influence the profile evolution in this regime, $a_{ic}$ is set to 0 to remove the effect of enhanced shear entirely.

5.2.3. Results of parameter estimation. The model parameters were estimated for the six shots simultaneously in the training set in order to describe accurately a variety of plasma scenarios with RAPTOR. The time evolution of the $T_{ei}$, $i$ and $U_{pl}$ profiles from the interpretative ASTRA transport simulations...
were fitted by the parameter estimation procedure using the cost function definition (21). In table 4 the estimated parameters at the final iteration step and the corresponding 95% confidence intervals are given for the training set. This table also includes the realistic initial guess of the parameters that was used at the first iteration step. The cost function \( J \) of the training set versus the iteration steps is depicted in figure 12.

At each iteration step during the learning process a new set of model parameters is computed which reduce the difference between measurements and predictions. The same model parameters are also used to calculate the cost function of the validation set. Overfitting is avoided by stopping the learning process when the cost function of the validation set starts increasing while the cost function of the learning set still decreases with each iteration step. The cost function of the validation set is shown in figure 13 and is in contrast to the cost function of the learning set not monotonically decreasing because of the slight increase around iteration step 7. Despite this non monotonic character however, the estimated model parameters at the final iteration step for the training set also yield the lowest value for the cost function of the validation set, which indicates that no overfitting during the learning process occurred. Moreover, the identified model parameters
yield predictions of the loop voltage $U_{pl}$ that are in good agreement with measurements, as can be seen from the low value of $\nu_{UU}$ in figures 12 and 13. The parameters in table 4 are therefore the optimal parameters to describe the TCV shots with RAPTOR and will be used to assess the quality of RAPTOR’s predictions of the temporal evolution of the $T_e$, $\iota$ and $U_{pl}$ profiles for shots not used in the parameter estimation process.

In figures 14–16 RAPTOR’s predictions of the temporal evolution of the $T_e$ and $\iota$ profiles are compared with the results of the ASTRA code for the three shots of the test set and when using both the initial guess of the model parameters at the first iteration step and the optimal model parameters at the final iteration step. The optimal model parameters decrease the cost function of the learning set by more than a factor five compared to their initial guess. From the definition of the cost function it follows that a lower value implies better agreement between the ASTRA interpretative results and RAPTOR. This can indeed be observed from the relative difference plots of the $T_e$ and $\iota$ profiles computed by ASTRA and RAPTOR at the first and final iteration step in figures 14–16. When using the identified model parameters the relative difference between the predictions of the $T_e$ profile of RAPTOR and ASTRA is at most 20% and the relative difference of the $\iota$ profile is below 10%.

The model parameters obtained by this technique are given in table 4. We note that the values correspond to our physics understanding of core electron transport. The neoclassical level, $\chi_{\text{neo}}$, is about one order of magnitude smaller than the anomalous level, confirming anomalous transport is dominating (trapped electron mode in this case [30]). We find that the sawtooth effects can be assimilated with a diffusion coefficient another order of magnitude larger than $c_{\text{ano}}$. This is consistent with significant profiles flattening observed in tokamak central plasmas. On the other hand the value of $c_{T_e} = 0.7$, is near the lower expected bound. Note that in these TCV L-mode plasmas, the power degradation is also smaller than the H-mode scaling law, $P \propto T_e^{0.55}$ [30], which would correspond to $c_{T_e} = 1.2$. However other parameters are involved with increasing input power, for example increased recycling and hence density. It will be interesting to perform a similar study with models including particle transport.

6. Outlook

In this paper a methodology was presented to estimate the parameters for a nonlinear control-oriented transport model. This was applied for estimating the parameters of the real-time transport code RAPTOR. A semi-empirical transport model was used to describe the electron heat diffusivity, which included the time-averaged sawtooth behavior by an ad hoc addition to the neoclassical conductivity and electron heat diffusivity. For future work it would be interesting to assess the effect of the choices of the cost function, weighing factors and
optimization routine on the quality of the predictions of the transport model.

There are other means to achieve good predictions using real-time capable transport models. One option, mentioned in the introduction, is to use neural network fits of the output of a quasilinear code [20]. Another option is the recent work on the non-stiffness of edge transport in L-mode tokamaks [30]. In this work it was shown that the average effect of sawteeth on density is different than on $T_e$, the inversion radius $\rho_{\text{inv}}$ being larger for $n_e$. To date a prescribed density profile is used in RAPTOR. The reason for this choice is that the most important nonlinear coupling between plasma profiles during a tokamak discharge stems from the electron temperature-dependent resistivity and bootstrap current and the $q$ profile-dependent confinement. The inclusion of a model for $n_e$ and the effect of sawteeth on this profile, as described in [30], would increase the quality of RAPTOR’s predictions. We would like to stress that the presented parameter estimation method is generic, and also completely different models for the neoclassical conductivity and electron heat diffusivity could be used as long as they appear as analytical functions of the plasma state variables ($q$, $T_e$, etc).

Capturing more physics with a control-oriented transport model will result in an increase of the number of unknown model parameters. Choosing all the right model parameters by hand will become a tedious, time consuming task and therefore the presented parameter estimation method would greatly aid with this process.

A wider operating regime of the tokamak can be spanned by the model by using data from a wider operating regime to fit the parameters. In this work the considered data of the TCV tokamak did not represent advanced plasma scenarios characterized by a very low or negative magnetic shear $s$. Also, more direct measurements of the $q$ profile, without requiring the intermediate step of interpretative transport simulations, should also result in a more accurate transport model. As a result some model parameters could not be identified. In the future, the control-oriented transport model in combination with the parameter estimation method can be used on other tokamaks as well in order to assess the performance of the developed methods. Finally, we mention the possibility to use the parameter estimation technique based on high-fidelity model simulations of future tokamaks (ITER, JT60-SA) to yield a fast, real-time capable 1D profile evolution model.

7. Conclusions

RAPTOR is a control-oriented tokamak profile evolution model designed to run in real-time during the plasma evolution, as well as for use in actuator trajectory optimization routines and controller design and testing. It uses semi-empirical models for the electron heat diffusivity $\chi_e$ and the electrical...
conductivity $\sigma_\parallel$ that contain unknown parameters, which need to be estimated in order to make reliable plasma profile predictions. A systematic method, based on the nonlinear least-squares theory, was developed to identify these model parameters.

A cost function is formulated reflecting the discrepancy between simulated and known, either measured or reconstructed, profile evolutions. The performance of the developed estimation method critically depends on the derivative of the cost function to the model parameters. In the developed estimation method this gradient is computed analytically, resulting in a faster convergence to the solution and higher performance as this analytical expression is exact.

It was shown that the parameter estimation method successfully recovers known model parameters from simulated data, demonstrating the correct implementation of the method. In reality, the measured plasma profiles are contaminated with noise. This situation was mimicked by adding white noise to simulation data. A statistical analysis of the identified model parameters was done for the noisy case, as parameter values could be affected strongly by seemingly insignificant variations in the data. In the noisy case the developed parameter estimation method does not recover the model parameters exactly, but the true model parameters are all covered by the calculated 95% confidence intervals. The statistical analysis also shows that, occasionally, model parameters can not be estimated independently: the effect of a specific model parameter can be compensated by another model parameter. The parameter estimation method was applied for shots of the TCV tokamak. The model parameters of $\chi_e$ and $\sigma_\parallel$ were estimated such that RAPTOR’s predictions are reliable for a variety of plasma conditions and match post-shot interpretative profile simulations using the ASTRA code. A set of parameters was found such that the difference between the experimental and simulated profiles was less than 10% in flat-top phase for the $\iota$ profile and 20% for the $T_e$ profile for the discharges considered.

Acknowledgments

This work has been supported in part by the Netherlands Organisation for Scientific Research and by the Swiss National Science Foundation.

References

[1] Gormezano C et al 2007 Nucl. Fusion 47 S285
[2] Moreau D et al 2008 Nucl. Fusion 48 106001
[3] Barton J E et al 2012 Nucl. Fusion 52 123018
[4] Maljaars E et al 2015 Nucl. Fusion 55 023001
[5] Felici F et al 2012 Plasma Phys. Control. Fusion 54 025002
[6] Xu C et al 2010 IEEE Trans. Plasma Sci. 38 163
[7] van Dongen J et al 2014 Plasma Phys. Control. Fusion 56 125008
[8] Felici F et al 2014 American Control Conf. pp 4816–23
[9] Felici F et al 2014 41st EPS Conf. on Plasma Physics (Berlin, Germany)
[10] Pereverzev G V et al 2002 Astra Automated System for Transport Analysis in a Tokamak Technical Report 5/98 IPP Report
[11] Artaud J et al 2010 Nucl. Fusion 50 043001
[12] Witrant E et al 2007 Plasma Phys. Control. Fusion 49 1075
[13] Ou Y et al 2007 Fusion Eng. Des. 82 1153
[14] Felici F et al 2011 Nucl. Fusion 51 083052
[15] Moreau D et al 2011 Nucl. Fusion 51 063009
[16] Moreau D et al 2015 Nucl. Fusion 55 063011
[17] Staebler G M et al 2007 Phys. Plasmas 14 055909
[18] Bourdelle C et al 2007 Phys. Plasmas 14 112501
[19] Rafiq T et al 2013 Phys. Plasmas 20 032506
[20] Citrin J et al 2015 Nucl. Fusion 55 092001

[21] Hinton F L et al 1976 Rev. Mod. Phys. 48 239
[22] Sauter O et al 1999 Phys. Plasmas 6 2834
[23] Sauter O et al 2002 Phys. Plasmas 9 5140
[24] Bard Y 1974 Nonlinear Parameter Estimation (New York: Academic)
[25] Nocedal J et al 2006 Numerical Optimization 2nd edn (Berlin: Springer)
[26] Aster R C et al 2005 Parameter Estimation and Inverse Problems (Amsterdam: Elsevier)
[27] Matsuda K 1989 IEEE Trans. Plasma Sci. 17 6
[28] Borovicka T et al 2012 Selecting representative data sets Advances in Data Mining Knowledge Discovery and Applications (Rijeka: In Tech) (doi:10.5772/50787)
[29] Hawkins D 2004 J. Chem. Inf. Comput. Sci. 44 1
[30] Sauter O et al 2014 Phys. Plasmas 21 055906