Phenomenological Modeling of Plasma Transport via Stochastic Filtering

Chao Xu, Yongsheng Ou and Eugenio Schuster

***Abstract*— The accuracy of first-principles predictive models for the evolution of plasma profiles is sometimes limited by the lack of understanding of the plasma transport phenomena. It is possible then to develop approximate transport models for the prediction of the plasma dynamics which are consistent with the available diagnostic data. This data-driven approach, usually referred to as phenomenological modeling, arises as an alternative to the more classical theory-driven approach. In this work we propose a stochastic filtering approach based on an extended Kalman filter to provide real-time estimates of poorly known or totally unknown transport coefficients. We first assume that plasma dynamics can be governed by tractable models obtained by first principles. However, the transport parameters are considered unknown and to-be-estimated. These estimates will be based solely on input-output diagnostic data and limited understanding of the transport physics. Numerical methods (e.g., finite differences) can be used to discretize the PDE models both in space and time to obtain finite-dimensional discrete-time state-space representations. The system states and to-be-estimated parameters are then combined into an augmented state vector. The resulting nonlinear state-space model is used for the design of an extended Kalman filter that provides real-time estimations not only of the system states but also of the unknown transport coefficients. Simulation results demonstrate the effectiveness of the proposed method for a benchmark transport model in cylindrical coordinates.**

I. Introduction

Mathematical modeling of plasma transport phenomena with modest complexity but capturing dominant dynamics is critical for plasma control design. Transport theories (classical, neoclassical and anomalous) produce, under necessary assumptions, strongly nonlinear models based on partial differential equations (PDEs). However, the complexity of these models often makes them not useful for control design since it is very challenging, if not impossible, to synthesize compact and reliable control strategies based on these complicated mathematical models. As an alternative, data-driven modeling techniques, including system identification [1] and data assimilations [2], have the potential to obtain practical, low-complexity, dynamic models for the control of plasmas systems.

Data-driven modeling techniques have been successfully used in the past to model plasma transport dynamics for active control design in nuclear fusion reactors. System identification using input and output (I/O) diagnostic data has been used to model the current profile dynamics in ASDEX Upgrade [3]. In the JET tokamak [4], a two-time-scale linear system is used to describe the dynamics of the magnetic and kinetic profiles around certain quasi-steady state trajectories, where system matrices can be identified from the experimental or simulation data using system identification algorithms in [1]. In the L-mode discharges of the JT-60U tokamak [5], diffusive and non-diffusive coefficients of the momentum transport equation of the toroidal rotation profile dynamics are estimated from transient data obtained by modulating the momentum source.

First-principles modeling of the plasma profile dynamics usually results in multiple-input-multiple-output (MIMO) infinite-dimensional transport models. Using the method of averaging over magnetic surfaces, the transport model can be formulated into one dimensional (1D) PDEs with respect to a variable indexing the magnetic surfaces [6], [7]. System identification often generates dynamic models fitting the input/output diagnostic data but does not take into account the physical structure of the transport model obtained by first principles. In this case the states of the identified models do not necessarily represent physical variables. In this work, we propose instead to use the tractable 1D PDE structure [7] of the first-principle model to estimate its transport coefficients using experimental data. Various numerical methods (such as the finite difference method [8]) can be used to obtain fully spatial-temporal discretized models in terms of given spatial nodes and sampling rates. For finite-dimensional discrete-time systems, stochastic filters (e.g., the Kalman filter) can be used to estimate the system states based on the input/output measurements. In order to be able to also estimate system parameters, such as the transport coefficients, it is possible to define an augmented state vector which includes both the original system states and these to-be-estimated parameters. The overall discrete-time model becomes nonlinear but stochastic filters (e.g., the extended Kalman filter) can still be used to estimate the augmented state vector.

The paper is organized as follows. We introduce a linear parabolic PDE system in Section II that retains the general structure of plasma transport models under the circular cylindrical approximation. Then, an explicit numerical discretization scheme [8] is derived based on the finite difference method over a given spatial–temporal grid. The stability of the numerical scheme is also discussed in this section. In Section III, we summarize the extended Kalman filter theory used in this work for the estimation of both system states and transport parameters. In Section IV, we test the performance of the proposed method in simulations. We close the paper by stating conclusions and potential research topics in Section V.

This work was supported by the NSF CAREER award program (ECCS-0645086).

C. Xu (chx205@lehigh.edu), Y. Ou and E. Schuster are with the Department of Mechanical Engineering and Mechanics, Lehigh University, 19 Memorial Drive West, Bethlehem, PA 18015, USA.

II. 1D Parabolic System and Discrete Schemes

Without loss generality, we consider the following parabolic system [9], [7]

$$\frac{\partial x}{\partial t}(\xi,t)=\frac{1}{\xi}\frac{\partial}{\partial \xi}\left[\xi\vartheta\frac{\partial x(\xi,t)}{\partial \xi}+Vx(\xi,t)\right]+S_{\rm IN}(\xi,t), \quad (1)$$

$$\frac{\partial x(0,t)}{\partial \xi}=0,\ \frac{\partial x(1,t)}{\partial \xi}=S_{\rm BC}(t),\ x(\xi,t_{\rm I})=x_0(\xi), \qquad (2)$$

where $x(\xi,t)$ represent a general plasma profile with respect to the normalized spatial coordinate $\xi\in[0,1]$ and time $t\in[t_{\rm I},t_{\rm F}]$. The parameters $\vartheta(\xi)$ and $V(\xi)$ are unknown and to-be-estimated based on observational data. Interior sources and boundary controls are denoted by $S_{\rm IN}(\xi,t)$ and $S_{\rm BC}(t)$, respectively. The initial distribution is denoted by $x_0(\xi)$. In order to simplify the presentation of the stochastic filtering approach for parameter estimation of 1D plasma transport models, in this work we assume without loss of generality that ϑ is a constant, $V(\xi)\equiv 0$, and the interior source can be represented as $S_{\rm IN}(\xi,t)=b_{\rm IN}(\xi)s_{\rm IN}(t)$.

In the rest of this section, we derive a discrete representation of the continuous PDE system (1)–(2) using an explicit scheme over the following spatial-temporal grid division:

$$0=\xi_0<\xi_1<\ldots<\xi_i<\ldots<\xi_M=1, \qquad (3)$$

$$t_0=t_{\rm I}<t_1<\ldots<t_j<\ldots<t_N=t_{\rm F}, \qquad (4)$$

where $\xi_i=ih$ and $t_j=t_0+jT$. The profile function is then rewritten as $x_i^j=x(\xi_i,t_j)$. The boundary conditions are discretized as

$$\frac{\partial x(\xi_0,t_j)}{\partial \xi}=\frac{-3x_0^j+4x_1^j-x_2^j}{2h}=0, \qquad (5)$$

$$\frac{\partial x(\xi_M,t_j)}{\partial \xi}=\frac{x_{M-2}^j-4x_{M-1}^j+3x_M^j}{2h}=S_{\rm BC}^j. \qquad (6)$$

Over the interior nodes $\xi_1,\ldots,\xi_{M-1}$, we obtain the following discrete schemes,

$$\begin{aligned}&\frac{x_i^{j+1}-x_i^j}{T}\\ =&\vartheta\frac{x_{i-1}^j-2x_i^j+x_{i+1}^j}{h^2}+\vartheta\frac{x_{i+1}^j-x_{i-1}^j}{2ih^2}+S_{{\rm IN},i}^j \qquad (7)\\ =&\frac{\vartheta}{h^2}\left[\frac{2i-1}{2i}x_{i-1}^j-2x_i^j+\frac{2i+1}{2i}x_{i+1}^j\right]+S_{{\rm IN},i}^j.\end{aligned}$$

We substitute (5)–(6) into (7) for $i=1$ and $i=M-1$, respectively. Then, we obtain the following discrete system:

$$\begin{cases} x_1^{j+1}=-\frac{4T\vartheta}{3h^2}\left[x_1^j-x_2^j\right]+x_1^j+TS_{{\rm IN},1}^j,\\ x_i^{j+1}=\frac{T\vartheta}{h^2}\left[\frac{2i-1}{2i}x_{i-1}^j-2x_i^j+\frac{2i+1}{2i}x_{i+1}^j\right]+x_i^j\\ \qquad +TS_{{\rm IN},i}^j,\\ x_{M-1}^{j+1}=\frac{2(M-2)T\vartheta}{3(M-1)h^2}x_{M-2}^j-\frac{2(M-2)T\vartheta}{3(M-1)h^2}x_{M-1}^j\\ \qquad +x_{M-1}^j+\frac{(2M-1)T\vartheta}{3(M-1)h}S_{\rm BC}^j+TS_{{\rm IN},M-1}^j.\end{cases} \qquad (8)$$

The system measurement is defined by

$$y(\xi,t)=\frac{1}{\xi}\frac{\partial x(\xi,t)}{\partial \xi}, \qquad (9)$$

which can be discretized as

$$\begin{aligned} y_1^{j+1}&=\frac{1}{2h^2}\left[x_2^{j+1}-x_0^{j+1}\right]=\frac{2}{3h^2}\left[x_2^{j+1}-x_1^{j+1}\right],\\ y_i^{j+1}&=\frac{1}{2ih^2}\left[x_{i+1}^{j+1}-x_{i-1}^{j+1}\right],\quad i=2,\ldots,M-2,\\ y_{M-1}^{j+1}&=\frac{2}{3(M-1)h^2}\left[x_{M-1}^{j+1}-x_{M-2}^{j+1}\right]+\frac{S_{\rm BC}^{j+1}}{3(M-1)h}.\end{aligned} \qquad (10)$$

We discuss now the stability of the discretized model with respect to the iteration index j (which is critical for effective estimation) because textbooks (e.g., [10]) on stability of finite difference schemes usually do not include boundary conditions in the analysis. Additionally, the cylindrical geometry makes the stability analysis more complicated than in Euclidean coordinates. The propagation matrix of the discrete scheme (8) is denoted by $\Phi=\omega A+I$, where $\omega=\frac{T\vartheta}{h^2}$, I is an identity matrix and A is the coefficient matrix of the discretization of the spatial derivatives in (8). By introducing a vector $\mathbf{x}^j$ as the collection of the unknowns $x_1^j,\ldots,x_{M-1}^j$, we can rewrite the discrete scheme (8) as $\mathbf{x}^{j+1}=\Phi\mathbf{x}^j+S^j$, where S^j represents the sources terms. Based on stability theory of linear systems [10], the numerical scheme is numerically stable if and only if all the eigenvalues of the propagation matrix Φ satisfy $|\lambda|<1$, where $\lambda\in\mathbb{C}$ solves the characteristic polynomial equation $\det(\lambda I-\Phi)=0$. We note that the matrix $(\lambda I-\Phi)$ takes a tridiagonal form

$$\lambda I-\Phi=\begin{bmatrix}\alpha_1 & \gamma_1 & & & \\ \beta_2 & \alpha_2 & \gamma_2 & & \\ & \cdots & \cdots & \cdots & \\ & & & \beta_{M-1} & \alpha_{M-1}\end{bmatrix}.$$

We carry out the LU decomposition, $\lambda I-\Phi=LU$, where L and U are defined by

$$L=\begin{bmatrix}l_1 & & & \\ \beta_2 & l_2 & & \\ & & \cdots & \\ & & \beta_{M-1} & l_{M-1}\end{bmatrix},\ U=\begin{bmatrix}1 & \mu_1 & & \\ & 1 & \mu_2 & \\ & & \cdots & \\ & & & 1\end{bmatrix},$$

with $l_1=\alpha_1$, $\mu_i=\gamma_i/l_i,(i=1,\ldots,M-2)$, $l_i=\alpha_i-\beta_i\mu_{i-1},(i=2,\ldots,M-1)$. Therefore, the λ–polynomial is determined by

$$p(\lambda,\omega)=\det(\lambda I-\Phi)=\det L=\Pi_{i=1}^{M-1}l_i.$$

By using algebraic tools (e.g., the discrete Routh–Hurwitz theorem or the Jury stability criterion [10]), we can obtain a stability condition for the discrete scheme in terms of the discretization parameter $\omega=\frac{T\vartheta}{h^2}$. Since the condition is expressed in terms of the unknown ϑ, we must make the time step T small enough to satisfy the stability condition for the whole possible range of ϑ.

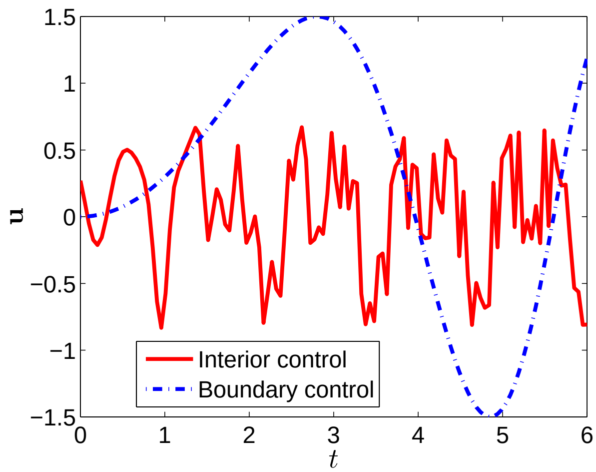

Fig. 1. Interior and boundary excitation signals defined by (20)–(21).

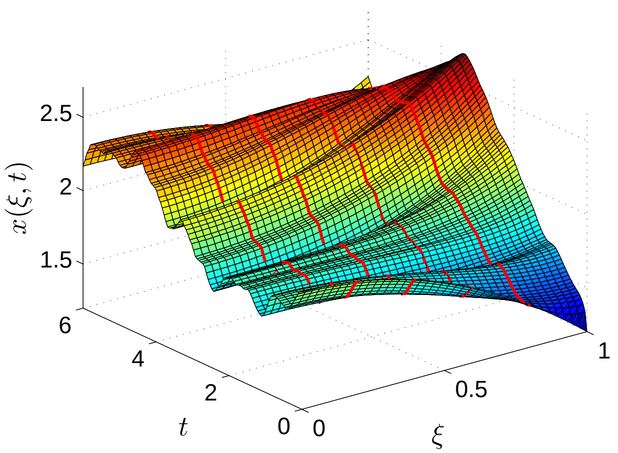

Fig. 2. Spatial-temporal evolution of the PDE system (1)–(2) with the excitation signals (20)–(21).

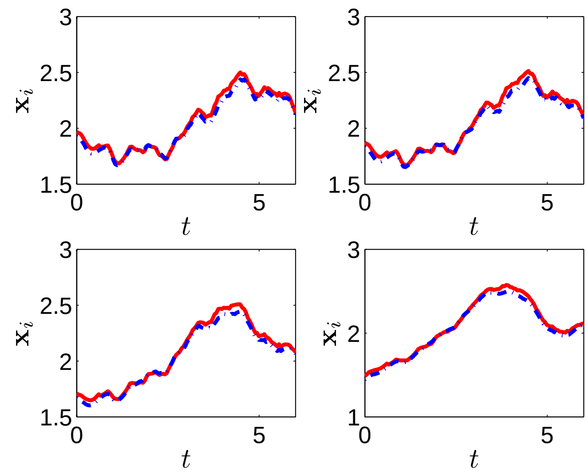

Fig. 3. Computational and estimated states (red solid v.s. blue dot lines).

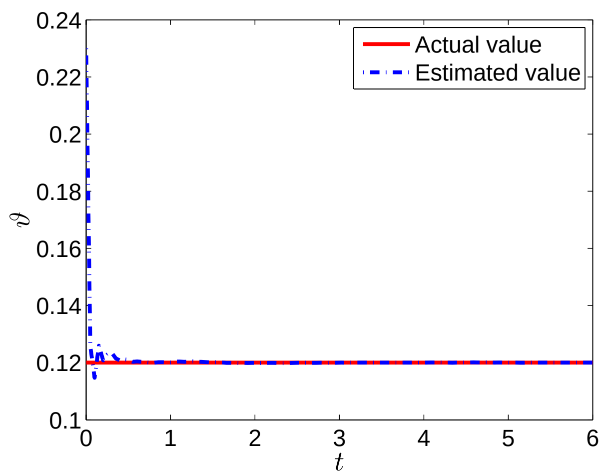

Fig. 4. Parameter estimation based on the extended Kalman filter.

III. The Extended Kalman Filter

By noting that $S^j_{\mathrm{IN},i} = b_{\mathrm{IN},i} s^j_{\mathrm{IN}}$, we can define

$$\mathbf{u}^j = \left[\begin{array}{cc} s^j_{\mathrm{IN}} & S^j_{\mathrm{BC}} \end{array}\right]^{\mathrm{T}}. \tag{11}$$

By introducing the augmented state and system output

$$\mathbf{z}^j = \left[\begin{array}{ccccc} x^j_1 & x^j_2 & \ldots & x^j_{M-1} & \vartheta^j \end{array}\right]^{\mathrm{T}}, \tag{12}$$

$$\mathbf{y}^j = \left[\begin{array}{cccc} y^j_1 & y^j_2 & \ldots & y^j_{M-1} \end{array}\right]^{\mathrm{T}}, \tag{13}$$

we can rewrite the discretized system (8)–(10) as the following nonlinear state space representation

$$\mathbf{z}^{j+1} = f(\mathbf{z}^j, \mathbf{u}^j) + \mathbf{w}^j, \quad \mathbf{y}^{j+1} = h(\mathbf{z}^{j+1}, \mathbf{u}^j) + \mathbf{v}^{j+1},$$

where f and h are the state and measurement mappings defined by the finite difference schemes in (8)–(10) and the equation for ϑ given by $\vartheta^{j+1} = \vartheta^j$. The disturbance input $\mathbf{w}$ and the measurement noise $\mathbf{v}$ are assumed to be white, zero–mean Gaussian random sequences, i.e., $\mathbf{w}^j \sim N(0, Q^j)$ and $\mathbf{v}^j \sim N(0, R^j)$, which satisfy the following properties

$$E\left[\mathbf{w}^{j_1} \cdot \mathbf{v}^{j_2}\right] = 0, \quad \forall j_1, j_2, \tag{14}$$

$$E\left[\mathbf{v}^{j_1} \cdot \mathbf{v}^{j_2}\right] = 0,\ E\left[\mathbf{w}^{j_1} \cdot \mathbf{w}^{j_2}\right] = 0, \quad \forall j_1 \neq j_2, \tag{15}$$

$$E\left[\mathbf{w}^{j} \cdot \mathbf{w}^{j}\right] = Q^j,\ E\left[\mathbf{v}^{j} \cdot \mathbf{v}^{j}\right] = R^j, \tag{16}$$

where Q^j and R^j are covariance matrices.

In the rest of this section we give a brief introduction of the extended Kalman filter which is the nonlinear version of the well-known Kalman filter in estimation theory. In the following equations, we use $\hat{\mathbf{z}}^{j+1|j}$ to represent the state propagation before the measurements are considered,

$$\hat{\mathbf{z}}^{j+1|j} = f(\hat{\mathbf{z}}^{j|j}, \mathbf{u}^j), \tag{17}$$

where $\hat{}$ represents the estimated value. We then compute the Jacobian matrices with respect to the current state $\hat{\mathbf{z}}^{j|j}$, the propagation state $\hat{\mathbf{z}}^{j+1|j}$ and the control input $\mathbf{u}^j$,

$$F^j = \left.\frac{\partial f(\mathbf{z}, \mathbf{u})}{\partial \mathbf{z}}\right|_{\hat{\mathbf{z}}^{j|j}, \mathbf{u}^j}, \quad H^{j+1} = \left.\frac{\partial h(\mathbf{z}, \mathbf{u})}{\partial \mathbf{z}}\right|_{\hat{\mathbf{z}}^{j+1|j}, \mathbf{u}^j}. \tag{18}$$

We are able to improve the propagation result in (17) by taking into account the measurement $\mathbf{y}^{j+1}$,

$$\hat{\mathbf{z}}^{j+1|j+1} = \hat{\mathbf{z}}^{j+1|j} + K^{j+1}\left[\mathbf{y}^{j+1} - h(\mathbf{z}^{j+1|j}, \mathbf{u}^j)\right], \tag{19}$$

where the gain K^{j+1} is determined as

$$P^{j+1|j} = F^j P^{j|j} \left(F^j\right)^{\mathrm{T}} + Q^j,$$
$$K^{j+1} = P^{j+1|j} \left(H^{j+1}\right)^{\mathrm{T}} \Big[H^{j+1} P^{j+1|j} \left(H^{j+1}\right)^{\mathrm{T}} + R^{j+1}\Big]^{-1},$$
$$P^{j+1|j+1} = \left(I - K^{j+1} H^{j+1}\right) P^{j+1|j}.$$

More details on the extended Kalman filter can be found in [2].

IV. A Numerical Example

A. Numerical simulation of the PDE system - dense grid

We first solve the PDE system (1)–(2) on a dense grid based on an implicit finite difference scheme. We consider a constant parameter $\vartheta = 0.12$ in the simulation. The interior actuation function is given by $b_{\mathrm{in}}(\xi) = 1 - \xi^4$, $(0 \le \xi \le 1)$ and the excitation signals (shown in Fig. 1) are chosen as

$$s_{\mathrm{in}}(t) = \frac{1}{3}\sin(5t) + \frac{1}{2}\cos\left[10t + \cos(5t^2)\right], \qquad (20)$$
$$S_{\mathrm{BC}}(t) = \frac{3}{2}\sin\left(\frac{1}{5}t^2\right). \qquad (21)$$

The spatial–temporal domain is given by $\Omega = \{(\xi, t) : 0 \le \xi \le 1, 0 \le t \le 6\}$. To start the simulation, the initial distribution is assumed as $x(\xi, 0) = x_0(\xi) = 2 - \frac{4}{5}\xi^2$. The simulation of system (1)–(2) is carried out over the grid nodes (3)–(4) with time step $T = T_d = 0.05$(s) and spatial step $h = h_d = 0.025$, where we use the subscript d to denote a dense grid division. The spatial–temporal evolution obtained from the numerical simulation is shown in Fig. 2 where the evolutionary trajectories corresponding to the values at the four finite-difference node points used for the extended Kalman filter are marked with red solid lines on the 3D surface.

B. The extended Kalman filter - sparse grid

We now let $h = h_s = 0.2$ and $T = T_s = 0.05$, where the subscript s denotes an sparse grid division. We use an explicit difference scheme to obtain a fourth order discrete system based on (8). The measurements defined by (9) are also taken at the same spatial nodes that are used to obtain the state propagation scheme (8). An initial guess is needed to start the estimation of both the system states and the transport parameter ϑ. A good guess of the initial states is important to obtain a correct estimation. If the initial estimates of the states are far from the actual values, the iteration of the extended Kalman filter can diverge soon after several propagation steps due to the linearization with respect to the estimated states. However, a relatively fair initial guess of the to-be-estimated parameter within the stability region of the finite difference scheme is good enough to robustly generate an accurate parameter estimation. The initial guesses for the states are perturbed values of the initial distribution $x(\xi, 0) = x_0(\xi) = 2 - \frac{4}{5}\xi^2$ at the four difference node points. The initial guess for the parameter is given by $\hat{\vartheta} = 0.23$. With the given initial settings, we use the extended Kalman filter to obtain both state and parameter estimations which are shown in Fig. 3 and Fig. 4, respectively. The parameter estimation can converge to the real value rapidly.

V. Conclusions

We consider a parameter estimation problem for a benchmark model in plasma transport, which is governed by a 1D parabolic PDE. The explicit scheme is then used to obtain a finite-dimensional discrete-time approximation based on the finite-difference discretization of the PDE system over a given spatial-temporal grid division. By including the unknown transport coefficient as an augmented state variable, we are able to reformulate the discrete-time linear system into an augmented nonlinear system. Then, the extended Kalman filtering technique is used to obtain real time estimations of both the system state and the transport coefficient based on the measurements.

In this work, we only consider the case where the to-be-estimated parameter is a constant in space. However, by parameterizing spatially varying parameters via a number of constants parameters distributed in space, it is possible to formulate the estimation problems of spatially distributed parameters within the framework discussed in this work. In order to avoid reducing T excessively in order to satisfy the stability condition for the proposed explicit discretization scheme, implicit *unconditionally stable* discretization schemes must be developed to obtain robust parameter estimations.

This work presents an alternative to the first-principles approach to the modeling of the plasma dynamics and transport phenomena by assimilating the experimental observations into transport PDE models with modest complexity. Since the assimilation of experimental data is carried out in real time, this method can be effectively integrated into a feedback plasma control system.

References

[1] L. Ljung, *System Identification: Theory for the User*. New Jersey: Prentice–Hall PTR, 1999.

[2] B. Anderson and J. Moore, *Optimal Filtering*. New Jersey: Prentice-Hall, 1979.

[3] Y.-S. Na, *Modelling of Current Profile Control in Tokamak Plasmas*. Fakultat fur Physik: Technische Universitat Munchen, 2003.

[4] D. Moreau, D. Mazon, A. Ariola, *et al.*, “A two-time-scale dynamic-model approach for magnetic and kinetic profile control in advanced tokamak scenarios on JET,” *Nuclear Fusion*, vol. 48, p. 106001, 2008.

[5] M. Yoshida, Y. Koide, H. Takenaga, H. Urano, N. Oyama, *et al.*, “Momentum transport and plasma rotation profile in toroidal direction in JT-60U L-mode plasmas,” *Nuclear Fusion*, vol. 47, pp. 856–863, 2007.

[6] F. Hinton and R. Hazeltine, “Theory of plasma transport in toroidal confinement systems,” *Reviews of Modern Physics*, vol. 48, pp. 239–308, 1976.

[7] J. Blum, *Numerical Simulation and Optimal Control in Plasma Physics*. New York: John Wiely & Sons, 1989.

[8] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations (2nd edition)*. Philadelphia: SIAM: Society for Industrial and Applied Mathematics, 2004.

[9] K. Gentle, “Dependence of heat pulse propagation on transport mechanisms: Consequences of nonconstant transport coefficients,” *Physics of Fluids*, vol. 32, pp. 1241–1258, 1992.

[10] P. Sarachik, *Principles of Linear Systems*. New York: Cambridge University Press, 1997.