Dynamic Optimization of Open-loop Input Signals for Ramp-up Current Profiles in Tokamak Plasmas[1]

Zhigang Ren[a], Chao Xu[a,2], Qun Lin[b], Ryan Loxton[b] and Kok Lay Teo[b]

[a] *State Key Laboratory of Industrial Control Technology and Institute of Cyber-Systems & Control, Zhejiang University, Hangzhou, Zhejiang, China.*
[b] *Department of Mathematics & Statistics, Curtin University, Perth, Western Australia, Australia.*

Abstract

Establishing a good current spatial profile in tokamak fusion reactors is crucial to effective steady-state operation. The evolution of the current spatial profile is related to the evolution of the poloidal magnetic flux, which can be modeled in the normalized cylindrical coordinates using a parabolic partial differential equation (PDE) called the magnetic diffusion equation. In this paper, we consider the dynamic optimization problem of attaining the best possible current spatial profile during the ramp-up phase of the tokamak. We first use the Galerkin method to obtain a finite-dimensional ordinary differential equation (ODE) model based on the original magnetic diffusion PDE. Then, we combine the control parameterization method with a novel time-scaling transformation to obtain an approximate optimal parameter selection problem, which can be solved using gradient-based optimization techniques such as sequential quadratic programming (SQP). This control parameterization approach involves approximating the tokamak input signals by piecewise-linear functions whose slopes and break-points are decision variables to be optimized. We show that the gradient of the objective function with respect to the decision variables can be computed by solving an auxiliary dynamic system governing the state sensitivity matrix. Finally, we conclude the paper with simulation results for an example problem based on experimental data from the DIII-D tokamak in San Diego, California.

Keywords: Nuclear Fusion, Current Profile Control, Control Parameterization, Time-scaling Transformation

1. Introduction

Fusion is a nuclear reaction in which multiple atomic nuclei collide under very high energy and join together to form a combined atomic nucleus that is lighter than the total mass of the reactants, with the excess mass converted into energy according to Einstein's mass-energy theory. Nuclear fusion reactions occur naturally in the core of the Sun and generate a continuous supply of energy for the universe. Nuclear fusion was first accomplished in the laboratory in the 1950s, and since then has shown considerable promise as a safe, clean and potentially inexhaustible energy production method. As such, it could become the best compromise between nature and the energy needs of mankind.

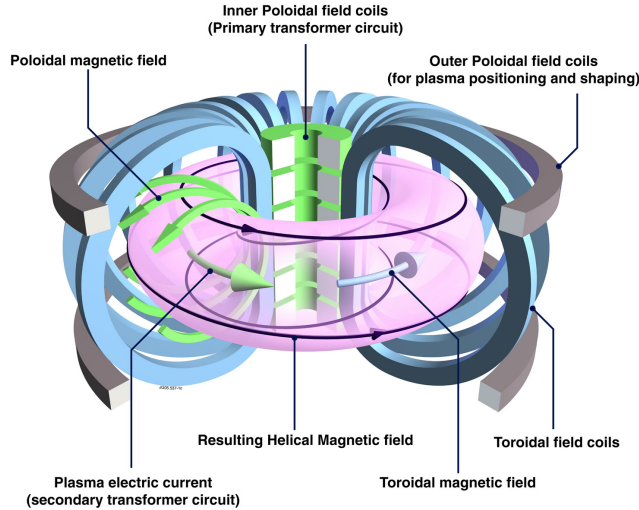[2]   Correspondence to: Chao Xu, Email: cxu@zju.edu.cn

Figure 1: Schematic of a tokamak chamber and magnetic profile (Source: EFDA-JET).

In the core of the Sun, gravitational forces can compress matter (mostly hydrogen) up to very high densities and temperatures, high enough to ignite fusion reactions. The Sun's gravitational field can maintain the thermonuclear reactions at a controlled and steady rate, by keeping the enormous thermal expansion forces balanced. However, fusion on Earth must be controlled by means other than gravity, as it is impossible to attain in the laboratory density levels near those in the center of the Sun. Hence, for nuclear fusion reactors, it is necessary to work at low gas densities, but at temperatures considerably higher than the temperature in the center of the Sun. At these extremely high temperatures, all matter is in the plasma state, which is considered the fourth state of mass. Fortunately, a torus-shaped device called a tokamak (shown in Figure 1) can be used to confine the plasma via a helical magnetic field.

Control engineering is considered one of the three critical technologies for achieving viable nuclear fusion power[3]. Accordingly, it has become an important area for multidisciplinary collaboration in the fusion research community. Many exciting research topics are surveyed in the book [2] and two special issues of IEEE Control Systems Magazine[4]. Among various challenging research issues, the control of the current profile in tokamak plasmas is known to be critical to plasma confinement, magnetohydrodynamic stability and effective steady-state operation (e.g., [17] and [22]).

The evolution in time of the current profile is related to the evolution of the poloidal magnetic flux, which is modeled by the magnetic diffusion equation, a parabolic partial differential equation (PDE) in the normalized cylindrical coordinate system. The problem of manipulating the current profile to achieve high performance while satisfying safety requirements has attracted considerable attention in the literature.

---

[3] Refer to E. Synakowski's presentation titled "Fusion Energy Research: On Our Science, Leverage and Credibility" at the University Fusion Association General Meeting, held during the 51st Annual Meeting of the American Physical Society Division of Plasma Physics (November 2-6, 2009, Atlanta, Georgia, USA).

[4] Refer to papers in the special issues titled "Control of Tokamak Plasmas: Part I" (October 2005) and "Control of Tokamak Plasmas: Part II" (April 2006) in IEEE Control Systems Magazine, organized by A. Pironti and M. Walker.
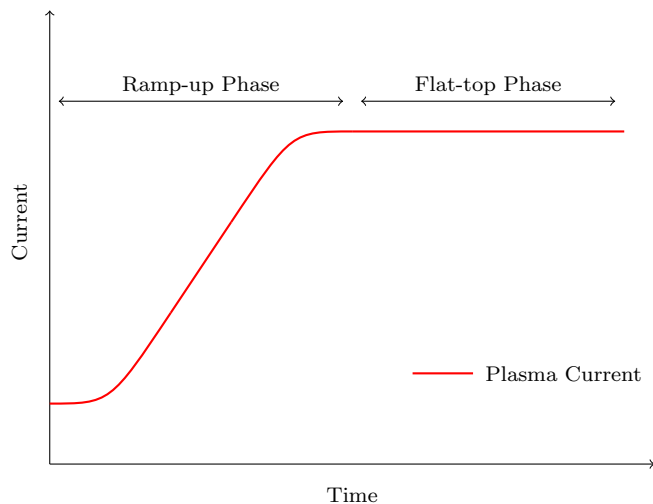
Figure 2: The plasma current evolution can be divided into two phases – ramp-up phase and flat-top phase.

The pioneering work in this area is due to Moreau, who in [15] derived empirical models for current profile evolutions using system identification techniques, and then used these models to synthesize a controller for safety factor profile manipulation. This approach, called the data-driven modeling method, does not rely on first principles to derive the model, but rather involves constructing the model by fitting observed data from experiments (e.g., JT-60 in Japan and DIII-D in the USA [16]), where the observed data is generated by input signals covering a sufficiently broad range of frequencies. As an alternative approach, it is also possible to use PDE models derived from first principles for the control and optimization of internal spatial profiles. This was a popular method in the 1970s and 1980s; one can refer, for example, to the books [4] and [26]. From the viewpoint of distributed parameter systems governed by PDEs, several advanced control and optimization approaches have been discussed based on new developments in internal diagnostics and plasma actuation. Simplified distributed parameter system models have been studied in [18] and [27] toward current profile control at DIII-D and Tore Supra, respectively. In order to achieve an optimal current evolution in the plasma discharges of the ramp-up phase, various dynamic optimization techniques are applied to obtain numerical solutions in an open loop fashion, such as [7], [8], [19], [28]. For online implementations of optimized actuation commands, feedback strategies (e.g.,[5, 6, 20, 29]) are necessary to attenuate external perturbations and system uncertainties. However, many challenging research problems still remain open.

The plasma transport is governed by a set of PDEs (magnetic flux, temperature and density profiles of both electrons and ions) that are nonlinearly coupled (with coupling strength depending on plasma scenarios). In the papers by Witrant [27] and Ou [18], the mathematical formulation for the distributed parameter system framework is discussed, for both the Tore Supra and the DIII-D experimental toka-maks. In these papers, only the magnetic flux PDE is considered and scaling laws are used to account (in an approximate fashion) for the effects of the other PDEs. In the present work, using the framework proposed by Witrant [27] and Ou [18], we consider the problem of obtaining a specific desired output current profile at the end of the current ramp-up phase, as shown in Figure 2. This problem can be

3

formulated as a finite-time dynamic optimization problem for the magnetic flux diffusion PDE system, which is a type of a bilinear infinite-dimensional system, i.e., the diffusivity control term appears in the second-order elliptic operator. We first discretize the PDE model over the state space using the finite element method to yield a finite-dimensional system of ODEs. The motivation for performing this discretization procedure is that, in general, it is much easier to solve the ODE-constrained optimization problem rather than the original PDE-constrained optimization problem. After using the finite element method to obtain a finite-dimensional system of ODEs, we then apply the control parameterization approach to obtain an optimal parameter selection problem [23]. This involves approximating the control function by a linear combination of temporal basis functions (usually simple characteristic functions, polynomials or splines) with the constant coefficients to be determined by numerical optimization procedures such as sequential quadratic programming (SQP). This methodology follows the *discretize-then-optimize* approach [24] rather than the *optimize-then-discretize* approach. More specifically, discretization techniques are first applied to derive an approximate optimization problem, which is then solved using various optimization techniques. The numerical convergence of suboptimal solutions generated by the discretize-then-optimize approach remains a challenging issue, much more difficult than the convergence analysis for the numerical discretization of PDEs [11, 14].

This paper is organized as follows. The formulation of the PDE optimization problem for the current profile system is introduced in Section 2. In Section 3, we use the Galerkin method to obtain a finite-dimensional model derived from the original PDE system. In Section 4, we use the control parameterization technique and the time-scaling transformation to approximate the control input variables by linear combinations of basis functions, thereby obtaining an approximate problem that can be solved using numerical optimization methods such as SQP. In Section 5, numerical results for a plasma discharge from the DIII-D tokamak (under shot #129412) are presented. Finally, in Section 6, we conclude the paper by summarizing our results and suggesting topics for further research.

2. Problem Formulation

The dynamic behavior of the magnetic-flux profile $\psi(\hat{\rho}, t)$ is described by the following parabolic PDE [28]:

$$\frac{1}{\vartheta_1(\hat{\rho})} \frac{\partial \psi(\hat{\rho}, t)}{\partial t} = \frac{u_1(t)}{\hat{\rho}} \frac{\partial}{\partial \hat{\rho}} \left[\hat{\rho} D(\hat{\rho}) \frac{\partial \psi(\hat{\rho}, t)}{\partial \hat{\rho}} \right] + \vartheta_2(\hat{\rho}) u_2(t), \tag{1}$$

with the Neumann boundary conditions

$$\frac{\partial \psi(0, t)}{\partial \hat{\rho}} = 0, \quad \frac{\partial \psi(1, t)}{\partial \hat{\rho}} = u_3(t), \tag{2}$$

where t denotes time; $\hat{\rho}$ denotes the normalized radius; and $\psi(\hat{\rho}, t)$ denotes the poloidal magnetic flux around the tokamak. Moreover, $\vartheta_1(\hat{\rho})$, $\vartheta_2(\hat{\rho})$ and $D(\hat{\rho})$ are given functions of $\hat{\rho}$ that depend on the particular tokamak device under consideration and can be identified offline using experimental data. The auxiliary functions $u_1(t)$, $u_2(t)$ and $u_3(t)$ depend on the total power $P(t)$, the total plasma current

$I(t)$, and the average density $\bar{n}(t)$ according to the following equations:

$$u_1(t) = \bar{n}(t)^{\frac{3}{2}} I(t)^{-\frac{3}{2}} P(t)^{-\frac{3}{4}}, \tag{3a}$$

$$u_2(t) = I(t)^{-1} P(t)^{\frac{1}{2}}, \tag{3b}$$

$$u_3(t) = \kappa I(t), \tag{3c}$$

where κ is a given constant. Note that $\bar{n}(t)$, $I(t)$, and $P(t)$ are the open-loop input signals for the physical actuators in the tokamak. Furthermore, the initial condition for the magnetic flux profile is given by

$$\psi(\hat{\rho}, 0) = \psi_0(\hat{\rho}). \tag{4}$$

For notational simplicity, we define

$$\psi_t(\hat{\rho}, t) = \frac{\partial \psi(\hat{\rho}, t)}{\partial t}, \quad \psi_{\hat{\rho}}(\hat{\rho}, t) = \frac{\partial \psi(\hat{\rho}, t)}{\partial \hat{\rho}}.$$

Then, the PDE system defined by equation (1) with boundary conditions (2) and initial condition (4) can be written as

$$\frac{\psi_t(\hat{\rho}, t)}{\vartheta_1(\hat{\rho})} = \frac{u_1(t)}{\hat{\rho}} \left[\hat{\rho} D(\hat{\rho}) \psi_{\hat{\rho}}(\hat{\rho}, t) \right]_{\hat{\rho}} + \vartheta_2(\hat{\rho}) u_2(t), \tag{5a}$$

$$\psi_{\hat{\rho}}(0, t) = 0, \quad \psi_{\hat{\rho}}(1, t) = u_3(t), \quad t \in [0, T], \tag{5b}$$

$$\psi(\hat{\rho}, 0) = \psi_0(\hat{\rho}), \quad \hat{\rho} \in [0, 1]. \tag{5c}$$

In [21], we proposed a numerical optimization approach that involves solving (5) using the finite element method. However, this approach sometimes struggles to handle the Neumann boundary conditions (5b). Thus, in this paper, we introduce a new variable $\omega(\hat{\rho}, t)$ defined by

$$\omega(\hat{\rho}, t) = \hat{\rho} D(\hat{\rho}) \frac{\partial \psi(\hat{\rho}, t)}{\partial \hat{\rho}}.$$

Then the PDE system (5) becomes

$$\frac{1}{\hat{\rho} D(\hat{\rho})} \omega_t(\hat{\rho}, t) = u_1(t) \left[\frac{\vartheta_1(\hat{\rho})}{\hat{\rho}} \omega_{\hat{\rho}}(\hat{\rho}, t) \right]_{\hat{\rho}} + u_2(t) \left[\vartheta_1(\hat{\rho}) \vartheta_2(\hat{\rho}) \right]_{\hat{\rho}}, \tag{6a}$$

$$\omega(0, t) = 0, \quad \omega(1, t) = D(1) u_3(t), \tag{6b}$$

$$\omega(\hat{\rho}, 0) = \hat{\rho} D(\hat{\rho}) \frac{\partial \psi_0(\hat{\rho})}{\partial \hat{\rho}}, \tag{6c}$$

where

$$\omega_t(\hat{\rho}, t) = \frac{\partial \omega(\hat{\rho}, t)}{\partial t}, \quad \omega_{\hat{\rho}}(\hat{\rho}, t) = \frac{\partial \omega(\hat{\rho}, t)}{\partial \hat{\rho}}.$$

In this new system, the boundary conditions (6b) are Dirichlet conditions rather than Neumann conditions.

Our aim is to choose the open-loop input signals $\bar{n}(t)$, $I(t)$ and $P(t)$ judiciously so that the output profile $\omega(\hat{\rho}, t)$ at the terminal time $t = T$ is as close as possible to a given target profile $\omega_d(\hat{\rho})$, which is related to the desired output poloidal flux derivative. Thus, the problem is to minimize the following objective functional:

$$J = \frac{1}{2} \int_0^1 [\omega(\hat{\rho}, T) - \omega_d(\hat{\rho})]^2 d\hat{\rho}. \tag{7}$$

The open-loop actuator input signals must satisfy the following bound constraints:

$$\bar{n}_{\min} \leq \bar{n}(t) \leq \bar{n}_{\max}, \quad I_{\min} \leq I(t) \leq I_{\max}, \quad P_{\min} \leq P(t) \leq P_{\max}, \tag{8}$$

where the lower and upper bounds $\bar{n}_{\min}$, $\bar{n}_{\max}$, $I_{\min}$, $I_{\max}$, $P_{\min}$, $P_{\max}$ are given constants. Furthermore, the actuator input signals must also satisfy the initial conditions

$$\bar{n}(0) = \bar{n}_0, \quad I(0) = I_0, \quad P(0) = P_0, \tag{9}$$

and the terminal conditions

$$\bar{n}_{\min}^T \leq \bar{n}(T) \leq \bar{n}_{\max}^T, \quad I_{\min}^T \leq I(T) \leq I_{\max}^T, \quad P_{\min}^T \leq P(T) \leq P_{\max}^T, \tag{10}$$

where the initial values $\bar{n}_0$, I_0, P_0 and the terminal bounds $\bar{n}_{\min}^T$, $\bar{n}_{\max}^T$, $I_{\min}^T$, $I_{\max}^T$, $P_{\min}^T$, $P_{\max}^T$ are given. Any vector-valued continuous function $\boldsymbol{\theta} = (\bar{n}, I, P) : [0, T] \to \mathbb{R}^3$ that satisfies constraints (8)-(10) is called an admissible control. Let Θ denote the class of all such admissible controls.

We now state our dynamic optimization problem formally as follows.

Problem P$_0$. *Given the PDE system (6), find an admissible control $\boldsymbol{\theta} = (\bar{n}, I, P) \in \Theta$ such that the cost functional (7) is minimized.*

3. Finite Element Approximation

In this section, we apply the Galerkin finite element scheme [25] to approximate the PDE model (6) by a finite-dimensional ODE system [1, 24]. This yields a new dynamic optimization problem governed by ODEs, for which an efficient computational technique will be developed in Section 4.

Let $\eta(\hat{\rho})$ be a trial function. Multiplying both sides of (6a) by $\hat{\rho}\eta(\hat{\rho})$ and then integrating the resulting equation over $[0, 1]$ gives

$$\int_0^1 \frac{\eta(\hat{\rho})}{D(\hat{\rho})} \omega_t(\hat{\rho}, t) d\hat{\rho} = \int_0^1 \hat{\rho}\eta(\hat{\rho}) \left[\frac{\vartheta_1(\hat{\rho})}{\hat{\rho}} u_1(t) \omega_{\hat{\rho}}(\hat{\rho}, t) \right]_{\hat{\rho}} d\hat{\rho} + \int_0^1 \hat{\rho}\eta(\hat{\rho}) \left[\vartheta_1(\hat{\rho})\vartheta_2(\hat{\rho}) \right]_{\hat{\rho}} u_2(t) d\hat{\rho}. \tag{11}$$

Using integration by parts to simplify the first integral on the right-hand side of (11), we obtain

$$\int_0^1 \frac{\eta(\hat{\rho})}{D(\hat{\rho})}\omega_t(\hat{\rho},t)d\hat{\rho} = \eta(1)\vartheta_1(1)u_1(t)\omega_{\hat{\rho}}(1,t) - \eta(0)\vartheta_1(0)u_1(t)\omega_{\hat{\rho}}(0,t)$$

$$- u_1(t)\int_0^1 \frac{\vartheta_1(\hat{\rho})}{\hat{\rho}}\omega_{\hat{\rho}}(\hat{\rho},t)[\eta(\hat{\rho}) + \hat{\rho}\eta'(\hat{\rho})]d\hat{\rho} \tag{12}$$

$$+ u_2(t)\int_0^1 \hat{\rho}\eta(\hat{\rho})\left[\vartheta_1(\hat{\rho})\vartheta_2(\hat{\rho})\right]_{\hat{\rho}} d\hat{\rho}.$$

Now, consider the partition $[(i-1)/N, i/N]$, $i = 1, 2, \ldots, N$, of the spatial domain [0,1]. We assume that the magnetic flux profile $\omega(\hat{\rho}, t)$ can be approximated by a linear combination of basis B-spline functions $\beta_i(\hat{\rho})$, $i = 1, 2, \ldots, N$, where $\beta_i(\hat{\rho})$ corresponds to the ith partition point. Mathematically,

$$\omega(\hat{\rho},t) \approx \omega^N(\hat{\rho},t) = \sum_{i=1}^N x_i(t)\beta_i(\hat{\rho}), \tag{13}$$

where $x_i(t)$, $i = 1, 2, \ldots, N$, are weighting functions. We impose the conditions $\beta_i(0) = 0$, $i = 1, 2, \ldots, N$, to ensure that the approximation (13) satisfies the left boundary condition in (6b). For the right boundary condition in (6b), we apply the well-known penalization method [3], which involves replacing

$$\omega(1,t) = D(1)u_3(t),$$

with

$$\omega(1,t) + \epsilon\omega_{\hat{\rho}}(1,t) = D(1)u_3(t), \tag{14}$$

where ϵ is a small perturbation coefficient. By substituting (14) into (12), we obtain

$$\int_0^1 \frac{\eta(\hat{\rho})}{D(\hat{\rho})}\omega_t(\hat{\rho},t)d\hat{\rho} = \frac{1}{\epsilon}\eta(1)\vartheta_1(1)u_1(t)[D(1)u_3(t) - \omega(1,t)] - \eta(0)\vartheta_1(0)u_1(t)\omega_{\hat{\rho}}(0,t)$$

$$- u_1(t)\int_0^1 \frac{\vartheta_1(\hat{\rho})}{\hat{\rho}}\omega_{\hat{\rho}}(\hat{\rho},t)[\eta(\hat{\rho}) + \hat{\rho}\eta'(\hat{\rho})]d\hat{\rho} \tag{15}$$

$$+ u_2(t)\int_0^1 \hat{\rho}\eta(\hat{\rho})\left[\vartheta_1(\hat{\rho})\vartheta_2(\hat{\rho})\right]_{\hat{\rho}} d\hat{\rho}.$$

Hence, by choosing $\eta(\hat{\rho}) = \beta_j(\hat{\rho})$, $j = 1, 2, \ldots, N$, as the trial functions and applying approximation (13), equation (15) becomes

$$\sum_{i=1}^N \left[\int_0^1 \frac{1}{D(\hat{\rho})}\beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho}\right]\dot{x}_i(t) = \frac{1}{\epsilon}\beta_j(1)\vartheta_1(1)u_1(t)\left[D(1)u_3(t) - \sum_{i=1}^N x_i(t)\beta_i(1)\right]$$

$$- u_1(t)\sum_{i=1}^N\left[\int_0^1 \frac{\vartheta_1(\hat{\rho})}{\hat{\rho}}\beta_i'(\hat{\rho})[\beta_j(\hat{\rho}) + \hat{\rho}\beta_j'(\hat{\rho})]d\hat{\rho}\right]x_i(t) + u_2(t)\int_0^1 \hat{\rho}\beta_j(\hat{\rho})\left[\vartheta_1(\hat{\rho})\vartheta_2(\hat{\rho})\right]_{\hat{\rho}} d\hat{\rho}. \tag{16}$$

We introduce the following notation:

$$A_{ij} = \int_0^1 \frac{1}{D(\hat{\rho})}\beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho},$$

$$B_{ij} = \int_0^1 \frac{\vartheta_1(\hat{\rho})}{\hat{\rho}}\beta_i(\hat{\rho})\beta_j'(\hat{\rho})d\hat{\rho},$$

$$C_{ij} = \int_0^1 \vartheta_1(\hat{\rho})\beta_i'(\hat{\rho})\beta_j'(\hat{\rho})d\hat{\rho},$$

$$d_j = \int_0^1 \hat{\rho}\beta_j(\hat{\rho})\left[\vartheta_1(\hat{\rho})\vartheta_2(\hat{\rho})\right]_{\hat{\rho}} d\hat{\rho},$$

$$e_j = \frac{1}{\epsilon}\beta_j(1)\vartheta_1(1)D(1),$$

$$F_{ij} = \frac{1}{\epsilon}\vartheta_1(1)\beta_i(1)\beta_j(1),$$

where $i, j = 1, \ldots, N$. Then equation (16) can be rewritten in matrix form as follows:

$$\boldsymbol{A}\dot{\boldsymbol{x}}(t) = -(\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{F})u_1(t)\boldsymbol{x}(t) + u_2(t)\boldsymbol{d} + u_1(t)u_3(t)\boldsymbol{e},$$

where

$$\boldsymbol{x}(t) = [x_1(t), \ldots, x_N(t)]^\top,$$

$$\boldsymbol{A} = [A_{ij}], \quad \boldsymbol{B} = [B_{ij}], \quad \boldsymbol{C} = [C_{ij}], \quad \boldsymbol{F} = [F_{ij}],$$

$$\boldsymbol{d} = [d_1, \ldots, d_N]^\top, \quad \boldsymbol{e} = [e_1, \ldots, e_N]^\top.$$

Recalling the initial condition (6c), we must have

$$\omega(\hat{\rho}, 0) = \hat{\rho}D(\hat{\rho})\frac{\partial\psi_0(\hat{\rho})}{\partial\hat{\rho}} = \sum_{i=1}^N x_i(0)\beta_i(\hat{\rho}). \tag{17}$$

We multiply both sides of (17) by $\beta_j(\hat{\rho})$ and then integrate over $\hat{\rho} \in [0,1]$ to obtain

$$\sum_{i=1}^N \left[\int_0^1 \beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho}\right] x_i(0) = \int_0^1 \hat{\rho}D(\hat{\rho})\beta_j(\hat{\rho})\frac{\partial\psi_0(\hat{\rho})}{\partial\hat{\rho}}d\hat{\rho}. \tag{18}$$

Define

$$\bar{A}_{ij} = \int_0^1 \beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho},$$

$$\bar{b}_j = \int_0^1 \hat{\rho}D(\hat{\rho})\beta_j(\hat{\rho})\frac{\partial\psi_0(\hat{\rho})}{\partial\hat{\rho}}d\hat{\rho},$$

where $i, j = 1, 2, \ldots, N$. Then equation (18) can be written as $\bar{\boldsymbol{A}}\boldsymbol{x}(0) = \bar{\boldsymbol{b}}$, where $\bar{\boldsymbol{A}} = [\bar{A}_{ij}]$ and $\bar{\boldsymbol{b}} = [\bar{b}_0, \bar{b}_1, \ldots, \bar{b}_N]^\top$. By following the same arguments as in [10], it can be shown that matrices $\boldsymbol{A}$ and $\bar{\boldsymbol{A}}$ are nonsingular. Consequently, under approximation (13), the PDE system (6) reduces to the following finite-dimensional ODE system:

$$\dot{\boldsymbol{x}}(t) = -\boldsymbol{A}^{-1}(\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{F})u_1(t)\boldsymbol{x}(t) + u_2(t)\boldsymbol{A}^{-1}\boldsymbol{d} + u_1(t)u_3(t)\boldsymbol{A}^{-1}\boldsymbol{e}, \tag{19a}$$

$$\boldsymbol{x}(0) = \bar{\boldsymbol{A}}^{-1}\bar{\boldsymbol{b}}. \tag{19b}$$

Substituting (3) into (19), we obtain

$$\dot{\boldsymbol{x}}(t) = -\bar{n}(t)^{\frac{3}{2}} I(t)^{-\frac{3}{2}} P(t)^{-\frac{3}{4}} \boldsymbol{A}^{-1} (\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{F}) \boldsymbol{x}(t)$$

$$+ I(t)^{-1} P(t)^{\frac{1}{2}} \boldsymbol{A}^{-1} \boldsymbol{d} + \kappa \bar{n}(t)^{\frac{3}{2}} I(t)^{-\frac{1}{2}} P(t)^{-\frac{3}{4}} \boldsymbol{A}^{-1} \boldsymbol{e}, \tag{20a}$$

$$\boldsymbol{x}(0) = \bar{\boldsymbol{A}}^{-1} \bar{\boldsymbol{b}}. \tag{20b}$$

Now, based on (13), consider the following expansion for the desired output profile $\omega_d(\hat{\rho})$:

$$\omega_d(\hat{\rho}) = \sum_{i=1}^{N} x_i^d \beta_i(\hat{\rho}), \tag{21}$$

where x_i^d, $i = 1, 2, \ldots, N$, are weighting coefficients. To obtain the values of x_i^d, we multiply both sides of (21) by $\beta_j(\hat{\rho})$ and then integrate over $\hat{\rho} \in [0, 1]$ to obtain

$$\int_0^1 \omega_d(\hat{\rho}) \beta_j(\hat{\rho}) d\hat{\rho} = \sum_{i=1}^{N} \left[\int_0^1 \beta_i(\hat{\rho}) \beta_j(\hat{\rho}) d\hat{\rho} \right] x_i^d. \tag{22}$$

Define

$$\hat{b}_j = \int_0^1 \omega_d(\hat{\rho}) \beta_j(\hat{\rho}) d\hat{\rho},$$

where $j = 1, 2, \ldots, N$. Then we can rewrite (22) as

$$\hat{\boldsymbol{b}} = \bar{\boldsymbol{A}} \boldsymbol{x}^d,$$

where $\hat{\boldsymbol{b}} = [\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_N]^\top$, and $\boldsymbol{x}^d = [x_1^d, x_2^d, \ldots, x_N^d]^\top$. Therefore, the weighting coefficients for the desired output profile are given by

$$\boldsymbol{x}^d = \bar{\boldsymbol{A}}^{-1} \hat{\boldsymbol{b}}. \tag{23}$$

Using the expansion (13), the actual output profile $\omega(\hat{\rho}, T)$ at the terminal time T is approximated as follows:

$$\omega(\hat{\rho}, T) = \sum_{i=1}^{N} x_i(T) \beta_i(\hat{\rho}). \tag{24}$$

Substituting (21) and (24) into the cost functional (7), we obtain

$$J(\bar{n}, I, P) = \frac{1}{2} \int_0^1 \left[\sum_{i=1}^{N} (x_i(T) - x_i^d) \beta_i(\hat{\rho}) \right]^2 d\hat{\rho}$$

$$= \frac{1}{2} \int_0^1 \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i(T) - x_i^d) \beta_i(\hat{\rho}) \beta_j(\hat{\rho}) (x_j(T) - x_j^d) d\hat{\rho}$$

$$= \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i(T) - x_i^d) \left[\int_0^1 \beta_i(\hat{\rho}) \beta_j(\hat{\rho}) d\hat{\rho} \right] (x_j(T) - x_j^d).$$

9

Hence, the cost functional (7) can be rewritten as

$$J(\bar{n}, I, P) = \frac{1}{2}[\boldsymbol{x}(T) - \boldsymbol{x}^d]^\top \bar{\boldsymbol{A}}[\boldsymbol{x}(T) - \boldsymbol{x}^d], \tag{25}$$

where $\boldsymbol{x}^d = \bar{\boldsymbol{A}}^{-1}\hat{\boldsymbol{b}}$ as in (23). Problem P_0, the original PDE dynamic optimization problem, is now approximated by the following ODE dynamic optimization problem, which we call Problem P_N.

Problem P_N. *Given the ODE system (20), find an admissible control $\boldsymbol{\theta} = (\bar{n}, I, P) \in \Theta$ such that the cost functional (25) is minimized.*

4. Numerical Solution Procedure

4.1. Piecewise-Linear Control Parameterization

To solve Problem P_N, we will apply the control parameterization method [9, 23], whereby the tokamak open-loop input signals $\bar{n}(t)$, $I(t)$, $P(t)$ are approximated by piecewise-linear functions. More specifically, we first subdivide the time horizon $[0, T]$ into p subintervals $[\gamma_{k-1}, \gamma_k)$, $k = 1, 2, \ldots, p$, where $\gamma_0 = 0$ and $\gamma_p = T$, and the interior knot points γ_k, $k = 1, 2, \ldots, p-1$, are free decision parameters. We impose the following constraints:

$$\tau_{\min} \leq \gamma_k - \gamma_{k-1} \leq \tau_{\max}, \quad k = 1, 2, \ldots, p. \tag{26}$$

Here, $\tau_{\min} > 0$ and $\tau_{\max} > 0$ are the minimum and maximum subinterval durations, respectively. We approximate the derivative of the vector-valued control function $\boldsymbol{\theta} = (\bar{n}, I, P)$ as follows:

$$\dot{\boldsymbol{\theta}}(t) \approx \boldsymbol{\sigma}^k, \quad t \in [\gamma_{k-1}, \gamma_k), \quad k = 1, 2, \ldots, p, \tag{27}$$

where $\boldsymbol{\sigma}^k = [\sigma_1^k, \sigma_2^k, \sigma_3^k]^\top$. Here, σ_1^k is an approximation of the derivative of $\bar{n}(t)$ on $[\gamma_{k-1}, \gamma_k)$, σ_2^k is an approximation of the derivative of $I(t)$ on $[\gamma_{k-1}, \gamma_k)$, and σ_3^k is an approximation of the derivative of $P(t)$ on $[\gamma_{k-1}, \gamma_k)$. We can express (27) as follows:

$$\dot{\boldsymbol{\theta}}(t) \approx \sum_{k=1}^{p} \boldsymbol{\sigma}^k \chi_{[\gamma_{k-1}, \gamma_k)}(t), \quad t \in [0, T], \tag{28}$$

where $\chi_{[\gamma_{k-1}, \gamma_k)} : \mathbb{R} \to \mathbb{R}$ is the indicator function defined by

$$\chi_{[\gamma_{k-1}, \gamma_k)}(t) = \begin{cases} 1, & \text{if } t \in [\gamma_{k-1}, \gamma_k), \\ 0, & \text{otherwise.} \end{cases}$$

According to (28), $\dot{\boldsymbol{\theta}}(t)$ is approximated by a piecewise-constant function with discontinuities at the internal knot points $\gamma_1, \gamma_2, \ldots, \gamma_{p-1}$. Thus, $\boldsymbol{\theta}(t)$ is piecewise-linear with jumps in the derivative at $\gamma_1, \gamma_2, \ldots, \gamma_{p-1}$. We introduce new state variables $\upsilon_{N+1}(t) = \bar{n}(t)$, $\upsilon_{N+2}(t) = I(t)$ and $\upsilon_{N+3}(t) = P(t)$

governed by the following dynamics (for $i = 1, 2, 3$):

$$\dot{v}_{N+i}(t) = \sum_{k=1}^{p} \sigma_i^k \chi_{[\gamma_{k-1}, \gamma_k)}(t), \quad t \in [0, T], \tag{29a}$$

$$v_{N+i}(0) = v_{N+i}^0, \tag{29b}$$

where $v_{N+1}^0 = \bar{n}_0$ is the initial value of $\bar{n}$, $v_{N+2}^0 = I_0$ is the initial value of I, and $v_{N+3}^0 = P_0$ is the initial value of P (recall the initial conditions (9)). Let $v_{N+1}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\gamma})$, $v_{N+2}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\gamma})$, $v_{N+3}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\gamma})$ denote the solution of system (29), where

$$\boldsymbol{\sigma} = [(\boldsymbol{\sigma}^1)^\top, (\boldsymbol{\sigma}^2)^\top, \ldots, (\boldsymbol{\sigma}^p)^\top]^\top,$$

$$\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_{p-1}]^\top.$$

Then, the dynamic system (20) becomes

$$\dot{\boldsymbol{x}}(t) = -v_{N+1}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{\frac{3}{2}} v_{N+2}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{-\frac{3}{2}} v_{N+3}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{F})\boldsymbol{x}(t)$$

$$+ v_{N+2}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{-1} v_{N+3}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{\frac{1}{2}} \boldsymbol{A}^{-1}\boldsymbol{d}$$

$$+ \kappa v_{N+1}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{\frac{3}{2}} v_{N+2}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{-\frac{1}{2}} v_{N+3}(t|\boldsymbol{\sigma}, \boldsymbol{\gamma})^{-\frac{3}{4}} \boldsymbol{A}^{-1}\boldsymbol{e}, \tag{30a}$$

$$\boldsymbol{x}(0) = \bar{\boldsymbol{A}}^{-1}\bar{\boldsymbol{b}}. \tag{30b}$$

Let $\boldsymbol{x}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\gamma})$ denote the solution of system (30). To determine $\boldsymbol{x}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\gamma})$, we can solve (29) and (30) sequentially over the subintervals $[\gamma_{k-1}, \gamma_k)$, $k = 1, 2, \ldots, p$.

Recall that the open-loop actuator input signals $\bar{n}(t)$, $I(t)$, and $P(t)$ must satisfy the bound constraints (8). Thus, we impose the following continuous state inequality constraints on the new state variables v_{N+1}, v_{N+2} and v_{N+3}:

$$\bar{n}_{\min} \leq v_{N+1}^p(t|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \bar{n}_{\max}, \quad t \in [0, T], \tag{31a}$$

$$I_{\min} \leq v_{N+2}^p(t|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq I_{\max}, \quad t \in [0, T], \tag{31b}$$

$$P_{\min} \leq v_{N+3}^p(t|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq P_{\max}, \quad t \in [0, T]. \tag{31c}$$

We also impose the following constraints corresponding to (10):

$$\bar{n}_{\min}^T \leq v_{N+1}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \bar{n}_{\max}^T, \tag{32a}$$

$$I_{\min}^T \leq v_{N+2}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq I_{\max}^T, \tag{32b}$$

$$P_{\min}^T \leq v_{N+3}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq P_{\max}^T. \tag{32c}$$

Clearly, since $v_{N+1}(t)$, $v_{N+2}(t)$ and $v_{N+3}(t)$ are piecewise-linear with break-points at $t = \gamma_1, \gamma_2, \ldots, \gamma_{p-1}$,

constraints (31) and (32) can be combined to yield

$$\bar{n}_{\min} \leq v_{N+1}^p(\gamma_k|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \bar{n}_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{33a}$$

$$I_{\min} \leq v_{N+2}^p(\gamma_k|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq I_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{33b}$$

$$P_{\min} \leq v_{N+3}^p(\gamma_k|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq P_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{33c}$$

$$\max(\bar{n}_{\min}, \bar{n}_{\min}^T) \leq v_{N+1}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \min(\bar{n}_{\max}, \bar{n}_{\max}^T), \tag{33d}$$

$$\max(I_{\min}, I_{\min}^T) \leq v_{N+2}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \min(I_{\max}, I_{\max}^T), \tag{33e}$$

$$\max(P_{\min}, P_{\min}^T) \leq v_{N+3}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) \leq \min(P_{\max}, P_{\max}^T). \tag{33f}$$

These constraints are special cases of the well-known canonical form in the optimal control literature (see [9]).

Now, under the approximation (28), the cost functional (25) becomes

$$J^p(\boldsymbol{\sigma}, \boldsymbol{\gamma}) = \frac{1}{2}[\boldsymbol{x}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) - \boldsymbol{x}^d]^\top \bar{\boldsymbol{A}}[\boldsymbol{x}^p(T|\boldsymbol{\sigma}, \boldsymbol{\gamma}) - \boldsymbol{x}^d]. \tag{34}$$

We now state the approximate optimal parameter selection problem as follows.

Problem P_N^p. *Given the ODE system (29)-(30), find a control parameter vector $\boldsymbol{\sigma}$ and a switching time vector $\boldsymbol{\gamma}$ such that the cost functional (34) is minimized subject to constraints (26) and (33).*

Convergence results showing that the optimal cost of Problem P_N^p converges to the optimal cost of Problem P_N as $p \to \infty$ are given in [11, 14].

4.2. Time-scaling Transformation

In the standard control parameterization method, $\tau_{\min} = \tau_{\max}$, i.e., the subinterval durations (and therefore the control switching points) are fixed. For more accurate results, the control switching points should be variable along with the control parameters [9]. However, it is well known that variable switching points cause computational difficulties [12]. To overcome these difficulties, we will employ the so-called *time-scaling transformation* [9, 12] to map the variable switching points to fixed points in a new time horizon. This yields a new optimization problem in which the switching times are fixed.

To apply the time-scaling transformation, we first introduce a new time variable $s \in [0, p]$. Then, we relate the new time variable s to the old time variable t through the following equations:

$$\frac{dt(s)}{ds} = \zeta_k, \quad s \in [k-1, k), \quad k = 1, 2, \ldots, p, \tag{35}$$

$$t(0) = 0, \tag{36}$$

$$t(p) = T, \tag{37}$$

where $\zeta_k = \gamma_k - \gamma_{k-1}$, $k = 1, 2, \ldots, p$, are new decision parameters. Note that equations (35)-(37) define $t = t(s)$ as a non-negative and non-decreasing piecewise-linear function of s.

Integrating (35)-(37) gives

$$t(s) = \int_0^s \frac{dt(\eta)}{d\eta} d\eta = \zeta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor) + \sum_{k=1}^{\lfloor s \rfloor} \zeta_k, \quad s \in [0, p].$$

Hence,

$$t(k) = \sum_{l=1}^k \zeta_l = \sum_{l=1}^k (\gamma_l - \gamma_{l-1}) = \gamma_k, \quad k = 0, 1, \ldots, p. \tag{38}$$

This shows that $s = k$ in the new time horizon corresponds to $t = t_k$ in the original time horizon.

Let $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \ldots, \zeta_p]^\top$. Clearly, the vector $\boldsymbol{\zeta} \in \mathbb{R}^p$ must satisfy the following constraints:

$$\tau_{\min} \leq \zeta_k \leq \tau_{\max}, \quad k = 1, 2, \ldots, p, \tag{39a}$$

$$\zeta_1 + \zeta_2 + \cdots + \zeta_p = T. \tag{39b}$$

Let $\tilde{\boldsymbol{x}}(s)$ denote $\boldsymbol{x}(t(s))$. Then under the time-scaling transformation (35)-(37), the dynamics (30) become

$$\dot{\tilde{\boldsymbol{x}}}(s) = -\zeta_k \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{F})\tilde{\boldsymbol{x}}(s) + \zeta_k \tilde{v}_{N+2}(s)^{-1} \tilde{v}_{N+3}(s)^{\frac{1}{2}} \boldsymbol{A}^{-1}\boldsymbol{d}$$

$$+ \kappa \zeta_k \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{1}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}\boldsymbol{e}, \quad s \in [k-1, k), \quad k = 1, 2, \ldots, p, \tag{40a}$$

$$\tilde{\boldsymbol{x}}(0) = \bar{\boldsymbol{A}}^{-1}\bar{\boldsymbol{b}}. \tag{40b}$$

Moreover, the dynamics (29) become (for $i = 1, 2, 3$):

$$\dot{\tilde{v}}_{N+i}(s) = \zeta_k \sigma_i^k, \quad s \in [k-1, k), \quad k = 1, 2, \ldots, p, \tag{41a}$$

$$\tilde{v}_{N+i}(0) = v_{N+i}^0. \tag{41b}$$

In view of (38), the canonical constraints (33) become

$$\bar{n}_{\min} \leq \tilde{v}_{N+1}(k) \leq \bar{n}_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{42a}$$

$$I_{\min} \leq \tilde{v}_{N+2}(k) \leq I_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{42b}$$

$$P_{\min} \leq \tilde{v}_{N+3}(k) \leq P_{\max}, \quad k = 0, 1, \ldots, p-1, \tag{42c}$$

$$\max(\bar{n}_{\min}, \bar{n}_{\min}^T) \leq \tilde{v}_{N+1}(p) \leq \min(\bar{n}_{\max}, \bar{n}_{\max}^T), \tag{42d}$$

$$\max(I_{\min}, I_{\min}^T) \leq \tilde{v}_{N+2}(p) \leq \min(I_{\max}, I_{\max}^T), \tag{42e}$$

$$\max(P_{\min}, P_{\min}^T) \leq \tilde{v}_{N+3}(p) \leq \min(P_{\max}, P_{\max}^T). \tag{42f}$$

Let $\tilde{\boldsymbol{x}}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$ and $\tilde{v}_{N+1}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$, $\tilde{v}_{N+2}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$, $\tilde{v}_{N+3}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$ denote the solutions of systems (40) and (41) corresponding to $\boldsymbol{\sigma} \in \mathbb{R}^{pr}$ and $\boldsymbol{\zeta} \in \mathbb{R}^p$.

Under the time-scaling transformation defined by (35)-(37), the cost functional (34) becomes

$$\tilde{J}^p(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = \frac{1}{2}[\tilde{\boldsymbol{x}}^p(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}) - \boldsymbol{x}^d]^\top \bar{\boldsymbol{A}}[\tilde{\boldsymbol{x}}^p(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}) - \boldsymbol{x}^d]. \tag{43}$$

The following optimal parameter selection problem is equivalent to Problem P_N^p.

Problem Q_N^p. *Given the ODE system (40)-(41), find a control parameter vector $\boldsymbol{\sigma}$ and a duration vector $\boldsymbol{\zeta}$ such that the cost functional (43) is minimized subject to the constraints (39) and (42).*

4.3. Solving Problem Q_N^p

Problem Q_N^p is an optimal parameter selection problem in the so-called canonical form [23]. In principle, such problems can be solved as nonlinear optimization problems using the SQP method. However, to do this, we need the gradients of the cost functional (43) and the canonical constraints (42) with respect to the decision parameters. Since (42) and (43) are implicit, rather than explicit, functions of $\boldsymbol{\sigma}$ and $\boldsymbol{\zeta}$, computing their gradients is a non-trivial task. We now show that these gradients can be computed using the sensitivity method described in [13].

We first define the state variations with respect to σ_i^k and ζ_k as follows:

$$\phi^{ki}(s|\boldsymbol{\sigma},\boldsymbol{\zeta}) = \frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k}, \tag{44}$$

and

$$\varphi^k(s|\boldsymbol{\sigma},\boldsymbol{\zeta}) = \frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \zeta_k}. \tag{45}$$

Furthermore, define

$$\begin{aligned}
\boldsymbol{g}(\tilde{\boldsymbol{x}}(s), \tilde{v}_{N+1}(s), \tilde{v}_{N+2}(s), \tilde{v}_{N+3}(s)) ={}& -\tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{F})\tilde{\boldsymbol{x}}(s) \\
&+ \tilde{v}_{N+2}(s)^{-1} \tilde{v}_{N+3}(s)^{\frac{1}{2}} \boldsymbol{A}^{-1}\boldsymbol{d} \\
&+ \kappa \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{1}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}\boldsymbol{e},
\end{aligned}$$

and

$$\boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta}) = \boldsymbol{g}(\tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta}), \tilde{v}_{N+1}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta}), \tilde{v}_{N+2}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta}), \tilde{v}_{N+3}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})).$$

Then

$$\begin{aligned}
\frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}} ={}& -\tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{F}), \\
\frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+1}} ={}& -\frac{3}{2} \tilde{v}_{N+1}(s)^{\frac{1}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{F})\tilde{\boldsymbol{x}}(s) \\
&+ \frac{3\kappa}{2} \tilde{v}_{N+1}(s)^{\frac{1}{2}} \tilde{v}_{N+2}(s)^{-\frac{1}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}\boldsymbol{e}, \\
\frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+2}} ={}& \frac{3}{2} \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{5}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{F})\tilde{\boldsymbol{x}}(s) \\
&- \tilde{v}_{N+2}(s)^{-2} \tilde{v}_{N+3}(s)^{\frac{1}{2}} \boldsymbol{A}^{-1}\boldsymbol{d} - \frac{\kappa}{2} \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{3}{4}} \boldsymbol{A}^{-1}\boldsymbol{e}, \\
\frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+3}} ={}& \frac{3}{4} \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{3}{2}} \tilde{v}_{N+3}(s)^{-\frac{7}{4}} \boldsymbol{A}^{-1}(\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{F})\tilde{\boldsymbol{x}}(s) \\
&+ \frac{1}{2} \tilde{v}_{N+2}(s)^{-1} \tilde{v}_{N+3}(s)^{-\frac{1}{2}} \boldsymbol{A}^{-1}\boldsymbol{d} - \frac{3\kappa}{4} \tilde{v}_{N+1}(s)^{\frac{3}{2}} \tilde{v}_{N+2}(s)^{-\frac{1}{2}} \tilde{v}_{N+3}(s)^{-\frac{7}{4}} \boldsymbol{A}^{-1}\boldsymbol{e},
\end{aligned}$$

where we have omitted the $\boldsymbol{\sigma}$ and $\boldsymbol{\zeta}$ arguments on the right-hand side for clarity. For each $l = 1, 2, \ldots, p,$

it follows from (40) that

$$\tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta}) = \tilde{\boldsymbol{x}}^p(l-1|\boldsymbol{\sigma},\boldsymbol{\zeta}) + \int_{l-1}^{s} \zeta_l \boldsymbol{g}(\eta|\boldsymbol{\sigma},\boldsymbol{\zeta})d\eta, \qquad s \in [l-1,l). \tag{46}$$

If $k \leq l$, then differentiating (46) with respect to σ_i^k yields

$$\begin{aligned}
\frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} &= \frac{\partial \tilde{\boldsymbol{x}}^p(l-1|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} \\
&+ \int_{l-1}^{s} \left\{ \zeta_l \frac{\partial \boldsymbol{g}(\eta|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}} \frac{\partial \tilde{\boldsymbol{x}}^p(\eta|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} + \zeta_l \frac{\partial \boldsymbol{g}(\eta|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+i}} \frac{\partial \tilde{v}_{N+i}^p(\eta|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} \right\} d\eta, \quad s \in [l-1,l).
\end{aligned} \tag{47}$$

Since σ_i^k is the ith control value on the subinterval $s \in [k-1,k)$, it does not affect the state at time instants before $s \in [k-1,k)$. Hence, if $k > l$, we can obtain

$$\frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} = \boldsymbol{0}, \qquad s \in [l-1,l). \tag{48}$$

By differentiating (47) with respect to s, we obtain

$$\frac{d}{ds}\left\{ \frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} \right\} = \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}} \frac{\partial \tilde{\boldsymbol{x}}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} + \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+i}} \frac{\partial \tilde{v}_{N+i}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k}, \quad s \in [l-1,l). \tag{49}$$

Moreover, from (40b),

$$\frac{\partial \tilde{\boldsymbol{x}}^p(0|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k} = \frac{\partial}{\partial \sigma_i^k}\{\bar{\boldsymbol{A}}^{-1}\bar{\boldsymbol{b}}\} = \boldsymbol{0}. \tag{50}$$

It follows from (48)-(50) that, for each $k = 1, 2, \ldots, p$ and $i = 1, 2, 3$, the state variation defined in (44) satisfies the following dynamic system:

$$\dot{\boldsymbol{\phi}}^{ki}(s) = \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\phi}^{ki}(s) + \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+i}} \frac{\partial \tilde{v}_{N+i}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \sigma_i^k}, \qquad s \in [l-1,l), \quad l = k, k+1, \ldots, p, \tag{51}$$

with the initial condition

$$\boldsymbol{\phi}^{ki}(s) = \boldsymbol{0}, \quad s \leq k-1. \tag{52}$$

Similarly, the state variation defined in (45) satisfies the following dynamic system:

$$\begin{aligned}
\dot{\boldsymbol{\varphi}}^k(s) &= \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\varphi}^k(s) + \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+1}} \frac{\partial \tilde{v}_{N+1}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \zeta_k} + \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+2}} \frac{\partial \tilde{v}_{N+2}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \zeta_k} \\
&+ \zeta_l \frac{\partial \boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \tilde{v}_{N+3}} \frac{\partial \tilde{v}_{N+3}^p(s|\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \zeta_k} + \delta_{kl}\boldsymbol{g}(s|\boldsymbol{\sigma},\boldsymbol{\zeta}), \quad s \in [l-1,l), \quad l = k, k+1, \ldots, p,
\end{aligned} \tag{53}$$

with the initial condition

$$\boldsymbol{\varphi}^k(s) = \boldsymbol{0}, \quad s \leq k-1, \tag{54}$$

where δ_{kl} denotes the Kronecker delta function. The next two results give explicit formulas for the derivatives $\partial \tilde{v}_{N+i}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\zeta})/\partial \sigma_i^k$ and $\partial \tilde{v}_{N+i}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\zeta})/\partial \zeta_k$ in (51) and (53).

Theorem 1. *For each $l = 1, \ldots, p$, the state variation of $\tilde{v}_{N+i}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\zeta})$ with respect to σ_i^k on the interval*

15

$[l-1, l]$ *is given by*

$$\frac{\partial \tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i} = \begin{cases} \zeta_l[s - (l-1)], & \text{if } k = l, \\ \zeta_k, & \text{if } k < l, \\ 0, & \text{if } k > l. \end{cases} \tag{55}$$

Proof. The theorem is proved by induction. For $l = 1$, it follows from (41) that

$$\tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta}) = v^0_{N+i} + \zeta_1 \sigma^1_i(s - 0) = v^0_{N+i} + \zeta_1 \sigma^1_i s, \quad s \in [0, 1].$$

Clearly, for all $s \in [0, 1]$,

$$\frac{\partial v^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i} = \begin{cases} \zeta_1 s, & \text{if } k = 1, \\ 0, & \text{if } k > 1, \end{cases}$$

which shows that (55) is satisfied for $l = 1$. Now, suppose that (55) holds for $l = q$. Then for all $s \in [q - 1, q]$,

$$\frac{\partial \tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i} = \begin{cases} \zeta_q[s - (q-1)], & \text{if } k = q, \\ \zeta_k, & \text{if } k < q, \\ 0, & \text{if } k > q. \end{cases}$$

For $l = q + 1$, we have from (41a),

$$\tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta}) = \tilde{v}^p_{N+i}(q|\boldsymbol{\sigma}, \boldsymbol{\zeta}) + \zeta_{q+1} \sigma^{q+1}_i(s - q), \quad s \in [q, q+1].$$

Hence, for all $s \in [q, q+1]$,

$$\frac{\partial \tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i} = \begin{cases} \zeta_{q+1}(s - q), & \text{if } k = q + 1, \\ \frac{\partial \tilde{v}^p_{N+i}(q|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i}, & \text{if } k < q + 1, \\ 0, & \text{if } k > q + 1. \end{cases}$$

Applying the inductive hypothesis yields

$$\frac{\partial \tilde{v}^p_{N+i}(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma^k_i} = \begin{cases} \zeta_{q+1}(s - q), & \text{if } k = q + 1, \\ \zeta_k, & \text{if } k < q + 1, \\ 0, & \text{if } k > q + 1. \end{cases}$$

This shows that (55) holds for $l = q + 1$. Thus, the result follows from mathematical induction. $\qquad\square$

Theorem 2. *For each $l = 1, \ldots, p$, the state variation of $\tilde{v}^p_{N+i}(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$ with respect to ζ_k on the interval*

$[l-1, l]$ *is given by*

$$\frac{\partial \tilde{v}_{N+i}^p(s|\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \zeta_k} = \begin{cases} \sigma_i^l[s - (l-1)], & \text{if } k = l, \\ \sigma_i^k, & \text{if } k < l, \\ 0, & \text{if } k > l. \end{cases}$$

Proof. Similar to the proof of Theorem 1. □

Using Theorems 1 and 2, we can solve (51)-(52) and (53)-(54) numerically to determine the state variations (44) and (45). The gradient of the canonical constraints (42) can then be immediately evaluated. For the gradient of the cost functional (43), we have

$$\frac{\partial \tilde{J}^p(\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \sigma_i^k} = [\tilde{\boldsymbol{x}}^p(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}) - \boldsymbol{x}^d]^\top \bar{\mathbf{A}} \boldsymbol{\phi}^{ki}(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}),$$

and

$$\frac{\partial \tilde{J}^p(\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \zeta_k} = [\tilde{\boldsymbol{x}}^p(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}) - \boldsymbol{x}^d]^\top \bar{\mathbf{A}} \boldsymbol{\varphi}^k(p|\boldsymbol{\sigma}, \boldsymbol{\zeta}),$$

where $\boldsymbol{\phi}^{ki}(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$ and $\boldsymbol{\varphi}^k(\cdot|\boldsymbol{\sigma}, \boldsymbol{\zeta})$ denote, respectively, the solutions of (51)-(52) and (53)-(54) corresponding to $(\boldsymbol{\sigma}, \boldsymbol{\zeta})$. These gradient formulas can be incorporated into existing nonlinear programming algorithms to solve Problem Q_N^p. In the next section, we validate this approach numerically.

Note that SQP and other nonlinear programming algorithms are designed to find local optima, and there is no way of guaranteeing that the solutions generated by such algorithms are globally optimal. This problem can be addressed by running the optimization algorithm from different starting points. The time-scaling transformation provides increased optimization flexibility, often allowing for a reduction in the number of decision variables needed to achieve a desired level of accuracy. With less decision variables, it becomes much easier to determine good starting points and thus mitigate the issues caused by local optima. One strategy that works well in practice is to first solve the problem with fixed switching times and then use the solution obtained as the starting point for the problem with variable switching times.

5. Numerical Simulations

This section reports computational results from applying the method described in Sections 3 and 4 to the example problem in reference [28]. For this example, the functions $D(\hat{\rho})$, $\vartheta_1(\hat{\rho})$, and $\vartheta_2(\hat{\rho})$ are constructed using experimental data from the DIII-D tokamak in San Diego, California; see Figures 3-5. Moreover, the initial magnetic flux profile is taken from shot #129412 from the DIII-D tokamak; see Figure 6. For the input bounds and initial conditions, we have $\bar{n}_{\min} = 1.0$ [10^{19}m^{-3}], $\bar{n}_{\max} = 2.95$ [10^{19}m^{-3}], $I_{\min} = 0.06$ [MA], $I_{\max} = 1.18$ [MA], $P_{\min} = 1.90$ [MW], $P_{\max} = 4.7$ [MW], $\bar{n}_0 = 2.0282$ [10^{19}m^{-3}], $I_0 = 0.7299$ [MA], $P_0 = 4.6773$ [MW]. Our simulation study was carried out within the MATLAB programming environment. The MATLAB code implements the gradient-based optimization procedure described in Section 4 by combining FMINCON with MATLAB's non-stiff differential equation solver ODE45 to integrate the state system (40) and the sensitivity systems (51)-(52)
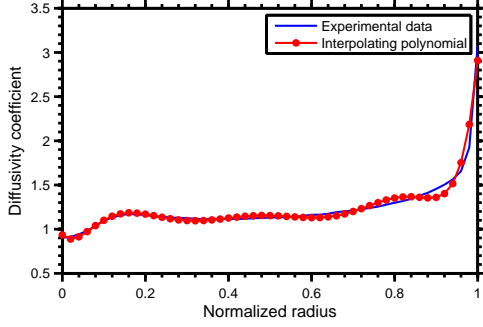
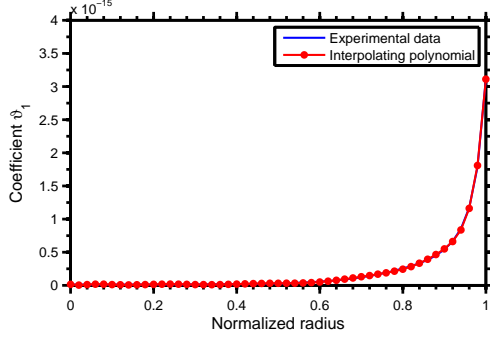Figure 3: Diffusivity coefficient function $D(\hat{\rho})$.

Figure 4: Coefficient function $\vartheta_1(\hat{\rho})$.

and (53)-(54). We considered two target profiles $\omega_d(\hat{\rho})$ in our simulations: the first target profile is generated using the experimental input data in [28]; the second target profile is generated using the experimental output flux data in [19] (see Figure 20(a) in [19]).

5.1. Model Validation

Our first numerical experiment was designed to validate the finite element scheme by simulating the approximate ODE model under the experimental input data in [28]. For the spatial discretization of the PDE system, we used the first-order basis functions $\beta_i(\hat{\rho})$, $i = 1, 2, \ldots, N$, defined by

$$
\beta_i(\hat{\rho}) = \begin{cases} 1 + N\hat{\rho} - i, & \text{if } \hat{\rho} \in \left[\frac{i-1}{N}, \frac{i}{N}\right], \\ 1 - N\hat{\rho} + i, & \text{if } \hat{\rho} \in \left[\frac{i}{N}, \frac{i+1}{N}\right], \\ 0, & \text{otherwise,} \end{cases}
\tag{56}
$$

where N is the number of subintervals in the spatial domain. A comparison of the results from our finite element scheme with $N = 10$ and the MATLAB function PDEPE is shown in Figure 7. The maximum error is less than 6%. Thus, by using just ten first-order spatial basis functions, we can capture over 94% of the original PDE dynamics, as shown in Figure 7. The numerical results in Sections 5.2 and 5.3 were also generated using first-order basis functions defined by equation (56).

5.2. Optimal Control for Target Profile 1

To apply the control parameterization technique, we subdivide the time interval $[0, T] = [0, 1.2]$ into p subintervals. The control input functions are approximated by piecewise-linear functions with break-points at $\gamma_1, \gamma_2, \ldots, \gamma_{p-1}$, where $\gamma_0 = 0$ and $\gamma_p = 1.2$. For our simulations with target profile 1, we set $N = 10$ as the number of spatial intervals and $\epsilon = 0.001$ as the right boundary penalization parameter. Moreover, we set the terminal bounds as $I_{\min}^T = I_{\max}^T = 1.18$ [MA], $\bar{n}_{\min}^T = P_{\min}^T = -\infty$, $\bar{n}_{\max}^T = P_{\max}^T = +\infty$. Thus, $\bar{n}(T)$ and $P(T)$ are free while the input current is subject to the terminal constraint $I(T) = 1.18$ [MA].

We first considered equidistant switching times in which $\tau_{\min} = \tau_{\max} = \frac{1.2}{p}$. The optimal controls and optimal flux profiles generated by our method for $p = 6, 8, 10$ are shown in Figure 8. The results show
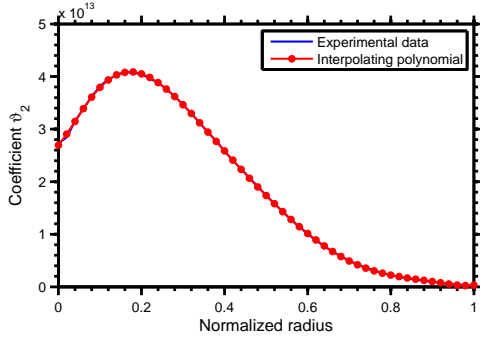
18

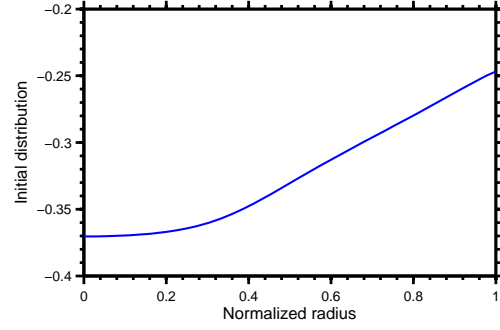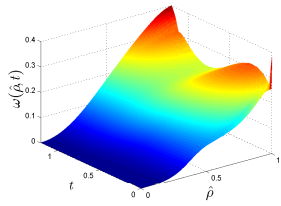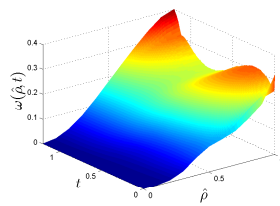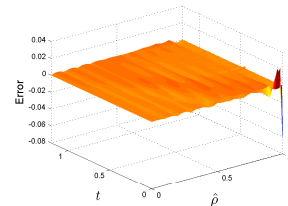Figure 5: Coefficient function $\vartheta_2(\hat{\rho})$.

Figure 6: The initial distribution $\psi_0(\hat{\rho})$.
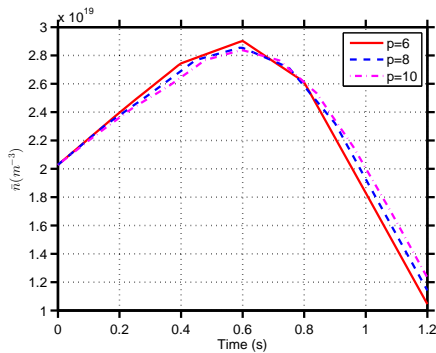

(a) PDEPE in MATLAB.

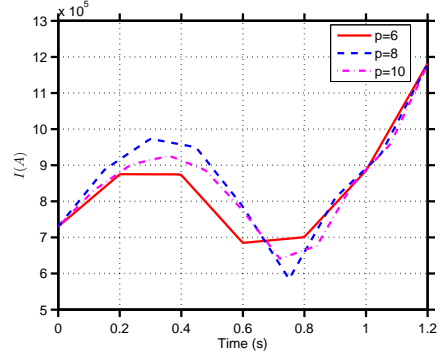(b) Finite element scheme in Section 3 ($N = 10$).

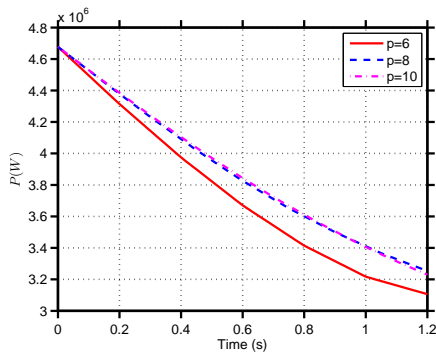(c) Numerical error: PDEPE in MATLAB versus our finite element method.

Figure 7: State trajectory corresponding to the experimental input data in [28].
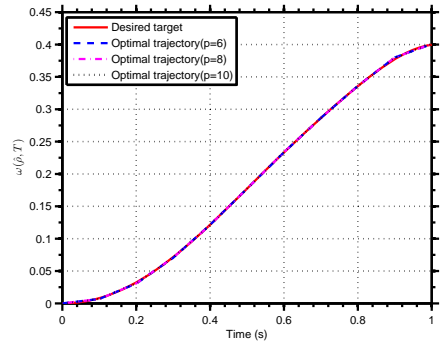

(a) Optimal control input $\bar{n}(t)$.

(b) Optimal control input $I(t)$.

(c) Optimal control input $P(t)$.

(d) Optimal ω-profile at the terminal time.

Figure 8: Optimal controls for target profile 1 (fixed switching times).

(a) Optimal control input $\bar{n}(t)$.

(b) Optimal control input $I(t)$.

(c) Optimal control input $P(t)$.
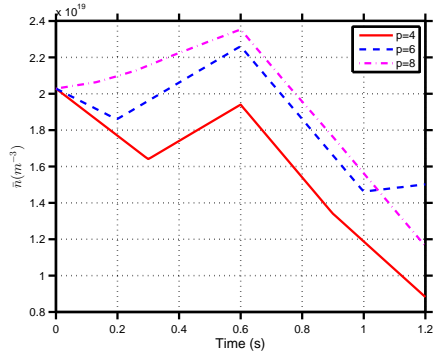
(d) Optimal ω-profile at the terminal time.
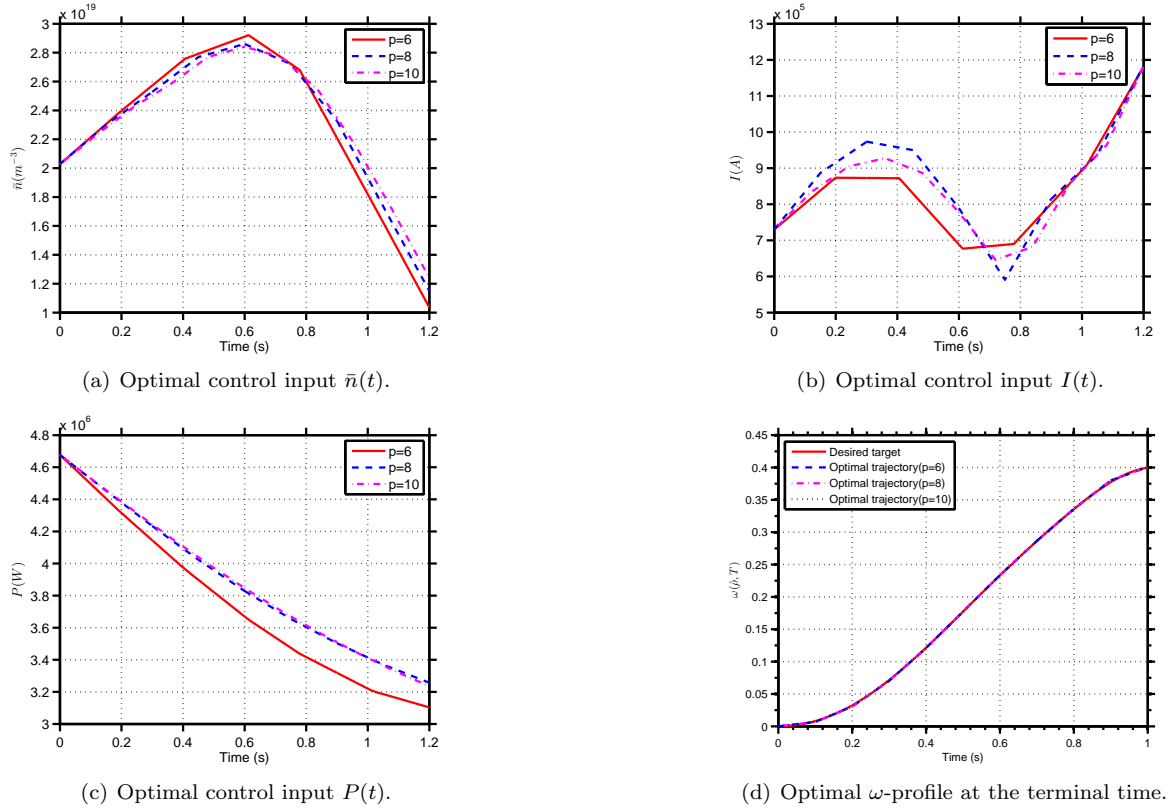
Figure 9: Optimal controls for target profile 1 (variable switching times).
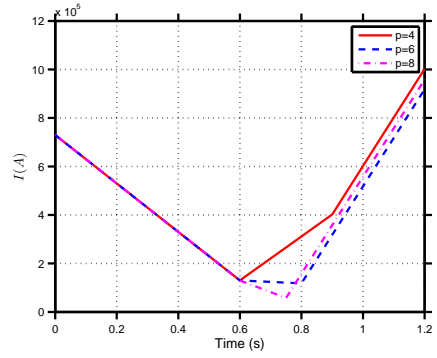
Table 1: Optimal cost values for target profile 1.

	Optimal Cost	
p	Fixed Switching Times	Variable Switching Times
6	7.1477×10^{-8}	6.6488×10^{-8}
8	6.1294×10^{-8}	6.1085×10^{-8}
10	5.6701×10^{-8}	5.6697×10^{-8}

that the proposed numerical optimization method can drive the final ω-profile to within close proximity of the desired target profile. Note that increasing p yields reduced matching error, as expected. However, increasing p from $p = 8$ to $p = 10$ does not result in any significant change in the objective functional value, despite a significant increase in the overall computation time.

We next used the time-scaling transformation to allow for variable switching times. We again considered $p = 6, 8, 10$ subintervals over a time horizon of $[0, 1.2]$, with $\tau_{\min} = 0.12$ and $\tau_{\max} = 0.2$. Our results are shown in Figure 9. Compared with the results in Figure 8, the final ω-profile is closer to the desired ω_d-profile. That is, the optimal solutions in Figure 9 have lower costs than the corresponding solutions in Figure 8, thus demonstrating that the time-scaling transformation is a useful tool for improving matching accuracy via adaptive optimization of the control break-points. Table 1 provides a comparison of the optimal costs for fixed versus variable switching times.
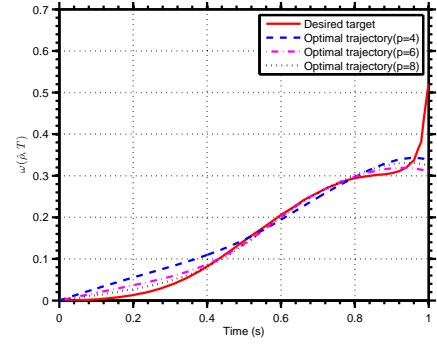
(a) Optimal control input $\bar{n}(t)$.

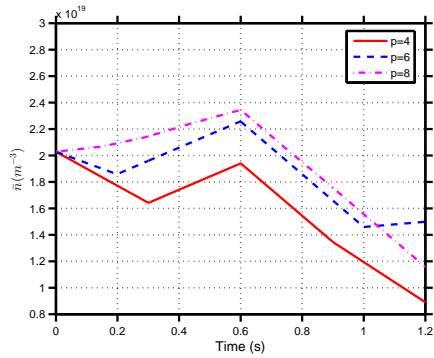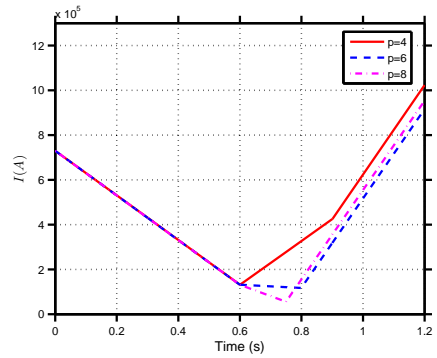(b) Optimal control input $I(t)$.
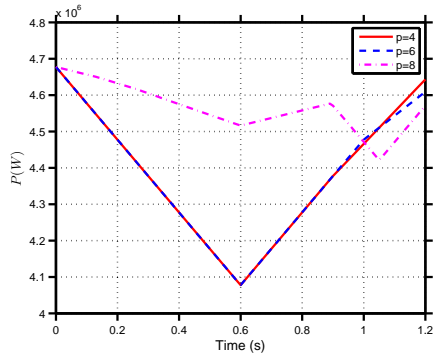
(c) Optimal control input $P(t)$.

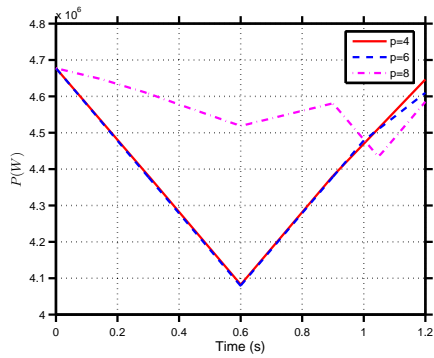(d) Optimal ω-profile at the terminal time.

Figure 10: Optimal controls for target profile 2 (fixed switching times).
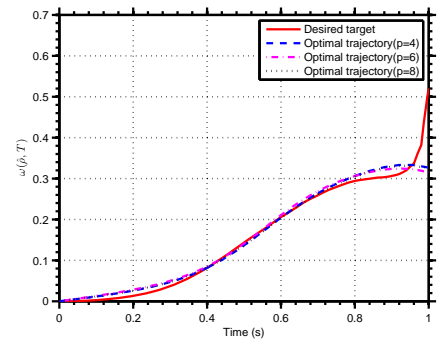

(a) Optimal control input $\bar{n}(t)$.

(b) Optimal control input $I(t)$.

(c) Optimal control input $P(t)$.

(d) Optimal ω-profile at the terminal time.

Figure 11: Optimal controls for target profile 2 (variable switching times).

Table 2: Optimal cost values for target profile 2.

| p | Optimal Cost | |
	Fixed Switching Times	Variable Switching Times
4	4.5588×10^{-4}	2.7403×10^{-4}
6	3.1073×10^{-4}	2.3765×10^{-4}
8	2.3757×10^{-4}	2.3466×10^{-4}

5.3. Optimal Control for Target Profile 2

For our simulations with target profile 2, we allowed the terminal states $\bar{n}(T)$, $I(T)$ and $P(T)$ to be free. As in Section 5.2, we used $\epsilon = 0.001$ as the boundary penalization parameter. Target profile 2 is more challenging numerically and thus we used $N = 25$ spatial intervals compared with $N = 10$ for target profile 1. For $p = 4, 6, 8$ and equidistant switching times, the optimal open-loop actuator input trajectories and associated terminal ω-profile are shown in Figure 10. From the simulation results, we can see that, as with target profile 1, the numerical optimization method is successful at driving the output profile $\omega(\hat{\rho}, T)$ towards the given target $\omega_d(\hat{\rho})$. The matching accuracy is especially high during the interval $[0, 0.9]$. The reason for the matching discrepancy near $\hat{\rho} = 1$ is because the given target is not reachable for the simplified PDE model. We next allowed the switching times to vary and applied the time-scaling transformation. We again considered $p = 4, 6, 8$ subintervals over a time horizon of $[0, 1.2]$. The optimal time evolutions for the three input signals $\bar{n}(t)$, $I(t)$, $P(t)$ and the corresponding output trajectory $\omega(\hat{\rho}, T)$ are shown in Figure 11. These results are improvements over the results in Figure 10, although there is still some matching discrepancy near $\hat{\rho} = 1$. In Table 2, the optimal costs are compared for fixed and variable switching times.

6. Conclusion and Future Work

This paper has presented an effective computational method for solving a finite-time optimal control problem arising during the ramp-up phase of a tokamak plasma. The method is based on a combination of the Galerkin finite element method and the control parameterization technique. Simulation results using experimental data from the DIII-D tokamak demonstrate that the method is effective at driving the plasma profile to a predefined desired profile at the terminal time. Nevertheless, there is still room for improvement. We list several opportunities for future research below.

- Our problem formulation in Section 2 only considers bound constraints on the control variables. However, real tokamaks may also be subject to state constraints. These constraints are more difficult to handle than control bounds, but effective computational methods do exist (see [9, 14]).

- Our numerical simulations focused on the ramp-up phase of the tokamak. In future work, we will apply our method to the flat-top phase for steady-state operation.

- The time-scaling transformation used in this work is the critial tool for dealing with optimal control problems with free terminal times such as time-optimal control problem. We will explore this challenging issue for tokamak operation in future work.

22

- We used fixed knot points for the spatial discretization in Section 3. More accurate results may be obtained by using an adaptive mesh refinement scheme.

- Our method follows the discretize-then-optimize approach, whereby the PDE system is first discretized to obtain a system of ODEs, and then computational optimal control techniques for ODEs are applied. An alternative approach would be to optimize the original PDE system directly, i.e., optimize-then-discretize.

References

[1] N. U. Ahmed and K. L. Teo. *Optimal Control of Distributed Parameter Systems*. North Holland, 1984.

[2] M. Ariola and A. Pironti. *Magnetic Control of Tokamak Plasmas*. Springer, 2008.

[3] B. F. Ben, B. Christine, and E. F. Henda. Dirichlet boundary control for a parabolic equation with a final observation I: A space-time mixed formulation and penalization. *Asymptotic Analysis*, 71(1-2):101–121, 2011.

[4] J. Blum. *Numerical Simulation and Optimal Control in Plasma Physics*. Wiley, 1989.

[5] M. D. Boyer, J. Barton, E. Schuster, T. C. Luce, J. R. Ferron, M. L. Walker, D. A. Humphreys, B. G. Penaflor, and R. D. Johnson. First-principles-driven model-based current profile control for the DIII-D tokamak via LQI optimal control. *Plasma Physics and Controlled Fusion*, 55(10):105007, 2013.

[6] F. Bribiesca-Argomedo, C. Prieur, E. Witrant, and S. Brémond. A strict control Lyapunov function for a diffusion equation with time-varying distributed coefficients. *IEEE Transactions on Automatic Control*, 58(2):290–303, 2013.

[7] J. V. Dongen, F. Felici, G. M. D. Hogeweij, P. Geelen, and E. Maljaars. Numerical optimization of actuator trajectories for ITER hybrid scenario profile evolution. *Plasma Physics and Controlled Fusion*, 56(12):125008, 2014.

[8] F. Felici and O. Sauter. Non-linear model-based optimization of actuator trajectories for tokamak plasma profile control. *Plasma Physics and Controlled Fusion*, 54(2):025002, 2012.

[9] Q. Lin, R. Loxton, and K. L. Teo. The control parameterization method for nonlinear optimal control: A survey. *Journal of Industrial and Management Optimization*, 10(1):275–309, 2013.

[10] Q. Lin, Y. H. Wu, R. Loxton, and S. Y. Lai. Linear B-spline finite element method for the improved Boussinesq equation. *Journal of Computational and Applied Mathematics*, 224(2):658–667, 2009.

[11] R. Loxton, Q. Lin, V. Rehbock, and K. L. Teo. Control parameterization for optimal control problems with continuous inequality constraints: New convergence results. *Numerical Algebra, Control and Optimization*, 2(3):571–599, 2012.

[12] R. Loxton, Q. Lin, and K. L. Teo. Switching time optimization for nonlinear switched systems: Direct optimization and the time-scaling transformation. *Pacific Journal of Optimization*, 10(3):537–560, 2014.

[13] R. Loxton, K. L. Teo, and V. Rehbock. Optimal control problems with multiple characteristic time points in the objective and constraints. *Automatica*, 44(11):2923–2929, 2008.

[14] R. Loxton, K. L. Teo, V. Rehbock, and K. F. C. Yiu. Optimal control problems with a continuous inequality constraint on the state and the control. *Automatica*, 45(10):2250–2257, 2009.

[15] D. Moreau et al. Real-time control of the q-profile in JET for steady state advanced tokamak operation. *Nuclear Fusion*, 43(9):870–882, 2003.

[16] D. Moreau et al. Plasma models for real-time control of advanced tokamak scenarios. *Nuclear Fusion*, 51(6):063009, 2011.

[17] M. Murakami et al. Progress toward fully noninductive, high beta conditions in DIII-D. *Physics of Plasmas*, 13(5):056106, 2006.

[18] Y. Ou, T. Luce, E. Schuster, J. Ferron, M. Walker, C. Xu, and D. Humphreys. Towards model-based current profile control at DIII-D. *Fusion Engineering and Design*, 82(5):1153–1160, 2007.

[19] Y. Ou, C. Xu, E. Schuster, T. Luce, J. Ferron, M. Walker, and D. Humphreys. Design and simulation of extremum-seeking open-loop optimal control of current profile in the DIII-D tokamak. *Plasma Physics and Controlled Fusion*, 50(11):115001, 2008.

[20] Y. Ou, C. Xu, E. Schuster, T. Luce, J. Ferron, M. Walker, and D. Humphreys. Optimal tracking control of current profile in tokamaks. *IEEE Transactions on Control Systems Technology*, 19(2):432–441, 2011.

[21] Z. Ren, C. Xu, Q. Lin, R. Loxton, K. L. Teo, and J. Chu. Finite element approximation and input parameterization for the optimal control of current profiles in tokamak plasmas. In *Proceedings of the 19th World Congress of the International Federation of Automatic Control*, volume 19, pages 7892–7897. IFAC, 2014.

[22] T. Taylor. Physics of advanced tokamaks. *Plasma Physics and Controlled Fusion*, 39(12B):B47–B79, 1997.

[23] K. L. Teo, C. J. Goh, and K. H. Wong. *A Unified Computational Approach to Optimal Control Problems*. Longman Scientific and Technical, 1991.

[24] K. L. Teo and Z. S. Wu. *Computational Methods for Optimizing Distributed Systems*. Academic Press, 1984.

[25] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer, 2006.

[26] P. Wang. Identification of distributed parameter systems: Non-computational aspects. In *Distributed Parameter Systems: Modelling and Identification*, pages 1–10. Springer, 1978.

[27] E. Witrant, E. Joffrin, S. Brémond, G. Giruzz, D. Mazon, O. Barana, and P. Moreau. A control-oriented model of the current profile in tokamak plasma. *Plasma Physics and Controlled Fusion*, 49(7):1075–1105, 2007.

[28] C. Xu, Y. Ou, J. Dalessio, E. Schuster, T. Luce, J. Ferron, M. Walker, and D. Humphreys. Ramp-up phase current-profile control of tokamak plasmas via nonlinear programming. *IEEE Transactions on Plasma Science*, 38(2):163–173, 2010.

[29] C. Xu, Y. Ou, and E. Schuster. Sequential linear quadratic control of bilinear parabolic PDEs based on POD model reduction. *Automatica*, 47(2):418–426, 2011.