

Dynamic Optimization of Tokamak Plasmas via Control Parameterization and the Time-scaling Transformation

Zhigang Ren^a, Chao Xu^a, Qun Lin^b, and Ryan Loxton^b

^aState Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China
E-mail: renzhigang@zju.edu.cn, cxu@zju.edu.cn

^bDepartment of Mathematics & Statistics, Curtin University, Perth, Australia
E-mail: q.lin@curtin.edu.au, r.loxton@curtin.edu.au

Abstract: Fusion nuclear reactions, in which multiple atomic nuclei collide to form a single atomic nucleus, can only occur at extremely high temperatures, where all matter is in the plasma state. In the majority of today's experimental fusion reactors, the fusion plasma is confined to a torus shape using a magnetic confinement system called a tokamak. The performance of a tokamak depends crucially on the current spatial profile, which is related to the poloidal magnetic flux. Accordingly, in this paper, we investigate a finite-time optimal control problem in which the aim is to drive the current spatial profile to within close proximity of a desired target profile, subject to a parabolic PDE governing the evolution of the poloidal magnetic flux. To solve this optimal control problem, we first use the finite element method to approximate the PDE model by an ODE model. Then, we apply the control parameterization and time-scaling techniques to obtain an approximate finite-dimensional optimization problem, which can be solved using sequential quadratic programming methods. Simulation results using experimental data from the DIII-D tokamak in San Diego, California demonstrate the effectiveness of the proposed approach.

Key Words: Nuclear Fusion, Current Profile Control, Finite Element Method, Control Parameterization, Time-scaling Transformation, Dynamic Optimization

1 Introduction

The evolution of the poloidal magnetic flux can be described by the following parabolic PDE [1]:

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

with the Neumann boundary conditions

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

and the initial condition

$$\psi(\hat{\rho}, 0) = \psi_0(\hat{\rho}), \quad (3)$$

where t denotes time; $\hat{\rho}$ denotes normalized radius; $\psi(\hat{\rho}, t)$ denotes the poloidal magnetic flux around the tokamak; and $\vartheta_1(\hat{\rho})$, $\vartheta_2(\hat{\rho})$ and $D(\hat{\rho})$ are given functions. 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:

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

$$u_2(t) = I(t)^{-1} P(t)^{\frac{1}{2}}, \quad (4b)$$

$$u_3(t) = \kappa I(t), \quad (4c)$$

where κ is a given constant. Note that $\bar{n}(t)$, $I(t)$, and $P(t)$ are the control inputs for the physical actuators.

The output toroidal current spatial profile is defined in terms of the magnetic flux as follows:

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

The current spatial profile is crucial to magnetohydrodynamic stability and efficient steady-state operation [2, 3]. In practical operation, the aim is to make the current profile $\omega(\hat{\rho}, t)$ match a given target profile $\omega_d(\hat{\rho})$ at the terminal time $t = T$. Thus, the problem is to minimize the following cost functional:

$$J(\bar{n}, I, P) = \frac{1}{2} \int_0^1 \Gamma_0 (\omega(\hat{\rho}, T) - \omega_d(\hat{\rho}))^2 d\hat{\rho} + \int_0^T (\Gamma_1 \bar{n}(t) + \Gamma_2 I(t) + \Gamma_3 P(t)) dt, \quad (5)$$

where Γ_0 , Γ_1 , Γ_2 and Γ_3 are non-negative weighting factors. The actuator inputs must satisfy the following physical bound constraints:

$$a_1 \leq \bar{n}(t) \leq b_1, a_2 \leq I(t) \leq b_2, a_3 \leq P(t) \leq b_3, \quad (6)$$

where a_i and b_i , $i = 1, 2, 3$, are given constants.

Using subscript notation for the derivatives, equation (1) with boundary conditions (2) and initial condition (3) can be written as

$$\frac{\psi_i(\hat{\rho}, t)}{\vartheta_1(\hat{\rho})} = \frac{u_1(t)}{\hat{\rho}} [\hat{\rho} D(\hat{\rho}) \psi_{\hat{\rho}}(\hat{\rho}, t)]_{\hat{\rho}} + \vartheta_2(\hat{\rho}) u_2(t), \quad (7a)$$

$$\psi_{\hat{\rho}}(0, t) = 0, \quad \psi_{\hat{\rho}}(1, t) = u_3(t), \quad t \in [0, T], \quad (7b)$$

$$\psi(\hat{\rho}, 0) = \psi_0(\hat{\rho}), \quad \hat{\rho} \in [0, 1]. \quad (7c)$$

We now state our optimal control problem formally as follows: *Given the PDE system (7), find control signals $\bar{n}(t)$, $I(t)$ and $P(t)$ to minimize the cost functional (5) subject to the constraints (6). We refer to this problem as Problem P.*

Problem P has been considered previously in references [1, 4]. In these references, discretize-then-optimize computational approaches are developed for computing the optimal control signals: first the parabolic PDE (1) is approximated by a system of ODEs (using the proper orthogonal decomposition method in [1] and the finite element

This work was partially supported by the National Natural Science Foundation of China through grants (61473253, 61320106009) and Development Program of China (863 Program) (2012AA041701).

method in [4]); then the resulting ODE optimal control problem is converted into a finite-dimensional optimization problem using the control parameterization method [5]. Reference [1] uses a piecewise-linear control parameterization scheme in which the control function is approximated by a piecewise-linear function defined in terms of a finite number of decision parameters. Reference [4] uses a piecewise-constant control parameterization scheme in which the control heights are the decision variables to be chosen optimally. The disadvantage of these existing control parameterization schemes is that they are based on an equally-spaced partition of the time horizon: the break-points/discontinuities of the piecewise-linear and piecewise-constant approximate controls are pre-specified at equidistant points. Equidistant break-points/discontinuities are unlikely to be optimal in terms of minimizing the current profile matching error. Thus, in this paper, we propose a superior approach in which the partition points are chosen optimally via numerical optimization techniques, along with the parameters defining the approximate controls. This requires the use of a novel time-scaling transformation, as directly optimizing variable partition points is well known to cause computational difficulties [6]. Numerical results show that optimizing the partition points leads to significant reductions in matching error and input energy compared with the methods in [1, 4].

2 Finite Element Approximation

Let $\eta(\hat{\rho})$ be a trial function. Multiplying both sides of (7a) by $\hat{\rho}\eta(\hat{\rho})$ and then integrating the resulting equation over $[0, 1]$ gives

$$\begin{aligned} & \int_0^1 \frac{\hat{\rho}\eta(\hat{\rho})\psi_t(\hat{\rho}, t)}{\vartheta_1(\hat{\rho})} d\hat{\rho} \\ &= \int_0^1 u_1(t)\eta(\hat{\rho}) \frac{\partial}{\partial \hat{\rho}} [\hat{\rho}D(\hat{\rho})\psi_{\hat{\rho}}(\hat{\rho}, t)] d\hat{\rho} \\ &+ \int_0^1 \hat{\rho}\eta(\hat{\rho})\vartheta_2(\hat{\rho})u_2(t)d\hat{\rho}. \end{aligned}$$

Using integration by parts and applying equation (7b), we obtain

$$\begin{aligned} & \int_0^1 \frac{\hat{\rho}\eta(\hat{\rho})\psi_t(\hat{\rho}, t)}{\vartheta_1(\hat{\rho})} d\hat{\rho} = D(1)\eta(1)u_1(t)u_3(t) \\ & - u_1(t) \int_0^1 \hat{\rho}D(\hat{\rho})\psi_{\hat{\rho}}(\hat{\rho}, t)\eta'(\hat{\rho})d\hat{\rho} \\ & + u_2(t) \int_0^1 \hat{\rho}\vartheta_2(\hat{\rho})\eta(\hat{\rho})d\hat{\rho}. \end{aligned} \quad (8)$$

We partition the spatial domain $[0,1]$ into N subintervals $I_i = [(i-1)\Delta, i\Delta]$, $i = 1, 2, \dots, N$, where $\Delta = 1/N$. Then, we assume that the magnetic flux profile $\psi(\hat{\rho}, t)$ can be approximated by a linear combination of basis B-spline functions $\beta_i(\hat{\rho})$, $i = 0, 1, \dots, N$, corresponding to this partition:

$$\psi(\hat{\rho}, t) \approx \sum_{i=0}^N x_i(t)\beta_i(\hat{\rho}), \quad (9)$$

where $x_i(t)$, $i = 0, 1, \dots, N$, are weighting functions. Substituting (9) into (8) yields

$$\begin{aligned} & \sum_{i=0}^N \int_0^1 \frac{\hat{\rho}\eta(\hat{\rho})}{\vartheta_1(\hat{\rho})} \dot{x}_i(t)\beta_i(\hat{\rho})d\hat{\rho} = D(1)\eta(1)u_1(t)u_3(t) \\ & - u_1(t) \sum_{i=0}^N \int_0^1 \hat{\rho}D(\hat{\rho})x_i(t)\beta'_i(\hat{\rho})\eta'(\hat{\rho})d\hat{\rho} \\ & + u_2(t) \int_0^1 \hat{\rho}\vartheta_2(\hat{\rho})\eta(\hat{\rho})d\hat{\rho}. \end{aligned} \quad (10)$$

By choosing $\eta(\hat{\rho}) = \beta_j(\hat{\rho})$, $j = 0, 1, \dots, N$, as the trial functions, equation (10) can be rewritten in matrix form as follows:

$$\mathbf{A}\dot{\mathbf{x}}(t) = -u_1(t)\mathbf{B}\mathbf{x}(t) + u_2(t)\mathbf{c} + u_1(t)u_3(t)\mathbf{d}, \quad (11)$$

where

$$\begin{aligned} \mathbf{x}(t) &= [x_0(t), x_1(t), \dots, x_N(t)]^\top, \\ \mathbf{A} &= [A_{ij}] = \left[\int_0^1 \frac{\hat{\rho}\beta_i(\hat{\rho})\beta_j(\hat{\rho})}{\vartheta_1(\hat{\rho})} d\hat{\rho} \right], \\ \mathbf{B} &= [B_{ij}] = \left[\int_0^1 \hat{\rho}D(\hat{\rho})\beta'_i(\hat{\rho})\beta'_j(\hat{\rho})d\hat{\rho} \right], \\ \mathbf{c} &= [c_j] = \left[\int_0^1 \hat{\rho}\vartheta_2(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho} \right], \\ \mathbf{d} &= [d_j] = [D(1)\beta_j(1)]. \end{aligned}$$

Recalling the initial condition (7c), we must have

$$\psi_0(\hat{\rho}) = \sum_{i=0}^N x_i(0)\beta_i(\hat{\rho}). \quad (12)$$

Multiplying both sides of (12) by $\beta_j(\hat{\rho})$ and then integrating over $[0, 1]$ gives

$$\sum_{i=0}^N \left[\int_0^1 \beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho} \right] x_i(0) = \int_0^1 \beta_j(\hat{\rho})\psi_0(\hat{\rho})d\hat{\rho}. \quad (13)$$

Define matrix $\bar{\mathbf{A}}$ and vector $\bar{\mathbf{b}}$ as follows:

$$\begin{aligned} \bar{\mathbf{A}} &= [\bar{A}_{ij}] = \left[\int_0^1 \beta_i(\hat{\rho})\beta_j(\hat{\rho})d\hat{\rho} \right], \\ \bar{\mathbf{b}} &= [\bar{b}_j] = \left[\int_0^1 \beta_j(\hat{\rho})\psi_0(\hat{\rho})d\hat{\rho} \right]. \end{aligned}$$

Then equation (13) can be written as

$$\bar{\mathbf{A}}\mathbf{x}(0) = \bar{\mathbf{b}}. \quad (14)$$

By following the same arguments as in [7], it can be shown that matrices \mathbf{A} and $\bar{\mathbf{A}}$ are nonsingular. Consequently, equations (4), (11), and (14) can be combined to yield the following approximate ODE system for PDE system (7):

$$\begin{aligned} \dot{\mathbf{x}}(t) &= -\bar{n}(t)^{\frac{3}{2}}I(t)^{-\frac{3}{2}}P(t)^{-\frac{3}{4}}\mathbf{A}^{-1}\mathbf{B}\mathbf{x}(t) \\ &+ P(t)^{\frac{1}{2}}I(t)^{-1}\mathbf{A}^{-1}\mathbf{c} \\ &+ \kappa\bar{n}(t)^{\frac{3}{2}}I(t)^{-\frac{1}{2}}P(t)^{-\frac{3}{4}}\mathbf{A}^{-1}\mathbf{d}, \end{aligned} \quad (15a)$$

$$\mathbf{x}(0) = \bar{\mathbf{A}}^{-1}\bar{\mathbf{b}}. \quad (15b)$$

Now, based on (9), we consider the following expansion for the desired output profile $\omega_d(\hat{\rho})$:

$$\omega_d(\hat{\rho}) = \sum_{i=0}^N x_i^d \beta_i'(\hat{\rho}), \quad (16)$$

where x_i^d , $i = 0, 1, \dots, N$, are weighting coefficients. To determine the weighting coefficients, we multiply both sides of (16) by $\beta_j'(\hat{\rho})$ and then integrate over $\hat{\rho} \in [0, 1]$ to obtain

$$\int_0^1 \omega_d(\hat{\rho}) \beta_j'(\hat{\rho}) d\hat{\rho} = \sum_{i=0}^N \left[\int_0^1 \beta_i'(\hat{\rho}) \beta_j'(\hat{\rho}) d\hat{\rho} \right] x_i^d. \quad (17)$$

Define matrix $\hat{\mathbf{A}}$ and vector $\hat{\mathbf{b}}$ as follows:

$$\hat{\mathbf{A}} = [\hat{A}_{ij}] = \left[\int_0^1 \beta_i'(\hat{\rho}) \beta_j'(\hat{\rho}) d\hat{\rho} \right],$$

$$\hat{\mathbf{b}} = [\hat{b}_j] = \left[\int_0^1 \omega_d(\hat{\rho}) \beta_j'(\hat{\rho}) d\hat{\rho} \right].$$

Then we can rewrite (17) as

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

where $\mathbf{x}^d = [x_0^d, x_1^d, \dots, x_N^d]^\top$. As with matrices \mathbf{A} and $\bar{\mathbf{A}}$, it can be shown that matrix $\hat{\mathbf{A}}$ is nonsingular. Therefore, the coefficients for the desired output profile are given by

$$\mathbf{x}^d = \hat{\mathbf{A}}^{-1} \hat{\mathbf{b}}. \quad (18)$$

Using the expansion (9), the actual output profile $\omega(\hat{\rho}, T)$ at the terminal time T is approximated as follows:

$$\omega(\hat{\rho}, T) = \frac{\partial \psi(\hat{\rho}, T)}{\partial \hat{\rho}} \approx \sum_{i=0}^N x_i(T) \beta_i'(\hat{\rho}). \quad (19)$$

Substituting (16) and (19) into the cost functional (5), we obtain

$$J(\bar{n}, I, P)$$

$$= \frac{1}{2} \Gamma_0 \sum_{i=0}^N \sum_{j=0}^N (x_i(T) - x_i^d) \hat{A}_{ij} (x_j(T) - x_j^d)$$

$$+ \int_0^T (\Gamma_1 \bar{n}(t) + \Gamma_2 I(t) + \Gamma_3 P(t)) dt \quad (20)$$

$$= \frac{1}{2} \Gamma_0 [\mathbf{x}(T) - \mathbf{x}^d]^\top \hat{\mathbf{A}} [\mathbf{x}(T) - \mathbf{x}^d]$$

$$+ \int_0^T (\Gamma_1 \bar{n}(t) + \Gamma_2 I(t) + \Gamma_3 P(t)) dt,$$

where $\mathbf{x}^d = \hat{\mathbf{A}}^{-1} \hat{\mathbf{b}}$ as in (18). Problem P, the original PDE optimal control problem, is now approximated by the following ODE optimal control problem, which we call Problem Q: *Given the ODE system (15), find control signals $\bar{n}(t)$, $I(t)$ and $P(t)$ to minimize the cost functional (20) subject to the constraints (6).*

3 Numerical Solution Procedure

3.1 Piecewise-linear Control Parameterization

To solve Problem Q, we subdivide the time horizon $[0, T]$ into p subintervals $[\gamma_{k-1}, \gamma_k]$, $k = 1, 2, \dots, p$, where $\gamma_0 =$

0 and $\gamma_p = T$, and the interior knot points γ_k , $k = 1, 2, \dots, p-1$, are free decision parameters. The following constraints are imposed on the subinterval lengths:

$$\tau_{\min} \leq \gamma_k - \gamma_{k-1} \leq \tau_{\max}, \quad k = 1, 2, \dots, p. \quad (21)$$

Here, $\tau_{\min} > 0$ and $\tau_{\max} > 0$ are the minimum and maximum subinterval durations, respectively. We approximate the derivatives of the control signals as follows:

$$\dot{\bar{n}}(t) \approx \sum_{k=1}^p \sigma_1^k \chi_{[\gamma_{k-1}, \gamma_k]}(t), \quad (22a)$$

$$\dot{I}(t) \approx \sum_{k=1}^p \sigma_2^k \chi_{[\gamma_{k-1}, \gamma_k]}(t), \quad (22b)$$

$$\dot{P}(t) \approx \sum_{k=1}^p \sigma_3^k \chi_{[\gamma_{k-1}, \gamma_k]}(t), \quad (22c)$$

where $\chi_{[\gamma_{k-1}, \gamma_k]} : \mathbb{R} \rightarrow \mathbb{R}$ 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 (22), the derivative of each control signal is approximated by a piecewise-constant function with discontinuities at the internal knot points $\gamma_1, \gamma_2, \dots, \gamma_{p-1}$. Thus, the control signals are piecewise-linear with jumps in the derivative at $\gamma_1, \gamma_2, \dots, \gamma_{p-1}$. We introduce new state variables $x_{N+1}(t) = \bar{n}(t)$, $x_{N+2}(t) = I(t)$ and $x_{N+3}(t) = P(t)$ governed by the following dynamics (for $i = 1, 2, 3$):

$$\dot{x}_{N+i}(t) = \sum_{k=1}^p \sigma_i^k \chi_{[\gamma_{k-1}, \gamma_k]}(t), \quad t \in [0, T], \quad (23a)$$

$$x_{N+i}(0) = x_{N+i}^0, \quad (23b)$$

where x_{N+1}^0 is the initial value of $\bar{n}(t)$, x_{N+2}^0 is the initial value of $I(t)$, and x_{N+3}^0 is the initial value of $P(t)$. Then, the dynamic system (15) becomes

$$\dot{\mathbf{x}}(t) = -x_{N+1}^{\frac{3}{2}}(t) x_{N+2}^{-\frac{3}{2}}(t) x_{N+3}^{-\frac{3}{4}}(t) \mathbf{A}^{-1} \mathbf{B} \mathbf{x}(t)$$

$$+ x_{N+2}^{-1}(t) x_{N+3}^{\frac{1}{2}}(t) \mathbf{A}^{-1} \mathbf{c}$$

$$+ \kappa x_{N+1}^{\frac{3}{2}}(t) x_{N+2}^{-\frac{1}{2}}(t) x_{N+3}^{-\frac{3}{4}}(t) \mathbf{A}^{-1} \mathbf{d}, \quad (24a)$$

$$\mathbf{x}(0) = \bar{\mathbf{A}}^{-1} \bar{\mathbf{b}}. \quad (24b)$$

Recall that the control variables $\bar{n}(t)$, $I(t)$, and $P(t)$ must satisfy the bound constraints (6). Thus, we impose the following continuous inequality constraints on the new state variables $x_{N+1}(t)$, $x_{N+2}(t)$ and $x_{N+3}(t)$:

$$a_i \leq x_{N+i}(t) \leq b_i, \quad t \in [0, T], \quad i = 1, 2, 3. \quad (25)$$

Clearly, since $x_{N+1}(t)$, $x_{N+2}(t)$ and $x_{N+3}(t)$ are piecewise-linear with break-points at $t = \gamma_1, \gamma_2, \dots, \gamma_{p-1}$, the continuous inequality constraints (25) are equivalent to the following constraints:

$$a_i \leq x_{N+i}(t_k) \leq b_i, \quad k = 0, 1, \dots, p, \quad i = 1, 2, 3. \quad (26)$$

Such constraints are special cases of the well-known canonical form in the optimal control literature (see [5]).

Now, under the approximation (22), the cost functional (20) becomes

$$J^p(\boldsymbol{\sigma}, \boldsymbol{\gamma}) = \frac{1}{2} \Gamma_0 [\mathbf{x}(T) - \mathbf{x}^d]^\top \hat{\mathbf{A}} [\mathbf{x}(T) - \mathbf{x}^d] + \sum_{i=1}^3 \sum_{k=1}^p \int_{\gamma_{k-1}}^{\gamma_k} \Gamma_i x_{N+i}(t) dt, \quad (27)$$

where $\boldsymbol{\sigma} = [\sigma_i^k]$ and $\boldsymbol{\gamma} = [\gamma_k]$. Our approximate problem, called Problem R, is stated as follows: *Given the ODE system (23)-(24), find a control parameter vector $\boldsymbol{\sigma}$ and a break-point vector $\boldsymbol{\gamma}$ such that the cost functional (27) is minimized subject to constraints (21) and (26).*

3.2 Time-scaling Transformation

The computational difficulties caused by variable partition points are well known [6]. To overcome these difficulties, we transform the old time variable $t \in [0, T]$ into a new time variable $s \in [0, p]$ through the following differential equation:

$$\frac{dt(s)}{ds} = \zeta_k, \quad s \in [k-1, k], \quad k = 1, 2, \dots, p, \quad (28)$$

with the boundary conditions

$$t(0) = 0, \quad t(p) = T, \quad (29)$$

where $\zeta_k = \gamma_k - \gamma_{k-1}$, $k = 1, 2, \dots, p$, are new decision parameters.

Integrating (28)-(29) gives

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

Hence,

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

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

Let $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \dots, \zeta_p]^\top$. Clearly, in view of (21), the vector $\boldsymbol{\zeta} \in \mathbb{R}^p$ must satisfy the following constraints:

$$\tau_{\min} \leq \zeta_k \leq \tau_{\max}, \quad k = 1, 2, \dots, p, \quad (30a)$$

$$\zeta_1 + \zeta_2 + \dots + \zeta_p = T. \quad (30b)$$

Let $\tilde{\mathbf{x}}(s) = \mathbf{x}(t(s))$ and $\tilde{x}_{N+i}(s) = x_{N+i}(t(s))$. Then under the time-scaling transformation (28)-(29), the ODE system (24) becomes

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(s) &= -\zeta_k \tilde{x}_{N+1}^{\frac{3}{2}}(s) \tilde{x}_{N+2}^{-\frac{3}{2}}(s) \tilde{x}_{N+3}^{-\frac{3}{4}}(s) \mathbf{A}^{-1} \mathbf{B} \tilde{\mathbf{x}}(s) \\ &\quad + \zeta_k \tilde{x}_{N+2}^{-1}(s) \tilde{x}_{N+3}^{\frac{1}{2}}(s) \mathbf{A}^{-1} \mathbf{c} \\ &\quad + \kappa \zeta_k \tilde{x}_{N+1}^{\frac{3}{2}}(s) \tilde{x}_{N+2}^{-\frac{1}{2}}(s) \tilde{x}_{N+3}^{-\frac{3}{4}}(s) \mathbf{A}^{-1} \mathbf{d}, \end{aligned} \quad (31a)$$

$$s \in [k-1, k], \quad k = 1, 2, \dots, p, \quad (31b)$$

$$\tilde{\mathbf{x}}(0) = \bar{\mathbf{A}}^{-1} \bar{\mathbf{b}}. \quad (31c)$$

Moreover, the ODEs (23) become (for $i = 1, 2, 3$):

$$\dot{\tilde{x}}_{N+i}(s) = \zeta_k \sigma_i^k, \quad s \in [k-1, k], \quad k = 1, 2, \dots, p, \quad (32a)$$

$$\tilde{x}_{N+i}(0) = x_{N+i}^0. \quad (32b)$$

Since $t(k) = \gamma_k$, the canonical constraints (26) become

$$a_i \leq \tilde{x}_{N+i}(k) \leq b_i, \quad k = 0, 1, \dots, p, \quad i = 1, 2, 3. \quad (33)$$

Under the time-scaling transformation defined by (28)-(29), the cost functional (27) becomes

$$\begin{aligned} \tilde{J}^p(\boldsymbol{\sigma}, \boldsymbol{\zeta}) &= \frac{1}{2} \Gamma_0 [\tilde{\mathbf{x}}(p) - \mathbf{x}^d]^\top \hat{\mathbf{A}} [\tilde{\mathbf{x}}(p) - \mathbf{x}^d] \\ &\quad + \sum_{i=1}^3 \sum_{k=1}^p \int_{k-1}^k \zeta_k \Gamma_i \tilde{x}_{N+i}(s) ds. \end{aligned} \quad (34)$$

The following dynamic optimization problem, which we call Problem S, is equivalent to Problem R: *Given the ODE system (31)-(32), find a control parameter vector $\boldsymbol{\sigma}$ and a subinterval duration vector $\boldsymbol{\zeta}$ such that the cost functional (34) is minimized subject to the constraints (30) and (33).*

3.3 Solving Problem S

To solve Problem S using existing nonlinear optimization techniques, we need the gradients of the cost functional (34) and the canonical constraints (33) with respect to the decision parameters [8]. We now show that these gradients can be computed by solving a set of auxiliary ODEs.

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

$$\phi^{ki}(s) = \frac{\partial \tilde{\mathbf{x}}(s)}{\partial \sigma_i^k}, \quad \varphi^k(s) = \frac{\partial \tilde{\mathbf{x}}(s)}{\partial \zeta_k}. \quad (35)$$

Next, define

$$\begin{aligned} &\mathbf{g}(\tilde{\mathbf{x}}(s), \tilde{x}_{N+1}(s), \tilde{x}_{N+2}(s), \tilde{x}_{N+3}(s)) \\ &= -\tilde{x}_{N+1}^{\frac{3}{2}}(s) \tilde{x}_{N+2}^{-\frac{3}{2}}(s) \tilde{x}_{N+3}^{-\frac{3}{4}}(s) \mathbf{A}^{-1} \mathbf{B} \tilde{\mathbf{x}}(s) \\ &\quad + \tilde{x}_{N+2}^{-1}(s) \tilde{x}_{N+3}^{\frac{1}{2}}(s) \mathbf{A}^{-1} \mathbf{c} \\ &\quad + \kappa \tilde{x}_{N+1}^{\frac{3}{2}}(s) \tilde{x}_{N+2}^{-\frac{1}{2}}(s) \tilde{x}_{N+3}^{-\frac{3}{4}}(s) \mathbf{A}^{-1} \mathbf{d}, \end{aligned}$$

and

$$\bar{\mathbf{g}}(s) = \mathbf{g}(\tilde{\mathbf{x}}(s), \tilde{x}_{N+1}(s), \tilde{x}_{N+2}(s), \tilde{x}_{N+3}(s)).$$

The next two results show how to compute the state variations defined in (35). These results can be proved in a similar manner to the corresponding results in [4, 5].

Theorem 1 *For each $k = 1, 2, \dots, p$ and $i = 1, 2, 3$, the state variation with respect to σ_i^k satisfies the following dynamic system:*

$$\dot{\phi}^{ki}(s) = \zeta_l \frac{\partial \bar{\mathbf{g}}(s)}{\partial \tilde{\mathbf{x}}} \phi^{ki}(s) + \zeta_l \frac{\partial \bar{\mathbf{g}}(s)}{\partial \tilde{x}_{N+i}} \frac{\partial \tilde{x}_{N+i}(s)}{\partial \sigma_i^k}, \quad (36a)$$

$$s \in [l-1, l], \quad l = k, k+1, \dots, p,$$

$$\phi^{ki}(s) = \mathbf{0}, \quad s \leq k-1, \quad (36b)$$

where

$$\frac{\partial \tilde{x}_{N+i}(s)}{\partial \sigma_i^k} = \begin{cases} \zeta_l (s - (l-1)), & \text{if } k = l, \\ \zeta_k, & \text{if } k < l, \\ 0, & \text{if } k > l. \end{cases}$$

Theorem 2 For each $k = 1, 2, \dots, p$, the state variation with respect to ζ_k satisfies the following dynamic system:

$$\begin{aligned} \dot{\varphi}^k(s) &= \zeta_l \frac{\partial \bar{g}(s)}{\partial \tilde{x}} \varphi^k(s) \\ &+ \zeta_l \sum_{i=1}^3 \frac{\partial \bar{g}(s)}{\partial \tilde{x}_{N+i}(s)} \frac{\partial \tilde{x}_{N+i}(s)}{\partial \zeta_k} + \delta_{kl} \bar{g}(s), \quad (37a) \\ s &\in [l-1, l], \quad l = k, k+1, \dots, p, \end{aligned}$$

$$\varphi^k(s) = 0, \quad s \leq k-1, \quad (37b)$$

where δ_{kl} denotes the Kronecker delta function, and

$$\frac{\partial \tilde{x}_{N+i}(s)}{\partial \zeta_k} = \begin{cases} \sigma_i^l(s - (l-1)), & \text{if } k = l, \\ \sigma_i^k, & \text{if } k < l, \\ 0, & \text{if } k > l. \end{cases}$$

Using Theorems 1 and 2, we can solve (36) and (37) numerically to determine the state variations (35). The gradient of the canonical constraints (33) can then be immediately evaluated. For the gradient of the cost functional (34), we have

$$\begin{aligned} \frac{\partial \tilde{J}^p(\sigma, \zeta)}{\partial \sigma_i^k} &= \Gamma_0 [\tilde{x}(p) - \mathbf{x}^d]^\top \hat{\mathbf{A}} \phi^{ki}(p) \\ &+ \sum_{l=1}^p \int_{l-1}^l \zeta_l \Gamma_i \frac{\partial \tilde{x}_{N+i}(s)}{\partial \sigma_i^k} ds, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \tilde{J}^p(\sigma, \zeta)}{\partial \zeta_k} &= \Gamma_0 [\tilde{x}(p) - \mathbf{x}^d]^\top \hat{\mathbf{A}} \varphi^k(p) \\ &+ \sum_{i=1}^3 \sum_{l=1}^p \int_{l-1}^l \zeta_l \Gamma_i \frac{\partial \tilde{x}_{N+i}(s)}{\partial \zeta_k} ds + \sum_{i=1}^3 \int_{k-1}^k \Gamma_i \tilde{x}_{N+i}(s) ds. \end{aligned}$$

These gradient formulas can be incorporated into existing nonlinear programming algorithms to solve Problem S.

4 Simulation Results

We now apply the computational method proposed in Sections 2 and 3 to an example. This example, which comes from reference [1], is based on experimental data from the DIII-D tokamak in San Diego, California. The functions $D(\hat{\rho})$, $\vartheta_1(\hat{\rho})$, and $\vartheta_2(\hat{\rho})$ in the PDE model (1) are given in reference [1]. The initial magnetic flux profile is taken from shot #129412 from the DIII-D tokamak [1]. For the spatial discretization of the PDE system, we use the first-order basis B-spline functions $\beta_i(\hat{\rho})$, $i = 0, 1, \dots, N$, defined by

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

where N is the number of subintervals in the spatial domain and $\Delta = 1/N$ is the subinterval length. In our numerical simulations, we used $N = 5$ subintervals, which proved to be sufficient for high-accuracy solutions.

In applying the control parameterization technique, we subdivide the time interval $[0, T] = [0, 1.2]$ into p

Method	ME ($\times 10^6$)	IE ($\times 10^6$)	Cost ($\times 10^6$)
(i)	1.4156	5.1441	6.5597
(ii)	1.8949	5.1889	7.0838
(iii)	1.3023	4.7619	6.0642

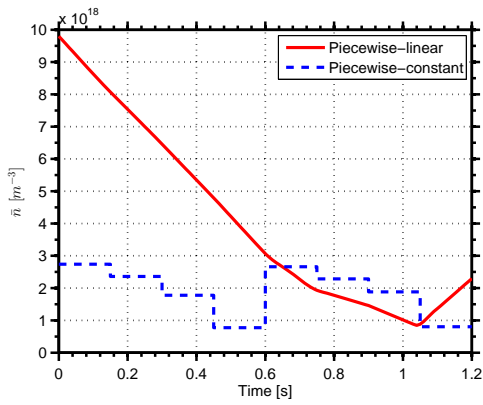
Table 1: Optimal cost values for the example in Section 4 (ME = Matching Error, IE = Input Energy).

subintervals. The control input functions are approximated by piecewise-linear functions with break-points at $\gamma_1, \gamma_2, \dots, \gamma_{p-1}$, where $\gamma_0 = 0$ and $\gamma_p = 1.2$. The lower and upper bounds in (6) are given by: $a_1 = 0.01 [10^{19} \text{m}^{-3}]$, $b_1 = 10.5 [10^{19} \text{m}^{-3}]$, $a_2 = 0.01 [\text{MA}]$, $b_2 = 10.9 [\text{MA}]$, $a_3 = 0.01 [\text{MW}]$, $b_3 = 10 [\text{MW}]$. Furthermore, the weighting factors in the objective functional (5) are $\Gamma_0 = 10^{-2}$, $\Gamma_1 = 10^{-25}$, $\Gamma_2 = 10^{-12}$, $\Gamma_3 = 10^{-12}$, and the desired target profile $\omega_d(\hat{\rho})$ is the same as in [1].

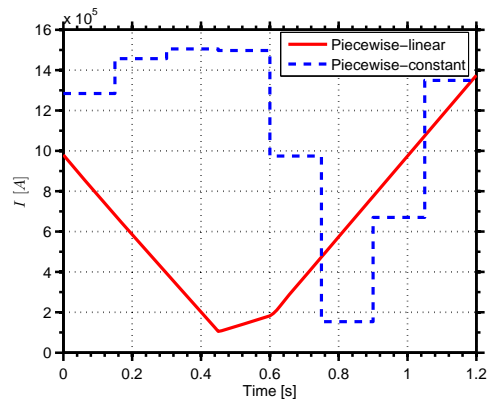
Using the MATLAB optimization software FMINCON, we implemented three different methods to solve Problem P: (i) piecewise-linear control parameterization with fixed partition points (method from [1]); (ii) piecewise-constant control parameterization with fixed partition points (method from [4]); and (iii) piecewise-linear control parameterization with variable partition points (the new method described in this paper). The solutions from methods (i) and (ii) are shown in Figure 1; the solution from method (iii) is shown in Figure 2. The optimal cost values are shown in Table 1. Note that our new method gives cost improvements of 7.6% and 14.4% over the methods in [1] and [4], respectively.

References

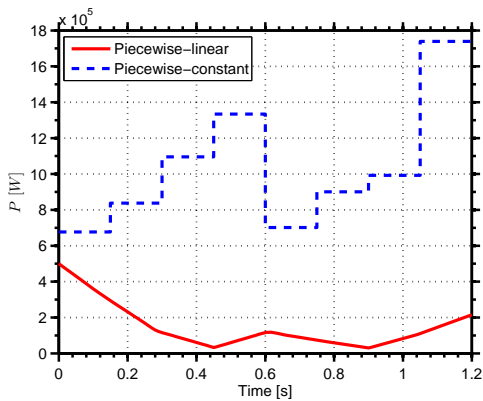
- [1] C. Xu, Y. Ou, J. Dalessio, E. Schuster, T. Luce, J. Ferron, M. Walker, and D. Humphreys, Ramp-up phase current-profile control of tokamak plasmas via nonlinear programming, *IEEE Transactions on Plasma Science*, 38(2): 163-173, 2010.
- [2] T. Taylor, Physics of advanced tokamaks, *Plasma Physics and Controlled Fusion*, 39(12B): B47-B79, 1997.
- [3] M. Murakami et al., Progress toward fully noninductive, high beta conditions in DIII-D, *Physics of Plasmas*, 13(5): 056106, 2006.
- [4] 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*, Cape Town, August 2014: 7892-7897.
- [5] R. Loxton, Q. Lin, 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.
- [6] 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.
- [7] 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.
- [8] M. S. Bazarara, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New Jersey: Wiley, 2006.



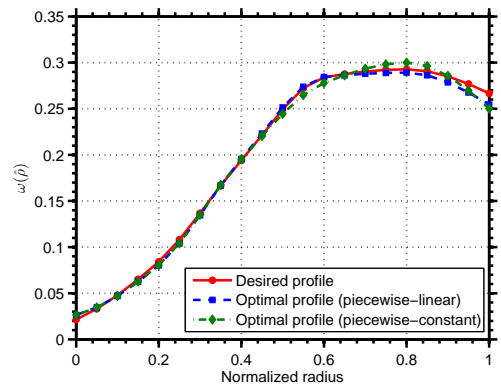
(a) Optimal control input $\bar{n}(t)$.



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

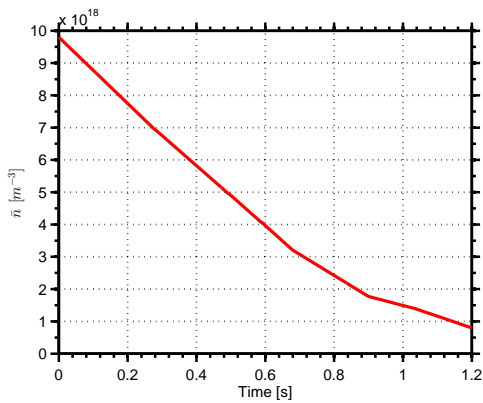


(c) Optimal control input $P(t)$.

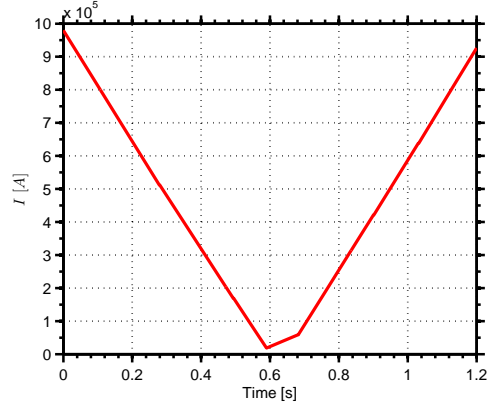


(d) Optimal ω -profile at the terminal time.

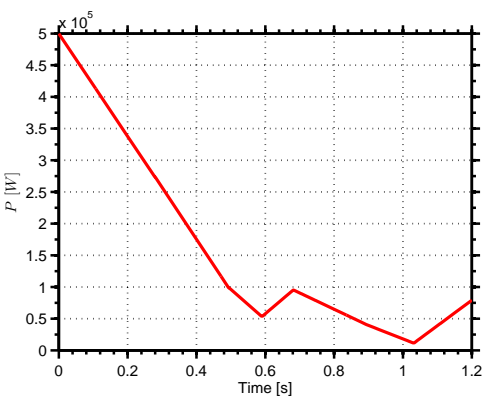
Fig. 1: Optimal solution in Section 4 for $p = 8$ (fixed partition points).



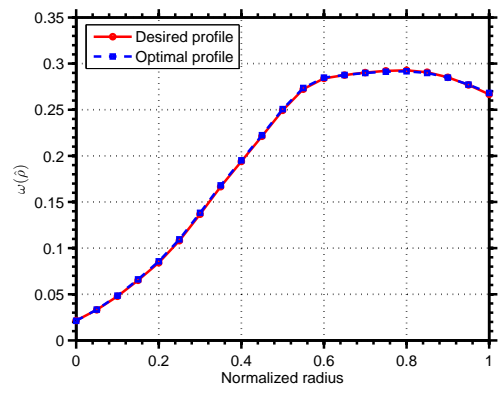
(a) Optimal control input $\bar{n}(t)$.



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



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



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

Fig. 2: Optimal solution in Section 4 for $p = 8$ (variable partition points).