Start-up phase plasma discharge design of a tokamak via control parameterization method*

Guo Shan(郭 珊)[a], Xu Ke(许 珂)[a], Xu Chao(许 超)[b]†, Ren Zhi-Gang(任志刚)[b], and Xiao Bing-Jia(肖炳甲)[c]

[a] *Department of Mathematics, Zhejiang University, Hangzhou* 310027, *China*
[b] *State Key Laboratory of Industrial Control Technology and Institute of Cyber-Systems & Control, Zhejiang University, Hangzhou* 310027, *China*
[c] *Institute of Plasma Physics, Chinese Academy of Sciences, Hefei* 230031, *China*

(Received 16 September 2014; revised manuscript received 29 October 2014; published online 19 January 2015)

The tokamak start-up is a very important phase during the process to obtain a suitable equalizing plasma, and its governing model can be described as a set of nonlinear ordinary differential equations (ODEs). In this paper, we first estimate the parameters in the original model and set up an accurate model to express how the variables change during the start-up phase, especially how the plasma current changes with respect to time and the loop voltage. Then, we apply the control parameterization method to obtain an approximate optimal parameters selection problem for the loop voltage design to achieve a desired plasma current target. Computational optimal control techniques such as the variational method and the costate method are employed to solve the problem, respectively. Finally, numerical simulations are performed and the results obtained via different methods are compared. Our numerical parameterization method and optimization procedure turn out to be effective.

Keywords: tokamak, start-up, optimal control, variational method

PACS: 52.55.Fa, 02.60.Pn **DOI:** 10.1088/1674-1056/24/3/035202

1. Introduction

Nowadays, nuclear fusion has been regarded as the most promising resource because it is a clean energy resource which has tremendous potential.[1] The tokamak is believed to be the most promising fusion reactor. In a tokamak, the toroidal field is produced by electromagnets that surround the torus, and the poloidal field is generated by a toroidal electric current that is carried by the plasma. This current is induced inside the plasma with secondary electromagnet sets.[2] The structure of a tokamak is shown in Fig. 1.

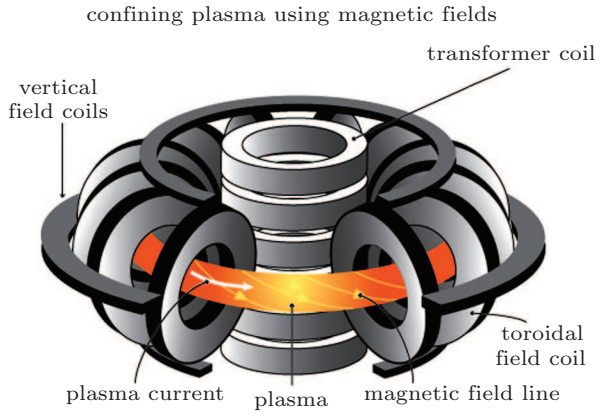


Fig. 1. (color online) Schematic diagram of a tokamak chamber.

Since the tokamak fusion reactor is a very complex device containing various connected subsystems with a large number of instabilities and during the working process it must deal with extremum temperature, high pressure, and great current, it has to face a lot of challenging control problems that urgently need to be solved. The control of the current profile in tokamak plasmas has been proved to be the enabler of advanced scenarios characterized by improved confinement, enhanced magnetohydrodynamic (MHD) stability, and possible steady-state operation.[3] There is a considerable interest in the start-up phase, as well as the tokamak breakdown process. This area is still far from completion.[4] In an operating fusion reactor, the electric field is applied to ionize the pre-fill gas and to ramp up the plasma current.[5] However, in the start-up of a reactor, either initially or after a temporary shutdown, the plasma will have to be heated to its operating temperature of greater than 10 keV. In designing the power supply system of a large tokamak, it is important to estimate the breakdown voltage and to decide the optimal voltage that we should supply to achieve successful start-up and a stable plasma current.[6] In this paper, we first study the properties of the zero-dimensional (0D) model, and in the future more complex models governed by partial differential equations (PDEs) can be used to generate more accurate results.

Now, we consider the most favourable situation that does not take into account impurities and escaping electrons. The model we use is described in Refs. [7]–[9], respectively. In this model, the term describing the average absorbed ECRH (electron cyclotron resonance heating) power is substituted into the

*Project supported by the National Natural Science Foundation of China (Grant Nos. 61104048 and 61473253) and the National High Technology Research and Development Program of China (Grant No. 2012AA041701).

†Corresponding author. E-mail: cxu@zju.edu.cn

© 2015 Chinese Physical Society and IOP Publishing Ltd

energy balance equation for electrons. The energy balance equation for ions also takes into account the absorbed additional ICRH (ion cyclotron resonance heating) power. The energy balance equations for electrons and ions in the 0D approximation are written as the following ordinary differential equations:

$$\begin{cases} \dfrac{3}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}(n_\mathrm{e}kT_\mathrm{e}) = P_\mathrm{OH} - P_\Delta - P_\mathrm{ion} - \dfrac{3}{2}\dfrac{n_\mathrm{e}kT_\mathrm{e}}{\tau_\mathrm{E}} + P_\mathrm{ECRH}, \\ \dfrac{3}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}(n_\mathrm{i}kT_\mathrm{i}) = P_\Delta - P_\mathrm{cx} - \dfrac{3}{2}\dfrac{n_\mathrm{i}kT_\mathrm{i}}{\tau_\mathrm{E}} + P_\mathrm{ICRH}. \end{cases} \tag{1}$$

The terms on the left-hand side denote the energy modification rates of electrons and ions. On the right-hand side, $(3/2)(n_\mathrm{e}kT_\mathrm{e}/\tau_\mathrm{E})$ and $(3/2)(n_\mathrm{i}kT_\mathrm{i}/\tau_\mathrm{E})$ are transport losses of energy and the other terms are the power inputs and power losses, whose specific meanings and expressions are listed in Table 1.

Table 1. Notations in the system.

	Meanings	Unit
n_e	concentrations of electrons	10^{20} m^{-3}
n_i	concentrations of ions, and $n_\mathrm{e} = n_\mathrm{i} = n$	10^{20} m^{-3}
n_0	concentration of neutrals	10^{20} m^{-3}
T_e	temperatures of electrons	keV
T_i	temperatures of ions	keV
k	1.381×10^{-23}, Boltzmann constant	J·K^{-1}
V_p	plasma column volume	m^3
V_v	vacuum vessel volume	m^3
P_OH	$I_\mathrm{p}^2R_\mathrm{p}/V_\mathrm{p}$, specific power of Ohmic heating	MW
R_p	$2\eta R/a^2$, plasma resistance	Ω
η	$1.65\times10^{-3}(\ln L_\mathrm{p})/T_\mathrm{e}^{1.5}$, Spitzer resistivity	Ω·m
P_Δ	$0.24[(T_\mathrm{e}-T_\mathrm{i})/T_\mathrm{e}^{1.5}]n_\mathrm{e}^2$, specific equilibration power between electrons and ions	MW
P_ion	$1.6\times10^{18}n_0n_\mathrm{e}S_\mathrm{i}W_\mathrm{i}$, specific energy losses through neutral gas ionization	MW
S_i	$2\times10^{-13}(T_\mathrm{e}/R_y)^{1/2}\exp(-R_y/T_\mathrm{e})(6+T_\mathrm{e}/R_y)^{-1}$ s, neutral gas ionization rate	m^3·s^{-1}
W_i	0.03	keV
R_y	0.0136	keV
P_cx	$2.4\times10^5n_\mathrm{e}n_0(T_\mathrm{i}-T_0)T_\mathrm{i}^{0.327}$ specific energy losses through charge exchange	MW
T_0	2.59×10^{-5}	keV
τ_E	energy confinement time	s
τ_p	particle confinement time, and assume that $\tau_\mathrm{E}=\tau_\mathrm{p}=\tau$	s
μ_0	$4\pi\times10^{-7}$, vacuum permeability	H·m^{-1}

The particle balance equations for electrons of concentration n_e and neutrals of concentration n_0 are governed by

$$\begin{cases} \dfrac{\mathrm{d}n_\mathrm{e}}{\mathrm{d}t} = n_0n_\mathrm{e}S_\mathrm{i} - \dfrac{n_\mathrm{e}}{\tau_\mathrm{p}}, \\ \dfrac{\mathrm{d}n_0}{\mathrm{d}t} = \dfrac{V_\mathrm{p}}{V_\mathrm{v}}\left(\dfrac{n_\mathrm{e}}{\tau_\mathrm{p}} - n_0n_\mathrm{e}S_\mathrm{i}\right). \end{cases} \tag{2}$$

Note that $\mathrm{d}n_\mathrm{e}/\mathrm{d}t$ and $\mathrm{d}n_0/\mathrm{d}t$ are the concentration modification rates of electrons and ions. $n_0n_\mathrm{e}S_\mathrm{i}$ allows for neutral gas ionization and $n_\mathrm{e}/\tau_\mathrm{p}$ is the transport loss of particles.

The electric circuit equation for the plasma current I_p is governed by

$$L_\mathrm{p}\frac{\mathrm{d}I_\mathrm{p}}{\mathrm{d}t} + R_\mathrm{p}I_\mathrm{p} = U, \tag{3}$$

which can be derived by Ohm's law directly. $L_\mathrm{p} = \mu_0R(\ln(8R/a)+l_\mathrm{i}/2-2)$ is the plasma column inductance and U is the loop voltage, which is considered as a control variable. Here, we set $l_\mathrm{i}\approx0.5$ (flat current profile), and the other notations used in Eqs. (1)–(3) are listed in Table 1.

In the system of Eqs. (1)–(3), k, R_y, W_i, T_0, l_i, μ_0 are physical constants and n, n_0, T_e, T_i, I_p are variables. In this paper, we will use the experimental parameters V_p, V_v, R, a, τ, L_p and the initial values of n, n_0, T_e, T_i, I_p which are parameters of the experimental advanced superconducting tokamak (EAST). EAST is the first shaped tokamak of mega-ampere scale to achieve plasma utilizing a fully superconducting poloidal field coil system located in Hefei, China.[10] Parameters of EAST that we have used are listed in Table 2.

Table 2. EAST parameters.

Variables	Value	Remarks
R	1.85 m	major radius
a	0.4 m	minor radius
P_ICRH	10 MW	ion cyclotron resonance heating power
P_ECRH	4 MW	electron cyclotron resonance heating power
I_p	1 MA	plasma current

2. Problem formulation and basic analysis in EAST

First of all, we rewrite Eqs. (1)–(3) as

$$\begin{cases} \dfrac{\mathrm{d}nT_\mathrm{e}}{\mathrm{d}t} = \alpha\dfrac{I_\mathrm{p}^2}{T_\mathrm{e}^{1.5}} - \beta\dfrac{T_\mathrm{e}-T_\mathrm{i}}{T_\mathrm{e}^{1.5}}n^2 - \gamma nn_0S_\mathrm{i}(T_\mathrm{e}) - \xi nT_\mathrm{e} \\ \qquad\triangleq f(n,n_0,T_\mathrm{e},T_\mathrm{i},I_\mathrm{p}), \\ \dfrac{\mathrm{d}nT_\mathrm{i}}{\mathrm{d}t} = \beta\dfrac{T_\mathrm{e}-T_\mathrm{i}}{T_\mathrm{e}^{1.5}}n^2 - \zeta nn_0(T_\mathrm{i}-T_0)T_\mathrm{i}^{0.327} - \xi nT_\mathrm{i} \\ \qquad\triangleq g(n,n_0,T_\mathrm{e},T_\mathrm{i}), \end{cases} \tag{4}$$

where $\alpha = 2.2\times10^3\ln L_\mathrm{p}/(kV_\mathrm{p})$, $\beta = 0.16k^{-1}$, $\gamma = 1.6\times10^{18}W_\mathrm{i}/1.5k$, $\xi=\tau^{-1}$, $\zeta = 1.6\times10^5k^{-1}$,

$$\begin{cases} \dfrac{\mathrm{d}n}{\mathrm{d}t} = n_0nS_\mathrm{i}(T_\mathrm{e}) - \dfrac{n}{\tau} \triangleq h(n,n_0,T_\mathrm{e}), \\ \dfrac{\mathrm{d}n_0}{\mathrm{d}t} = -\dfrac{V_\mathrm{p}}{V_\mathrm{v}}\left(n_0nS_\mathrm{i}(T_\mathrm{e}) - \dfrac{n}{\tau}\right) \triangleq -\dfrac{V_\mathrm{p}}{V_\mathrm{v}}h(n,n_0,T_\mathrm{e}), \end{cases} \tag{5}$$

and

$$L_\mathrm{p}\frac{\mathrm{d}I_\mathrm{p}}{\mathrm{d}t} + R_\mathrm{p}(T_\mathrm{e})I_\mathrm{p} = U, \tag{6}$$

where $R_\mathrm{p}(T_\mathrm{e}) = 3.3\times10^3R(\ln L_\mathrm{p})/a^2T_\mathrm{e}^{1.5}$.

Because $n_0(n) = -(V_\mathrm{p}/V_\mathrm{v})n + C$, we then remove the equation for the ion concentration $\mathrm{d}n_0$, to obtain a new system including four equations and transform the system (1)–(3)

to the following compact form:

$$\begin{cases} \dfrac{\mathrm{d}n}{\mathrm{d}t} = h(n, n_0(n), T_\mathrm{e}) = F_1(n, T_\mathrm{e}), \\ \dfrac{\mathrm{d}T_\mathrm{e}}{\mathrm{d}t} = \dfrac{f(n, n_0(n), T_\mathrm{e}, T_\mathrm{i}, I_\mathrm{p}) - T_\mathrm{e} h(n, n_0(n), T_\mathrm{e})}{n} \\ \qquad = F_2(n, T_\mathrm{e}, T_\mathrm{i}, I_\mathrm{p}), \\ \dfrac{\mathrm{d}T_\mathrm{i}}{\mathrm{d}t} = \dfrac{g(n, n_0(n), T_\mathrm{e}, T_\mathrm{i}) - T_\mathrm{i} h(n, n_0(n), T_\mathrm{e})}{n} \\ \qquad = F_3(n, T_\mathrm{e}, T_\mathrm{i}), \\ \dfrac{\mathrm{d}I_\mathrm{p}}{\mathrm{d}t} = \dfrac{U - R_\mathrm{p}(T_\mathrm{e}) I_\mathrm{p}}{L_\mathrm{p}} = F_4(T_\mathrm{e}, I_\mathrm{p}, U). \end{cases} \tag{7}$$

For notational simplicity, we let $\boldsymbol{X}(t) = [n(t), T_\mathrm{e}(t), T_\mathrm{i}(t), I_\mathrm{p}(t)]^\mathrm{T} \in \mathbb{R}^4$, then the system can be expressed as

$$\dot{\boldsymbol{X}}(t) = \boldsymbol{F}(t, \boldsymbol{X}(t), U), \quad t \geq 0. \tag{8}$$

To figure out how the system works we first calculate values of the four variables for $t \in [0, 2]$. Figure 2 shows how the variables change when $U = 6$ V, $P_\mathrm{ECRH} = 4$ MW, $P_\mathrm{ICRH} = 10$ MW with initial values $[n\ T_\mathrm{e}\ T_\mathrm{i}\ I_\mathrm{p}] = [2 \times 10^{-5}\ 1 \times 10^{-3}\ 3 \times 10^{-5}\ 1]$. It is obvious that in this situation I_p will increase rapidly and reach 1×10^6 A at about 0.904 s. Now, we focus on determining the control of the loop voltage U to optimize other variables, particularly the plasma current I_p. In this paper, we only discuss a simple situation in which we separate the whole progress into two steps. The first step is to set $U = 6$ V which is actually the upper bound and $P_\mathrm{ECRH} = 4$ MW, $P_\mathrm{ICRH} = 10$ MW until $I_\mathrm{p} = 1 \times 10^6$ A at the final time instant $t = 0.904$ s. The second step is to control U so that I_p can stay at 1×10^6 A. Since the first step is stable, the main control purpose is to maintain the plasma current after $t = 0.904$ s. For convenience, we reset the initial time $t = 0$ s at the beginning of the second step, as well as initial condition

$$\boldsymbol{X}(0) = \boldsymbol{X}_0 = [0.001195, 4989, 380.6, 1 \times 10^6]^\mathrm{T}. \tag{9}$$

Thus, the altered problem is to minimize the following objective function:

$$J(U) = \int_0^{t_\mathrm{f}} [(I_\mathrm{p}(t) - I_\mathrm{p}^*(t)]^2 \mathrm{d}t + \lambda \int_0^{t_\mathrm{f}} U^2(t) \mathrm{d}t. \tag{10}$$

The control U must satisfy $0\ \mathrm{V} \leq U \leq 6\ \mathrm{V}$ and $I_\mathrm{p}^* = 1 \times 10^6$ A, t_f is the final time, λ is the weight of control cost. Now, we may state our problem formally as follows.

Problem 1 Given the dynamic system (8) with initial condition (9), find an admissible control $U \in [0\ \mathrm{V}, 6\ \mathrm{V}]$ such that the cost functional (10) is minimized.

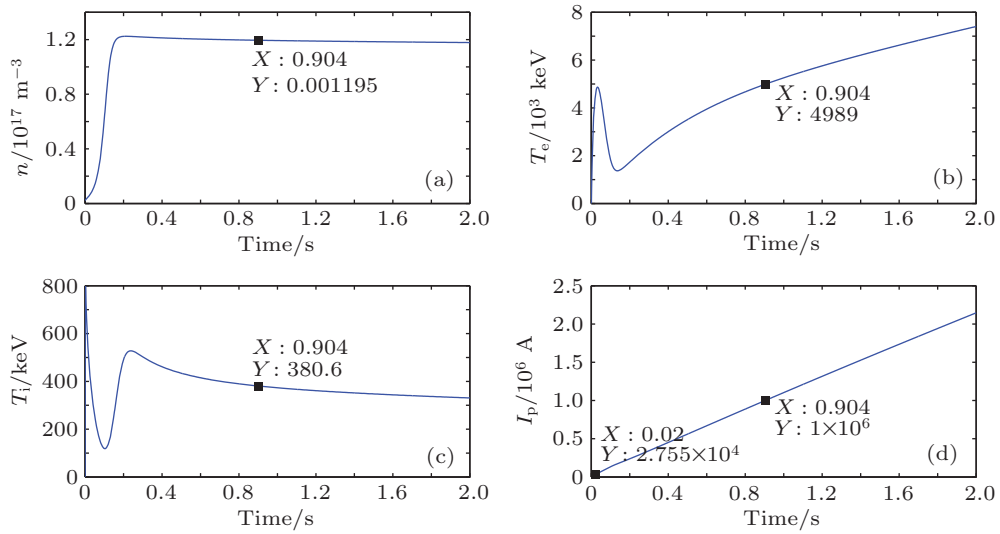


Fig. 2. (color online) Variable values in the first 2 s ($U = 6$ V). (a) n, (b) T_e, (c) T_i, and (d) I_p.

3. Numerical solution methods

3.1. Control parameterization

To solve Problem 1, we will apply the control parametrization method introduced in Refs. [11]–[13]. Here we use two different methods.

Method 1 The control is approximated by a piecewise constant function with possible discontinuities at a set of fixed switching points.

The values of the piecewise constant functions then become decision variables to be selected optimally. In this paper, we will divide the time horizon $[0, t_\mathrm{f}]$ into 10 subintervals $[t_{k-1}, t_k)$, $k = 1, 2, 3, \ldots, 10$ and $t_0, t_2, \ldots, t_{10}$ is equally placed time points such that

$$0 = t_0 < t_1 < t_2 < \cdots < t_9 < t_{10} = t_\mathrm{f}, \tag{11}$$

and the control function U is approximated by a constant value

on each subinterval denoted by σ^k, then U can be expressed as

$$U(t) \approx \sum_{k=1}^{10} \sigma^k \chi_{[t_{k-1},t_k)}(t), \tag{12}$$

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

$$\chi_{[t_{k-1},t_k)}(t) = \begin{cases} 1, & \text{if } t \in [t_{k-1},t_k), \\ 0, & \text{otherwise}, \end{cases} \tag{13}$$

and σ^k satisfies $0\ \text{V} \le \sigma^k \le 6\ \text{V}$. For convenience, we use $\boldsymbol{\sigma}$ to denote $[\sigma^1, \sigma^2, \ldots, \sigma^{10}]$.

When we finish the control parameterization procedure, the dynamic system (8) becomes

$$\begin{aligned} \dot{\boldsymbol{X}}(t) &= \sum_{k=1}^{10} \boldsymbol{F}(t, \boldsymbol{X}(t), \sigma^k) \chi_{[t_{k-1},t_k)}(t) \\ &= \hat{\boldsymbol{F}}_1(t, \boldsymbol{X}(t), \boldsymbol{\sigma}), \quad t \in [0, t_\text{f}], \end{aligned} \tag{14}$$

and the cost functional (10) becomes

$$\begin{aligned} \hat{J}_1(\boldsymbol{\sigma}) &= \int_0^{t_\text{f}} [I_\text{p}(t|\boldsymbol{\sigma}) - I_\text{p}^*(t)]^2 \,\text{d}t \\ &\quad + \lambda \sum_{k=1}^{10} \int_0^{t_\text{f}} [\sigma^k \chi_{[t_{k-1},t_k)}(t)]^2 \,\text{d}t. \end{aligned} \tag{15}$$

Now we can state the approximate parameter selection problem as follows.

Problem 2 Given the dynamic system (14) with initial condition (9), find admissible control parameter vector $\boldsymbol{\sigma} \in [0\ \text{V}, 6\ \text{V}]$ such that the cost functional (15) is minimized.

Method 2 The control is approximated by a polynomial $U(t) = \xi_1 t^3 + \xi_2 t^2 + \xi_3 t + \xi_4$, where $\xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{R}$, are the parameters to be optimized.

Now ξ_1, ξ_2, ξ_3, ξ_4 are decision variables to be selected optimally. The dynamic system (8) becomes

$$\begin{aligned} \dot{\boldsymbol{X}}(t) &= \boldsymbol{F}(t, \boldsymbol{X}(t), U), \\ &= \hat{\boldsymbol{F}}_2(t, \boldsymbol{X}(t), \xi_1, \xi_2, \xi_3, \xi_4), \quad t \in [0, t_\text{f}], \end{aligned} \tag{16}$$

and the cost functional (10) becomes

$$\begin{aligned} \hat{J}_2 &= \int_0^{t_\text{f}} [I_\text{p}(t) - I_\text{p}^*(t)]^2 \,\text{d}t \\ &\quad + \lambda \int_0^{t_\text{f}} (\xi_1 t^3 + \xi_2 t^2 + \xi_3 t + \xi_4)^2 \,\text{d}t. \end{aligned} \tag{17}$$

When using Method 2, we must ensure that $U(t) \in [0\ \text{V}, 6\ \text{V}]$. Therefore, we add some inequality constraints generated by the sampling method. In this paper, we will divide the time horizon $[0, t_\text{f}]$ into 20 subintervals $[t_{j-1}, t_j), j = 1, 2, 3, \ldots, 20$ and $t_0, t_2, \ldots, t_{20}$ are equally placed time points. Then the inequality constraints can be described as $0\ \text{V} \le \xi_1 t_j^3 + \xi_2 t_j^2 + \xi_3 t_j + \xi_4 \le 6\ \text{V}$ where $j = 1, 2, \ldots, 20$.

Now, we can state the approximate parameter selection problem as follows.

Problem 3 Given the dynamic system (16) with initial condition (9), find admissible control parameters ξ_1, ξ_2, ξ_3, ξ_4 such that the cost functional (17) is minimized.

However, in some situation, it will be more convenient for the numerical computation if we transform the integral part of the cost functional to a terminal state of the dynamic system. Here, we define a new variable X^J satisfying

$$\begin{cases} \dfrac{\text{d}X^J}{\text{d}t} = (I_\text{p} - I_\text{p}^*)^2 + \lambda U^2, \\ X^J(0) = 0. \end{cases} \tag{18}$$

Notice that $X^J(t_\text{f})$ equals cost functional J in Eq. (10), and we add X^J to the dynamic system (8). We introduce the following new state variable vector:

$$\begin{cases} \boldsymbol{X}_n(t) = [n(t), T_\text{e}(t), T_\text{i}(t), I_\text{p}(t), X^J(t)]^\text{T} \in \mathbb{R}^5, \\ \dot{\boldsymbol{X}}_n(t) = \boldsymbol{F}_n(t, \boldsymbol{X}_n(t), U), \quad t \in [0, t_\text{f}], \end{cases} \tag{19}$$

with the initial condition

$$\boldsymbol{X}_n(0) = [1.195 \times 10^{-3}, 4989, 380.6, 1 \times 10^6, 0]^\text{T}. \tag{20}$$

Then, the cost functional (10) becomes

$$J_n = X^J(t_\text{f}). \tag{21}$$

Now we can state the approximate parameter selection problem as follows.

Problem 4 Given the dynamic system (19) with initial condition (20), find admissible controls $U \in [0\ \text{V}, 6\ \text{V}]$ such that the cost functional (21) is minimized.

Similarly we can apply Methods 1 and 2 to Problem 4 to obtain other approximate problems by changing the dynamic systems (14) and (16) into

$$\dot{\boldsymbol{X}}_{n1}(t) = \boldsymbol{F}_{n1}(t, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma}), \tag{22}$$

$$\dot{\boldsymbol{X}}_{n2}(t) = \boldsymbol{F}_{n2}(t, \boldsymbol{X}_{n2}(t), \xi_1, \xi_2, \xi_3, \xi_4). \tag{23}$$

Meanwhile, we change the cost functional Eqs. (15) and (17) into a terminal time representation

$$J_{n1} = X^J(t_\text{f}), \tag{24}$$

$$J_{n2} = X^J(t_\text{f}). \tag{25}$$

Problem 5 Given the dynamic system (22) with initial condition (20), find admissible control parameter vector $\boldsymbol{\sigma} \in [0\ \text{V}, 6\ \text{V}]$ such that the cost functional (24) is minimized.

Problem 6 Given the dynamic system (23) with initial condition (20), find admissible control parameters ξ_1, ξ_2, ξ_3, ξ_4 such that the cost functional (25) is minimized.

3.2. Variational method

As problems 2, 3, 5, and 6 are optimal parameter selection problems in a canonical form then we can solve them as nonlinear optimization problems using the active-set[14] or SQP method.[15] Algorithm 1 below is the optimization algorithm for solving the approximate problem.

Algorithm 1

Step 1 Give an initial guess;

Step 2 Compute objective value and gradient value at the current points;

Step 3 If it is optimal, algorithm ends. Otherwise, go to Step 4;

Step 4 Calculate search direction;

Step 5 Determine optimal step length;

Step 6 Generate new parameter vector, then go back to Step 2.

There is a nonlinear optimization solver FMINCON in MATLAB which can automatically carry out Steps 3–6. In spite of this, we need to supply the gradient formula of the cost functional. Here, we solve Problem 5 using the variational method.[16] We define the state variation with respect to $\boldsymbol{\sigma}$ as follows:

$$\boldsymbol{\Gamma}(t|\boldsymbol{\sigma}) = \frac{\partial \boldsymbol{x}(t|\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}. \tag{26}$$

From Eq. (22), we have

$$\boldsymbol{X}_{n1}(t) = \boldsymbol{X}_{n1}(0) + \int_0^t \boldsymbol{F}_{n1}(s, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma})\,\mathrm{d}s. \tag{27}$$

Thus, we take the derivative with respect to $\boldsymbol{\sigma}$ to obtain

$$\begin{aligned}\frac{\partial \boldsymbol{X}_{n1}(t)}{\partial \boldsymbol{\sigma}} &= \frac{\partial \boldsymbol{X}_{n1}(0)}{\partial \boldsymbol{\sigma}} + \int_0^t \frac{\partial \boldsymbol{F}_{n1}(s, \boldsymbol{X}_{n1}(s), \boldsymbol{\sigma})}{\partial \boldsymbol{X}_{n1}} \frac{\partial \boldsymbol{X}_{n1}(s)}{\partial \boldsymbol{\sigma}}\,\mathrm{d}s \\ &\quad + \int_0^t \frac{\partial \boldsymbol{F}_{n1}(s, \boldsymbol{X}_{n1}(s), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}\,\mathrm{d}s. \end{aligned} \tag{28}$$

By differentiating with respect to time, we can obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{\partial \boldsymbol{X}_{n1}}{\partial \boldsymbol{\sigma}}\right\} &= \frac{\partial \boldsymbol{F}_{n1}(t, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma})}{\partial \boldsymbol{X}_{n1}} \frac{\partial \boldsymbol{X}_{n1}(t)}{\partial \boldsymbol{\sigma}} \\ &\quad + \frac{\partial \boldsymbol{F}_{n1}(t, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}. \end{aligned} \tag{29}$$

Thus, $\boldsymbol{\Gamma}(\cdot|\boldsymbol{\sigma})$ is the solution of the following variational system:

$$\begin{aligned}\dot{\boldsymbol{\Gamma}}(t|\boldsymbol{\sigma}) &= \frac{\partial \boldsymbol{F}_{n1}(t, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma})}{\partial \boldsymbol{X}_{n1}} \boldsymbol{\Gamma}(t|\boldsymbol{\sigma}) \\ &\quad + \frac{\partial \boldsymbol{F}_{n1}(t, \boldsymbol{X}_{n1}(t), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}, \end{aligned} \tag{30}$$

with the initial condition

$$\boldsymbol{\Gamma}(0) = \frac{\partial \boldsymbol{X}_{n1}(0)}{\partial \boldsymbol{\sigma}} = 0. \tag{31}$$

Now we can use the chain rule to compute the gradient of the cost functional with respect to the decision variables

$$\frac{\partial J_{n1}(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} = \frac{\partial X^J_{n1}(T|\boldsymbol{\sigma})}{\partial \boldsymbol{X}_{n1}} \boldsymbol{\Gamma}(T|\boldsymbol{\sigma}) + \frac{\partial X^J_{n1}(T|\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}. \tag{32}$$

By integrating with the nonlinear programming algorithm in MATLAB using this gradient, then Problem 5 can be solved numerically.

3.3. Costate method

Next we introduce the costate method to compute the cost functional gradient. Here we compute Problems 2 and 3 using this method.

According to classical optimal control theory,[11] the Hamiltonian of Problem 2 can be defined by

$$\begin{aligned} &H(t, \boldsymbol{X}(t), \boldsymbol{\sigma}, \boldsymbol{\mu}) \\ &= (I_{\mathrm{p}}(t|\boldsymbol{\sigma}) - I^*_{\mathrm{p}})^2 + \lambda \sum_{k=1}^{10} \left[\sigma^k \chi_{[t_{k-1}, t_k)}(t)\right]^2 \\ &\quad + \boldsymbol{\mu}^{\mathrm{T}} \hat{F}_1(t, \boldsymbol{X}(t), \boldsymbol{\sigma}), \end{aligned} \tag{33}$$

where $\boldsymbol{\mu} \in \mathbb{R}^4$ is called the costate vector. Using the maximum principle, the costate system is defined by

$$\begin{cases} \dot{\boldsymbol{\mu}}(t) = -\left[\dfrac{\partial H(t, \boldsymbol{X}(t|\boldsymbol{\sigma}), \boldsymbol{\sigma}, \boldsymbol{\mu}(t))}{\partial \boldsymbol{X}}\right]^{\mathrm{T}}, \quad t \in [0, t_{\mathrm{f}}], \\ \boldsymbol{\mu}(t_{\mathrm{f}}) = \boldsymbol{0}. \end{cases} \tag{34}$$

Let $\boldsymbol{\mu}(\cdot|\boldsymbol{\sigma})$ denote the solution of this costate system, then we can compute the cost functional gradient

$$\frac{\partial \hat{J}_1(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} = \int_0^{t_{\mathrm{f}}} \frac{\partial H(t, \boldsymbol{X}(t|\boldsymbol{\sigma}), \boldsymbol{\sigma}, \boldsymbol{\mu}(t|\boldsymbol{\sigma}))}{\partial \boldsymbol{\sigma}}\,\mathrm{d}t. \tag{35}$$

We can use Simpson's rule to evaluate this integral term in Eq. (35). Let q be an even integer. Choose a partition $\{t_k\}_{k=0}^q$ such that

$$t_0 = 0, \quad t_k - t_{k-1} = h, \quad t_q = t_{\mathrm{f}},$$

then

$$\begin{aligned} &\int_0^{t_{\mathrm{f}}} \frac{\partial H(t)}{\partial \boldsymbol{\sigma}}\,\mathrm{d}t \\ &\approx \frac{h}{3}\left[\frac{\partial H(t_0)}{\partial \boldsymbol{\sigma}} + 4\frac{\partial H(t_1)}{\partial \boldsymbol{\sigma}} + 2\frac{\partial H(t_2)}{\partial \boldsymbol{\sigma}} + 4\frac{\partial H(t_3)}{\partial \boldsymbol{\sigma}}\right. \\ &\quad \left. + 2\frac{\partial H(t_4)}{\partial \boldsymbol{\sigma}} + \cdots + 4\frac{\partial H(t_{q-1})}{\partial \boldsymbol{\sigma}} + \frac{\partial H(t_q)}{\partial \boldsymbol{\sigma}}\right]. \end{aligned} \tag{36}$$

The Hamiltonian corresponding to Problem 3 is defined as

$$\begin{aligned} H(t, \boldsymbol{X}, \xi, \boldsymbol{\mu}) &= (I_{\mathrm{p}} - I_{P_0})^2 + \lambda(\xi_1 t^3 + \xi_2 t^2 + \xi_3 t + \xi_4)^2 \\ &\quad + \boldsymbol{\mu}^{\mathrm{T}} \hat{\boldsymbol{F}}_2(t, \boldsymbol{X}(t), \xi_1, \xi_2, \xi_3, \xi_4), \end{aligned} \tag{37}$$

and the rest of the computation is similar to Eqs. (34)–(36).

With the computation of gradient of the cost functional, Problems 2 and 3 can be solved numerically.

4. Comparison of results

4.1. Method 1

In simulation, we set the time interval $[0,1]$ and equally divide it into 10 subintervals. Then, we optimize the loop voltage $U(l), l = 1,2,\ldots,10$. The numerical results of Method 1 are shown in Figs. 3 and 4. The curve of I_p during the whole start-up phase (both the first part in which $U = 6$ V and the second part in which optimal control is applied are included) is shown in Fig. 5. From these figures we can see that the optimal control is quite effective.

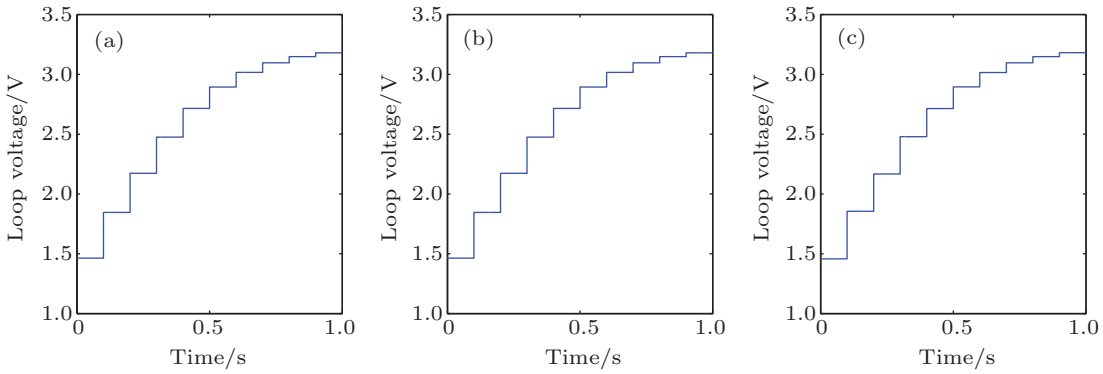


Fig. 3. (color online) Optimization results of piecewise-constant loop voltage. (a) Optimation without gradient, (b) optimation with gradient in variational method, (c) optimation with gradient in costate method.

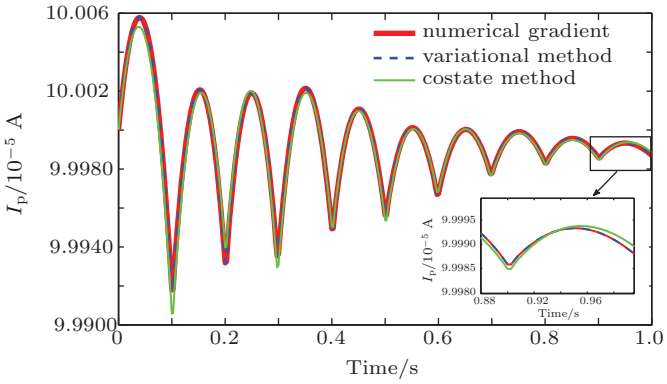


Fig. 4. (color online) Optimization results of I_p with respect to piecewise-constant loop voltage.

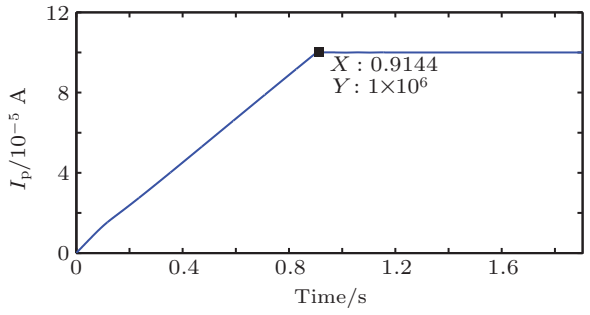


Fig. 5. (color online) Optimization results of I_p during the whole period.

We can see the I_p curves fluctuate around the desired value (1 MA) and tend to be stable with the time increases. We set the same initial guess for the loop voltage parameters σ_0^k, $k = 1,2,\ldots,10$. The results of parameters obtained using three different methods, i.e., numerical gradient (NG), variational method (VM), and costate method (CM) are almost the same, while the computational time is significantly different. Note that the NG method means we do not supply the gradient of the objective function and the software compute gradients using finite differences. Table 3 illustrates the comparison of computing time and optimized results. The costate method (CM) spends the least time to obtain the similar minimum object value.

Table 3. M1 computation time and optimized results.

Method	Object value	Time/s
NG	47291.7	2103.53
VM	47291.8	321.13
CM	48628.9	61.56

The advantage of M1 is that the loop voltage is piecewise-constant, which is easy to control. However, the smoothness of I_p cannot be ensured and the computation time is relatively long.

4.2. Method 2

The loop voltage is in the parameterization form of 3-order polynomial about time t. Then, it can be expressed as $U(t) = \xi_1 t^3 + \xi_2 t^2 + \xi_3 t + \xi_4$, where ξ_1, ξ_2, ξ_3, and ξ_4 are optimal parameters. The numerical calculations of Method 2 are shown in Figs. 6 and 7. Table 4 illustrates the comparison of computing time and optimized results. The weight in the figures stands for the weight (λ) of the control cost in the cost functional. It has influence on the optimal cost functional value, but has little impact on the optimal control parameters.

Because there are only four parameters, the computing time is much less than that of Method 1 and the two methods (CM and VM) have some minor differences in values and time.

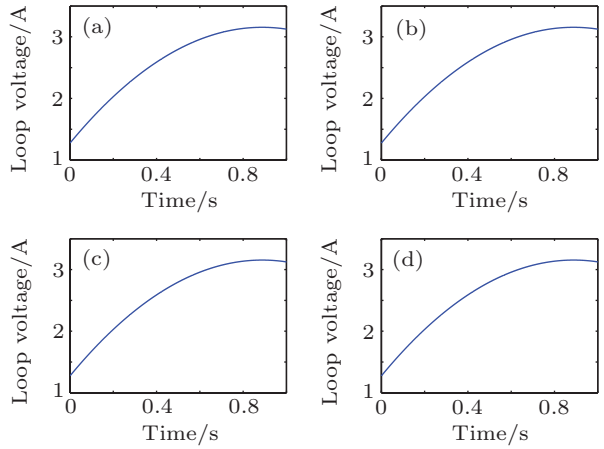


Fig. 6. (color online) Optimization results of 3-order polynomial loop voltage. (a) VM with weight = 0, (b) VM with weight = 2000, (c) CM with weight = 0, (d) CM with weight = 2000.

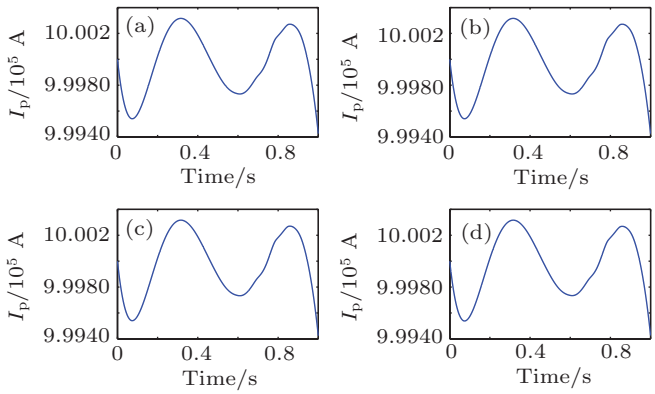


Fig. 7. (color online) Optimization results of I_p with respect to 3-order polynomial loop voltage. (a) Variational method with weight = 0, (b) variational method with weight = 2000, (c) costate method with weight = 0, (d) costate method with weight = 2000.

Table 4. M2 computing time and optimized results.

Method	Object value	Time/s
VM weight=0	56229.8	11.73
VM weight=2000	70357.6	12.35
CM weight=0	56869.1	10.99
CM weight=2000	71029.4	10.51

Though the object value of M2 is slightly larger than M1, it is much faster.

5. Conclusion and remark

We present an effective computational method for plasma current design during the start-up phase in a tokamak. The simulation results listed above have demonstrated that both of the control parametrization methods are effective. Besides, during the simulation progress, we have compared the variational method with the costate method to compute the cost functional gradient. It turns out that the costate method is much faster than the variational method while the final cost functional is quite similar, especially when we use Method 1, which spends longer time since it has more parameters to be decided and more numerical ODEs to be computed. Even though, there is still a great deal of room for improvement. For instance, in this paper we have designed the current curve during the whole time interval. However, in many situations, we only care about the states of the terminal time and do not need to track a fixed current curve. In our future work, we will consider another situation in which we only set the terminal state constraints and it will generate the optimal current curve automatically without any additional constraints.

Remark 1 While using the variational method, it may need to compute the dynamic system and the variational system separately. This means that we need to compute the dynamic system first and then use the interpolation of the variable values to compute the variational system, because the ODEs may be singular and give wrong results if we put them together.

Remark 2 When it comes to the costate method, $X^J(t)$ is not suitable to be added, because we will encounter a problem when computing $\partial H/\partial X^J$.

References

[1] Schultz K R 2006 *IEEE Control Syst.* **26** 32
[2] Mukhovatov V S and Shafranov V D 1971 *Nucl. Fusion* **11** 605
[3] Xu C, Ou Y, Dalessio J, Schuster E, Luce T C, Ferron J R, Walker M L and Humphreys D A 2010 *IEEE Trans. Plasma Sci.* **38** 163
[4] Knoepful H 1985 *Tokamak Start-up: Problems and Scenarios Related to the Transient Phases of a Thermonuclear Fusion Reactor* (New York: Springer) pp. 7–43
[5] Papoular R 1976 *Nucl. Fusion* **16** 37
[6] Ikeda K 2007 *Nucl. Fusion* **47** preface
[7] Lloyd B, Carolan P G and Warrick C D 1996 *Plasma Phys. Control. Fusion* **38** 1627
[8] Azizov E A, Barkalov A D, Gladush G G, *et al.* 2003 *Problems of Atomic Science and Technology. Series: Plasma Physics* **1** 49
[9] Barkalov A D and Gladush G G 2007 *Plasma Devices and Operations* **15** 185
[10] Leuer J A, Xiao B J, Humphreys D A, Walker M L, Hyatt A W, Jackson G L, Mueller D, Penaflor B G, Piglowski D A, Johnson R D and others 2010 *Fusion Science and Technology* **57** 48
[11] Teo K L, Goh C J and Wong K H 1991 *A Unified Computational Approach to Optimal Control Problems* (New York: Longman Scientific and Technical) pp. 99–122
[12] Lin Q, Loxton R and Teo K L 2013 *Journal of the Operations Research Society of China* **1** 275
[13] Yu X, Ren Z G and Xu C 2014 *Chin. Phys. B* **23** 040201
[14] Byrd R H, Gilbert J C and Nocedal J 2000 *Math. Program.* **89** 149
[15] Nocedal J and Wright S J 2006 *Numerical Optimization* (New York: Springer) pp. 526–572
[16] Loxton R, Teo K L and Rehbock V 2008 *Automatica* **44** 2923