POD-Based Optimal Control of Current Profile in Tokamak Plasmas via Nonlinear Programming

C. Xu, J. Dalessio, Y. Ou, E. Schuster, T.C. Luce, J.R. Ferron, M.L. Walker and D.A. Humphreys

***Abstract*— In a magnetic fusion reactor, the achievement of a certain type of plasma current profiles, which are compatible with magnetohydrodynamic (MHD) stability at high plasma pressure, is key to enabling high fusion gain and noninductive sustainment of the plasma current for steady-state operation. The evolution in time of the current profile is related to the evolution of the spatial derivative of the poloidal flux profile, which is modeled in normalized cylindrical coordinates using a partial differential equation (PDE) usually referred to as the magnetic diffusion equation. The dynamics of the plasma poloidal flux profile can be modified by three actuators: the total plasma current, the non-inductive power, and the average plasma density. These three actuators, which are constrained not only in value and rate but also in their initial and final values, are used to drive the poloidal flux profile, or equivalently the current profile, as close as possible to a desired target profile at a specific final time. To solve this constrained finite-time open-loop optimal control problem, model reduction based on proper orthogonal decomposition (POD) is combined with sequential quadratic programming (SQP) in an iterative fashion. The use of a low dimensional dynamical model reduces the computational effort, and therefore the time required to solve the optimization problem, which is critical for a potential implementation of a receding horizon control strategy.**

I. Introduction

Nuclear fusion is the process by which two nuclei fuse together to form a heavier nucleus. This process is accompanied by a release of energy, which is the result of the mass "lost" in the reaction. The amount of released energy is given by Einstein's famous equation (derived in 1905 as a part of his special theory of relativity), $E = (M_r - M_p)c^2$, where E is the energy, M_r the mass of the reactant nuclei, M_p the mass of the product nuclei, and c the speed of light.

To make a fusion reaction possible, a certain amount of energy is required to bring two repellant nuclei carrying positive charges sufficiently close. To overcome the Coulomb barrier, the kinetic energy of the nuclei is increased by heating. The temperature required for a thermonuclear fusion reaction to take place is around 100 million degrees. At much lower temperatures (about 10 thousand degrees), the electrons and nuclei separate and create an ionized gas called plasma, also known as the fourth state of matter.

This work was supported in part by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA), the NSF CAREER award program (ECCS-0645086), and DoE contract number DE-FC02-04ER54698. C. Xu (chx205@lehigh.edu), J. Dalessio, Y. Ou and E. Schuster are with the Department of Mechanical Engineering and Mechanics, Lehigh University, 19 Memorial Drive West, Bethlehem, PA 18015, USA. T. C. Luce, J. R. Ferron, M.L. Walker and D.A. Humphreys are with the DIII-D tokamak, General Atomics, San Diego, California, USA.

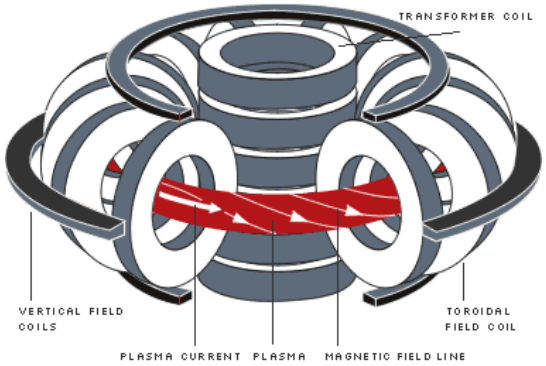

Fig. 1. The geometry of a tokamak device. Source: Max-Planck-Institut für Plasmaphysik (IPP).

An intangible doughnut-shaped bottle created by magnetic lines is used to confine the high-temperature plasma. One type of magnetic confinement device is called Tokamak, an acronym for the Russian words Toroidalnaya Kamera ee Magnitnaya Katushka (toroidal chamber with magnetic coils), which was invented in the Soviet Union in the late 1950s. An assembly of (toroidal) coils produces a magnetic field in the direction of the torus, to which is added the (poloidal) magnetic field created by an intense axial current flowing in the plasma itself (Fig. 1). A combined field is therefore created, in which the magnetic field lines twist their way around the tokamak to form a helical structure. It is possible to use the poloidal component B_{pol} of the helicoidal magnetic lines to define nested toroidal surfaces corresponding to constant values of the poloidal magnetic flux. The poloidal flux at a point P is the total magnetic flux through the surface S bounded by the toroidal ring passing through P, i.e., $\int B_{pol} dS$.

In a tokamak, the achievement of a suitable (toroidal) current profile plays an important role in enabling high fusion gain and noninductive sustainment of the plasma current for steady-state operation (see, e.g., [1], [2], [3]). The evolution in time of the current profile is related to the evolution of the spatial derivative of the poloidal flux profile. Therefore, we can control the current profile by controlling the poloidal flux profile. The time evolution of the poloidal flux profile is modeled in normalized cylindrical coordinates using a partial differential equation (PDE) usually referred to as the magnetic diffusion equation. The dynamics of the plasma poloidal flux profile can be modified by three actuators: the total plasma current, the non-inductive power, and the average plasma density. These physical actuators enter the magnetic diffusion equation as interior, boundary, and diffusivity control terms.

One possible approach to current profile control focuses on creating the desired current profile during the plasma current ramp-up and early flattop phases with the aim of maintaining this target profile during the subsequent phases of the discharge. Since the actuators that are used to achieve the desired target profile are constrained, experiments have shown that some of the desirable target profiles may not be achieved for all arbitrary initial conditions. Therefore, a perfect matching of the desirable target profile may not be physically possible. In practice, the objective is to achieve the best possible approximate matching at a prespecified time t_f during the early flattop phase of the total plasma current pulse. Thus, such matching problem can be formulated as a finite-time optimal control problem for the magnetic diffusion PDE [4].

Optimal control of PDE systems has been studied extensively (e.g., [5], [6], [7] and references therein), including both direct and indirect approaches. The indirect method studies the necessary condition for optimality (i.e., the Pontryagin maximum principle) and the corresponding numerical algorithms used for the minimization of the associated cost functionals, while the direct method treats system states and controls as independent variables to view the original optimal control problems as constrained PDE-based optimization problems, which can be handled by the sequential quadratic programming (SQP) method [8].

To overcome the high dimensionality of the problem, reduced order modeling (ROM) techniques play a very important role in dealing with control of infinite dimensional dynamical systems. The Proper Orthogonal Decomposition (POD) method is an efficient ROM technique used to obtain low dimensional dynamical systems (LDDS's) from data ensembles that arise from numerical simulation or experimental observation. The POD method has been widely used and proved successful to discover coherent structures from complex physical processes (e.g., [9], [10]) and to control PDE systems (e.g., [11]).

In this paper, we combine POD and SQP to compute the open-loop optimal control sequences in the time interval $[t_0, t_f]$ that minimize the quadratic error between the actual poloidal magnetic flux profile at time t_f and a desired target profile. This work is aimed at saving long trial-and-error periods of time currently spent by fusion experimentalists trying to manually adjust the time evolutions of the actuators to achieve the desired current profile at some prespecified time. Simulation results show the effectiveness of this approach.

The paper is organized as follows. The optimal control problem for the current profile system is introduced in Section II. The POD method to obtain a reduced order model is discussed in Section III. In Section IV, the Galerkin projection method based on a test function set composed by dominant POD modes is also discussed. In Section V, the procedure for the POD-LDDS-based optimization is stated, and a brief introduction to SQP optimization theory is presented. Simulation studies are presented in Section VI. The paper is closed in Section VII by stating conclusions and future research remarks.

II. Statement of the Optimal Control Problems

The poloidal flux system can be modeled by the following parabolic PDE over $\mathcal{Q} = [0,1] \times [t_0, t_f]$ (see, e.g., [3]):

$$\begin{cases} \dfrac{1}{\Theta(x)}\dfrac{\partial y}{\partial t} = \dfrac{1}{x^\gamma}\dfrac{\partial}{\partial x}\left(x^\gamma \left(\dfrac{\bar{n}(t)}{I(t)\sqrt{P_{tot}(t)}}\right)^{\frac{3}{2}} \Gamma(x)\dfrac{\partial y}{\partial x}\right), \\ \qquad\qquad + \dfrac{\chi(x)}{\Theta(x)}\dfrac{\sqrt{P_{tot}(t)}}{I(t)} \\ y(x,0) = y_0(x), \quad x \in \Omega = [0,1] \\ \left[\dfrac{\partial y}{\partial x}\right]_{x=0} = 0, \quad \left[\dfrac{\partial y}{\partial x}\right]_{x=1} = I(t), \end{cases} \tag{1}$$

where $y(x,t)$ is the poloidal flux, and $\gamma = d-1$ is a constant related to the geometry of the problem. The dimension index $d = 2$ represents a cylindrical geometry and $d = 1$ a slab geometry. The system parameters $\Theta(x), \Gamma(x)$ and $\chi(x)$ are positive and differentiable with respect to the spatial coordinate. In system (1), we have three physical actuators $\bar{n}(t)$, $I(t)$ and $P_{tot}(t)$, which represent the average density, total plasma current and total power of the non-inductive heating source, respectively. Due to physical limitations, such as magnitude and rate saturations, the three physical actuators must satisfy the following constraints:

1 Magnitude saturations for the three actuators: $Lb_{\bar{n}} \leq \bar{n}(t) \leq Ub_{\bar{n}}$, $Lb_I \leq I(t) \leq Ub_I$, $Lb_{\sqrt{P_{tot}}} \leq \sqrt{P_{tot}(t)} \leq Ub_{\sqrt{P_{tot}}}$, where $Lb_{(\cdot)}$ and $Ub_{(\cdot)}$ are given physical bounds;

2 Rate saturations for the average density $\bar{n}(t)$ and the total plasma current $I(t)$: $lb_{\bar{n}} \leq \frac{d\bar{n}(t)}{dt} \leq ub_{\bar{n}}$, $lb_I \leq \frac{dI(t)}{dt} \leq ub_I$, where $lb_{(\cdot)}$ and $ub_{(\cdot)}$ are given physical bounds;

3 Fixed initial and final values for both the average density $\bar{n}(t)$ and the total plasma current $I(t)$: $\bar{n}(t_0) = \bar{n}_0$, $\bar{n}(t_f) = \bar{n}_f$, $I(t_0) = I_0$, $I(t_f) = I_f$.

In this paper, we use the set of actuators $\Xi = (\bar{n}, I, P_{tot})$ satisfying the indicated constraints to compute an open loop optimal input sequences for the poloidal flux system (1), such that the system output at $t = t_f$, i.e., $y(x,t_f)$, can reach a desired profile $y^*(x)$ as closely as possible, i.e., $\|y(x,t_f) - y^*(x)\|_{L^2(\Omega)} \leq \varepsilon$, $(\varepsilon \geq 0)$. The optimal control problem can be stated as follows:

$$\begin{cases} \min\limits_{\bar{n}, I, P_{tot}} \mathcal{J} = \displaystyle\int_0^1 |y(x,t_f) - y^*(x)|^2 dx \\ \text{s.t.} \quad \text{system}\,(1), \text{ and constraints.} \end{cases} \tag{2}$$

In order to take into account the constraints for $I(t)$ and $\bar{n}(t)$, reducing at the same time the computational effort, we propose simplified models for these control inputs, which are parameterized in terms of a few to-be-optimized parameters, i.e.,

$$I(t) = \frac{I_f + I_0}{2} + \frac{I_f - I_0}{2}\text{erf}\left(\frac{t-\mu_I}{\sqrt{2}\sigma_I}\right), \tag{3}$$

$$\bar{n}(t) = \frac{\bar{n}_f + \bar{n}_0}{2} + \frac{\bar{n}_f - \bar{n}_0}{2}\text{erf}\left(\frac{t-\mu_{\bar{n}}}{\sqrt{2}\sigma_{\bar{n}}}\right), \tag{4}$$

where $\sigma_I, \sigma_{\bar{n}}, \mu_I, \mu_{\bar{n}}$ are the to-be-optimized shape parameters and $\mathrm{erf}(\cdot)$ is the error function defined by [12]

$$\mathrm{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-\tau^2) d\tau. \tag{5}$$

Remark 1: Modeling $I(t)$ and $\bar{n}(t)$ as in (3) and (4) has some physical advantages. The evolutions of these control inputs can start and stop at given values very smoothly, without much change rate. The maximum change rate takes place at $t = \mu_I$ and $t = \mu_{\bar{n}}$ respectively.

III. Proper Orthogonal Decomposition Method

The set $\mathcal{V} = \mathrm{span}\{y_1, \cdots, y_n\} \subset \mathbb{R}^m$ refers to a data ensemble consisting of n snapshots $y_j = y(\bar{x}, t_j)$, for $j = 1, \ldots, n$ and where $\bar{x}$ denotes m equidistant discrete points in the spatial domain $[0, 1]$, of system (1) obtained either from experiments or simulations. Let $\{\psi_k\}_{k=1}^d$ be an orthonormal basis of the data ensemble $\mathcal{V}$, where $d = \dim \mathcal{V} \leq m$. We then project each of the snapshots onto the basis ψ_k,

$$y_j = \sum_{k=1}^{d} (y_j^T \psi_k) \psi_k, \quad j = 1, \cdots, n. \tag{6}$$

The goal of the POD method is to find a subset of the orthonormal basis $\{\psi_k\}_{k=1}^d$ such that for some predefined $1 \leq l \leq d$ the following average index is minimized

$$\begin{aligned} &\min_{\{\psi_k\}_{k=1}^l} \frac{1}{n} \sum_{j=1}^{n} \left\| y_j - \sum_{k=1}^{l} (y_j^T \psi_k) \psi_k \right\|^2, \\ &\quad \text{subject to} \quad \psi_i^T \psi_j = \delta_{ij}, 1 \leq i \leq l, 1 \leq j \leq i, \\ &\quad \text{where} \quad \|y\| = \sqrt{y^T y}. \end{aligned} \tag{7}$$

The solution of (7) can be found in the literature, e.g., [9], [13]. Defining the correlation matrix $K \in \mathbb{R}^{n \times n}$ as $K_{ij} = \frac{1}{n}(y_j^T y_i)$, for $i, j = 1, \ldots, n$, it follows the following singular value decomposition result [13]:

Theorem 1: Let $\lambda_1 > \ldots > \lambda_l > \ldots > \lambda_d > 0$ denote the positive eigenvalues of the correlation matrix K and $v_1, \ldots, v_l, \ldots, v_d$ the associated eigenvectors, where $d = \mathrm{rank}(K)$. Then, the POD basis functions take the form

$$\psi_k = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^{n} (v_k)_j y_j = \frac{1}{\sqrt{\lambda_k}} Y v_k, \ (k = 1, \ldots, d), \tag{8}$$

where $(v_k)_j$ is the j-th component of the eigenvector v_k and $Y = (y_1, \cdots, y_n)$ is the collection matrix of all the snapshots. Moreover, the error (energy ratio) associated with the approximation with the first l POD modes is

$$\varepsilon_l = \frac{1}{n} \sum_{j=1}^{n} \left\| y_j - \sum_{k=1}^{l} (y_j^T \psi_k) \psi_k \right\|^2 = \sum_{k=l+1}^{d} \lambda_k. \tag{9}$$

IV. POD/Galerkin Method

We denote the input functions by $u_0(t) = \left(\frac{\bar{n}(t)}{I(t)\sqrt{P_{tot}(t)}} \right)^{\frac{3}{2}}$, $u_1(t) = \frac{\sqrt{P_{tot}(t)}}{I(t)}$, $u_2(t) = I(t)$, and let $\psi_i \in V_{POD}$ be the test function, where $V_{POD} = \mathrm{span}\{\psi_1, \ldots, \psi_l\}$ is the test function space spanned by the POD modes. Then, we can obtain the weak form of (1)

$$\begin{aligned} &\frac{\partial}{\partial t} \int_0^1 x^\gamma \frac{y(x,t)}{\Theta(x)} \psi_i(x) dx \\ &= u_0(t) \int_0^1 \frac{\partial}{\partial x} \left(x^\gamma \Gamma(x) \frac{\partial y(x,t)}{\partial x} \right) \psi_i(x) dx \\ &\quad + u_1(t) \int_0^1 x^\gamma \frac{\chi(x)}{\Theta(x)} \psi_i(x) dx. \end{aligned} \tag{10}$$

We integrate the second term by parts and take into account the boundary conditions to obtain

$$\begin{aligned} &\tfrac{\partial}{\partial t} \textstyle\int_0^1 x^\gamma \frac{y(x,t)}{\Theta(x)} \psi_i(x) dx + u_0(t) \int_0^1 x^\gamma \Gamma(x) \frac{\partial y}{\partial x} \psi_i'(x) dx \\ &\quad = u_0(t) u_2(t) \Gamma(1) \psi_i(1) + u_1(t) \textstyle\int_0^1 x^\gamma \frac{\chi(x)}{\Theta(x)} \psi_i(x) dx. \end{aligned} \tag{11}$$

We substitute the Galerkin approximation $y(x,t) \approx z(x,t) = \sum_{j=1}^{l} z_j(t) \psi_j(x)$ into the weak form (11) to obtain

$$\begin{aligned} &\sum_{j=1}^{l} a(\psi_i, \psi_j) \frac{dz_j(t)}{dt} + u_0(t) \sum_{j=1}^{l} b(\psi_i, \psi_j) z_j(t) \\ &= u_0(t) u_2(t) \Gamma(1) \psi_i(1) + u_1(t) \left(\frac{\chi}{\Theta}, \psi_i \right)_{L^2}, \\ &\qquad (i = 1, \ldots, l), \end{aligned} \tag{12}$$

where $z_1, \ldots, z_l$ are unknown functions and

$$a(\psi_i, \psi_j) = \int_0^1 x^\gamma \frac{1}{\Theta(x)} \psi_i(x) \psi_j(x) dx, \tag{13}$$

$$b(\psi_i, \psi_j) = \int_0^1 x^\gamma \Gamma(x) \psi_i'(x) \psi_j'(x) dx, \tag{14}$$

$$\left(\frac{\chi}{\Theta}, \psi_i \right)_{L^2} = \int_0^1 x^\gamma \frac{\chi(x)}{\Theta(x)} \psi_i(x) dx. \tag{15}$$

The initial values for $z_1, \ldots, z_l$ are uniquely determined by the following relation: $\sum_{i=1}^{l} z_i(0) \psi_i(x) \approx y(x, 0)$, i.e.,

$$z_i(0) = (y(x,0), \psi_i(x))_{L^2}, \quad i = 1, \ldots, l. \tag{16}$$

We denote $Z = (z_1, \ldots, z_l)^T$ the vector of the unknown functions, and write (12) in the matrix form

$$\mathcal{M} \dot{Z} + u_0(t) \mathcal{K} Z = l_1 u_1(t) + l_2 u_0(t) u_2(t), \tag{17}$$

where $\mathcal{M} = [a(\psi_i, \psi_j)]_{i,j=1,\ldots,l}$, $\mathcal{K} = [b(\psi_i, \psi_j)]_{i,j=1,\ldots,l}$, and $l_1 = [(\chi \Theta^{-1}, \psi_i)_{L^2}]_{i=1,\ldots,l}$, $l_2 = [\Gamma(1)\psi_i(1)]_{i=1,\ldots,l}$. We multiply both sides with $\mathcal{M}^{-1}$ to obtain

$$\dot{Z} = u_0(t) L_0 Z + L_1 u_1(t) + L_2 u_0(t) u_2(t), \tag{18}$$

where $L_0 = -\mathcal{M}^{-1}\mathcal{K}$, $L_1 = \mathcal{M}^{-1} l_1$ and $L_2 = \mathcal{M}^{-1} l_2$.

V. POD-LDDS-based Optimization

In this paper, we take advantage of the POD model reduction technique to solve the optimization problem iteratively.

A. Algorithm

To start the optimization algorithm, we assume an initial guess for the control sequences to simulate the PDE system (1). Then, we use the simulation data ensemble to generate POD basis functions and the LDDS (18) for the PDE (1). Then we solve a finite dimensional nonlinear optimization problem to obtain optimal control sequences. We update the POD modes, and the LDDS, after generating a new simulation data ensemble with the obtained optimal control sequences. The algorithm can be summarized as follows:

1 Set $k = 0$ to give the initial guess for the control sequences, $\bar{n}^{(k)}$, $I^{(k)}$ and $(P_{tot})^{(k)}$;

2 Simulate the PDE system (1) to obtain data ensemble $Y^{(k)}$;

3 Generate the POD modes $V_{POD}^{(k)}$ and the LDDS from the data ensemble,

$$\frac{dZ^{(k)}}{dt} = f_{LDDS}^{(k)}\left(Z^{(k)}, \bar{n}^{(k)}, I^{(k)}, (P_{tot})^{(k)}\right); \quad (19)$$

4 Solve the following optimization problem for the LDDS,

$$\begin{cases} \min\limits_{Z^{(k)},\Xi^{(k)}} \mathcal{F}\left(Z^{(k)}, \Xi^{(k)}\right) \\ \quad \triangleq \left\| \sum\limits_{i=1}^{l} z_i^{(k)}(t_f)\varphi_i^{(k)}(x) - y^*(x) \right\|, \\ \text{subject to}: \quad e\left(Z^{(k)}, \Xi^{(k)}\right) = 0, \\ e\left(Z^{(k)}, \Xi^{(k)}\right) \triangleq \dfrac{dZ^{(k)}}{dt} - f_{LDDS}^{(k)}\left(Z^{(k)}, \Xi^{(k)}\right), \\ \Xi^{(k)} = \left(\bar{n}^{(k)}, I^{(k)}, (P_{tot})^{(k)}\right); \end{cases} \quad (20)$$

5 Go back to Step 2 and stop the iteration if the data ensemble satisfies

$$\|Y^{(k+1)} - Y^{(k)}\| \le \varepsilon; \quad (21)$$

Otherwise, continue iteration until the error criterion is satisfied.

B. Sequential Quadratic Programming

For each iteration (k), we now have a constrained nonlinear programming (NLP) problem (20). We define $X^{(k)} \triangleq \left(Z^{(k)}, \Xi^{(k)}\right)$, and rewrite (20) as

$$\begin{cases} \min\limits_{X^{(k)}} \mathcal{F}\left(X^{(k)}\right) \\ \text{s.t.} \quad e\left(X^{(k)}\right) = 0. \end{cases} \quad (22)$$

If $\widehat{\lambda^{(k)}}$ is the Lagrange multiplier corresponding to a local minimizer $\widehat{X^{(k)}}$ of (22), then the Lagrangian $\mathcal{L}\left(X^{(k)}, \lambda^{(k)}\right) = \mathcal{F}\left(X^{(k)}\right) + \lambda^{(k)} \cdot e\left(X^{(k)}\right)$ satisfies $\mathcal{L}\left(X^{(k)}, \widehat{\lambda^{(k)}}\right) = \mathcal{F}\left(X^{(k)}\right)$ for all admissible $X^{(k)}$. Then, we can equivalently rewrite the constrained NLP problem (22) as

$$\begin{cases} \min\limits_{X^{(k)}} \mathcal{L}\left(X^{(k)}, \widehat{\lambda^{(k)}}\right) = \mathcal{F}\left(X^{(k)}\right) + \widehat{\lambda^{(k)}} \cdot e\left(X^{(k)}\right) \\ \quad \text{s.t.} \quad e\left(X^{(k)}\right) = 0, \end{cases} \quad (23)$$

since (23) also has $\widehat{X^{(k)}}$ as a local minimizer. The multiplier $\widehat{\lambda^{(k)}}$ is unknown but an algorithm can approximate it as $\widehat{X^{(k)}}$ is approximated. To obtain this local optimizer pair $\left(\widehat{X^{(k)}}, \widehat{\lambda^{(k)}}\right)$, we use the SQP method [8] which approximates the NLP (23) with a sequence of quadratic programming (QP) problems. At each step j, a local model of the optimization problem is constructed around the current solution $\left(X^{(k,j)}, \lambda^{(k,j)}\right)$ by using a quadratic approximation of the objective functional and a linear approximation of the constraint equation. Then the original NLP (23) becomes a QP problem whose solution yields a step toward the solution of the original problem:

$$\mathrm{QP}^{(k,j)}: \begin{cases} \min\limits_{p^{(k,j)}} \mathcal{L}\left(X^{(k,j)} + p^{(k,j)}, \lambda^{(k,j)}\right) \\ \approx \mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) \\ \quad + \nabla\mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) \cdot p^{(k,j)} \\ \quad + \dfrac{1}{2} p^{(k,j)} \cdot \nabla^2\mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) p^{(k,j)}, \\ \text{s.t. } e\left(X^{(k,j)} + p^{(k,j)}\right) \approx e\left(X^{(k,j)}\right) \\ \quad + \nabla e\left(X^{(k,j)}\right) \cdot p^{(k,j)} = 0. \end{cases} \quad (24)$$

To solve each $\mathrm{QP}^{(k,j)}$ in (24), we introduce another Lagrangian functional as follows

$$l\left(p^{(k,j)}, \lambda^{(k,j)}, \mu^{(k,j)}\right) = \mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) + \nabla\mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) \cdot p^{(k,j)} + \frac{1}{2} p^{(k,j)} \cdot \nabla^2\mathcal{L}\left(X^{(k,j)}, \lambda^{(k,j)}\right) p^{(k,j)} \quad (25)$$

$$+ \mu^{(k,j)} \cdot \left\{ e\left(X^{(k,j)}\right) + \nabla e\left(X^{(k,j)}\right) \cdot p^{(k,j)} \right\}. \quad (26)$$

Then, we can formulate the following unconstrained optimization problem

$$\min_{\{p^{(k,j)}, \mu^{(k,j)}\}} l\left(p^{(k,j)}, \mu^{(k,j)}\right), \quad (27)$$

and the associated first order optimality conditions can be obtained by forcing the first order derivatives to be zero, i.e.,

$$\begin{pmatrix} \nabla^2\mathcal{L} & (\nabla e)^T \\ \nabla e & 0 \end{pmatrix} \begin{pmatrix} p^{(k,j)} \\ \mu^{(k,j)} \end{pmatrix} = - \begin{pmatrix} \nabla\mathcal{L} \\ e \end{pmatrix}. \quad (28)$$

We assume that the Hessian $\nabla^2\mathcal{L}(X^{(k,j)}, \lambda^{(k,j)})$ is positive definite and the minimizer is well–defined. Therefore, by solving $p^{(k,j)}, \mu^{(k,j)}$ from (28), the update of $\left(X^{(k,j)}, \lambda^{(k,j)}\right)$ can be implemented as follows,

$$X^{(k,j+1)} = X^{(k,j)} + p^{(k,j)}, \quad (29)$$

$$\lambda^{(k,j+1)} = \lambda^{(k,j)} + \mu^{(k,j)}. \quad (30)$$

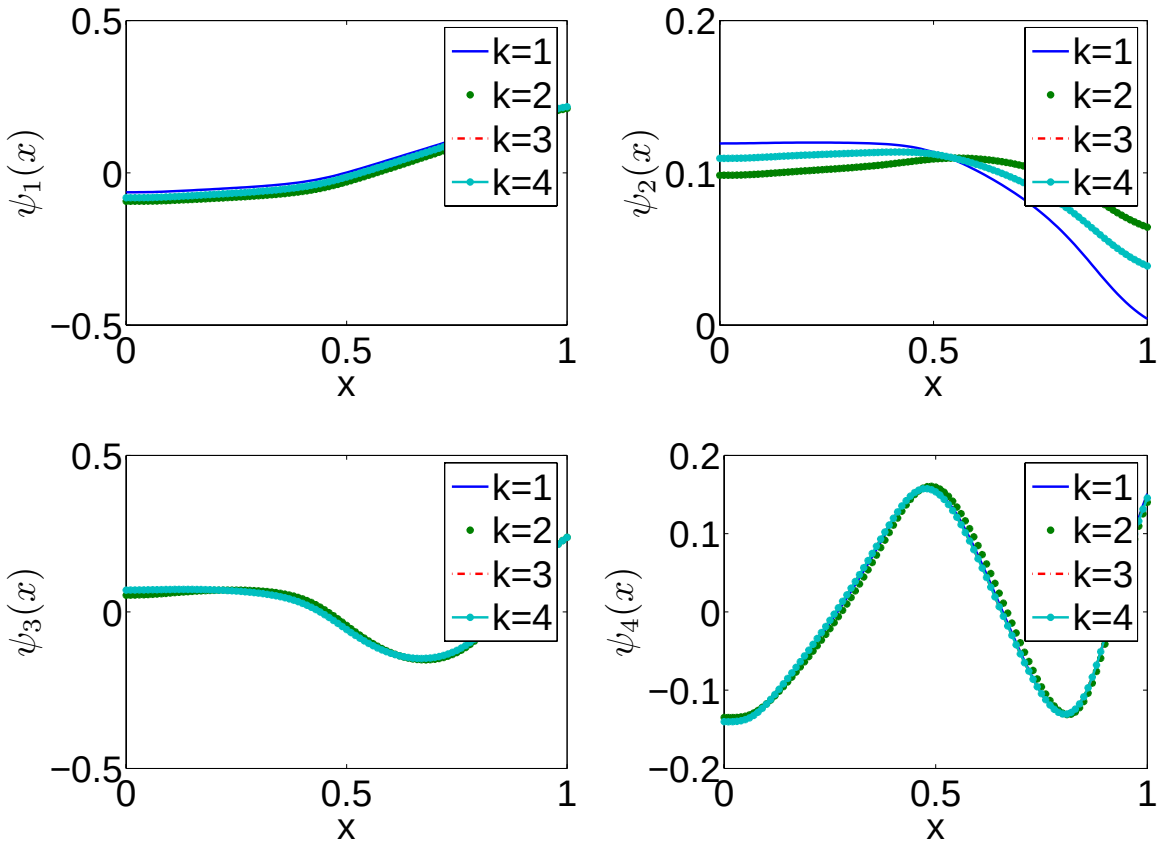

Fig. 2. Iterations of POD modes.

It is interesting to note that (28)-(30) can also be seen as the solution via Newton's method of $\nabla\mathcal{L} = 0$ and $e = 0$, the optimality condition and constraint of (23) respectively.

VI. Numerical Illustrations

In this section, we implement the POD-LDDS-based optimization procedure for the poloidal flux system in a simulation of a DIII-D plasma. The system parameters in the PDE (1) can be given as interpolation polynomials

$$\begin{aligned}\Theta(x) &= 255.1x^6 - 682.5x^5 \\ &\quad + 688.4x^4 - 322.3x^3 + 69.5x^2 - 5.8x + 0.5, \\ \Gamma(x) &= 172.9812x^6 - 458.3779x^5 + 446.8464x^4 \\ &\quad - 190.9911x^3 + 31.4997x^2 - 0.2664x + 0.9213, \\ \chi(x) &= -504.0x^6 + 1331.9x^5 - 1227.6x^4 + 437.1x^3 \\ &\quad - 26.2x^2 - 14.3x - 2.9,\end{aligned}$$

and the initial and desired profiles are given by $y(x, t_0) = -0.17x^3 + 0.35x^2 - 0.06x - 0.37$ and $y^*(x) = -1.72x^3 + 3.86x^2 - 0.77x - 0.23$ respectively.

We discretize the time domain $[0, 1.2]$ into seven subintervals and parameterize the actuator P_{tot} as eight discrete points, i.e., $P_{tot}(t_i), (i = 1, 2, \ldots, 8)$. We use linear interpolation to approximate the function $P_{tot}(t)$ over $(t_i, t_{i+1}), (i = 1, 2, \ldots, 7)$. The functions $\bar{n}(t)$ and $I(t)$ are parameterized by $\mu_I, \mu_{\bar{n}}, \sigma_I$ and $\sigma_{\bar{n}}$ in expressions (3)–(4). In total, we have twelve parameters to be optimized subject to the following physical constraints:

$$0.45 \le \mu_{\bar{n}} \le 0.75, \quad 0.45 \le \mu_I \le 0.75, \tag{31}$$

$$0.09 \le \sigma_{\bar{n}} \le 0.15, \quad 0.09 \le \sigma_I \le 0.15, \tag{32}$$

$$0.5 \le P_{tot}(t_k) \le 20, \quad k = 1, 2, \ldots, 8. \tag{33}$$

We take $\bar{n}_0 = 2$, $\bar{n}_f = 2.7$, $I_0 = 0.709229$, $I_f = 1.18774$, and set the initial guess $\left(\mu_I^{(0)}, \sigma_I^{(0)}, \mu_{\bar{n}}^{(0)}, \sigma_{\bar{n}}^{(0)}\right) = (0.6, 0.12, 0.6, 0.1)$, $\sqrt{P_{tot}(t_i)} = 3$, $(i = 1, 2, \ldots, 8)$. We can note that the POD modes (Fig. 2), and optimized sequences $\bar{n}^{(k)}$ (Fig. 3), $I^{(k)}$ (Fig. 4), ${P_{tot}}^{(k)}$ (Fig. 5) also show a fast convergence. The iterative procedure is therefore stopped after the fourth iteration. The evolution of the final profile $y^{(k)}(t_f, x)$ iteration after iteration, which is shown in Fig. 6, is obtained by simulating the PDE system (1) using the optimized sequences $\left(\bar{n}^{(k)}, I^{(k)}, {P_{tot}}^{(k)}\right)$, $(k = 1, 2, 3, 4)$. In Fig. 7, the converged open-loop optimal control sequences $\bar{n}(t)$, $I(t)$, $P_{tot}(t)$, reconstructed after the fourth iteration, are used to simulate the the poloidal flux evolution governed by the PDE system (1).

Remark 2: The fact that the convergent final profile can not match the desired target profile perfectly does not imply a limitation of the POD-LDDS-based optimization procedure but a consequence of the lack of precise reachability for the parabolic system under the actuator constraints and assumptions (3)-(4).

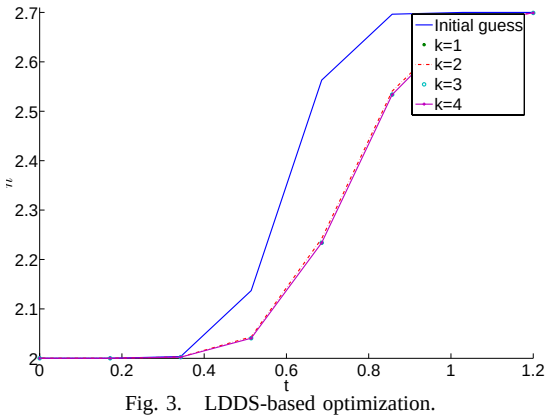

Fig. 3. LDDS-based optimization.

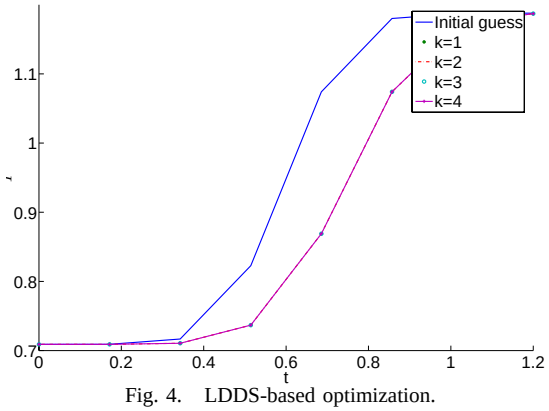

Fig. 4. LDDS-based optimization.

VII. Conclusions and Future Research

In this paper, we solve the open–loop optimal current profile control problem arising in tokamak plasmas by using POD model reduction and nonlinear programming. Numerical studies demonstrate that the iterative POD-LDDS-optimization procedure is effective and reduces computational effort when compared with PDE-optimization.

The POD-LDDS-optimization procedure involves two iterations. One is the POD-LDDS iteration (k) in (29) (where the model is updated as the input sequences change) and the other is the SQP iteration (j) in (32) (where a series of QP problems are solved). For the SQP iteration, convergence has been theoretically proved (see, e.g., [8]). However, the convergence of the POD-LDDS iteration is still an open problem and a part of our future work.

References

[1] E. Witrant *et al.*, "A control-oriented model of the current profile in tokamak plasma," *Plasma Phys. Control. Fusion*, vol. 49, pp. 1075–1105, 2007.

[2] Y. Ou *et al.*, "Extremum-seeking finite-time optimal control of plasma current profile at the DIII-D tokamak," *Proceedings of the 2007 American Control Conference*, 2007.

[3] ——, "Towards model-based current profile control at DIII-D," *Fusion Engineering and Design*, vol. 82, pp. 1153–1160, 2007.

[4] J. Blum, *Numerical simulation and optimal control in plasma physics*. John Wiley &. Sons, 1988.

[5] P. Neittaanmaki and D. Tiba, *Optimal control of nonlinear parabolic systems: theory, algorithms, and applications*. Marcel Dekker, Inc., 1994.

[6] V. Arnautu and P. Neittaanmaki, *Optimal control from theory to computer programs*. Kluwer Academic Publishers, 1993.

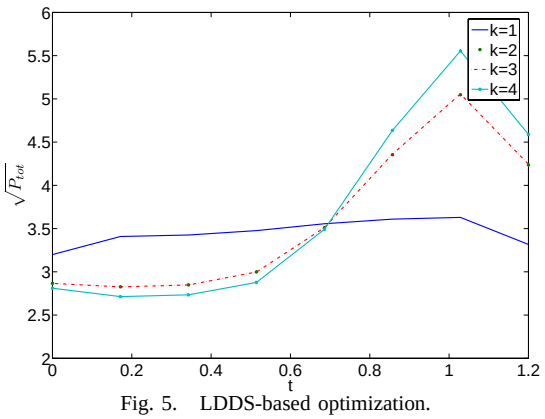

Fig. 5. LDDS-based optimization.

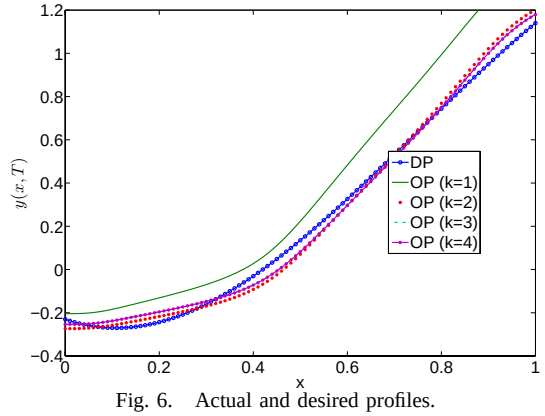

Fig. 6. Actual and desired profiles.

[7] R. Pytlak, *Numerical Methods for Optimal Control Problems with State Constraints*. Springer, 1999.

[8] J. Nocedal and S. J. Wright, *Numerical Optimization (2nd edition)*. New York: Springer, 2006.

[9] P. Holmes, J. Lumley, and G. Berkooz, *Turbulence, coherent structures, dynamical systems and symmetry*. New York: Cambridge University Press, 1996.

[10] M. Bergmann, L. Cordier, and J. Brancher, "Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced-order model," *Physics of Fluids*, vol. 17, pp. 097 101(1–21), 2005.

[11] K. Kunisch and S. Volkwein, "Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition," *Journal of Optimization Theory and Applications*, vol. 102, pp. 345–371, 1999.

[12] O. Kallenberg, *Foundations of Modern Probability (Springer Series in Statistics)*. Springer, 2002.

[13] K. Kunisch and S. Volkwein, "Galerkin proper orthogonal decomposition methods for parabolic problems," *Numerische Mathematik*, vol. 90, pp. 117–148, 2001.

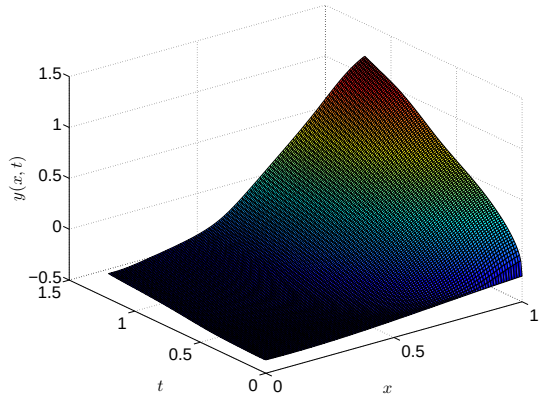

Fig. 7. System evolution with optimization control.