Optimal Control for Quasi-Static Evolution of Plasma Equilibrium in Tokamaks

Holger Heumann, Jacques Blum

OPTIMAL CONTROL FOR QUASI-STATIC EVOLUTION OF PLASMA EQUILIBRIUM IN TOKAMAKS

HOLGER HEUMANN AND JACQUES BLUM

Abstract. We present a new approach to the optimization of plasma scenarios in tokamaks. We formulate this task as an optimal control problem and use numerical methods for optimization problems with partial differential equation (PDE) constraints. The latter are discretized by linear finite elements and implicit Euler time stepping. Due to the free-boundary setting, we have introduced a new linearization of the non-linear equations that is consistent with the discretization. It is only this consistency, overlooked in previous approaches for the even simpler static setting, that guarantees that the method converges to the optimum. We give various numerical tests and simulation that illustrate our new method.

1. Introduction

A tokamak is an experimental device that aims at establishing the conditions to start nuclear fusion reactions. It uses a magnetic field to confine the plasma. Basic control mechanism for running such a device are externally applied voltages in various coils surrounding the main chamber, beam injection or wave heating devices. The control of plasma in tokamaks relies until today mainly on empirical experience and feed-back control systems [2]. Todays experimental tokamaks such as Tore Supra in Cadarache, JET (Joint European Torus) at Culham or TCV (Tokamak à Configuration Variable) in Lausanne use a very complex settings of such feed-back control system. It is highly questionable that such feed-back control systems can be transferred one-to-one to newly designed tokamaks. Hence, in view of the upcoming ITER tokamak, the control of plasma scenarios has recently gained priority attention within the research community.

With this work we would like to promote a modern feed-forward approach to the control of plasma in tokamaks. Feed-forward approaches, also called *optimal control* approaches in engineering and mathematical literature, predict control parameters prior to the actual experiment. Feed-forward approaches for the control of plasma date back to the pioneering work in [3]. But while in [3] even the supercomputers of that time were able to tackle only simplified stationary problems, we have now the computational power to tackle the entire discharge scenario. We would like to stress that nevertheless feed-back control will be inevitable to run tokamak devices.

As in any other optimal control formulation we have to rely on the assumption that the whole process is accurately described by a set of (partial) differential equations. In our case it is the widely accepted Grad-Hogan model [10, 11] for the direct simulation of plasma evolution. This model decomposes the evolution of plasma in a tokamak into two main subproblems:

Date: May 7, 2014, submitted to SIAM Journal on Scientific Computing.

- the two dimensional axisysmmetric equilibrium problem together with Maxwell's equation in coils and conducting structures that determine the evolution of the poloidal flux for given pressure;
- the one dimensional transport equations together with the resistive diffusion equation for the poloidal flux that describe the evolution of density and temperature in the direction perpendicular to the poloidal flux lines.

We refer to [4] or [17] for more details on that topic, but need to highlight that both the appropriate models for transport and the numerical treatment of the two coupled problems are presently the subject of intense research [7, 15, 16, 6].

In this work we will concentrate on the equilibrium part and extend the ideas of the static case in [3] to a quasi-static setting. We formulate a cost function that penalizes deviations of the plasma from the desired state. State-of-the-art methods for constrained optimization, so-called *sequential quadratic programming (SQP)* [21, Chapter 18], are used to minimize this cost function under the constraint of quasistatic evolution of plasma equilibrium. It is nowadays widely accepted that only accurately evaluated sensitivities of the direct non-linear model can guarantee robustness and liability of the optimization procedure. The free plasma boundary setting of the direct problem leads in here to some subtle discretization schemes that have been overlooked in similar earlier work on the static case [3, 13].

The outline of the paper is the following: In Section 2 we recall the Grad-Shafranov-Schlüter equation for plasma equilibrium and supplement these with Maxwell's equations in the coils and other conducting structures. This results in a non-linear free boundary problem that is a static problem if we prescribe currents in the coils and quasi-static if we prescribe voltages in the coil circuits. We present a variational formulation and formulate the corresponding inverse problems. The first aims at identifying optimal currents that ensure that the stationary plasma equilibrium has a certain desired shape. The second aims at identifying optimal evolution of voltages that ensure that the transient plasma equilibrium follows a desired shape trajectory. In Section 3 we present the main numerical methods. We use a standard finite element method with linear Lagrangian basis functions and focus on specific points, due to the non-linearity introduced by the definition of the plasma domain. This is followed by numerical tests and simulations that validate our implementation.

2. Optimal Control of Plasma Equilibrium

2.1. **Plasma Equilibrium as Free-Boundary Problem.** The essential equations for describing plasma equilibrium in a tokamak are force balance

$$\operatorname{grad} p = \mathbf{J} \times \mathbf{B}\,, \tag{1}$$

and the magneto-quasistatic approximation of Maxwell's equations with Ohm's law:

$$\partial_t \mathbf{B} = \operatorname{curl} \mathbf{E}_{src} - \operatorname{curl} \frac{1}{\sigma}\mathbf{J}\,, \quad \operatorname{div} \mathbf{B} = 0\,, \quad \operatorname{curl} \frac{1}{\mu}\mathbf{B} = \mathbf{J}\,. \tag{2}$$

Here, p is the kinetic pressure, $\mathbf{B}$ the magnetic field and $\mathbf{J}$ the current density and $\mathbf{E}_{src}$ an externally applied electric field. μ and σ are the magnetic permeability and the electrical conductivity. Hence for the resistive timescale the plasma is in equilibrium and (1) holds at each instant of time. Only the Faraday's law in (2) drives the dynamics.

As it can easily be seen from (1), both the magnetic field $\mathbf{B}$ and the current density $\mathbf{J}$ lie on surfaces of constant pressure p. This property is the major idea behind the tokamak concept for fusion: *A magnetic cage confines particles and the particle pressure gradient is balanced by the magnetic pressure gradient.* Hence, in light of this, the fusion community prefers to work with a reformulation of (1) and (2) that uses unknowns that are constant on such constant pressure surfaces. We would like to refer to standard text books, e.g. [23, Chapter 3], [8, Chapter 6] and [17, Chapter 4], for a detailed derivation and restate here only the final equations.

Under the assumption of perfect axial symmetry, it is convenient to put (1) and (2) in a cylindrical coordinate system (r, φ, z) and to consider only a meridian section of the tokamak. The primal unknowns are the *poloidal magnetic flux* $\psi = \psi(r, z)$, the pressure $p = p(\psi)$ and the *poloidal current flux* $f = f(\psi)$. The poloidal magnetic flux $\psi := r\mathbf{A} \cdot \mathbf{e}_\varphi$ is the scaled toroidal component of the vector potential $\mathbf{A}$, i.e. $\mathbf{B} = \operatorname{curl} \mathbf{A}$ and $\mathbf{e}_\varphi$ the unit vector for φ. The poloidal current flux $f = r\mathbf{B} \cdot \mathbf{e}_\varphi$ is the scaled toroidal component of the magnetic field. We introduce $\Omega_\infty = [0, \infty] \times [-\infty, \infty]$, the positive half plane, to denote the meridian plane that contains the cross section of the tokamak device. The geometry of the tokamak determines the various subdomains (see Figure 1):

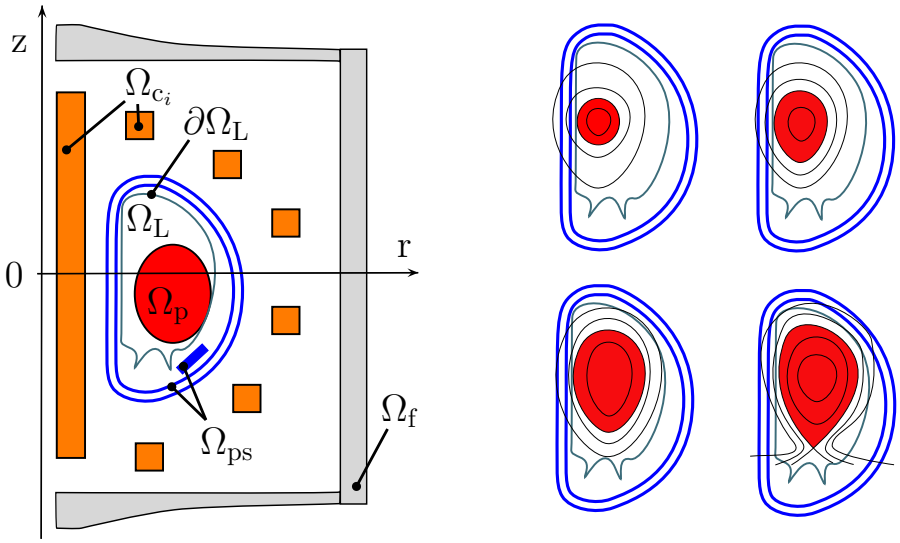


FIGURE 1. Left: Geometric description of the poloidal cross section of the tokamak device. Right: Sketch for characteristic plasma shapes during the so-called ramp-up phase. The ψ-isolines are indicated by black lines. In the beginning (first three pictures) the plasma touches the limiter wall and becomes more and more elongated, while finally it moves into the contact-free configuration.

- the domain $\Omega_f \subset \Omega_\infty$ corresponds to those parts that are made of iron, the iron core and return limbs;
- the domain $\Omega_{c_i} \subset \Omega_\infty$ corresponds to the N poloidal field coils, where each coil Ω_{c_i} has n_i wire turns, total resistance R_i and cross section S_i;

- the domain $\Omega_{\rm ps} \subset \Omega_\infty$ corresponds to the passive structures, with conductivity σ;
- the domain $\Omega_{\rm L} \subset \Omega_\infty$, bounded by the limiter, corresponds to the domain that is accessible by the plasma.

Then, the equilibrium of plasma in a tokamak has to satisfy the following non-linear boundary value problem:

$$\begin{aligned} L\psi(r,z,t) &= j(r,z) \quad \text{in} \quad \Omega\,; \\ \psi(0,z,t) &= 0\,; \\ \lim_{\|(r,z)\|\to+\infty} \psi(r,z,t) &= 0\,; \\ \psi(r,z,0) &= \psi^0(r,z)\,, \end{aligned} \tag{3}$$

where L is a non-linear second-order elliptic differential operator

$$L\psi := -\frac{\partial}{\partial r}\left(\frac{1}{\mu(\psi)r}\frac{\partial \psi}{\partial r}\right) - \frac{\partial}{\partial z}\left(\frac{1}{\mu(\psi)r}\frac{\partial \psi}{\partial z}\right) := -\nabla\cdot\left(\frac{1}{\mu(\psi)r}\nabla\psi\right), \tag{4}$$

with

$$\mu(\psi) = \mu_{\rm f}(|\nabla\psi|^2 r^{-2}) \begin{cases} \geq \mu_0 & \text{in } \Omega_{\rm f} \\ = \mu_0 & \text{else}\,. \end{cases} \tag{5}$$

Here, ∇ is the 2D gradient in the (r,z)-plane. The current density j is a non-linear function of ψ:

$$j = \begin{cases} rp'(\psi) + \frac{1}{\mu_0 r} f f'(\psi) & \text{in } \Omega_{\rm p}(\psi)\,; \\ \frac{n_i V_i(t)}{R_i S_i} - 2\pi \frac{n_i^2}{R_i S_i^2} \int_{\Omega_{{\rm c}_i}} \frac{\partial \psi}{\partial t}\, dr dz =: \frac{I_i(\psi)}{S_i} & \text{in } \Omega_{{\rm c}_i}\,; \\ -\frac{\sigma}{r}\frac{\partial \psi}{\partial t} & \text{in } \Omega_{\rm ps}\,; \\ 0 & \text{elsewhere}\,, \end{cases} \tag{6}$$

where the domain $\Omega_{\rm p}(\psi)$, the *domain of the plasma*, is the largest subdomain of $\Omega_{\rm L}$ that is bounded by a ψ-isoline that is closed in $\Omega_{\rm L}$, more precisely,

$$\Omega_{\rm p}(\psi) = \{(r,z) \in \Omega_{\rm L}\ , \psi(r,z) \geq \max\left(\max_{(r,z)\in\partial\Omega_{\rm L}}(\psi(r,z)), \max_{(r_X,z_X)\in\Omega_{\rm L}}(\psi(r_{\rm X},z_{\rm X}))\right)\}$$

where $(r_{\rm X}, z_{\rm X})$ denotes the coordinates of the saddle points, also called *X-points*, of ψ. The different characteristic shapes of $\Omega_{\rm p}(\psi)$ are illustrated in Figure 1: the boundary of $\Omega_{\rm p}(\psi)$ either touches the boundary of $\Omega_{\rm L}$, the limiter, or the boundary contains one or more corners (X-points of ψ). The current density j is non-linear in ψ due to the non-linear functions p' and $f\,f'$ and the definition of the plasma domain $\Omega_{\rm p}(\psi)$. While $\Omega_{\rm p}(\psi)$ is fully determined for a given ψ, the two functions p' and $f\,f'$ are not determined by the model (3)-(6). In general, (3)-(6) is augmented by the so-called transport and diffusion equations, that determine p' and $f\,f'$. In this work, we will assume that, up to some scaling coefficient λ, the functions p' and $f\,f'$ are known. For the current profile j in the plasma domain $\Omega_{\rm p}$ we will consider the a priori model:

$$j = \lambda j_{\rm p}(r,\psi_{\rm N}) := \lambda h(r) g(\psi_{\rm N}) := \lambda\left(\beta\frac{r}{R_0} + (1-\beta)\frac{R_0}{r}\right)(1-\psi_{\rm N}^\alpha)^\gamma \quad \text{in } \Omega_{\rm p}\,, \tag{7}$$

with R_0 the major radius of the vacuum chamber and $\alpha, \beta, \gamma \in \mathbb{R}$ given parameters. We refer to [20] for a physical interpretation of these parameters. The parameter β

is the poloidal beta at $r = R_0$, whereas α and γ describe the peakage of the current profile. The normalized poloidal flux $\psi_{\mathrm{N}}(r,z)$ is

$$\psi_{\mathrm{N}}(r,z) = \frac{\psi(r,z) - \psi_{\mathrm{ax}}(\psi)}{\psi_{\mathrm{bnd}}(\psi) - \psi_{\mathrm{ax}}(\psi)}\,. \tag{8}$$

with the scalar values ψ_{ax} and ψ_{bnd} being the flux values at the *magnetic axis* and at the boundary of the plasma:

$$\begin{aligned} \psi_{\mathrm{ax}}(\psi) &:= \psi(r_{\mathrm{ax}}, z_{\mathrm{ax}})\,, \\ \psi_{\mathrm{bnd}}(\psi) &:= \max\left(\max_{(r,z)\in\partial\Omega_{\mathrm{L}}} \psi(r,z), \max_{(r_{\mathrm{X}},z_{\mathrm{X}})\in\Omega_{\mathrm{L}}} \psi(r_{\mathrm{X}}, z_{\mathrm{X}}) \right)\,, \end{aligned} \tag{9}$$

where $(r_{\mathrm{ax}}, z_{\mathrm{ax}}) = (r_{\mathrm{ax}}(\psi), z_{\mathrm{ax}}(\psi))$ and $(r_{\mathrm{x}}, z_{\mathrm{x}})$=$(r_{\mathrm{x}}(\psi), z_{\mathrm{x}}(\psi))$ are the global maximum and the saddle points of ψ in Ω_{L}. For convenience we introduce also the coordinates $(r_{\mathrm{bnd}}, z_{\mathrm{bnd}}) = (r_{\mathrm{bnd}}(\psi), z_{\mathrm{bnd}}(\psi))$ of the point that determines the plasma boundary, i.e.

$$(r_{\mathrm{bnd}}, z_{\mathrm{bnd}}) = \operatorname*{arg\,max}_{(r,z)\in\partial\Omega_{\mathrm{L}}} \psi(r,z) \quad \text{or} \quad (r_{\mathrm{bnd}}, z_{\mathrm{bnd}}) = \operatorname*{arg\,max}_{(r_{\mathrm{X}},z_{\mathrm{X}})\in\Omega_{\mathrm{L}}} \psi(r_{\mathrm{X}}, z_{\mathrm{X}}).$$

The coefficient λ is determined by forcing that

$$I_{\mathrm{P}} = \int_{\Omega(\psi)} j(r,z)\,dr\,dz \tag{10}$$

holds for some given *total plasma current* I_{P}.

In the following we will refer to (3)-(6) as the *quasi-static free-plasma-boundary problem*. The static case, the *static free-plasma-boundary problem* is the special case, where $\frac{\partial\psi}{\partial t}$ is set to zero in Ω_{c_i} and the currents I_i in the coils are assumed to be known input parameters. The first lines of (3) and (6) are known as the *Grad-Shafranov-Schlüter* equation [12, 22, 19].

2.2. **Variational Formulation on the Truncated Domain.** Let $\Omega \subset \Omega_\infty$ be a sufficiently large semi-circle of radius ρ_Γ, that is centered at the origin and contains the geometry of the tokamak. The boundary $\partial\Omega$ splits into $\Gamma_{r=0} := \{(r,z)\,,\, r = 0\}$ and $\Gamma = \partial\Omega \setminus \Gamma_{r=0}$. The variational formulation of (3) uses the following Sobolev space:

$$V := \left\{ \psi : \Omega \to \mathbb{R}, \int_\Omega \psi^2 r\,dr\,dz < \infty, \int_\Omega (\nabla\psi)^2 r^{-1}\,dr\,dz < \infty \right\} \cap C^0(\Omega). \tag{11}$$

Then we define

- two mappings $\mathsf{A} : V \times V \to \mathbb{R}$ and $\mathsf{J}_{\mathrm{p}} : V \times V \to \mathbb{R}$ that are linear in the last argument:

$$\begin{aligned} \mathsf{A}(\psi,\xi) &:= \int_\Omega \frac{1}{\mu(\psi) r} \nabla\psi \cdot \nabla\xi\,dr\,dz \\ \mathsf{J}_{\mathrm{p}}(\psi,\xi) &:= \int_{\Omega_{\mathrm{p}}(\psi)} j_{\mathrm{p}}(r, \psi_{\mathrm{N}}(\psi(r,z), \psi_{\mathrm{ax}}, \psi_{\mathrm{bnd}}))\xi\,dr\,dz \end{aligned} \tag{12}$$

- $N+1$ bilinear forms $\mathsf{j}_{\mathrm{cv}}, \mathsf{j}_{\mathrm{c}_i} : V \times V \to \mathbb{R}$

$$\begin{aligned} \mathsf{j}_{\mathrm{cv}}(\psi,\xi) &:= -\int_{\Omega_{\mathrm{ps}}} \frac{\sigma}{r}\psi\xi \, dr dz \\ \mathsf{j}_{\mathrm{c}_i}(\psi,\xi) &:= -2\pi \frac{n_i^2}{R_i S_i^2} \left(\int_{\Omega_{\mathrm{c}_i}} \psi \, dr dz\right)\left(\int_{\Omega_{\mathrm{c}_i}} \xi \, dr dz\right) \end{aligned} \tag{13}$$

- N bilinear mappings $\ell_{\mathrm{c}_i} : \mathbb{R} \times V \to \mathbb{R}$:

$$\ell_{\mathrm{c}_i}(V,\xi) := \frac{n_i V}{R_i S_i} \int_{\Omega_{\mathrm{c}_i}} \xi \, dr dz \tag{14}$$

- a bilinear form $\mathsf{C} : V \times V$ on Γ, that accounts for the boundary conditions at infinity [1]:

$$\begin{aligned} \mathsf{C}(\psi,\xi) :=& \frac{1}{\mu_0}\int_\Gamma \psi(\mathbf{P}_1) N(\mathbf{P}_1)\xi(\mathbf{P}_1) dS_1 \\ &+ \frac{1}{2\mu_0}\int_\Gamma\int_\Gamma (\psi(\mathbf{P}_1)-\psi(\mathbf{P}_2)) M(\mathbf{P}_1,\mathbf{P}_2)(\xi(\mathbf{P}_1)-\xi(\mathbf{P}_2)) dS_1 dS_2. \end{aligned} \tag{15}$$

with

$$\begin{aligned} M(\mathbf{P}_1,\mathbf{P}_2) &= \frac{k_{\mathbf{P}_1,\mathbf{P}_2}}{2\pi(r_1 r_2)^{\frac{3}{2}}}\left(\frac{2-k^2_{\mathbf{P}_1,\mathbf{P}_2}}{2-2k^2_{\mathbf{P}_1,\mathbf{P}_2}} E(k_{\mathbf{P}_1,\mathbf{P}_2}) - K(k_{\mathbf{P}_1,\mathbf{P}_2})\right) \\ N(\mathbf{P}_1) &= \frac{1}{r_1}\left(\frac{1}{\delta_+}+\frac{1}{\delta_-}-\frac{1}{\rho_\Gamma}\right) \text{ and } \delta_\pm = \sqrt{r_1^2+(z_1-\rho_\Gamma)^2}\,, \end{aligned}$$

where $\mathbf{P}_i = (r_i, z_i)$ and K and E the complete elliptic integrals of first and second kind, respectively and

$$k_{\mathbf{P}_j,\mathbf{P}_k} = \sqrt{\frac{4 r_j r_k}{(r_j-r_k)^2+(z_j-z_k)^2}}\,.$$

We refer to [13, Chapter 2.4] for the details of the derivation. The bilinear form $\mathsf{C}(\cdot,\cdot)$ follows basically from the so called *uncoupling procedure* in [9] for the usual coupling of boundary integral and finite element methods. The Green's function that is used in the derivation of the boundary integral methods for our problem was used earlier in finite difference methods for the Grad-Shafranov-Schlüter equations [18].

We introduce $\dot{\psi}$ for the derivative $\frac{\partial \psi}{\partial t}$ and derive the following variational formulation of the *quasi-static free-plasma-boundary problem*: Let the voltages $V_1(t), \dots V_N(t)$ and the initial data ψ^0 and λ^0 be given. Find $\psi(t) \in V$ and $\lambda(t) \in R$, $t \in [0,T]$ such that for all $\xi \in V$:

$$\begin{aligned} \mathsf{A}(\psi,\xi) - \lambda\, \mathsf{J}_{\mathrm{p}}(\psi,\xi) - \mathsf{j}_{\mathrm{cv}}(\dot{\psi},\xi) - \sum_{i=1}^N \mathsf{j}_{\mathrm{c}_i}(\dot{\psi},\xi) + C(\psi,\xi) &= \sum_{i=1}^N \ell_{\mathrm{c}_i}(V_i,\xi), \\ I_{\mathrm{p}} - \lambda\, \mathsf{J}_{\mathrm{p}}(\psi,1) = 0, \quad \psi(0) = \psi^0, \quad \lambda(0) &= \lambda^0\,. \end{aligned} \tag{16}$$

The implicit Euler method on $0 := t_0 < t_0 + \Delta t_1 = t_1 < \dots t_{n-1} + \Delta t_n = t_n = T$ gives the following semi-discretization for $\psi_{(k)}$ and $\lambda_{(k)}$ approximating $\psi(t_k)$ and

$\lambda(t_k)$:

$$\begin{aligned}
&\Delta t_k \mathsf{A}(\psi_{(k)},\xi) - \Delta t_k \lambda_{(k)}\, \mathsf{J}_{\mathrm{p}}(\psi_{(k)},\xi) - \mathsf{j}_{\mathrm{cv}}(\psi_{(k)}-\psi_{(k-1)},\xi) \\
&\qquad - \sum_{i=1}^{N} \mathsf{j}_{\mathrm{c}_i}(\psi_{(k)}-\psi_{(k-1)},\xi) + \Delta t_k C(\psi_{(k)},\xi) = \Delta t_k \sum_{i=1}^{N} \ell_{\mathrm{c}_i}(V_i(t_k),\xi), \\
&\qquad I_{\mathrm{p}}(t_k) - \lambda_{(k)}\, \mathsf{J}_{\mathrm{p}}(\psi_{(k)},1) = 0, \quad \psi_{(0)} = \psi^0, \quad \lambda_{(0)} = \lambda^0 .
\end{aligned} \tag{17}$$

To define Newton-type methods for (16) and (17) we need to calculate all the directional derivatives $D_\psi \mathsf{A}(\psi,\xi)(\widetilde{\psi})$, $D_\psi \mathsf{J}_{\mathrm{p}}(\psi,\xi)(\widetilde{\psi})$, $D_\psi \mathsf{j}_{\mathrm{cv}}(\psi,\xi)(\tilde{\psi})$, $D_\psi \mathsf{j}_{\mathrm{c}_i}(\psi,\xi)(\tilde{\psi})$ and $D_\psi C(\psi,\xi)(\tilde{\psi})$. This calculation is simple for the bilinear mappings $\mathsf{j}_{\mathrm{c}_i}$, j_{cv} and the non-linear mapping A (see (5)). The remaining derivatives are given in [5, p. 238].

2.3. **The Inverse Problem.** The inverse problem that corresponds to the equations (3)-(6) is the problem of determining external voltages such that the solution of (16) has certain prescribed properties. We will state this problem more rigorously as *optimal control problem.*

Let $\Gamma_{\mathrm{desi}}(t) \subset \Omega_{\mathrm{L}}$ denote the evolution of a closed line, contained in the domain Ω_{L} that is either smooth and touches the limiter at one point or has at least one corner. The former case prescribes a desired plasma boundary that touches the limiter. The latter case aims at a plasma with X-point that is entirely in the interior of Ω_{L}. Further let $(r_{\mathrm{desi}}(t), z_{\mathrm{desi}}(t)) \in \Gamma_{\mathrm{desi}}(t)$ and $(r_1(t), z_1(t)), \dots (r_{N_{\mathrm{desi}}}(t), z_{N_{\mathrm{desi}}}(t)) \in \Gamma_{\mathrm{desi}}(t)$ be $N_{\mathrm{desi}}+1$ points on that line. We define a *quadratic* functional $K(\psi)$ that evaluates to zero if $\Gamma_{\mathrm{desi}}(t)$ is an $\psi(t)$-iso-line, i.e. if $\psi(t)$ is constant on $\Gamma_{\mathrm{desi}}(t)$:

$$K(\psi) := \frac{1}{2} \int_0^T \left(\sum_{i=1}^{N_{\mathrm{desi}}} \big(\psi(r_i(t), z_i(t), t) - \psi(r_{\mathrm{desi}}(t), z_{\mathrm{desi}}(t), t)\big)^2 \right) dt. \tag{18}$$

Another functional, that will serve as *regularization*, is

$$R(V_1, \dots, V_N) := \sum_{i=1}^{N} \frac{w_i}{2} \int_0^T (V_i(t))^2\, dt, \tag{19}$$

that penalizes the strength of the voltages V_i and represents the energetic cost of the coil system. The coefficients $w_i \geq 0$ are called *regularization weights.*

Problem 1. Inverse Quasi-Static Free-Plasma-Boundary Problem. *We assume: that the geometric quantities Ω_{c_i} (S_i), $1 \leq i \leq N$, Ω_{ps}, Ω_{f}, Ω_{L} are given; that the physical quantities R_i and N_i, $1 \leq i \leq N$, and σ, μ_0 and μ_f are given; that the function $j_{\mathrm{p}} : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ is given; that the total plasma current $I_{\mathrm{p}}(t)$ is given and that the initial data ψ^0 and λ^0 is given. Define the control vector $\mathbf{u}$, with components $\mathbf{u}_i = V_i$, $0 < i \leq N$, the state vector $\mathbf{y} := (\psi, \lambda)$ and functional $J(\mathbf{y},\mathbf{u}) := K(\psi) + R(V_1, \dots, V_N)$. Solve the following non-linear optimal control problem:*

$$\min_{\mathbf{y},\mathbf{u}} J(\mathbf{y},\mathbf{u}) \quad \textit{subject to } (16)$$

The inverse problem that corresponds to the static case of equations (3)-(6) is the problem of finding externally applied currents I_i in the coils that correspond to a given shape and position of the plasma. The numerical solution of this inverse

problem, the *inverse static free-plasma-boundary problem* was previously treated in e.g. [3] or [13].

3. Numerical Methods

3.1. **Fast Algorithms for Optimal Control Problems.** The inverse Problem 1 is a *constrained optimization* problem, where the constraints are a set of partial differential equations that model the plasma equilibrium in a tokamak. In the mathematical literature such type of problems are called *optimal control* problems. The following section is a short summary on algorithms for general optimal control problems.

We consider the following optimal control problem

$$\min_{\mathbf{u},\mathbf{y}} \frac{1}{2}\langle \mathbf{u}, \mathbf{H}\mathbf{u}\rangle + \frac{1}{2}\langle \mathbf{y}, \mathbf{K}\mathbf{y}\rangle \quad \text{s.t} \qquad \mathbf{B}(\mathbf{y}) = \mathbf{F}(\mathbf{u}), \tag{20}$$

where $\frac{1}{2}\langle \cdot, \mathbf{H}\cdot\rangle$ and $\frac{1}{2}\langle \cdot, \mathbf{K}\cdot\rangle$, are quadratic cost functions and $\mathbf{y}$ and $\mathbf{u}$ the so-called state and control variables. In our setting $\mathbf{y}$ will be the variable that describes the plasma and $\mathbf{u}$ will be the externally applied currents or voltages. In a continuous setting, where the state $\mathbf{y}$, e.g. the flux ψ is a function of space, $< \cdot, \mathbf{H}\cdot >$ and $< \cdot, \mathbf{K}\cdot >$ are weighted L^2-innerproducts, and $\mathbf{B}(\mathbf{y})$ and $\mathbf{F}(\mathbf{u})$ are non-linear operators with the variational formulation (16) or (17). In the discrete setting, e.g. a finite element framework, where $\mathbf{y}$ are finitely many variables, $< \cdot, \mathbf{H}\cdot >$ and $< \cdot, \mathbf{K}\cdot >$ are weighted Euclidian innerproducts and $\mathbf{B}(\mathbf{y})$ and $\mathbf{F}(\mathbf{u})$ are proper approximations of the non-linear operators in the variational formulation (16) or (17).We refer to [14, Chapter 1] for more general formulations with non-linear cost functions and non-separating constraints. *Sequential Quadratic Programming (SQP)* is one of the most effective methods for non-linear constrained optimization with significant non-linearities in the constraints [21, Chapter 18]. SQP methods find a numerical solution by generating iteration steps that minimize quadratic cost functions subject to linear constraints. The Lagrange function formalism in combination with Newton-type iterations is one approach to derive the SQP-methods: the Lagrangian for (20) is

$$L(\mathbf{y}, \mathbf{u}, \mathbf{p}) = \frac{1}{2}\langle \mathbf{u}, \mathbf{H}\mathbf{u}\rangle + \frac{1}{2}\langle \mathbf{y}, \mathbf{K}\mathbf{y}\rangle + \langle \mathbf{p}, \mathbf{B}(\mathbf{y}) - \mathbf{F}(\mathbf{u})\rangle \tag{21}$$

and the solution of (20) is a stationary point of this Lagrangian:

$$\begin{aligned} \mathbf{K}\mathbf{y} + D_{\mathbf{y}}\mathbf{B}^T(\mathbf{y})\mathbf{p} &= 0, \\ \mathbf{H}\mathbf{u} - D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u})\mathbf{p} &= 0, \\ \mathbf{B}(\mathbf{y}) - \mathbf{F}(\mathbf{u}) &= 0 \end{aligned} \tag{22}$$

The superscript T indicates the adjoint operator, which corresponds to matrix transposition in the finite dimensional case. The second line in (22) corresponds to the optimality condition for the gradient of the reduced cost functional $\frac{1}{2}\langle \mathbf{u}, \mathbf{H}\mathbf{u}\rangle + \frac{1}{2}\langle \mathbf{y}(\mathbf{u}), \mathbf{K}\mathbf{y}(\mathbf{u})\rangle$, where $\mathbf{y}(\mathbf{u})$ is implicitly defined by $\mathbf{B}(\mathbf{y}(\mathbf{u})) = \mathbf{F}(\mathbf{u})$. This is the main reason for which gradient type methods for a corresponding unconstrained optimization problem for the reduced cost function are too expensive: one evaluation of the gradient requires the very expensive solution of the non-linear problem in the third line of (22).

A quasi-Newton method for solving (22) are iterations of the type

$$\begin{pmatrix} \mathbf{K} & 0 & D_{\mathbf{y}}\mathbf{B}^T(\mathbf{y}^k) \\ 0 & \mathbf{H} & -D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k) \\ D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k) & -D_{\mathbf{u}}\mathbf{F}(\mathbf{u}^k) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1}-\mathbf{y}^k \\ \mathbf{u}^{k+1}-\mathbf{u}^k \\ \mathbf{p}^{k+1}-\mathbf{p}^k \end{pmatrix} = - \begin{pmatrix} \mathbf{K}\mathbf{y}^k + D_{\mathbf{y}}\mathbf{B}^T(\mathbf{y}^k)\mathbf{p}^k \\ \mathbf{H}\mathbf{u}^k - D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k)\mathbf{p}^k \\ \mathbf{B}(\mathbf{y}^k) - \mathbf{F}(\mathbf{u}^k) \end{pmatrix} \tag{23}$$

which corresponds to the following quadratic optimization problem with linear constraints:

$$\begin{aligned} &\min_{\mathbf{u}^{k+1},\mathbf{y}^{k+1}} \frac{1}{2}\langle \mathbf{u}^{k+1}, \mathbf{H}\mathbf{u}^{k+1}\rangle + \frac{1}{2}\langle \mathbf{y}^{k+1}, \mathbf{K}\mathbf{y}^{k+1}\rangle \\ \text{s.t} \quad &\mathbf{B}(\mathbf{y}^k) + D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)(\mathbf{y}^{k+1}-\mathbf{y}^k) = \mathbf{F}(\mathbf{u}^k) + D_{\mathbf{u}}\mathbf{F}(\mathbf{u}^k)(\mathbf{u}^{k+1}-\mathbf{u}^k), \end{aligned} \tag{24}$$

We call (23) a quasi-Newton method since we omit the second order derivatives of $\mathbf{B}$ and $\mathbf{F}$. If the linear systems in (23) become too large, we could pursue the usual Schur complement approach: and obtain the following linear system for the increment $\Delta\mathbf{u}^k := \mathbf{u}^{k+1} - \mathbf{u}^k$

$$\mathbf{M}(\mathbf{y}^k, \mathbf{u}^k)\Delta\mathbf{u}^k = \mathbf{h}(\mathbf{y}^k, \mathbf{u}^k) \tag{25}$$

with

$$\begin{aligned} \mathbf{M}(\mathbf{y}^k, \mathbf{u}^k) &:= \mathbf{H} + D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k)D_{\mathbf{y}}\mathbf{B}^{-T}(\mathbf{y}^k)\mathbf{K}D_{\mathbf{y}}\mathbf{B}^{-1}(\mathbf{y}^k)D_{\mathbf{u}}\mathbf{F}(\mathbf{u}^k) \\ \mathbf{h}(\mathbf{y}^k, \mathbf{u}^k) &:= -\mathbf{H}\mathbf{u}^k - D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k)D_{\mathbf{y}}\mathbf{B}^{-T}(\mathbf{y}^k)\mathbf{K}D_{\mathbf{y}}\mathbf{B}^{-1}(\mathbf{y}^k)(\mathbf{r}(\mathbf{y}^k, \mathbf{u}^k)) \\ \mathbf{r}(\mathbf{y}^k, \mathbf{u}^k) &:= -\mathbf{B}(\mathbf{y}^k) + D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)\mathbf{y}^k + \mathbf{F}(\mathbf{u}^k). \end{aligned}$$

We will use iterative methods, e.g. the conjugate gradient methods, to solve (25). Since, in our case the number of control variables will be small, we can expect convergence within very few iterations. In each iteration step, we still have to solve the two linear systems corresponding to $D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)$ and $D_{\mathbf{y}}\mathbf{B}^T(\mathbf{y}^k)$. These are usually referred to as linearized direct and adjoint problems. Once we know $\Delta\mathbf{u}^k$ we could compute $\mathbf{y}^{k+1}$ and $\mathbf{p}^{k+1}$ by:

$$\begin{aligned} D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)\mathbf{y}^{k+1} &= \mathbf{r}(\mathbf{u}^k, \mathbf{y}^k) + D_{\mathbf{u}}\mathbf{F}(\mathbf{u}^k)\Delta\mathbf{u}^k \\ D_{\mathbf{y}}\mathbf{B}^T(\mathbf{y}^k)\mathbf{p}^{k+1} &= -\mathbf{K}\mathbf{y}^{k+1}. \end{aligned} \tag{26}$$

Fortunately, the conjugate gradient (CG) method for (25) allows for direct computation of $\mathbf{y}^{k+1}$ and $\mathbf{u}^{k+1}$ (see Algorithm 1).

We would like to highlight that the SQP-method relies on proper derivatives of the non-linear operators $\mathbf{B}$ and $\mathbf{F}$. In our case $\mathbf{F}$ is affine, hence the derivative of $\mathbf{B}$ remains the most difficult part. When it comes to the derivation of fully discrete schemes there are two choices:

(1) discretize the derivative of the continuous operator;
(2) compute the derivative of a discretized version of the continuous operator.

With the first approach, we cannot guarantee that the discretization yields a linear operator that is a derivative, hence we cannot guarantee that the numerical solution is the solution of the discrete optimization problem. Therefore it is nowadays considered to be important to follow the second approach.

Algorithm 1

Set $\Delta\mathbf{u}\leftarrow 0$;
Set $\mathbf{y}^{k+1} \leftarrow D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)^{-1}\mathbf{r}(\mathbf{u}^k, \mathbf{y}^k)$
Set $\mathbf{p}^{k+1} \leftarrow -D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)^{-T}\mathbf{K}\mathbf{y}^{k+1}$
Set $\mathbf{R}\leftarrow \mathbf{H}\mathbf{u}^k - D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k)\mathbf{p}^{k+1}$
Set $\mathbf{P}\leftarrow -\mathbf{R}$
while $\mathbf{R} \neq 0$ **do**
  $\mathbf{a} \leftarrow D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)^{-1}D_{\mathbf{u}}\mathbf{F}(\mathbf{u}^k)\mathbf{P}$
  $\mathbf{b} \leftarrow D_{\mathbf{y}}\mathbf{B}(\mathbf{y}^k)^{-T}\mathbf{a}$
  $\mathbf{M}\leftarrow \mathbf{H}\mathbf{P} + D_{\mathbf{u}}\mathbf{F}^T(\mathbf{u}^k)\mathbf{b}$
  $\alpha \leftarrow \frac{\mathbf{R}^T\mathbf{R}}{\mathbf{P}^T\mathbf{M}}$
  $\Delta\mathbf{u} \leftarrow \Delta\mathbf{u} + \alpha\mathbf{P}$
  $\mathbf{y}^{k+1} \leftarrow \mathbf{y}^{k+1} + \alpha\mathbf{a}$
  $\mathbf{p}^{k+1} \leftarrow \mathbf{p}^{k+1} - \alpha\mathbf{b}$
  $\widetilde{\mathbf{R}} \leftarrow \mathbf{R} + \alpha\mathbf{M}$
  $\beta \leftarrow \frac{\widetilde{\mathbf{R}}^T\widetilde{\mathbf{R}}}{\mathbf{R}^T\mathbf{R}}$
  $\widetilde{\mathbf{P}} \leftarrow -\widetilde{\mathbf{R}} + \beta\mathbf{P}$
  $\mathbf{R} \leftarrow \widetilde{\mathbf{R}}$, $\mathbf{P} \leftarrow \widetilde{\mathbf{P}}$
end while
$\mathbf{u}^{k+1} \leftarrow \mathbf{u}^k + \Delta\mathbf{u}$

3.2. **The Galerkin Discretization.** We introduce a triangulation Ω_h of the domain Ω that resolves the subdomains $\Omega_{\mathrm{L}}, \Omega_{\mathrm{f}}, \Omega_{c_i}, \Omega_{\mathrm{ps}}$ and use standard H^1-conforming Lagrangian finite elements with nodal degrees of freedom.

Let $\lambda_i(r,z)$ denote the Lagrangian basis functions associated to the vertices of the mesh. The finite element approximation ψ_h of ψ is:

$$\psi_h(r,z,t) = \sum_i \psi_i(t)\lambda_i(r,z) \text{ with } \psi_i \in \mathbb{R}. \tag{27}$$

The *domain of the plasma* $\Omega_{\mathrm{p}}(\psi_h)$ of a finite element function ψ_h is bounded by a continuous, piecewise straight, closed line. The critical points $(r_{\mathrm{ax}}(\psi_h), z_{\mathrm{ax}}(\psi_h))$ and $(r_{\mathrm{bnd}}(\psi_h), z_{\mathrm{bnd}}(\psi_h))$ are the coordinates of certain vertices of the mesh. The saddle point of a piecewise linear functions ψ_h is some vertex (r_0, z_0) with the following property: if $(r_1,z_1),(r_2,z_2)\ldots(r_n,z_n)$, denote the counterclockwise ordered neighboring vertices the sequence of discrete gradients $\psi_0-\psi_1$, $\psi_0-\psi_2\ldots\psi_0-\psi_n$ changes at least four times the sign.

To discretize the non-linear and linear operators in (16) we split the integrations over the entire domain Ω in sums of integrations over the elements T of the mesh Ω_h and apply some quadrature. Except for the bilinear form $\mathsf{C}(\psi,\xi)$ (see (15)), this procedure involves basically two different types of integrals:

$$\int_T f(r,\psi_h,\lambda_i)drdz \quad \text{and} \quad \int_{T\cap\Omega_{\mathrm{p}}(\psi_h)} g(r,\psi_h,\lambda_i)drdz, \tag{28}$$

where the two integrands $T\ni(r,z)\mapsto f(r,\psi_h(r,z),\lambda_i(r,z))$ and $T\cap\Omega_{\mathrm{p}}(\psi)\ni(r,z)\mapsto g(r,\psi_h(r,z),\lambda_i(r,z))$ are smooth functions on the integration domains. The second type of integrals appears in J_{P} due to the fact that the mesh does not resolve the

boundary of the plasma domain Ω_{p}. In any case we will use the centers of gravity

$$\begin{aligned} \mathbf{b}_T &:= (r_T, z_T) \\ \mathbf{b}_T(\psi_h) &:= (r_T(\psi_h), z_T(\psi_h)) := (r_{T\cap\Omega_{\mathrm{p}}(\psi_h)}, z_{T\cap\Omega_{\mathrm{p}}(\psi_h)}) \end{aligned} \tag{29}$$

of the integration domains T or $T \cap \Omega_{\mathrm{p}}(\psi_h)$ as quadrature points. The barycenter for the second type of integrals depends itself on ψ_h. Our choice of quadrature rule introduces a consistency error of order $O(h^2)$, where h is the diameter of the triangle, i.e. the quadrature is exact for linear integrands.

For a triangle T with vertex coordinates $\mathbf{a}_i, \mathbf{a}_j, \mathbf{a}_k$ the center of gravity corresponds to the barycenter:

$$(r_T, z_T) = \frac{1}{3}(\mathbf{a}_i + \mathbf{a}_j + \mathbf{a}_k). \tag{30}$$

If the domain of integration is $T \cap \Omega_{\mathrm{p}}(\psi_h)$, we have to distinguish the two cases. $T \cap \Omega_{\mathrm{p}}(\psi_h)$ is either a triangle or a quadrilateral. Without loss of generality we assume that $\mathbf{a}_i, \mathbf{a}_j, \mathbf{a}_k$ is a counterclockwise ordering of the vertex coordinates of T and that $\partial\Omega_{\mathrm{p}}(\psi)$ intersects ∂T at two points $\mathbf{m}_k$ and $\mathbf{m}_j$ at the edges opposite to the vertices $\mathbf{a}_k$ and $\mathbf{a}_j$ (See Figure 2). The barycentric coordinates of the intersecting

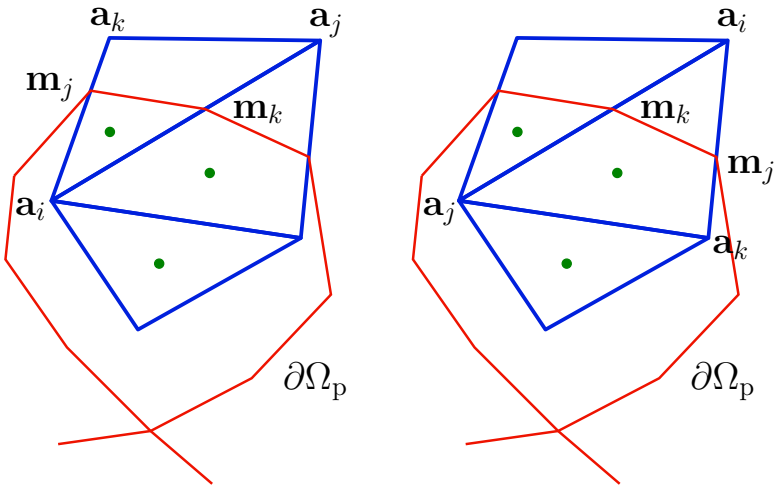


FIGURE 2. Integration over $T \cap \Omega_{\mathrm{p}}(\psi_h)$. The green dots indicate the location of the quadrature point. The integration domain $T \cap \Omega_{\mathrm{p}}(\psi_h)$ is either a) empty, b) the whole element T, c) a triangular domain or d) or quadrilateral domain.

points $\mathbf{m}_k$ and $\mathbf{m}_j$ are functions of ψ_h:

$$\begin{aligned} 1 - \lambda_i(\mathbf{m}_k) = \lambda_j(\mathbf{m}_k) = \lambda_j(\mathbf{m}_k(\psi_h)) = \frac{\psi_{\mathrm{bnd}}(\psi_h) - \psi_i}{\psi_j - \psi_i} \\ 1 - \lambda_i(\mathbf{m}_j) = \lambda_k(\mathbf{m}_j) = \lambda_k(\mathbf{m}_j(\psi_h)) = \frac{\psi_{\mathrm{bnd}}(\psi_h) - \psi_i}{\psi_k - \psi_i}, \end{aligned} \tag{31}$$

and, clearly, we have $\lambda_k(\mathbf{m}_k) = \lambda_j(\mathbf{m}_j) = 0$.

If $T \cap \Omega_{\mathrm{p}}(\psi_h)$ is a *triangle* and $\mathbf{a}_i$ that vertex of T that is contained in $T \cap \Omega_{\mathrm{p}}(\psi_h)$ (See Figure 2, left) we find:

$$(r_T(\psi_h), z_T(\psi_h)) = \mathbf{a}_i + \frac{1}{3}\lambda_j(\mathbf{m}_k)(\mathbf{a}_j - \mathbf{a}_i) + \frac{1}{3}\lambda_k(\mathbf{m}_j)(\mathbf{a}_k - \mathbf{a}_i) \tag{32}$$

and

$$|T \cap \Omega_{\mathrm{p}}(\psi_h)| = |T|\lambda_j(\mathbf{m}_k)\lambda_k(\mathbf{m}_j). \tag{33}$$

If $T \cap \Omega_{\mathrm{p}}(\psi_h)$ is a *quadrilateral* and $\mathbf{a}_i$ that vertex of T that is *not* contained in $T \cap \Omega_{\mathrm{p}}(\psi_h)$ (See Figure 2, right) we find:

$$\begin{aligned}(r_T(\psi_h), z_T(\psi_h)) \\ = \mathbf{a}_i + \frac{1}{3}\frac{1-\lambda_j^2(\mathbf{m}_k)\lambda_k(\mathbf{m}_j)}{1-\lambda_j(\mathbf{m}_k)\lambda_k(\mathbf{m}_j)}(\mathbf{a}_j - \mathbf{a}_i) + \frac{1}{3}\frac{1-\lambda_j(\mathbf{m}_k)\lambda_k^2(\mathbf{m}_j)}{1-\lambda_j(\mathbf{m}_k)\lambda_k(\mathbf{m}_j)}(\mathbf{a}_k - \mathbf{a}_i)\end{aligned} \tag{34}$$

and

$$|T \cap \Omega_{\mathrm{p}}(\psi_h)| = |T|\left(1 - \lambda_j(\mathbf{m}_k)\lambda_k(\mathbf{m}_j)\right). \tag{35}$$

3.3. **The Discrete Optimal Control Problem.** We introduce the subscript h to denote the fully discrete version of the operators involved in (17) that we obtain when we use the quadrature rules from the previous section 3.2. The fully discrete formulation of the dynamic free-plasma-boundary problem is: Find $(\psi_h^1, \lambda^1), \ldots, (\psi_h^n, \lambda^n)$, such that

$$\begin{aligned}&\Delta t_k \mathsf{A}_h(\psi_h^k, \xi) - \Delta t_k \lambda^k \mathsf{J}_{\mathrm{P},h}(\psi_h^k, \xi) - \mathsf{j}_{\mathrm{cv,h}}(\psi_h^k - \psi_h^{k-1}, \xi) \\ &\quad - \sum_{i=1}^{N} \mathsf{j}_{\mathrm{c}_i,h}(\psi_h^k - \psi_h^{k-1}, \xi) + \Delta t_k C(\psi_h^k, \xi) = \Delta t_k \sum_{i=1}^{N} \ell_{\mathrm{c}_i,h}(V_i(t_k), \xi), \\ &I_{\mathrm{p}}(t_k) - \lambda^k \mathsf{J}_{\mathrm{p},h}(\psi_h^k, 1) = 0, \quad \psi_h^0 = \psi^0, \quad \lambda^{(0)} = \lambda(t_0).\end{aligned} \tag{36}$$

To formulate the *discrete* version of the optimal control Problem 1 we define a discrete cost function $K_h(\{\psi_h^k\}_{k=1}^n)$:

$$K_h(\{\psi_h^k\}_{k=1}^n) = \sum_{k=1}^{n}\left(\frac{\Delta t_k}{2}\sum_{i=1}^{N_{\mathrm{desi}}}\left(\psi_h^k(r_i, z_i) - \psi_h^k(r_{\mathrm{desi}}(t_k), z_{\mathrm{desi}}(t_k))\right)^2\right) \tag{37}$$

and a discrete regularization function:

$$R_h(\{V_1(t_k)\}_{k=1}^n, \ldots, \{V_N(t_k)\}_{k=1}^n) = \sum_{i=1}^{N}\frac{w_i}{2}\sum_{k=1}^{n}\Delta t_k V_i^2(t_k)\,. \tag{38}$$

Problem 2. Discrete Inverse Dynamic Free-Plasma-Boundary Problem. *We assume that the geometric quantities Ω_{c_i} (S_i), $1 \le i \le N$, Ω_{ps}, Ω_{f}, Ω_{L} are given; that the physical quantities R_i and N_i, $1 \le i \le N$, and σ, μ_0 and μ_f are given; that the function $j_{\mathrm{p}} : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ is given; that the external sources $\{I_{\mathrm{p}}(t_k)\}_{k=0}^n$ are given, that the initial data ψ^0, λ^0 is given, Then we define the control vector $\bar{\mathbf{u}} := (\{V_1(t_k)\}_{k=1}^n, \ldots, \{V_N(t_k)\}_{k=1}^n)$, the state vector $\bar{\mathbf{y}} := (\psi_h^1, \lambda^1, \ldots, \psi_h^n, \lambda^n)$ and the functional*

$$J_h(\bar{\mathbf{y}}, \bar{\mathbf{u}}) := K_h(\{\psi_h^k\}_{k=1}^n) + R_h(\{V_1(t_k)\}_{k=1}^n, \ldots, \{V_N(t_k)\}_{k=1}^n). \tag{39}$$

Solve the following discrete non-linear optimal control problem:

$$\min_{\bar{\mathbf{y}}, \bar{\mathbf{u}}} J_h(\bar{\mathbf{y}}, \bar{\mathbf{u}}) \quad \textit{subject to } (36)$$

As it is indicated in the previous sections, we have to pay particular attention to the discretization of our non-linear constraints. We need to ensure that, loosely speaking, the linear constraints of the quadratic optimization problem during the SQP iterations are accurate derivatives of the non-linear constraints. Only then, the SQP-method guarantees convergence to the optimal solution.

Except for $\mathsf{J}_{\mathrm{P},h}(\psi_h^k,\xi)$ the derivatives for the non-linear operators in (36) are straight-forward and simple calculations. The non-linear operator $\mathsf{J}_{\mathrm{P},h}(\psi_h^k,\xi)$ involves quadrature points that depend on the numerical solution ψ_h and we have to pay particular attention to the derivative of this operator.

3.4. **The Derivative of the Discrete Current $\mathsf{J}_{\mathrm{P},h}$.** The discrete non-linear current is given by

$$\mathsf{J}_{\mathrm{P},h}(\psi_h,\lambda_m)=\sum_T \mathsf{J}^T_{\mathrm{P},h}(\psi_h,\lambda_m):=\sum_T |T\cap\Omega_{\mathrm{p}}(\psi_h)|\, j_{\mathrm{p}}(\mathbf{b}_T(\psi_h))\,\lambda_m(\mathbf{b}_T(\psi_h)),$$

where $j_{\mathrm{p}}(\mathbf{b}_T(\psi_h))=j_{\mathrm{p}}(r_T(\psi_h),\psi_{\mathrm{N}}(\psi_h(\mathbf{b}_T(\psi_h))),\psi_{\mathrm{ax}}(\psi_h),\psi_{\mathrm{bnd}}(\psi_h)))$. To find the derivative

$$D_\psi \mathsf{J}_{\mathrm{P},h}(\psi_h,\lambda_m)(\lambda_n)=\frac{d}{d\psi_n}\mathsf{J}_{\mathrm{P},h}(\psi_h,\lambda_m)$$

we compute the derivative for each term of the sum. Computing the derivative of each term is a tedious application of chain and product rules. We distinguish three different cases: $T\cap\Omega_{\mathrm{p}}(\psi_h)=0$, $T\cap\Omega_{\mathrm{p}}(\psi_h)=T$ and $T\cap\Omega_{\mathrm{p}}(\psi_h)\subset T$. With a slightly abuse of notation we identify ψ_{bnd} and ψ_{ax} with the corresponding finite element expansion coefficient and use the Dirac deltas $\delta_{n,\mathrm{bnd}}$ and $\delta_{n,\mathrm{ax}}$.

(1) $T\cap\Omega_{\mathrm{p}}(\psi_h)=0$:

$$\frac{d}{d\psi_n}\mathsf{J}^T_{\mathrm{P},h}(\psi_h,\lambda_m)=0\quad\forall n,m;$$

(2) $T\cap\Omega_{\mathrm{p}}(\psi_h)=T$:

$$\begin{aligned}\frac{d}{d\psi_n}\mathsf{J}^T_{\mathrm{P},h}(\psi_h,\lambda_m)=&|T|\frac{\partial j_{\mathrm{p}}(r_T,\psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial\psi_{\mathrm{N}}}\frac{\partial\psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial\psi_h}\lambda_n(\mathbf{b}_T)\lambda_m(\mathbf{b}_T)\\&+|T|\frac{\partial j_{\mathrm{p}}(r_T,\psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial\psi_{\mathrm{N}}}\frac{\partial\psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial\psi_{\mathrm{bnd}}}\delta_{n,\mathrm{bnd}}\lambda_m(\mathbf{b}_T)\\&+|T|\frac{\partial j_{\mathrm{p}}(r_T,\psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial\psi_{\mathrm{N}}}\frac{\partial\psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial\psi_{\mathrm{ax}}}\delta_{n,\mathrm{ax}}\lambda_m(\mathbf{b}_T)\end{aligned}$$

(3) $T\cap\Omega_{\mathrm{p}}(\psi_h)\subset T$: Without loss of generality we adopt the notation from section 3.2, introduce $\lambda_j^k=\lambda_j(\mathbf{m}_k)$ $\lambda_k^j=\lambda_k(\mathbf{m}_j)$ use $\mathbf{b}_T$ to denote $\mathbf{b}_T(\psi_h)$. We define $Ar=|T|\lambda_j^k\lambda_k^j$ if $T\cap\Omega_{\mathrm{p}}(\psi_h)$ is a triangle and $Ar=|T|(1-\lambda_j^k\lambda_k^j)$ if $T\cap\Omega_{\mathrm{p}}(\psi_h)$ is a quadrilateral. We find

$$\frac{d}{d\psi_n}\mathsf{J}^T_{\mathrm{P},h}(\psi_h,\lambda_m)=A_n(\psi_h,\lambda_m)+C_n(\psi_h,\lambda_m)+T_n(\psi_h,\lambda_m)$$

with

- the derivative related to the area $|T \cap \Omega_{\mathrm{p}}(\psi_h)|$:

$$\begin{aligned} A_n(\psi_h, \lambda_m) =& s|T| \left(\frac{\partial \lambda_j^k}{\partial \psi_n} \lambda_k^j + \lambda_k^j \frac{\partial \lambda_j^k}{\partial \psi_n} \right) j_{\mathrm{p}}(\mathbf{b}_T)\, \lambda_m(\mathbf{b}_T) \\ &+ s|T| \left(\frac{\partial \lambda_j^k}{\partial \psi_{\mathrm{bnd}}} \lambda_k^j + \lambda_k^j \frac{\partial \lambda_j^k}{\partial \psi_{\mathrm{bnd}}} \right) \delta_{n,\mathrm{bnd}}\, j_{\mathrm{p}}(\mathbf{b}_T)\, \lambda_m(\mathbf{b}_T), \end{aligned}$$

  where $s = 1$ if $|T \cap \Omega_{\mathrm{p}}(\psi_h)|$ is a triangle and $s = -1$ else.
- the derivative related to the current $j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))$:

$$\begin{aligned} C_n(\psi_h, \lambda_m) =& Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial \psi_{\mathrm{N}}} \frac{\partial \psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial \psi_h} \lambda_n(\mathbf{b}_T) \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial \psi_{\mathrm{N}}} \frac{\partial \psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial \psi_h} \nabla \psi_h(\mathbf{b}_T) \cdot \frac{\partial \mathbf{b}_T}{\partial \psi_n} \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial \psi_{\mathrm{N}}} \frac{\partial \psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial \psi_h} \nabla \psi_h(\mathbf{b}_T) \cdot \frac{\partial \mathbf{b}_T}{\partial \psi_{\mathrm{bnd}}} \delta_{n,\mathrm{bnd}} \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial \psi_{\mathrm{N}}} \frac{\partial \psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial \psi_{\mathrm{bnd}}} \delta_{n,\mathrm{bnd}} \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial \psi_{\mathrm{N}}} \frac{\partial \psi_{\mathrm{N}}(\mathbf{b}_T)}{\partial \psi_{\mathrm{ax}}} \delta_{n,\mathrm{ax}} \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial r} \frac{dr_T}{d\psi_n} \lambda_m(\mathbf{b}_T) \\ &+ Ar \frac{\partial j_{\mathrm{p}}(r_T, \psi_{\mathrm{N}}(\mathbf{b}_T))}{\partial r} \frac{\partial r_T}{\partial \psi_{\mathrm{bnd}}} \delta_{n,\mathrm{bnd}} \lambda_m(\mathbf{b}_T) \end{aligned}$$

- the derivative related to the test function $\lambda_i(\mathbf{b}_T)$:

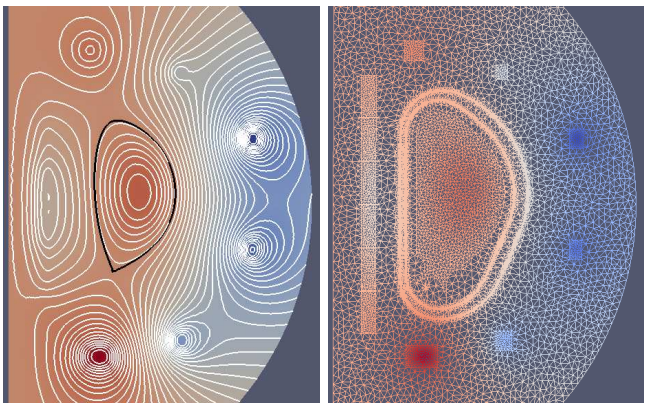
$$T_n(\psi_h, \lambda_m) = Ar \left(j_{\mathrm{p}}(\mathbf{b}_T)\, \nabla \lambda_m(\mathbf{b}_T) \cdot \frac{\partial \mathbf{b}_T}{\partial \psi_n} + j_{\mathrm{p}}(\mathbf{b}_T)\, \nabla \lambda_m(\mathbf{b}_T) \cdot \frac{\partial \mathbf{b}_T}{\partial \psi_{\mathrm{bnd}}} \delta_{n,\mathrm{bnd}} \right)$$

The derivatives of ψ_{N} follow easily from (8). We would like to stress that the Galerkin matrix $D_\psi \mathbf{J}_{\mathrm{P},h}^T(\psi_h, \lambda_m)(\lambda_n)$ can be assembled in a fairly standard, i.e. element wise, fashion, provided we compute in a preprocessing step the following information for each element: we know for each element T to which of the three different cases it belongs. If an element T belongs to the last case we need to know the barycententric coordinates of the intersection points $\lambda_k(\mathbf{m}_j)$ and $\lambda_j(\mathbf{m}_k)$, the barycenter $\mathbf{b}_T(\psi_h)$ and the derivatives $\frac{\partial \lambda_k(\mathbf{m}_j)}{\partial \psi_i}$, $\frac{\partial \lambda_k(\mathbf{m}_j)}{\partial \psi_j}$, $\frac{\partial \lambda_k(\mathbf{m}_j)}{\partial \psi_k}$, $\frac{\partial \lambda_k(\mathbf{m}_j)}{\partial \psi_{\mathrm{bnd}}}$, $\frac{\partial \lambda_j(\mathbf{m}_k)}{\partial \psi_i}$, $\frac{\partial \lambda_j(\mathbf{m}_k)}{\partial \psi_j}$, $\frac{\partial \lambda_j(\mathbf{m}_k)}{\partial \psi_k}$, $\frac{\partial \lambda_j(\mathbf{m}_k)}{\partial \psi_{\mathrm{bnd}}}$ and $\frac{\partial \mathbf{b}_T}{\partial \psi_i}$, $\frac{\partial \mathbf{b}_T}{\partial \psi_j}$, $\frac{\partial \mathbf{b}_T}{\partial \psi_k}$, $\frac{\partial \mathbf{b}_T}{\partial \psi_{\mathrm{bnd}}}$. All this information can be easily computed for given ψ_h, ψ_{bnd} and ψ_{ax} using the formulas (31), (32) and (34). All the terms that contain the Dirac deltas $\delta_{n,\mathrm{bnd}}$ or $\delta_{n,\mathrm{ax}}$ lead to non-local entries in the stiffness matrix. They connect all the coefficients of ψ_{bnd} and ψ_{ax} with all coefficients that are associated to vertices of elements that are intersected by the boundary $\partial\Omega_{\mathrm{P}}(\psi_h)$.

4. Numerical Tests

We present three different numerical tests, that verify that we use correct gradients in our implementation. In the following calculation we use a ITER-like geometry and an equlibrium plasma (see Fig. 3) that corresponds to the currents

in the table of Figure 3. The total plasma current is $I_{\rm P} = 15 \times 10^6 A$ and the four parameters for the current profile (7) are $R_0 = 6.2m$, $\alpha = 2.0$, $\beta = 0.5978$ and $\gamma = 1.395$.



coil	I_i
U3	-1.400×10^6
U2	-9.500×10^6
U1	-20.388×10^6
L1	-20.388×10^6
L2	-9.000×10^6
L3	3.564×10^6
P1	5.469×10^6
P2	-2.266×10^6
P3	-6.426×10^6
P4	-4.820×10^6
P5	-7.504×10^6
P6	17.240×10^6

FIGURE 3. Left: The plasma (flux lines and flux intensity) that correspond to the currents in the table. Center: The triangulation of the ITER-geometry. Right: Data for coils. Coils U1-U3 and L1-L3 are upper and lower position control coils. Coils P1-P6 are the poloidal field coils.

4.1. **Convergence of finite differences of $\mathsf{J}_{\mathrm{P},h}$ to $D_\psi \mathsf{J}_{\mathrm{P},h}$.** Let $\mathbf{J}(\psi_h)$ denote the Galerkin matrix of the derivative $D_\psi \mathsf{J}_{\mathrm{P},h}(\psi_h, \lambda_m)(\lambda_n)$ of the bilinear form $\mathsf{J}_{\mathrm{P},h}(\psi_h, \lambda_m)$. In a first test we verify that for given ψ_h and perturbation $\delta\psi_h = \sum_n \lambda_n \mathbf{v}_n$ our implementation yields:

$$E_{\rm FD}(\varepsilon) := \frac{\left\| \varepsilon^{-1} \begin{pmatrix} \mathsf{J}_{\mathrm{P},h}(\psi_h + \varepsilon\delta\psi_h, \lambda_1) - \mathsf{J}_{\mathrm{P},h}(\psi_h, \lambda_1) \\ \mathsf{J}_{\mathrm{P},h}(\psi_h + \varepsilon\delta\psi_h, \lambda_2) - \mathsf{J}_{\mathrm{P},h}(\psi_h, \lambda_2) \\ \vdots \end{pmatrix} - \mathbf{J}(\psi_h)\mathbf{v} \right\|}{\|\mathbf{J}(\psi_h)\mathbf{v}\|} = O(\epsilon). \tag{40}$$

The perturbation increment $\delta\psi_h$ is randomly chosen. In Table 1 we monitor this relative error and observe, as expected, first order convergence.

4.2. **Convergence of the Sensitivities.** We consider the variational formulation of the static free-plasma-boundary problem for given currents $I_1, \dots I_N$: Find ψ such that

$$\begin{aligned} \mathsf{A}(\psi, \xi) - \lambda\, \mathsf{J}_{\rm P}(\psi, \xi) + C(\psi, \xi) &= \sum_{i=1}^{N} \ell_{\mathrm{c}_i}(\frac{R_i}{n_i} I_i, \xi), \\ I_{\rm p} - \lambda\, \mathsf{J}_{\mathrm{p},h}(\psi, 1) &= 0\,. \end{aligned} \tag{41}$$

When we solve this non-linear problem with a *real* Newton method, the convergence theory for Newton methods asserts that we will have second order convergence, when we use accurate derivatives of the non-linear operators. But, since in many cases also with inaccurate derivatives one observes super-linear convergence, this second order convergence is an inappropriate indicator for accurate or inaccurate derivatives.

$\varepsilon_i = 0.5^i$	$E_{\mathrm{FD}}(\varepsilon_i)$	rate	$E_{\mathrm{RN}}(\varepsilon_i)$	rate	$E_{\mathrm{QN}}(\varepsilon_i)$	rate
0	0.0150819		11647.2		2928.07	
1	0.0075431	0.99958	5485.42	1.08631	1125.83	1.37897
2	0.0037716	0.99998	3362.31	0.70615	1182.56	−0.07086
3	0.0018859	0.99997	1936.28	0.79616	846.373	0.48247
4	0.0009429	0.99997	694.294	1.47967	149.343	2.50266
5	0.0004715	0.99998	4.26942	7.34536	268.207	−0.84472
6	0.0002357	0.99999	9.65561	−1.17733	145.894	0.87843
7	0.0001179	0.99998	2.39544	2.01108	70.5145	1.04893
8	0.0000589	0.99997	0.60127	1.99421	34.6608	1.02462
9	0.0000295	0.99973	0.15064	1.99689	17.1804	1.01254
10	0.0000147	1.00045	0.03770	1.99843	8.55259	1.00633
11	0.0000074	0.99865	0.00943	1.99919	4.26687	1.00318
12	0.0000037	0.99134	0.00236	1.99959	2.13108	1.00159
13	0.0000019	0.92854	0.00059	1.99982	1.06495	1.00081
14	0.0000011	0.870235	0.00015	1.99993	0.53233	1.00043
15	0.0000004	1.568283	0.00003	2.00002	0.26613	1.00028

TABLE 1. Convergence and convergence rate $(\frac{log(E_{\dots}(\varepsilon_{i+1}))-log(E_{\dots}(\varepsilon_i))}{log(\varepsilon_{i+1})-log(\varepsilon_i)})$ of the finite difference error E_{FD} and the errors E_{RN} and E_{QN} of the real Newton method described in Section 3.2 and the quasi-Newton method in [3, 13].

A better indicator is the following test: Let $\psi_h(I_1, \dots, I_n)$ denote the discrete solution for given control parameters $I_1, \dots I_N$ and $I_1(\varepsilon), \dots, I_N(\varepsilon)$ is some perturbation of the data. A *real* Newton method that determines $\psi_h(I_1(\varepsilon), \dots, I_n(\varepsilon))$ and uses $\psi_h^0(I_1(\varepsilon), \dots, I_n(\varepsilon)) = \psi_h(I_1, \dots, I_n)$ as initial guess, yields already in the first iteration an approximation $\psi_h^1(I_1(\varepsilon), \dots, I_n(\varepsilon))$ that is a second order perturbation of the solution.

Here, we will solve the non-linear static free-plasma-boundary problem (41) with the Newton method outlined in Section 3.2 for the data given in the table in Figure 3. As perturbation we choose a random incremental current for each coil and scale it with $\varepsilon = 0.5^0, \dots 0.5^{15}$. In Table 1 we monitor the error

$$E_{\mathrm{RN}}(\varepsilon) = \|\mathbf{v}(\varepsilon) - \mathbf{v}^1(\varepsilon)\|, \tag{42}$$

where $\mathbf{v}(\varepsilon)$ and $\mathbf{v}^1(\varepsilon)$ are the coefficients of the solution $\psi_h(I_1(\varepsilon), \dots, I_n(\varepsilon))$ and the first Newton iterate $\psi_h^1(I_1(\varepsilon), \dots, I_n(\varepsilon))$. We get second order convergence, which shows that we use accurate derivatives. In contrast, the results of the Newton-like method that follows from the discretization of the continuous Newton method in [3, 13], yields only first order convergence. This second approach does not use an accurate derivative.

4.3. **Quasi-Static Free-Plasma-Boundary Problem.** In this example we study the inverse quasi-static free-plasma-boundary problem defined in Problem 2. For demonstration purposes we present here a first example with 5 timesteps with timestep length $\tau = 1s$, the plasma current varies in time $I_{\mathrm{P}}(t_1) = 692350.62A$, $I_{\mathrm{P}}(t_2) = 884463A$, $I_{\mathrm{P}}(t_3) = 1077116.1A$, $I_{\mathrm{P}}(t_4) = 1269026.7A$, $I_{\mathrm{P}}(t_5) = 1461458.2A$, the parameters of the current profile (7) are constant: $\alpha = 2$, $\beta = 0.09$, $\gamma = 1.95$, $R_0 = 6.211m$. The coils have total resistance $R_i = 5 \times 10^{-4}\Omega$ and wire turns n_i

Newton iteration	0	1	2	3	4
K_h	0.149969	0.010203	0.009634	0.009631	0.009631
R_h	4×10^{-6}	0.010082	0.009447	0.009448	0.009448
rel. residuum		5.37051	0.006889	1×10^{-5}	1×10^{-9}
CG iterations		60	58	60	60

TABLE 2. Dynamic free-plasma-boundary problem: convergence history.

coil	n_i	step 1	step 2	step 3	step 4	step 5
U3	553.0	597.146	601.843	595.93	552.652	432.379
U2	553.0	−749.273	−780.675	−728.05	−545.037	−124.193
U1	553.0	265.837	112.625	5.3857	−156.643	−345.023
L1	553.0	−140.032	−258.246	−333.515	−387.679	−322.303
L2	553.0	−296.528	−226.159	−117.576	53.0662	310.864
L3	553.0	386.966	426.462	452.892	447.854	367.449
P1	248.6	891.221	915.511	911.834	842.517	647.715
P2	115.2	−548.666	−442.296	−356.817	−247.353	−110.528
P3	185.9	−426.476	−490.581	−649.557	−819.796	−966.274
P4	168.9	−985.828	−967.609	−1068.51	−1150.05	−1157.72
P5	216.8	−342.551	−184.267	−50.9846	87.5257	192.708
P6	459.4	347.911	409.129	452.483	462.886	391.759

TABLE 3. Dynamic free-plasma-boundary problem: Computed optimal voltages.

as indicated in Table 3. We use 65 points to prescribe the desired boundary in each time step and deploy the SQP method to determine the optimal solution. The quadratic subproblems within the SQP iterations are solved by the CG method outlined in Section 3.1.

In this experiment the SQP-iterations are stopped if the relative residuum, the scaled righthand side in (23), is smaller than 10^{-7}. The stopping criterion for the conjugate gradient method is a residuum that is smaller than 10^{-7} or that the number of iterations exceeds the number of control parameters, i.e. 60. The Table 2 summarizes the convergence. The computed optimal voltages are shown in Table 3, the corresponding evolution of the plasma is depicted in Figure 4. We see a perfect agreement with the desired evolution of the boundary.

Finally, we would like to show first results for a so-called ramp-up scenario, where the plasma evolves from a small circular to a large elongated plasma. We simulate $60s$. The computed coil voltages are depicted in Figure 5. Then, if we use those as data to solve the direct problem we observe that the plasma boundary follows indeed the prescribed trajectory (see Figure 6).

5. Conclusions and Perspectives

The models that describe the evolution of plasma in tokamak devices are highly non-linear. Already a direct simulation, meaning a simulation that shows only how the plasma will evolve for certain given control inputs is a demanding task.

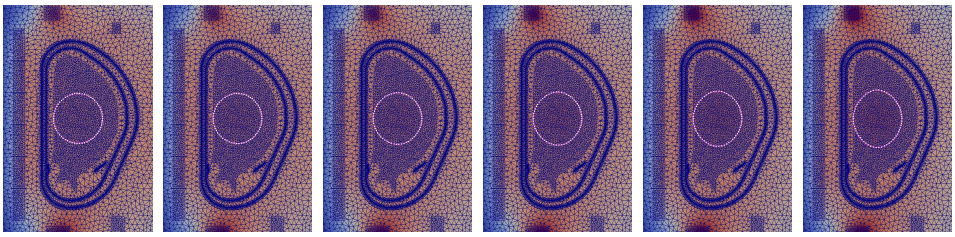


FIGURE 4. Dynamic free-plasma-boundary problem: the plasma boundary (magenta) follows the prescribed boundary (white points).

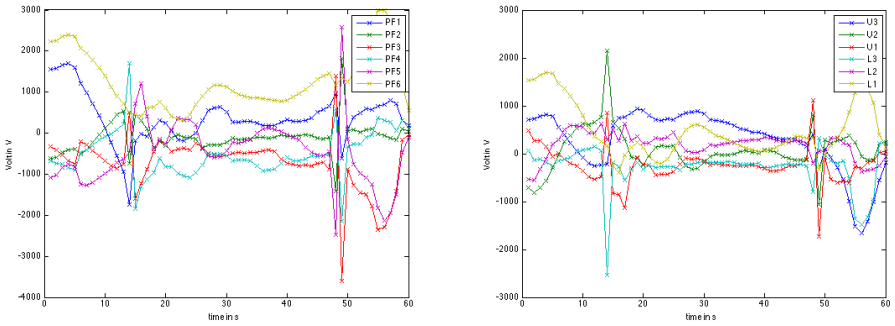


FIGURE 5. Instationary Equilibrium: The optimal voltages.

Elongated plasmas are known to suffer from serious physical instabilities [23, Section 6.15], and therefore also numerical methods will be very likely to be unstable.

On the other hand, the real pivotal question for the control of plasma is the task to find those control parameters that ensure that the plasma evolves through certain prescribed states. From the engineering point of view it would be beneficial to provide the majority of control as feed-forward control that can be determined a priori, and to minimize the amount of feed-back control that needs to be calculated during the operation. Until today, the calculation of the feed-forward control is mainly tackled by linear ODE models that have the flavor of electrical circuits. These ODE models determine the coil voltages that correspond to the coil currents of precomputed static equilibria. Particular feed-back controller are then necessary to ensure that the currents in the system are indeed close to the precomputed ones. The utility of this approach is underlined by the fact that control engineers have today the knowledge to operate the present tokamaks.

Nevertheless we believe that the computation of feed-forward control can be done much more accurately. In this work we formulated an optimal control problem, that uses the force balance (1) and Maxwell's equations (2) as constraints, to compute directly the optimal feed-forward control. This approach is much more consistent, since we use the full nonlinear PDE model. Then, in the solution of this optimization problem, it is Newton's method that determines consistently the linearization, that provide the linear approximation of the non-linear model.

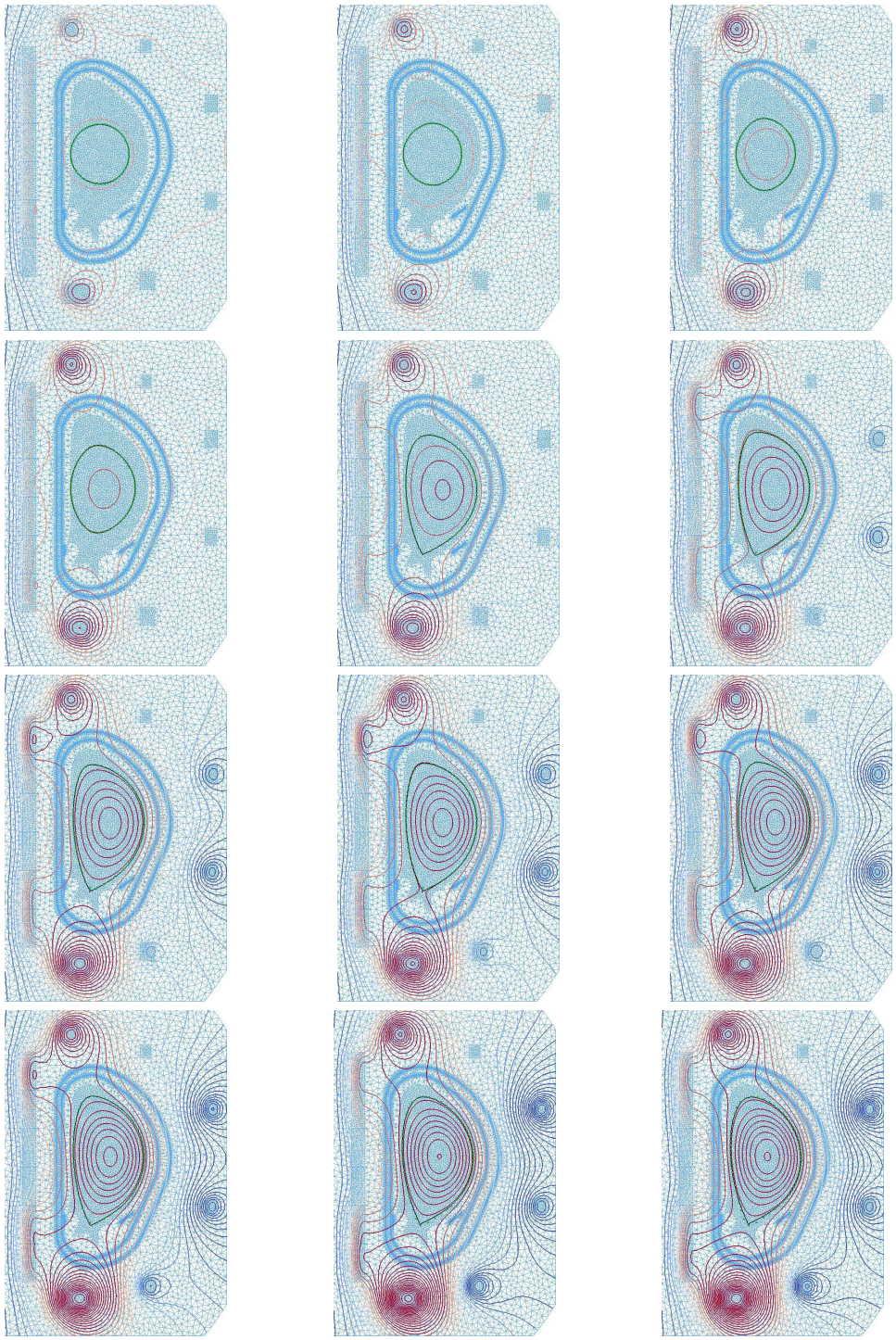


FIGURE 6. Optimal control for a ramp-up scenario: the plasma boundary (green) follows the prescribed boundary (black points), snapshots at $t = 0, 2, 6, 10, 20, 30, 40, 45, 50, 54, 58, 60s$.

We presented here the two basic ingredients, the finite element method and sequential quadratic programming, to set up such an optimal control approach for plasma evolution in tokamaks. The finite element method allows for an easy treatment of free plasma boundary and the decaying conditions at infinity. Further, we can, almost automatically, compute linearizations and the adjoint operators that are required for the optimization with sequential quadratic programming. In the future we will have to augment the non-linear model (3)-(6) with further equations, namely the resitive diffusion and transport equations, that provide a more complete description. Our preliminary simulations provide a fairly promising proof of concepts.

References

[1] R. Albanese, J. Blum, and O. Barbieri. On the solution of the magnetic flux equation in an infinite domain. In *EPS. 8th Europhysics Conference on Computing in Plasma Physics (1986)*, 1986.

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

[3] J. Blum. *Numerical simulation of the plasma equilibrium in a tokamak*. Wiley/Gauthier-Villars, 1987.

[4] J. Blum and J. Le Foll. Plasma equilibrium evolution at the resistive diffusion timescale. *Computer Physics Reports*, 1(7-8):465–494, 1984.

[5] J. Blum, J. Le Foll, and B. Thooris. The self-consistent equilibrium and diffusion code SCED. *Computer Physics Communications*, 24:235 – 254, 1981.

[6] J.A. Crotinger. Corsica; a comprehensive simulation of toroidal magnetic-fusion devices. Technical report ucrl-id-126284, Lawrence Livermore National Laboratory, 1997.

[7] E. Fable, C. Angioni, A. A. Ivanov, K. Lackner, O. Maj, S. Yu, Medvedev, G. Pautasso, and G. V. Pereverzev. A stable scheme for computation of coupled transport and equilibrium equations in tokamaks. *Nuclear Fusion*, 53(3), 2013.

[8] J. P. Freidberg. *Ideal Magnetohydrodynamics*. Plenum US, 1987.

[9] G.N. Gatica and G.C. Hsiao. The uncoupling of boundary integral and finite element methods for nonlinear boundary value problems. *J. Math. Anal. Appl.*, 189(2):442–461, 1995.

[10] H. Grad and J. Hogan. Classical diffusion in a tokomak. *Phys. Rev. Lett.*, 24:1337–1340, Jun 1970.

[11] H. Grad, P. N. Hu, and D. C. Stevens. Adiabatic evolution of plasma equilibrium. *Proc Natl Acad Sci U S A.*, 72:3789–3793, 1975.

[12] H. Grad and H. Rubin. Hydromagnetic equilibria and force-free fields. *Proceedings of the 2nd UN Conf. on the Peaceful Uses of Atomic Energy*, 31:190, 1958.

[13] Virginie Grandgirard. *Modélisation de l'équilibre d'un plasma de tokamak*. PhD thesis, l'Université de Franche-Comté, 1999.

[14] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

[15] M. Honda. Simulation technique of free-boundary equilibrium evolution in plasma ramp-up phase. *Computer Physics Communications*, 181(9):1490–1500, SEP 2010.

[16] A. A. Ivanov, R. R. Khayrutdinov, S. Yu. Medvedev, and Yu. Yu Poshekhonov. The SPIDER code - solution of direct and inverse problems for free boundary tokamak plasma equilibrium. Technical Report 39, Keldysh Institute, 2009.
[17] Stephen Jardin. *Computational methods in plasma physics*. Boca Raton, FL : CRC Press/Taylor & Francis, 2010.
[18] K. Lackner. Computation of ideal MHD equilibria. *Computer Physics Communications*, 12(1):33 – 44, 1976.
[19] R. Lüst and A. Schlüter. Axialsymmetrische magnetohydrodynamische Gleichgewichtskonfigurationen. *Z. Naturforsch. A*, 12:850–854, 1957.
[20] J.L. Luxon and B.B. Brown. Magnetic analysis of non-circular cross-section tokamaks. *Nuclear Fusion*, 22(6):813, 1982.
[21] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
[22] V. D. Shafranov. On magnetohydrodynamical equilibrium configurations. *Soviet Journal of Experimental and Theoretical Physics*, 6:545, 1958.
[23] J. Wesson. *Tokamaks*. The International Series of Monographs in Physics. Oxford University Press, 2004.

E-mail address: holger.heumann@inria.fr

CASTOR TEAM, INRIA, 2004, Route des Lucioles, BP 93, 06560 Sophia Antipolis, France

E-mail address: jblum@math.unice.fr

Laboratoire J.-A. Dieudonné, Université de Nice, Parc Valrose, 06108 Nice cedex 2, France & CASTOR TEAM, INRIA, 2004, Route des Lucioles, BP 93, 06902 Sophia Antipolis, France