

# Gradient-based optimization of 3D MHD equilibria

Elizabeth J. Paul <sup>1</sup>,<sup>†</sup>, Matt Landreman <sup>2</sup> and Thomas Antonsen, Jr. <sup>2</sup>

<sup>1</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

<sup>2</sup>Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, MD 20742, USA

(Received 17 December 2020; revised 26 February 2021; accepted 2 March 2021)

Using recently developed adjoint methods for computing the shape derivatives of functions that depend on magnetohydrodynamic (MHD) equilibria (Antonsen *et al.*, *J. Plasma Phys.*, vol. 85, issue 2, 2019; Paul *et al.*, *J. Plasma Phys.*, vol. 86, issue 1, 2020), we present the first example of analytic gradient-based optimization of fixed-boundary stellarator equilibria. We take advantage of gradient information to optimize figures of merit of relevance for stellarator design, including the rotational transform, magnetic well and quasi-symmetry near the axis. With the application of the adjoint method, we reduce the number of equilibrium evaluations by the dimension of the optimization space ( $\sim 50$ – $500$ ) in comparison with a finite-difference gradient-based method. We discuss regularization objectives of relevance for fixed-boundary optimization, including a novel method that prevents self-intersection of the plasma boundary. We present several optimized equilibria, including a vacuum field with very low magnetic shear throughout the volume.

**Key words:** fusion plasma, plasma confinement, plasma simulation

## 1. Introduction

The improved confinement of modern stellarators is largely attributed to the numerical optimization of the magnetic field. With the technique pioneered by Nührenberg & Zille (1988), the shape of the plasma boundary,  $S_p$ , is optimized for certain properties of the magnetohydrodynamic (MHD) equilibrium,

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = \mu_0 \nabla p, \quad \text{in } V_p, \quad (1.1a)$$

$$\mathbf{B} \cdot \hat{\mathbf{n}} = 0, \quad \text{on } S_p, \quad (1.1b)$$

where  $S_p = \partial V_p$ ,  $p(\psi)$  is the prescribed pressure profile and  $\hat{\mathbf{n}}$  is the unit normal. For the past three decades, stellarator optimization has largely proceeded with this fixed-boundary approach with codes such as STELLOPT (Lazerson *et al.* 2020) and ROSE (Drevlak *et al.* 2018), resulting in the W7-X (Beidler *et al.* 1990), HSX (Anderson *et al.* 1995) and NCSX (Zarnstorff *et al.* 2001) configurations. We use the term ‘fixed-boundary’ to describe this

<sup>†</sup> Email address for correspondence: [epaul@princeton.edu](mailto:epaul@princeton.edu)

approach as the equilibrium is computed with a specified boundary, although the boundary is varied throughout the optimization.

Although there has been significant success with this approach, there are several ways to improve the performance of fixed-boundary optimization. Specifically, the incorporation of derivative information could be transformative. Fixed-boundary optimization has previously relied on derivative-free methods, such as genetic algorithms (Miner *et al.* 2001), differential evolution (Mynick, Pomphrey & Ethier 2002) and Brent's algorithm (Drevlak *et al.* 2018), or derivative-based methods with finite-difference gradients. While global derivative-free algorithms may prevent the optimization from terminating in local minima, they are only effective for smaller problems (Nocedal & Wright 2006). Gradient-based optimization with finite-difference gradients requires excessive function evaluations in high-dimensional spaces and suffers from error that enters due to the choice of step size. If the step size is too small, the error is dominated by round-off error, and if it is too large it is dominated by nonlinearity (Sauer 2012). Owing to the requirement of excessive function evaluations and the unreliability of the gradient information, gradient-based optimization with finite-difference derivatives is not always effective. In this work, we present the first example of analytic gradient-based, fixed-boundary optimization of stellarator equilibria.

With a local gradient descent approach, each iteration reduces to a one-dimensional line search (Nocedal & Wright 2006); thus, the further incorporation of derivative information eliminates restrictions on the size of the optimization space. There is some evidence from the machine learning community that overparameterization of the space can accelerate optimization (Oymak & Soltanolkotabi 2018). Therefore, it is possible that increasing the Fourier resolution of the plasma boundary may similarly eliminate local minima.

There are several ways that this derivative information can be obtained. For sufficiently simple figures of merit, the objective can be analytically differentiated and implemented by hand. Alternatively, the derivatives can be obtained programmatically using automatic differentiation tools. When a given objective function depends on the solution of a set of equations, such as the MHD equilibrium equations (1.1), the derivatives can be obtained more efficiently using an adjoint method. With this technique, the solution of only one additional equation, known as the adjoint equation, is required. Once the adjoint solution is obtained, the derivative of a given objective can be obtained with respect to any optimization parameter, eliminating the need to solve a perturbed set of equations. In this way, the cost associated with obtaining a high-dimensional gradient is reduced significantly. In this work, we employ adjoint-based gradients of functions which depend on the MHD equilibrium equations. This adjoint method results from a generalized self-adjointness property of the MHD force operator (Antonsen, Paul & Landreman 2019). This technique has been demonstrated for computing the shape derivatives of several figures of merit relevant for stellarator design, including the magnetic well, rotational transform, and magnetic ripple (Paul *et al.* 2020). Each of these objective functions have been included in modern stellarator designs (Beidler *et al.* 1990; Anderson *et al.* 1995; Zarnstorff *et al.* 2001; Henneberg *et al.* 2019). We also employ analytic gradients for objectives that do not depend on the MHD equilibrium equations, such as the volume and properties of the surface curvatures.

There have been several recent applications of derivative information to other related problems in stellarator design. The FOCUS (Zhu *et al.* 2018) and FOCUSADD (McGreivy, Hudson & Zhu 2021) codes optimize coil shape to be consistent with a given plasma boundary with gradients obtained from analytic and automatic differentiation methods, respectively. Our work is distinct from the FOCUS approach, as we use gradients to

optimize properties of the equilibrium rather than using gradients to optimize coils in order to match the boundary of a given equilibrium. Adjoint methods have recently been applied to directly optimize coils for quasi-symmetry near the magnetic axis in a vacuum field (Giuliani *et al.* 2020). In contrast, our approach can be applied to optimize equilibria with arbitrary pressure. Furthermore, we have developed adjoint methods for direct optimization of coil shapes for properties of an MHD equilibrium (Antonsen *et al.* 2019; Paul *et al.* 2020), although its application to optimization is not presented in this work. Adjoint methods have also been used to optimize the local magnetic field for neoclassical properties (Paul *et al.* 2019) and the coil winding surface for properties of the current potential (Paul *et al.* 2018).

We discuss the new gradient-based fixed-boundary optimization tool in § 2. In § 3, we present regularization terms for fixed-boundary optimization, including a constraint on the curvature of the boundary that prevents self-intersection. In § 4, we present several optimization demonstrations, including obtaining a low magnetic shear configuration (§ 4.1), a configuration with a magnetic well (§ 4.2), and a configuration with quasi-symmetry near the magnetic axis (§ 4.3). We conclude in § 5.

## 2. Overview of ALPOpt optimization tool

As with the STELLOPT and ROSE codes, the ALPOpt<sup>1</sup> tool optimizes the boundary of VMEC (Hirshman & Whitson 1983) equilibria. The VMEC code obtains solutions of (1.1) under the assumption of nested toroidal magnetic surfaces. The plasma boundary is described by a set of Fourier coefficients of the cylindrical coordinates,  $\{R_{m,n}^c, Z_{m,n}^s\}$ ,

$$R(\theta, \phi) = \sum_{m,n} R_{m,n}^c \cos(m\theta - nN_p\phi), \quad (2.1a)$$

$$Z(\theta, \phi) = \sum_{m,n} Z_{m,n}^s \sin(m\theta - nN_p\phi), \quad (2.1b)$$

where  $\theta$  is a poloidal angle,  $\phi$  is the cylindrical toroidal angle and  $N_p$  is the number of field periods. Therefore, the optimization space is taken to be the set of coefficients  $\{R_{m,n}^c, Z_{m,n}^s\}$ . The pressure profile  $p(\psi)$  and another function of flux, either the rotational transform  $\iota(\psi)$  or enclosed toroidal current  $I_T(\psi)$ , are prescribed and fixed. (For all of the examples in this work,  $I_T(\psi)$  is fixed.) The optimization code interfaces with VMEC through python, and optimization is performed with the scipy<sup>2</sup> and NLOPT (Johnson 2014) packages.

The optimization tool takes advantage of the adjoint method for obtaining the shape gradient of MHD equilibria (Antonsen *et al.* 2019; Paul *et al.* 2020). Here the perturbation to the magnetic field due to a perturbation of the boundary is expressed as

$$\delta\mathbf{B}_1 = \nabla \times (\boldsymbol{\xi}_1 \times \mathbf{B}) - \delta\iota_1(\psi)\nabla\psi \times \nabla\phi, \quad (2.2)$$

where the displacement vector satisfies  $\boldsymbol{\xi}_1 \cdot \hat{\mathbf{n}} = \delta\mathbf{x} \cdot \hat{\mathbf{n}}$  on the boundary for a given normal perturbation to the surface and  $\delta\iota_1(\psi)$  is the perturbation to the rotational transform profile that may arise due to the constraint of fixed  $I_T(\psi)$ . This perturbed magnetic field can be related to an *adjoint* perturbed magnetic field,  $\delta\mathbf{B}_2$ , through a generalized self-adjointness

<sup>1</sup>See <https://github.com/ejpaul/ALPOpt>

<sup>2</sup>See <https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html>

relation,

$$\int_{V_p} d^3x (\boldsymbol{\xi}_1 \cdot \mathbf{F}_2 - \boldsymbol{\xi}_2 \cdot \mathbf{F}_1) - 2\pi \int_{V_p} d\psi (\delta I_{T,2}(\psi)\delta l_1(\psi) - \delta I_{T,1}(\psi)\delta l_2(\psi)) - \frac{1}{\mu_0} \int_{S_p} d^2x \hat{\mathbf{n}} \cdot (\boldsymbol{\xi}_2 \delta \mathbf{B}_1 - \boldsymbol{\xi}_1 \delta \mathbf{B}_2) \cdot \mathbf{B} = 0, \quad (2.3)$$

where the linearized force operator is

$$\mathbf{F}_{1,2} = \frac{(\nabla \times \delta \mathbf{B}_{1,2}) \times \mathbf{B}}{\mu_0} + \frac{(\nabla \times \mathbf{B}) \times \delta \mathbf{B}_{1,2}}{\mu_0} + \nabla (\boldsymbol{\xi}_{1,2} \cdot \nabla p). \quad (2.4)$$

The adjoint approach is as follows: rather than directly computing the perturbed magnetic field  $\delta \mathbf{B}_1$  by perturbing the boundary of a fixed-boundary equilibrium or solving a set of linearized equilibrium equations, the adjoint magnetic field is computed. The adjoint magnetic field has no perturbation to the boundary ( $\boldsymbol{\xi}_2 \cdot \hat{\mathbf{n}} = 0$ ) but may have a perturbation to the toroidal current profile ( $\delta I_{T,2}(\psi)$ ) or a bulk force perturbation ( $\mathbf{F}_2$ ). In this work, we consider figures of merit whose derivatives can be computed with a perturbation to the toroidal current profile or a bulk force perturbation which takes the form of the gradient of a scalar function of flux or the divergence of an anisotropic pressure tensor.

Rather than consider a set of linearized equations, we add a small perturbation to a nonlinear VMEC equilibrium in the form of a perturbation to the pressure profile, toroidal current profile, or an anisotropic pressure tensor. In the case of the addition of a pressure tensor, the ANIMEC (Cooper *et al.* 1992) code is used to evaluate the adjoint equilibrium. The resulting shape gradient,  $\mathcal{G}$ , of a given objective function  $f$ , is defined through,

$$\delta f(S_p; \delta \mathbf{x}) = \int_{S_p} d^2x \delta \mathbf{x} \cdot \hat{\mathbf{n}} \mathcal{G}. \quad (2.5)$$

Here  $\delta f(S_p; \delta \mathbf{x})$  is the shape derivative of  $f$  with respect to a normal perturbation of the surface,  $\delta \mathbf{x} \cdot \hat{\mathbf{n}}$ . Given the shape gradient, which quantifies the local sensitivity to normal perturbations of the surface, the derivatives with respect to the parameters  $\Omega = \{R_{m,n}^c, Z_{m,n}^s\}$  are computed,

$$\frac{\partial f}{\partial \Omega} = \int_{S_p} d^2x \frac{\partial \mathbf{x}}{\partial \Omega} \cdot \hat{\mathbf{n}} \mathcal{G}. \quad (2.6)$$

### 2.1. Managing code failures

The equilibrium code may return with an error for a given plasma boundary. There are many reasons for these failures, such as the flux coordinate Jacobian becoming ill-conditioned or the number of iterations exceeding the maximum prescribed value. When such a failure is experienced during the optimization, the objective function is set to an arbitrarily large value (e.g.  $10^{12}$ ) to enforce an effective constraint. The approach of assigning a very large fictitious objective value at unevaluable points is common in the optimization literature (Rasheed, Hirsh & Gelsey 1997; Emmerich *et al.* 2002) and is employed in the STELLOPT code (Lazerson *et al.* 2020).

As encountering such unevaluable points makes the parameter space non-smooth, it is prudent to try to avoid code failures by placing additional constraints on the parameter space. Such a technique has been employed in the optimization of aircraft by checking

that simulation outputs match physical model assumptions, such as the drag coefficient being non-negative (Gelsey 1995; Gelsey, Schwabacher & Smith 1998). In § 3, we present constraints which prevent the surface from self-intersecting or from obtaining large curvature, leading to an ill-conditioned Jacobian and code failure. Even in the presence of such constraints, unevaluable points may still be present. Here the objective function becomes discontinuous, which is problematic for optimizers which assume function continuity. In this work we employ the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method with an Armijo–Wolfe line search, which assumes  $C^2$  continuity of the objective function (Nocedal & Wright 2006). Although convergence of the BFGS algorithm is not guaranteed for non-smooth problems, it has been observed that a reasonable approximation to the optimum is most often achieved (Lemaréchal 1982). With the standard Armijo–Wolfe line search method, the gradient need not be evaluated unless the objective function satisfies the sufficient decrease criterion. Thus the gradient does not need to be evaluated at unevaluable points. In this way, as long as the line search never returns a point where the objective is not differentiable, the BFGS method is well defined. Although there are specialized quasi-Newton methods for non-smooth objective functions (Lewis & Overton 2013), we have obtained acceptable results with a standard BFGS method.

### 3. Regularization terms

#### 3.1. Preventing surface self-intersection

We now describe a constraint which prevents self-intersection of the plasma boundary. Given a surface described by the cylindrical coordinates  $R(\theta, \phi)$  and  $Z(\theta, \phi)$ , self-intersection of the boundary may occur if either:

- (a)  $R(\theta, \phi) < 0$  at any point; or
- (b) the planar curve  $\mathbf{x}_{\phi_0}(\theta) = \{R(\theta, \phi_0), Z(\theta, \phi_0)\}$  is self-intersecting for any  $\phi_0$ .

Condition (a) can be avoided using a penalty objective of the form,

$$f_R(S_P) = \frac{1}{A_P} \int_{S_P} d^2x \exp\left(-\frac{(R - R_{\min})}{w_R}\right), \tag{3.1}$$

where  $w_R$  is a weight which sets the gradient length scale of the objective,  $R_{\min}$  is the minimum allowable major radius and  $A_P$  is the area of  $S_P$ .

Condition (b) can be avoided by introducing a constraint on the *global radius of curvature* (Gonzalez & Maddocks 1999; Walker 2016) of each of the curves,  $\mathbf{x}_{\phi_0}(\theta)$ , which will now be defined. The radius of the unique circumcircle containing any three points  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ , and  $\mathbf{x}_3$  (figure 1), can be computed from

$$\rho(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{|\mathbf{x}_1 - \mathbf{x}_2||\mathbf{x}_1 - \mathbf{x}_3||\mathbf{x}_2 - \mathbf{x}_3|}{4\mathcal{A}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)}, \tag{3.2}$$

where  $\mathcal{A}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$  is the area of a triangle with vertices  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ , and  $\mathbf{x}_3$ . Under the assumption that these points lie on a non-self-intersecting smooth curve,  $\mathcal{C}$  such that  $\mathbf{x}_1 = \mathbf{x}(l_1)$ ,  $\mathbf{x}_2 = \mathbf{x}(l_2)$ , and  $\mathbf{x}_3 = \mathbf{x}(l_3)$ , then the radius of the circumcircle satisfies  $\rho(\mathbf{x}(l_1), \mathbf{x}(l_2), \mathbf{x}(l_2)) \leq \rho(\mathbf{x}(l_1), \mathbf{x}(l_2), \mathbf{x}(l_3))$ . This limiting case can be computed from

$$\rho(\mathbf{x}(l_1), \mathbf{x}(l_2), \mathbf{x}(l_2)) = \lim_{l_2 \rightarrow l_3} \rho(\mathbf{x}(l_1), \mathbf{x}(l_2), \mathbf{x}(l_3)) = \frac{|\mathbf{x}(l_1) - \mathbf{x}(l_2)|}{2\sqrt{1 - \left(\hat{\mathbf{t}}(\mathbf{x}(l_2)) \cdot \frac{\mathbf{x}(l_1) - \mathbf{x}(l_2)}{|\mathbf{x}(l_1) - \mathbf{x}(l_2)|}\right)^2}}, \tag{3.3}$$

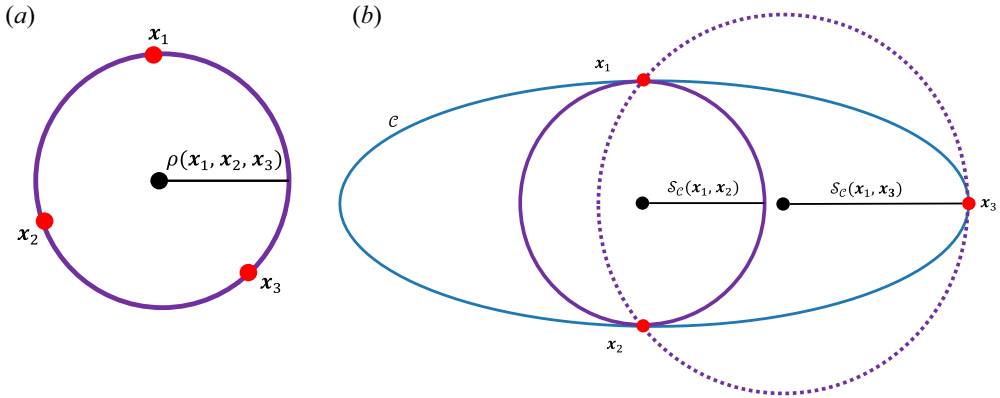


FIGURE 1. (a) The circumcircle containing  $x_1, x_2$  and  $x_3$  with radius  $\rho(x_1, x_2, x_3)$ . (b) The points  $x_1, x_2$ , and  $x_3$  lie on a closed elliptical curve  $\mathcal{C}$  (blue). The point tangent radius for points  $x_1$  and  $x_2$  and points  $x_1$  and  $x_3$  are displayed. The global radius of curvature at  $x_1$  is  $\mathcal{S}_C(x_1, x_2)$ .

where  $\hat{t}$  is the unit tangent. This is the radius of the circle that passes through  $x(l_1)$  and is tangent to  $\mathcal{C}$  at  $x(l_2)$ , which we define as the self-contact function between these points,

$$\mathcal{S}_C(x_1, x_2) = \frac{|x_1 - x_2|}{2\sqrt{1 - \left(\hat{t}(x_2) \cdot \frac{(x_1 - x_2)}{|x_1 - x_2|}\right)^2}}, \tag{3.4}$$

the radius of the so-called point-tangent circle (figure 1).

In the limit that  $l_2, l_3 \rightarrow l_1$ , we obtain

$$\lim_{l_2, l_3 \rightarrow l_1} \rho(x(l_1), x(l_2), x(l_3)) = \rho_C(x(l_1)), \tag{3.5}$$

where  $\rho_C(x(l_1))$  is the local radius of curvature of  $\mathcal{C}$  at  $x(l_1)$ . We define the global radius of curvature as

$$\rho_G(x_1) = \min_{x_2, x_3 \in \mathcal{C}} \rho(x_1, x_2, x_3) = \min_{x_2 \in \mathcal{C}} \mathcal{S}_C. \tag{3.6}$$

This can be thought of as a generalization of the local radius of curvature, as we have the inequality  $0 \leq \rho_G(x) \leq \rho_C(x)$ . At a given point, the minimizer in (3.6) will either be a point where  $(x_1 - x_2) \cdot \hat{t}(x_2) = 0$  or  $\rho_C(x_2) = \mathcal{S}_C(x_1, x_2)$  (Smutny 2004). In other words, at a given point  $x_1$ , the point on the curve that has the smallest self-contact function containing  $x_1$  will be where the local radius of curvature is equal to the point-tangent radius or the displacement of the points  $(x_1 - x_2)$  is orthogonal to the curve at  $x_2$ .

If the inequality

$$\min_{x_1 \in \mathcal{C}} \rho_G(x_1) = \min_{x_1, x_2 \in \mathcal{C}} \mathcal{S}_C(x_1, x_2) \geq \mathcal{R} \tag{3.7}$$

is satisfied, this implies that the curve can be ‘thickened’ to a tube without self-contact surrounding  $\mathcal{C}$  such that at any given point, the cross-section of the tube is a circle in the plane perpendicular to the local tangent vector of radius  $\mathcal{R}$  (Gonzalez & Maddocks 1999). For this reason, the concept of the global radius of curvature has been employed for the shape optimization of finite-thickness knots (Gonzalez & Maddocks 1999; Carlen *et al.* 2005; Walker 2016).

We enforce the inequality constraint (3.7) to prevent the self-intersection of every planar curve  $\mathbf{x}_{\phi_0}(\theta)$  with a penalty function of the form

$$f_S = \frac{1}{A_P} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta \left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi} \right| \int_0^{2\pi} d\theta' \exp \left( - \frac{(\mathcal{S}_C(\mathbf{x}_\phi(\theta), \mathbf{x}_\phi(\theta')) - \mathcal{R})}{w_S} \right), \tag{3.8}$$

where  $\mathcal{R}$  is the minimum allowable global radius of curvature and  $w_S$  is a weight function. In addition to preventing self-intersection of the boundary, this objective improves the regularity of the surface by reducing the curvature of the poloidal cross-sections. A demonstration of this penalty function is presented in § 3.3.

### 3.2. Surface curvature constraint

While constraining the self-contact function reduces the curvature of the cross-sections of the plasma boundary, we include additional terms which improve the regularity of the plasma boundary. First, we place an effective constraint on the magnitude of the principal curvatures,

$$|\kappa_1| \leq \kappa_{\max,1} \quad |\kappa_2| \leq \kappa_{\max,2}, \tag{3.9a,b}$$

by including a penalty function of the form,

$$f_\kappa = \frac{\int_{S_P} d^2x \left( \exp((|\kappa_1|^2 - \kappa_{\max,1}^2)/w_k^2) + \exp((|\kappa_2|^2 - \kappa_{\max,2}^2)/w_k^2) \right)}{A_P}. \tag{3.10}$$

A similar constraint on the principal curvatures is employed in the ROSE code (Drevlak *et al.* 2018). We assume the convention that  $\kappa_1 \leq \kappa_2$  and  $\kappa_{1,2} < 0$  indicates concavity.

We additionally add a penalty on the smallest principal curvature,

$$f_{\bar{\kappa}} = \frac{\int_{S_P} d^2x \kappa_1^2}{A_P}. \tag{3.11}$$

As large concavity is associated with coil complexity (Paul *et al.* 2018), this regularization becomes critical in the two-staged optimization approach. Also, regularity of the boundary improves the convergence of the VMEC code. If the boundary is not close to being a star domain, indicating that there exists a coordinate axis such that the line segment connecting the axis and any point on the boundary is contained within the boundary, then the solver may fail to initialize a guess for the magnetic axis. For highly shaped boundaries, the coordinate surfaces may begin to overlap. Thus, the inclusion of these regularization terms prevents code failures during optimization and improves the convergence toward the optimum. This advantage of curvature penalization is highlighted in the magnetic well optimization in § 4.2.

### 3.3. Constrained volume optimization

To demonstrate the procedure described in § 3.1 for avoiding self-intersection, we define an objective function,

$$f(S_P) = V_P + \lambda_S f_S + \lambda_R f_R + \lambda_\kappa f_\kappa, \tag{3.12}$$

where  $V_P$  is the volume enclosed by the plasma boundary. For this example, we take  $\lambda_S = \lambda_R = 10^3$ ,  $R_{\min} = \mathcal{R} = 0.2$  m,  $w_R = w_S = 10^{-2}$ ,  $\lambda_\kappa = 1$ ,  $\kappa_{\max,1} = \kappa_{\max,2} = 10$  m<sup>-1</sup>

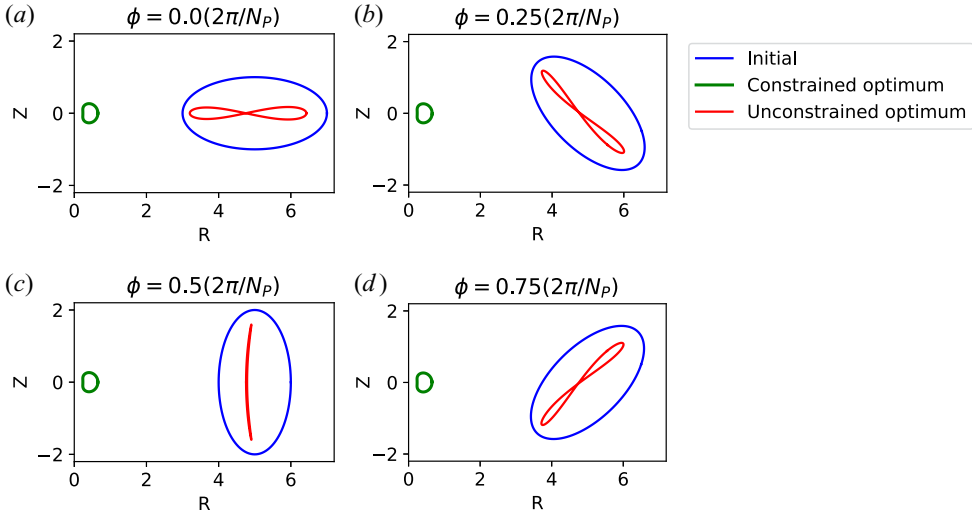


FIGURE 2. Cross-sections of the initial (blue) and optimized surfaces obtained by minimizing the objective function given by (3.12). The surface obtained without the constraints ( $\lambda_R = \lambda_S = 0$ ) becomes self-intersecting (red), and the surface obtained with the constraints remains regular (green).

and  $w_k = 1$ . Here the volume is computed using a surface integral upon application of the divergence theorem,

$$V_p = \int d^3x = \int_{S_p} d^2x Z \hat{n} \cdot \hat{Z}, \tag{3.13}$$

to avoid discretizing the volume.

We begin with a stellarator whose boundary is defined by a rotating ellipse,

$$R(\theta, \phi) = R_0 + 0.5a (\cos(\theta - N_p\phi) + \cos(\theta)) - 0.5b (\cos(\theta - N_p\phi) - \cos(\theta)) \tag{3.14a}$$

$$Z(\theta, \phi) = 0.5b (\sin(\theta) + \sin(\theta - N_p\phi)) + 0.5a (\sin(\theta) - \sin(\theta - N_p\phi)), \tag{3.14b}$$

where  $a$  is the semi-major axis and  $b$  is the semi-minor axis. For this example, we take  $R_0 = 5$ ,  $a = 2$ ,  $b = 1$  and  $N_p = 3$ . We optimize  $f(S_p)$  with respect to modes  $m \leq 3$  and  $|n| \leq 3$  using the BFGS algorithm provided by SciPy. (This optimization algorithm is used for all demonstrations in this work.)

We compare the surfaces obtained with and without ( $\lambda_S = \lambda_R = 0$ ) the constraint terms, shown in figure 2. Without the constraint, the boundary begins to self-intersect. Although the volume is not well-defined for this surface, the discretized surface integral given by (3.13) is reduced from its initial value of 197.39 to 4.18 m<sup>3</sup> by minimizing the projection of the normal vector in the  $\hat{Z}$  direction. With the constraint, the surface is reduced to an axisymmetric torus with a nearly circular cross-section centred at a major radius  $R_0 = 0.43$  m with an averaged minor radius of 0.24 m and a volume of 0.50 m<sup>3</sup>.

In figure 3 we display the value of the global radius of curvature (3.6) for the initial and final surfaces. On the initial surface,  $\rho_G$  is minimized at the endpoints of the semi-major axis where the global radius of curvature matches the local radius of curvature.



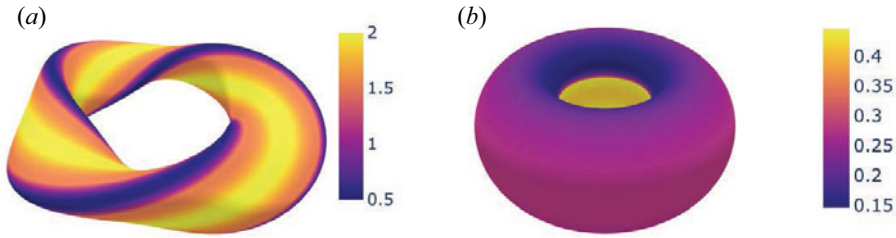


FIGURE 3. The global radius of curvature (3.6) of the (a) initial surface (3.14) and (b) surface optimized for minimal volume with regularity constraints (3.12).

The maximum value of  $\rho_G$  is obtained in the region near the endpoint of the ellipse’s semi-minor axis, where the point-tangent radius is minimized for  $x_2$  on the opposite side of the semi-minor axis. The optimized boundary features decreased values of  $\rho_G$  on the outboard side, where  $\rho_G$  matches the local radius of curvature. We note that the minimum value of  $\mathcal{S}_C$  on the optimized slightly violates the constraint at 0.15 m owing to the penalty formulation.

#### 4. Optimization demonstrations

##### 4.1. Target rotational transform

We now discuss a demonstration of optimization to obtain a target rotational transform profile. We consider a target function which quantifies the difference between the rotational transform and a desired profile,  $\iota_{\text{target}}(\psi)$ ,

$$f_i = \frac{1}{2} \int_{V_p} d\psi (\iota(\psi) - \iota_{\text{target}}(\psi))^2, \tag{4.1}$$

in order to obtain a low shear stellarator. The gradient of this objective function is obtained by computing an equilibrium with a perturbed toroidal current profile,

$$\delta I_T(\psi) = w(\psi)(\iota(\psi) - \iota_{\text{target}}(\psi)), \tag{4.2}$$

as described in Antonsen *et al.* (2019). We present a benchmark problem in [appendix A](#) to demonstrate that the gradient-based optimizer can converge to a known minimum of an objective involving  $f_i$  in a 2D space.

Although the VMEC code assumes nested surfaces such that the magnetic field is integrable, the solution to Laplace’s equation with the prescribed boundary  $S_p$  will not generally have continuously nested surfaces. We can, however, aim to obtain a vacuum field with a value of the rotational transform that improves the integrability of the ‘real’ vacuum field. (Note that although we can run VMEC with prescribed  $\iota(\psi)$  rather than  $I_T(\psi)$ , it will not generally be a vacuum field.) We choose  $\iota_{\text{target}}$  to be the ‘most irrational’ noble between  $p/q = 0/1$  and  $p'/q' = 1/2$  (Greene, MacKay & Stark 1986),

$$\iota_{\text{target}} = \frac{p + \gamma p'}{q + \gamma q'} \approx 0.381966, \tag{4.3}$$

where  $\gamma = (\sqrt{5} + 1)/2$  is the golden mean. This is a noble irrational, indicating that its path in the Farey tree is eventually alternating. At a given level in the Farey tree, the

path that changes direction is said to be ‘most irrational’ if a cantorus with this rotation number has the smallest flux of field line trajectories (Meiss 1992). Thus, the value of  $\iota_{\text{target}}$  is chosen to have the smallest flux for rotation numbers between  $p/q = 0/1$  and  $p'/q' = 1/2$ . According to the Kolmogorov–Arnold–Moser (KAM) theorem, invariant circles with sufficiently irrational frequencies persist under small perturbations. Thus, choosing such a value for the rotational transform is likely to result in a large volume of magnetic surfaces. We remark that moderate shear may be desirable for certain stellarator design studies. Our adjoint formalism is quite flexible, and the choice of  $\iota_{\text{target}}$  could be modified accordingly.

We define our objective function to be

$$f(S_P) = f_i + \lambda_R f_R + \lambda_S f_S, \quad (4.4)$$

where  $\lambda_S = \lambda_R = 10^3$ ,  $R_{\min} = \mathcal{R} = 0.2$  m and  $w_R = w_S = 10^{-2}$ . We begin with a surface given by a rotating ellipse with a non-planar axis,

$$\begin{aligned} R(\theta, \phi) = & R_0 + R_{0,1} \cos(\phi) + 0.5a (\cos(\theta - N_P\phi) + \cos(\theta)) \\ & - 0.5b (\cos(\theta - N_P\phi) - \cos(\theta)) \end{aligned} \quad (4.5a)$$

$$\begin{aligned} Z(\theta, \phi) = & -R_{0,1} \sin(\phi) + 0.5b (\sin(\theta) + \sin(\theta - N_P\phi)) \\ & + 0.5a (\sin(\theta) - \sin(\theta - N_P\phi)), \end{aligned} \quad (4.5b)$$

with  $a = 2$ ,  $b = 1$ ,  $N_P = 3$ ,  $R_0 = 5$  and  $R_{0,1} = -0.5$ . We consider a vacuum field with  $p(\psi) = I_T(\psi) = 0$ .

To investigate the benefit of increasing the dimensionality of the optimization space, we optimize with respect to the boundary harmonics  $m \leq 3$ ,  $|n| \leq 3$ . We then use this result to optimize with respect to  $m \leq 4$ ,  $|n| \leq 4$  and then with respect to  $m \leq 5$ ,  $|n| \leq 5$ . In comparison with the result of the initial optimization in the low-dimensional space, we are able to reduce the objective function by 70% (figure 4). We obtain a rotational transform profile which very closely matches the target value with an objective value of  $f_i = 4.93 \times 10^{-10}$ . If the optimization space is further increased, the optimum is reduced by less than 1%. For the further analysis of the optimization in this section, we present results from the optimization with respect to the modes  $m \leq 5$  and  $|n| \leq 5$ .

We note that we were not able to reduce the gradient norm to the requested tolerance of  $10^{-8}$ . This result can be attributed to approximations made in computing the gradient with the adjoint method. As discussed in Antonsen *et al.* (2019) and Paul *et al.* (2020), the adjoint equations are derived under the assumption that the VMEC code satisfies MHD force balance (1.1). However, there is always some residual error in force balance due to discretization error and the assumption of continuously nested flux surfaces. For these calculations, we converged to a force balance tolerance of  $10^{-11}$  with 99 flux surfaces and mode numbers  $|m| \leq 11$ ,  $|n| \leq 11$ . These resolution parameters were chosen to strike a balance between an accurate adjoint solution and efficiency of the optimization. Preliminary calculations indicate that increasing the resolution parameters over what was used in this work does not cause a significant change in the optimum. Furthermore, the linear adjoint solution is approximated by adding a small perturbation to the nonlinear MHD force balance. This technique is effectively a forward-difference approximation of the adjoint equation, which introduces an error that scales with the magnitude of the perturbation. When there are small errors in the gradient, the computed gradient may no longer be a descent direction, making convergence difficult near the optimum (Dekeyser 2014).

Nonetheless, we effectively eliminate the magnetic shear throughout the volume. In figure 5 we display the initial and optimized boundaries (solid) along with an interior

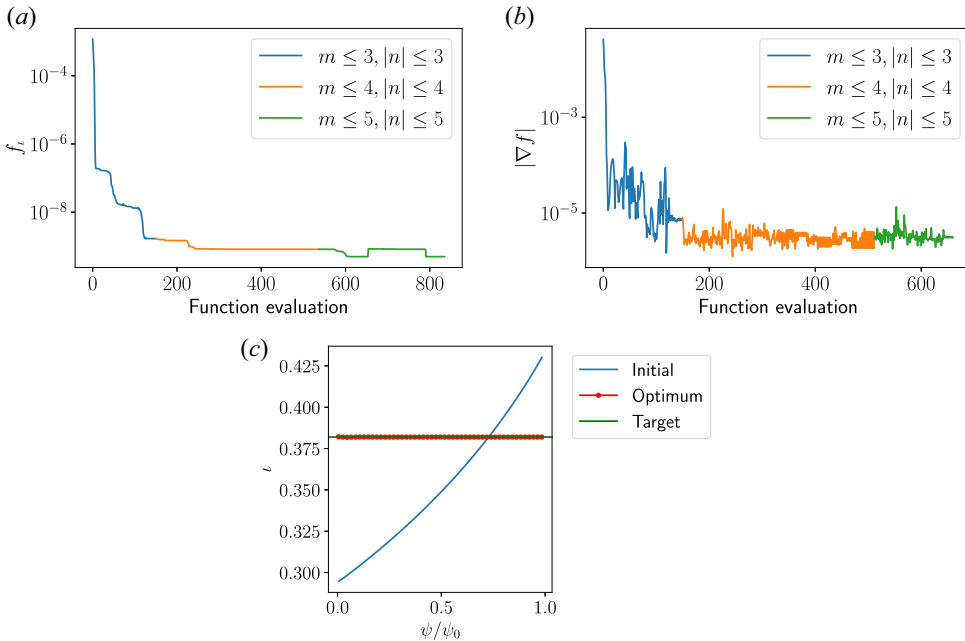


FIGURE 4. (a) The  $f_i$  objective value (4.1), (b) the L2 norm of the  $f(S_P)$  objective (4.4) gradient and (c) the initial and final profiles of the rotational transform. In (a,b) the optimization was performed in a staged approach: first with respect to modes  $m \leq 3$  and  $|n| \leq 3$ , then with respect to  $m \leq 4$  and  $|n| \leq 4$  and, finally, with respect to  $m \leq 5$  and  $|n| \leq 5$ .

surface at  $\psi/\psi_0 = 0.06$  (dashed) and the magnetic axis (star). We can consider some of the features of the optimized surface in consideration of the expression for the rotational transform near the magnetic axis (Mercier 1964; Helander 2014), which indicates that rotating ellipticity and torsion of the axis contribute to the on-axis transform in a vacuum field. We see that the interior surfaces become slightly more elliptical in order to increase  $\iota$  near the axis, whereas the ellipticity of the boundary is slightly decreased, becoming more square. However, the torsion of the magnetic axis is maintained.

To analyse the effect of low shear on the integrability of the field, we compute the vacuum field using the SPEC code (Hudson *et al.* 2012). The SPEC calculations are performed with a single volume with Beltrami parameter  $\mu = 0$  such that the magnetic field satisfies  $\nabla \times \mathbf{B} = 0$  with a Neumann boundary condition (1.1b). In figure 6 we show a Poincaré section for the initial field, which has a small island chain at the  $\iota = 3/8$  resonance. With the optimized boundary, we eliminate this island chain and obtain a large volume of nested surfaces.

#### 4.2. Magnetic well

We next consider an objective function which aims to obtain a magnetic well,

$$f_w = \frac{\int_{V_P} d\psi (w_1(\psi)V'(\psi) - w_2(\psi)V'(\psi))}{\int_{V_P} d\psi (w_1(\psi)V'(\psi) + w_2(\psi)V'(\psi))}, \tag{4.6}$$

with  $w_1(\psi) = \exp(-(\psi/\psi_0 - 1.0)^2/0.1^2)$  and  $w_2(\psi) = \exp(-(\psi/\psi_0)^2/0.1^2)$ . When  $V''(\psi) < 0$ , a magnetic well is said to be present, which provides a stabilizing term in the Mercier criterion for interchange modes (Mercier & Luc 1974). We can consider  $f_w$

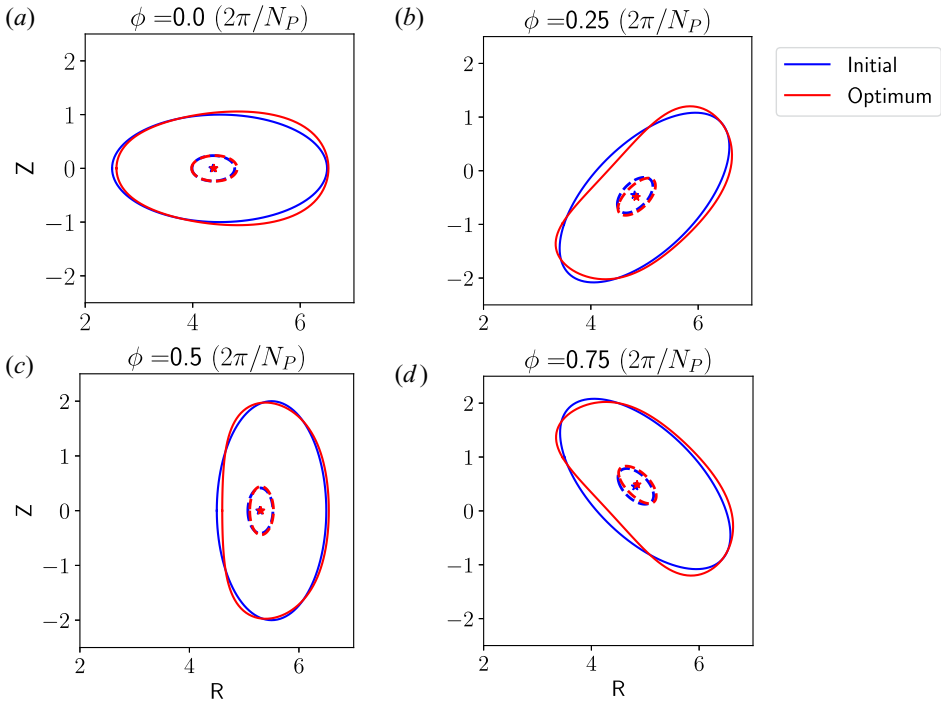


FIGURE 5. Poloidal cross-sections of the initial boundary (4.5) and boundary which optimizes (4.4) (solid) along with the surface at  $\psi/\psi_0 = 0.06$  (dashed) and magnetic axis (star).

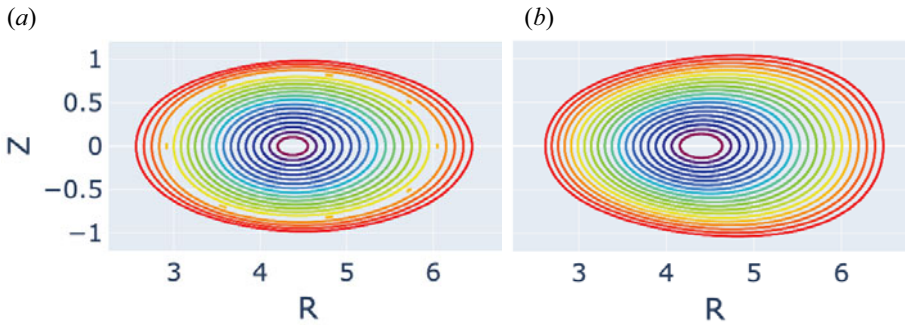


FIGURE 6. Poincaré section computed from the SPEC vacuum field using the initial boundary (4.5) and boundary which minimizes (4.4): (a) initial; (b) optimized.

to be a normalized finite-difference approximation of  $V''(\psi)$ ; thus minimization of  $f_w$  is performed in order to achieve a magnetic well. A similar objective is employed in the ROSE code (Drevlak *et al.* 2018), computed from integration along a field line. Some experimental observations point toward a correlation between the existence of a magnetic hill and the onset of fluctuations (Castellano *et al.* 2002). Nonetheless, we remark that the existence of a magnetic well does not always imply ideal MHD stability (Watanabe *et al.* 2005).

$\lambda_\kappa$	$\kappa_{\max}$	$\max( \kappa_1 ,  \kappa_2 )$	$f_w$
$10^2$	$7 \text{ m}^{-1}$	$6.55 \text{ m}^{-1}$	-0.21
$10^3$	$7 \text{ m}^{-1}$	$6.36 \text{ m}^{-1}$	-0.07
$10^0$	$8 \text{ m}^{-1}$	$7.84 \text{ m}^{-1}$	-0.19
$10^0$	$10 \text{ m}^{-1}$	$9.95 \text{ m}^{-1}$	-0.16

TABLE 1. Parameters used for the optimization of the magnetic well and the resulting values of the optimum objective function (4.8).

The gradient of this objective function is obtained by computing an equilibrium with a perturbed pressure profile,

$$\delta p(\psi) = \frac{w_1(\psi) - w_2(\psi) - f_w(w_1(\psi) + w_2(\psi))}{\int_{V_p} d^3x (w_1(\psi) + w_2(\psi))}, \tag{4.7}$$

as described in Paul *et al.* (2020).

We define our objective function to be

$$f(S_p) = f_w + \frac{\lambda_V}{2} (V_p - V_p^{\text{init}})^2 + \lambda_S f_S + \lambda_R f_R + \lambda_\kappa f_\kappa + \lambda_{\bar{\kappa}} f_{\bar{\kappa}}, \tag{4.8}$$

with  $\lambda_S = 10^{-3}$ ,  $\lambda_R = 10^2$ ,  $w_S = w_R = 10^{-2}$ ,  $\kappa_{\max} \equiv \kappa_{\max,1} = \kappa_{\max,2} = 7 \text{ m}^{-1}$ ,  $w_\kappa = 1$ ,  $\lambda_{\bar{\kappa}} = \lambda_V = 1$  and  $\lambda_\kappa = 10^2$ . Here  $V_p^{\text{init}}$  is the volume of the initial surface. To determine the importance of the maximum curvature regularization term, we optimize with four sets of parameters as described in table 1. The values  $\lambda_\kappa = 10^2$  and  $\kappa_{\max} = 7 \text{ m}^{-1}$  were chosen to balance the curvature and well metrics.

In addition to the regularization terms, we include a term in the objective function which penalizes a change in the volume, as an increase in the inverse aspect ratio can yield a Shafranov shift (§ 3.7 in Wesson & Campbell 2011). This shift in the flux surfaces toward a smaller major radius implies that the volume of a flux surface increases less rapidly than its cross-sectional area. Furthermore, the flux through a surface increases more rapidly than its cross-sectional area because the geometric centre is moving into a region of increased field strength (assuming the field is mostly toroidal). Thus, the volume increases less rapidly than the flux, and a negative value of  $V''(\psi)$  can be achieved (Taylor 1965).

We begin with a boundary given by a rotating ellipse (A 2) with  $a = 2$ ,  $b = 1$ ,  $R_0 = 5$  and  $R_{0,1} = 0$ . We optimize the boundary with respect to the modes  $m \leq 10$  and  $|n| \leq 10$ . The equilibrium is computed with  $p(\psi) = 0$  and  $I_T(\psi) = 0$  such that a vacuum field is considered. We arrive at the boundary given in figure 7. With the addition of the volume constraint, the aspect ratio remains roughly constant (3.54 for the initial boundary and 3.46 for the optimized boundary), so that the well is provided by the shaping of the boundary rather than the Shafranov shift that arises due to the inverse aspect ratio. We note that the optimized surface features triangularity that rotates from outward-pointing with horizontal elongation to inward-pointing with vertical elongation.

To demonstrate the effect of the curvature regularization terms, we examine the optimization with  $\lambda_\kappa = \lambda_{\bar{\kappa}} = 0$ . We converge to the boundary shown in figure 8. As can be seen, regions of large curvature are exhibited, including a dimple-like feature on the inboard side and a concave ‘pinching’ feature. Again the aspect ratio remains roughly constant (3.46).

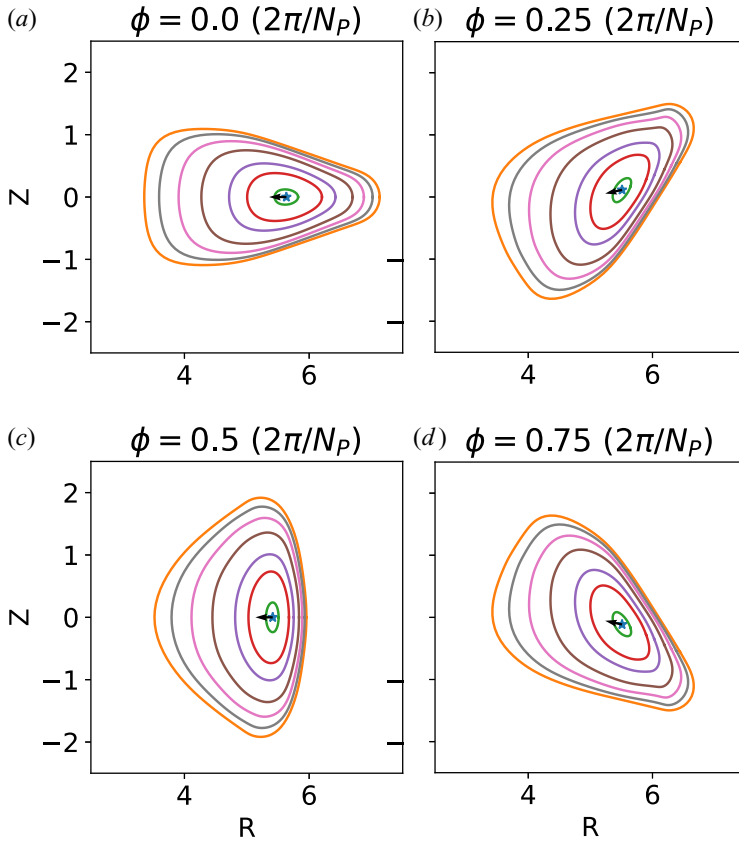


FIGURE 7. Cross-sections of the boundary (orange) optimized for the magnetic well (4.8) with regularization on the curvature ( $\lambda_\kappa = 10^2$ ,  $\lambda_{\bar{\kappa}} = 1$ ). Several interior surfaces are shown, along with the location of the magnetic axis (blue star) and axis normal vector (black arrow). The optimized boundary features rotating triangularity and elongation.

We can understand the geometric dependence of the magnetic well by considering the expression for  $V''(\psi)$  on the axis that arises from the near-axis expansion in the inverse coordinate representation (Landreman & Jorge 2020),

$$V''(\psi) \propto \int_0^{2\pi} d\varphi_B \frac{1}{B_0^4} [3(B_{1s}^2 + B_{1c}^2) - 4B_0 B_{20}], \tag{4.9}$$

where  $\varphi_B$  is the Boozer toroidal angle and we have made the assumption of a vacuum field. The field strength near the axis is expanded as

$$B = B_0(\varphi_B) + r [B_{1c}(\varphi_B) \cos(\vartheta_B) + B_{1s}(\varphi_B) \sin(\vartheta_B)] + r^2 [B_{20}(\varphi_B) + B_{2c} \cos(2\vartheta_B) + B_{2s} \sin(2\vartheta_B)] + O(r^3), \tag{4.10}$$

where  $r \propto \sqrt{\psi}$  is the effective minor radius and  $\vartheta_B$  is the Boozer poloidal angle.

As can be seen from (A34) in Landreman & Sengupta (2019), the poloidally independent shift in the flux surface in the normal direction,  $X_{20}$ , increases proportional to  $B_{20}$ . Here, the normal vector is  $\hat{n} = \mathbf{r}''(l)/|\mathbf{r}''(l)|$ , where  $\mathbf{r}(l)$  is the position vector along the magnetic axis parameterized by the length. In the axisymmetric limit, a positive value of  $X_{20}$  indicates a

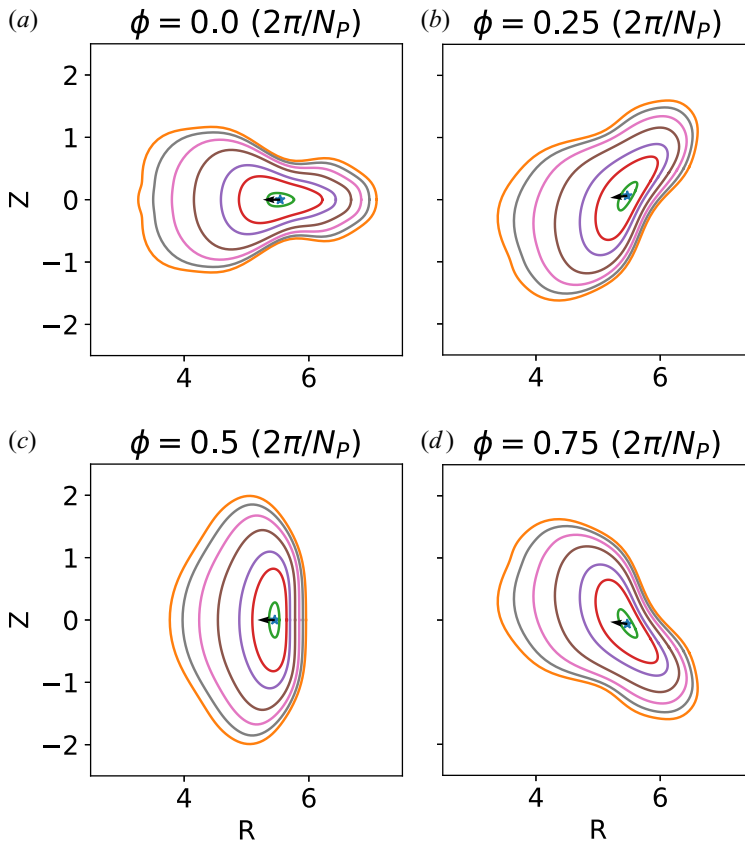


FIGURE 8. Cross-sections of the boundary optimized for the magnetic well (4.8) without regularization on the principal curvature ( $\lambda_\kappa = \lambda_{\bar{\kappa}} = 0$ ). Several interior surfaces are shown, along with the location of the magnetic axis (blue star) and axis normal vector (black arrow). The optimized boundary features a ‘dimple’ feature at the  $\phi = 0$  plane and regions of increased concavity.

net shift of the surfaces’ geometric centre toward a smaller major radius. This correlation between  $X_{20}$  and  $B_{20}$  is consistent with the statement that a Shafranov shift provides a magnetic well. In figures 7 and 8 we display the magnetic axis (blue star) and normal vector (black arrow) along with the shapes of several magnetic surfaces for the optimized configurations. While both the unconstrained and constrained optima feature a net shift in the geometric centre of the flux surfaces in the normal direction, this is achieved with vastly different shaping of the boundary. Interestingly, both configurations feature negative triangularity at the  $\phi = 0.5 (2\pi/N_p)$  plane.

We note that  $B_{20}$  arises owing to other shaping components of the surface (Landreman & Sengupta 2019). In axisymmetry, negative values of  $B_{2c}$  contribute to positive (outward-pointing) triangularity ( $X_{2c} < 0$ ) ((B11) in Landreman 2021). Assuming stellarator symmetry, negative values of  $B_{2c}$  contribute to positive values of  $B_{20}$  if the surface is vertically elongated, and positive values of  $B_{2c}$  contribute to positive values of  $B_{20}$  if the surface is horizontally elongated. This implies that positive triangularity coupled with vertical elongation or negative triangularity coupled with horizontal elongation contributes to the magnetic well in axisymmetry. These trends are consistent with the

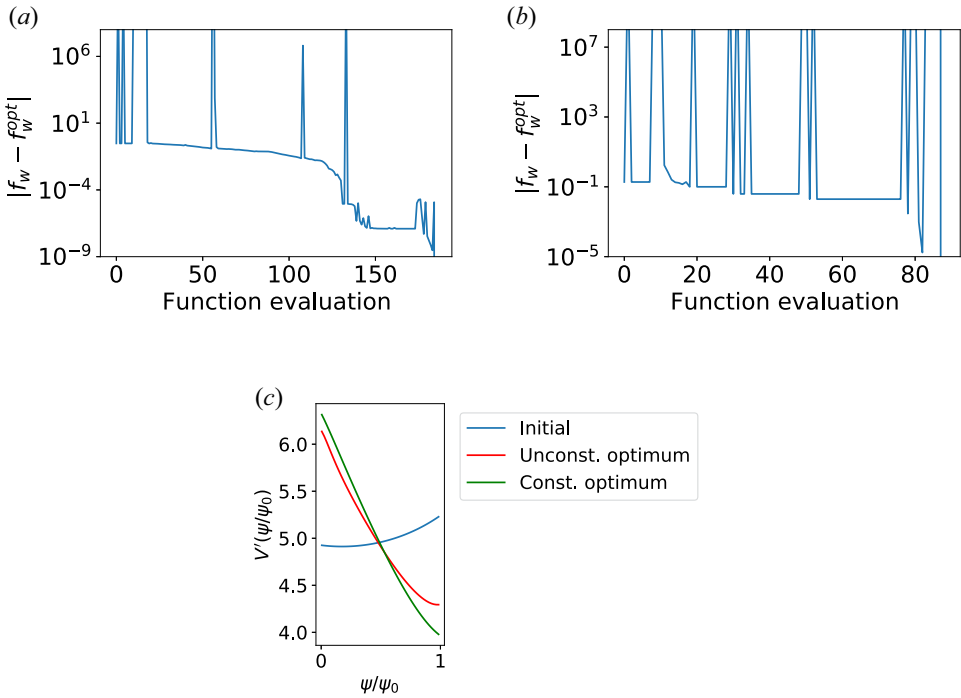


FIGURE 9. Convergence of the optimization of the objective function (4.8) (a) with the inclusion of the curvature constraints ( $\lambda_\kappa = 10^2$ ,  $\lambda_{\bar{\kappa}} = 1$ ) and (b) without the curvature constraints ( $\lambda_\kappa = \lambda_{\bar{\kappa}} = 0$ ). (c) With the inclusion of the curvature constraint, we achieve a deeper magnetic well throughout the volume.

stability analysis of oblate plasmas with negative triangularity (Pogutse & Yurchenko 1982; Kesner, Ramos & Gang 1995; Medvedev *et al.* 2015). In three dimensions, the connection between  $B_{20}$  and  $B_{2c}$  is more complicated (Landreman & Sengupta 2019, (A41) and (A42)). Nonetheless, we note that the magnetic well near the axis arises at second order in the expansion parameter, which includes the effects of ellipticity and triangularity. Thus, it is not surprising that the optimized boundary with curvature constraints features rotating ellipticity and triangularity.

In comparing the convergence of the optimization with and without the curvature constraints (figure 9), we see that the presence of the curvature objective prevents some failures of the VMEC code. Each function evaluation that resulted in a VMEC failure is visualized as a ‘spike’ that extends above  $|f_w - f_w^{\text{opt}}| = 1$ , as we artificially set the value of the objective function to  $10^{12}$  at these points. Although the constrained optimization still features some VMEC failures (12 versus 19), the inclusion of the curvature constraints improves the convergence toward the optimum, and a deeper magnetic well is achieved throughout the volume.

### 4.3. Quasi-symmetry

We next consider optimization for quasi-symmetry near the magnetic axis. This is quantified through the objective,

$$f_{QS} = \frac{\int_{V_p} d^3x w(\psi)(B - \bar{B})^2}{\int_{V_p} d^3x B^2}, \tag{4.11}$$



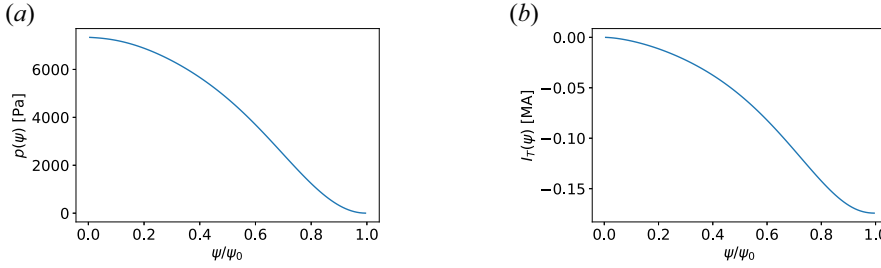


FIGURE 10. (a) The pressure and (b) the integrated toroidal current profiles used for the optimization of quasi-symmetry on the axis (4.13).

where  $\bar{B} = \int_{V_p} d^3x w(\psi) B$  and we take  $w(\psi) = \exp(-(\psi/\psi_0)^2/0.1^2)$ . This objective aims to make the field strength constant on the magnetic axis, which is a feature of quasi-axisymmetric and quasi-helically symmetric equilibria. A similar objective function has also been included in optimization for energetic particle confinement (Drevlak *et al.* 2014). The gradient of this objective function is obtained by computing an equilibrium with the addition of an anisotropic pressure tensor,  $\nabla \cdot \mathbf{P} = p_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + p_{\perp} (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$  with

$$p_{\parallel}(\psi, B) = \frac{w(\psi)(B - \bar{B})^2 - f_{QS}w(\psi)B^2}{\int_{V_p} d^3x w(\psi)B^2}, \tag{4.12a}$$

$$p_{\perp} = p_{\parallel} - B \frac{\partial p_{\parallel}(\psi, B)}{\partial B}, \tag{4.12b}$$

as described in Paul *et al.* (2020).

We take our objective function to be

$$f(S_p) = f_{QS} + \lambda_i f_i + \lambda_{\bar{\kappa}} f_{\bar{\kappa}} + \lambda_{\kappa} f_{\kappa} + \lambda_S f_S + \lambda_{\mathcal{R}} f_{\mathcal{R}}, \tag{4.13}$$

where  $\iota_{\text{target}}(\psi)$  is taken to be the initial rotational transform profile. We also take  $\lambda_i = 1$ ,  $\lambda_{\bar{\kappa}} = 10^{-3}$ ,  $\lambda_{\kappa} = 5 \times 10^{-2}$ ,  $\kappa_{\text{max},1} = \kappa_{\text{max},2} = 10 \text{ m}^{-1}$ ,  $w_{\kappa} = 1$ ,  $w_S = w_{\mathcal{R}} = 10^{-1}$ ,  $\lambda_{\mathcal{R}} = 10^3$ ,  $\lambda_S = 10^{-1}$ ,  $\mathcal{R} = 0.5 \text{ m}$  and  $R_{\text{min}} = 0.2 \text{ m}$ . The additional constraint on the rotational transform is required to prevent the surface from becoming axisymmetric to reduce the toroidal ripple on the axis.

We begin with a rotating ellipse boundary with torsion, given by

$$R(\theta, \phi) = R_0 + R_{0,1} \cos(\phi) + R_{0,2} \cos(2\phi) + 0.5a (\cos(\theta - N_p\phi) + \cos(\theta)) - 0.5b (\cos(\theta - N_p\phi) - \cos(\theta)) \tag{4.14a}$$

$$Z(\theta, \phi) = R_{0,1} \sin(\phi) + Z_{0,2} \cos(2\phi) + 0.5b (\sin(\theta) + \sin(\theta - N_p\phi)) + 0.5a (\sin(\theta) - \sin(\theta - N_p\phi)), \tag{4.14b}$$

with  $R_0 = 5$ ,  $R_{0,1} = R_{0,2} = 0.3$ ,  $a = 2$ ,  $b = 1$  and  $N_p = 3$ . We perform optimization at  $\beta = 1.2 \%$  with the profiles shown in figure 10. We optimize with respect to the modes  $m \leq 10$  and  $|n| \leq 10$  in a staged approach: first optimizing with respect to  $m \leq 3$  and  $|n| \leq 3$ , then with respect to  $m \leq 6$  and  $|n| \leq 6$ , then with respect to the full set of modes.

We display the magnetic field strength on the initial and optimized boundaries in figure 11. We note the initial configuration features a large toroidal variation of the field strength, with increased field strength near the ‘corners’ that arise owing to the large

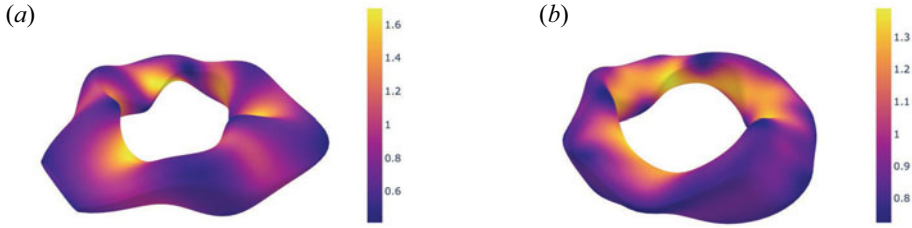


FIGURE 11. The magnetic field strength on (a) the initial boundary (4.14) and (b) the boundary optimized for quasi-symmetry on the axis (4.13).

torsion of the axis. Although the axis ripple figure of merit (4.11) reduces the toroidal variation of the field strength on the axis, which must vanish in both quasi-axisymmetry and quasi-helical symmetry, we find that the optimized magnetic field is driven closer to quasi-axisymmetry.

We remark that a design based solely on the objective (4.11) may be limited in that quasi-symmetry is not targeted away from the axis. We therefore quantify the departure from quasi-symmetry in the two configurations in two additional ways. The first is through the figure of merit,

$$f_{TP}(\psi) = \frac{\left\langle \left| \frac{\partial B}{\partial \theta} \frac{\partial}{\partial \phi} (\mathbf{B} \cdot \nabla B) - \frac{\partial B}{\partial \phi} \frac{\partial}{\partial \theta} (\mathbf{B} \cdot \nabla B) \right| \right\rangle_{\psi}}{\langle B^3 \rangle_{\psi} / r_{\text{minor}}(\psi)}, \tag{4.15}$$

where  $r_{\text{minor}}(\psi) = \sqrt{A(\psi)}/\pi$  is the averaged minor radius of a flux surface where  $A(\psi)$  is the toroidally averaged cross-sectional area. (This is the same definition of the minor radius used in the VMEC code.) The flux-surface average of a quantity  $Q$  is

$$\langle Q \rangle_{\psi} = \frac{\int_0^{2\pi} d\theta \int_0^{2\pi} d\phi Q \sqrt{g}}{\int_0^{2\pi} d\theta \int_0^{2\pi} d\phi \sqrt{g}}, \tag{4.16}$$

where  $\sqrt{g} = (\nabla\psi \times \nabla\theta \cdot \nabla\phi)^{-1}$  is the flux coordinate Jacobian. Equation (4.15) is a normalized figure of merit which employs the triple product form for quasi-symmetry (Helander 2014; Rodriguez, Helander & Bhattacharjee 2020),

$$\nabla\psi \times \nabla B \cdot \nabla (\mathbf{B} \cdot \nabla B) = 0, \tag{4.17}$$

and allows us to quantify the quasi-symmetry error without specifying the helicity of the symmetry. We also perform a Boozer coordinate  $(\psi, \vartheta_B, \varphi_B)$  transformation (Sanchez *et al.* 2000) to obtain the Fourier harmonics of the field strength,

$$B(\psi, \vartheta_B, \varphi_B) = \sum_{m,n} B_{m,n} \cos(m\vartheta_B - n\varphi_B). \tag{4.18}$$

The deviation from quasi-axisymmetry can then be quantified as

$$f_{QA}(\psi) = \sqrt{\frac{\sum_{m,n \neq 0} B_{m,n}(\psi)^2}{\sum_{m,n} B_{m,n}(\psi)^2}}. \tag{4.19}$$

Although quasi-symmetry was only targeted on the axis, we reduce the quasi-symmetry error throughout the volume with respect to both metrics without introducing a reduction in the rotational transform (figure 12).

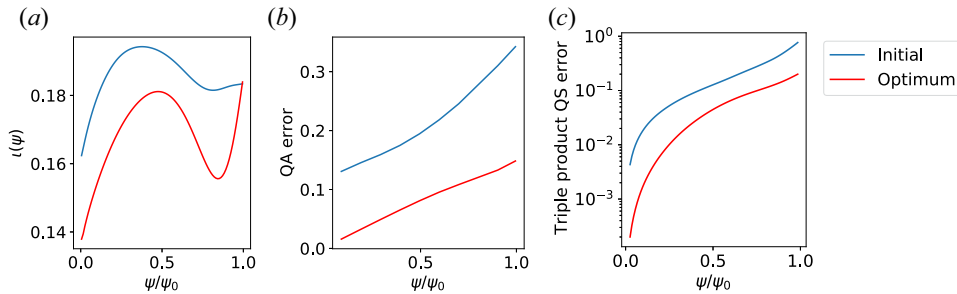


FIGURE 12. (a) The rotational transform, (b) the quasi-axisymmetric error (4.19) and (c) the triple product error (4.15) for the initial boundary (4.14) and boundary optimized for small magnetic ripple on the axis (4.13).

## 5. Conclusions

In this work, we have provided the first example of gradient-based, fixed-boundary optimization of stellarator equilibria. We provide examples of several equilibria obtained with the ALPOpt tool, including a vacuum field with ultra-low magnetic shear. Furthermore, we have identified regularization terms for fixed-boundary optimization that prevent self-intersection and reduce the surface curvature. These regularization terms improve the convergence toward the optimum and may reduce the required coil complexity. Another approach to reducing coil complexity would be to incorporate metrics related to the properties of the current potential solution on a uniformly offset winding surface (Carlton-Jones, Paul & Dorland 2020).

The availability of derivative information enables the optimization in high-dimensional spaces. We present the optimization of the boundary with respect to the set of boundary harmonics  $m \leq 10$ ,  $|n| \leq 10$  (441 modes). This dimensionality is significantly larger than that of previous optimized efforts. For example, ESTELL was optimized with respect to the modes  $m \leq 4$ ,  $|n| \leq 4$  (Drevlak *et al.* 2013) and NCSX was optimized with respect to  $m \leq 6$  and  $|n| \leq 4$  (Zarnstorff *et al.* 2001). Such an increase in the optimization space may enable further refinement of an optimum, as demonstrated in § 4.1.

Gradient information of equilibrium quantities is obtained using an adjoint method (Antonsen *et al.* 2019; Paul *et al.* 2020) which requires solving a modified equilibrium problem. For the objectives presented in this work, the rotational transform (§ 4.1), the magnetic well (§ 4.2) and near-axis quasi-symmetry (§ 4.3), the adjoint equilibrium problem requires the addition of a perturbation to the toroidal current profile, pressure profile and the addition of an anisotropic pressure tensor of a specific form. For other figures of merit, the adjoint equilibrium problem requires the addition of a bulk force of a different form, such as an anisotropic pressure that cannot be handled by the variational principle employed by the ANIMEC code (Cooper *et al.* 1992). In this case, rather than approximating the linearized adjoint problem by adding a perturbation to the nonlinear equilibrium solution, a linearized equilibrium solution can be computed. We have demonstrated this technique for computing the shape gradient of the magnetic well for axisymmetric equilibria (Paul 2020). In generalizing this approach to 3D equilibria, there are additional challenges that arise. As regular singular points occur at every surface where the rotational transform resonates with a mode included in the spectrum for the solution vector, additional care must be taken in regularizing the equations. To avoid these difficulties associated with 3D MHD equilibria, the adjoint equations for a vacuum or force-free equilibrium model could instead be considered.

We have presented several example configurations obtained with the set of figures of merit for which derivative information is available. Although similar objective functions can be minimized with derivative-free methods using codes such as STELLOPT and ROSE, the availability of derivative information is known to improve convergence toward the optimum, especially in high-dimensional spaces. A direct comparison between the adjoint approach and derivative-free methods is left for future work. Upon further advancement of adjoint methods, we hope that these numerical advances will enable the identification of equilibria of experimental relevance. To conclude, we anticipate many extensions of this work, such as applying the same adjoint principle for free-boundary optimization.

### Acknowledgements

The authors would like to acknowledge C. Zhu for the development of a Python interface with the VMEC code and S. R. Hudson for enlightening discussions and help with the SPEC calculations.

*Editor Per Helander thanks the referees for their advice in evaluating this article.*

### Funding

This work was supported by the US Department of Energy through grants DE-FG02-93ER-54197 and DE-FC02-08ER-54964. This work was also supported by a grant from the Simons Foundation (560651, ML). Some of the computations presented in this paper have used resources at the National Energy Research Scientific Computing Center (NERSC).

### Declaration of interests

The authors report no conflict of interest.

### Appendix A. Target rotational transform optimization

In this section, we perform a benchmark to demonstrate the convergence of the gradient-based optimization to a known minimum in a 2D space using the following objective function,

$$\tilde{f}_i = \frac{1}{2} (f_i - 0.06)^2, \quad (\text{A } 1)$$

defined with  $t_{\text{target}} = 0.618034$  with the definition in (4.1). Although this target function is not physically relevant, as there remains residual error in the rotational transform profile, it allows us to easily identify a local optimum where  $f_i = 0.06$ .

We begin with the boundary,

$$R(\theta, \phi) = R_0 + R_{1,0} \cos(\theta) + R_{1,1} \cos(\theta - N_P \phi), \quad (\text{A } 2a)$$

$$Z(\theta, \phi) = Z_{1,0} \sin(\theta) + Z_{1,1} \sin(\theta - N_P \phi), \quad (\text{A } 2b)$$

with  $R_0 = 5$ ,  $R_{1,0} = 1.5$ ,  $R_{1,1} = 0.6$ ,  $Z_{1,0} = 1.5$  and  $Z_{1,1} = -0.6$ . We consider a vacuum field with profiles  $p(\psi) = I_T(\psi) = 0$ . We identify the minimum of  $\tilde{f}_i$  with respect to  $R_{1,1}$  and  $Z_{1,1}$  by performing a scan over the local space. We achieve convergence to this optimum in nine function evaluations (six BFGS iterations). In figure 13 we present the convergence of the objective function in the 2D optimization space, with the optimum denoted by the blue star. We are able to reduce the L2 gradient norm to  $5.13 \times 10^{-13}$ .

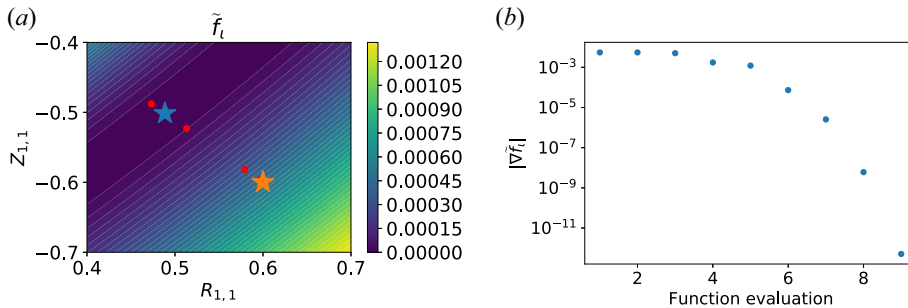


FIGURE 13. (a) Convergence of BFGS optimization beginning at the orange star toward optimum function value at  $R_{1,1} = 0.46$  and  $Z_{1,1} = -0.47$  (blue star). Each function evaluation is denoted by a red dot. (b) Convergence of the L2 norm of the gradient of the objective function (A 1).

## REFERENCES

- ANDERSON, F. S. B., ALMAGRI, A. F., ANDERSON, D. T., MATTHEWS, P. G., TALMADGE, J. N. & SHOHEH, J. L. 1995 The Helically Symmetric eXperiment, (HSX) goals, design and status. *Fusion Technol.* **27** (3T), 273–277.
- ANTONSEN, T., PAUL, E. J. & LANDREMAN, M. 2019 Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. *J. Plasma Phys.* **85** (2), 905850207.
- BEIDLER, C., GRIEGER, G., HERRNEGGER, F., HARMMEYER, E., KISSLINGER, J., LOTZ, W., MAASSBERG, H., MERKEL, P., NÜHRENBURG, J., RAU, F., *et al.* 1990 Physics and engineering design for Wendelstein VII-X. *Fusion Technol.* **17** (1), 148–168.
- CARLEN, M., LAURIE, B., MADDOCKS, J. H. & SMUTNY, J. 2005 Biarcs, global radius of curvature, and the computation of ideal knot shapes. In *Physical and Numerical Models in Knot Theory: Including Applications to the Life Sciences*, pp. 75–108. World Scientific.
- CARLTON-JONES, A., PAUL, E. J. & DORLAND, W. 2020 Computing the shape gradient of stellarator coil complexity with respect to the plasma boundary. *J. Plasma Phys.* [arXiv:2011.03702](https://arxiv.org/abs/2011.03702).
- CASTELLANO, J., JIMÉNEZ, J. A., HIDALGO, C., PEDROSA, M. A., FRAGUAS, A. L., PASTOR, I., HERRANZ, J., ALEJALDRE, C. & TJ-II TEAM 2002 Magnetic well and instability thresholds in the TJ-II stellarator. *Phys. Plasmas* **9** (2), 713–716.
- COOPER, W. A., HIRSHMAN, S. P., MERAZZI, S. & GRUBER, R. 1992 3D magnetohydrodynamic equilibria with anisotropic pressure. *Comput. Phys. Commun.* **72** (1), 1.
- DEKEYSER, W. 2014 Optimal plasma edge configurations for next-step fusion reactors. PhD thesis, Katholieke Universiteit Leuven.
- DREVLAK, M., BEIDLER, C. D., GEIGER, J., HELANDER, P. & TURKIN, Y. 2018 Optimisation of stellarator equilibria with ROSE. *Nucl. Fusion* **59** (1), 016010.
- DREVLAK, M., BROCHARD, F., HELANDER, P., KISSLINGER, J., MIKHAILOV, M., NÜHRENBURG, C., NÜHRENBURG, J. & TURKIN, Y. 2013 ESTELL: a quasi-toroidally symmetric stellarator. *Contrib. Plasma Phys.* **53** (6), 459–468.
- DREVLAK, M. C., GEIGER, J., HELANDER, P. & TURKIN, Y. 2014 Fast particle confinement with optimized coil currents in the W7-X stellarator. *Nucl. Fusion* **54** (7), 073002.
- EMMERICH, M., GIOTIS, A., ÖZDEMİR, M., BÄCK, T. & GIANNAKOGLU, K. 2002 Metamodel-assisted evolution strategies. In *International Conference on Parallel Problem Solving From Nature*, pp. 361–370. Springer.
- GELSEY, A. 1995 Intelligent automated quality control for computational simulation. *AI EDAM* **9** (5), 387–400.
- GELSEY, A., SCHWABACHER, M. & SMITH, D. 1998 Using modeling knowledge to guide design space search. *Artif. Intell.* **101** (1–2), 35–62.

- GIULIANI, A., WECHSUNG, F., CERFON, A., STADLER, G. & LANDREMAN, M. 2020 Single-stage gradient-based stellarator coil design: optimization for near-axis quasi-symmetry. [arXiv:2010.02033](https://arxiv.org/abs/2010.02033).
- GONZALEZ, O. & MADDOCKS, J. H. 1999 Global curvature, thickness, and the ideal shapes of knots. *Proc. Natl Acad. Sci.* **96** (9), 4769–4773.
- GREENE, J. M., MACKAY, R. S. & STARK, J. 1986 Boundary circles for area-preserving maps. *Physica D* **21** (2–3), 267–295.
- HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77** (8), 087001.
- HENNEBERG, S. A., DREVLAK, M., NÜHRENBURG, C., BEIDLER, C. D., TURKIN, Y., LOIZU, J. & HELANDER, P. 2019 Properties of a new quasi-axisymmetric configuration. *Nucl. Fusion* **59** (2), 026014.
- HIRSHMAN, S. P. & WHITSON, J. C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26** (12), 3553–3568.
- HUDSON, S. R., DEWAR, R. L., DENNIS, G., HOLE, M. J., MCGANN, M., VON NESSI, G. & LAZERSON, S. 2012 Computation of multi-region relaxed magnetohydrodynamic equilibria. *Phys. Plasmas* **19** (11), 112502.
- JOHNSON, S. G. 2014 The NLOpt nonlinear-optimization package, <http://github.com/stevengj/nlopt>.
- KESNER, J., RAMOS, J. J. & GANG, F.-Y. 1995 Comet cross-section tokamaks. *J. Fusion Energy* **14** (4), 361–371.
- LANDREMAN, M. 2021 Figures of merit for stellarators near the magnetic axis. *J. Plasma Phys.* **87** (1), 905870112.
- LANDREMAN, M. & JORGE, R. 2020 Magnetic well and Mercier stability of stellarators near the magnetic axis. *J. Plasma Phys.* **86** (5), 905860510.
- LANDREMAN, M. & SENGUPTA, W. 2019 Constructing stellarators with quasisymmetry to high order. *J. Plasma Phys.* **85** (6), 815850601.
- LAZERSON, S., SCHMITT, J., ZHU, C. & STELLOPT DEVELOPERS 2020 STELLOPT. <https://github.com/PrincetonUniversity/STELLOPT>.
- LEMARÉCHAL, C. 1982 Numerical experiments in nonsmooth optimization. IASA Workshop on Progress in Non-differentiable Optimization. 61–84. Laxenburg, Austria.
- LEWIS, A. S. & OVERTON, M. L. 2013 Nonsmooth optimization via quasi-newton methods. *Math. Program.* **141** (1), 135–163.
- MCGREIVY, N., HUDSON, S. R. & ZHU, C. 2021 Optimized finite-build stellarator coils using automatic differentiation. *Nucl. Fusion* **61** (2), 026020.
- MEDVEDEV, S. Y., KIKUCHI, M., VILLARD, L., TAKIZUKA, T., DIAMOND, P., ZUSHI, H., NAGASAKI, K., DUAN, X., WU, Y., IVANOV, A. A., *et al.* 2015 The negative triangularity tokamak: stability limits and prospects as a fusion energy system. *Nucl. Fusion* **55** (6), 063013.
- MEISS, J. D. 1992 Symplectic maps, variational principles, and transport. *Rev. Mod. Phys.* **64** (3), 795.
- MERCIER, C. 1964 Equilibrium and stability of a toroidal magnetohydrodynamic system in the neighbourhood of a magnetic axis. *Nucl. Fusion* **4** (3), 213.
- MERCIER, C. & LUC, H. 1974 *The MHD approach to the problem of plasma confinement in closed magnetic configurations*. Lectures in Plasma Physics. Commission of the European Communities.
- MINER, W. H. JR., Valanju, P. M., Hirshman, S. P., Brooks, A. & Pomphrey, N. 2001 Use of a genetic algorithm for compact stellarator coil design. *Nucl. Fusion* **41** (9), 1185.
- MYNICK, H. E., POMPHREY, N. & ETHIER, S. 2002 Exploration of stellarator configuration space with global search methods. *Phys. Plasmas* **9** (3), 869–876.
- NOCEDAL, J. & WRIGHT, S. J. 2006 *Numerical Optimization*. Springer.
- NÜHRENBURG, J. & ZILLE, R. 1988 Quasi-helically symmetric toroidal stellarators. *Phys. Lett. A* **129** (2), 113–117.
- OYMAK, S. & SOLTANOLKOTABI, M. 2018 Overparameterized nonlinear learning: gradient descent takes the shortest path? [arXiv:1812.10004](https://arxiv.org/abs/1812.10004).
- PAUL, E. J. 2020 Adjoint methods for stellarator shape optimization and sensitivity analysis. [arXiv:2005.07633](https://arxiv.org/abs/2005.07633).

- PAUL, E. J., ABEL, I. G., LANDREMAN, M. & DORLAND, W. 2019 An adjoint method for neoclassical stellarator optimization. *J. Plasma Phys.* **85** (5), 795850501.
- PAUL, E. J., ANTONSEN, T., LANDREMAN, M. & COOPER, W. A. 2020 Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. Part 2. Applications. *J. Plasma Phys.* **86** (1), 905860103.
- PAUL, E. J., LANDREMAN, M., BADER, A. & DORLAND, W. 2018 An adjoint method for gradient-based optimization of stellarator coil shapes. *Nucl. Fusion* **58** (7), 076015.
- POGUTSE, O. & YURCHENKO, E. 1982 *Reviews of Plasma Physics*. Consultants Bureau.
- RASHEED, K., HIRSH, H. & GELSEY, A. 1997 A genetic algorithm for continuous design space search. *Artif. Intell. Eng.* **11** (3), 295–305.
- RODRIGUEZ, E., HELANDER, P. & BHATTACHARJEE, A. 2020 Necessary and sufficient conditions for quasisymmetry. *Phys. Plasmas* **27** (6), 062501.
- SANCHEZ, R., HIRSHMAN, S. P., WARE, A. S., BERRY, L. A. & SPONG, D. A. 2000 Ballooning stability optimization of low-aspect-ratio stellarators. *Plasma Phys. Control. Fusion* **42** (6), 641.
- SAUER, T. 2012 *Numerical Analysis*. Pearson.
- SMUTNY, J. 2004 Global radii of curvature, and the biarc approximation of space curves: in pursuit of ideal knot shapes. PhD thesis, EPFL.
- TAYLOR, J. B. 1965 Simple toroidal magnetic field with negative  $V''$ . *Phys. Fluids* **8** (6), 1203–1205.
- WALKER, S. W. 2016 Shape optimization of self-avoiding curves. *J. Comput. Phys.* **311**, 275–298.
- WATANABE, K. Y., SAKAKIBARA, S., NARUSHIMA, Y., FUNABA, H., NARIHARA, K., TANAKA, K., YAMAGUCHI, T., TOI, K., OHDACHI, S., KANEKO, O., *et al.* 2005 Effects of global MHD instability on operational high beta-regime in LHD. *Nucl. Fusion* **45** (11), 1247.
- WESSON, J. & CAMPBELL, D. J. 2011 *Tokamaks*. International Series of Monographs on Physics, vol. 149. Oxford University Press.
- ZARNSTORFF, M. C., BERRY, L. A., BROOKS, A., FREDRICKSON, E., FU, G. Y., HIRSHMAN, S., HUDSON, S., KU, L. P., LAZARUS, E., MIKKELSEN, D., *et al.* 2001 Physics of the compact advanced stellarator NCSX. *Plasma Phys. Control. Fusion* **43** (12A), A237.
- ZHU, C., HUDSON, S. R., SONG, Y. & WAN, Y. 2018 New method to design stellarator coils without the winding surface. *Nucl. Fusion* **58**, 016008.