

# An adjoint-based method for optimising MHD equilibria against the infinite- $n$ , ideal ballooning mode

Rahul Gaur<sup>1,†</sup>, Stefan Buller<sup>1</sup>, Maximilian E. Ruth<sup>2</sup>, Matt Landreman<sup>1</sup>, Ian G. Abel<sup>1</sup> and William D. Dorland<sup>1,3</sup>

<sup>1</sup>Institute for Research in Electronics and Applied Physics, University of Maryland, College Park 20740, MD, USA

<sup>2</sup>Center for Applied Mathematics, Cornell University, Ithaca 14850, NY, USA

<sup>3</sup>Department of Physics, University of Maryland, College Park 20740, MD, USA

(Received 15 February 2023; revised 6 September 2023; accepted 7 September 2023)

We demonstrate a fast adjoint-based method to optimise tokamak and stellarator equilibria against a pressure-driven instability known as the infinite- $n$  ideal ballooning mode. We present three finite- $\beta$  (the ratio of thermal to magnetic pressure) equilibria: one tokamak equilibrium and two stellarator equilibria that are unstable against the ballooning mode. Using the self-adjoint property of ideal magnetohydrodynamics, we construct a technique to rapidly calculate the change in the eigenvalue, a measure of ideal ballooning instability. Using the SIMSOPT optimisation framework, we then implement our fast adjoint gradient-based optimiser to minimise the eigenvalue and find stable equilibria for each of the three originally unstable equilibria.

**Keywords:** fusion plasma, plasma instabilities, plasma confinement

## 1. Introduction

Magnetic confinement is currently considered the most promising way (Baalrud *et al.* 2020) to achieve the United States' goal of building a pilot fusion power plant that generates net electricity before 2040 (National Academies of Sciences *et al.* 2021). Most advanced fusion reactor designs today are based on two main designs that use magnetic confinement: tokamaks and stellarators. These devices work by using strong magnetic fields to keep a hot, dense plasma at their centre. The main difference between tokamaks and stellarators lies in their geometric shape. Tokamaks are symmetric about a fixed axis, whereas stellarators are not. Due to the difference in axisymmetry, the tokamak equilibria are two-dimensional (2D) axisymmetric and the stellarator equilibria are three-dimensional (3D).

For a fixed magnetic field strength, the power density  $P$  of a fusion device scales as  $\beta^2$ , where  $\beta$  is the ratio of the plasma pressure to the magnetic pressure. Since

† Email address for correspondence: [rgaur@umd.edu](mailto:rgaur@umd.edu)

current tokamaks and stellarators are low- $\beta$  devices, one way to improve the efficiency of these devices is to increase the operating  $\beta$ . However, doing so creates a large pressure gradient from the centre to the edge of the device, which is a source of a variety of pressure-driven, magnetohydrodynamic (MHD) and kinetic instabilities. An important category of pressure-driven instabilities are the high- $n$  ballooning modes.

In tokamaks, there have been numerous studies that have used the ideal ballooning mode to determine the plasma beta limit. This led to the development of codes such as EPED (Snyder *et al.* 2007) that can determine edge pressure profile with reasonable accuracy. In stellarators, the ideal ballooning mode might not cause a significant loss of energy, but Tang, Connor & Hastie (1980) and Aleynikova *et al.* (2018) have shown that it is related to a kinetic instability known as the kinetic ballooning mode (KBM). Recent articles have shown the detrimental effects of KBM turbulence on stellarators for finite beta values (Aleynikova *et al.* 2018; McKinney *et al.* 2021). The KBM is similar to the ideal ballooning mode, albeit with additional kinetic effects. Therefore, the ideal ballooning mode could be used as a proxy for KBM stability.

Numerous studies have been conducted to optimise tokamaks (Miller & Moore 1979; Bernard & Moore 1981) and stellarators (Sanchez *et al.* 2000a; Gates *et al.* 2017) against ideal MHD instabilities. However, such calculations can be computationally costly and time consuming. In this paper, we use the self-adjoint property of ideal MHD to devise an adjoint-based method that speeds up the optimisation of 2D and 3D equilibria against the infinite- $n$ , ideal ballooning mode. Adjoint-based methods have been extensively used for aeronautical design (Giles & Pierce 2000) and recently in the context of stellarator optimisation (see Paul, Landreman & Antonsen (2021) and references therein). Using this technique, we can speed up the process of finding equilibria that are stable against the ballooning mode.

The remainder of this paper is divided as follows: in §2, we briefly describe the fundamentals of a general 3D MHD equilibrium and follow it with the details of the VMEC equilibrium solver (Hirshman & Whitson 1983) in §2.1. Using VMEC, we obtain and present the details of one 2D axisymmetric equilibrium in §2.2 and two 3D equilibria in §§2.3 and 2.4. In §3, we present the physical, mathematical and numerical details used to solve the infinite- $n$ , ideal ballooning equation. We then analyse the susceptibility of the chosen local equilibria to the ideal ballooning instability. In §3.3, we explain the self-adjoint property of the ideal ballooning equation. We briefly explain how the ballooning eigenvalue can be used as a proxy to stabilise the equilibria against the KBM. Using the self-adjoint technique, we formulate an adjoint method which we elucidate and test in §4. In §5, we present the details of the overall optimisation process and present an adjoint-based optimiser using the SIMSOPT (Landreman *et al.* 2021) framework. In the penultimate section, we present our results, comparing the optimised stable equilibria with their initial, unstable counterparts. Finally, in §7 we summarise our work and discuss possible ways in which it can be extended.

## 2. Ideal MHD equilibrium

In this section, we start with the general form of a 3D, divergence-free magnetic field. We use this form to represent the magnetic field in tokamaks and stellarators. After that, we briefly describe the steady-state, ideal MHD, force-balance equation. In §2.1, we explain the details of solving the ideal MHD force-balance equation using the VMEC code. Finally, we present the details of three MHD equilibria in §§2.2–2.4 that we will use throughout this study.

A divergence-free magnetic field  $\mathbf{B}$  can be written in the Clebsch form (D’haeseleer *et al.* 2012)

$$\mathbf{B} = \nabla\alpha_t \times \nabla\psi_p. \quad (2.1)$$

The form (2.1) is generally used for tokamak equilibria. For stellarators, we use the following equivalent representation

$$\mathbf{B} = \nabla\psi \times \nabla\alpha_s. \quad (2.2)$$

We focus on solutions whose magnetic field lines lie on closed nested toroidal surfaces, known as flux surfaces. For tokamaks, we label the flux surfaces with their enclosed poloidal flux  $\psi_p$  whereas for stellarators, we use the enclosed toroidal flux  $\psi$ . On each flux surface, lines of constant  $\alpha_t$  and  $\alpha_s$  coincide with the magnetic field lines in tokamaks and stellarators, respectively. Thus, the variables  $\alpha_t$  and  $\alpha_s$  are known as field line labels.

To facilitate the calculation of various physical quantities from a general equilibrium solver, we use multiple coordinate systems. We use the right-handed cylindrical coordinate system  $(R, \zeta, Z)$  where  $R$  and  $Z$  are the radial and vertical distances from the origin and  $\zeta$  is the azimuthal angle around the symmetry axis. We also define a curvilinear coordinate system comprising the PEST coordinates  $(\psi_p, \zeta, \theta)$  where  $\psi_p$  is the flux surface label,  $\zeta$  is the cylindrical azimuthal angle and  $\theta$  is the ‘straight-field-line’ poloidal angle (D’haeseleer *et al.* 2012) such that  $\alpha_t = \zeta - q(\psi_p)(\theta - \theta_0)$ . Similarly, for 3D equilibria, we use the coordinate system  $(\psi, \zeta, \theta)$  and  $\alpha_s = \theta - \iota(\psi)(\zeta - \zeta_0)$ . The pitch of the magnetic field line on a flux surface is described by the safety factor

$$q(\psi) = \frac{1}{\iota(\psi)} \equiv \frac{d\psi}{d\psi_p} = \frac{1}{(2\pi)^2} \oint d\zeta \oint d\theta \frac{\mathbf{B} \cdot \nabla\zeta}{\mathbf{B} \cdot \nabla\theta}, \quad (2.3)$$

where  $\iota$ , the rotational transform, is the inverse of the safety factor. Using (2.1) or (2.2) for the magnetic field, one has to solve the steady-state ideal MHD force balance equation in the absence of a background flow

$$\mathbf{j} \times \mathbf{B} = \nabla p, \quad (2.4)$$

where  $p$  is the plasma pressure, and  $\mathbf{j}$  is the plasma current given by the Ampere’s law

$$\mu_0 \mathbf{j} = \nabla \times \mathbf{B}, \quad (2.5)$$

where  $\mu_0$  is the coefficient of permeability in vacuum. For axisymmetric equilibria, simplifying (2.4) yields the Grad–Shafranov equation (Shafranov 1957; Grad & Rubin 1958). For a general 3D equilibrium, such an equation does not exist. However, we can solve for the axisymmetric and 3D cases using a general numerical equilibrium solver. In the following section, we explain how we use the numerical solver VMEC<sup>1</sup> to obtain both 2D axisymmetric and 3D equilibria.

<sup>1</sup>The main idea of this work is independent of the equilibrium solver. Our technique should also work with any other equilibrium solver.

2.1. Numerical equilibrium solver

We generate numerical equilibria using the 3D equilibrium code VMEC (Hirshman & Whitson 1983). The code works by minimising the integral

$$W = \int \left( \frac{p}{\gamma - 1} + \frac{B^2}{2\mu_0} \right) dV, \tag{2.6}$$

subject to multiple geometric constraints (Kruskal & Kulsrud 1958). For our study, we used the fixed-boundary mode of VMEC. The fixed-boundary mode takes the shape of the boundary surface denoted by the cylindrical coordinates  $R_b$  and  $Z_b$  in terms of the Fourier-decomposed poloidal ( $\Theta$ ) and toroidal ( $\zeta$ ) modes

$$\left. \begin{aligned} R_b &= \sum_n \sum_m \hat{R}_b(m, n) \exp(i(m\Theta - n\zeta)), \\ Z_b &= \sum_n \sum_m \hat{Z}_b(m, n) \exp(i(m\Theta - n\zeta)), \end{aligned} \right\} \tag{2.7}$$

where  $m$  and  $n$  are integers. We also provide VMEC with the coefficients of the polynomials representing the global radial pressure  $p(s)$  and the rotational transform  $\iota(s)$  as a function of the normalised toroidal flux  $s$ , and the total toroidal or poloidal flux enclosed by the boundary. The poloidal angle  $\Theta$  used by VMEC is related to the straight field line  $\theta$  by the following equation

$$\Theta = \theta + \Lambda, \tag{2.8}$$

where

$$\Lambda = \sum_n \sum_m \hat{\Lambda}(m, n) \exp(i(m\Theta - n\zeta)). \tag{2.9}$$

For a boundary shape, pressure, rotational transform and enclosed toroidal flux, it solves for the flux surfaces. To achieve this, it uses a steepest descent method to minimise the energy integral in (2.6) on each surface for fixed  $p$ ,  $\iota$ , and enclosed flux subject to various topological constraints imposed by the ideal MHD. In a more compact form, VMEC solves

$$\min_{R, Z, \Lambda} W[R, Z, \Lambda; p, \iota, \psi(s = 1)], \quad \text{s.t. } R(s = 1) = R_b, Z(s = 1) = Z_b. \tag{2.10}$$

After running the code, we obtain the shape of the flux surfaces, the magnetic field and a set of important physical quantities. The characteristic physical quantities that we use in this work are defined as follows.

- (i) The total enclosed toroidal flux by the boundary  $\psi_b = \int dV \mathbf{B} \cdot \nabla \zeta$ .
- (ii) The normalising magnetic field  $B_N = \psi_b / (\pi a_N^2)$  where  $a_N = \sqrt{\mathcal{A}_b / \pi}$  is the effective minor radius and  $\mathcal{A}_b$  is the average area enclosed by the boundary (averaged over  $\zeta$ ).
- (iii) The ratio of the total plasma pressure to the magnetic pressure on the magnetic axis  $\beta_{ax} = 2\mu_0 p(s = 0) / B_N^2$ .
- (iv) The aspect ratio  $A$  and the normalised minor radius  $a_N$  of the device.
- (v) The radius of curvature of the boundary  $R_c(\theta) = (d^2R/dZ^2) / (1 + (dR/dZ)^2)^{3/2}$  where  $R$  and  $Z$  are the cylindrical coordinates used to parametrise the boundary.
- (vi) The volume-averaged, normalised plasma pressure  $\langle \beta \rangle = \int dV p / \int dV B^2$  where  $dV$  is the differential volume element.
- (vii) The total enclosed toroidal current  $j_\zeta = |\int dV (\mathbf{j} \cdot \nabla \zeta)|$ .

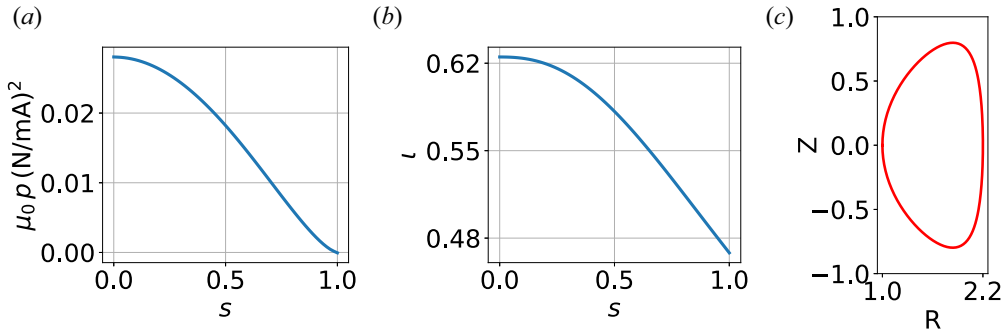


FIGURE 1. Plots of the inputs to the VMEC code for the DIII-D-like design: (a) the pressure, (b) the rotational transform as a function of the normalised toroidal flux  $s$  and (c) the cross section of the boundary.

---

$\beta_{ax}(\%)$	$\langle\beta\rangle(\%)$	$j_z(\text{MA})$	$\langle B\rangle(\text{T})$	$\bar{\iota}$	$A$	$a_N(\text{m})$	$B_N(\text{T})$	$\psi_{\text{LCFS}}(\text{Tm}^2)$
14	7.6	0.616	0.677	0.568	2.42	0.68	0.679	1.0

---

TABLE 1. Relevant physical quantities for the DIII-D-like equilibrium.

(viii) The volume-averaged magnetic field  $\langle B \rangle = \int dV B / \sqrt{V}$  where  $dV$  is the differential volume element.

(ix) The mean rotational transform  $\bar{\iota} = \int ds \iota / \int ds$ .

Using VMEC, we generate three equilibria: an axisymmetric equilibrium with a DIII-D-like boundary shape and two 3D equilibria, modified NCSX- and modified Henneberg-QA. In the following sections, we provide important details for each of these equilibria.

### 2.2. 2D axisymmetric equilibrium

In this study, the first equilibrium that we choose is a high- $\beta$ , axisymmetric, DIII-D-like equilibrium with a negative triangularity boundary: a boundary that looks like an inverted-D. Negative triangularity equilibria have generally been found to have enhanced confinement Marinoni *et al.* (2019) and stability against edge localised modes. We then choose a negative triangularity equilibrium from Gaur *et al.* (2023) where it is shown to be unstable against the ideal ballooning mode. With this equilibrium as an initial state, we run our ideal ballooning stability optimisation to find a stable equilibrium while maintaining a boundary shape with negative triangularity. The input pressure, the rotational transform and the boundary shape profile for this equilibrium are shown in figure 1.

Using these inputs, we run VMEC to obtain the global MHD equilibrium. Our optimiser sometimes finds solutions that meet the ideal ballooning stability constraints in a trivial fashion. For example, the optimiser may give us a large aspect ratio or decrease the minor radius causing the volume averaged  $B$  to increase, causing the  $\beta$  to decrease. To avoid these trivial solutions, we have to impose additional constraints on important characteristic physical quantities to prevent them from changing significantly. The values of the relevant physical quantities for this equilibrium are provided in table 1.

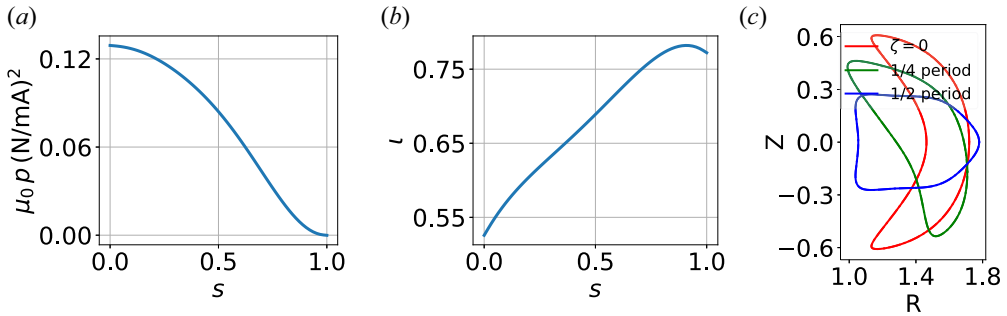


FIGURE 2. Plots of the inputs to the VMEC code for the modified NCSX design: (a) the pressure, (b) the rotational transform as a function of the normalised toroidal flux  $s$  and (c) the cross section of the boundary. Note the large negative shear until  $s \approx 0.85$ .

$\beta_{\text{ax}}(\%)$	$\langle\beta\rangle(\%)$	$j_{\zeta}(\text{MA})$	$\langle B\rangle(\text{T})$	$\bar{i}$	$A$	$a_N(\text{m})$	$B_N(\text{T})$	$\psi_{\text{LCFS}}(\text{m}^2)$
10.9	5.9	0.305	1.59	0.68	4.36	0.32	1.54	0.514

TABLE 2. Important physical quantities for the modified NCSX equilibrium.

### 2.3. Modified NCSX equilibrium

The second equilibrium we select is the 3D equilibrium for the NCSX design (Zarnstorff *et al.* 2001; Fu *et al.* 2007). This equilibrium is designed to have a hidden symmetry known as quasisymmetry (Boozer 1983; Garren & Boozer 1991) where the strength of the magnetic field  $|\mathbf{B}|$  does not change along a special coordinate. Quasisymmetry<sup>2</sup> is a useful property because it ensures orbit confinement, which helps improve energetic particle confinement, a major issue in stellarators. The pressure, rotational transform and boundary shape profile for this equilibrium are shown in figure 2.

The values of equilibrium-dependent quantities output by VMEC are presented in table 2.

### 2.4. Modified Henneberg-QA

The final equilibrium we study is the modified Henneberg-QA design (Henneberg *et al.* 2019). This equilibrium is also designed to have quasisymmetry for a wide variety of pressure profiles. The pressure, rotational transform and boundary shape profile for this equilibrium are shown in figure 3.

For reasons explained in the previous section, we impose additional constraints on some of the physical quantities. The values of these equilibrium-dependent parameters that we use as constraints in § 6 are presented in table 3.

In the following section, we describe the ideal ballooning stability and analyse these equilibria by measuring their instability against the ideal ballooning mode.

## 3. The infinite- $n$ ideal ballooning mode

In this section, we present the details of the infinite- $n$  ideal ballooning mode. In § 3.1, we briefly describe its theoretical foundation and mathematical formulation. In § 3.2, we present a numerical technique used to solve the ideal ballooning equation. In § 3.3, we

<sup>2</sup>We do not optimise equilibria for quasisymmetry in this work.

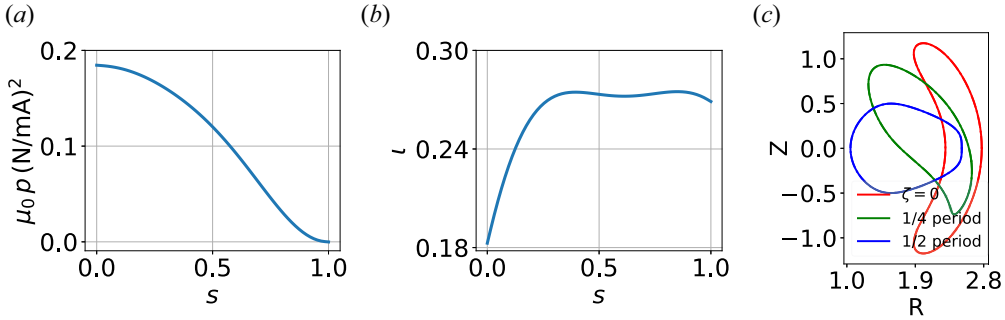


FIGURE 3. Plots of the inputs to the VMEC code for the modified Henneberg-QA design: (a) the pressure, (b) the rotational transform as a function of the normalised toroidal flux  $s$  and (c) the cross section of the boundary.

$\beta_{ax}(\%)$	$\langle\beta\rangle(\%)$	$j_{\zeta}(\text{MA})$	$\langle B\rangle(\text{T})$	$\bar{i}$	$A$	$a_N(\text{m})$	$B_N(\text{T})$	$\psi_{\text{LCFS}}(\text{Tm}^2)$
6.5	3.56	0.204	2.5	0.263	3.37	0.60	2.35	2.67

TABLE 3. Relevant physical quantities for the modified Henneberg-QA design.

describe the mathematical properties that we use to formulate an adjoint-based method and accelerate optimisation against the ideal ballooning mode. In the final section, we describe how optimisation against the ideal ballooning mode can speed up optimisation against an electromagnetic mode seen in kinetic plasma turbulence, known as the KBM.

### 3.1. Physical and mathematical description

This work involves a detailed analysis of three equilibria against an important MHD instability, the infinite- $n$  ideal ballooning instability (Connor, Hastie & Taylor 1979; Dewar & Glasser 1983): a field-aligned, pressure-driven Alfvén wave that grows when the destabilising pressure gradient in the region of ‘bad’ curvature exceeds the stabilising effect of field-line bending. Using magnetic field unit vector  $\mathbf{b} = \mathbf{B}/B$ , the region of ‘bad’ curvature is defined as a region of a flux surface where  $(\mathbf{b} \cdot \nabla \mathbf{b}) \cdot \nabla p > 0$ .

The ideal ballooning equation

$$\frac{1}{\mathcal{J}} \frac{\partial}{\partial \theta} \left( \frac{|\nabla \alpha_t|^2}{\mathcal{J} B^2} \frac{\partial \hat{X}}{\partial \theta} \right) + 2 \frac{1}{B^4} \frac{d(\mu_0 p)}{d\psi} \left[ \mathbf{B} \times \nabla \left( \mu_0 p + \frac{B^2}{2} \right) \cdot \nabla \alpha_t \right] \hat{X} = -\rho \omega^2 \frac{|\nabla \alpha_t|^2}{B^2} \hat{X}, \tag{3.1}$$

is a second-order eigenvalue differential equation that calculates the perturbation  $\hat{X}(\theta)$  along the ballooning coordinate  $\theta$  and its eigenvalue (or the square of the growth rate)  $-\omega^2$ . In (3.1),  $\rho$  is the plasma mass density,  $\mathcal{J} = (\mathbf{B} \cdot \nabla \theta)^{-1}$  and the rest of the terms are defined in §2. This equation is solved subject to the following condition on the eigenfunction

$$\lim_{\theta \rightarrow \pm\infty} \hat{X}(\theta; \psi, \alpha_t, \theta_0) = 0, \tag{3.2}$$

where  $\theta_0$  is the ballooning parameter.<sup>3</sup> Details of the derivation of the ideal ballooning equation are given in Connor *et al.* (1979) and Dewar & Glasser (1983).

The ballooning equation balances the stabilising field-line bending term and the destabilising pressure gradient with the inertia of the resulting Alfvén wave, oscillating with a frequency  $\omega$ . Note that (3.1) depends on  $\psi$  (or  $\psi_p$ ) as a parameter, and we can compute the coefficients from the equilibrium quantities on each surface. Before solving (3.1) numerically, we normalise and write the ballooning equation on a field line (fixed  $\alpha_t$ ) as

$$\frac{d}{d\theta} g \frac{d\hat{X}}{d\theta} + c\hat{X} = \hat{\lambda}f\hat{X}, \tag{3.3}$$

where

$$\left. \begin{aligned} g &= (\mathbf{b} \cdot \nabla_N \theta) \frac{|\nabla_N \alpha_t|^2}{B/B_N}, \\ c &= \frac{2}{(\mathbf{b} \cdot \nabla_N \theta)} \frac{1}{(B/B_N)^4} \frac{d(\mu_0 p/B_N^2)}{d\psi_N} \left[ \mathbf{b} \times \nabla_N \left( \frac{2\mu_0 p + B^2}{2B_N^2} \right) \cdot \nabla_N \alpha_t \right], \\ f &= \frac{1}{(\mathbf{b} \cdot \nabla_N \theta)} \frac{|\nabla_N \alpha_t|^2}{(B/B_N)^3}, \\ \hat{\lambda} &= - \left( \frac{\omega a_N}{v_A} \right)^2, \quad v_A = \frac{B_N}{\sqrt{4\pi\rho}}, \end{aligned} \right\} \tag{3.4}$$

where  $v_A$  is the Alfvén speed and the values and definitions of the effective minor radius  $a_N$  and the normalising magnetic field  $B_N$  are the normalising length and magnetic field strength, respectively, defined in §2.1. The normalised gradient  $\nabla_N = a_N \nabla$ . In the next section, we present the numerical procedure used to solve the ideal ballooning equation.

### 3.2. Numerical implementation and eigenvalues of the selected equilibria

In this section, we briefly discuss the numerical technique used to solve the ballooning equation (3.3). Our numerical technique is virtually identical to that used by Sanchez *et al.* (2000b) in their ballooning solver COBRAVMEC. After briefly explaining the details of our solver, we present the maximum eigenvalue as a function of the normalised toroidal flux  $s$  for the three equilibria we presented in §2.

The ideal ballooning equation is a second-order ordinary differential equation with real-valued coefficients. This eigenvalue equation is discretised using a second-order accurate, central-finite-difference scheme

$$g_{j+1/2} \frac{(\hat{X}_{j+1} - \hat{X}_j)}{\Delta\theta^2} - g_{j-1/2} \frac{(\hat{X}_j - \hat{X}_{j-1})}{\Delta\theta^2} + (c_j - \hat{\lambda}f_j)\hat{X}_j = 0, \quad j = 0 \dots N - 1 \tag{3.5}$$

subject to the boundary conditions

$$\hat{X}_0 = 0, \quad \hat{X}_N = 0, \tag{3.6}$$

where  $N$  is an odd number of uniformly spaced points in the ballooning space,  $\theta_b$  is a finite user-selected value that determines the extent of the eigenfunction such that  $\theta_j \in [-\theta_b, \theta_b]$

<sup>3</sup>In the context of infinite- $n$  ideal ballooning mode analyses, there is a value of the ballooning parameter  $\theta_0$  at which the ballooning mode is the least stable. To find this value, we treat  $\theta_0$  as a free parameter and scan its values to find  $\theta_0$  for which  $\omega^2$  is the most negative.



and  $\Delta\theta = \theta_{j+1} - \theta_j$ . First-order derivatives are evaluated at half points  $j - 1/2, j - 3/2$  instead of grid points to ensure numerical stability. For a fixed poloidal and toroidal resolution, the time taken by our solver is proportional to  $\theta_b$ . Therefore, it is important to find the right balance between speed and accuracy. Throughout this work, we chose the domain limit  $\theta_b = 5\pi$  for all calculations, as we found it to be a sufficient range to capture the ballooning eigenfunction. We observed that the values  $\theta_b > 5\pi$  made a relatively small difference from the value obtained of  $\hat{\lambda}$ . The discrete set of (3.5) is written in the form of a matrix equation

$$\mathbf{A}\hat{\mathbf{X}} = \hat{\lambda}\hat{\mathbf{X}}, \tag{3.7}$$

where the exact matrix  $\mathbf{A}$  is provided in Appendix A. We then solve (3.7) to find the largest eigenvalue using an Arnoldi iterative scheme using the `scipy.linalg.eigs` solver in Python and refine the accuracy of the largest eigenvalue in the grid spacing  $\Delta\theta$  using variational refinement

$$\hat{\lambda} = \frac{\int_{-\theta_b}^{\theta_b} d\theta \left( c|\hat{\mathbf{X}}|^2 - g \left| \frac{d\hat{\mathbf{X}}}{d\theta} \right|^2 \right)}{\int_{-\theta_b}^{\theta_b} d\theta f|\hat{\mathbf{X}}|^2}, \tag{3.8}$$

where the derivative  $d\hat{\mathbf{X}}/d\theta$  is calculated using a fourth-order accurate finite-difference scheme and the integral is performed using a fourth-order accurate Simpson’s rule (1/3 rule) with `scipy.integrate.simps`.<sup>4</sup> Note that we only solve for and refine the largest eigenvalue of (3.3) and not the entire eigenvalue spectrum.

### 3.3. Properties of the ideal ballooning equation

The ideal ballooning (3.3) is a linear equation that can be written as

$$\mathcal{L}\hat{\mathbf{X}} = \hat{\lambda}\hat{\mathbf{X}}, \tag{3.9}$$

where the linear operator

$$\mathcal{L} \equiv \frac{1}{f} \frac{d}{d\theta} g \frac{d}{d\theta} + \frac{c}{f}, \tag{3.10}$$

and the coefficients  $g, c, f$  are real-valued functions along a field line. Mathematically, the solutions of (3.9) form the basis for the Hilbert space equipped with the following inner product

$$\langle \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2 \rangle = \int_{-\infty}^{\infty} d\theta \hat{\mathbf{X}}_1^* \hat{\mathbf{X}}_2, \tag{3.11}$$

and are square integrable, i.e.  $\langle \hat{\mathbf{X}}, \hat{\mathbf{X}} \rangle < \infty$ . Owing to the self-adjoint nature of ideal MHD (Freidberg 2014), for solutions  $\hat{\mathbf{X}}_1$  and  $\hat{\mathbf{X}}_2$  of (3.9) the operator  $\mathcal{L}$  satisfies the following property:

$$\langle \mathcal{L}\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2 \rangle = \langle \hat{\mathbf{X}}_1, \mathcal{L}\hat{\mathbf{X}}_2 \rangle, \tag{3.12}$$

where we have used the boundary condition  $\lim_{\theta \rightarrow \pm\infty} \hat{\mathbf{X}}_1 = \lim_{\theta \rightarrow \pm\infty} \hat{\mathbf{X}}_2 = 0$ . Using (3.12), one can show that all eigenvalues  $\hat{\lambda}$  (3.1) will be real numbers. Therefore,  $\omega = \pm i\sqrt{\hat{\lambda}}$  will be purely real, an oscillating mode, or purely imaginary, a growing mode. We

<sup>4</sup>Our Python code is freely available at [github.com/rahulgaur104/ideal-ballooning-optimizer](https://github.com/rahulgaur104/ideal-ballooning-optimizer).

refer to oscillating modes as stable and to growing modes as unstable. We use this powerful property in § 4 to formulate an adjoint method, a technique that can speed up the calculation of the gradient of  $\hat{\lambda}_{\max}$  on each flux surface.

### 3.4. Relation to the KBM

In this section, we explain how the ideal ballooning equation is directly related to an important mode of gyrokinetic plasma turbulence known as the KBM. Unlike ideal MHD which is a fluid theory, a single fluid with properties that vary in configuration space, a gyrokinetic model takes into account the distribution of different ion and electron species in both configuration and velocity space. Using a gyrokinetic model, Tang *et al.* (1980) and Aleynikova *et al.* (2018) have shown that for devices with a large aspect ratio, modes with wavenumbers  $k_y \ll 1/\rho_i$ , where  $\rho_i = eB/m_i$  is the ion gyroradius and  $m_i$  is the ion mass, the gyrokinetic model can be reduced to the ideal ballooning equation with corrections that depend on  $k_y \rho_i$ :

$$\frac{d}{d\theta} g \frac{d\hat{X}}{d\theta} + c\hat{X} = \omega(\omega_{*,s} - \omega)f\hat{X}, \quad (3.13)$$

where

$$\omega_{*,s} = c_0 k_y \rho_i, \quad (3.14)$$

where  $c_0$  is a constant on a flux surface and  $k_y \rho_i$  is the normalised wavenumber of the mode in the long wavelength limit, i.e.  $k_y \rho_i \rightarrow 0$ , we recover the ideal ballooning equation exactly. This means that an ideal ballooning unstable mode is also unstable to the KBM. In fact, the KBM is an ideal ballooning mode with kinetic effects.

Using a simple mixing-length argument, one can qualitatively argue that the turbulence heat flux diffusion is

$$D \sim \frac{\text{Im}(\omega)}{k_y^2}, \quad (3.15)$$

where  $\text{Im}(\omega)$  is the imaginary part of  $\omega$  also called the growth rate. This implies that low-wavenumber turbulence has the highest rate of diffusion and leads to poor plasma confinement. Hence, even if ideal ballooning unstable modes do not lead to disruption in stellarators, they could lead to a large heat flux transport through the KBM channel. Since calculating KBM growth rates using a microstability code is expensive, one can use ideal ballooning stability as a necessary condition for KBM stability to optimise against the KBM.<sup>5</sup> For tokamaks, this is one of the fundamental ideas currently used in the EPED code (Snyder *et al.* 2007) to predict the plasma pressure profile in the pedestal region.

In summary, in this section, we have explained the mathematical, physical and numerical methods used to solve the ideal ballooning equation. We have also explained the self-adjoint property of the ideal ballooning equation and the crucial link between the ideal ballooning mode and KBM. In the next section, we use the self-adjoint property of the ideal ballooning equation to outline and test the adjoint method.

## 4. Developing an adjoint method

In this section, we derive and explain the process of calculating the gradients of the ideal ballooning eigenvalue  $\hat{\lambda}$  using an adjoint method. Using this method, we find the maximum eigenvalue  $\hat{\lambda}_{\max}$  on each flux surface. We then elucidate how this technique is

<sup>5</sup>Note that ideal ballooning stability is a necessary but not sufficient condition for KBM stability. A mode can be stable against the ideal ballooning mode, but unstable against the KBM.

faster than the conventional gradient-based method and illustrate this by plotting gradients from a typical optimisation run and calculating the speed-up.

To find  $\hat{\lambda}_{\max}$  on each flux surface, we need the gradient of the eigenvalue of a general function  $\mathcal{H}$  such that  $\mathcal{H}$  is maximised if and only if  $\hat{\lambda} = \hat{\lambda}_{\max}$ . Mathematically, this problem can be defined as follows

$$\max \mathcal{H}(\hat{\lambda}, \tilde{\mathbf{p}}, \hat{\mathbf{p}}), \quad \text{s.t. } \mathcal{G}(\hat{\lambda}, \hat{X}, \tilde{\mathbf{p}}, \hat{\mathbf{p}}) \equiv \mathcal{L}\hat{X} - \hat{\lambda}\hat{X} = 0, \tag{4.1}$$

where  $\hat{\lambda}$  is the eigenvalue,  $\hat{X}$  is the eigenfunction,  $\tilde{\mathbf{p}}$  is the state vector that contains all the equilibrium parameters such as the boundary shape and  $\hat{\mathbf{p}} = (\alpha_t, \theta_0)$  is a vector that contains the parameters of the ideal ballooning equation,  $\mathcal{H}$  is an objective function and  $\mathcal{G}$  is the ideal ballooning operator. To maximise  $\mathcal{H}$  on a flux surface for a given equilibrium, i.e. for a fixed  $\tilde{\mathbf{p}}$ , we need the gradient

$$\frac{d\mathcal{H}}{d\hat{\mathbf{p}}} = \left. \frac{\partial \mathcal{H}}{\partial \hat{\lambda}} \right|_{\hat{\mathbf{p}}} \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} + \left. \frac{\partial \mathcal{H}}{\partial \hat{\mathbf{p}}} \right|_{\hat{\lambda}}. \tag{4.2}$$

The most expensive term to calculate in (4.2) is the gradient of the eigenvalue  $\lambda$ .<sup>6</sup> To obtain that, we expand the equation  $d\mathcal{G}/d\hat{\mathbf{p}} = 0$  at a constant  $\tilde{\mathbf{p}}$

$$-\frac{\partial \mathcal{G}}{\partial \hat{\lambda}} \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} = \frac{\partial \mathcal{G}}{\partial \hat{X}} \frac{\partial \hat{X}}{\partial \hat{\mathbf{p}}} + \left. \frac{\partial \mathcal{G}}{\partial \hat{\mathbf{p}}} \right|_{\hat{X}, \hat{\lambda}}. \tag{4.3}$$

This equation can be explicitly written with the help of (3.10)

$$\frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} \hat{X} = (\mathcal{L} - \hat{\lambda}) \frac{\partial \hat{X}}{\partial \hat{\mathbf{p}}} + \frac{\partial \mathcal{L}}{\partial \hat{\mathbf{p}}} \hat{X}. \tag{4.4}$$

To simplify (4.4) further, we multiply it by the eigenfunction  $\hat{X}^*$  and integrate it throughout the domain  $\theta \in [-\theta_b, \theta_b]$ . Upon doing that, we use the self-adjoint property (3.12) and work through the algebra (given in Appendix B) to obtain the adjoint relation

$$\frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} = \frac{\int_{-\theta_b}^{\theta_b} d\theta \left( \frac{\partial c}{\partial \hat{\mathbf{p}}} |\hat{X}|^2 - \left. \frac{\partial g}{\partial \hat{\mathbf{p}}} \right| \frac{d\hat{X}}{d\theta} \right)^2 - \hat{\lambda} \frac{\partial f}{\partial \hat{\mathbf{p}}} |\hat{X}|^2}{\int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2}. \tag{4.5}$$

To obtain  $\partial \hat{\lambda} / \partial \hat{\mathbf{p}}$  using a central-finite-difference scheme, one has to solve the ideal ballooning equation  $2n_{\hat{\mathbf{p}}} = 4$  times at each optimisation step, where  $n_{\hat{\mathbf{p}}}$  is the length of the vector  $\hat{\mathbf{p}}$ . However, using the adjoint relation (4.5), we only have to solve it once per optimisation step, as long as we can calculate the gradients of geometry-related quantities  $g$ ,  $c$  and  $f$  four times. Since gradients of  $g$ ,  $c$  and  $f$  can be calculated roughly two orders of magnitude faster than solving the ideal ballooning equation, we speed up the gradient calculation by approximately a factor of four. Therefore, we use the adjoint relation (4.5)

<sup>6</sup>Note that the derivative of an eigenvalue is only well-defined when the eigenvalue is isolated. Optimisation problems with stringent penalty terms can lead the optimiser to points with multiplicity (Lewis & Overton 1996).

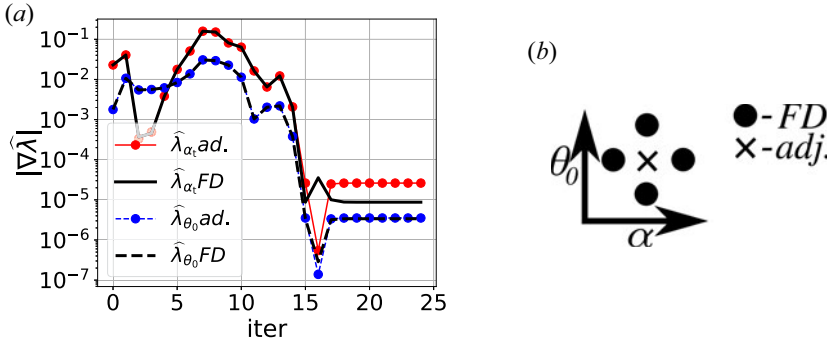


FIGURE 4. (a) Comparison between the gradients of eigenvalue  $\hat{\lambda}_{\alpha_i} = \partial \hat{\lambda} / \partial \alpha_i$  and  $\hat{\lambda}_{\theta_0} = \partial \hat{\lambda} / \partial \theta_0$  obtained using a finite-difference scheme against those obtained using an adjoint method. The quantity iter is the number of iterations taken by the local optimiser on a flux surface before finding  $\hat{\lambda}_{\max}$ . The gradients match well for around four orders of magnitude. The discrepancy between the adjoint and finite difference  $\hat{\lambda}_{\alpha_i}$  is due to the finite resolution of the VMEC run. (b) Different grids used to calculate the gradient of the eigenvalue  $\hat{\lambda}$  on a flux surface. A finite-difference scheme requires four points, whereas an adjoint method only requires one point. This gives us a four times speed-up.

to calculate the gradient of  $\hat{\lambda}$ . In this study, we choose

$$\mathcal{H}(\hat{\lambda}, \tilde{\mathbf{p}}, \hat{\mathbf{p}}) = \hat{\lambda}. \tag{4.6}$$

Applying this fact to (4.2) and using (4.5),

$$\frac{d\mathcal{H}}{d\hat{\mathbf{p}}} = \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} = \frac{\int_{-\theta_b}^{\theta_b} d\theta \left( \frac{\partial c}{\partial \hat{\mathbf{p}}} |\hat{X}|^2 - \frac{\partial g}{\partial \hat{\mathbf{p}}} \left| \frac{d\hat{X}}{d\theta} \right|^2 - \hat{\lambda} \frac{\partial f}{\partial \hat{\mathbf{p}}} |\hat{X}|^2 \right)}{\int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2}. \tag{4.7}$$

This relation gives us the derivative of the ballooning objective function at any point  $\hat{\mathbf{p}} = (\alpha_i, \theta_0)$ . Note that in this work we use (4.7) combined with a local, gradient-based algorithm to find  $\hat{\lambda}_{\max}$  on a flux surface. However, this method is valid and, under appropriate conditions, can be extended to the equilibrium parameters  $\tilde{\mathbf{p}}$ . The details of the extended adjoint method are given in Appendix B.1.

In the next section, we present comparison between adjoint and the regular finite-difference-based methods and present tests comparing the gradients of  $\hat{\lambda}$  from these two methods.

#### 4.1. Comparing adjoint gradients with a finite-difference method

In this section, we first compare the values of the gradients of  $\hat{\lambda}_{\max}$  from the adjoint method with their values obtained using a central-finite-difference method. We take a typical optimisation loop in the modified NCSX case and show a gradient comparison in figure 4. As you can see, the gradients obtained using an adjoint method match well with the gradients obtained with a finite-difference method.

To show the computational speedup, we also compare the time taken by an adjoint method with the regular finite-difference-based method. For the 30 iterations shown in

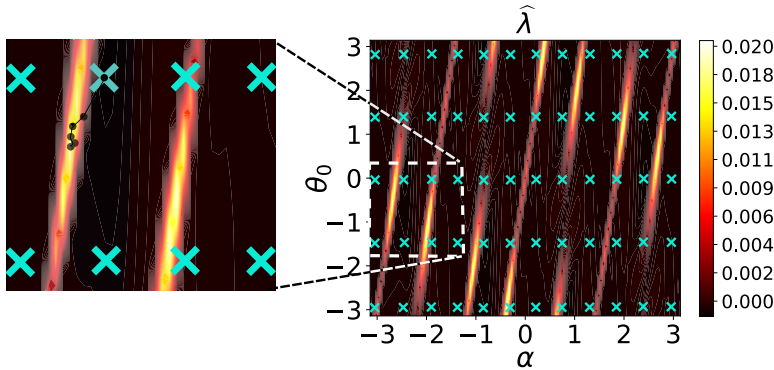


FIGURE 5. Process of finding the globally maximum eigenvalue  $\hat{\lambda}_{\max}$  on the flux surface  $s = 0.8$  of the NCSX equilibrium. We start by first finding the maximum  $\hat{\lambda}$  on a discrete grid (marked by crosses) of  $\alpha_t$  and  $\theta_0$ . From the discrete maximum  $\hat{\lambda}$ , we search for the global maximum eigenvalue using a local optimiser. In the inset, we show the approximate path taken by the optimiser to reach  $\hat{\lambda}_{\max}$ . This process is repeated for all the flux surfaces.

figure 4(a), the adjoint method was about 4 times faster than the finite-difference method. Indeed, the most expensive part of the gradient calculation is the ballooning solver. As shown in the illustration in figure 4(b), for a second-order accurate central-difference scheme, an adjoint method only needs a single call to the ballooning solver, whereas the finite-difference solver needs four. In principle, a speed-up factor of up to four should be possible.

### 5. Details of the optimisation process

In this section, we explain the optimisation process to find equilibria that are stable against the ideal ballooning mode. In § 5.1, we describe the process of using an adjoint method to find the maximum eigenvalue  $\hat{\lambda}_{\max}$  on each flux surface. In § 5.2, we explain how we use  $\hat{\lambda}_{\max}$  and other penalty terms to construct the overall objective function  $\mathcal{F}$ . Finally, in § 5.3, we explain how we search for ballooning stable equilibria while minimising  $\mathcal{F}$  using the SIMSOPT framework.

#### 5.1. Finding $\hat{\lambda}_{\max}$ on each flux surface

To calculate the ballooning objective function we find the maximum  $\hat{\lambda}$  on each flux surface. To do that, we solve (3.3) on several flux surfaces, multiple field lines on each surface and numerous values of  $\theta_0$  on each field line. We calculate  $\hat{\lambda}_{\max}$  on  $n_s = 16$  flux surfaces for each equilibrium. For the 3D equilibria, we scan  $n_{\alpha_t} = 42$  field lines in the range  $\alpha_t = [-\pi, \pi)$ . Since all field lines are identical in a 2D axisymmetric equilibrium, we scan only one field line, i.e.  $n_{\alpha_t} = 1$  for the 2D equilibrium. On each field line, we scan  $n_{\theta_0} = 21$  values of  $\theta_0$  in the range  $\theta_0 = [-\pi/2, \pi/2)$ . The maximum  $\hat{\lambda}$  from a coarse grid scan gives us a value close to the global maximum. From the maximum  $\hat{\lambda}$  of the coarse grid, we launch a local gradient-based optimiser to find the global maximum eigenvalue. This process is explained using the illustration figure 5. Using this process, we obtain  $\hat{\lambda}_{\max}$  as a function of the normalised toroidal flux  $s$ . Figure 6 shows the plot of  $\hat{\lambda}_{\max}$  against  $s$  for the three chosen equilibria.

For each new equilibrium, on all  $n_s = 16$  flux surfaces, the local optimiser takes an average of 20 iterations to find  $\hat{\lambda}_{\max}$ . Moreover, as described in figure 4(b), at each step, the

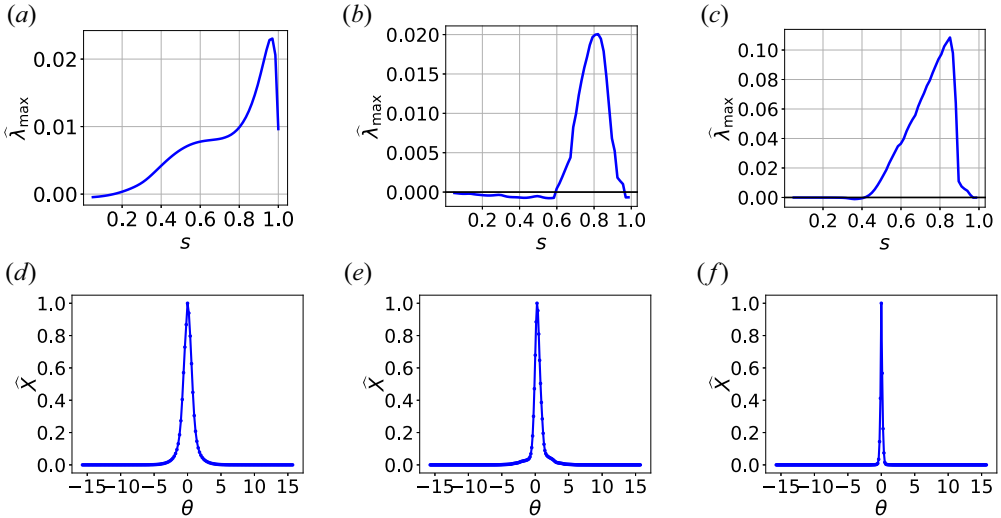


FIGURE 6. Plots of  $\hat{\lambda}_{\max}$  against the normalised toroidal flux  $s$  for the three chosen equilibria in (a–c) and the eigenfunctions  $\hat{X}$  at the maximum  $\hat{\lambda}_{\max}$  for each equilibrium in the (d–f): (a) DIII-D  $\hat{\lambda}_{\max}$ , (b) NCSX  $\hat{\lambda}_{\max}$ , (c) Henneberg-QA  $\hat{\lambda}_{\max}$ , (d) DIII-D  $\hat{X}(s = 0.95)$ , (e) NCSX  $\hat{X}(s = 0.8)$  and (f) Henneberg-QA  $\hat{X}(s = 0.84)$ . The decay of the eigenfunction could be a result of the Anderson localisation of ballooning modes, as discussed by Redi *et al.* (2002).

use of a finite-difference method requires four evaluations of the eigenvalue  $\hat{\lambda}$ . This means that on average, we have to call the ballooning solver 1280 ( $16 \times 20 \times 4$ ) times. This is a computationally expensive step that we speed up using our adjoint-based method.

### 5.2. Finding ballooning-stable equilibria

Once we have found  $\hat{\lambda}_{\max}$ , we seek an equilibrium stable to the ideal ballooning mode by minimising  $\hat{\lambda}_{\max}$  on each flux surface. To do so, we need to define an objective function that depends on  $\hat{\lambda}_{\max}$  such that minimising the objective function should allow us to achieve a stable equilibrium. Moreover, during optimisation, once a flux surface is stabilised against the ideal ballooning mode, our objective function should ignore that particular surface. This would be useful as we do not want to penalise a stable equilibrium. To this end, we design the following ideal ballooning objective function

$$f_{\text{ball}} = \sum_{j=1}^{\text{ns}} \text{ReLU}(\hat{\lambda}_{\max,j} - \hat{\lambda}_{\text{th},j}), \tag{5.1}$$

where ns is the total number of surfaces and

$$\text{ReLU}(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ x, & \text{if } x > 0, \end{cases} \tag{5.2}$$

is the rectified linear unit operator: an operator that sets all the non-positive values to zero and  $\hat{\lambda}_{\text{th},j}$  is the threshold below which we declare a surface ideal ballooning stable. The value  $\hat{\lambda}_j = 0$  on the  $j$ th surface implies marginal stability but we choose  $\hat{\lambda}_{\text{th},j} = 0.0001$  to ensure that all surfaces are slightly away from marginal ideal ballooning stability. An equilibrium is ideal ballooning stable if  $f_{\text{ball}} = 0$ .

It is also important to prevent the optimiser from minimising  $f_{\text{ball}}$  in a trivial way. For example, for 2D equilibria, going to a larger aspect ratio value stabilises the ideal ballooning mode. For 3D equilibria, the optimiser can sometimes reduce the minor radius, which, for a fixed toroidal flux, causes the magnetic field to increase. This lowers the overall  $\beta$  and consequently the unstable curvature drive term. Similarly, if we allow rotational transforms to increase freely, the optimiser can sometimes create large gradients of  $\iota$ , generating large currents which are suboptimal. To avoid achieving such trivial solutions and uninteresting equilibria, we add a combination of the following penalty terms to the optimiser:

- (i)  $f_{\text{asp}} = (A - A_0)$  to penalise any deviation from the aspect ratio of the initial equilibrium;
- (ii)  $f_{\text{minr}} = (a_N - a_{N0})$  to penalise any deviation from the minor radius of the initial equilibrium;
- (iii)  $f_{\langle B \rangle} = (\langle B \rangle - \langle B \rangle_0)$  to penalise any deviation of the volume-averaged magnetic field from its value in the initial equilibrium;
- (iv)  $f_{R_c} = \int d\theta \text{ReLU}(-R_c)$ , where  $R_c$  is the radius of curvature of the boundary; this term penalises any boundary shapes that are curved into the plasma

Using the ballooning objective function (5.1) and one or more of the penalty terms described previously, we can get the overall objective function  $\mathcal{F}$ . Given a vector of input parameters  $\mathbf{p}$ , our goal is to solve

$$\min_{\mathbf{p}} \mathcal{F}(\mathbf{p}), \quad \text{s.t. } f_{\text{ball}} = 0. \tag{5.3}$$

We achieve this with the help of the SIMSOPT (Landreman *et al.* 2021) package. The implementation details of the optimisation are described in the next subsection.

### 5.3. Optimisation with the SIMSOPT package

In this subsection, we discuss the implementation-related details of an adjoint ballooning solver with the SIMSOPT (Landreman *et al.* 2021) package. First, we briefly explain how an optimisation problem can be solved using SIMSOPT. Next, we go into the details of how we solve the ideal ballooning optimisation and how the use of an adjoint method can speed up this process.

The SIMSOPT package is a optimisation framework containing a suite of codes that, along with the VMEC code, have been used to optimise 3D equilibria for various properties such as energetic fast-particle confinement, quasisymmetry, simpler magnetic coil geometry and neoclassical transport. The user specifies the input parameters (also referred to as degrees of freedom (DoFs)) and the objective function  $\mathcal{F}$  and SIMSOPT can perform a gradient-based or gradient-free nonlinear least squares optimisation.

As an example, let us construct an optimisation problem to stabilise an equilibrium while penalising the change in the aspect ratio and the minor radius of the boundary

$$\mathcal{F} = f_{\text{asp}}^2 + f_{\text{minr}}^2 + f_{\text{ball}}^2. \tag{5.4}$$

To do so, we use gradient-based optimisation in SIMSOPT, where one calculates  $\partial\mathcal{F}/\partial\mathbf{p}$  to update the parameter vector at the  $i$ th iteration,  $\mathbf{p}_i$  as

$$\mathbf{p}_{i+1} = \mathbf{f} \left( \mathbf{p}_i, \frac{\partial\mathcal{F}}{\partial\mathbf{p}_i} \right). \tag{5.5}$$

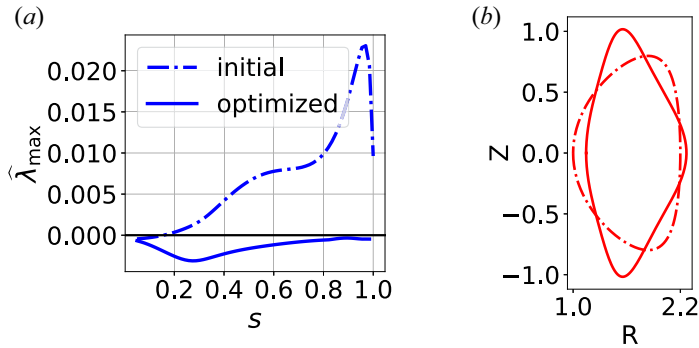


FIGURE 7. (a) Maximum eigenvalue  $\hat{\lambda}_{\max}$  of the initial and optimised DIII-D-like equilibrium. The optimised equilibrium is stable. (b) Boundary shape of the initial and final equilibria. Note the negative triangularity of the initial equilibrium and the positive triangularity of the optimised equilibrium.

This is done until the optimiser reaches a local minima, i.e. a region in the parameter space where  $\partial\mathcal{F}/\partial\mathbf{p} = 0$  or the relative change in the gradient is small enough. Typically, one has to evaluate the gradient of  $\mathcal{F}$  hundreds of times during an optimisation loop before finding a local minimum. In this study, evaluating  $f_{\text{ball}}$  is the most expensive step. Because the speed of the optimisation is limited by the rate at which we can compute  $f_{\text{ball}}$ , we have used an adjoint method to calculate  $\hat{\lambda}_{\max}$  which gives us  $f_{\text{ball}}$ .

## 6. Results

In this section, we present the results of our study. In § 6.1, we compare the initial and optimised 2D axisymmetric equilibrium. In §§ 6.2 and 6.3, we do the same for the modified NCSX and modified Henneberg-QA equilibria, respectively. We also compare the values of relevant physical quantities in the initial and optimised equilibria.

### 6.1. Stabilising the DIII-D-like equilibrium

For the 2D axisymmetric case, we start with a high- $\beta$  equilibrium with a negative triangularity boundary. Due to axisymmetry, the 2D boundary does not depend on the toroidal angle  $\zeta$ . In other words, only  $n = 0$  are needed in (2.7) to represent the boundary of a 2D equilibrium. Therefore, the number of modes needed to specify a 2D boundary is much lower than that for a typical 3D boundary. In this problem, we pick the six largest Fourier modes as our DoFs:  $\hat{R}_b(0, 1)$ ,  $\hat{R}_b(0, 2)$ ,  $\hat{R}_b(0, 3)$ ,  $\hat{Z}_b(0, 1)$ ,  $\hat{Z}_b(0, 3)$ ,  $\hat{Z}_b(0, 5)$ . The full objective function is

$$\mathcal{F} = f_{\text{asp}}^2 + f_{R_c}^2 + f_{\text{minr}}^2 + f_{(B)}^2 + f_{\text{ball}}^2, \tag{6.1}$$

where all terms except  $f_{\text{ball}}$  are penalty terms to prevent the optimiser from producing a trivial solution. After this, we start with the negative triangularity equilibrium described in § 2.2 and run SIMSOPT to find multiple equilibria that are stable against the ideal ballooning mode, i.e. equilibria with  $f_{\text{ball}} = 0$ . We present one of the optimised equilibria in figure 7. We also compare the values of equilibrium-dependent quantities in table 4.

We observe that the optimiser is moving toward a positive triangularity equilibrium, indicating that, for the similar values of the relevant parameters (given in table 4)



Equilibrium	$\beta_{ax}(\%)$	$\langle\beta\rangle(\%)$	$j_{\zeta}(\text{MA})$	$\langle B\rangle(\text{T})$	$\bar{t}$	$A$	$a_N(m)$	$B_N(T)$	$\psi_{\text{LCFS}}(Tm^2)$
Initial	14.0	7.6	0.616	0.677	0.568	2.42	0.68	0.679	1.0
Optimised	13.7	7.3	0.728	0.686	0.568	2.42	0.68	0.681	1.0

TABLE 4. Comparison between relevant physical quantities of the initial and optimised DIII-D equilibrium.

$\hat{R}_b(n, m)$	$\hat{Z}_b(n, m)$
([1, 4], 0)	([1, 4], 0)
([-3, 3], 1)	([-3, 3], 1)
([-3, 3], 2)	([-3, 3], 2)
([-2, 2], 3)	([-2, 2], 3)
([-2, 2], 4)	([-2, 2], 4)
([-2, 2], 5)	([-2, 2], 5)
([-1, 1], 6)	([-1, 1], 6)

TABLE 5. Boundary shape DoFs for the NCSX case.

positive triangularity high- $\beta$  equilibria are more stable than their negative triangularity counterparts. Our findings are consistent with recent observations by Davies, Dickinson & Wilson (2022), Nelson, Paz-Soldan & Saarelma (2022) and Saarelma *et al.* (2021) that negative triangularity equilibria are more unstable against the ideal ballooning mode compared with positive triangularity equilibria. This behaviour prevents the formation of a steep pressure gradient, which limits the operational beta value of negative triangularity equilibria but also prevents the occurrence of edge localised modes.

### 6.2. Stabilising the NCSX equilibrium

The first 3D equilibrium we optimise is an unstable NCSX equilibrium. Since the boundary has a 3D shape, we have to use both toroidal and poloidal modes to change its shape. For this demonstration, we choose 72 boundary modes listed in table 5 as DoFs. In table 5,  $[i, j]$  denotes all integers between  $i$  and  $j$  (including  $i$  and  $j$ ). The total number of DoFs, 72, is much larger than the axisymmetric case. After choosing these DoFs, we choose the following objective function

$$\mathcal{F} = 0.5f_{\text{minr}}^2 + 0.5f_{(B)}^2 + 50f_{\text{ball}}^2, \tag{6.2}$$

We run SIMSOPT with this configuration to obtain multiple equilibria with  $f_{\text{ball}} = 0$ . We have plotted a comparison of one of these equilibria with the initial equilibrium in figure 8.

The optimiser successfully stabilised the equilibrium through boundary shaping. From the change in the boundary shape, it is evident that the ideal ballooning stability is sensitive to the boundary shape for the NCSX equilibrium. We present a comparison of the important equilibrium-dependent quantities for the initial and optimised equilibria in table 6.

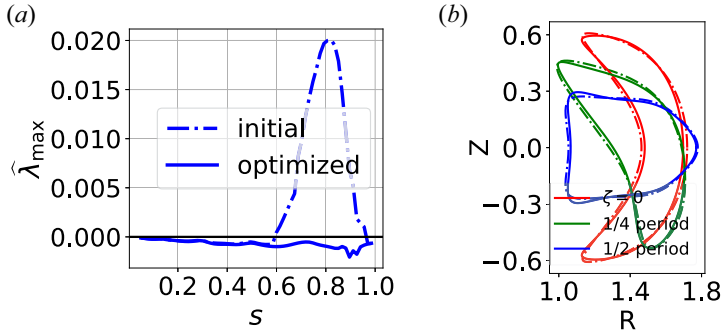


FIGURE 8. (a) Maximum eigenvalue  $\hat{\lambda}_{\max}$  of the initial and optimised modified NCSX equilibrium. (b) Boundary shape of the initial and final equilibria at three different values of the toroidal angle  $\zeta$ . The dotted curves correspond to the initial cross sections whereas the solid curves are the final cross sections. Note how sensitive the maximum eigenvalue is to the boundary shape.

Equilibrium	$\beta_{\text{ax}}(\%)$	$\langle\beta\rangle(\%)$	$j_{\zeta}(\text{MA})$	$\langle B\rangle(\text{T})$	$\bar{r}$	$A$	$a_N(m)$	$B_N(\text{T})$	$\psi_{\text{LCFS}}(\text{Tm}^2)$
Initial	10.9	5.9	0.305	1.596	0.68	4.36	0.325	1.54	0.514
Optimised	9.8	5.3	0.361	1.686	0.68	4.49	0.317	1.62	0.514

TABLE 6. Comparison between relevant physical quantities of the initial and optimised NCSX equilibrium.

$\hat{R}_b(n, m)$	$\hat{Z}_b(n, m)$
$([1, 4], 0)$	$([1, 4], 0)$
$([-3, 3], 1)$	$([-3, 3], 1)$
$([-3, 3], 2)$	$([-3, 3], 2)$
$([-2, 2], 3)$	$([-2, 2], 3)$
$([-2, 2], 4)$	$([-2, 2], 4)$
$([-2, 2], 5)$	$([-2, 2], 5)$
$([-1, 1], 6)$	$([-1, 1], 6)$

TABLE 7. Boundary shape DoFs for the modified Henneberg-QA case. We have chosen the exact same coefficients as the NCSX case.

### 6.3. Stabilising the modified Henneberg-QA equilibrium

As a final example, we present a modified Henneberg-QA equilibrium. In this example, we allow the boundary coefficients given in table 7 to change freely.

For this problem, we choose the following objective function

$$\mathcal{F} = 0.1f_{\text{minr}}^2 + 0.2f_{(B)}^2 + 50f_{\text{ball}}^2. \tag{6.3}$$

After choosing the DoFs and the objective function, we run SIMSOPT and obtain multiple stable equilibria. We compare the stabilised equilibrium with the initial, unstable equilibrium in figure 9 and the equilibrium-dependent quantities in table 8. We can see

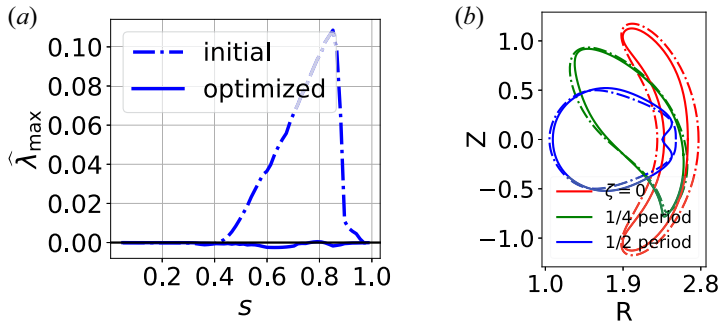


FIGURE 9. (a) Maximum eigenvalue  $\hat{\lambda}_{\max}$  of the initial and optimised modified Henneberg-QA equilibrium. (b) Boundary shape of the initial and final equilibria at three different positions of the toroidal angle  $\zeta$ . The dotted curves correspond to the initial cross sections whereas the solid curves are the final cross sections.

Equilibrium	$\beta_{\text{ax}}(\%)$	$\langle \beta \rangle(\%)$	$j_{\zeta}(\text{MA})$	$\langle B \rangle(T)$	$\bar{i}$	$A$	$a_N(m)$	$B_N(T)$	$\psi_{\text{LCFS}}(Tm^2)$
Initial	6.7	3.56	0.204	2.5	0.263	3.37	0.601	2.35	2.677
Optimised	4.7	2.57	0.204	2.9	0.263	3.58	0.562	2.69	2.677

TABLE 8. Comparison between relevant physical quantities of the initial and optimised modified Henneberg-QA equilibrium.

in figure 9(a) that the boundary shape deforms in a way to reduce the curvature on the outboard side, stabilising the ideal ballooning mode.

### 7. Summary and conclusions

We began this work by briefly explaining the various curvilinear coordinate systems that we used to fully define a general 3D ideal MHD equilibrium. In § 2, we generated three different equilibria, one axisymmetric 2D and two 3D, and briefly described their properties.

Upon generating the equilibria, in § 3, we have provided a physical description of the infinite- $n$  ideal ballooning mode and explain the numerical methods used to calculate the maximum eigenvalues on a given flux surface. Using these numerical methods, we evaluated the stability of the three equilibria against the ideal ballooning mode. We also described the self-adjoint property of the ideal ballooning mode and its relation to the KBM.

Using the self-adjoint property explained in § 3, we have developed an adjoint method in § 4 and explained how to use it to speed up the calculation of the maximum ballooning eigenvalue  $\hat{\lambda}_{\max}$  on each surface. To demonstrate the efficiency and accuracy of the adjoint method, we have also presented a comparison of gradients between an adjoint method and a simple finite-difference scheme. We found that the adjoint method is up to four times faster than the finite-difference scheme.

In § 5, we describe the details of the overall optimisation process and how we accomplish that using the SIMSOPT code. After implementing the optimisation, we

presented the results in § 6. We have presented the specific details of the objective function and the DoFs for each equilibrium and stabilised the initial, ideal ballooning unstable equilibria. We have briefly described the physical mechanisms that stabilise the ideal ballooning mode.

This work presents many avenues for future research. A key step forward is to extend our technique to include all equilibrium-dependent parameters  $\tilde{\boldsymbol{p}}$  as explained in [Appendix B.1](#). One could also use the exact same method to optimise stellarators and tokamaks against low- $n$ , unstable ideal MHD modes, as explained in [Appendix B.2](#). Since solving for low- $n$  ideal MHD codes is much more computationally expensive, the advantage of using an adjoint method would be even greater. Another possible direction would be to use an adjoint method to get derivatives of the ballooning eigenvalue with respect to the plasma shape. Finally, one could use the ideal ballooning optimiser as a tool that could help optimise an equilibrium against KBMs and help us look for equilibrium-dependent proxies for MHD or kinetic instabilities.

### Acknowledgements

We thank Professor D. Bindel, Dr A. Bader and Dr B. Faber for helpful discussions and encouragement. R.G. thanks R. Conlin for his valuable suggestions at the APS-DPP 2022 conference.

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

### Funding

This work was supported by the US Department of Energy, Office of Science and Office of Fusion Energy Sciences under award numbers DE-SC0018429 and DE-FG02-93ER54197. This research used resources from the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility and the Stellar cluster at Princeton University. This material is based on work supported by the National Science Foundation Graduate Research Fellowship under grant number DGE-1650441.

### Declaration of interests

The authors report no conflict of interest.

### Data availability statement

The Zenodo link containing all the equilibria and optimisation code used in this paper is freely available at <https://zenodo.org/record/8092287>.

### Appendix A. The discretised ideal ballooning equation

After applying the boundary conditions  $\hat{X}_0 = \hat{X}_N = 0$  to (3.5), we can rewrite the ballooning equation as

$$(\mathbf{A} - \lambda \mathbf{I})\hat{X} = 0, \tag{A1}$$

with  $\mathbb{I}$  being the identity matrix and the asymmetric tridiagonal matrix  $\mathbf{A}$  with the following form:

$$\begin{bmatrix}
 \frac{g_{1/2} + g_{3/2} - (\Delta\theta)^2 c_1}{(\Delta\theta)^2 f_1} & -\frac{g_{3/2}}{(\Delta\theta)^2 f_1} & 0 & 0 & \dots \\
 -\frac{g_{3/2}}{(\Delta\theta)^2 f_2} & \frac{g_{3/2} + g_{5/2} - (\Delta\theta)^2 c_2}{(\Delta\theta)^2 f_2} & -\frac{g_{5/2}}{(\Delta\theta)^2 f_2} & 0 & \dots \\
 & & & \ddots & \\
 & & & & \ddots \\
 0 & 0 & 0 & 0 & \dots \\
 0 & 0 & 0 & 0 & \dots \\
 & 0 & & 0 & \\
 & 0 & & 0 & \\
 & \vdots & & \vdots & \\
 & \vdots & & \vdots & \\
 \frac{g_{N-5/2} + g_{N-3/2} - (\Delta\theta)^2 c_{N-2}}{h^2 f_{N-2}} & & -\frac{g_{N-3/2}}{(\Delta\theta)^2 f_{N-2}} & & \\
 -\frac{g_{N-3/2}}{(\Delta\theta)^2 f_{N-1}} & & \frac{g_{N-3/2} + g_{N-1/2} - (\Delta\theta)^2 c_{N-1}}{(\Delta\theta)^2 f_{N-1}} & & 
 \end{bmatrix}. \tag{A2}$$

**Appendix B. Details of the adjoint ideal ballooning calculation**

In this appendix, we derive (4.5) starting with (3.3). To do that, we use the self-adjoint property of ideal MHD given in (3.12) as well as the Dirichlet boundary conditions satisfied by the eigenfunction,

$$\hat{X}(\theta = \pm\theta_b) = \hat{X}^*(\theta = \pm\theta_b) = 0. \tag{B1}$$

Defining  $\mathbf{p} = (\tilde{\mathbf{p}}, \hat{\mathbf{p}})$  as the union of all the parameters of the problem, we start by taking the derivative  $\partial/\partial\mathbf{p}$  of (3.3),

$$\frac{d}{d\theta} \frac{\partial g}{\partial \mathbf{p}} \frac{d\hat{X}}{d\theta} + \frac{d}{d\theta} g \frac{d}{d\theta} \frac{\partial \hat{X}}{\partial \mathbf{p}} + \frac{\partial c}{\partial \mathbf{p}} \hat{X} + c \frac{\partial \hat{X}}{\partial \mathbf{p}} = \frac{\partial \hat{\lambda}}{\partial \mathbf{p}} f \hat{X} + \hat{\lambda} f \frac{\partial \hat{X}}{\partial \mathbf{p}} + \hat{\lambda} \hat{X} \frac{\partial f}{\partial \mathbf{p}}. \tag{B2}$$

Multiplying with  $\hat{X}^*$  on both sides, integrating throughout the domain,

$$\begin{aligned}
 & \int_{-\theta_b}^{\theta_b} d\theta \hat{X}^* \frac{d}{d\theta} \frac{\partial g}{\partial \mathbf{p}} \frac{d\hat{X}}{d\theta} + \int_{-\theta_b}^{\theta_b} \hat{X}^* \frac{d}{d\theta} g \frac{d}{d\theta} \frac{\partial \hat{X}}{\partial \mathbf{p}} + \int_{-\theta_b}^{\theta_b} d\theta \frac{\partial c}{\partial \mathbf{p}} |\hat{X}|^2 + \int_{-\theta_b}^{\theta_b} d\theta c \hat{X}^* \frac{\partial \hat{X}}{\partial \mathbf{p}} \\
 & = \frac{\partial \hat{\lambda}}{\partial \mathbf{p}} \int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2 + \hat{\lambda} \int_{-\theta_b}^{\theta_b} d\theta \frac{\partial f}{\partial \mathbf{p}} |\hat{X}|^2 + \hat{\lambda} \int_{-\theta_b}^{\theta_b} d\theta f \hat{X}^* \frac{\partial \hat{X}}{\partial \mathbf{p}}. \tag{B3}
 \end{aligned}$$

Using integration by parts, (B1) and rearranging (B3), we can write

$$\int_{-\theta_b}^{\theta_b} d\theta \left( \frac{\partial c}{\partial \mathbf{p}} |\hat{X}|^2 - \frac{\partial g}{\partial \mathbf{p}} \left| \frac{d\hat{X}}{d\theta} \right|^2 - \hat{\lambda} \frac{\partial f}{\partial \mathbf{p}} |\hat{X}|^2 \right) - \frac{\partial \hat{\lambda}}{\partial \mathbf{p}} \int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2 = \int_{-\theta_b}^{\theta_b} d\theta \left( -\frac{d}{d\theta} g \frac{d\hat{X}^*}{d\theta} - c\hat{X}^* + \hat{\lambda} f \hat{X}^* \right) \frac{\partial \hat{X}}{\partial \mathbf{p}}. \tag{B4}$$

Due to the self-adjoint property of ideal MHD, the right-hand side of (B4) is zero. The rest of the equation can be arranged so that

$$\frac{\partial \hat{\lambda}}{\partial \mathbf{p}} = \frac{\int_{-\theta_b}^{\theta_b} d\theta \left( \frac{\partial c}{\partial \mathbf{p}} |\hat{X}|^2 - \frac{\partial g}{\partial \mathbf{p}} \left| \frac{d\hat{X}}{d\theta} \right|^2 - \hat{\lambda} \frac{\partial f}{\partial \mathbf{p}} |\hat{X}|^2 \right)}{\int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2}. \tag{B5}$$

Therefore, to calculate  $\partial \hat{\lambda} / \partial \mathbf{p}$ , we only need the gradients of the geometric coefficients  $g, c, f$ , the eigenfunction  $\hat{X}$  and the eigenvalue  $\hat{\lambda}$  of the ballooning equation; we have to solve the ideal ballooning equation only once. This speeds up the optimisation loop, as it is much faster to obtain the gradient of the geometric coefficients than to solve the ballooning equation multiple times.

B.1. Extending our adjoint-based technique to equilibrium-dependent DoFs

In this paper, we have used an adjoint method to find the maximum eigenvalue  $\hat{\lambda}_{\max}$  with respect to the parameters  $\hat{\mathbf{p}}$  of the ballooning equation. It is possible to extend our method to minimise  $f_{\text{ball}}$  to the equilibrium parameters  $\tilde{\mathbf{p}}$  under the appropriate conditions. We define the problem and find the pertinent conditions in this appendix. We want to find

$$\min f_{\text{ball}}(\hat{\lambda}_{\max}, \tilde{\mathbf{p}}, \hat{\mathbf{p}}), \quad \text{s.t. } \mathcal{G}(\hat{\lambda}, \hat{X}, \tilde{\mathbf{p}}, \hat{\mathbf{p}}) \equiv \mathcal{L}\hat{X} - \hat{\lambda}\hat{X} = 0, \tag{B6}$$

where all symbols are defined in § 4 and the ballooning objective function  $f_{\text{ball}}$  is defined in (5.1). To minimise  $f_{\text{ball}}$  with respect to the equilibrium parameters, we need

$$\frac{df_{\text{ball}}}{d\tilde{\mathbf{p}}} = \left. \frac{\partial f_{\text{ball}}}{\partial \hat{\lambda}_{\max}} \right|_{\tilde{\mathbf{p}}, \hat{\mathbf{p}}} \frac{\partial \hat{\lambda}_{\max}}{\partial \tilde{\mathbf{p}}} + \left. \frac{\partial f_{\text{ball}}}{\partial \hat{\mathbf{p}}} \right|_{\hat{\lambda}_{\max}, \tilde{\mathbf{p}}} \frac{\partial \hat{\mathbf{p}}}{\partial \tilde{\mathbf{p}}} + \left. \frac{\partial f_{\text{ball}}}{\partial \tilde{\mathbf{p}}} \right|_{\hat{\lambda}_{\max}, \hat{\mathbf{p}}}. \tag{B7}$$

The most expensive term to calculate in (B7) is the gradient of the eigenvalue  $\lambda$ . To obtain that, we take the derivative of the operator  $\mathcal{G}$  with respect to  $\lambda$ ,

$$-\frac{\partial \mathcal{G}}{\partial \hat{\lambda}} \frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} = \frac{\partial \mathcal{G}}{\partial \hat{\lambda}} \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} + \frac{\partial \mathcal{G}}{\partial \hat{X}} \frac{\partial \hat{X}}{\partial \hat{\mathbf{p}}} + \frac{\partial \mathcal{G}}{\partial \hat{X}} \frac{\partial \hat{X}}{\partial \tilde{\mathbf{p}}} + \frac{\partial \mathcal{G}}{\partial \hat{\mathbf{p}}} + \frac{\partial \mathcal{G}}{\partial \tilde{\mathbf{p}}}. \tag{B8}$$

We also express  $\hat{\lambda}$  around a point  $\mathbf{p}_0 = (\tilde{\mathbf{p}}_0, \hat{\mathbf{p}}_0)$  in the state space as a Taylor series,

$$\hat{\lambda} = \hat{\lambda}(\mathbf{p}_0) + \frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} \cdot \delta \tilde{\mathbf{p}} + \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} \cdot \delta \hat{\mathbf{p}} + \frac{\partial}{\partial \tilde{\mathbf{p}}} \frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} : \delta \tilde{\mathbf{p}} \delta \tilde{\mathbf{p}} + \frac{\partial}{\partial \hat{\mathbf{p}}} \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} : \delta \hat{\mathbf{p}} \delta \hat{\mathbf{p}} + O(|\delta \mathbf{p}|^3), \tag{B9}$$

and assume that the optimiser takes a step size  $|\delta\mathbf{p}|$  that is smaller than the radius of convergence of Taylor series (B9),

$$\left| \frac{\partial}{\partial \tilde{\mathbf{p}}} \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} : \delta \tilde{\mathbf{p}} \delta \tilde{\mathbf{p}} \right| \ll \left| \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} \cdot \delta \tilde{\mathbf{p}} \right|, \quad \left| \frac{\partial}{\partial \hat{\mathbf{p}}} \frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} : \delta \hat{\mathbf{p}} \delta \hat{\mathbf{p}} \right| \ll \left| \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} \cdot \delta \hat{\mathbf{p}} \right|. \quad (\text{B10})$$

Using (B9) and (B10)

$$\frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} = \frac{\partial \hat{\lambda}}{\partial \hat{\mathbf{p}}} + \frac{\partial \hat{\lambda}}{\partial \tilde{\mathbf{p}}} \cdot \frac{\partial \tilde{\mathbf{p}}}{\partial \hat{\mathbf{p}}}. \quad (\text{B11})$$

Next, we use the fact that  $\partial \hat{\lambda} / \partial \hat{\mathbf{p}} = 0$  at  $\hat{\lambda} = \hat{\lambda}_{\max}$ , and that our choice of  $f_{\text{ball}}$  only explicitly depends on  $\hat{\lambda}_{\max}$ . Using the explicit form of the linear operator from (3.1), we multiply (B8) by  $X^*$  and integrate throughout the domain, to rewrite (B7) as

$$\frac{df_{\text{ball}}}{d\mathbf{p}} = \sum_{j=1}^{\text{ns}} \text{ReLU}'(\hat{\lambda}_{\max,j}) \frac{\int_{-\theta_b}^{\theta_b} d\theta \left( \frac{\partial c}{\partial \mathbf{p}} |\hat{X}|^2 - \frac{\partial g}{\partial \mathbf{p}} \left| \frac{d\hat{X}}{d\theta} \right|^2 - \hat{\lambda} \frac{\partial f}{\partial \mathbf{p}} |\hat{X}|^2 \right)}{\int_{-\theta_b}^{\theta_b} d\theta f |\hat{X}|^2}, \quad (\text{B12})$$

where  $\mathbf{p} = (\tilde{\mathbf{p}}, \hat{\mathbf{p}})$  is the union of all the parameters of the problem and  $\text{ReLU}'^7$  is the derivative of the ReLU operator such that

$$\text{ReLU}'(x) = \begin{cases} 0, & \text{if } x < 0, \\ 1, & \text{if } x > 0. \end{cases} \quad (\text{B13})$$

Calculating the derivative of the geometric coefficients  $g$ ,  $c$  and  $f$  with respect to the equilibrium parameter vector  $\tilde{\mathbf{p}}$  is not straightforward in VMEC and may lack the requisite accuracy for an adjoint method to work. However, an equilibrium solver such as DESC (Dudt & Kolemen 2020; Conlin *et al.* 2022; Panici *et al.* 2022) that is designed to calculate these gradients along with the geometric coefficients accurately may enable us to utilise the full potential of this adjoint-based method. Since the speed up obtained with an adjoint method is linearly proportional to the length of the vector  $\tilde{\mathbf{p}}$ , using (B12) we can, in principle, speed up the calculation of  $df_{\text{ball}}/d\tilde{\mathbf{p}}$  by an order of magnitude for 2D axisymmetric equilibria and by two orders or magnitude for 3D equilibria.

### B.2. Extending our adjoint technique to low- $n$ , ideal MHD solvers

Note that this process can be applied to any ideal MHD eigenvalue solver. For fluctuations that are not confined to a flux surface, one can solve for a fluctuation of the form

$$X = \sum_{m,n} \hat{X}_{m,n}(\psi) e^{i(m\theta - n\zeta)}, \quad (\text{B14})$$

where  $X = \{X_\psi, X_\alpha\}$  are components of fluctuation  $X$  perpendicular to the equilibrium magnetic field line, and  $m$  and  $n$  are the poloidal and toroidal mode numbers, respectively. We solve for  $\hat{X}(\psi)$ , using codes such as ELITE and GATO (Bernard, Helton & Moore

<sup>7</sup>The derivative of the ReLU operator is not well-defined at  $x = 0$ . We may have to replace it with an activation function that is continuous with a well-defined derivative. For example, we could use the logistic function  $1/(1 + e^{-cx})$  with a large positive real number  $c$ .

1981) for axisymmetric equilibria or TERPSICHORE (Anderson *et al.* 1990), CAS-3D (Schwab 1993) for 3D equilibria. For GATO, TERPSICHORE, and CAS-3D, the ideal MHD energy principle is used to solve the matrix equation

$$\mathcal{A}X = \lambda\mathcal{B}X, \quad (\text{B15})$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are real symmetric matrices. Currently, solving such an equation using these codes takes at least a few minutes. For such a problem, we can repeat the process explained at the beginning of this appendix to obtain the gradient,

$$\frac{\partial\lambda}{\partial\mathbf{p}} = X^T \left( \frac{\partial\mathcal{A}}{\partial\mathbf{p}} - \lambda \frac{\partial\mathcal{B}}{\partial\mathbf{p}} \right) X / X^T \mathcal{B}X, \quad (\text{B16})$$

for all modes. Equation (B16) is similar to the Hellman–Feynman theorem (Hellmann 1933; Feynman 1939). For axisymmetric equilibria, combining gradient information with fast equilibrium solvers such as EFIT (Lao *et al.* 1985) can help mitigate real-time disruption. One could also couple this adjoint approach with an optimiser to find low- $n$ , ideal MHD stable equilibria.

#### REFERENCES

- ALEYNIKOVA, K., ZOCCO, A., XANTHOPOULOS, P., HELANDER, P. & NÜHRENBERG, C. 2018 Kinetic ballooning modes in tokamaks and stellarators. *J. Plasma Phys.* **84**, 745840602.
- ANDERSON, D.V., COOPER, W.A., GRUBER, R., MERAZZI, S. & SCHWENN, U. 1990 TERPSICHORE: a three-dimensional ideal magnetohydrodynamic stability program. In *Scientific Computing on Supercomputers II* (ed. J.T. Devreese & E.V. Camp), vol. 159, pp. 159–174. Plenum Press, New York.
- BAALRUD, S., FERRARO, N., GARRISON, L., HOWARD, N., KURANZ, C., SARFF, J., SCIME, E. & SOLOMON, W. 2020 A community plan for fusion energy and discovery plasma sciences. Report of the 2019–2020 American Physical Society Division of Plasma Physics Community Planning Process.
- BERNARD, L.C., HELTON, F.J. & MOORE, R.W. 1981 GATO: an MHD stability code for axisymmetric plasmas with internal separatrices. *Comput. Phys. Commun.* **24**, 377–380.
- BERNARD, L.C. & MOORE, R.W. 1981 Systematic optimization of tokamaks for ideal magnetohydrodynamic stability. *Phys. Rev. Lett.* **46**, 1286.
- BOOZER, A.H. 1983 Transport and isomorphic equilibria. *Phys. Fluids* **26**, 496–499.
- CONLIN, R., DUDT, D.W., PANICI, D. & KOLEMEN, E. 2022 The DESC stellarator code suite part II: Perturbation and continuation methods. *J. Plasma Phys.* **89**, 955890305.
- CONNOR, J.W., HASTIE, R.J. & TAYLOR, J.B. 1979 High mode number stability of an axisymmetric toroidal plasma. *Proc. R. Soc. Lond. A* **365**, 1.
- DAVIES, R., DICKINSON, D. & WILSON, H.R. 2022 Kinetic ballooning modes as a constraint on plasma triangularity in commercial spherical tokamaks. *Plasma Phys. Control. Fusion*. **64**, 105001.
- DEWAR, R.L. & GLASSER, A.H. 1983 Ballooning mode spectrum in general toroidal systems. *Phys. Fluids* **26**, 3038.
- D'HAESELEER, W.D., HITCHON, W.N.G., CALLEN, J.D. & SHOHET, J.L. 2012 *Flux Coordinates and Magnetic Field Structure: A Guide to a Fundamental Tool of Plasma Theory*. Springer Science & Business Media.
- DUDT, D.W. & KOLEMEN, E. 2020 DESC: a stellarator equilibrium solver. *Phys. Plasmas* **27**, 102513.
- FEYNMAN, R.P. 1939 Forces in molecules. *Phys. Rev.* **56**, 340.
- FREIDBERG, J.P. 2014 *Ideal MHD*. Cambridge University Press.
- FU, G.Y., ISAEV, M., KU, L., MIKHAILOV, M., REDDI, M.H., SANCHEZ, R., SUBBOTIN, A., COOPER, W.A., HIRSHMAN, S.P., MONTICELLO, D.A., *et al.* 2007 Ideal magnetohydrodynamic stability of the NCSX. *Fusion Sci. Technol.* **51**, 218–231.



- GARREN, D.A. & BOOZER, A.H. 1991 Magnetic field strength of toroidal plasma equilibria. *Phys. Fluids B* **3**, 2805–2821.
- GATES, D.A., BOOZER, A.H., BROWN, T., BRESLAU, J., CURRELI, D., LANDREMAN, M., LAZERSON, S.A., LORE, J., MYNICK, H., NEILSON, G.H., *et al.* 2017 Recent advances in stellarator optimization. *Nucl. Fusion* **57**, 126064.
- GAUR, R., ABEL, I.G., DICKINSON, D. & DORLAND, W.D. 2023 Microstability of  $\beta \sim 1$  tokamak equilibria. *J. Plasma Phys.* **89**, 905890112.
- GILES, M.B. & PIERCE, N.A. 2000 An introduction to the adjoint approach to design. *Flow Turbul. Combust.* **65**, 393–415.
- GRAD, H. & RUBIN, H. 1958 Proceedings of the Second United Nations International Conference on the Peaceful uses of Atomic Energy, p. 400. Geneva : 1958: UN.
- HELLMANN, H. 1933 Zur rolle der kinetischen elektronenenergie für die zwischenatomaren kräfte. *Z. Phys.* **85** (3).
- 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**, 026014.
- HIRSHMAN, S.P. & WHITSON, J.C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26**, 3553.
- KRUSKAL, M.D. & KULSRUD, R.M. 1958 Equilibrium of a magnetically confined plasma in a toroid. *Phys. Fluids* **1**, 265.
- LANDREMAN, M., MEDASANI, B., WECHSUNG, F., GIULIANI, A., JORGE, R. & ZHU, C. 2021 SIMSOPT: a flexible framework for stellarator optimization. *J. Open Source Softw.* **6**, 3525.
- LAO, L.L., JOHN, H.S., STAMBAUGH, R.D., KELLMAN, A.G. & PFEIFFER, W. 1985 Reconstruction of current profile parameters and plasma shapes in tokamaks. *Nucl. Fusion* **25**, 1611.
- LEWIS, A.S. & OVERTON, M.L. 1996 Eigenvalue optimization. *Acta Numer.* **5**, 149–190.
- MARINONI, A., AUSTIN, M.E., HYATT, A.W., WALKER, M.L., CANDY, J., CHRYSTAL, C., LASNIER, C.J., MCKEE, G.R., ODSTRČIL, T., PETTY, C.C., *et al.* 2019 H-mode grade confinement in L-mode edge plasmas at negative triangularity on DIII-D. *Phys. Plasmas* **26**, 042515.
- MCKINNEY, I.J., PUESCHEL, M.J., FABER, B.J., HEGNA, C.C., ISHIZAWA, A. & TERRY, P.W. 2021 Kinetic-ballooning-mode turbulence in low-average-magnetic-shear equilibria. *J. Plasma Phys.* **87**, 905870311.
- MILLER, R.L. & MOORE, R.W. 1979 Shape optimization of tokamak plasmas to localized magnetohydrodynamic modes. *Phys. Rev. Lett.* **43**, 765.
- NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, MEDICINE, *et al.* 2021 *Bringing fusion to the U.S. grid.*
- NELSON, A.O., PAZ-SOLDAN, C. & SAARELMA, S. 2022 Prospects for H-mode inhibition in negative triangularity tokamak reactor plasmas. *Nucl. Fusion* **62**, 096020.
- PANICI, D., CONLIN, R., DUFT, D.W. & KOLEMEN, E. 2022 The DESC stellarator code suite part I: Quick and accurate equilibria computations. *J. Plasma Phys.* **89**, 955890303.
- PAUL, E.J., LANDREMAN, M. & ANTONSEN, T. 2021 Gradient-based optimization of 3D MHD equilibria. *J. Plasma Phys.* **87**, 905870214.
- REDI, M.H., JOHNSON, J.L., KLASKY, S., CANIK, J., DEWAR, R.L. & COOPER, W.A. 2002 Anderson localization of ballooning modes, quantum chaos and the stability of compact quasi-axisymmetric stellarators. *Phys. Plasmas* **9**, 1990–1996.
- SAARELMA, S., AUSTIN, M.E., KNOLKER, M., MARINONI, A., PAZ-SOLDAN, C., SCHMITZ, L. & SNYDER, P.B. 2021 Ballooning instability preventing the H-mode access in plasmas with negative triangularity shape on the DIII-D tokamak. *Plasma Phys. Control. Fusion* **63**, 105006.
- SANCHEZ, R., HIRSHMAN, S.P., WARE, A.S., BERRY, L.A. & SPONG, D.A. 2000a Ballooning stability optimization of low-aspect-ratio stellarators. *Plasma Phys. Control. Fusion* **42**, 641.
- SANCHEZ, R., HIRSHMAN, S.P., WHITSON, J.C. & WARE, A.S. 2000b COBRA: an optimized code for fast analysis of ideal ballooning stability of three-dimensional magnetic equilibria. *J. Comput. Phys.* **161**, 576.
- SCHWAB, C. 1993 Ideal magnetohydrodynamics: global mode analysis of three-dimensional plasma configurations. *Phys. Fluids B* **5**, 3195–3206.

- SHAFRANOV, V.D. 1957 On equilibrium magnetohydrodynamic configurations. *Zh. Eksp. Teor. Fiz.* **33**, 710.
- SNYDER, P.B., BURRELL, K.H., WILSON, H.R., CHU, M.S., FENSTERMACHER, M.E., LEONARD, A.W., MOYER, R.A., OSBORNE, T.H., UMANSKY, M., WEST, W.P., *et al.* 2007 Stability and dynamics of the edge pedestal in the low collisionality regime: physics mechanisms for steady-state ELM-free operation. *Nucl. Fusion* **47**, 961.
- TANG, W.M., CONNOR, J.W. & HASTIE, R.J. 1980 Kinetic-ballooning-mode theory in general geometry. *Nucl. Fusion* **20**, 1439.
- 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**, A237.