J. Plasma Phys. (2026), *vol*. 92, E72 © The Author(s), 2026. Published by Cambridge University Press.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited. doi:10.1017/S0022377826101652

Spectrally accurate, reverse-mode differentiable bounce-averaging algorithm and its applications

Kaya Unalmis[1], Rahul Gaur[2], Rory Conlin[3], Dario Panici[2] and Egemen Kolemen[2,4,5]

[1]Electrical and Computer Engineering, Princeton University, Princeton, New Jersey, USA
[2]Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey, USA
[3]Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland, USA
[4]Andlinger Center for Energy and the Environment, Princeton University, Princeton, New Jersey, USA
[5]Princeton Plasma Physics Laboratory, Princeton, New Jersey, USA

Corresponding authors: Kaya Unalmis, kunalmis@alumni.princeton.edu; Egemen Kolemen, ekolemen@princeton.edu

(Received 28 October 2025 UTC; revision received 16 February 2026 UTC; accepted 31 March 2026 UTC)

We present a fast, spectrally (exponentially) accurate, automatically differentiable bounce-averaging algorithm that is used to simplify kinetic models. Using this algorithm, implemented in the `DESC` stellarator optimisation suite, we can perform efficient optimisation of many objectives to improve stellarator performance, such as the effective ripple ϵ_{eff} metric for the neoclassical transport coefficient in the low collisionality regime and proxies for energetic particle confinement. For the first time, we optimise a finite-beta stellarator to directly reduce neoclassical ripple transport using reverse-mode differentiation. This ensures the computational cost of differentiation is independent of the number of controllable parameters.

Key words: fusion plasma

1. Introduction

Stellarators, first conceived by Spitzer Jr (1958), represent a distinct approach to magnetic confinement fusion that offers unique advantages over tokamaks. These toroidal devices achieve plasma confinement through external magnetic fields rather than through plasma current, providing greater design flexibility and operational stability. The absence of a continuous toroidal symmetry allows for magnetic field optimisation through boundary shaping, which helps minimise the net toroidal current and thereby avoids current-driven plasma disruptions that plague tokamak operation (Helander 2014).

The design of optimal stellarator configurations is a complex optimisation problem involving hundreds of degrees of freedom. Traditional optimisation approaches have evolved significantly over the past few decades. `VMEC` (Variational Moments

Equilibrium Code), developed by Hirshman & Whitson (1983), served as the foundation for stellarator optimisation. Building upon VMEC, several frameworks have emerged: STELLOPT (Spong *et al*. 1998; Lazerson *et al*. 2020), which implements a suite of physics-based optimisation criteria, ROSE (Drevlak *et al*. 2018), which focuses on coil optimisation and engineering constraints, and more recently, SIMSOPT (Landreman *et al*. 2021). In DESC (Dudt & Kolemen 2020; Conlin *et al*. 2023; Dudt *et al*. 2023; Panici *et al*. 2023; Dudt *et al*. 2025), unlike previous optimisers, it is not necessary to re-solve the magnetohydrodynamic (MHD) force balance equation at each optimisation step (Conlin *et al*. 2024). Additional objectives that depend on equilibrium force balance can be optimised simultaneously while ensuring force balance.

Traditional approaches to stellarator optimisation rely on finite difference techniques. Such techniques are low-order accurate and hinder the ability of the optimiser to find good solutions. Furthermore, finite difference techniques require computing the objective function multiple times to estimate the derivative in the direction of each optimisable parameter; this is infeasible when the number of parameters is large. In contrast, adjoint methods and automatic differentiation (Sapienza *et al*. 2025) can compute derivatives with respect to all optimisable parameters with a computational cost that is comparable to the cost of a single objective function evaluation. These methods have greatly improved the ability to solve inverse design problems for stellarators (Antonsen, Paul & Landreman 2019; Paul *et al*. 2019, 2021; Dudt *et al*. 2025).

We present an automatically differentiable bounce-averaging algorithm that is used to simplify kinetic models such as drift and gyrokinetics that study phenomena at time scales longer than the bounce orbit time. This algorithm has been implemented in the DESC optimisation suite. Previous works (Matsuda & Stewart 1986; Nemov *et al*. 1999, 2008; Kernbichler *et al*. 2016; Petrov & Harvey 2016; Lazerson *et al*. 2020; Velasco *et al*. 2020, 2021*b*) have used bounce-averaging to accelerate the solution of Fokker–Planck equations. However, such works are incompatible with automatic differentiation. Moreover, their computation is discretised with lower-order accuracy than in this work. This work enables fast optimisation to improve stellarator performance with exponential accuracy.

In § 2, we present an application of bounce-averaging to compute the neoclassical transport coefficient in the low collisionality regime where the transport coefficients increase with decreasing collision frequency. Then, in § 3, we describe our numerical methods to compute bounce-averaged objectives for optimisation. In § 4, we apply this framework to optimise against neoclassical transport. In § 5, we conclude this work and explain how it can be extended.

2. Neoclassical model of plasma

Our study concerns configurations where magnetic field lines lie on closed, nested toroidal surfaces, known as flux surfaces. We label these surfaces with their enclosed toroidal flux ψ. Such a divergence-free magnetic field may be written in the Clebsch form (D'haeseleer *et al*. 2012), showing that curves of constant (ψ, α) trace field lines,

$$\boldsymbol{B} = \nabla\psi \times \nabla\alpha. \tag{2.1}$$

The dynamics of a magnetised hot plasma differ significantly from that of an unmagnetised fluid. Unlike isotropic hard-sphere collisions that govern the behaviour of an uncharged fluid, a plasma behaves differently in directions perpendicular and

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

parallel to the magnetic field lines because of Coulomb collisions. In magnetised plasmas, particles traverse helical trajectories, gyrating around magnetic field lines and drifting across them. The classical transport model assumes a simplistic view of particle collisions and does not adequately incorporate the effects of these drifts. To properly account for these drifts, trapped and passing particles, and the magnetic geometry, we use the neoclassical transport theory.

There are three fundamental length and time scales relevant to magnetised plasmas. The time scales correspond to the particle transit frequency $v_{\mathrm{th,s}}/L_B$, where $v_{\mathrm{th,s}} = (2T_s/m_s)^{1/2}$ is the thermal speed, the Coulomb collision frequency $\nu_{ss'} \propto T^{-3/2}$ and the gyration frequency $\Omega_s = Z_s e|B|/(m_s c)$, where s, s' are the species of interest, $Z_s e$ is the charge, m_s is the mass and c is the speed of light. For each time scale, the corresponding length scales are the gradient scale length of the magnetic field L_B, the mean free path λ_{mfp} and the gyroradius $r_{\mathrm{gyro},s} = v_{\mathrm{th},s}/\Omega_s$. In a magnetised plasma,

$$\nu_{ss'} \sim \frac{v_{\mathrm{th,s}}}{L_B} \ll \Omega_s, \tag{2.2}$$

$$\lambda_{\mathrm{mfp}} \sim L_B \gg r_{\mathrm{gyro},s}. \tag{2.3}$$

Using a random walk estimate, we can calculate the classical heat transport coefficient in the perpendicular direction as $D_\perp \sim \nu_{ss'} r^2_{\mathrm{gyro},s} \sim T^{-1/2}$ (Helander & Sigmar 2005), whereas, using neoclassical theory, we have $\Delta r \sim r_{\mathrm{gyro},s}|B|/|B|_{\mathrm{poloidal}}$ with $|B|$ and $|B|_{\mathrm{poloidal}}$ given by the total and poloidal magnetic field strength, respectively. The transport coefficient is then $D_\perp \sim \nu_{ss'} r^2_{\mathrm{gyro},s}|B|^2/|B|^2_{\mathrm{poloidal}} \sim T^{-1/2}|B|^2/|B|^2_{\mathrm{poloidal}}$. Note that the ratio $|B|/|B|_{\mathrm{poloidal}}$ strongly depends on the magnetic field geometry and significantly affects the regime of neoclassical transport.

Magnetised plasmas can be weakly or strongly collisional as defined by the collisionality $\nu_\star = L_B/\lambda_{\mathrm{mfp}}$. In a strongly collisional plasma, particles undergo frequent collisions without covering significant distances along a magnetic field line, i.e. $\nu_\star \gg 1$. Conversely, in a weakly collisional plasma, particles can traverse a significant distance before colliding, i.e. $\nu_\star \ll 1$. Stellarator plasmas in practical applications tend to be weakly collisional.

Based on the stellarator geometry, the weak collisionality regime can be further partitioned into the banana or plateau regime according to the reciprocal of the aspect ratio $\epsilon \sim \iota^{-1}|B|_{\mathrm{poloidal}}/|B|$, where ι is the rotational transform (Helander 2014). Most stellarators lie in the regime where $\nu_\star \ll \epsilon^{3/2}$. This categorisation is illustrated in figure 1.

The standard neoclassical theory first enabled computation of the neoclassical transport coefficients in the low collisionality regime for a simplified model of the magnetic field. This analysis was later extended to stellarator magnetic fields (Kovrizhnykh 1984; Ochs 2025). The following section outlines this process in one regime of interest to stellarator equilibrium optimisation.

2.1. *Effective ripple*

In the low collision limit $\nu_\star \ll \epsilon^{3/2}$, the neoclassical model studies the plasma distribution f determined by a simplified Boltzmann equation known as the drift-kinetic equation. For a particle with mass m, let $\boldsymbol{v}_\parallel$ and $\boldsymbol{v}_\perp$ be the velocity parallel and orthogonal, respectively, to the unit vector magnetic field $\boldsymbol{b}$. In the drift-kinetic equation, the velocity space may be parametrised with three independent coordinates: the total energy E, the magnetic moment $\mu = m|v_\perp|^2/(2|B|)$ and the gyrophase angle.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

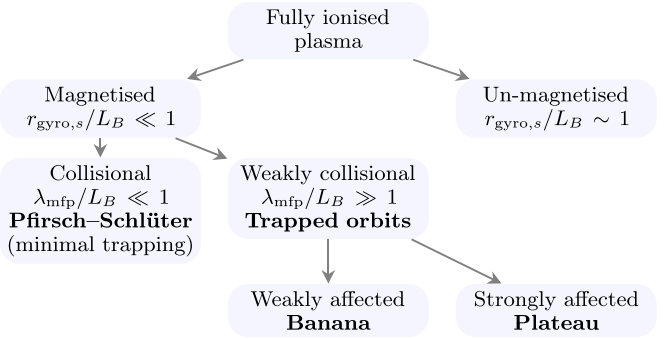


FIGURE 1. A schematic categorising neoclassical transport.

In this treatment, the equation is averaged over the gyrophase angle. We seek a steady-state solution and linearise the distribution of guiding centres $f = f_0 + f_1$ into a background f_0 that is Maxwellian in velocity and a higher order correction f_1. Thus, the background is parametrised in velocity space with E and the higher order correction with (E, μ). The linearised drift-kinetic equation reduces to the following partial differential equation (PDE) (Abel *et al*. 2013):

$$\mathcal{C}[f] = \boldsymbol{v}_{\mathrm{D}s} \cdot \nabla f_0 + |v_\parallel| \boldsymbol{b} \cdot \nabla f_1, \tag{2.4}$$

$$\boldsymbol{v}_{\mathrm{D}s} = \frac{|v_\parallel|^2}{\Omega_s} \boldsymbol{b} \times (\boldsymbol{b} \cdot \nabla \boldsymbol{b}) + \frac{|v_\perp|^2}{2\Omega_s} \frac{\boldsymbol{b} \times \nabla |B|}{|B|} + \boldsymbol{v}_{\mathrm{Baños}}. \tag{2.5}$$

The electric field is neglected in this section as we focus on the $1/\nu$ collisionality regime.

To reduce neoclassical transport, one may minimise the radial particle flux density,

$$\Gamma = \int \mathrm{d}^3 \boldsymbol{v} \; f_1 \boldsymbol{v}_{\mathrm{D}s} \cdot \nabla \psi. \tag{2.6}$$

Appendix B shows a derivation to extract a dimensionless quantity Γ_0 (2.8) for the optimisation objective which is proportional to the flux surface average $\langle \Gamma \rangle$ (B13).

$$\langle \Gamma \rangle = -\Gamma_0 \frac{2^{3/2} \pi c^2}{3^2 e^2 m^{3/2}} \int_0^\infty \mathrm{d}E \; \frac{E^{5/2}}{\nu} \frac{\partial f_0}{\partial \psi} \tag{2.7}$$

$$\Gamma_0 = \left(\int_0^{2\pi} \mathrm{d}\alpha \int_{|B|_{\min}}^{|B|_{\max}} \frac{\mathrm{d}\varrho}{\varrho^3} \sum_w \frac{I_1^2}{I_2} \right) \left(\int_0^{2\pi} \mathrm{d}\alpha \int_{\zeta_1}^{\zeta_2} \frac{\mathrm{d}\zeta}{\boldsymbol{B} \cdot \nabla \zeta} \right)^{-1} \tag{2.8}$$

$$I_1(\psi, \alpha, \varrho, w) = \int_{\zeta_1(w)}^{\zeta_2(w)} \frac{\mathrm{d}\zeta}{\boldsymbol{B} \cdot \nabla \zeta} (1 - |B|/\varrho)^{1/2} (4\varrho/|B| - 1) \, |\nabla \psi| \kappa_{\mathrm{G}} \tag{2.9}$$

$$I_2(\psi, \alpha, \varrho, w) = \int_{\zeta_1(w)}^{\zeta_2(w)} \frac{\mathrm{d}\zeta}{\boldsymbol{B} \cdot \nabla \zeta} (1 - |B|/\varrho)^{1/2} \tag{2.10}$$

The quantity κ_{G} is the geodesic curvature of the field line (B5) and the velocity space coordinate ϱ is defined as

$$\varrho = E/\mu. \tag{2.11}$$

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

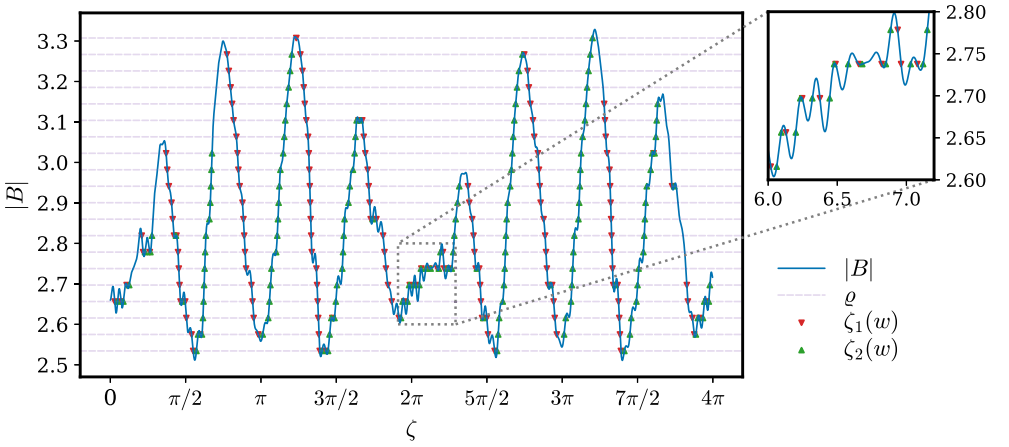


FIGURE 2. Bounce points within $(\zeta_1, \zeta_2) = (0, 4\pi)$ on the field line $(\psi, \alpha) = (\psi_{\text{plasma boundary}}, 0)$ for a mesh of ϱ values on a W7-X stellarator. For a given ϱ marked by a horizontal line, $|v_\parallel| = 0$ at the bounce points marked by triangles. The plasma distribution vanishes in the hypograph of $|B|$.

The index w labels the well with boundaries $\zeta_1(w)$ and $\zeta_2(w)$ where a bouncing particle is trapped. These boundaries are referred to as bounce points. Only the particles which are trapped within the interval $[\zeta_1, \zeta_2]$ are considered so that $\zeta_1 \leqslant \min_w \zeta_1(w)$ and $\max_w \zeta_2(w) \leqslant \zeta_2$. An illustration is shown in figure 2.

In axisymmetric configurations, integration along the field line for a single poloidal transit between two global maxima of $|B|$ is sufficient for convergence of Γ_0. On irrational magnetic surfaces, it is sufficient to integrate along a single field line (D'haeseleer *et al*. 2012, § 4.9). On a rational or near-rational surface in non-axisymmetric configurations, it is necessary to integrate along multiple field lines until the surface is covered sufficiently.

The effective ripple modulation amplitude ϵ_{eff} is related to Γ_0 as follows:

$$\epsilon_{\text{eff}}^{3/2} = \frac{\pi}{2^{7/2}} \frac{(B_0 R_0)^2}{\langle |\nabla \psi| \rangle^2} \Gamma_0, \tag{2.12}$$

where B_0 is a background magnetic field typically chosen to be $|B|_{\max}$ and R_0 is the average major radius of the stellarator. A reason ϵ_{eff} is preferred to Γ_0 as an optimisation objective is that the latter vanishes near the magnetic axis, which reduces the ability to distinguish between good and bad configurations. Since ϵ_{eff} depends only on geometry, reducing it by varying the plasma boundary can reduce the radial neoclassical loss of trapped particles.

3. Algorithm

We briefly describe a few fundamental parts of our algorithm. Section 1 discusses the bounce integral in more detail. In § 2, we describe efficient quadrature used for these integrals. Section 3 discusses our inverse method to solve the ideal MHD equation. In §§ 4–6, we describe our approach to obtain data along field lines.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

To motivate the need for an efficient algorithm, let us estimate the computational cost of bounce-averaging with a blunt approach to the computation. After discretising to N_s field lines, where each field line is followed over N_w magnetic wells for each of N_ϱ pitch angles, there will be $\mathcal{O}(N_s N_w N_\varrho)$ bounce integrals. With N_q quadrature points each, the integrand is evaluated at $\mathcal{O}(N_s N_w N_\varrho N_q) \sim 10^8$ points. Furthermore, the path of integration is unknown *a priori* because the field lines move during optimisation. Finding the position of the field lines on a known grid may involve N_i Newton iterations for each point. With N_c spectral coefficients used to approximate the map on which that root-finding is done, the cost grows to $\mathcal{O}(N_c N_i N_s N_w N_\varrho N_q)$. Moreover, the memory required to reverse-mode differentiate the objective grows linearly with the problem size.

3.1. *Bounce integral*

The bounce integral of x may be written as a time-weighted integral over the trajectory of the particle along its bounce orbit (Mackenbach *et al*. 2023*b*, § 2). Since the dynamics parallel to the field lines dominate, the particle trajectory may be approximated to follow field lines by parametrising time t as the distance along a field-line following coordinate. Since the magnetic moment is an adiabatic invariant for which the gyro-average of $\mathrm{d}\mu/\mathrm{d}t$ is approximately zero, the pitch angle of a bouncing particle stays nearly constant over the time scale to complete bounce orbits when energy is conserved. Labelling the boundaries $\zeta_1(w)$ and $\zeta_2(w)$ of magnetic well w where the parallel velocity vanishes, using the streamline property in curvilinear coordinates,

$$|v_\parallel| \mathrm{d}t = \frac{\mathrm{d}\zeta}{\boldsymbol{b} \cdot \nabla \zeta}, \tag{3.1}$$

and $|v_\parallel|^2 = (2E/m)(1 - |B|/\varrho)$, then allows writing the integral as follows:

$$\overline{x}(\psi, \alpha, \varrho, w) = \frac{m^{1/2}}{(2E)^{1/2}} \int_{\zeta_1(w)}^{\zeta_2(w)} \frac{\mathrm{d}\zeta}{\boldsymbol{b} \cdot \nabla \zeta} (1 - |B|/\varrho)^{-1/2} x. \tag{3.2}$$

More generally, integrals between bounce points involve a map g, smooth in ζ, weighted by a map with behaviour matching $|v_\parallel|^\eta$ near the bounce points,

$$\int_{\zeta_1(w)}^{\zeta_2(w)} \mathrm{d}\zeta \, |v_\parallel|^\eta g(\psi, \alpha, \zeta, \varrho), \quad \eta \in \{-1, 1\}. \tag{3.3}$$

3.2. *Quadrature*

Gaussian quadrature approximates $\int_{-1}^{1} \mathrm{d}\zeta \, \varsigma g(\zeta) \approx \sum_{i=1}^{N_q} \sigma_i g(\zeta_i)$ for some weight ς positive and continuous in the interior by approximating g with its Hermite interpolation polynomial and choosing σ_i, ζ_i to avoid evaluating the derivative (Süli & Mayers 2003). For integrable (3.3), we can construct such a quadrature for ς matching the non-polynomial behaviour of $|v_\parallel|^\eta$ or, more generally, employ a change of variable whose Jacobian vanishes slowly near singularities such that the integrand can then be approximated by a low-degree polynomial. In the latter approach, the transformation should also be mild enough to prevent unnecessary clustering of quadrature points that would increase the condition number of the problem.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

Our transformation for bounce integrals defines z such that $a_1(w, a_2[z]) = \zeta$,

$$a_1 : \begin{cases} \mathbb{N} \times [-1, 1] \to \mathbb{R}, \\ w, z \mapsto (z+1)[\zeta_2(w) - \zeta_1(w)]/2 + \zeta_1(w), \end{cases} \tag{3.4}$$

$$a_2 : \begin{cases} [-1, 1] \to [-1, 1], \\ z \mapsto \sin(\pi z/2), \end{cases} \tag{3.5}$$

$$\int_{\zeta_1(w)}^{\zeta_2(w)} \mathrm{d}\zeta \ |v_\parallel|^\eta g(\zeta) \approx \frac{\zeta_2(w) - \zeta_1(w)}{2} \sum_{i=1}^{N_q} \sigma_i |v_\parallel|^\eta g(a_1(w, a_2[z_i])). \tag{3.6}$$

When neither bounce point is on a local maximum of the potential, the midpoint scheme in z (3.7) is exponentially accurate,[1]

$$\sigma_i = \pi \sin\left[\pi(2i-1)/(2N_q)\right]/N_q, \quad a_2[z_i] = \cos\left[\pi(2i-1)/(2N_q)\right]. \tag{3.7}$$

If, in addition, $\eta = 1$, then (3.8) yields faster convergence,

$$\sigma_i = \pi \sin\left[\pi i/(N_q+1)\right]/(N_q+1), \quad a_2[z_i] = \cos\left[\pi i/(N_q+1)\right]. \tag{3.8}$$

In general, Gauss–Legendre quadrature in z is exponentially accurate if (3.3) is integrable. Figures 3, 4, 5 and 6 illustrate the convergence.

It is often of interest to integrate a nonlinear combination of bounce integrals over ϱ. Such integrands can be non-smooth in ϱ due to the logarithmic divergence (Calvo *et al*. 2017, § 4) of (3.2) as ϱ approaches the value of $|B|$ at any local maxima within or at the bounce points. The Alpert (1999) quadratures are high-order accurate for these singularities.

We compare the following quadratures in their ability to compute elliptic integrals (3.9), (3.10), which are similar to bounce integrals in a simple stellarator geometry.

(i) Midpoint scheme.
(ii) Simpson's 1/3 in the interior completed by a midpoint scheme.
(iii) Double exponential (DE) tanh − sinh.
(iv) Implicitly weighted Gauss–Chebyshev of the first (GC_1) (3.7) or second kind (GC_2) (3.8).
(v) Gauss–Legendre (GL_1) or Gauss–Lobatto–Legendre (GL_2) each composed with the sin transformation in (3.5). Compared with item (iv), this quadrature offers more resolution near the boundary and less in the interior.

$$\int_{-\arcsin k}^{\arcsin k} \mathrm{d}\zeta \ (k^2 - \sin^2\zeta)^{-1/2} \underset{\text{(C11)}}{=} 2K(k) \tag{3.9}$$

$$\int_{-\arcsin k}^{\arcsin k} \mathrm{d}\zeta \ (k^2 - \sin^2\zeta)^{1/2} \underset{\text{(C12)}}{=} 2E(k) + 2(k^2-1)K(k) \tag{3.10}$$

To further benchmark the quadratures in a magnetic field with ripples, we show two more cases that model particles trapped in the following wells in figures 5 and 6.

[1] By the Euler–Maclaurin expansion, the order of convergence is at least the number of continuous derivatives along the field line of $|B|$ and the surface geometry. Their analyticity then implies exponential convergence (Trefethen & Weideman 2014). Although these quantities can be non-analytic (Cary & Shasharina 1997), the equilibria produced by DESC yield flux surfaces on which their truncated Fourier series converge fast and these analytic series generate the data.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

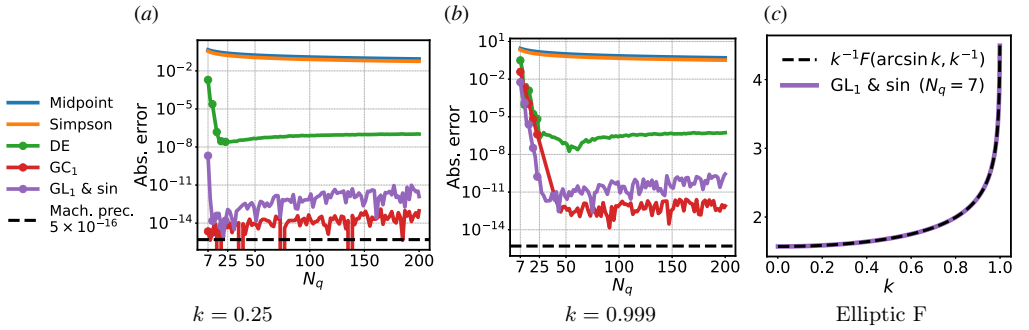


FIGURE 3. Convergence of quadratures for (3.9). GC_1 and GL_1 show spectral convergence whereas midpoint, Simpson and double exponential quadratures hit floating point plateaus early.

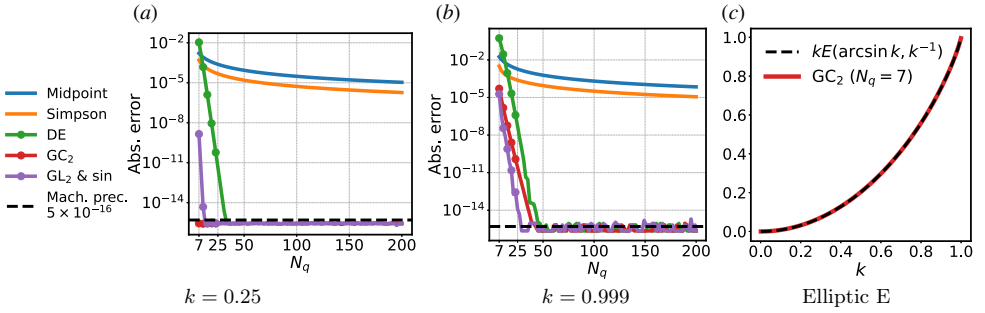


FIGURE 4. Convergence of quadratures for (3.10). GC_2 and GL_2 show spectral convergence.

3.3. *Inverse method*

In this section, we briefly discuss how we find stellarator equilibria. At static equilibrium, the ideal MHD equations that approximate the behaviour of the plasma reduce to

$$\boldsymbol{B} \cdot \nabla \boldsymbol{B} = \nabla \left(p + |B|^2/2\right), \tag{3.11}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{3.12}$$

which describes a balance between the plasma pressure p, magnetic field pressure $|B|^2$ and the effect of field line curvature $\boldsymbol{B} \cdot \nabla \boldsymbol{B}$. We solve the ideal MHD equation using an inverse method. The computational domain is a solid torus in curvilinear flux coordinates (ρ, θ, ζ), where $\rho = (\psi/\psi_{\text{plasma boundary}})^{1/2}$ and (θ, ζ) are angles on a doubly-periodic surface. Here, Λ and ω are to be determined maps that relate the angles (θ, ζ) that parametrise a given plasma boundary to the Clebsch angle,

$$\alpha = \theta + \Lambda - \iota(\zeta + \omega). \tag{3.13}$$

Fourier–Zernike series parametrised in flux coordinates (ρ, θ, ζ) are chosen to approximate Λ, ω and the map to a cylindrical coordinate system (R, ϕ, Z) in the lab frame,

$$R(\rho, \theta, \zeta)\hat{\boldsymbol{R}}(\phi) + Z(\rho, \theta, \zeta)\hat{\boldsymbol{Z}}. \tag{3.14}$$

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

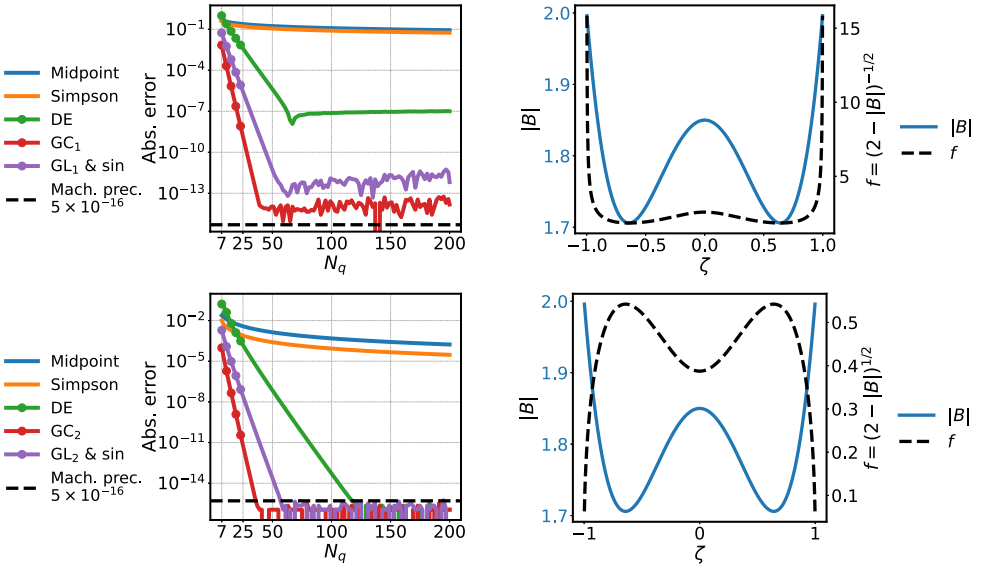


FIGURE 5. Convergence of quadratures for the quantity labeled by f.

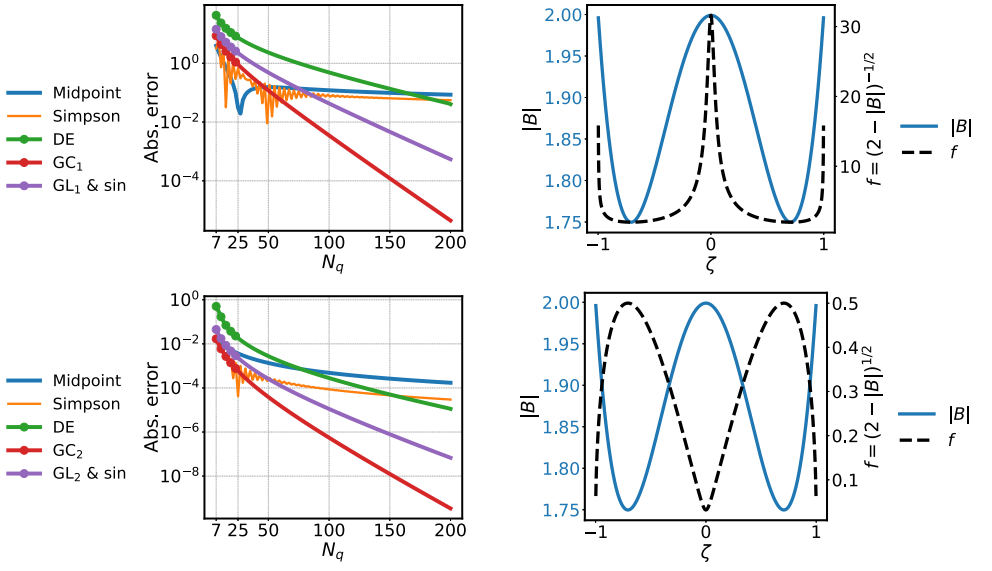


FIGURE 6. Convergence of quadratures for the quantity labelled by f. In the top row, the integrand becomes nearly non-integrable as the parallel velocity nearly vanishes at $\zeta = 0$. In either case, splitting the quadrature there recovers fast convergence.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

It can be shown from (2.1) that $\nabla\theta \times \nabla\zeta \neq \mathbf{0}$ implies

$$\boldsymbol{B} \cdot \nabla\theta = -[\nabla\psi \cdot (\nabla\theta \times \nabla\zeta)] \left(\frac{\partial\alpha}{\partial\zeta}\right)_{\psi,\theta}, \tag{3.15}$$

$$\boldsymbol{B} \cdot \nabla\zeta = +[\nabla\psi \cdot (\nabla\theta \times \nabla\zeta)] \left(\frac{\partial\alpha}{\partial\theta}\right)_{\psi,\zeta}. \tag{3.16}$$

Thus, we find equilibria by searching for a combination (R, Z, Λ, ω) to reduce the force balance error (3.11) at a set of collocation points using pseudo-spectral methods, subject to constraints on the pressure profile and the rotational transform or toroidal current profile. This boundary value problem is then solved as a minimisation problem using a trust-region method. In an optimisation constrained by force balance, varying (R, Z, Λ, ω) changes the magnetic field and (3.14) such that (3.13) remains valid.

Two advantages of this inverse approach for optimisation with bounce-averaged objectives are stated as follows.

(i) The variables (θ, ζ) on the boundary surface may be constructed so that maps parametrised in these coordinates are spectrally condensed (Hirshman & Breslau 1998; Hindenlang, Plunk & Maj 2025). Consequently, maps parametrised in (ρ, θ, ζ) in the plasma volume tend to have spectral expansions that converge more rapidly.

(ii) Force balance and other geometric objectives are best computed on a particular grid in (ρ, θ, ζ), which is fixed throughout optimisation. This ensures the spectral basis for (R, Z, Λ, ω) can be precomputed, avoiding 'off-grid' interpolation of a three-dimensional basis that bottlenecks pseudo-spectral codes (Boyd 2013, § 10.7). Furthermore, if the coordinate system varied throughout the optimisation, then so does the optimal grid for interpolation and quadrature. To preserve spectral accuracy, a pseudo-spectral code must first find this optimal grid and compute the basis there. This 'moving-grid' interpolation is doubly expensive in optimisation because the mentioned tasks must also be differentiated, which consumes significant memory.

These qualities enable faster generation of magnetic field data, which we discuss in the following section.

3.4. *Fast interpolation*

In this section, we outline our method for fast interpolation.

The Zernike basis concentrates the frequency transform of smooth maps on discs at lower frequencies than geometry-agnostic tensor-product bases. Boyd & Yu (2011) show the required number of spectral coefficients is typically half that of Fourier–Chebyshev. This ensures an optimisation that varies a finite number of coefficients in the Fourier–Zernike series for (R, Z, Λ, ω) at a time has more freedom compared with expansions in other bases. However, the Zernike basis is expensive to evaluate.

Our algorithm computes the Fourier–Zernike basis once prior to optimisation on a uniform $K_\theta \times K_\zeta$ grid in $(\theta, N_{\mathrm{FP}}\zeta) \in [0, 2\pi)^2$ on each surface. Any smooth, periodic map g required by the objective is computed from (R, Z, Λ, ω) on this grid and interpolated with a fast Fourier transform (FFT). The resulting Fourier

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

series are evaluated using type-2 non-uniform FFTs with computational cost that is linearithmic in $K_\theta K_\zeta$ plus linear in the number of points to evaluate (Barnett *et al*. 2019; Barnett 2021; Hsuan Shih *et al*. 2021),

$$g_{k_\theta k_\zeta} = \frac{c_{k_\theta}}{4\pi^2} \iint_{[0,2\pi)^2} \mathrm{d}\theta \, \mathrm{d}(N_{\mathrm{FP}}\zeta) \, g(\theta, N_{\mathrm{FP}}\zeta) \mathrm{e}^{-\mathrm{i}k_\theta \theta} \mathrm{e}^{-\mathrm{i}k_\zeta N_{\mathrm{FP}}\zeta}, \tag{3.17}$$

$$g(\alpha, \zeta) = \sum_{k_\theta=0}^{\lfloor K_\theta/2 \rfloor} \sum_{k_\zeta=-\lfloor K_\zeta/2 \rfloor}^{\lceil K_\zeta/2 \rceil - 1} \mathrm{Real}\left(g_{k_\theta k_\zeta} \mathrm{e}^{\mathrm{i}k_\theta \theta(\alpha,\zeta)} \mathrm{e}^{\mathrm{i}k_\zeta N_{\mathrm{FP}}\zeta} \right), \tag{3.18}$$

$$c_{k_\theta} = 1 \quad \text{if } k_\theta \in \{0, K_\theta/2\}\,; \text{ else } 2.$$

3.5. *Map to the mesh of field lines*

Evaluating maps along field lines requires finding the position of the field lines on some grid. To identify the coordinate θ at a given point (α, ζ) one may solve (3.13). To avoid repeating that inversion everywhere our objective demands, we compute the spectral projection $\{a_{xy}\}$ of the map $\alpha, \zeta \mapsto \theta - \alpha$ onto a tensor-product basis $\{b_{xy}\}$ that is orthogonal with respect to some weight ς,

$$a_{xy} \sim \iint \mathrm{d}\alpha \, \mathrm{d}\zeta \, (\theta - \alpha) \, \varsigma b^*_{xy}(\alpha, \zeta). \tag{3.19}$$

The Fourier–Chebyshev basis defined on the field period $(\alpha, N_{\mathrm{FP}}\zeta) \in [0, 2\pi)^2$ is chosen for reasons discussed by Mason & Handscomb (2002, §§ 5.5, 5.6, 6.3.4) and Boyd (2013, § 4.5),

$$b_{xy}(\alpha, \zeta) = \mathrm{e}^{\mathrm{i}x\alpha} \cos(y \arccos[N_{\mathrm{FP}}\zeta/\pi - 1]). \tag{3.20}$$

On each flux surface, (3.13) is solved on a tensor-product grid of size $X \times Y$ on the Fourier nodes across field lines and the Chebyshev nodes along field lines using Newton iteration with a backtracking line search. The series (3.19) is computed by interpolating $\theta - \alpha$ on that grid with a discrete cosine transform along field lines, followed by an FFT across field lines. The convergence of the series is illustrated in figure 7.

To extend the map beyond a single field period, we use

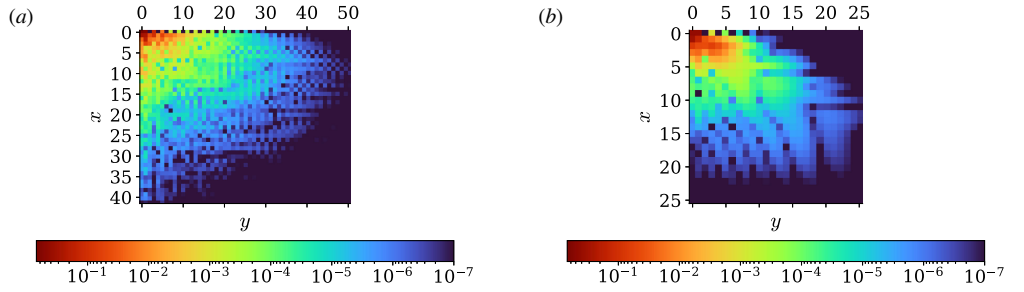
$$\theta \equiv \alpha_{\mathrm{mod}} + \sum_{x=0}^{\lfloor X/2 \rfloor} \sum_{y=0}^{Y-1} \mathrm{Real}\left(a_{xy} b_{xy}(\alpha_{\mathrm{mod}}, \zeta_{\mathrm{mod}}) \right) \pmod{2\pi}, \tag{3.21}$$

$$\alpha_{\mathrm{mod}} = \alpha_{\mathrm{shift}} \bmod (2\pi),$$
$$\alpha_{\mathrm{shift}} = \alpha + \iota \lfloor N_{\mathrm{FP}}\zeta/(2\pi) \rfloor 2\pi/N_{\mathrm{FP}},$$
$$\zeta_{\mathrm{mod}} = \zeta \bmod (2\pi/N_{\mathrm{FP}}).$$

The equivalence (3.21) is due to $\alpha + \iota\zeta = \alpha_{\mathrm{shift}} + \iota\zeta_{\mathrm{mod}}$ and uniqueness of solutions to (3.13). Figure 8 shows an illustration. The construction and evaluation of this series is accelerated with partial summation (Boyd 2013, § 10).

This approach avoids issues that result from changing the basis for (R, Z, Λ, ω) at each optimisation step (Appendix D).

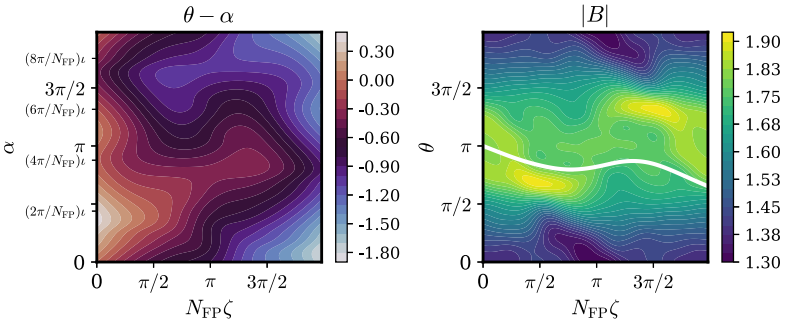
https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press



$|a_{xy}|$ on the plasma boundary of an NCSX stellarator with $N_{\mathrm{FP}} = 3$.

$|a_{xy}|$ on the plasma boundary of a Heliotron stellarator with $N_{\mathrm{FP}} = 19$.

FIGURE 7. Convergence of the spectral projection of α, $\zeta \mapsto \theta - \alpha$ onto the Fourier–Chebyshev basis (3.21). Equation (3.13) was solved to error $\leqslant 10^{-7}$. Note that if $\omega \to \Lambda/\iota$, then $\theta - \alpha \to \iota\zeta$, so the spectral width reduces to one parameter. Thus, if the optimiser is motivated to match higher frequency spectral coefficients of ω with Λ/ι, then field lines can be tracked at lower resolution.



$\alpha, \zeta \mapsto \theta - \alpha$ is shown on the left. On the right, $\alpha = \pi$ is shown in white.

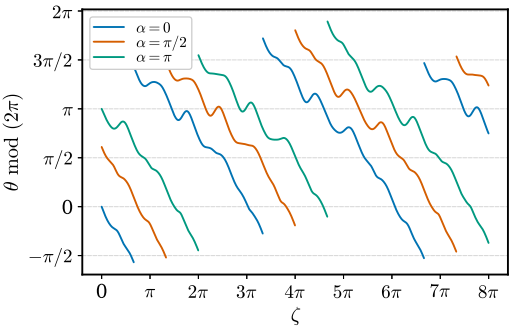


FIGURE 8. θ on the plasma boundary of an NCSX stellarator with $N_{\mathrm{FP}} = 3$.

3.6. *Jacobian of the map to the mesh of field lines*

In this section, we explain how we accelerate the iterative solve discussed in the previous section throughout optimisation. To bypass differentiating the iterative solve, we write the tangent and adjoint methods directly (Sapienza *et al*. 2025,

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

§§ 3.3.3, 3.9.2). For this task, we leverage the implicit function theorem to differentiate solutions $\theta^\star$ to (3.13) with respect to the optimisable parameters, denoted here with $\boldsymbol{x}_{\text{opt}}$. Define

$$f : \boldsymbol{x}_{\text{opt}}, \theta \mapsto \theta + \Lambda - \iota(\zeta + \omega) - \alpha. \tag{3.22}$$

Let $(\boldsymbol{x}^\star_{\text{opt}}, \theta^\star)$ satisfy $f(\boldsymbol{x}^\star_{\text{opt}}, \theta^\star) = 0$,

$$\frac{\partial f}{\partial \theta}(\boldsymbol{x}^\star_{\text{opt}}, \theta^\star) = 1 + \frac{\partial(\Lambda - \iota\omega)}{\partial \theta}(\boldsymbol{x}^\star_{\text{opt}}, \theta^\star) \underset{(3.13)}{=} \left(\frac{\partial \alpha}{\partial \theta}\right)_{\psi,\zeta}(\boldsymbol{x}^\star_{\text{opt}}, \theta^\star). \tag{3.23}$$

In the (ψ, α, ζ) covariant basis, the only non-zero component of the non-vanishing magnetic field is (3.16), so the derivative (3.23) is invertible. By the implicit function theorem, $\theta^\star$ is a continuously differentiable map of $\boldsymbol{x}_{\text{opt}}$ and $f(\boldsymbol{x}_{\text{opt}}, \theta^\star(\boldsymbol{x}_{\text{opt}})) = 0$ near $\boldsymbol{x}^\star_{\text{opt}}$. Moreover,

$$\frac{\partial \theta^\star}{\partial \boldsymbol{x}_{\text{opt}}}(\boldsymbol{x}_{\text{opt}}) = -\left[\frac{\partial f}{\partial \theta}(\boldsymbol{x}_{\text{opt}}, \theta^\star(\boldsymbol{x}_{\text{opt}}))\right]^{-1} \frac{\partial f}{\partial \boldsymbol{x}_{\text{opt}}}(\boldsymbol{x}_{\text{opt}}, \theta^\star(\boldsymbol{x}_{\text{opt}})). \tag{3.24}$$

Thus, we differentiate directly through the solution $\theta^\star$. Likewise, after updating $\boldsymbol{x}_{\text{opt}}$, we use (3.24) to warm start the next Newton iteration at an initial value that is correct to first order.

4. Optimisation for reduced neoclassical transport

We present an optimisation starting from a finite-beta helically omnigenous (OH) equilibrium. Finite-beta refers to the non-zero ratio of plasma pressure and magnetic pressure. We target flux surfaces near the boundary to reduce the effective ripple while maintaining reasonable elongation and curvature. With weights, $w_{\text{A}}, w_{\text{C}}, w_{\text{E}}, w_{\text{O}}, w_{\text{R}}$, the objective (4.1) is minimised while ensuring ideal MHD force balance (3.11) is maintained,

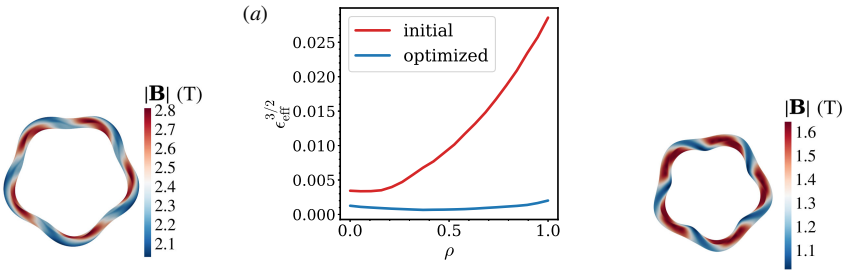
$$w_{\text{A}} f^2_{\text{aspect}} + w_{\text{C}} f^2_{\text{curv}} + w_{\text{E}} f^2_{\text{elongation}} + w_{\text{O}} f^2_{\text{omni}} + w_{\text{R}} f^2_{\text{ripple}}. \tag{4.1}$$

The initial equilibrium along with the definitions of the curvature and elongation objectives are provided by Gaur *et al*. (2025*a*) and Gaur (2024). The results are presented in figure 9. The optimisation took less than two hours with a GPU (NVIDIA Corporation 2020).

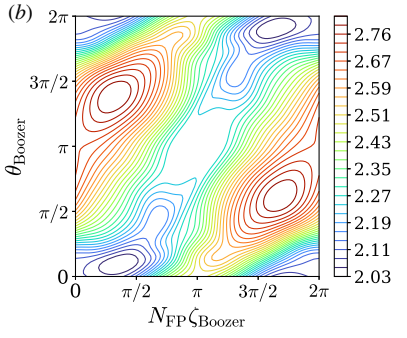
The omnigeneity objective is based on the work of Dudt *et al*. (2024), where it was shown that optimising for omnigeneity can in turn reduce the effective ripple. Directly optimising to reduce the effective ripple instead has the advantage that the optimiser is not biased towards a user-specified omnigenous field. For example, in Gaur *et al*. (2025*b*), we used this property to optimise for an umbilic boundary, while maintaining a low effective ripple without biasing the optimiser towards an omnigenous field with a specific helicity.

It should be noted that the assumptions used to derive the effective ripple increase in validity as the magnetic field becomes more omnigenous. For example, bounce-averaging assumes the radial orbit width is small compared with the magnetic field gradient scale length $\Delta r \ll L_B$. When finite orbit width effects become dominant (d'Herbemont *et al*. 2022), particles may traverse 'potato' orbits requiring a more global treatment (Satake, Okamoto & Sugama 2002). Hence, there is utility in optimisation that uses both objectives, either simultaneously or in succession.

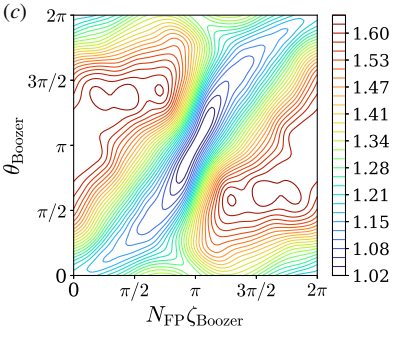
https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press



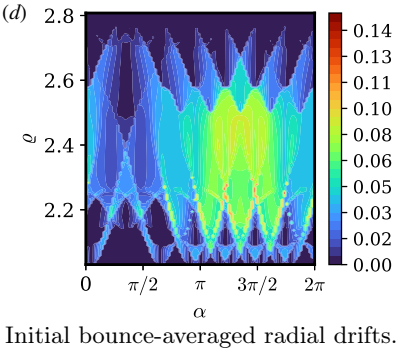
Neoclassical transport optimization.



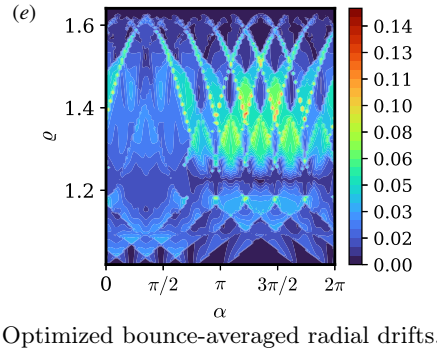
Initial equilibrium $|B|$.



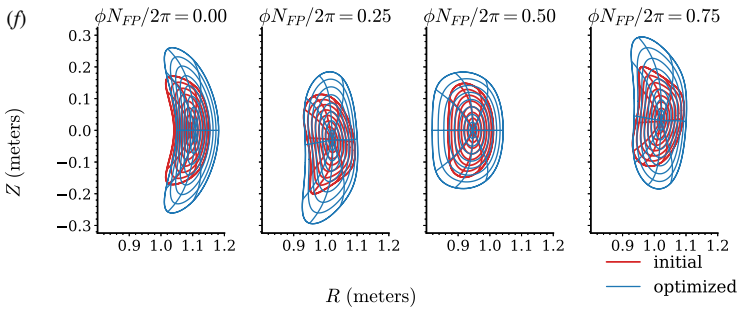
Optimized equilibrium $|B|$.



Initial bounce-averaged radial drifts.



Optimized bounce-averaged radial drifts.



Cross sections are compared at constant toroidal angles in the lab frame.

FIGURE 9. An OH transport optimisation. Panels (*b*) and (*c*) are shown in Boozer coordinates (D'haeseleer *et al*. 2012). Panels (*d*) and (*e*) show bounce-averaged radial drifts, summed in magnitude over $\zeta \in (0, 2\pi)$. The size of the region with large drifts appears reduced.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

5. Conclusions

In this work, we optimised a finite-beta configuration to directly reduce neoclassical transport using reverse-mode differentiation. More generally, we upgraded the DESC stellarator optimisation suite for fast, accurate, automatically differentiable bounce-averaging. We discussed how we perform moving-grid interpolation without sacrificing spectral accuracy. This accuracy ensures that changes in the objective due to small changes in controllable parameters reflect genuine improvement or degradation rather than noise due to error. Therefore, optimisation is more likely to be successful.

Our algorithm enables optimisation for many objectives to improve stellarator performance. These include maximisation of the second adiabatic invariant (Helander 2014, § 3.7; Rodríguez *et al*. 2024), energetic particle confinement (Nemov *et al*. 2008; Velasco *et al*. 2021*a*) and proxies for gyrokinetic turbulence such as the available energy (Mackenbach, Proll & Helander 2022; Mackenbach *et al*. 2023*a*). We have added all these objectives to DESC. Some of these objectives have previously had limited use in optimisation due to expensive computational requirements or difficulty finding desirable configurations. Further demonstration of optimisation with them remains as future work.

Acknowledgements

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

Funding

This work is funded through the SciDAC program by the US Department of Energy, Office of Fusion Energy Science, and Office of Advanced Scientific Computing Research under contract numbers DE-AC02-09CH11466, DE-SC0022005, and Field Work Proposal No. 1019. This work was also funded by the Peter B. Lewis Fund for Student Innovation in Energy and the Environment. This research used the resources of the Della computing cluster at Princeton University.

Declaration of interests

The authors declare no competing interests.

Author Contributions

Kaya Unalmis: Conceptualization; Methodology; Formal analysis; Investigation; Software; Validation; Visualization; Data curation; Writing - original draft; Writing - review & editing.

Rahul Gaur: Formal analysis; Software; Writing - review & editing.

Rory Conlin: Methodology; Software.

Dario Panici: Writing - review & editing.

Egemen Kolemen: Supervision; Writing - review & editing; Funding acquisition; Project administration.

Appendix A. Open source code

The implementation and supplementary information describing how correctness of automatic differentiation is enforced may be found in the DESC repository (Dudt *et al*. 2025). The implementation uses accelerated linear algebra XLA and Google's JAX library (Bradbury *et al*. 2018). JIT compilation in JAX compiles the code at the start of an optimisation. Optimisation may be accelerated on CPUs, GPUs and TPUs.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

Appendix B. Effective ripple

We explain a short derivation of ϵ_{eff}, similar to the one used by Nemov *et al.* (1999).

To obtain an explicit expression for Γ, we will bounce-integrate the drift-kinetic equation (2.4). Applying this operator to the drift-kinetic equation discretises the spatial coordinate ζ to a set of integral equations labelled by the magnetic well index w.

The collision operator in (2.4) is chosen to capture pitch angle scattering,

$$\mathcal{C}[f] = \nu m \frac{|v_\parallel|}{|B|} \frac{\partial}{\partial \mu} \left(\mu |v_\parallel| \frac{\partial f}{\partial \mu} \right). \tag{B1}$$

These derivatives are at fixed position and energy. The collision frequency ν depends only on the energy of the particle. The velocity ratio is $|v_\parallel|/|v| = (1 - |B|/\varrho)^{1/2}$. The nullspace of this collision operator contains velocity-isotropic distributions, so $\mathcal{C}[f_0 + f_1] = \mathcal{C}[f_1]$. In this form, (2.4) is the linearised Lorentz-gas Fokker–Planck equation (Goldston & Rutherford 1995, § 13).

In weakly collisional plasmas, the collision frequency is small compared with the particle bounce frequency. Consequently, fluctuations due to collisions homogenise along field lines rapidly, implying that the spatial variation in the plasma distribution along field lines in any particular magnetic well is small. Therefore, we approximate f_0 and f_1 to be spatially uniform along field lines in any particular magnetic well,

$$\nabla f = \left(\frac{\partial f}{\partial \psi} \right)_{\alpha,\zeta,E,\mu} \nabla \psi + \left(\frac{\partial f}{\partial \alpha} \right)_{\psi,\zeta,E,\mu} \nabla \alpha + \left(\frac{\partial f}{\partial \zeta} \right)_{\psi,\alpha,E,\mu} \nabla \zeta, \tag{B2}$$

$$|\nabla f| \gg |(\partial f/\partial \zeta) \nabla \zeta|. \tag{B3}$$

Nested flux surfaces (2.1) then imply the parallel drift $\boldsymbol{v}_{\text{Baños}}$ and the parallel spatial derivative of f_1 will be negligible in the bounce-integrated drift-kinetic equation,

$$\begin{aligned} \overline{\mathcal{C}[f_1]} &= \nu m \frac{\partial}{\partial \mu} \mu \int \frac{\mathrm{d}\zeta}{\boldsymbol{b} \cdot \nabla \zeta} \frac{|v_\parallel|}{|B|} \frac{\partial f_1}{\partial \mu} \\ &= \nu m \frac{\partial}{\partial \mu} \mu \frac{\partial f_1}{\partial \mu} \overline{|v_\parallel|^2/|B|} \\ &= \frac{\partial f_0}{\partial \psi} \overline{\boldsymbol{v}_{\mathrm{D}s} \cdot \nabla \psi}. \end{aligned} \tag{B4}$$

To write the last relation (B4), we assume there are sufficiently many passing particles so that f_0 is independent of α.[2] We proceed to invert the collision operator. First, label the geodesic curvature of the field line,

$$\kappa_{\mathrm{G}} = [\boldsymbol{b} \times (\boldsymbol{b} \cdot \nabla \boldsymbol{b})] \cdot \frac{\nabla \psi}{|\nabla \psi|} = \frac{\boldsymbol{b} \times \nabla |B|}{|B|} \cdot \frac{\nabla \psi}{|\nabla \psi|}. \tag{B5}$$

The second equality is a consequence of ideal MHD force balance (3.11). Now, the primitive with respect to μ of the bounce-integrated radial drift velocity is identified

[2]The claim $|\overline{(\partial f_0/\partial \alpha) \boldsymbol{v}_{\mathrm{D}s} \cdot \nabla \alpha}| \ll |\overline{\boldsymbol{v}_{\mathrm{D}s} \cdot \nabla f_0}|$ requires care because $|\nabla \alpha|$ grows unbounded when the magnetic shear is non-zero. If the distribution has variation across field lines, we assume it is captured by the higher order correction f_1.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

as follows:

$$\frac{\partial}{\partial \mu}\overline{|v_\parallel|\beta} = \overline{\boldsymbol{v}_{\mathrm{D}s} \cdot \nabla \psi}, \tag{B6}$$

$$\frac{\partial \beta}{\partial \mu} = \frac{\boldsymbol{v}_{\mathrm{D}s} \cdot \nabla \psi}{|v_\parallel|} = (|v|^2 |v_\parallel|^{-1} + |v_\parallel|) \frac{|\nabla \psi| \kappa_{\mathrm{G}}}{2\Omega_s}, \tag{B7}$$

$$\beta = -(3|v|^2 |v_\parallel| + |v_\parallel|^3) \frac{m |\nabla \psi| \kappa_{\mathrm{G}}}{6 \Omega_s |B|}. \tag{B8}$$

Inverting the μ derivative in (B4) completes the inversion of the collision operator,

$$\nu m \frac{\partial}{\partial \mu}\left(\mu \frac{\partial f_1}{\partial \mu}\overline{|v_\parallel|^2/|B|}\right) = \frac{\partial}{\partial \mu}\left(\frac{\partial f_0}{\partial \psi}\overline{|v_\parallel|\beta}\right)$$

$$\frac{\partial f_1}{\partial \mu} = \frac{\partial f_0}{\partial \psi} \frac{\overline{|v_\parallel|\beta}}{\nu m \mu \overline{|v_\parallel|^2/|B|}}. \tag{B9}$$

To compute (2.6), we will use the (E, μ) parametrisation of velocity space,

$$\int \mathrm{d}^3 \boldsymbol{v} = \frac{2\pi}{m^2} |B| \int_0^\infty \mathrm{d}E \int_0^{E/|B|} \frac{\mathrm{d}\mu}{|v_\parallel|} \tag{B10}$$

$$= \frac{2^{1/2}\pi}{m^{3/2}} |B| \int_0^\infty \mathrm{d}E \, E^{1/2} \int_{|B|}^\infty \frac{\mathrm{d}\varrho}{\varrho^2 (1 - |B|/\varrho)^{1/2}}. \tag{B11}$$

The plasma distribution vanishes where $\mu \geqslant E/|B|$, so the integration region was truncated. Using (B10), applying integration by parts in the μ coordinate and enforcing the boundary condition $\lim_{\mu \to 0} f_1 = 0$ at fixed energy, the radial particle flux density (2.6) can be written in terms of known quantities as follows:

$$\Gamma = -\int \mathrm{d}^3 \boldsymbol{v} \, |v_\parallel| \beta \frac{\partial f_1}{\partial \mu}. \tag{B12}$$

To make optimisation efficient, the flux surface average of the radial particle flux density is of interest to minimise. This is the average on an infinitesimal volume covering the surface,

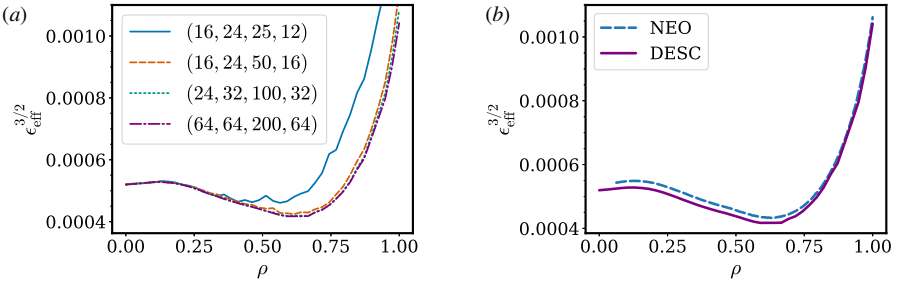
$$\langle \Gamma \rangle = \left(\int \frac{\mathrm{d}s}{|\nabla \psi|} \Gamma\right) \left(\int \frac{\mathrm{d}s}{|\nabla \psi|}\right)^{-1}. \tag{B13}$$

Here, $\mathrm{d}s$ is the differential surface area Jacobian. As (B12) enables computing the radial particle flux density through a quotient of bounce integrals along the magnetic field line (B9), it is more tractable to also compute the flux surface average along the field line,

$$\langle \Gamma \rangle = \left(\int_0^{2\pi} \mathrm{d}\alpha \int_{\mathbb{R}} \frac{\mathrm{d}\zeta}{\boldsymbol{B} \cdot \nabla \zeta} \Gamma\right) \left(\int_0^{2\pi} \mathrm{d}\alpha \int_{\mathbb{R}} \frac{\mathrm{d}\zeta}{\boldsymbol{B} \cdot \nabla \zeta}\right)^{-1}. \tag{B14}$$

We proceed to extract a dimensionless quantity Γ_0 for the optimisation objective. First, we use (B11) and (B14) to remove the spatial dependence in the boundary of the velocity integral,

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press



Five field lines are followed for 100 field periods. The legend shows the resolution (X, Y, N_ϱ, N_q). $K_\theta = K_\zeta = 33$.

NEO-DESC comparison.

FIGURE 10. Resolution scan for ϵ_{eff} on the W7-X equilibrium in the DESC repository.

$$\langle \Gamma \rangle = -\frac{2\pi}{m^2}\left(\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\infty} \mathrm{d}E\, E \int_{\mathbb{R}} \frac{\mathrm{d}\zeta}{\boldsymbol{b}\cdot\nabla\zeta} \int_{|B|}^{\infty} \frac{\mathrm{d}\varrho}{\varrho^2}\, \beta \frac{\partial f_1}{\partial \mu}\right)\left(\int_0^{2\pi} \mathrm{d}\alpha \int_{\mathbb{R}} \frac{\mathrm{d}\zeta}{\boldsymbol{B}\cdot\nabla\zeta}\right)^{-1}$$
$$= -\frac{2\pi}{m^2}\left(\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\infty} \mathrm{d}E\, E \int_{|B|_{\min}}^{|B|_{\max}} \frac{\mathrm{d}\varrho}{\varrho^2} \sum_w \overline{|v_\parallel|\beta} \frac{\partial f_1}{\partial \mu}\right)\left(\int_0^{2\pi} \mathrm{d}\alpha \int_{\mathbb{R}} \frac{\mathrm{d}\zeta}{\boldsymbol{B}\cdot\nabla\zeta}\right)^{-1}. \tag{B15}$$

Here, $|B|_{\min}$ and $|B|_{\max}$ are the min and max values on the flux surface. The integral was truncated at $|B|_{\max}$ as $f_1 = 0$ for passing particles. Now, changing coordinates in (B8),

$$\beta = -\frac{(2mE^3)^{1/2}c}{3\varrho e|B|}(1 - |B|/\varrho)^{1/2}(4\varrho/|B| - 1)|\nabla\psi|\kappa_{\mathrm{G}} \tag{B16}$$

and using the new partition for the velocity integral (B15), the expression (B14) may be approximated using a sum over all wells in the interval $[\zeta_1, \zeta_2]$ (2.7).

B.1 Resolution scan for the neoclassical transport coefficient

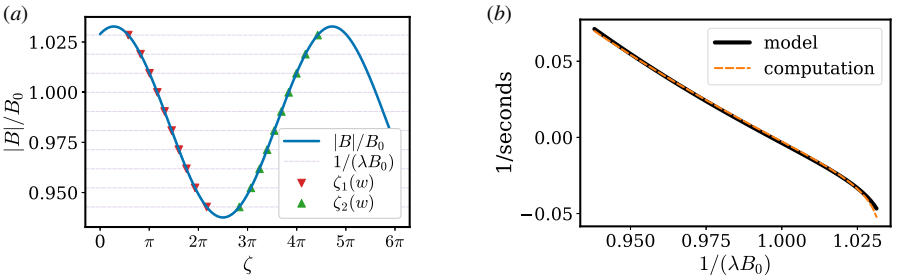
Figure 10(*a*) presents a resolution scan for ϵ_{eff}. Figure 10(*b*) compares the result to the NEO code (Nemov *et al*. 1999), which uses a finite difference technique and requires transforming to Boozer coordinates. For the bounce integrals, NEO employs an explicit Runge–Kutta scheme which has algebraic convergence of order $1 + \eta/2$ for (3.3).

We mention some performance benchmarking of our algorithm in what follows. Computing $\epsilon_{\text{eff}} : \mathbb{R}^{4074} \to \mathbb{R}^{10}$ and its derivative, on ten flux surfaces, following each field line for 75 field periods, with resolution $(K_\theta, K_\zeta, X, Y, N_\varrho, N_q) = (32, 32, 32, 32, 100, 32)$ was profiled to take less than one and two seconds, respectively, with a CPU (Intel Corporation 2019), with peak memory consumption near one gigabyte. These computations are at least an order of magnitude faster with a GPU.

Appendix C. Bounce-averaged drifts in a shifted-circle model

In a shifted-circle model for plasma equilibrium, we can obtain analytical expressions for bounce-averaged drifts. We further verify our algorithm with this model in figure 11.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press



For a pitch marked by a horizontal line, $|v_\parallel| = 0$ at the points marked by triangles.

The bounce-averaged binormal drift in the configuration in figure 11a is compared.

FIGURE 11. Comparison of our shifted-circle model for the binormal drift with the result computed by our algorithm. The minor difference in panel (*b*) is because the shifted-circle model is accurate to $\mathcal{O}(\epsilon^2)$.

In this model, the magnetic field can be written as

$$\boldsymbol{B} = \nabla\alpha \times \nabla\chi = F\nabla\phi + \frac{\mathrm{d}\chi}{\mathrm{d}\rho}\frac{\rho}{R_0}\nabla\vartheta, \tag{C1}$$

where $\alpha = \phi - \iota^{-1}\vartheta$, χ is the poloidal flux, F is the enclosed poloidal current, R_0 is the average major radius and ρ is a radial coordinate. To lowest order, the Grad–Shafranov equation has the constant solution $F = F_0$. To the next order, the pressure gradient is

$$\mathrm{d}p/\mathrm{d}\rho = -F_0 R^{-2}(\mathrm{d}F/\mathrm{d}\rho). \tag{C2}$$

To first order, the poloidal magnetic field can be ignored so that the field satisfies $|B| = B_0(1 - \epsilon\cos\vartheta)$ and $\boldsymbol{b}\cdot\nabla\vartheta = G_0(1 - \epsilon\cos\vartheta)$, where $\epsilon \ll 1$ is the reciprocal of the aspect ratio. Here, B_0 and G_0 are constants. In this model, the global shear $\hat{s}$, normalised pressure gradient and integrated local shear are

$$\hat{s} = -\rho\iota^{-1}(\mathrm{d}\iota/\mathrm{d}\rho), \tag{C3}$$

$$\alpha_{\mathrm{MHD}} = -2^{-1}\iota^{-2}(\mathrm{d}p/\mathrm{d}\rho), \tag{C4}$$

$$\mathtt{gds21} = (\mathrm{d}\chi/\mathrm{d}\rho)(\mathrm{d}\iota^{-1}/\mathrm{d}\rho)\nabla\chi\cdot\nabla\alpha = -\hat{s}(\hat{s}\vartheta - |B|^{-4}\alpha_{\mathrm{MHD}}\sin\vartheta) + \mathcal{O}(\epsilon). \tag{C5}$$

The binormal, geometric part of the $\nabla|B|$ drift is

$$(\nabla|B|)_{\mathrm{drift}} = |B|^{-3}(\boldsymbol{B}\times\nabla|B|)\cdot\nabla\alpha = f_2(-\hat{s} + \cos\vartheta - \hat{s}^{-1}\mathtt{gds21}\sin\vartheta). \tag{C6}$$

The binormal, geometric part of the curvature drift is

$$\begin{aligned}\mathtt{cvdrift} &= |B|^{-3}[\boldsymbol{B}\times\nabla(p + |B|^2/2)]\cdot\nabla\alpha \\ &= (\nabla|B|)_{\mathrm{drift}} + f_3|B|^{-3}(\mathrm{d}p/\mathrm{d}\rho) \\ &= f_2(-\hat{s} + \cos\vartheta + \hat{s}\vartheta\sin\vartheta - B_0^{-4}\alpha_{\mathrm{MHD}}\sin^2\vartheta) + f_3 B_0^{-2}\alpha_{\mathrm{MHD}} + \mathcal{O}(\epsilon).\end{aligned} \tag{C7}$$

The scalars f_2 and f_3 contain some constants. The bounce-averaged drift is

$$\langle v_{\mathrm{D}}\rangle = \left(\int_{\vartheta_1}^{\vartheta_2}\frac{\mathrm{d}\vartheta}{\boldsymbol{b}\cdot\nabla\vartheta}|v_\parallel|^{-1}\right)^{-1}\int_{\vartheta_1}^{\vartheta_2}\frac{\mathrm{d}\vartheta}{\boldsymbol{b}\cdot\nabla\vartheta}\left[|v_\parallel|\mathtt{cvdrift} + \frac{|v_\perp|^2}{2|v_\parallel|}(\nabla|B|)_{\mathrm{drift}}\right]. \tag{C8}$$

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

As used by Connor, Hastie & Martin (1983) and shown by Hegna (2015), in the limit of a large aspect ratio shifted-circle model, the parallel speed of a particle with a fixed energy is $|v_\parallel| = (2E/m)^{1/2}(2\epsilon\lambda B_0)^{1/2}(k^2 - \sin^2(\vartheta/2))^{1/2}$, where

$$k^2 = 2^{-1}[(1 - \lambda B_0)(\epsilon\lambda B_0)^{-1} + 1] \tag{C9}$$

parametrises the pitch angle λ. Using these simplifications and $|v_\perp|^2/2 = E/m - |v_\parallel|^2/2$,

$$\begin{aligned}\langle v_D \rangle = \left(\int_{-2\arcsin k}^{2\arcsin k} \frac{\mathrm{d}\vartheta}{\boldsymbol{b}\cdot\boldsymbol{\nabla}\vartheta} (2\epsilon\lambda B_0)^{-1/2}(k^2 - \sin^2(\vartheta/2))^{-1/2} \right)^{-1} \\ \int_{-2\arcsin k}^{2\arcsin k} \frac{\mathrm{d}\vartheta}{\boldsymbol{b}\cdot\boldsymbol{\nabla}\vartheta} \big[(2\epsilon\lambda B_0)^{1/2}(k^2 - \sin^2(\vartheta/2))^{1/2} \mathtt{cvdrift} \\ - 2^{-1/2}(\epsilon\lambda B_0)^{1/2}(k^2 - \sin^2(\vartheta/2))^{1/2}(\boldsymbol{\nabla}|B|)_{\mathrm{drift}} \\ + 2^{-3/2}(\epsilon\lambda B_0)^{-1/2}(k^2 - \sin^2(\vartheta/2))^{-1/2}(\boldsymbol{\nabla}|B|)_{\mathrm{drift}} \big].\end{aligned} \tag{C10}$$

The following identities simplify (C10). The incomplete elliptic integrals are converted to complete elliptic integrals using the reciprocal-modulus transformation in (C11) and (C12) (Olver *et al*. 2024). Here, K and E are complete elliptic integrals of the first and second kind, respectively,

$$\mathsf{I}_0 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{-1/2} = 4K(k), \tag{C11}$$

$$\mathsf{I}_1 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{1/2} = 4[E(k) + (k^2 - 1)K(k)], \tag{C12}$$

$$\mathsf{I}_2 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{-1/2}\vartheta \sin\vartheta = 16[E(k) + (k^2 - 1)K(k)], \tag{C13}$$

$$\mathsf{I}_3 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{1/2}\vartheta \sin\vartheta = \frac{16}{9}[(4k^2 - 2)E + (3k^4 - 5k^2 + 2)K], \tag{C14}$$

$$\mathsf{I}_4 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{-1/2}\sin^2\vartheta = \frac{16}{3}[(2k^2 - 1)E + (1 - k^2)K], \tag{C15}$$

$$\begin{aligned}\mathsf{I}_5 &= \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{1/2}\sin^2\vartheta \\ &= \frac{32}{30}[2(1 - k^2 + k^4)(E + (k^2 - 1)K) - (1 - 3k^2 + 2k^4)k^2 K],\end{aligned} \tag{C16}$$

$$\mathsf{I}_6 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{-1/2}\cos\vartheta = 8E - 4K, \tag{C17}$$

$$\mathsf{I}_7 = \int_{-2\arcsin k}^{2\arcsin k} \mathrm{d}\vartheta \; (k^2 - \sin^2(\vartheta/2))^{1/2}\cos\vartheta = \frac{4}{3}[(2k^2 - 1)E - (k^2 - 1)K]. \tag{C18}$$

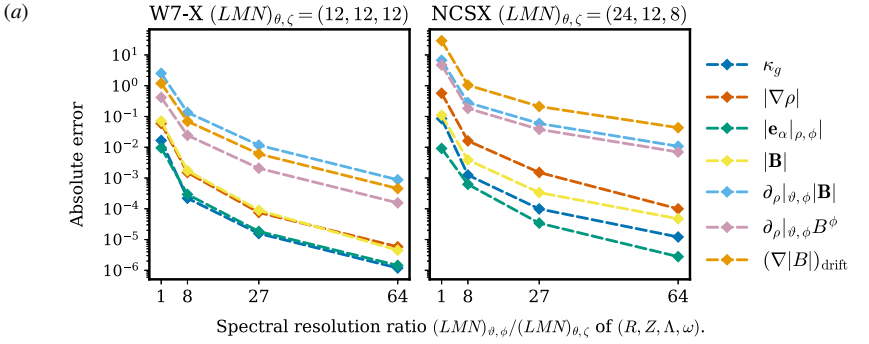
https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

Using these formulae with $\epsilon_\lambda = 2\epsilon\lambda B_0$, to lowest order, the bounce-averaged drift is

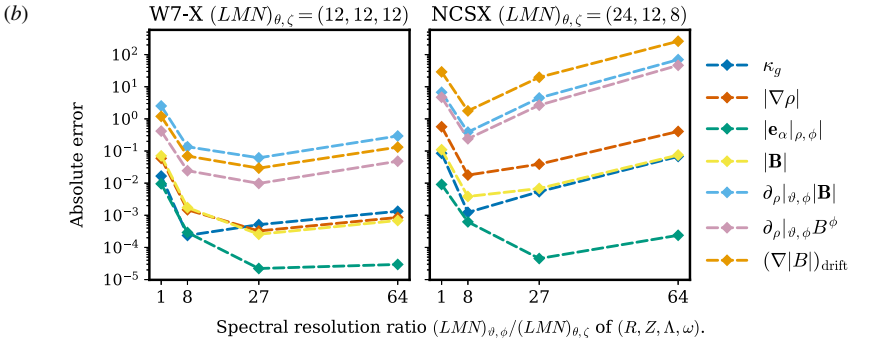
$$\langle v_D \rangle = \frac{1}{\mathsf{I}_0}\left[f_3 \frac{\alpha_{\mathrm{MHD}}}{B_0^2}\epsilon_\lambda \mathsf{I}_1 - \frac{f_2}{2}\left(\hat{s}(\mathsf{I}_0 + \epsilon_\lambda \mathsf{I}_1 - \mathsf{I}_2 - \epsilon_\lambda \mathsf{I}_3) + \frac{\alpha_{\mathrm{MHD}}}{B_0^4}(\mathsf{I}_4 + \epsilon_\lambda \mathsf{I}_5) - (\mathsf{I}_6 + \epsilon_\lambda \mathsf{I}_7) \right) \right]. \tag{C19}$$

Appendix D. Issues with changing the spectral basis to straight field line coordinates

In figure 12, we show that parametrising our basis for (R, Z, Λ, ω) in straight field line coordinates $(\vartheta, \phi) = (\theta + \Lambda, \zeta + \omega)$ (Grimm, Greene & Johnson 1976) is inefficient and ill-conditioned. Therefore, we use the approach discussed in the main text instead.



Equation (3.13) was solved to error $\leqslant 10^{-10}$ on the tensor-product of the optimal concentric sampling grid (Ramos-López *et al.* 2016) in $(\rho, \vartheta) \in [0, 1] \times [0, 2\pi)$ and a uniform grid in $\phi \in [0, 2\pi/N_{\mathrm{FP}})$. (R, Z, Λ, ω) were interpolated to a Fourier–Zernike series in (ρ, ϑ, ϕ) with maximum mode numbers $(L, M, N)_{\vartheta,\phi}$ on this grid. The interpolation used a 1.5× over-sampled, in both ρ and ϑ, weighted least-squares fit to improve conditioning for the Zernike series, followed by an FFT in ϕ. (The optimal grid for interpolation to a Zernike series does not coincide with the optimal grid for quadrature to project onto the Zernike basis because the Zernike basis is not a tensor-product basis. Interpolation with a weighted least-squares fit was chosen because the interpolation grid is sparser than the quadrature grid.) Each quantity was then computed on a uniform grid in (ρ, ϑ, ϕ). Quadrature required to compute a quantity in the plot was done on an over-sampled grid to account for nonlinearity in the computation from (R, Z, Λ, ω). Zernike polynomials were evaluated with stable Jacobi polynomial recurrence relations using the algorithm in Elmacioglu *et al.* (2025).



This is the same demonstration as figure 12a except equation (3.13) is solved to error 10^{-7}.

FIGURE 12. Error induced by changing the Fourier–Zernike basis for (R, Z, Λ, ω) from flux coordinates (θ, ζ) to the straight field line coordinates (ϑ, ϕ). Fitting at the resolution that obtains the error of 10^{-4} Tesla in $|B|$ on the NCSX stellarator in panel (a) took 10 minutes with a CPU (Intel Corporation 2019).

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

REFERENCES

ABEL, I.G., PLUNK, G.G., WANG, E., BARNES, M., COWLEY, S.C., DORLAND, W. & SCHEKOCHIHIN, A.A. 2013 Multiscale gyrokinetics for rotating tokamak plasmas: fluctuations, transport and energy flows. *Rep. Prog. Phys.* **76**, 116201.

ALPERT, B.K. 1999 Hybrid gauss-trapezoidal quadrature rules. *SIAM J. Sci. Comput.* **20**, 1551–1584.

ANTONSEN, T., PAUL, E.J. & LANDREMAN, M. 2019 Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. *J. Plasma Phys.* **85**, 905850207.

BARNETT, A.H. 2021 Aliasing error of the $\exp(\beta\sqrt{1-z^2})$ kernel in the nonuniform fast Fourier transform. *Appl. Comput. Harmon. A* **51**, 1–16.

BARNETT, A.H., MAGLAND, J. & AF KLINTEBERG, L. 2019 A parallel nonuniform fast Fourier transform library based on an 'exponential of semicircle' kernel. *Siam J. Sci. Comput.* **41**, C479–C504.

BOYD, J.P. 2013 *Chebyshev and Fourier Spectral Methods: Second Revised Edition*, chap. 4.5, 10. Dover Publications.

BOYD, J.P. & YU, F. 2011 Comparing seven spectral methods for interpolation and for solving the Poisson equation in a disk: Zernike polynomials, Logan–Shepp ridge polynomials, Chebyshev–Fourier series, cylindrical Robert functions, Bessel–Fourier expansions, square-to-disk conformal mapping and radial basis functions. *J. Comput. Phys.* **230**, 1408–1438.

BRADBURY, J. *et al.* 2018 JAX v0.7.2: composable transformations of Python+NumPy programs. Available at: https://github.com/jax-ml/jax.

CALVO, I., PARRA, F.I., VELASCO, J.L. & ALONSO, J.A. 2017 The effect of tangential drifts on neoclassical transport in stellarators close to omnigeneity. *Plasma Phys. Control. Fusion* **59**, 055014.

CARY, J.R. & SHASHARINA, S.G. 1997 Omnigenity and quasihelicity in helical plasma confinement systems. *Phys. Plasmas* **4**, 3323–3333.

CONLIN, R., DUDT, D.W., PANICI, D. & KOLEMEN, E. 2023 The DESC stellarator code suite. Part 2. Perturbation and continuation methods. *J. Plasma Phys.* **89**, 955890305.

CONLIN, R., KIM, P., DUDT, D.W., PANICI, D. & KOLEMEN, E. 2024 Stellarator optimization with constraints. *J. Plasma Phys.* **90**, 905900501.

CONNOR, J.W., HASTIE, R.J. & MARTIN, T.J. 1983 Effect of pressure gradients on the bounce-averaged particle drifts in a tokamak. *Nucl. Fusion* **23**, 1702.

DREVLAK, M., BEIDLER, C.D., GEIGER, J., HELANDER, P. & TURKIN, Y. 2018 Optimisation of stellarator equilibria with ROSE. *Nucl. Fusion* **59**, 016010.

DUDT, D.W., CONLIN, R., PANICI, D. & KOLEMEN, E. 2023 The DESC stellarator code suite part 3: quasi-symmetry optimization. *J. Plasma Phys.* **89**, 955890201.

DUDT, D., CONLIN, R., PANICI, D., UNALMIS, K., ELMACIOGLU, Y.G., GAUR, R., KIM, P. & KOLEMEN, E. 2025 DESC. Zenodo. Available at: https://github.com/PlasmaControl/DESC.

DUDT, D.W., GOODMAN, A.G., CONLIN, R., PANICI, D. & KOLEMEN, E. 2024 Magnetic fields with general omnigenity. *J. Plasma Phys.* **90**, 905900120.

DUDT, D.W. & KOLEMEN, E. 2020 DESC: A stellarator equilibrium solver. *Phys. Plasmas* **27**, 102513.

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*, chap 4.9, 5.2, 6.6. Springer Science & Business Media.

D'HERBEMONT, V., PARRA, F.I., CALVO, I. & VELASCO, J.L. 2022 Finite orbit width effects in large aspect ratio stellarators. *J. Plasma Phys.* **88**, 905880507.

ELMACIOGLU, Y.G., CONLIN, R., DUDT, D.W., PANICI, D. & KOLEMEN, E. 2025 Zernipax: a fast and accurate Zernike polynomial calculator in Python. *Appl. Math. Comput.* **505**, 129534.

GAUR, R. 2024 Omnigenous equilibria with enhanced stability: dataset and analysis files. Available at: https://doi.org/10.5281/zenodo.13887566.

GAUR, R., CONLIN, R., DICKINSON, D., PARISI, J.F., PANICI, D., DUDT, D., KIM, P., UNALMIS, K., DORLAND, W.D. & KOLEMEN, E. 2025*a* Omnigenous stellarators with improved ideal and kinetic ballooning stability. *Plasma Phys. Control. Fusion* **67**, 125015.

GAUR, R., PANICI, D.G., ELDER, T.M., LANDREMAN, M., UNALMIS, K., ELMACIOGLU, Y.G., DUDT, D.W., CONLIN, R. & KOLEMEN, E. 2025*b* Omnigenous umbilic stellarators. *J. Plasma Phys.* **91**, E151.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

GOLDSTON, R.J. & RUTHERFORD, P.H. 1995 *Introduction to Plasma Physics*, chap. 13. Taylor & Francis.

GRIMM, R.C., GREENE, J.M. & JOHNSON, J.L. 1976 Computation of the magnetohy-drodynamic spectrum in axisymmetric toroidal confinement systems. *Methods Comput. Phys.* **16**, 253.

HEGNA, C.C. 2015 The effect of three-dimensional fields on bounce averaged particle drifts in a tokamak. *Phys. Plasmas* **22**, 072510.

HELANDER, P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields. *Rep. Prog. Phys.* **77**, 087001.

HELANDER, P. & SIGMAR, D.J. 2005 *Collisional Transport in Magnetized Plasmas*, chap. 1, 11, vol. 4. Cambridge University Press.

HINDENLANG, F., PLUNK, G.G. & MAJ, O. 2025 Computing mhd equilibria of stellarators with a flexible coordinate frame. *Plasma Phys. Control. Fusion* **67**, 045002.

HIRSHMAN, S.P. & BRESLAU, J. 1998 Explicit spectrally optimized Fourier series for nested magnetic surfaces. *Phys. Plasmas* **5**, 2664–2675.

HIRSHMAN, S.P. & WHITSON, J.C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26**, 3553.

HSUAN SHIH, Y., WRIGHT, G., ANDÉN, J., BLASCHKE, J., BARNETT, A.H. 2021 CUFINUFFT: a load-balanced GPU library for general-purpose nonuniform FFTs. arXiv:2102.08463.

Intel Corporation 2019 Intel core i7-9750h.

KERNBICHLER, W., KASILOV, S.V., KAPPER, G., MARTITSCH, A.F., NEMOV, V.V., ALBERT, C. & HEYN, M.F. 2016 Solution of drift kinetic equation in stellarators and tokamaks with broken symmetry using the code NEO-2. *Plasma Phys. Control Fusion* **58**, 104001.

KOVRIZHNYKH, L.M. 1984 Neoclassical theory of transport processes in toroidal magnetic confinement systems, with emphasis on non-axisymmetric configurations. *Nucl. Fusion* **24**, 851.

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.

LAZERSON, S., SCHMITT, J., ZHU, C., BRESLAU, J. & STELLOPT Developers, All 2020 Stellopt. Available at: https://doi.org/10.11578/dc.20180627.6.

MACKENBACH, R.J.J., PROLL, J.H.E., WAKELKAMP, R. & HELANDER, P. 2023*a* The available energy of trapped electrons: a nonlinear measure for turbulent transport. *J. Plasma Phys.* **89**, 905890513.

MACKENBACH, R.J.J., DUFF, J.M., GERARD, M.J., PROLL, J.H.E., HELANDER, P. & HEGNA, C.C. 2023*b* Bounce-averaged drifts: equivalent definitions, numerical implementations, and example cases. *Phys. Plasmas* **30**, 093901.

MACKENBACH, R.J.J., PROLL, J.H.E. & HELANDER, P. 2022 Available energy of trapped electrons and its relation to turbulent transport. *Phys. Rev. Lett.* **128**, 175001.

MASON, J.C. & HANDSCOMB, D.C. 2002 *Chebyshev Polynomials*, chap. 5.5, 5.6, 6.3.4. Chapman and Hall/CRC.

MATSUDA, Y. & STEWART, J.J. 1986 A relativistic multiregion bounce-averaged fokker-planck code for mirror plasmas. *J. Comput. Phys.* **66**, 197–217.

NEMOV, V.V., KASILOV, S.V., KERNBICHLER, W., HEYN, M.F. 1999 Evaluation of $1/\nu$ neoclassical transport in stellarators. *Phys. Plasmas* **6**, 4622–4632.

NEMOV, V.V., KASILOV, S.V., KERNBICHLER, W. & LEITOLD, G.O. 2008 Poloidal motion of trapped particle orbits in real-space coordinates. *Phys. Plasmas* **15**, 052501.

NVIDIA Corporation 2020 NVIDIA A100 tensor core GPU. Available at: https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/PB-10577-001_v02.pdf.

OCHS, I.E. 2025 Bounce-averaged theory in arbitrary multi-well plasmas: solution domains and the graph structure of their connections. *J. Plasma Phys.* **91**, E123.

OLVER, F.W.J. *et al.* 2024 NIST digital library of mathematical functions. Available at: https://dlmf.nist.gov/.

PANICI, D., CONLIN, R., DUDT, D.W., UNALMIS, K., KOLEMEN, E. 2023 The DESC stellarator code suite. Part 1. Quick and accurate equilibria computations. *J. Plasma Phys.* **89**, 955890303.

PAUL, E.J., ABEL, I.G., LANDREMAN, M., DORLAND, W. 2019 An adjoint method for neoclassical stellarator optimization. *J. Plasma Phys.* **85**, 795850501.

PAUL, E.J., LANDREMAN, M. & ANTONSEN T. 2021 Gradient-based optimization of 3d mhd equilibria. *J. Plasma Phys.* **87**, 905870214.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press

PETROV, Y.V., HARVEY, R.W. 2016 A fully-neoclassical finite-orbit-width version of the cql3d fokker–planck code. *Plasma Phys. Control. Fusion* **58**, 115001.

RAMOS-LÓPEZ, D., SÁNCHEZ-GRANERO, M.A., FERNÁNDEZ-MARTÍNEZ, M. & MARTÍNEZ-FINKELSHTEIN, A. 2016 Optimal sampling patterns for Zernike polynomials. *Appl. Math. Comput.* **274**, 247–257.

RODRÍGUEZ, E., HELANDER, P. & GOODMAN, A.G. 2024 The maximum-j property in quasi-isodynamic stellarators. *J. Plasma Phys.* **90**, 905900212.

SAPIENZA, F. *et al.* 2025 Differentiable programming for differential equations: a review. arXiv:2406.09699.

SATAKE, S., OKAMOTO, M., SUGAMA, H. 2002 Lagrangian neoclassical transport theory applied to the region near the magnetic axis. *Phys. Plasmas* **9**, 3946–3960.

SPITZER Jr., L. 1958 The stellarator concept. *Phys. Fluids* **1**, 253–264.

SPONG, D.A., HIRSHMAN, S.P., WHITSON, J.C., BATCHELOR, D.B., CARRERAS, B.A., LYNCH, V.E. & ROME, J.A. 1998 J^* optimization of small aspect ratio stellarator/tokamak hybrid devices. *Phys. Plasmas* **5**, 1752–1758.

SÜLI, E. & MAYERS, D.F. 2003 *An Introduction to Numerical Analysis*. Cambridge University Press.

TREFETHEN, L.N., WEIDEMAN, J.A.C. 2014 The exponentially convergent trapezoidal rule. *SIAM Rev.* **56**, 385–458.

VELASCO, J.L., MULAS, S., CAPPA, A. *et al.* 2021*a* A model for the fast evaluation of prompt losses of energetic ions in stellarators. *Nucl. Fusion* **61**, 116059.

VELASCO, J.L., CALVO, I., PARRA, F.I., D'HERBEMONT, V., SMITH, H.M., CARRALERO, D., ESTRADA, T. & the W7-X Team 2021*b* Fast simulations for large aspect ratio stellarators with the neoclassical code knosos. *Nucl. Fusion* **61**, 116013.

VELASCO, J.L., CALVO, I., PARRA, F.I. & GARCÍA-REGAÑA, J.M. 2020 KNOSOS: a fast orbit-averaging neoclassical code for stellarator geometry. *J. Comput. Phys.* **418**, 109512.

https://doi.org/10.1017/S0022377826101652 Published online by Cambridge University Press