

MRX: A differentiable 3D MHD equilibrium solver without nested flux surfaces

Tobias Blickhan,^{1,*} Julianne Stratton,¹ and Alan A. Kaptanoglu^{1,2}

¹*Courant Institute of Mathematical Sciences, New York University*

²*William E. Boeing Department of Aeronautics and Astronautics, University of Washington*

This article introduces a new 3D magnetohydrodynamic (MHD) equilibrium solver, MRX, based on relaxation of a magnetic field to a lower energy equilibrium state via admissible variations. We describe the mathematical theory behind this method and discuss what can be transferred to the fully discrete setting of a relaxation code. Our code is designed to address a number of traditional challenges to 3D MHD equilibrium solvers: enforcing physical constraints such as divergence-free magnetic field, high-order convergence, handling strongly shaped polar geometries, and controlling topology changes of the magnetic field. Building on the JAX framework, we address questions of computational efficiency on modern computing architectures, user accessibility, and differentiability at each step. The solver is verified on manufactured vacuum solutions and against a VMEC equilibrium of a quasi-axisymmetric stellarator. It is further demonstrated on a finite- β NCSX configuration, where we study h and p refinement, preservation of helicity, reconnection events and seeded island chains. Finally, an inexact Newton method enables the computation of high-resolution finite- β equilibria with islands and chaos in minutes on a single GPU.

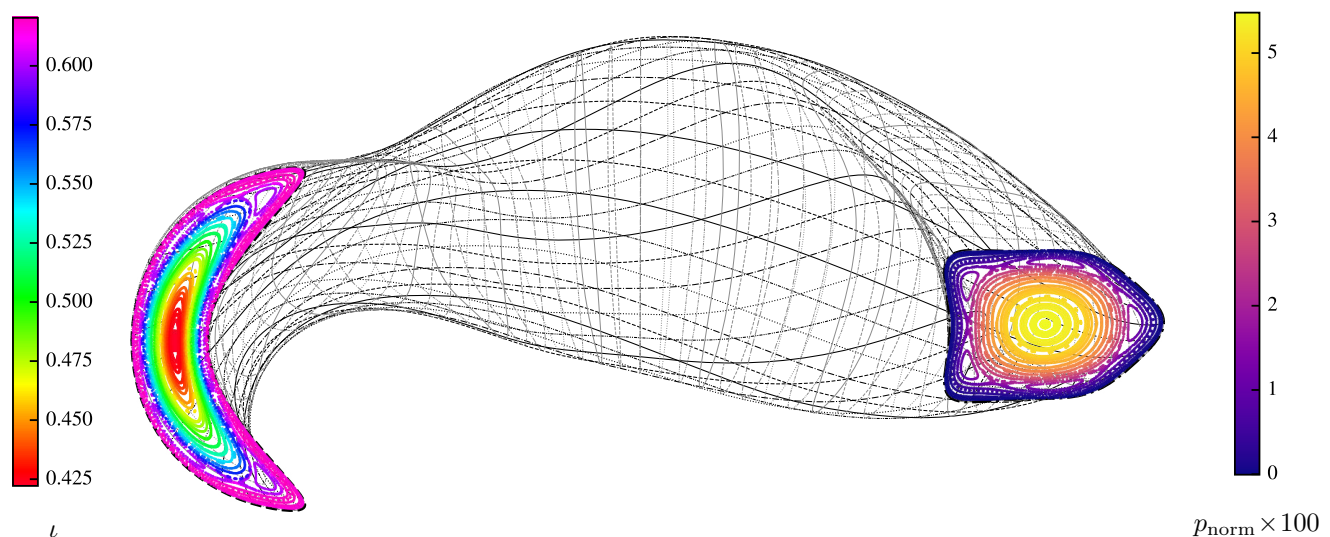


FIG. 1: Visualization of a non-nested NCSX equilibrium with MRX, colored by rotational transform ι (left) and normalized pressure (right). Starting from a nested configuration, a large $3/5$ island chain has formed through magnetic reconnection.

1. INTRODUCTION

Fusion energy offers a promising path for a clean, reliable, and sustainable source of energy to power global infrastructure, but significant challenges remain. The dynamics of plasmas are often modeled by magnetohydrodynamic (MHD) models, which are partial differential equations (PDEs) that treat the plasma as a single fluid.

For many approximately steady-state space and laboratory plasmas, the equilibrium is well-described by the

time-independent limit of the MHD equations with zero equilibrium flow. Three-dimensional, static, ideal magnetohydrodynamic equilibrium is given by the force balance equation coupled with Maxwell's equations.

This constitutes a search for $B : \Omega \rightarrow \mathbb{R}^3$, with a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ such that,

$$J \times B = \text{grad } p, \quad \text{curl } B = J, \quad \text{div } B = 0, \quad (1)$$

together with suitable boundary conditions. We will refer to (1) as the magnetohydrostatic (MHS) problem. Throughout this work, we take units so that vacuum permeability $\mu_0 = 1$. Plasma pressure is denoted by p .

Fig. 1 illustrates an equilibrium magnetic field configuration for a stellarator geometry. The 3D geometry leads

* tobias.blickhan@nyu.edu

to a mix of magnetic surfaces and island chains.

The boundary conditions we use in this work are $B \cdot n = J \times n = 0$ on $\partial\Omega$, where n is the unit vector normal to the boundary. Physically, this means $\partial\Omega$ is a magnetic flux surface in vacuum. MRX handles more general boundary conditions of the form $(B - B_{\text{vac}}) \cdot n = 0$ as well, cf. Sections 4.2 and 8.2.

In this introduction section, we discuss the applications of the MHS problem in 1.1 and how different assumptions on the magnetic topology affect solution strategies in 1.2, 1.3, and 1.4. We comment on numerical evidence for the existence of three-dimensional equilibria in 1.5. Existing codes are presented in 1.6, our contributions in 1.7, and the outline of the remaining presentation in 2.

1.1. Applications

It is hard to overstate the importance of computing realistic magnetic equilibria. MHS equilibria form the foundation of the design of magnetic-confinement fusion devices by facilitating: (i) plasma state reconstructions from experimental data [1], (ii) MHD and other stability calculations [2], (iii) neoclassical [3] and gyrokinetic [4] transport calculations, and (iv) the initialization of extended MHD time-dependent codes such as JOREK3D [5], M3D-C1 [6] and NIMSTELL [7].

Interpreting and analyzing diagnostic data at every time snapshot requires reconstructing the current state of the equilibrium at each step. Besides computing physical quantities, these reconstructions are crucial for the performance of plasma control systems that are integral to fusion device design. These systems can allow for real-time instability and heat flux control [8].

1.2. Symmetry

Finding a solution to (1) is very challenging analytically and there are open questions about the existence of MHS solutions in toroidal geometry without the assumption of axisymmetry. The Grad conjecture [9] states that smooth MHS solutions with nested toroidal pressure surfaces can only exist in the presence of symmetry (axial, helical, or by reflection) and remains open [10, Conjecture 1].

An important class of fusion experiments, tokamaks, are based on magnetic confinement in axisymmetric toroidal geometry. In this case, nested flux surfaces are guaranteed, and can be labeled by a magnetic flux function ψ . Subsequently, the MHS equation (1) reduces to the Grad-Shafranov equation, a two-dimensional, elliptic, nonlinear partial differential equation. Various reliable solvers for the Grad-Shafranov equation are in use, see [11] and references therein.

1.3. Nested flux surfaces

Toroidal devices without axisymmetry are called stellarators. Nested flux surfaces are not guaranteed in a 3D toroidal volume Ω , and therefore the full 3D equations must be solved. Nonetheless, the most commonly used codes to compute stellarator equilibria, such as VMEC [12–14], GVEC [15, 16], and DESC [17] assume that nested magnetic flux surfaces exist and therefore the surfaces can be labeled by the toroidal magnetic flux ψ . The existence of nested flux surfaces implies the foliation of the plasma volume into nested tori. Magnetic field lines (i.e. integral curves of the vector field B) lie on these tori and $\text{grad } p$ is orthogonal to them, as can be seen from taking the dot product $B \cdot (J \times B - \text{grad } p) = 0$.

This implies also that the pressure $p = p_{\text{profile}}(\psi)$ and the rotational transform $\iota = \iota_{\text{profile}}(\psi)$ are flux functions. The rotational transform describes the number of poloidal rotations per single toroidal rotation of a magnetic field line. Nested flux surfaces are often a useful assumption because it simplifies the problem, establishes a convenient flux coordinate system, and reflects the expectation that stellarators with good confinement should exhibit large volumes of nested flux surfaces. Force balance generally can only hold in a very weak sense, as singular currents will appear at rational values of $\iota_{\text{profile}}(\psi)$ [18, 19]. This will be discussed in detail in Section 5.8.

1.4. Magnetic islands

As stellarator optimization and experiments increasingly scale to more realistic geometries representing fusion-scale devices, there is an increasing need for finding equilibria with islands and chaos. The former are closed tubes (that are non-nested with nearby flux surfaces) formed by magnetic field lines with toroidal geometry, the latter are volume-filling field lines. A five-fold island chain can be seen in Fig. 1, near the plasma edge.

Modeling magnetic islands and chaos is important for capturing the dynamics of real experiments, where these features often play a large part in transport and divertor operation [20]. The presence of uncontrolled magnetic islands and chaotic regions primarily leads to a decrease in the quality of confinement of the plasma. However, the experimental device Wendelstein-7X is engineered to take advantage of a particular magnetic island chain at its edge, paired with an island divertor to allow for a controlled release of heat from the plasma [21]. Sophisticated design of such island divertors [22] or non-resonant divertors [23, 24] is crucial, as the divertor nominally controls critical device properties including: the large heat fluxes to the material surfaces, impurity fluxes, plasma detachment, helium removal, etc [25, 26]. Resolving island regions can also help to initialize extended time-dependent MHD codes, since the formation of islands is very slow at fusion-relevant resistivity.

1.5. Numerical evidence for the existence of 3D toroidal MHD equilibria

There is substantial numerical and experimental evidence that the three-dimensional toroidal MHD equilibrium landscape is not effectively single-valued. Cooper et al. found both an axisymmetric equilibrium and a spontaneously bifurcated helical-core equilibrium for the same fixed boundary and prescribed pressure and rotational-transform profiles, with comparable MHD energies [27]. Panici et al. subsequently recovered both branches with DESC using deflation and force-balance errors below 1%, and also obtained 25 distinct non-axisymmetric near-axis-constrained equilibria that share the same magnetic axis, on-axis transform, and first-order quasisymmetry while differing substantially in magnetic shear and outer geometry [28]. For genuinely stellarator-shaped boundaries, Thun et al. constructed continuous families of low-force-residual equilibria at fixed boundary and rotational transform while varying only a pressure multiplier [29], while the MRxMHD variational formulation can provide family of relaxed three-dimensional equilibria selected by regional magnetic helicities, fluxes, and pressure jumps [30]. Experimentally, LHD exhibits hysteresis between healed and penetrated magnetic-island states as the resonant field, plasma beta, or poloidal flow is varied, indicating history-dependent long-lived states with different magnetic topology over overlapping parameter ranges [31]. Thus, although strict nonuniqueness for identical, fully specified non-axisymmetric stellarator data has not been established generally, the available evidence strongly supports multiple equilibria or long-lived near-equilibria separated by small changes in pressure, helicity-related constraints, topology, or residual force balance. We will illustrate strong numerical evidence for this hypothesis across the results presented in this manuscript.

1.6. Existing 3D MHD equilibrium codes

Essentially all 3D MHD equilibrium codes boil down to an optimization problem for minimizing the volume-integrated energy or volume-integrated square of the MHS residual, Equation (1).

We have already mentioned a class of 3D MHD equilibrium codes that assume nested flux surfaces. The solution is then determined by the flux coordinate system together with a handful of scalar functions, which becomes the optimization target (c.f. equation (26)).

a. Breaking nested flux surfaces Beyond this class of methods, there are a few other codes that can produce 3D toroidal MHD equilibria with islands, which fall broadly into the class of codes relying on MRxMHD with stepped pressure profiles (SPEC, SPECTRE [32, 33]), Picard-like fixed point iterations (PIES, HINT [34–36]) and magnetic relaxation codes (SIESTA [37], and the MRX code presented here). We give a summary of relevant 3D MHS

codes we are aware of in Table I. Some of these codes can perform free-boundary calculations, but the table summary and our focus in this work is on fixed-boundary equilibria.

The MRxMHD stepped pressure approach relies on solving Taylor relaxation problems in nested domains with pressure jumps at the interfaces. It has the advantage of reducing the problem to coupled Beltrami solves, but this is a very different physical model than magnetic relaxation, and convergence with respect to the number of subdomains is generally unclear in 3D toroidal geometry.

The PIES and HINT approaches are general approaches for non-nested equilibria, but each exhibit potential numerical disadvantages. PIES is a Picard iteration on the magnetic field, and such fixed-point schemes converge at a linear rate. In practice this rate degrades and under-relaxation of the current-profile updates is required for stable iteration [55]. A Newton variant of PIES [56] halves the iteration count under the assumption of nested flux surfaces.

HINT instead relaxes artificial dissipative MHD equations toward a steady state, alternating a pressure update at fixed B with a field update at fixed p . Since the field is advanced explicitly on an Eulerian grid, a CFL condition limits the step size while the relaxation must run over the much longer dissipative timescale, which is costly.

Out of the relaxation codes, SIESTA is the most similar code to that proposed in this work.

b. Added flexibility from finite elements Existing codes based on Fourier expansion in the angular variables do not support nonuniform angular meshes. In contrast, finite element (FE) codes can provide nonuniform meshes and local mesh refinement for resolving the small-scale features near magnetic islands (radial refinement) or divertor regions (poloidal refinement), while retaining a coarse representation elsewhere.

c. Structure preservation Classical magnetic relaxation comes from considering ideal admissible variations that monotonically minimize the energy and preserve the magnetic helicity and topology. However, many relaxation codes such as HINT and SIESTA cannot numerically guarantee that the divergence-free field properties and fixed magnetic helicity hold.

This brings into question in what sense numerical convergence is ever truly achieved. As we will discuss in Section 5.8, a perfect discrete representation of the continuous system is infeasible. At the same time, it is essential to have precise control of the approximation and errors made through discretization.

We propose to address this issue, as in recent work [40, 57], by using structure-preserving mixed FE methods.

1.7. Our contributions

In this work, we present a new numerical code to solve magnetic relaxation problems in realistic geometries us-

TABLE I: Three-dimensional MHD equilibrium codes for stellarator geometry, ordered by year of first publication. Codes restricted to Cartesian geometry used to study the Parker problem are omitted, such as [38–41]. Also omitted are codes for time-dependent MHD, e.g. JOEK3D [5], M3D-C1 [6] and NIMSTELL [7].

Code	Repository	Year	Method	Discretization	Back-end	Islands?
BETA / BETAS [42–45]	–	1978	relaxation (momentum)	FD $\times$ FR ²	Fortran	
└ NSTAB [46]	–	1992	energy minimization	FD $\times$ FR ²	Fortran	
Chodura & Schlüter [47]	–	1981	relaxation (CG)	FD (not field-aligned)	Fortran	✓
└ NEAR [48]	–	1985	relaxation (CG)	FD $\times$ FR ²	Fortran	
VMEC [12, 49]	gh:hiddenSymmetries/VMEC2000	1983	energy minimization	FD $\times$ FR ²	Fortran	
└ VMEC++ [13]	gh:proximafusion/vmecpp	2025	energy minimization	FD $\times$ FR ²	C++	
└ VMEX	gh:uwplasma/vmex	2026	energy minimization	FD $\times$ FR ²	JAX	
PIES [50]	–	1986	field-line Picard	FD $\times$ FR ²	Fortran	✓
HINT [51, 52]	–	1989	resistive time-stepping	FD (not field-aligned)	Fortran	✓
└ HINT2 [35, 36]	–	2006	resistive time-stepping	FD (not field-aligned)	Fortran	✓
SIESTA [37]	gh:ORNL-Fusion/SIESTA	2011	relaxation (Newton)	FD $\times$ FR ²	Fortran	✓
SPEC [53]	gh:PrincetonUniversity/SPEC	2012	stepped pressure	Chebyshev $\times$ FR ²	Fortran	✓
└ SPECTRE [33]	gl:spectre-eq/spectre	2026	stepped pressure	Chebyshev $\times$ FR ²	Numba	✓
DESC [17, 54]	gh:PlasmaControl/DESC	2020	nested collocation	Zernike $\times$ FR	JAX	
GVEC [15]	gh:gvec-group/gvec	2025	energy minimization	B-splines $\times$ FR ²	Fortran	
This work	gh:ToBlick/mrx	2026	relaxation (Newton)	B-spline FE	JAX	✓

ing a mixed finite element method built on top of the JAX computational framework.

a. Goals and scope The eventual goal is to design a new MHD equilibrium solver that can: (1) produce robust 3D MHS solutions with islands and chaos via magnetic relaxation, (2) scale on modern GPUs and provide differentiable objectives, and (3) numerically reflect the relevant structural properties to desired precision by using structure-preserving finite elements.

In this initial work, we propose only to solve the problem where $B \cdot n = 0$ is prescribed on a fixed computational boundary. The focus is on the mathematical properties and convergence of the code.

We focus on the case of solid toroidal geometry. There are no structural obstacles to transfer this work to arbitrarily-shaped domains, including ones with multiple cavities and tunnels.

Compared to previous magnetic relaxation methods for MHS, our approach also differs in the way the pressure is treated. We follow here the approach common in the hydrodynamics literature [58], cf. Section 5.4.

b. Structure preserving finite elements As will be discussed in Section 6, the finite element framework we employ guarantees the preservation of crucial features of the continuous problem after discretization. Among these are the preservation of a divergence-free magnetic field to machine accuracy, as well as helicity preservation and an energy dissipation equality to the order of nonlinear solver tolerances.

Computational electrodynamics in general and magnetohydrodynamics in particular are rich with geometric structure, as we will discuss in Section 3. It is by now well-understood how this structure can be retained in the corresponding discrete problems thanks to advances in mixed finite element methods [59, 60]. Even seem-

ingly benign problems such as the computation of a vector potential A such that $\text{curl } A = B$ given B can pose difficulties after discretization [60, Chapter 5].

Preserving these quantities requires a suitable choice of discrete vector spaces. While it is impossible to preserve all features of the continuous problem in a finite-dimensional approximation, it is possible to use approximation spaces where some identities (e.g. as $\text{div } B = 0$) hold to machine precision, while others (e.g. $J = \text{curl } B$) hold only up to the order of the scheme.

Remark. It is worth noting that the equations that hold only approximately in the discrete approximation are those that only hold approximately in nature, too. In other words, $\text{div } B = 0$ and $\text{curl } H = J$ can be considered exact laws of nature, while $\mu_0 H = B$ is a constitutive law - the value of μ_0 is determined experimentally and this linear relation is the zeroth order approximation of the magnetization of the plasma (this is the Bohr–Van Leeuwen theorem [61, §52] – the approximation is a very good one in real plasmas).

Structure-preserving FE methods have been applied to model MHD phenomena in a number of works [62–67] and also in magnetic relaxation codes [40, 41, 57]. However, to our knowledge, there are no works that tackle the relaxation problem in the practically relevant toroidal geometry using structure-preserving finite elements.

c. Code framework We built the codebase for this work on the JAX framework [68, 69], a tracing just-in-time compiler for generating high-performance accelerator code from pure Python and Numpy programs. The benefit of this is threefold.

(1) The compiled code is highly performant on CPUs, GPUs, and TPUs. It allows us to run MRX on the latest computational hardware available by leveraging the abstractions of the compilers in the OpenXLA framework.

(2) JAX supports automatic differentiation, i.e. the computation of gradients of functions with regard to their input arguments by tracing the primitive operations encountered throughout function evaluation. This allows, for example, the computation of derivatives of equilibrium fields with regard to geometry inputs without resorting to costly finite difference approximations. This is a very attractive feature for PDE-constrained optimization and other applications, e.g. incorporation of MRX into stellarator optimization routines, a subject of future work.

(3) Our code is open-source and highly accessible as all dependencies can be installed via the `pip` package manager after cloning the code repository¹. We also have interfaces for initializing equilibria from GVEC, VMEC, and DESC solutions, and initializing coordinate mappings from `map2disc` [70] instead of a nested flux surface solution.

2. OUTLINE

In this document, we will introduce the theoretical concepts one-by-one, followed by numerical experiments to demonstrate and validate their impact on the code. The following list covers the remaining sections of this work.

- 3 Notation and formalization of the problem statement.
- 4 Numerical verification of vacuum equilibria. These problems can be solved by eigendecomposition. No relaxation is involved.
- 5 The magnetic relaxation algorithm in continuous form and its structure-preserving discretization. (Non)convergence to equilibria in the continuous and the discrete problem. Newton-type descent methods.
- 6 Numerical study on the performance and impact of various hyperparameters in the relaxation algorithm.
- 7 Reconnection events, purposeful or spurious. Numerical studies that show the emergence of large islands in the computed equilibria.
- 8 Summary and pointers to future work.

3. FUNCTION SPACES AND THEIR APPROXIMATION

At the center of MRX stands the de Rham complex of function spaces and its faithful discrete approximation. This approximation yields important tools such as

the discrete Hodge-Helmholtz decomposition, the crucial identities $\text{curl grad} = \text{div curl} = 0$, discretely exact, and the identification of harmonic fields. The latter constitute the solutions to MHD equilibria with $J = p = 0$ which are called vacuum solutions.

A thorough treatment of the mathematical theory is given in [59, Section 2.2] or [71, Chapter 7.5].

We will discuss the properties of the de Rham complex in 3.1, the encoding of geometry in 3.2, and the crucial Hodge-Helmholtz decomposition in 3.3. In 3.4 and 3.5, we present the discretization strategy and some practical implementation choices.

3.1. The de Rham complex

Let Ω denote a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$. Denote the boundary of Ω by $\partial\Omega$ and the outward unit normal of a vector at $x \in \partial\Omega$ as $n(x)$.

The Hilbert space of square-integrable, vector-valued functions on this domain is denoted $L^2(\Omega; \mathbb{R}^3)$. We equip all of the following function spaces with the standard inner product, the induced norm, and we write the volume average of a function with the same bracket. For $u, v \in L^2(\Omega; \mathbb{R}^3)$,

$$\langle u, v \rangle_\Omega := \int_\Omega u \cdot v \, dx, \quad \|u\|_\Omega := \sqrt{\langle u, u \rangle_\Omega}. \quad (2)$$

The subscript names the domain of integration; the point-wise norm $|u|^2(x)$ is denoted with single lines. We also introduce the volume average for $f \in L^1(\Omega)$

$$\langle f \rangle_\Omega := \frac{1}{|\Omega|} \int_\Omega f \, dx. \quad (3)$$

Errors against a reference solution u^* are reported relative to it, $\|u - u^*\|_{\text{rel}} := \|u - u^*\|_\Omega / \|u^*\|_\Omega$. The magnetic pressure is $|B(x)|^2/2$ and the magnetic energy $\mathcal{E}(B) = \|B\|_\Omega^2/2$; in every relaxation run the initial field is normalized to $\|B\|_\Omega = 1$, so that $\mathcal{E} = 1/2$ and $\langle |B|^2 \rangle_\Omega = 1/|\Omega|$. The plasma β is the ratio of the kinetic to the magnetic pressure: $\beta_{\text{vol}} := \langle p \rangle_\Omega / \langle |B|^2/2 \rangle_\Omega = 2\langle p, 1 \rangle_\Omega / \|B\|_\Omega^2$. We report pressure values normalized by magnetic pressure, i.e. $p_{\text{norm}}(x) = p(x) / \langle |B|^2/2 \rangle_\Omega$.

The force residual is normalized as in GVEC [16] as

$$\|F\|_{\text{norm}}^2 := \frac{\|J \times B - \text{grad } p\|_\Omega^2}{\|\text{grad } |B|^2/2\|_\Omega^2}. \quad (4)$$

Definition 1 (Function spaces). *The spaces of vector fields on Ω with weak grad, curl and divergence are defined as:*

$$H^1(\Omega; \mathbb{R}) := \{p \in L^2(\Omega; \mathbb{R}) : \text{grad } p \in L^2(\Omega; \mathbb{R}^3)\}, \quad (5)$$

$$H^{\text{curl}}(\Omega; \mathbb{R}^3) := \{E \in L^2(\Omega; \mathbb{R}^3) : \text{curl } E \in L^2(\Omega; \mathbb{R}^3)\}, \quad (6)$$

$$H^{\text{div}}(\Omega; \mathbb{R}^3) := \{B \in L^2(\Omega; \mathbb{R}^3) : \text{div } B \in L^2(\Omega; \mathbb{R})\}. \quad (7)$$

¹ github.com/ToBlick/mrx

The corresponding spaces of vector fields with homogeneous Dirichlet boundary conditions are defined as:

$$H_0^1(\Omega; \mathbb{R}) := \{p \in H^1(\Omega; \mathbb{R}) : p|_{\partial\Omega} = 0\}, \quad (8)$$

$$H_0^{\text{curl}}(\Omega; \mathbb{R}^3) := \{E \in H^{\text{curl}}(\Omega; \mathbb{R}^3) : E \times n|_{\partial\Omega} = 0\}, \quad (9)$$

$$H_0^{\text{div}}(\Omega; \mathbb{R}^3) := \{B \in H^{\text{div}}(\Omega; \mathbb{R}^3) : B \cdot n|_{\partial\Omega} = 0\}. \quad (10)$$

We will, from now on, write $L^2(\Omega)$ for short for both vectorial and scalar spaces.

Remark. The expressions $B \cdot n|_{\partial\Omega} = 0$ should strictly speaking be understood in the sense of a trace operator, see [60, Lemma 2.1.1].

The 3D de Rham complex can be written as:

$$0 \rightarrow H^1 \xrightarrow{\text{grad}} H^{\text{curl}} \xrightarrow{\text{curl}} H^{\text{div}} \xrightarrow{\text{div}} L^2 \rightarrow 0. \quad (11)$$

It is a special case of a closed Hilbert complex

$$0 \rightarrow V^0 \xrightarrow{d^0} V^1 \xrightarrow{d^1} V^2 \xrightarrow{d^2} V^3 \rightarrow 0 \quad (12)$$

and satisfies the following properties for all k : (i) $\text{range } d^k \subset V^{k+1}$, (ii) $d^{k+1} \circ d^k = 0$, and (iii) the range of d^k is closed in V^{k+1} [72, Section 3.1.3].

For the de Rham complex, (ii) describes the central vector calculus identities $\text{curl grad} = \text{div curl} = 0$. We anticipate already that (iii) is crucial to retain during discretization in order to arrive at a well-posed discrete problem [60, Theorem 4.15].

We call elements of the space V^k a k -form. More details will be given in Section 3.3.

Each d^k has a Hilbert adjoint, the *co-differential* $d_{k+1}^* : V^{k+1} \rightarrow V^k$, defined by

$$\langle d^k u, v \rangle_{V^{k+1}} = \langle u, d_{k+1}^* v \rangle_{V^k}, \quad (13)$$

for all $u \in V^k$, $v \in V^{k+1}$ where $V_*^{k+1} \subset V^{k+1}$ is the domain of d_{k+1}^* . These domains are proper subspaces that carry the boundary conditions the differentials do not: for the de Rham complex (11), $d_1^* = -\text{div}$ on $H_0(\text{div})$, $d_2^* = \text{curl}$ on $H_0(\text{curl})$, and $d_3^* = -\text{grad}$ on H_0^1 . The co-differentials form the dual complex

$$0 \leftarrow V^0 \xleftarrow{d_1^*} V^1 \xleftarrow{d_2^*} V^2 \xleftarrow{d_3^*} V^3 \leftarrow 0, \quad (14)$$

which is again a closed Hilbert complex: $d_k^* \circ d_{k+1}^* = 0$ and $\text{range } d_{k+1}^* = (\ker d^k)^\perp$ is closed.

3.2. Domain and mapping

The domain of interest in this work is $\Omega \subset \mathbb{R}^3$, the volume of space that encloses the plasma. This domain is typically shaped in a moderately complicated manner, cf. Fig. 1. We will define approximation spaces in the logical domain, and therefore need to define appropriate operations to translate between the logical and physical domains.

a. Coordinate systems The approach we take in this work is to describe Ω as the image of the logical domain $\hat{\Omega} = [0, 1]^3$ after application of the mapping $\Phi : \hat{\Omega} \rightarrow \mathbb{R}^3$. We will denote the coordinates on the logical domain by $(r, \theta, \zeta) = \hat{x} \in \hat{\Omega}$. The map Φ is assumed to be a C^1 diffeomorphism everywhere except at $r = 0$. The Jacobian matrix of the mapping Φ is defined as $D\Phi$ with entries $(D\Phi(\hat{x}))_{ij} = \partial\Phi_i(\hat{x})/\partial\hat{x}_j$.

The columns of $D\Phi(\hat{x})$ are parallel to the unit vectors in the (r, θ, ζ) directions at position $\hat{x}$.

We denote $\Phi(\hat{x}) = x = (x_1, x_2, x_3)$, these are points in physical space. We also introduce a cylindrical coordinate system $(R, \phi, Z) = (\sqrt{x_1^2 + x_2^2}, \arctan_2(x_2, x_1), x_3)$.

b. Pull-back and push-forward Points in the logical domain $\hat{x} \in \hat{\Omega}$ and functions defined in logical coordinates such as $\hat{f} : \hat{\Omega} \rightarrow \mathbb{R}$ carry a hat. The hat decorator is omitted on coordinate names r, θ, ζ , the map Φ with its Jacobian $D\Phi$ and basis functions Λ_i^k . Through the following push-forward and pull-back operations, we can associate to every function $\hat{f}$ on the logical domain a function $f : \Omega \rightarrow \mathbb{R}$ on the physical domain.

Definition 2 (Push-forward). *For functions $\hat{f}, \hat{\rho} : \hat{\Omega} \rightarrow \mathbb{R}$ we define two push-forward operations to the physical domain as follows: For all $x \in \Omega$,*

$$f(x) := (\Phi_*^0 \hat{f})(x) := \hat{f}(\Phi^{-1}(x)), \quad (15)$$

$$\rho(x) := (\Phi_*^3 \hat{\rho})(x) := \frac{\hat{\rho}(\Phi^{-1}(x))}{\det D\Phi(\Phi^{-1}(x))}. \quad (16)$$

We furthermore define two push-forward operators on vector fields $\hat{E}, \hat{B} : \hat{\Omega} \rightarrow \mathbb{R}^3$ as:

$$E(x) := (\Phi_*^1 \hat{E})(x) := (D\Phi(\Phi^{-1}(x)))^{-T} \hat{E}(\Phi^{-1}(x)), \quad (17)$$

$$B(x) := (\Phi_*^2 \hat{B})(x) := \frac{D\Phi(\Phi^{-1}(x)) \hat{B}(\Phi^{-1}(x))}{\det D\Phi(\Phi^{-1}(x))}. \quad (18)$$

The corresponding pull-back operations are obtained when replacing the mapping Φ with its inverse Φ^{-1} .

Lemma 3 ([71], Theorem 6.4.4). *The operations grad, div and curl are natural with respect to the push-forward under C^1 -diffeomorphisms, that is, the push-forward of the gradient/curl/divergence is the gradient/curl/divergence of the push-forward.*

Note that even when Φ does not describe an orthogonal mapping, it crucially retains boundary conditions. The proof is given in Appendix A.

Lemma 4 (Pull-backs preserve boundary conditions). *When $\hat{f} \in H_0^1(\hat{\Omega})$, $\hat{E} \in H_0^{\text{curl}}(\hat{\Omega})$, and $\hat{B} \in H_0^{\text{div}}(\hat{\Omega})$, then $\Phi_*^0 \hat{f} \in H_0^1(\Omega)$, $\Phi_*^1 \hat{E} \in H_0^{\text{curl}}(\Omega)$, and $\Phi_*^2 \hat{B} \in H_0^{\text{div}}(\Omega)$.*

c. Stellarator geometry All calculations in this work take place in solid toroidal geometry. All mappings share a polar singularity, i.e.

$$\Phi(r, \theta, \zeta) \xrightarrow{r \rightarrow 0} x_{\text{axis}}(\zeta) \quad \forall \theta. \quad (19)$$

The limit does not depend on θ so the Jacobian of the map is not invertible at $r = 0$. The handling of this singularity is discussed in Section 3.4.

Given functions $R : \hat{\Omega} \rightarrow \mathbb{R}_{>0}$ and $Z : \hat{\Omega} \rightarrow \mathbb{R}$ (the cylindrical coordinates of the image point, hence the letters), we obtain a basic stellarator map as follows:

$$\Phi : \begin{bmatrix} r \\ \theta \\ \zeta \end{bmatrix} \mapsto \begin{bmatrix} R(r, \theta, \zeta) \cos(2\pi\zeta/n_{\text{fp}}) \\ -R(r, \theta, \zeta) \sin(2\pi\zeta/n_{\text{fp}}) \\ Z(r, \theta, \zeta) \end{bmatrix}, \quad (20)$$

so that $\hat{\Omega}$ covers one of the n_{fp} field periods and all fields are n_{fp} -periodic in the toroidal angle.

In this work, the functions R and Z are represented using spline discretizations introduced in Section 3.4.

Remark. More general maps using e.g. map2disc [70] or maps with additional symmetries (e.g. stellarator symmetry) have also been implemented in the MRX code.

d. Interface with nested flux surface solvers The functions $R, Z : \hat{\Omega} \mapsto \mathbb{R}$ are typically obtained from nested-flux surface solvers such as VMEC and GVEC. In MRX, the map is represented as a C^1 polar B-spline map on the 0-form basis of the de Rham sequence.

We emphasize here that we do not assume that Φ describes a flux coordinate system. All that is required is that $\Phi(\hat{\Omega}) = \Omega$. In general, the magnetic axis and the axis of the coordinate system, $\Phi(\hat{x})|_{r=0}$, need not coincide.

For every equilibrium solve, the mapping Φ is held constant. In contrast, for codes that operate under a nested flux surface assumption, the mapping is itself the optimization objective and changes during solving for equilibrium. Anticipating future use of this code for stellarator optimization (in which the boundary shape and therefore the mapping changes each iteration), we have also reimplemented map2disc's main functionality within MRX, with differentiable JAX code. Future work will focus on incorporating MRX into stellarator optimization routines.

It is of course possible to use such an optimized map for Φ . Doing so is beneficial for several reasons: as we expect features of the solution to align with the grid and we can use the physics to derive preconditioning strategies in flux-aligned coordinates.

We show some of the grids employed in MRX in the illustrative Fig. 2.

3.3. Hodge decomposition

The L^2 -orthogonal Hodge-Helmholtz decomposition expresses a vector field as the sum of a gradient field, a divergence-free field, and a harmonic component. In fact, vector fields in $L^2(\Omega)$ admit two L^2 -orthogonal Hodge-Helmholtz decompositions [71, Section 7.5.5], corresponding to the decomposition of V^1 and V^2 intro-

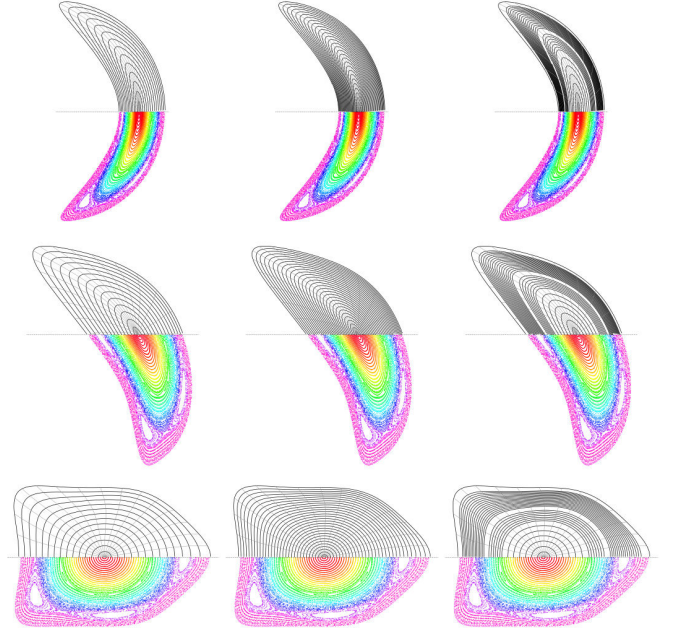


FIG. 2: Different meshes that can be used in MRX at $\zeta = \text{const.}$ slices. We show here also Poincaré sections of the reconnected field on these meshes (c.f. Section 7). The island width agrees to 2% across the three meshes.

duced in Section 3.1:

$$\begin{aligned} L^2(\Omega) &= \text{grad } H^1(\Omega) \oplus \text{curl } H_0^{\text{curl}}(\Omega) \oplus \mathfrak{H}^1(\Omega) \\ &= \text{curl } H^{\text{curl}}(\Omega) \oplus \text{grad } H_0^1(\Omega) \oplus \mathfrak{H}^2(\Omega). \end{aligned} \quad (21)$$

The orthogonality relations here are understood with respect to the L^2 inner product.

In terms of differential and co-differential, the Hodge decomposition reads

$$\begin{aligned} V^k &= \text{range } d^{k-1} \oplus \text{range } d_{k+1}^* \oplus \mathfrak{H}^k, \\ \mathfrak{H}^k &= \ker d^k \cap \ker d_k^*. \end{aligned} \quad (22)$$

The two decompositions of vector fields in $L^2(\Omega)$ in particular correspond to 1-forms and 2-forms.

Definition 5 (Harmonic forms). *The spaces $\mathfrak{H}^k$, $k \in \{0, 1, 2, 3\}$, are called harmonic spaces; elements $\mathfrak{h}_k$ of $\mathfrak{H}^k(\Omega)$ are harmonic forms. The L^2 -projection onto $\mathfrak{H}^k$ is denoted by $\Pi^{\mathfrak{H}^k}$. The harmonic spaces of the de Rham complex with homogeneous boundary conditions are denoted by $\mathfrak{H}_0^k(\Omega)$.*

Lemma 6 ([71, Corollary 7.5.4]). *$\text{grad } \mathfrak{h}_0 = 0$ in Ω : the elements of $\mathfrak{H}^0(\Omega)$ are the functions constant on every connected component of Ω . The elements of $\mathfrak{H}^1(\Omega)$ and $\mathfrak{H}^2(\Omega)$ are vector fields with $\text{curl } \mathfrak{h} = \text{div } \mathfrak{h} = 0$, with zero normal trace in $\mathfrak{H}^1(\Omega)$ and zero tangential trace in $\mathfrak{H}^2(\Omega)$, and $\mathfrak{H}^3(\Omega) = \{0\}$. Moreover $\dim \mathfrak{H}^k(\Omega) = \dim \mathfrak{H}_0^{3-k}(\Omega)$, the two spaces being isomorphic.*

All harmonic forms lie in the kernel of the Hodge-Laplace operator with corresponding boundary conditions.

Definition 7 (Hodge-Laplace operator). *The Hodge-Laplacian on $L^k : V^k \rightarrow V^k$ is*

$$L^k = d_{k+1}^* d^k + d^{k-1} d_k^*, \quad (23)$$

with domain $\{u \in V^k \cap V_*^k : d^k u \in V_*^{k+1}, d_k^* u \in V^{k-1}\}$.

For the de Rham complex (11), L^0 and L^3 are the scalar Laplacian $-\text{div grad}$, while L^1 and L^2 are the vector Laplacians $\text{curl curl} - \text{grad div}$. The boundary conditions differ due to the domains of the co-differentials.

By construction L^k is self-adjoint and positive semi-definite, $\langle L^k u, u \rangle_\Omega = \|d^k u\|_\Omega^2 + \|d_k^* u\|_\Omega^2$, its kernel is $\mathfrak{H}^k$ by definition, and it is invertible on $(\mathfrak{H}^k)^\perp$ by the Hodge decomposition (22).

As a consequence of de Rham's theorem [71, Section 7.5], the dimensions of the harmonic spaces are the Betti numbers of the domain. That is, $\dim \mathfrak{H}^0(\Omega)$ is the number of connected components of Ω (always one in this work), $\dim \mathfrak{H}^1(\Omega)$ is the number of tunnels (or handles) in Ω , and $\dim \mathfrak{H}^2(\Omega)$ is the number of cavities.

For a solid torus, the form of all domains in this work, $\dim \mathfrak{H}^1(\Omega) = 1$ and $\dim \mathfrak{H}^2(\Omega) = 0$. In the plasma physics community, elements of $\mathfrak{H}_0^0$ are usually referred to as vacuum fields; they describe magnetic fields with zero normal boundary trace that induce no current. For an axisymmetric toroidal domain, $\mathfrak{h}_1 \propto R^{-1} e_\phi$.

3.4. Discretization

MRX is built on B-spline basis functions defined on the logical domain $\hat{\Omega}$, mapped to the physical domain Ω via Φ .

a. Finite element exterior calculus The finite element framework we will follow, known as Finite Element Exterior Calculus (FEEC), crucially retains important features of the de Rham complex introduced in Section 3 after discretization such as the identities $\text{curl grad} = \text{div curl} = 0$. These relations hold to machine precision at any discretization resolution, they are topological (incidence) operations. Furthermore, this approach comes with certain guarantees about numerical stability, in particular, the inf-sup stability criterion for various saddle point problems that arise is satisfied [72, Theorem 3.8] and the approximation spaces can be decomposed via a discrete Hodge-Helmholtz decomposition.

Various excellent references exist on the theory behind FEEC and its connection to cohomology theory [59, 72]. The exposition we provide here for the sake of self-containedness will be presented on the level of the matrix-vector equations that arise after discretization.

The k -forms ($k \in \{0, 1, 2, 3\}$) are represented as piecewise polynomial functions valued in $\mathbb{R}$ (for $k \in \{0, 3\}$) and

$\mathbb{R}^3$ (for $k \in \{1, 2\}$). We denote these discrete function spaces by $V_h^k, k \in \{0, 1, 2, 3\}$.

These spaces are constructed such that the images of the (continuous) grad, curl and div operators are contained in the subsequent space: $f_h \in V_h^0$ implies $\text{grad } f_h \in V_h^1$, $A_h \in V_h^1$ implies $\text{curl } A_h \in V_h^2$, and $B_h \in V_h^2$ implies $\text{div } B_h \in V_h^3$.

Example 8 (Vector spaces that do not form a complex). *This fact importantly does not hold in any FE method. Consider, for example the space V_h^0 consisting of piecewise linear functions and $V_h^{\text{vec}} = V_h^0 \otimes V_h^0 \otimes V_h^0$, a vector-valued approximation space where every component is an element of V_h^0 . For $f \in V_h^0$, $\text{grad } f$ is a discontinuous function, hence $\text{grad } f \notin V_h^{\text{vec}}$.*

The spaces with (essential) homogeneous boundary conditions are denoted $V_0^k \forall k \in \{0, 1, 2, 3\}$. The basis functions that span V^k are denoted $\{\Lambda_i^k\}_i$. We denote the L^2 -projection operator to the space V^k by $\Pi^k : \Pi^k u = u_h := \sum_i u_i \Lambda_i^k$ where $\sum_j u_j \langle \Lambda_i^k, \Lambda_j^k \rangle_\Omega = \langle u, \Lambda_i^k \rangle_\Omega \forall i$.

b. Spline finite elements In this work, the spaces $\{V^k\}_k$ are spanned by a cartesian product of B-splines. B-splines are defined by their order p and a vector of knot points. They are piecewise polynomial functions, C^∞ between knot points and C^{p-m} at knot points, where m is the multiplicity of the knot point. In particular, given degrees $(p_r, p_\theta, p_\zeta) \in \mathbb{N}^3$, resolutions $(n_r, n_\theta, n_\zeta) \in \mathbb{N}^3$ such that $n_c > p_c, c \in \{r, \theta, \zeta\}$, let $\{N_{c,i}^0\}_{i=1}^{n_c}$ denote one-dimensional B-splines of degree p_c . The basis functions $\{\Lambda_i^0\}_{i=1}^{n_r n_\theta n_\zeta}$ are then defined via cartesian products and a flattening of their indices via $\text{vec} : \mathbb{N}^3 \rightarrow \mathbb{N}, (j, k, l) \mapsto i$. Evaluated at the point $\hat{x} = (r, \theta, \zeta)$, this yields $\Lambda_i^0(\hat{x})|_{i=\text{vec}(j,k,l)} = N_{r,j}^0(r) N_{\theta,k}^0(\theta) N_{\zeta,l}^0(\zeta)$. When we denote the space spanned by one-dimensional B-splines of degree p by S^p ,

$$V_h^0 = \bigotimes_{c \in \{r, \theta, \zeta\}} S^{p_c}, \quad V_h^1 = \begin{bmatrix} S^{p_r-1} \otimes S^{p_\theta} \otimes S^{p_\zeta} \\ S^{p_r} \otimes S^{p_\theta-1} \otimes S^{p_\zeta} \\ S^{p_r} \otimes S^{p_\theta} \otimes S^{p_\zeta-1} \end{bmatrix},$$

$$V_h^2 = \begin{bmatrix} S^{p_r} \otimes S^{p_\theta-1} \otimes S^{p_\zeta-1} \\ S^{p_r-1} \otimes S^{p_\theta} \otimes S^{p_\zeta-1} \\ S^{p_r-1} \otimes S^{p_\theta-1} \otimes S^{p_\zeta} \end{bmatrix}, \quad V_h^3 = \bigotimes_{c \in \{r, \theta, \zeta\}} S^{p_c-1}. \quad (24)$$

Here n_c is the number of splines of V^0 in direction c : in the periodic angles this is also the number of cells, while the clamped radial direction has $n_r - p$ cells. Throughout this work, we use $h_r := 1/n_r$ to denote a characteristic cell size.

For more detailed references regarding spline finite elements in computational electromagnetics, we refer to [73, 74].

Remark. In the interest of avoiding notational clutter, we will drop the subscript $_h$ in the following.

c. Polar splines The problems we consider occur naturally in toroidal geometry. As a result, the mapping Φ is singular at $r = 0$. We modify the cartesian product structure of the spline basis functions because regularity

requirements on the axis imply linear dependence of some of the splines with support at $r = 0$: The pull-back of any (physical) 0-form f will lead to $\hat{f} : \partial_\theta \hat{f}|_{r=0} = \text{constant}$. The requirement that all elements of V^0 satisfy this relation leads to a number of linear constraints, analogously for $k > 0$. A detailed reference for these constructions is [75, Chapter 5].

The number of splines affected depends on the desired regularity around the axis. Polar splines are introduced in [76–78] and used for MHD simulations in [66]. In the interest of brevity, we refer to these references for further details.

Example 9 (Continuously differentiable polar splines). *For elements of V^0 and V_0^0 , C^0 continuity demands that in every plane of constant ζ , the coefficients of all n_θ splines that are clamped at $r = 0$, the inner ring of basis functions, are equal. Hence, $\dim V^{\text{polar},0} = ((n_r - 1)n_\theta + 1)n_\zeta$. To obtain C^1 regularity, the two inner rings of basis functions are affected, a total of $2n_\theta$ functions. In this case, three C^1 -compatible basis functions are added to replace them, and $\dim V^{\text{polar},0} = ((n_r - 2)n_\theta + 3)n_\zeta$.*

In order to use polar splines in toroidal domains, it is sufficient that there exists a smooth diffeomorphism between Ω and a circular axisymmetric torus.

This is the case for every geometry produced by a nested flux surface solver such as VMEC or GVEC. These solvers represent the domain as a family of poloidal cross-sections swept around the torus, each the smooth image of the unit disc, and their nested-surface ansatz requires $\det D\Phi$ not to change sign for $r > 0$, so Φ is a local diffeomorphism away from the axis. It is injective because distinct flux surfaces do not intersect and the plasma does not intersect itself.

All maps used in this work are represented by the same splines that are used to span V^0 .

3.5. Implementation

We now give a very brief overview over some implementation choices.

a. Matrices By rescaling the splines of lower order in V^k , $k > 0$, it is possible to ensure that the degree-of-freedom to degree-of-freedom representation of grad, curl, and div operators are incidence matrices, i.e. they only hold entries ± 1 , independent of the geometry. We will denote these matrices by $\mathbb{D}_k$ ($k = 0$: grad, $k = 1$: curl, $k = 2$: div). It holds that $\mathbb{D}_{k+1}\mathbb{D}_k = 0$, independent of geometry and mesh size.

Furthermore, we denote the k -form mass matrix by $\mathbb{M}_k$ with entries $(\mathbb{M}_k)_{ij} = \langle \Lambda_i^k, \Lambda_j^k \rangle_\Omega$. The mass matrices depend on the metric through the pull-back operation, as they must. The integrals over Ω in this definition are pulled back to integrals over $\hat{\Omega}$ and evaluated using Gauss-Legendre quadrature in every knot interval (cell) of the spline spaces. The quadrature grid thus mimics the tensor-structure of the basis functions.

All operators (except the second variation bilinear form in the last section 5.9) in this work can be expressed as combinations of these basic matrices. For example, stiffness-like matrices are given by $\mathbb{D}_k^T \mathbb{M}_{k+1} \mathbb{D}_k$.

For $k = 0$, this represents the bilinear form $\langle \text{grad } \Lambda_i^0, \text{grad } \Lambda_j^0 \rangle_\Omega$, for $k = 1$ it is a curl-curl operator, and for $k = 2$, it is the bilinear form corresponding to $-\text{grad-div}$.

A weak derivative operator that maps k -form degrees of freedom to $k - 1$ -form degrees of freedom is expressed by $(-1)^k \mathbb{M}_{k-1}^{-1} \mathbb{D}_{k-1}^T \mathbb{M}_k$; we write it with a tilde, $\widetilde{\text{grad}}$, $\widetilde{\text{curl}}$, $\widetilde{\text{div}}$ for the three weak derivatives. It is called weak as the divergence/curl/gradient only holds in a weak L^2 sense, while the incidence relations (gradient of a 0-form is a 1-form, curl of a 1-form is a 2-form, divergence of a 2-form is a 3-form) hold point-wise.

The full Hodge-Laplace operators combine weak and strong derivatives, reflecting the differential/co-differential formulation of the continuous Hodge Laplacian:

$$\mathbb{L}_k := \mathbb{D}_k^T \mathbb{M}_{k+1} \mathbb{D}_k + \mathbb{M}_k \mathbb{D}_{k-1} \mathbb{M}_{k-1}^{-1} \mathbb{D}_{k-1}^T \mathbb{M}_k, \quad (25)$$

where it is understood that terms with $k < 0$ or $k > 3$ are omitted.

This discrete form exactly resembles the definition of L^k as $L^k = d_{k+1}^* d^k + d^{k-1} d_k^*$.

b. Matrix-free implementation None of these matrices are ever assembled. They are applied through the tensor-product structure of the spline bases: a k -form is evaluated on the $(p+1)^3$ Gauss-Legendre quadrature points of every cell by three one-dimensional contractions, multiplied by the metric coefficients stored on the quadrature grid, and contracted back. The cost of one operator apply is therefore proportional to the number of quadrature points, and the memory footprint to the geometry stored on them, a few floats per point from the symmetric metric tensor (6 unique entries), its inverse (6), and the Jacobian determinant (1).

c. Solvers and preconditioners Every linear system is solved by a Krylov method on the matrix-free operators: preconditioned conjugate gradients for symmetric positive definite operators, MINRES for indefinite symmetric operators.

We apply a custom-built preconditioner exploiting the sequence properties and tensor structure of our approximation space. This preconditioning strategy will be explained in a forthcoming publication.

The stopping criterion of every solve is the residual measured in the norm of its space, where the mass matrix preconditioner serves as a cheap proxy for the inverse mass-weighted norm.

The Hodge-Laplacian is only invertible on the subspace orthogonal to harmonic forms. We hence first compute the harmonic forms for every geometry and then deflate them from every solve, i.e. project both the right-hand side and solution of $L^k u = f$ to be orthogonal to every harmonic form.

d. Precision The code runs in three modes. In float64 every operation is in double precision. In float32, used for computational efficiency and needed for running on specialized architectures like TPUs, the Krylov iterations run in single precision and the residual can only be resolved to $\sqrt{\epsilon_{32}} \approx 3.5 \times 10^{-4}$. In mixed mode, the Krylov iterations run in float32 while the residual of the outer iterative refinement is formed in float64 and the correction is accumulated in float64, which reaches float64 tolerances at float32 efficiency.

Default tolerances are 10^{-10} in float64, 10^{-8} in mixed precision and 3.5×10^{-4} in float32.

Remark. All timings cited are on a single H100 GPU. Note that setup times involve host-side compilation, which might be slowed by concurrent jobs taking up resources.

4. VACUUM NUMERICS

The first verification experiment we perform concerns the computation of a harmonic (vacuum) field in stellarator geometry, described in 4.1. We verify the computations against an analytical solution in 4.2 and one computed from VMEC in 4.3.

4.1. Geometry

Both GVEC and VMEC supply $(r, \theta, \zeta) \mapsto (R, Z)(r, \theta, \zeta)$ as a discretized function: a spline in r ($p = 1$ for the VMEC case) and Fourier expansion in θ, ζ .

The spline coefficients of the MRX map are the L^2 projection of this series onto the map's tensor-product spline space, a cheap operation as it de-couples by direction. The map Φ is then assembled as in (20).

The initial magnetic field is obtained as the discrete curl of a vector potential, $B_0 = \text{curl } A_0$. From the equilibrium's toroidal and poloidal fluxes $r \mapsto \Psi_t(r), \Psi_p(r)$ and the angle stream function $(r, \theta, \zeta) \mapsto \lambda(r, \theta, \zeta)$ we obtain the closed-form reference 1-form

$$\hat{A}_0(r, \theta, \zeta) = 2\pi \left(-\frac{\lambda \partial_r \Psi_t(r)}{2\pi}, \Psi_t(r), -\frac{\Psi_p(r)}{n_{\text{fp}}} \right), \quad (26)$$

in the logical frame, whose curl is the equilibrium field in the Clebsch form of the nested-surface solver.

The discrete field is $B_0 = \text{curl } A_0$, so B_0 is divergence-free to machine precision. The potential $A_0 \in V^1$ is fixed by prescribing the line integrals of $\hat{A}_0$ over one set of cell-sized intervals (Greville histopolation). The toroidal flux is the circulation of A_0 along the wall, a sum of these prescribed line integrals.

Our numerical experiments are run on the geometry of the Landreman-Paul precise quasi-axisymmetric (QA)

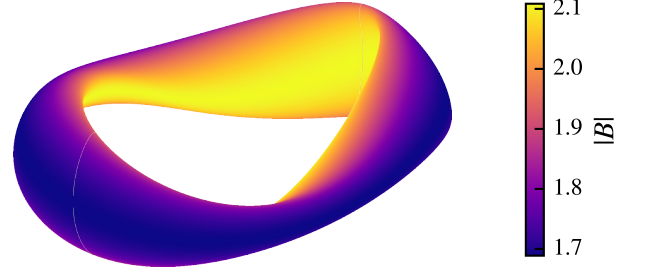


FIG. 3: Vacuum case: Magnitude $|B|$ of the vacuum magnetic field on the boundary of the QA device. Spline resolution $(32 \times 64 \times 32)$, degree $p = 3$. A faint seam is visible after one field period; this is only a visualization artifact, from mapping the plasma with field-period-symmetry.

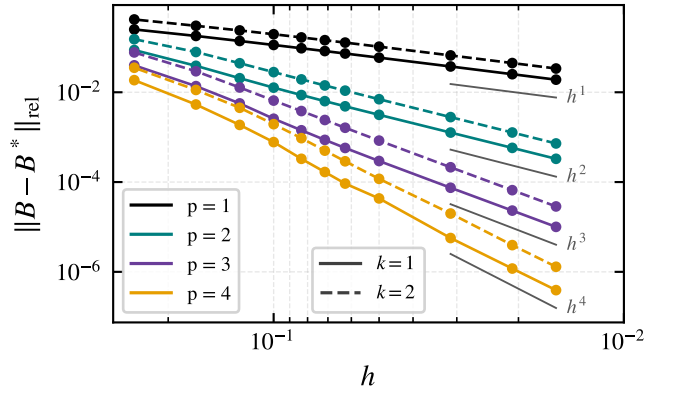


FIG. 4: Convergence of the discrete vacuum field to the analytic solution B^* with respect to resolution: $k = 1$ is the scalar-potential route $B = \text{grad } f + c\mathbf{h}_1$, $k = 2$ the vector-potential route $B = \text{curl } A$. The gray lines indicate the theoretical convergence rate h^p , which is observed for all cases.

device [79], shown in Fig. 3.²

4.2. Manufactured vacuum solution

We consider the Landreman-Paul precise QA design with $n_{\text{fp}} = 2$ and aim to solve the following prob-

² The original low-res input file for this configuration can be found at: https://github.com/hiddenSymmetries/simsopt/blob/master/tests/test_files/input.LandremanPaul2021_QA_lowres

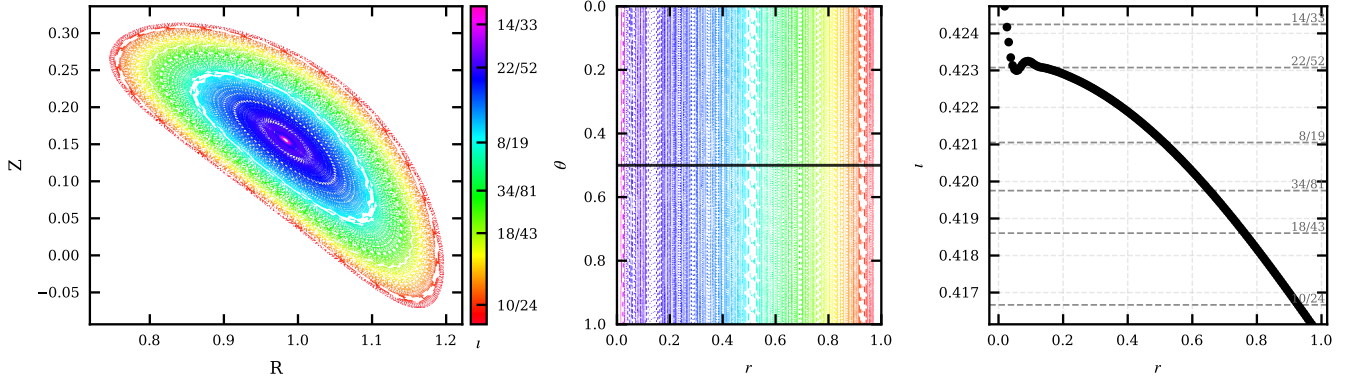


FIG. 5: Vacuum case, left: Poincaré ($\zeta = 0.25$) cross-section of the vacuum field of the QA device in the physical frame with field lines colored by rotational transform ι . Center: The same plot in the logical frame. Right: the rotational transform of every traced line, one marker per line, against the logical radius r at which the line crosses $\theta = 0.5$ (the horizontal line in the center plot).

lem: Find B such that $\text{div } B = \text{curl } B = 0$ in Ω and $(B - B_{\text{vac}}) \cdot n = 0$ on $\partial\Omega$. The idea here is that we can specify an analytic vacuum field as a boundary condition, and validate (h, p) convergence of our code to the analytic vacuum field in the whole volume. The field B_{vac} is chosen as

$$B_{\text{vac}}(R, \phi, Z) = \frac{e_\phi}{R} + \text{grad} (R^2 \cos(n_{\text{fp}} \phi)) \quad (27)$$

in cylindrical coordinates with $\phi \in [0, 2\pi]$. All computations take place in the toroidal coordinates (r, θ, ζ) .

There are two distinct ways to solve this problem, corresponding to the two Hodge decompositions of the vector-valued $L^2(\Omega)$ space, i.e. expressing the magnetic field as a 1-form H or a 2-form B .

a. Solution for $k = 1$ When we approximate the magnetic field as $H \in V^1$, then it holds that

$$H = c\mathfrak{h}_1 + \text{grad } f \quad (28)$$

where $\mathfrak{h}_1$ denotes the harmonic field, i.e. the single element of $\mathfrak{H}^1$, $f \in V^0$, and $c \in \mathbb{R}$ is a scale parameter. Note that $\mathfrak{h}_1 \perp_{L^2} \text{grad } f$ by the discrete Hodge-Helmholtz decomposition property of the FEEC spaces.

To satisfy $\text{div } H = 0$ and the normal boundary conditions, we can solve the Neumann problem

$$\langle \text{grad } f, \text{grad } v \rangle_\Omega = \langle B_{\text{vac}}, \text{grad } v \rangle_\Omega \quad \forall v \in V^0. \quad (29)$$

This constitutes the inversion of a $k = 0$ Hodge-Laplacian operator with natural boundary conditions. The scale c is set by matching the circulation along a closed toroidal curve, which in logical coordinates is obtained by following the vector field $(0, 0, 1)$.

b. Solution for $k = 2$ When approximating the vacuum field as $B \in V^2$, there is no explicit harmonic contribution as $\mathfrak{H}^2 = \{0\}$. Instead $B = \text{curl } A$, where the vector potential $A \in V^1$ solves the curl-curl problem

$$\langle \text{curl } A, \text{curl } w \rangle_\Omega = \langle B_{\text{vac}}, \text{curl } w \rangle_\Omega \quad \forall w \in V^1, \quad (30)$$

i.e. a $k = 1$ Hodge-Laplacian problem with natural boundary conditions. The toroidal flux is carried by the boundary circulation of A and there is no harmonic contribution.

Results are presented in Fig. 4. The meshes have $(m, 2m, m)$ splines per direction with $m = n + p$ and $n = 4, \dots, 64$ radial cells; the 1-form space then has $6m^3 - 12m^2 + 5m$ and the 2-form space $6m^3 - 12m^2 + 2m$ degrees of freedom, from 3,400 ($n = 8, p = 1$) to 1,830,000 ($n = 64, p = 4$).

Up to $n = 16$ a run takes 1.5 to 5 minutes at every degree, of which the two solves take seconds to a minute and the rest is the construction of the sequence and its harmonic forms together with the JAX compilation. At the finest meshes the solves dominate: at $n = 64$ the scalar-potential solve takes under a minute at every degree, while the vector-potential solve takes 0.7, 3.5, 8 and 23 minutes for $p = 1$ to 4.

4.3. Convergence to a VMEC equilibrium

Next, we compare the discrete harmonic 2-form $\mathfrak{h}_2 \in \mathfrak{H}_0^2$ against the vacuum field B_{VMEC} of a VMEC equilibrium of the same QA design. To match the scale of B_{VMEC} , we compute the circulation $c \in \mathbb{R}$ and report

$$\|B_{\text{VMEC}} - c\mathfrak{h}_2\|_\Omega / \|B_{\text{VMEC}}\|_\Omega. \quad (31)$$

Unlike the manufactured case there is no exact target: we measure agreement with another computed equilibrium, limited by the accuracy of the independent solve. Against a low-resolution VMEC reference (75 flux surfaces, 8/8 poloidal/toroidal modes) the difference saturates at $\approx 8 \times 10^{-5}$. Recomputing the reference at higher resolution (201 flux surfaces, 16/12 modes) lowers that floor to $\approx 5 \times 10^{-5}$, which all three degrees reach at the finest meshes (Fig. 6). The solution is shown in Fig. 5, and illustrates essentially nested flux surfaces, consistent

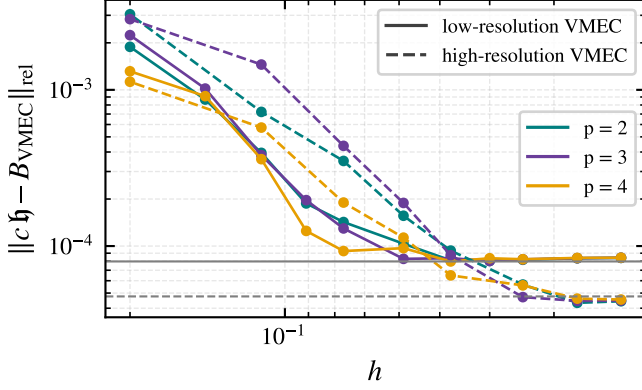


FIG. 6: Agreement $\|B_{\text{VMEC}} - c h_2\|_{\Omega} / \|B_{\text{VMEC}}\|_{\Omega}$ of the discrete harmonic 2-form with the high-resolution VMEC vacuum field. The discrepancy decreases with resolution at every p and saturates at $\approx 4.5 \times 10^{-5}$, the truncation of the reference. Against the low-resolution reference the difference plateaus earlier, at 8.4×10^{-5} .

with a similar calculation in SPECTRE [33, Fig. 5b] (at most, there is a very small 10/24 island chain). We expect further convergence of the floor will be limited both by resolution and by the assumption of nested flux surfaces in VMEC.

With the current MRX code implementation, the harmonic form of the largest reference meshes (1.8 to 3.9×10^5 degrees of freedom) is computed in 20 to 70 seconds after a setup of about two minutes.

5. MAGNETIC RELAXATION

We now move on to describing finite- β MHS solutions obtained by magnetic relaxation. We introduce its history in 5.1, and formulate it as a constrained minimization problem in 5.2. We discuss incompressible variations in 5.3 and clarify the treatment of pressure in 5.4. A brief discussion of regularization in 5.5 leads to an outline of the actual discrete relaxation scheme in 5.6. We include explanation of relaxation diagnostics in 5.7, a clarification on convergence studies in 5.8, and end with Newton's method in 5.9.

5.1. Relation to previous work

The fact that magnetic equilibria are stationary points of the energy with respect to admissible variations has been known for many decades.

a. Energy principle In [80] and [81], two different derivations of admissible variations are given. The latter authors credit Lundquist [82] with the development of this energy principle. The same papers also derive the stability conditions of equilibria based on the eigenvalues

of $\delta^2 \mathcal{E}$. Alternative derivations of the stability condition are given in [83, 84].

b. Helicity The observation that the helicity is conserved under admissible variations is usually credited to Woltjer [85], who also pointed out that the fields $B : \text{curl } B \times B = \lambda B$ with constant $\lambda \in \mathbb{R}$ are minima of the magnetic energy subject to the constraint of constant helicity. The term helicity was coined by Moffatt [86], who also connected it to the degree of entanglement of field lines. The term magnetic relaxation was also popularized by the fluid dynamics community, where the same process is used to study stationary solutions of Euler's equations [87, 88]. The review article [58] and textbook [89] give an overview of magnetic relaxation including energy bounds, the Arnold inequality, and specific topological considerations.

c. Numerics The first works to attempt a constrained minimization of $\mathcal{E}$ as a means to compute magnetic equilibria numerically were, to our knowledge, the works by Chodura and Schlüter [47] and the BETA code [42, 43]. A more recent application of this idea is the SIESTA code [37]. The principle of magnetic relaxation has also been investigated in the astrophysics community in recent years as a tool to study the Parker conjecture [38–41, 90, 91].

5.2. Constrained minimization

Consider a magnetic field $B \in H_0^{\text{div}}(\Omega)$ with $\text{div } B = 0$.

Definition 10 (Energy functional). *The magnetic energy $\mathcal{E} : H_0^{\text{div}}(\Omega) \rightarrow \mathbb{R}$ of the field B is given by*

$$\mathcal{E}(B) := \frac{1}{2} \|B\|_{\Omega}^2. \quad (32)$$

a. Admissible variations Ideal perturbations of B conserve the magnetic topology. They are defined as follows:

Definition 11 (Admissible variations). *Given $v \in H_0^{\text{div}}(\Omega) \cap C^1(\Omega)$, we define admissible variations $v \mapsto \delta_v B$ by:*

$$\delta_v B := \text{curl}(v \times B). \quad (33)$$

Proposition 12 (First variation of the energy). *For $B \cdot n = v \cdot n = 0$ on $\partial\Omega$ and $J = \text{curl } B$,*

$$\delta_v \mathcal{E}(B) = \langle B, \delta_v B \rangle_{\Omega} = -\langle J \times B, v \rangle_{\Omega}. \quad (34)$$

Proof. From the definition of $\delta_v B$ and integration by parts, $\delta_v \mathcal{E}(B) = \langle B, \text{curl}(v \times B) \rangle_{\Omega} = \langle \text{curl } B, v \times B \rangle_{\Omega}$. The boundary term vanishes since $((v \times B) \times B) \cdot n = (v \cdot B)(B \cdot n) - |B|^2(v \cdot n) = 0$. Finally $\langle J, v \times B \rangle_{\Omega} = -\langle J \times B, v \rangle_{\Omega}$. $\square$

b. Arnold's inequality As was mentioned in the introduction, to be useful, magnetic relaxation schemes should avoid relaxing to the vacuum equilibrium $B \propto \mathfrak{h}$.

To construct such a magnetic relaxation scheme, we will recall the following lower bound on the energy, often referred to as Arnold's theorem [92].

Definition 13 (Generalized helicity). *Let $A \in H_0^{\text{curl}}(\Omega)$ with $\text{curl } A = (\text{Id} - \Pi^{\mathfrak{H}_0^2})B$ and $\text{div } A = 0$. The generalized helicity of B is*

$$\mathcal{H}(B) := \langle A, B + \Pi^{\mathfrak{H}_0^2} B \rangle_{\Omega}. \quad (35)$$

Its value does not depend on the gauge of A ; the gauge $\text{div } A = 0$ is fixed for concreteness. When $\mathfrak{H}^2(\Omega) \neq \{0\}$, this is no longer the case and the definition must be adjusted.

Proposition 14 (Helicity bounds energy). *When $\mathfrak{H}^2(\Omega) = \{0\}$, the magnetic energy is bounded from below by the generalized helicity, with λ_{Ω} a constant that depends only on Ω :*

$$\frac{4\mathcal{E}(B)}{\sqrt{\lambda_{\Omega}}} \geq \mathcal{H}(B). \quad (36)$$

Proof. Write $B_{\mathfrak{H}} := \Pi^{\mathfrak{H}_0^2} B$. The Poincaré inequality [93, Proposition 4.1] and Helmholtz-Hodge decomposition (21) give

$$\|A\|_{\Omega}^2 \leq \frac{1}{\lambda_{\Omega}} \|\text{curl } A\|_{\Omega}^2 = \frac{1}{\lambda_{\Omega}} \|B\|_{\Omega}^2. \quad (37)$$

Hence,

$$\sqrt{\lambda_{\Omega}} \langle A, B + B_{\mathfrak{H}} \rangle_{\Omega} \leq \sqrt{\lambda_{\Omega}} \|A\|_{\Omega} (\|B\|_{\Omega} + \|B_{\mathfrak{H}}\|_{\Omega}),$$

by the Poincaré and Cauchy-Schwarz inequalities, and

$$\leq 2\sqrt{\lambda_{\Omega}} \|A\|_{\Omega} \|B\|_{\Omega} \leq 2\|B\|_{\Omega}^2 = 4\mathcal{E}(B),$$

since $\|B_{\mathfrak{H}}\|_{\Omega} \leq \|B\|_{\Omega}$. $\square$

Conserving the generalized helicity provides a barrier to relaxing to the trivial equilibrium. A relaxation trajectory is visually illustrated in Fig. 7. Resistive dissipation can be used to purposefully induce topological changes, but it is important to control any change in helicity that occurs. We justify below why admissible variations allow for sufficient control over this generalized helicity loss.

Proposition 15 (Preservation properties of admissible variations). *Admissible variations of the magnetic field preserve $\text{div } B = 0$ and the harmonic component $\Pi^{\mathfrak{H}_0^2} B$, since $\text{div } \text{curl} = 0$ and $\text{curl}(v \times B) \perp \mathfrak{H}_0^2$ by the Hodge-Helmholtz decomposition, and they leave the generalized helicity of Definition 13 unchanged: with $A \in H_0^{\text{curl}}(\Omega)$, $\text{curl } A = B - \Pi^{\mathfrak{H}_0^2} B$,*

$$\delta_v \mathcal{H}(B) = \delta_v \int_{\Omega} A \cdot (B + \Pi^{\mathfrak{H}_0^2} B) dx = 0. \quad (38)$$

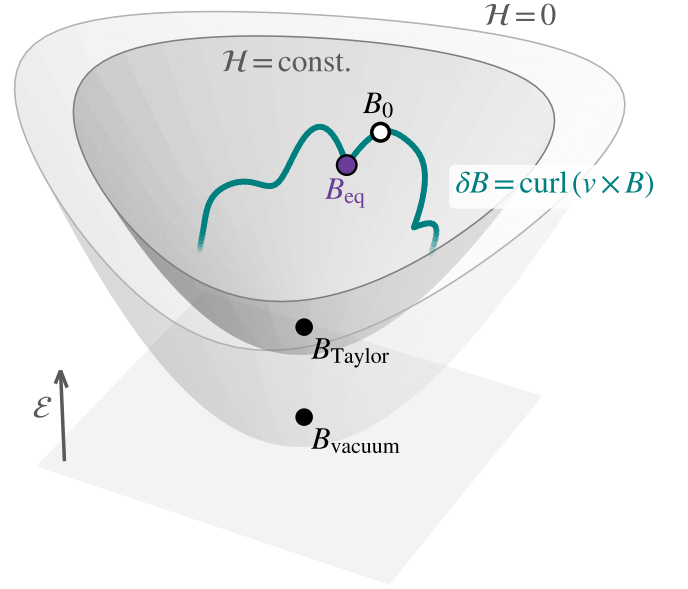


FIG. 7: A sketch of magnetic relaxation as a constrained minimization problem. Field configuration energy $\mathcal{E}$ is displayed on the z -axis. Constant helicity $\mathcal{H}$ sheets are plotted in gray. The lowest energy states at fixed helicity are Taylor states, $\text{curl } B_{\text{Taylor}} = \lambda B_{\text{Taylor}}$. By Arnold's inequality (36), lowest energy within one sheet grows with $|\mathcal{H}|$ and the vacuum field $\text{curl } B_{\text{vacuum}} = 0$ is the lowest energy state. The orbit generated by admissible variations (teal) satisfies the stricter constraint of constant field topology within one helicity sheet. Relaxation from B_0 descends along this orbit to the nearest minimum, B_{eq} . Smooth local minima of the energy along this orbit satisfy $\langle J \times B, v \rangle_{\Omega} = 0$. The orbit is plotted with kinks to demonstrate the regularity issues discussed in Section 5.8: at a kink the energy is not differentiable along the orbit, and a minimum there need not be force-free.

Proof. Let $B_{\mathfrak{H}} := \Pi^{\mathfrak{H}_0^2} B$. For the vector potential, $\delta_v A = v \times B + \text{grad } \varphi$ for some $\varphi \in H^1(\Omega)$ depending on the gauge. Hence,

$$\begin{aligned} \delta \mathcal{H}(v) &= \langle \delta_v A, B + B_{\mathfrak{H}} \rangle_{\Omega} + \langle A, \delta_v B \rangle_{\Omega} \\ &= \langle v \times B + \text{grad } \varphi, B + B_{\mathfrak{H}} \rangle_{\Omega} \end{aligned} \quad (39)$$

$$\begin{aligned} &+ \langle A, \text{curl}(v \times B) \rangle_{\Omega} \\ &= \langle v \times B + \text{grad } \varphi, B + B_{\mathfrak{H}} \rangle_{\Omega} \\ &+ \langle B - B_{\mathfrak{H}}, v \times B \rangle_{\Omega}. \end{aligned} \quad (40)$$

Since $B \times B = 0$ and $\text{grad } \varphi \perp_{L^2(\Omega)} \{B, B_{\mathfrak{H}}\}$ for all $\varphi \in H^1(\Omega)$, all terms vanish except $\pm \langle B_{\mathfrak{H}}, v \times B \rangle_{\Omega}$, which cancel one another. $\square$

Remark. For more information on generalized helicities, we refer to [89, Section III.7] and [40, 94].

c. Resistivity and reconnection Occasionally, it is useful in a magnetic relaxation scheme to induce some re-

connection. Resistivity is an effect that introduces magnetic diffusion, allowing for reconnection of field lines.

Physically, resistivity proportional to a constant parameter $\eta \geq 0$ modifies the variations to

$$\delta_v B = \text{curl}(v \times B - \eta J), \quad (41)$$

with $J \in H_0^{\text{curl}}(\Omega)$.

Lemma 16 (Resistive variations). *For $\delta_v B = \text{curl}(v \times B - \eta J)$ with $J \in H_0^{\text{curl}}(\Omega)$,*

$$\delta_v \mathcal{E}(B) = -\langle J \times B, v \rangle_\Omega - \eta \|J\|_\Omega^2, \quad (42)$$

and

$$\delta_v \mathcal{H}(B) = -2\eta \langle J, B \rangle_\Omega. \quad (43)$$

See Appendix A for the proof.

Proposition 17 (Helicity loss under resistivity). *With the variations of Lemma 16,*

$$|\delta_v \mathcal{H}(B)|^2 \leq 8\eta \mathcal{E}(B) |\delta_v \mathcal{E}(B)|, \quad (44)$$

an upper bound on $|\delta \mathcal{H}|$ of order $\sqrt{\eta}$.

While magnetic diffusion can drastically change the small-scale topology of the field, the global helicity can persist through much longer time scales. The proof is given in [89, Remark 7.19].

d. Gradient descent Setting $v = J \times B$ leads to $\delta_v \mathcal{E}(B) = -\|J \times B\|_\Omega^2 \leq 0$, a decrease in energy that formally terminates once $\text{curl} B = \lambda B$, a pressure-less equilibrium of the form of a Beltrami field.

5.3. Incompressible variations

Clearly, we need to generalize this picture to include finite pressure; we want to satisfy $J \times B = \nabla p$ at equilibrium, not $J \times B = 0$. This is achieved with the Leray projection.

Definition 18 (Leray projection). *The Leray projection of $u \in H^{\text{div}}(\Omega)$ is*

$$\Pi^{\text{Leray}} u := u - \text{grad } p, \text{ where } -\Delta p = -\text{div } u, \quad (45)$$

with homogeneous Neumann boundary conditions on p ; by construction $\text{div } \Pi^{\text{Leray}} u = \text{div}(u - \text{grad } p) = 0$.

Corollary 19 (Force balance at a stationary point). *For $v = \Pi^{\text{Leray}}(J \times B)$,*

$$\delta_v \mathcal{E}(B) = -\langle J \times B, J \times B - \text{grad } p \rangle_\Omega \quad (46)$$

$$= -\|J \times B - \text{grad } p\|_\Omega^2, \quad (47)$$

as $J \times B - \text{grad } p$ is divergence-free and hence L^2 -orthogonal to $\text{grad } p$; a stationary point of this descent satisfies $J \times B = \text{grad } p$.

More generally, the Hodge–Helmholtz decomposition yields

Corollary 20 (Decomposition of the Lorentz force). *By the Hodge–Helmholtz decomposition (21), for any B the Lorentz force $J \times B = \text{curl } B \times B$ decomposes as*

$$J \times B = \text{grad } p + \text{curl } \omega + c\mathfrak{h}, \quad (48)$$

with $\text{curl } \mathfrak{h} = \text{div } \mathfrak{h} = 0$, $\|\mathfrak{h}\|_\Omega = 1$ and $c \in \mathbb{R}$. The three components are L^2 -orthogonal to one another, hence $\|J \times B - \text{grad } p\|_\Omega \leq \|J \times B\|_\Omega$.

Another option is to solve $\text{curl } \text{curl } \omega = \text{curl}(J \times B)$ and set $c = \langle \mathfrak{h}, J \times B \rangle_\Omega$. This step is analogous to the computation of the potential A from B .

We then obtain $J \times B - \text{grad } p$ as $\text{curl } \omega + c\mathfrak{h}$. Note that $c \ll 1$ close to equilibria where $J \times B$ is approximately $\text{grad } p$, which is L^2 -orthogonal to $\mathfrak{h}$.

Proposition 21 (Admissible variations are push-forwards along volume-preserving flows). *Let $\mathcal{F}_t$ be the flow of a velocity field v_t , $d\mathcal{F}_t(x)/dt = v_t(\mathcal{F}_t(x))$, $\mathcal{F}_0(x) = x$, and let $B_t(\mathcal{F}_t(x)) := D\mathcal{F}_t(x)B_0(x)$ be the push-forward of B_0 . Then $\mathcal{F}_t$ is volume-preserving if and only if $\text{div } v_t = 0$, and $\partial_t B_t = \text{curl}(v_t \times B_t)$: the admissible variations of Definition 11 are the infinitesimal push-forwards along volume-preserving flows [89].*

Proof. Let $\mathcal{F}_t$ denote such a map and define $B_t(\mathcal{F}_t(x)) := D\mathcal{F}_t(x)B_0(x)$, the push-forward of an initial field B_0 with this flow. When $d\mathcal{F}_t(x)/dt = v_t(\mathcal{F}_t(x))$, $\mathcal{F}_0(x) = x$, then $\mathcal{F}_t$ is volume-preserving if and only if $\text{div } v_t = 0$ (by Liouville's formula) and we find $0 = d(B_t \circ \mathcal{F}_t - D\mathcal{F}_t B_0)/dt|_{t=0} = \partial_t B_t + DB_t v_t - Dv_t B_t = \partial_t B_t - \text{curl}(v_t \times B_t)$. An adapted calculation holds for compressible variations as well, see [95, Lemma C.1]. $\square$

Remark. It has been conjectured in [89, Section I.9] and shown in [96] that helicity is in fact the only integral invariant of volume-preserving maps given certain regularity assumptions.

5.4. Pressure

For Grad-Shafranov solvers as well as 3D nested equilibrium solvers that solve for the flux function ψ , the dependence of p and ι as functions of ψ is an input to the method. The physical fields $p(x)$ and $\iota(x)$ are then given by $p(x) = p_{\text{profile}}(\psi(x))$ and $\iota(x) = \iota_{\text{profile}}(\psi(x))$. With enough experience, one can use the inputs p_{profile} and ι_{profile} to shape the outputs p and ι . Even in this case, the final pressure in physical space is also an output of the numerical method, since it will vary as ψ varies. In three-dimensional configurations, this procedure breaks down as B is no longer determined by ψ and the toroidal flux value alone.

In our case, the pressure p (the same letter as the spline degree, which does not appear in this context) is

not treated as a dynamical variable and function of the density as in e.g. [37, 47, 48], but rather as a Lagrange multiplier. We prescribe an initial pressure and rotational transform only through the initial condition B_0 . At each step, it is the solution to the elliptic problem Eq. (45) with homogeneous Neumann boundary conditions, with the advantage that it enforces the important property that force balance holds at an energy minimum, Eq. (47). In other words, the final pressure profile will depend on the initial condition and the trajectory (the hysteresis) of the relaxation.

Remark. This choice is common among helicity-preserving relaxation methods in fluid dynamics but to our knowledge has not been used in previous magnetic relaxation codes.

For practical purposes, one way to replicate the profile design is to prescribe a p and ι profile using a code that assumes nested flux surfaces and then use the output of this code as an input to a relaxation run, potentially with a small perturbation to escape the (local) minimum with nested flux surfaces. As the relaxation process starts close to an equilibrium, we can expect that $\|B - B_0\|_\Omega$ is small, where B denotes the state of the magnetic field throughout the relaxation. The same is true for $\|p - p_0\|_\Omega$ and $\|\iota - \iota_0\|_\Omega$, giving us control of p and ι through the initial condition. In this case, the lack of direct control over the profiles might actually be a useful feature; the prescribed p_{profile} and ι_{profile} are roughly respected, and some modifications and flattening of the profiles can be achieved where islands and chaos open up in the solution.

5.5. Regularization

There are a priori only very weak regularity guarantees on $J \times B$ and this can lead to grid-scale oscillations in v in a numerical scheme. For this reason, it is beneficial to apply a smoothing operation to v , as has been done in [47] and also is done in the analytic treatment of the relaxation problem in [97, Equation 2.1].

We can let $v = (\text{Id} - \varepsilon \Delta)^{-\gamma} (J \times B - \text{grad } p)$ ($\varepsilon > 0$) in order to dampen high-frequency oscillations.

We note the following solvability result:

Lemma 22 ([72], Section 6.2.3). *For any $F \in H_0^{\text{div}}(\Omega)$ satisfying $F \perp \mathfrak{H}_0^2$ and $\text{div } F = 0$, there exists $u \in H_0^{\text{div}}(\Omega)$ such that $\text{curl curl } u = F$, $\text{div } u = 0$, and $u \perp \mathfrak{H}_0^2$.*

By Lemma 22, $\text{div } v = \text{div} (\text{Id} - \varepsilon \Delta)^{-\gamma} \Pi^{\text{Leray}} (J \times B) = 0$ and

$$\begin{aligned} \delta_v \mathcal{E}(B) &= -\langle (\text{Id} - \varepsilon \Delta)^{-\gamma} (J \times B - \text{grad } p), J \times B \rangle_\Omega \quad (49) \\ &= -\|J \times B - \text{grad } p\|_{H_{0,\varepsilon}^{-\gamma}(\Omega)}^2 = -\|v\|_{H_{0,\varepsilon}^\gamma(\Omega)}^2, \end{aligned}$$

i.e. the energy dissipation equals an ε -weighted Sobolev norm of order $-\gamma$ of the force residual.

5.6. Discrete relaxation

We now provide the discrete counterparts to the continuous relaxation. We simplify to the case $\gamma = 1$ without loss of generality.

a. Algorithm Relaxation is performed by solving the following system of equations:

Definition 23 (Relaxation step). *Let $B^n \in V_0^2$, $\delta t, \varepsilon \in \mathbb{R}_{>0}$. The relaxation step from B^n to B^{n+1} is the solution of*

$$J = \widetilde{\text{curl}} B^n, \quad (50)$$

$$p = (\text{div } \widetilde{\text{grad}})^{-1} \text{div} (J \times B^n) \quad (51)$$

$$v = (\text{Id} + \varepsilon \text{curl curl})^{-1} (J \times B^n - \widetilde{\text{grad}} p), \quad (52)$$

$$E = \Pi_0^1(v \times B^n) - \sigma \Pi_0^1 B^n, \quad (53)$$

$$B^{n+1} = B^n + \delta t \text{curl } E, \quad (54)$$

where the optional scalar $\sigma \in \mathbb{R}$ is chosen such that $\langle E, B^n + B^{n+1} \rangle_\Omega = 0$.

The operations $\widetilde{\text{grad}}$, $\widetilde{\text{curl}}$, $\widetilde{\text{div}}$ are the weak derivatives as described in Section 3.5. They are the discrete version of the co-differential operators from Equation (13).

The velocity smoothing constitutes a Hodge-Laplace problem for $k = 2$ and the Leray projection one for $k = 3$. The alternative path via $v = \text{curl } \omega + \mathfrak{ch}$ requires the inversion of a $k = 1$ Hodge-Laplacian:

$$\omega = (\widetilde{\text{curl}} \text{curl})^{-1} \widetilde{\text{curl}} (J \times B^n), \quad \widetilde{\text{div}} \omega = 0, \quad (55)$$

$$v = \text{curl} (\text{Id} + \varepsilon \widetilde{\text{curl}} \text{curl})^{-1} \omega + \frac{\langle J \times B^n, \mathfrak{h} \rangle_\Omega}{\|\mathfrak{h}\|_\Omega^2} \mathfrak{h} \quad (56)$$

All problems are posed with essential boundary conditions. Their discrete well-posedness is shown in [72, Sections 6.2.2–4].

All together, a single relaxation step thus constitutes the solution of two Hodge-Laplace equations and a number of mass matrix inversions.

b. Discrete helicity preservation We now discuss how the continuous helicity preservation from Proposition 15 reflects in the discrete setting.

Definition 24 (Discrete vector potential and helicity). *The discrete vector potential $A \in V_0^1$ of $B \in V_0^2$ is the solution of the Hodge-Laplace problem for $k = 1$ with essential boundary conditions [72, Section 6.2.2],*

$$\widetilde{\text{curl}} \text{curl } A - \text{grad } q = \widetilde{\text{curl}} B, \quad q = \widetilde{\text{div}} A, \quad A \perp_{L^2} \mathfrak{H}_0^1, \quad (57)$$

which gives $\text{curl } A = B - \Pi^{\mathfrak{H}_0^2} B$. The discrete helicity is that of Definition 13, $\mathcal{H} = \langle A, B + \Pi^{\mathfrak{H}_0^2} B \rangle_\Omega$.

The following identity holds for any $E \in V_0^1$ when $B^{n+1} = B^n + \delta t \text{curl } E$, the proof is given in Appendix A.

Lemma 25 (Discrete helicity balance). *Let $B^n \in V_0^2$, $E \in V_0^1$ and $B^{n+1} = B^n + \delta t \operatorname{curl} E$. Then*

$$\mathcal{H}^{n+1} - \mathcal{H}^n = 2\delta t \langle E, B^n \rangle_\Omega + \delta t^2 \langle E, \operatorname{curl} E \rangle_\Omega. \quad (58)$$

The continuous flow conserves the helicity because $E = v \times B$ is point-wise orthogonal to B and the second term, time-integration error of the explicit step, vanishes. Discretely, the projection of $v \times B$ onto V_0^1 to obtain E leads to a residual of order h^p in the interior.

Corollary 26 (Discrete helicity remainder). *For $E_\sigma = \Pi_0^1(v \times B^n) - \sigma \Pi_0^1 B^n$ as in (50),*

$$\begin{aligned} \frac{\mathcal{H}^{n+1} - \mathcal{H}^n}{2\delta t} &= \langle \Pi_0^1(v \times B^n) - v \times B^n, B^n - \Pi_0^1 B^n \rangle_\Omega \\ &\quad - \sigma \|\Pi_0^1 B^n\|_\Omega^2 + \frac{\delta t}{2} \langle E_\sigma, \operatorname{curl} E_\sigma \rangle_\Omega. \end{aligned} \quad (59)$$

Equation 59 is a quadratic function in σ . It can be solved for σ to make the right-hand side vanish and restore exact discrete helicity preservation at the cost of a few matrix-vector products and one additional mass matrix inversion.

An alternative method to exactly preserve helicity discretely is the introduction of an auxiliary variable $H = \Pi_0^1 B$ together with an implicit midpoint method for time-stepping. For the configurations in this work, $H \times n = 0$ is not a reasonable assumption, as the magnetic field does not vanish on the computational boundary. Projecting B to the V_0^1 space can lead to spurious oscillations. The works [40, 57] that use the auxiliary variable formulation assume either fully periodic or no-slip boundary conditions.

The advantage of the auxiliary variable approach is that the problem stays variational. In contrast, $\sigma \neq 0$ changes the energy identity by $\sigma \delta t \langle J, B \rangle_\Omega$. Given that σ is usually of the order 10^{-12} , it is in practice not enough to break energy dissipation.

We emphasize here that the preservation of global helicity is not a constraint that is as strong as the preservation of field-line topology that is present in the continuous problem. Indeed, as we will see in the numerical experiments, grid-scale islands can appear and vanish throughout relaxation after discretization.

c. Discrete energy evolution The following result is proven in Appendix A.

Proposition 27 (Discrete energy evolution). *Solutions to (50) satisfy*

$$\begin{aligned} \frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{\delta t} &= -\|v\|_\Omega^2 - \varepsilon \|\widetilde{\operatorname{curl} v}\|_\Omega^2 - \sigma \langle J, B^n \rangle_\Omega \\ &\quad + \frac{\delta t}{2} \|\operatorname{curl} E_\sigma\|_\Omega^2 \end{aligned} \quad (60)$$

with $E_\sigma = \Pi_0^1(v \times B^n) - \sigma \Pi_0^1 B^n$.

In particular, for $\sigma = 0$ the energy decreases for $0 < \delta t < 2(\|v\|_\Omega^2 + \varepsilon \|\widetilde{\operatorname{curl} v}\|_\Omega^2) / \|\operatorname{curl} E_\sigma\|_\Omega^2$.

Lemma 28 (Exact line search). *Since the magnetic energy is quadratic in the increment $\delta B = \operatorname{curl} E^n$, the line search $\min_{\delta t} \mathcal{E}(B^n + \delta t \operatorname{curl} E^n)$ has the closed-form solution*

$$\delta t_* = -\frac{\langle B^n, \operatorname{curl} E^n \rangle_\Omega}{\|\operatorname{curl} E^n\|_\Omega^2}. \quad (61)$$

We clip this time-step so that the velocity advects the field by at most a fraction C of a cell in one step: with $\hat{v}^i$ the logical components of v and h_i the knot spacing of direction i .

$$\delta t = \min \left\{ \delta t_*, \frac{\text{CFL}}{\max_{\hat{x} \in \Omega, i} |\hat{v}^i(\hat{x})|/h_i} \right\} \quad (62)$$

with $\text{CFL} < 1$. We pick $\text{CFL} = 0.5$ in all numerical experiments. Empirically, the cap binds only in the first steps of a relaxation from a rough initial field.

d. L-BFGS We can store the velocities from previous steps in order to build an L-BFGS descent path. We denote the memory size by m . Note that the L-BFGS approximation of the hessian is not aware of the curvature of the constraint (constant topology) manifold. While it is built only from admissible descent directions, a true second-order method needs to take the second constrained variation into account as discussed in Sec. 5.9. With $m = 1$, we recover the Polak–Ribière conjugate gradient method, also used in [47]. We use Powell’s restart [98] when consecutive forces are far from orthogonal and fall back to gradient descent ($m = 0$) for this step to avoid the descent stalling.

e. Chunking The relaxation loop runs internally in an inner, compiled loop of chunks, embedded in an outer loop that is pure Python. Between outer loop steps, any callback routines can run, such as computing and storing diagnostics. As a result, host-device synchronization happens once per chunk rather than once per step. This offers a good compromise between speed and flexibility. The fastest performance can be achieved running the entire relaxation in one chunk, but the overhead from the outer loop is typically small (1–5% of wall time).

The full descent state, the field, the pressure, the warm starts of the solves and the optimizer’s history is written to disk at every chunk and can be restarted from any checkpoint with the same geometry, mesh, degree and precision, so that a relaxation can be branched off from or resumed at a later time.

5.7. Diagnostics

The following diagnostics are central to our numerical experiments.

a. Force residual We recall that following GVEC’s convention, we normalize the force balance by the magnetic pressure gradient.

$$\|F\|_{\text{norm}}^2 = \|J \times B - \operatorname{grad} p\|_\Omega^2 / \|\operatorname{grad} |B|^2/2\|_\Omega^2.$$

b. Rotational transform The rotational transform ι is defined as the fraction of poloidal versus toroidal rotations along a field line.

For each traced field line, ι is fitted as the slope of the unwrapped poloidal against the toroidal angle over the whole trace.

c. Poincaré plots To compute integral curves we integrate $dx_t/dt = B(x_t)$ with $x_{t=0} = x_0$. This is done in logical coordinates: the integral curve there is $\hat{x}_t = \Phi^{-1} \circ x_t$, so $d\hat{x}_t/dt = D\Phi^{-1}(x_t) dx_t/dt = D\Phi^{-1}(x_t) B(x_t)$ and $\hat{x}_{t=0} = \Phi^{-1}(x_0)$. Using the definition of $B \in V^2$ and $H \in V^1$ through their logical representatives $\hat{B}$ and $\hat{H}$,

$$\begin{aligned} \frac{d}{dt} \hat{x}_t &= (\det D\Phi(\hat{x}_t))^{-1} \hat{B}(\hat{x}_t) \\ &= D\Phi^{-1} D\Phi^{-T} \hat{H}(\hat{x}_t). \end{aligned} \quad (63)$$

Since we want intersections of the integral curves with planes $\zeta = \text{constant}$, we re-parametrize by ζ , i.e. we divide by the ζ -component of the right-hand side:

$$\frac{d}{d\zeta} \hat{x} = \frac{\hat{B}(\hat{x})}{\hat{B}_\zeta(\hat{x})}, \quad (64)$$

where $\hat{B}_\zeta$ is the ζ -component of the logical vector $\hat{B}$ (and analogously $\hat{H}$). The ζ -component of the right-hand side is then unity, so ζ is the independent variable and we record crossings at fixed- ζ planes, stepping forward period by period. In our numerical experiments field lines are integrated with a fixed-step fifth-order explicit (Tsitouras') Runge–Kutta scheme from the `diffirax` [99] package.

Remark. This is in contrast to, e.g. the SIMSOPT suite, where field lines are traced in cartesian coordinates and plane crossings are detected in a secondary post-processing step via root-finding. We have found the re-parametrization with ζ to be significantly faster.

d. Island size The width of an island chain at $\iota = n_{\text{fp}} n/m$, where n counts the chain's toroidal turns per field period so that its toroidal mode number is $n_{\text{fp}} n$, is measured on the traced field lines. Among the lines whose fitted rotational transform lies within 2×10^{-3} of the rational, we take the largest excursion $\max_c r_c - \min_c r_c$ of the logical radius over all crossings c on all planes. A line inside the island near its separatrix spans the full island width, a nested surface spans nothing, so the measure is zero when the chain is closed.

Widths are quoted in logical r and in units of the radial cell $h_r = 1/n_r$.

5.8. Convergence

In this last section, we discuss the topic of convergence of magnetic relaxation methods. We now review the physical mechanism of current sheets forming on rational surfaces, what is known about the continuous problem from the analysis side, and how these difficulties lead to finite resolution effects.

a. Current sheets Ideal relaxation in the continuous limit keeps the topology of the field lines, so the equilibrium it tends to retains the rotational transform of the initial field on every surface. On a surface with rational transform $\iota = n/m$ the field lines close. Force balance across such a surface has no smooth solution in general: the perpendicular current $B \times \text{grad } p/|B|^2$ has a divergence that the parallel Pfirsch-Schlüter current must close. For the (m, n) harmonic the closing current is proportional to $1/(\iota m - n)$ [18].

The ideal displacement that would align the field is discontinuous across the surface for the same reason. At rational surfaces, B is tangentially discontinuous, hence the current curl B is a singular sheet. This mechanism is what the Grad conjecture [9, 10] and the Parker problem [87, 100] are concerned with.

b. The continuous problem The well-posedness of the magnetic relaxation is largely an open problem in analysis. It is known that the regularized relaxation is globally well-posed on the flat torus for $\gamma > d/2 + 1$ and its velocity tends to zero [97, 101]. However, the limit need not be smooth and generic fields cannot relax to a smooth equilibrium under their topological constraint [95, 102].

c. Discretization floor For any discretization, a real current sheet would be a sub-grid quantity. After a fast initial descent, any relaxation run will form them and the force residual stalls near the discretization floor where the forming current sheets can no longer be resolved.

This floor is not a stationary point of the discrete energy, at which the projected force would vanish. The optimization landscape becomes very flat along the directions of modifying the sheet-scale structures. At this stage any progress in decreasing the energy comes at the price of numerical reconnection. One can constrain the topology conservation more by using a smaller time-step or a tighter solver tolerance, but true convergence is not possible.

The consequence is that a relaxation run for practical purposes should aim to reach this floor and terminate instead of spending large amounts of computational effort for little gain.

The rise of the force residual can be pictured as taking long steps along a narrow, curved valley of the energy landscape. The quadratic model is accurate along the valley floor, yet any misalignment sends the step up the valley's sides, where the gradient (force residual) grows even though the energy still decreases. The regularized line search of Section 5.9 shortens the step once the sides are felt.

5.9. Newton's method

Far away from the discretization floor and close to a nested equilibrium state (in the weak sense), the optimization problem is well-behaved. We can make use of second-order information of the energy functional.

a. Second variation The hessian approximation L-BFGS builds is blind to the curvature of the constraint manifold. We now derive the correct expression for the second admissible variation.

Proposition 29 (Second variation of the energy). *Along the flows $\mathcal{F}_t$ of Proposition 21, with $\delta_v B = \text{curl}(v \times B)$ and $\delta_{u,v}^2 B = \text{curl}(u \times \delta_v B)$, the second variation of the energy is*

$$\delta_{u,v}^2 \mathcal{E}(B) = \langle \delta_u B, \delta_v B \rangle_\Omega + \frac{1}{2} \langle B, \text{curl}(v \times \delta_u B) \rangle_\Omega + \frac{1}{2} \langle B, \text{curl}(u \times \delta_v B) \rangle_\Omega. \quad (65)$$

Proof. Expand the push-forward $B_t = \exp(\mathcal{F}_t)B$ in t :

$$B_t = B + t \delta_v B + \frac{t^2}{2} \delta_{v,v}^2 B + O(t^3) \quad (66)$$

and substitute into $\mathcal{E}(B_t) = \frac{1}{2} \|B_t\|_\Omega^2$. The second-order contribution to the energy is $\frac{t^2}{2} (\langle \delta_v B, \delta_v B \rangle_\Omega + \langle B, \delta_{v,v}^2 B \rangle_\Omega)$, adding the contributions from u and v gives (65). $\square$

The Newton step minimizes the quadratic approximation of $\mathcal{E}(B_t)$: find u with $\text{div } u = 0$ such that

$$\delta_{u,v}^2 \mathcal{E}(B) = \langle J \times B - \text{grad } p, v \rangle_\Omega \quad \forall \text{div } v = 0. \quad (67)$$

The constraint $\text{div } u = 0$ can easily be incorporated by formulating the problem in terms of $\omega, c : \text{curl } \omega + c\mathfrak{h} = u$.

At this stage, we drop the harmonic contribution $c\mathfrak{h}$ as $J \times B \approx \text{grad } p$ is almost orthogonal to $\mathfrak{h}$ close to equilibria. The final equation to solve for $\omega \in V_0^1$ is thus

$$\widetilde{\text{curl}} \delta^2 \mathcal{E}(B) \text{curl } \omega = \widetilde{\text{curl}}(J \times B). \quad (68)$$

b. Solution strategy It remains to efficiently solve Equation (68) at every relaxation step.

Lemma 30 (The Newton system). *The system (68) is symmetric and singular. Every gradient $\text{grad } \varphi$, $\varphi \in V_0^0$, lies in its kernel since $\text{curl grad} = 0$, and the right-hand side is orthogonal to these gradients, as $\langle \widetilde{\text{curl}}(J \times B), \text{grad } \varphi \rangle_\Omega = \langle J \times B, \text{curl grad } \varphi \rangle_\Omega = 0$.*

The Gauss-Newton part $\langle \delta_u B, \delta_v B \rangle_\Omega$ of (65) has the further kernel of the velocities with $\text{curl}(v \times B) = 0$, the symmetries of B , on which the right-hand side vanishes as well by Proposition 12.

In practice, we rely on an inexact Newton method [103, Section 7.1] with a fixed number of MINRES iterations applied. The line search along the resulting direction is capped at the Newton length $\delta t = 1$.

The condition number of $\widetilde{\text{curl}} \delta^2 \mathcal{E}(B) \text{curl}$ is $\mathcal{O}(10^5)$, so to achieve anything with a fixed budget of MINRES iterations, we require a decent preconditioner. We make use of the fact that for stellarators, the magnetic field is mostly harmonic ($\|B - c\mathfrak{h}\|_\Omega / \|B\|_\Omega = 0.04$ on the NCSX case). As $\mathfrak{h}$ is constant throughout relaxation, we only

need to compute this approximate inverse once. More details on the preconditioning are given in Appendix 2 and in a forthcoming paper.

Overall, our procedure is similar to that of SIESTA, described in [37, Section IX.A] and [104].

c. Time-stepping near the discretization floor Once the force residual is small enough, the only way to further reduce the magnetic energy is by re-ordering the field near resonant surfaces and current sheets (c.f. Section 5.8). The quadratic model of Newton's approximation breaks down here and following the analytic line search of Equation (61) leads to significant reconnection from time-steps that are too large. While this decreases the energy and hence Newton's method is doing its job, we have to remember that the energy minimization is only a means to find (local) minima where $\delta_v \mathcal{E} = 0$.

We therefore modify the criterion to choose δt to penalize further amplification of current sheets and use the regularized energy $\mathcal{E} + \frac{\varepsilon}{2} \|J\|_\Omega^2$ instead, with the step direction unchanged.

Lemma 31 (Regularized line search). *The exact minimizer of $\mathcal{E} + \frac{\varepsilon}{2} \|J\|_\Omega^2$ along the increment $\delta B = \text{curl } E$ of a direction u with force F is*

$$\delta t_* = \frac{\langle F, u \rangle_\Omega - \varepsilon \langle J, \widetilde{\text{curl}} \delta B \rangle_\Omega}{\|\delta B\|_\Omega^2 + \varepsilon \|\widetilde{\text{curl}} \delta B\|_\Omega^2}, \quad \varepsilon = C h_r^2. \quad (69)$$

It replaces the step of Lemma 28. The effect is that relaxation stops once grid-scale structures emerge. We set $C = 0.1$, determined by a hyperparameter sweep shown in Appendix 2.

6. RELAXATION NUMERICS

We now formulate a robust group of experiments to test the numerics, hyperparameters, computational efficiency, and convergence of the proposed methodology. We introduce a benchmark example NCSX in 6.1 and show results for a quasi-Newton descent method in 6.2. Helicity drift/preservation are discussed in 6.3. We end with some nuances of Newton's method and its limit states in 6.4.

The exact specifications of all numerical experiments that are reported in this work are given in Appendix 1. As before, all reported timings are for a single H100 GPU.

6.1. NCSX configuration

NCSX was an experiment at Princeton Plasma Physics Laboratory, designed to determine whether a compact QA stellarator can combine ideal confinement properties of tokamaks with the functionality of stellarators [105]. Although the project was not fully realized, this configuration remains a useful tool for analysis of high- β design possibilities.

The equilibrium is the 1i383 NCSX reference configuration computed by VMEC with 49 flux surfaces, $m_{\text{pol}} = 9$ poloidal and $n_{\text{tor}} = 5$ toroidal modes, three field periods, major radius 1.42 m, minor radius 0.33 m, aspect ratio 4.4, on-axis field 1.47 T and a volume-averaged β of 4.25% [105]. Its rotational transform rises from 0.393 on the axis to 0.655 at the edge. The map Φ is the L^2 projection of the file's (R, Z) series onto the C^1 polar spline space, and the initial field is the discrete curl of the histopolated potential of Section 4.1. The rotational transform crosses both the $\iota = \frac{1}{2}$ and $\iota = \frac{3}{5}$ surfaces. This is the geometry from Fig. 1.³

All runs in the following branch off from the reference configuration $(n_r, n_\theta, n_\zeta) = 16 \times 32 \times 32$, $p = 2$, mixed precision.

Remark. The seconds per step quoted in the tables are the steady stepping rate over compiled chunks after the first, hence excluding just-in-time compilation. Wall-time axes and the time columns include this first chunk.

The diagnostics evaluated between chunks (helicity, pressures, check-pointing) are excluded from both. This overhead is small (1–5%) compared to the full run time and would not be present if the goal was simply to produce the final relaxed state without documenting the descent.

6.2. L-BFGS

We now show a number of different configurations that the relaxation code can be run in, and demonstrate the impacts of the different hyperparameter choices. A comprehensive list of choices available to the user is shown in Appendix 2.

We want to emphasize that all hyper-parameters in the code have interpretable values in terms of quantities of interest (helicity, plasma beta, etc) or discretization choices (grid scale, solver tolerances, etc).

a. Smoothing The hyperparameter ε that controls the strength of the smoothing is set to $\varepsilon = 0.02 h_r^2$ to offset the grid-dependence of the Laplacian spectrum. The constant 0.02 was determined by a hyperparameter sweep shown in Appendix 1.

From Fig. 8, it is evident smoothing the velocity before advecting the magnetic field is a necessity. For $\gamma = 0$, the force decreases to 2×10^{-7} at step 2200 and then grows to 2×10^{-5} over the remaining 18,000 steps while the helicity drifts by 8×10^{-4} . Grid-scale oscillations take over, and the run is unusable for the purposes of this work, its final field fully chaotic, the nested initial condition broken down entirely.

The need of regularization of magnetic relaxation dynamics is well understood in the continuous case [97]. To

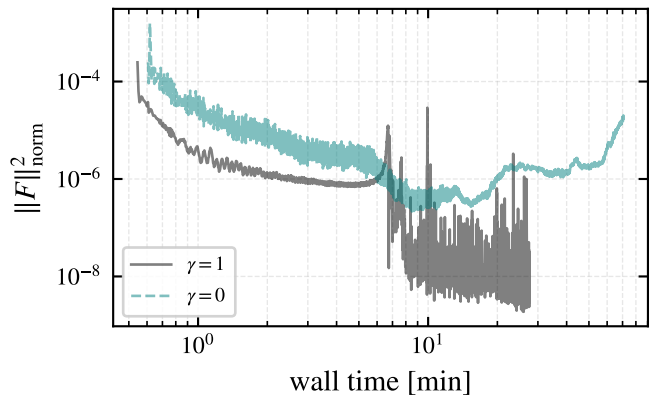


FIG. 8: Impact of velocity smoothing: $\gamma = 1$ (black, solid), $\gamma = 0$ (teal, dashed) for 20,000 steps. Without smoothing, grid-scale oscillations lead to a complete breakdown of the run.

our knowledge, the only work that systematically studied this numerically so far is [106], but only for a one-dimensional toy model.

6.3. Helicity preservation

Next, we verify the helicity balance of Section 5.6. Per step the plain explicit step on B drifts by the remainder (59), which can be compensated by solving for the correction σ . This requires operating at high numerical precision, since the drift of a single-precision field is of the order of its own machine epsilon.

Table II shows that double precision allows to conserve the helicity to 1.7×10^{-15} , ten orders of magnitude below the explicit step's 10^{-5} . In mixed precision, the corrected step is exact to 1.9×10^{-7} , the rounding of the stored helicity. The removed energy and force residual differ between the corrected and the explicit step by less than 2%.

TABLE II: Helicity preservation, with and without correction term, throughout L-BFGS descent for different choices of precision.

precision	$ \Delta\mathcal{H}/\mathcal{H} $	
	plain step	corrected step
mixed	1.00×10^{-5}	1.86×10^{-7}
float64	1.03×10^{-5}	1.73×10^{-15}

Remark. We do, however, see that the effects here are much milder than those shown in [40, Section 4.2.2], where the scheme without auxiliary variables dissipated the helicity entirely in only 1,000 steps, presumably due to the choice of larger time-steps.

³ The high-res data file for this configuration can be found at https://github.com/landreman/regcoil/blob/master/equilibria/wout_1i383_1.4m.nc

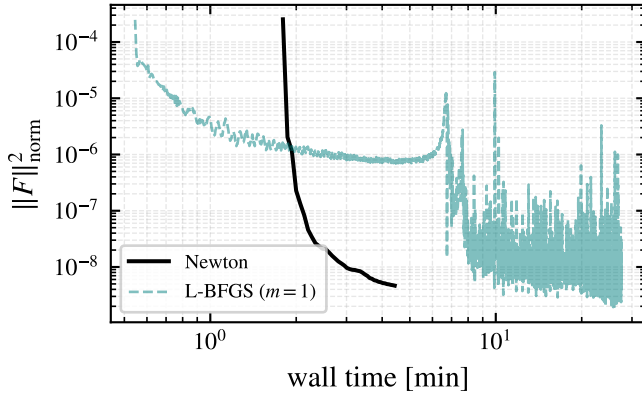


FIG. 9: Squared force residual against wall time for L-BFGS($m = 1$) (teal, dashed) and Newton's method (black, solid). Newton's method reaches the discretization floor in a fraction of the time and does not exhibit the same oscillatory behavior of $\|F\|^2$.

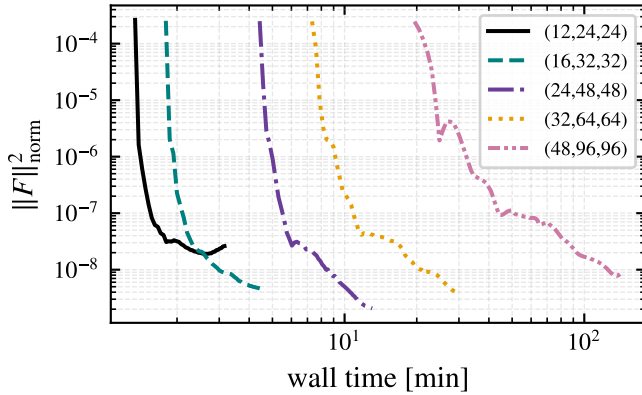


FIG. 10: Newton's method at different resolutions: most evident is the computational cost increase as $\propto n^{2.12}$ (Table III). All resolutions share a floor as more and more current sheets are resolved as h shrinks.

6.4. Newton's method: convergence to the floor

The addition of Newton's method speeds the method significantly and allows us to reach the discretization floor with normalized square force residuals $< 10^{-8}$. Results are shown in Table III and Fig. 9.

Computations at reasonable resolutions (e.g. $24 \times 48 \times 48$, 149,856 degrees of freedom) reach $< 10^{-8}$ in less than ten minutes.

The cost per step increases as $n_r^{2.1}$, below the rate of total degrees of freedom $\propto n_r^3$. The residual floor itself does not move appreciably with the resolution, but we note that helicity preservation improves, as expected, like n^p ($p = 2$).

In Appendix 2, we show Poincaré sections of the relaxed states at three resolutions. Only grid-scale features differ between them, while p and ι profiles and force residuals are in agreement across resolutions.

7. RECONNECTION

Magnetic reconnection is what allows the field line topology to change. It leads to the opening and closing of magnetic islands near resonant surfaces and current sheets. Reconnection happens either on grid-scale through finite-resolution effects or can be actively enabled by adding resistivity (η).

We discuss resistive reconnection in 7.1 and the targeted seeding of magnetic islands in 7.2. Hysteresis (the dependence of the relaxed state on the initial condition) is discussed in 7.3, and the pressure within magnetic island regions in 7.4.

7.1. Resistive reconnection

It is possible to induce magnetic topology changes by interjecting the relaxation process with a single resistive solve of the form $(\text{Id} - \eta \Delta) B^{n+1} = B^n$.

a. Controlling the amount of reconnection The natural question to ask is how to select η . We do so by assigning a certain helicity budget to the reconnection event.

Corollary 32 (Resistive dose). *The size of η can be chosen to target a given relative change in helicity: by Lemma 16,*

$$\frac{\delta \mathcal{H}}{\mathcal{H}} \approx -2\eta \frac{\langle J, B \rangle_\Omega}{\langle A, B \rangle_\Omega}. \quad (70)$$

For the modest helicity changes within the experiments reported here, we empirically observe that island width increases approximately with the square root of relative helicity change.

With resistive reconnection studies, it is possible to relax a state for a given time, store the reached state closer to equilibrium and perform a reconnection step. The reconnection step drops energy at the rate $-\eta \|J\|_\Omega^2$, the squared current norm at the rate $-2\eta \|\text{curl } J\|_\Omega^2$, and changes the helicity by the desired signed rate $-2\eta \langle J, B \rangle_\Omega$.

Plasma β typically drops along with J . Since the energy is essentially unchanged and $\text{grad } p = J \times B$, the relative drop of β to first order in η is that of J i.e. $-\eta \|\text{curl } J\|_\Omega^2 / \|J\|_\Omega^2$. This is assuming that the pressure-carrying perpendicular current scales with total J . Hence,

$$\frac{\delta \beta / \beta}{\delta \mathcal{H} / \mathcal{H}} \approx \frac{1}{2} \frac{\|\text{curl } J\|_\Omega^2}{\|J\|_\Omega^2} \frac{\langle A, B \rangle_\Omega}{\langle \text{curl } A, \text{curl } B \rangle_\Omega} \quad (71)$$

In the numerical experiments this factor is 5 to 7: a 1% helicity change drops β by 5 to 7%.

The reconnected field will sit at a much larger force balance, from where the relaxation can be continued to a (potentially lower) energy state that was previously topologically inaccessible. This new state might also allow for a lower force residual.

TABLE III: Comparing the relaxation performance for L-BFGS and Newton's method.

method	resolution	steps	s/step	time	$\ F\ _{\text{norm}}^2$	steps to $\ F\ _{\text{norm}}^2$			time to $\ F\ _{\text{norm}}^2$ [min]			$\Delta\mathcal{H}/\mathcal{H}$
			[s]	[min]	floor $\times 10^{-9}$	10^{-6}	10^{-7}	10^{-8}	10^{-6}	10^{-7}	10^{-8}	$\times 10^{-6}$
L-BFGS($m=1$)	$16 \times 32 \times 32$	5000	0.32	27.5	1.88	320	1190	1481	2.3	6.7	8.3	-10.6
Newton	$12 \times 24 \times 24$	40	2.80	3.2	19.1	3	6	—	1.4	1.6	—	-20.2
	$16 \times 32 \times 32$	40	4.05	4.4	4.66	4	6	19	2.0	2.1	3.0	-10.2
	$24 \times 48 \times 48$	60	8.76	13.0	2.06	5	9	30	5.0	5.6	8.7	-5.1
	$32 \times 64 \times 64$	80	16.66	29.3	4.03	8	14	51	9.3	10.9	21.2	-2.4
	$48 \times 96 \times 96$	140	53.12	142.6	7.75	16	29	121	32.8	44.3	125.7	-0.9
fit $\propto n_r$			$n_r^{2.12}$	$n_r^{2.76}$		$n_r^{1.17}$	$n_r^{1.16}$	$n_r^{1.70}$	$n_r^{2.26}$	$n_r^{2.42}$	$n_r^{3.38}$	$n_r^{-2.19}$

In the limit of infinitely many reconnection steps, all current is dissipated and the vacuum field $\mathbf{h}$ is recovered.

Remark. MRX supports check-pointing similar to standard optimization libraries. It is possible to launch another relaxation run starting from the stored state without any reconnection step and compare the results.

In Fig. 11, we show a relaxation run with four reconnection events, targeted to change the current helicity by 2.5% each time. We see, as expected, a stepped curve with helicity practically constant between reconnection events as well as a stepped decrease in β .

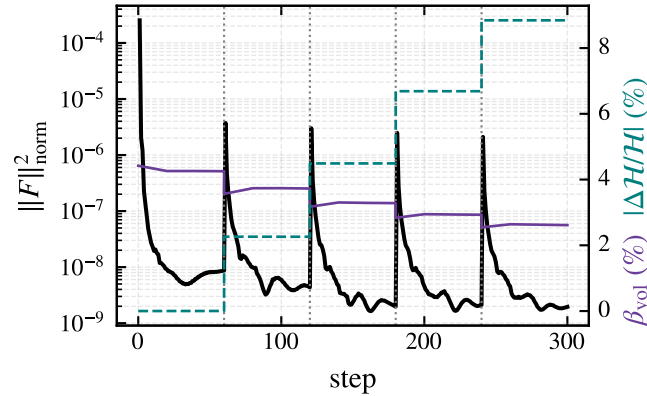


FIG. 11: Relaxation of the NCSX configuration with Newton's method and four resistive solves (dotted lines), each with a budget of 2.5% of the helicity. Between resistive solves, the helicity is constant to 10^{-7} . Poincaré sections are shown in Fig. 12.

In Fig. 12, we plot Poincaré sections of the field just before the first and the second reconnection event as well as at the end of the relaxation, i.e. after 60, 120 and 300 Newton steps. Islands increase in size between the reconnection events, and the final equilibrium exhibits a large 3/5 island chain.

7.2. Island seeding

An island chain at the resonance $r : |\iota(r)| = n_{\text{fp}} n/m$ can be seeded in the initial condition by perturbing the toroidal component of the reference potential,

$$\hat{A}_\zeta \mapsto \hat{A}_\zeta + \frac{a |\partial_r \Psi_t(r_0)|}{m} g(r) \cos(2\pi(m\theta - n\zeta)), \quad (72)$$

where $g(r) = \exp(-(r-r_0)^2/w^2) \frac{1-r^2}{1-r_0^2}$, r_0 is the resonant surface of the equilibrium, w the radial width of the seed, and $a = |\delta \hat{B}^r|/|\hat{B}^\zeta|$ at r_0 . The factor $(1-r^2)$ leaves the wall trace unchanged. Since the seed is added to the potential, $\text{div } B = 0$ holds exactly as well. However, helicity, β , and force balance are all modified slightly.

The pendulum estimate gives an island of full width $\approx \sqrt{8a n_{\text{fp}}/\pi m} |\partial_r \iota|$ in r . This is using the single-harmonic result from [107, Section 2.11] with χ_{mn} therein replaced by $a = |\delta \hat{B}^r|/|\hat{B}^\zeta|$, and our conventions for θ and ζ .

We exemplify this by seeding an island chain at the initial condition at two resonances: 3/5 for one case and 1/2 for the second.

The resulting Poincaré sections are shown in Fig. 13 and diagnostics in Table IX, compared to a run with no initial seeding. We see that the topology set by the seeded initial condition is largely preserved. Interestingly, our choice of seed for the 3/5 island chain differs from that formed via resistive reconnection by a phase shift. The island chains seeded here also appear more stochastized and exhibit more small scale spatial structure at the end of relaxation. Despite the differences appearing in the Poincaré plots, including the obvious differences appearing in the island chains, the final helicity, energy, volume-averaged- β , and force balance are in very large agreement; the final energies agree to five digits, the final helicities to 4 digits. From these volume-averaged quantities, one cannot distinguish between these very different magnetic fields.

The relative differences in energy, helicity, plasma beta and force residual between the three runs are small. The energies agree to 2×10^{-7} initially and to 10^{-5} or better at the end; the helicities to 3×10^{-5} initially and to 10^{-4} at the end. Plasma β agrees to 10^{-4} initially and to 5×10^{-3} at the end. The initial force residuals agree to

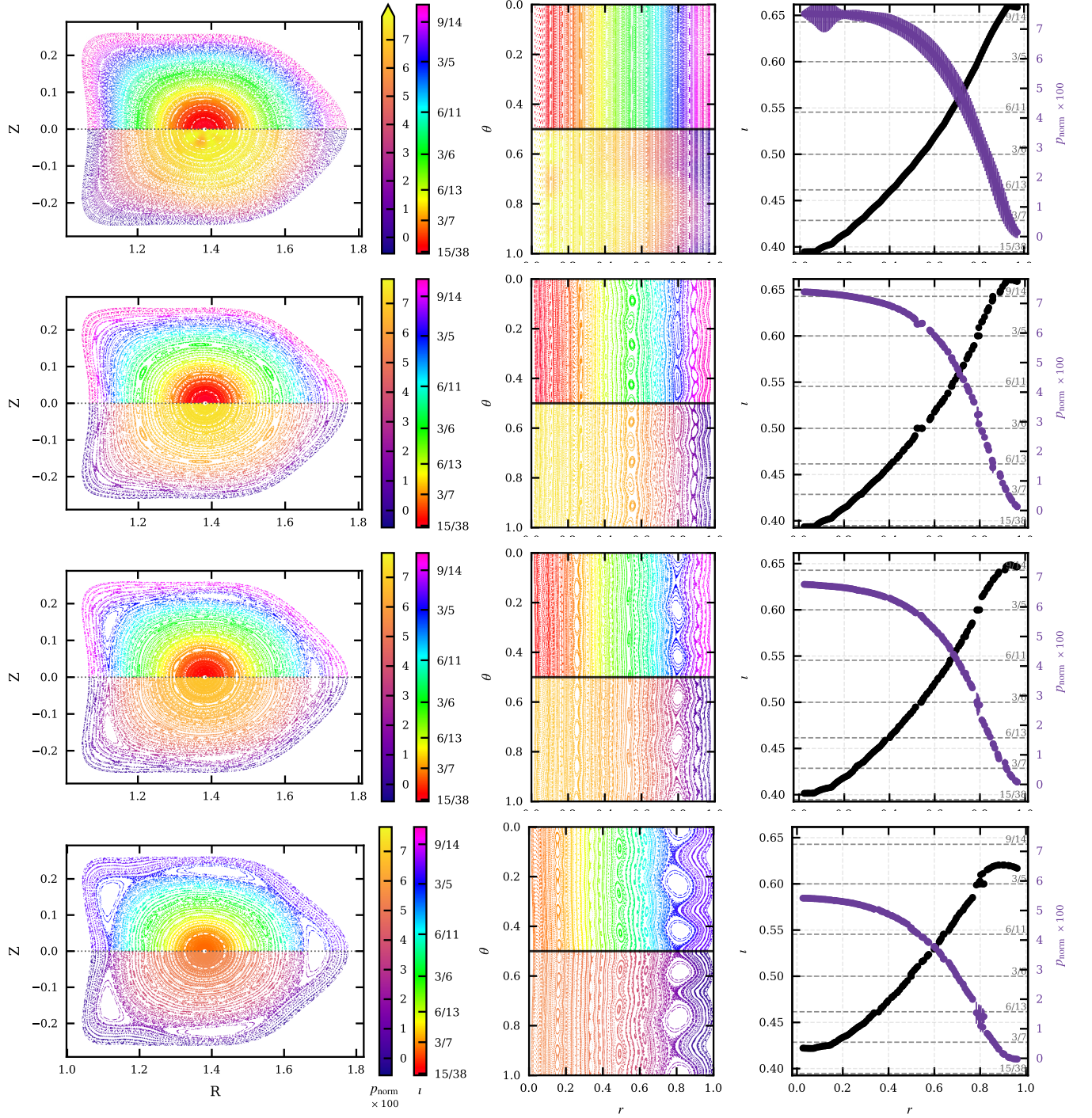


FIG. 12: Reconnection experiment: Poincaré sections at $\zeta = 0.5$ of a field period. From top to bottom: the initial condition, the state just before the first and the second reconnection step, and the final state. The $3/5$ chain, absent from the initial field, opens to $0.8 h_r$ before any resistivity is applied. It then grows to $1.5 h_r$ after the first resistive solve, 2.0 and $2.5 h_r$ after the second and third, and reaches $3.0 h_r$ in the final configuration.

1%, at the floor they differ by a factor of 1.1 ($\iota = 1/2$) and 2.8 ($\iota = 3/5$) from the unseeded run.

7.3. Hysteresis

The previous section showed that the final relaxation state can be heavily influenced by the initial condition be-

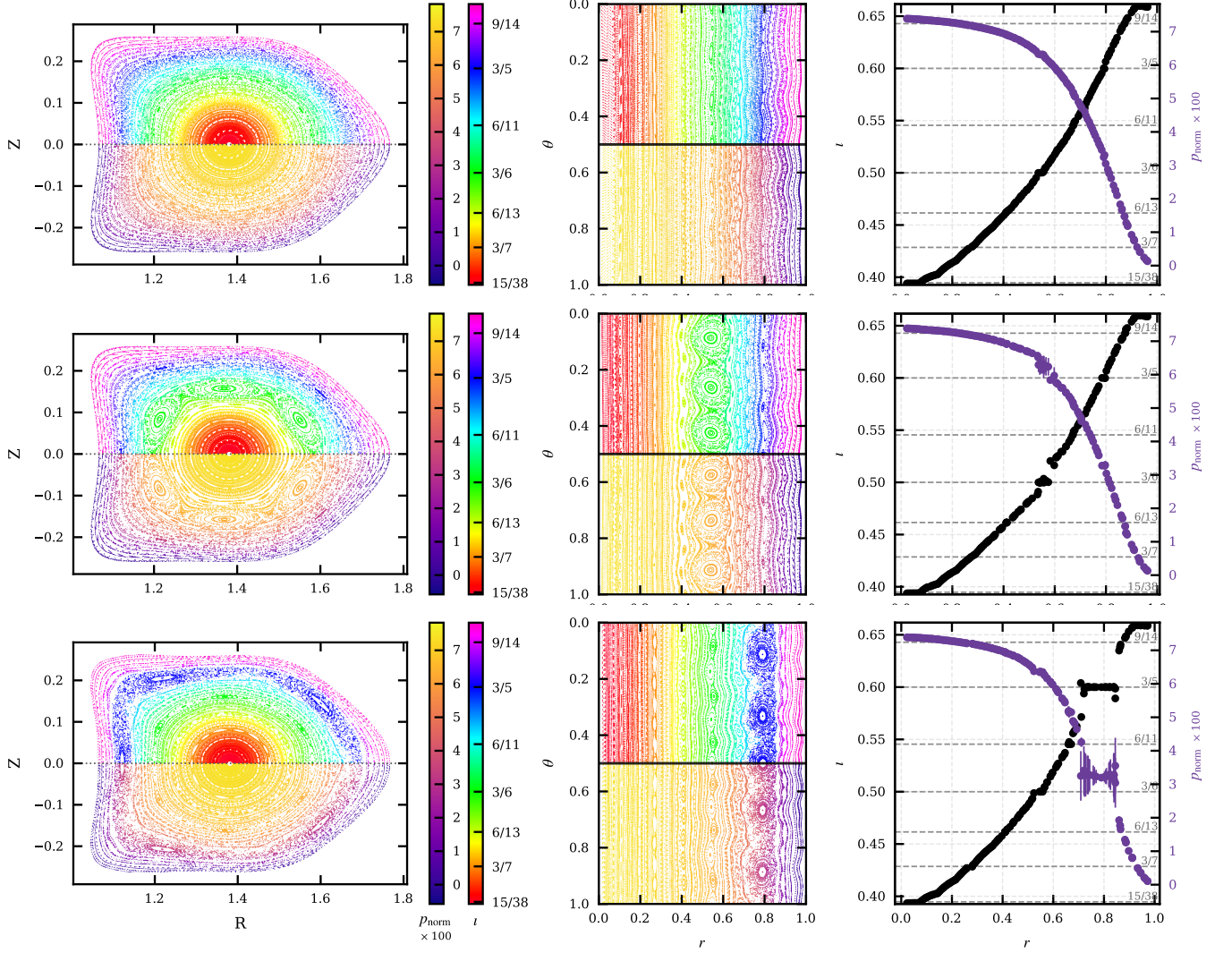


FIG. 13: Island seeding. Poincaré sections of the states of lowest force residual, reached in about 40 Newton steps (Table IX), at $\zeta = 0.5$. The tiles are as in Fig. 5. Initial conditions with no island seed (top), seeded at $\iota = 1/2$ (center) and at $\iota = 3/5$ (bottom) with amplitude $a = 10^{-2}$. The seeded chains, $2.7 h_r$ and $2.1 h_r$ wide, survive the Newton relaxation within $\pm 10\%$ of their initial width.

cause the initial state and its topological features can be frozen in, up to grid-scale reconnection processes. The next question we address is how strongly the relaxed state is determined by the initial condition the relaxation starts from. Our intuition was that if some resistive reconnection was introduced to climb the topological barriers from the island seeding, the initial states would converge to the same final equilibrium. This is similar in spirit to a real experimental start-up phase; the plasma begins in a cold resistive state, and heats up to $\sim \text{keV}$ temperatures where resistivity is very small and the equilibrium can get frozen-in.

We investigated this question using the following experiment. Starting from the same three initial conditions from Fig. 13, i.e. no seed, a seed at $\iota = 1/2$, and one at $\iota = 3/5$, we perform 300 Newton steps with

TABLE IV: Diagnostics for the seeded reconnection runs from Fig. 14 after 300 Newton steps.

seed at (ι, r)	$\Delta\mathcal{E}/\mathcal{E} \times 10^4$	$\Delta\mathcal{H}/\mathcal{H} \times 10^2$	$\beta_{\text{vol}}[\%]$	$\ F\ _{\text{norm}}^2 \times 10^9$
—	−3.88	−10.42	2.43	1.97
(1/2, 0.544)	−3.90	−10.42	2.43	1.98
(3/5, 0.794)	−3.90	−10.39	2.42	1.62

four resistive steps. The helicity budget is slightly higher for this experiment, 3% per reconnection, and leads to $|\Delta\mathcal{H}/\mathcal{H}| = 10.4\%$ at the final state.

Results are shown in Fig. 14. The resulting fields largely agree, with a big island chain near $r = 0.8$. As shown in Table IV, the relative changes in helicity, energy

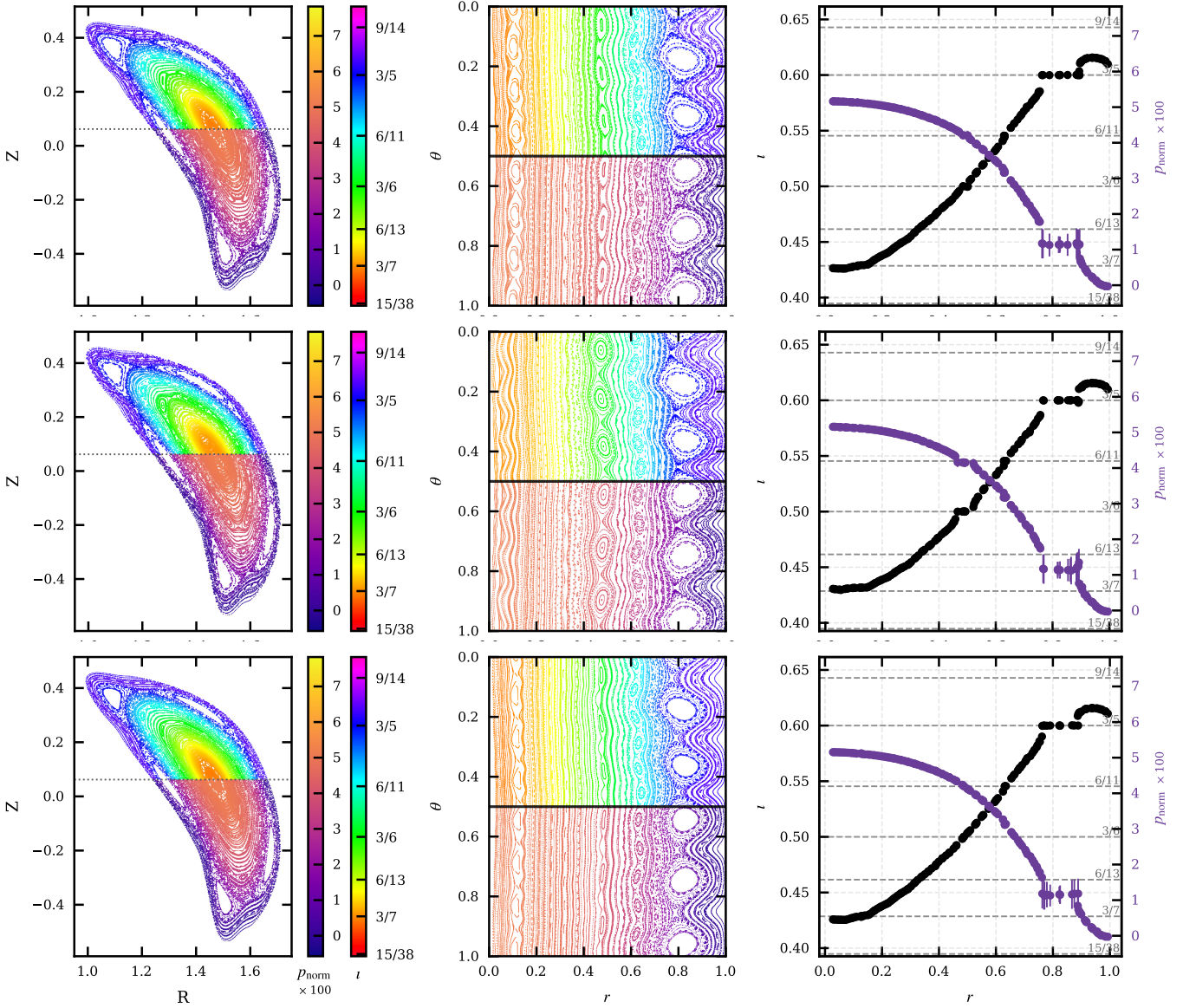


FIG. 14: Island seeding with reconnection. Poincaré sections of final states after 300 Newton steps and four resistive solves, at $\zeta = 0.25$.

and plasma β differ by $< 1\%$ relatively to one another. The force residua are all $< 2 \times 10^{-9}$.

The island widths over relaxation time are shown in Table V. We see that they converge to one another with every reconnection step. The only exception is the seeded $1/2$ island chain, which retains a width $> 1.0 h_r$ until the end of the run.

This is a very encouraging result. Given the freedom to re-organize the magnetic field, different initial conditions relax to similar final configurations. Again, the final energy, helicity, volume-averaged- β , and force balance are in very precise agreement.

The discrepancy $\Delta B := \max_{\text{seed}} \|B_{\text{seed}} - B_{\text{unseeded}}\|_{\text{rel}}$ drops from initially 9.1×10^{-4} to 1.8×10^{-4} at the final state. Every reconnection roughly halves the discrepancy

between the seeded runs and the unseeded reference.

In terms of other diagnostics, the three runs start within 1% of one another in the force residual and, as in Sec. 7.2, within 2×10^{-7} , 3×10^{-5} and 10^{-4} in energy, helicity and plasma beta. After 300 steps and four reconnections they agree to 2×10^{-6} in energy, to 4×10^{-4} in helicity and to 5×10^{-3} in plasma beta.

7.4. Island pressure

Typically, pressure profiles are flattened near the O-point of a magnetic island, although this depends on the width of the island; smaller islands are better able to maintain a pressure gradient across them [108].

TABLE V: Widths of the 3/5, 1/2 and 3/7 island chains in the seeded reconnection runs, in units of the radial cell $h_r = 1/16$.

chain	seed	initial	before reconnection event				final
			1	2	3	4	
3/5	–	0.0	0.8	1.6	2.2	2.7	3.4
	3/5	2.3	1.9	1.2	1.6	2.5	3.3
	1/2	0.0	0.7	1.7	2.2	2.7	3.4
1/2	–	0.0	0.9	0.4	0.4	0.4	0.9
	3/5	0.0	1.0	0.9	1.0	0.8	0.6
	1/2	2.6	2.4	2.0	1.6	1.6	1.5
3/7	–	0.0	0.3	0.5	0.3	0.6	0.9
	3/5	0.0	0.0	0.0	0.2	0.6	0.6
	1/2	0.0	0.4	0.4	0.4	0.8	0.6

If we look at the pressure profiles in Figures 12, 13, and 14, we see that the spread of p along field-lines (the error bars in these plots) is largest for those field-lines that lie in magnetic islands. The flattening is most clear in the large islands shown in Fig. 14. However, the spread of p is a sign that p retains a spatial gradient in r .

This is expected. A discrepancy $p|_{r_{\text{in}}} - p|_{r_{\text{out}}}$ between the island’s inner (high pressure) and outer (low pressure) edge leads to a parallel force residual, as $B \cdot F = -B \cdot \text{grad} p$. If we denote by ℓ the length a field line needs to move from r_{in} to r_{out} , this parallel force residual is of order $(p|_{r_{\text{in}}} - p|_{r_{\text{out}}})/\ell$. The length ℓ is large, tens of toroidal transits in our examples. This is in line with the Cary–Hanson formula [109, Equations (44-45)], which predicts one full rotation around an O-point every $4/(m|\partial_r \psi|w)$ toroidal transits, with m the poloidal mode number and w the island width. For this reason, the parallel force is small, and its impact on the volume force residual is smaller still by the island’s volume fraction; it is an indirect route to address the pressure flattening by minimizing the full volume-averaged force residual, so this process is rarely complete at the end of the relaxation.

8. DISCUSSION AND OUTLOOK

In this work, we presented a magnetic relaxation method in the spirit of [47] and [37] in order to compute MHS equilibria without assuming nested flux surfaces. We demonstrated some of the capabilities of this approach, focusing on the structure-preserving properties, including the crucial conservation of helicity and $\text{div} B = 0$.

Open questions and future work are discussed in 8.1 and 8.2.

To summarize, this work has demonstrated:

- High resolution convergence studies and relaxation runs.

- A principled mathematical approach, based on the discrete Hodge decomposition on non-simply connected domains. Control over the amount of reconnection and helicity preservation.
- Very computationally efficient non-nested equilibrium calculations (e.g. finite- β NCSX runs with $16 \times 32 \times 32$ resolution can be converged to 10^{-6} normalized force balance in 2 minutes and to the floor in 4, see Table III).
- Differentiable non-nested equilibrium calculations, including for future incorporation into stellarator optimization routines.

8.1. Open questions about convergence

Questions remain about what constitutes true convergence. We know any faithful relaxation method will move towards states with singular current sheets. In this work, we chose to modify our line search to halt relaxation once the discretization floor is reached rather than continue to reduce the magnetic energy at the cost of grid-scale reconnection, leading to an increase in force balance residual. Numerous other regularization choices are possible and are the subject of future work.

Our numerical results also point to the potential existence of lots of nearby equilibria or “almost-equilibria” (qualifying the fact that no solutions are ever at machine-precision normalized force balance), supporting the discussion in Sec. 1.5. Our various experiments seem to support this conclusion; Fig. 12 relax to states with different island sizes determined by the initial seeding, but all satisfy final normalized force balance to 10^{-8} with essentially identical final energy and helicity, and the resistive solutions in Fig. 14 mostly converge but do not perfectly do so, despite having essentially identical relaxation dynamics (final absolute energy, helicity, volume-averaged- β , and force residual). Nonetheless, it is a useful observation that when the topological barriers are reduced, the convergence of the final state, with respect to the initial state, is substantially improved. It remains plausible from these results and the previous literature that the number of nearby equilibria from small changes in the helicity, volume-averaged- β , or further convergence towards force balance may be very large. This opens up future work that should clarify how 3D non-nested equilibrium codes should be run to match experimental data.

8.2. Future work

There are several useful directions for future work.

a. Outer loop optimization There are many optimization, stability, and control problems to solve for nuclear fusion devices, many of which depend on the MHS solution. In the context of stellarator optimization, the computation of MHD equilibria comprises a constraint

within an outer optimization loop. In particular, when $Q : H_0^{\text{div}}(\Omega) \rightarrow \mathbb{R}$ is a given function that measures the quality of a magnetic field configuration for the sake of some application (e.g. quality of particle confinement, engineering feasibility), then the full optimization problem reads:

$$\min_{B \in H_0^{\text{div}}(\Omega)} Q(B) \quad \text{s.t.} \quad \text{curl } B \times B = \text{grad } p, \quad \text{div } B = 0.$$

Given a parametrized mapping $\Phi_\alpha : \hat{\Omega} \rightarrow \Omega$, we can also pose the problem

$$\begin{aligned} \min_{\alpha \in \mathbb{R}^{n_r n_\theta n_\zeta}} Q\left((\Phi_\alpha)_*^2 \hat{B}\right) \quad \text{s.t.} \quad \text{div } B = 0 \\ \text{and } \text{curl}((\Phi_\alpha)_*^2 \hat{B}) \times (\Phi_\alpha)_*^2 \hat{B} = \text{grad } p, \end{aligned} \quad (73)$$

where Φ_α is a C^1 diffeomorphism. An example for this is a spline map of the form $[\Phi_\alpha]_j = \sum_i \mathbf{a}_{ij} \Lambda_i^0$, where $j \in \{1, 2, 3\}$ and $\mathbf{a} \in \mathbb{R}^{n_r n_\theta n_\zeta \times 3}$. All experiments in this work are done using such maps.

To solve this shape optimization problem, we require access to gradient information of the objective with respect to the optimization parameters, $\{\partial_\alpha(\Phi_\alpha)_*^2 \hat{B}\}_i$. JAX's automatic differentiation tools provide automatic, highly efficient gradients. With this functionality in place, the presented code could be used as a back-end for equilibrium calculation in SIMSOPT [110] and other stellarator optimization suites.

b. TPU support The fact that TPUs do not support any float64 representation natively is a significant challenge for runs that aim to resolve a very small force residual. There are many applications, such as the optimization problems above, where very small ($\|F\|_\Omega^2 < 10^{-8}$) residual is not needed, but this manuscript was focused on demonstrating convergence, scaling, and structure-preservation. This is an open problem at the moment, potentially solvable by emulating double precision with several float32 variables. This is, however, non-trivial to implement without performance hits.

c. Free boundary solves Free boundary solves with coils are the most robust check for designing the magnetic fields of a stellarator experiment. We are also able to replace the boundary condition $B \cdot n = 0$ by $(B - B_{\text{vac}}) \cdot n = 0$ when we set $v = 0$ on the boundary. This can be done with a penalty method and will allow us to move beyond the case where the computational boundary is a conducting wall or a flux surface [111].

d. More general maps All maps in this work are of the form from Equation (20). Future work will illustrate that we have also implemented (1) the capability to exploit further symmetries like stellarator symmetry to save computational cost and (2) extended the code to maps that are more general (e.g. knotted devices) and not dependent on an initial flux surface solution, using a re-implementation of map2disc [70] in JAX within MRX.

e. Comparison of fixed-boundary codes for finite- β We noted in the introduction and Table I that HINT,

SPECTRE, SIESTA are in use for fixed-boundary finite- β stellarator equilibrium calculations. A forthcoming publication will perform comparisons between these codes and MRX.

ACKNOWLEDGMENTS

We would like to thank Omar Maj for pointing out the work of Chodura & Schlüter and many enlightening discussions on the topic over the last years. The authors would also like to thank Robert Babin, Martin Campos-Pinto, Mark Ciaciosa, Yaman Güçlü, Christopher Hansen, Florian Hindenlang, Elizabeth Paul, Francesco Patrizi, Benjamin Peherstorfer, Stefan Possanner, Matteo Raviola, Georg Stadler, and Vlad Vicol for valuable discussions.

This work was supported through grants from the Simons Foundation under award 560651. We gratefully acknowledge the support of a Google TPU Research Award for providing funding and TPU computing resources.

USE OF LANGUAGE MODELS

Large language models developed by Anthropic PBC and OpenAI PBC were used throughout this work as coding assistants, for implementing and refactoring parts of the MRX code base and its test suite, for setting up, bookkeeping, and analyzing the numerical experiments, and for generating figures and tables from the run records. They were also used to proofread the manuscript and to check derivations. All code and text was reviewed and edited by the authors, who take full responsibility for the content of this article.

Appendix A: Proofs

Proof of Lemma 4. Since $\Phi^{-1}(\partial\Omega) = \partial\hat{\Omega}$, it holds for all $x \in \partial\Omega$ that

$$(\Phi_*^0 \hat{f})(x) = \hat{f}(\Phi^{-1}(x)) = 0. \quad (\text{A1})$$

Next, for $\hat{x} \in \partial\hat{\Omega}$,

$$D\Phi(\hat{x})^T E(\Phi(\hat{x})) = \hat{E}(\hat{x}) = \hat{E}_r(\hat{x}) \hat{e}_r, \quad (\text{A2})$$

with $\hat{E}_r \in \mathbb{R}$ and $\hat{e}_r$ denoting the unit vector in r -direction. This implies $\partial_\theta \Phi \cdot (E \circ \Phi) = \partial_\zeta \Phi \cdot (E \circ \Phi) = 0$. The tangent vector at Φ is given by $\tau = \tau_1 \partial_\theta \Phi + \tau_2 \partial_\zeta \Phi$, hence $\tau \cdot (E \circ \Phi) = 0$. Lastly, for $\hat{x} \in \partial\hat{\Omega}$,

$$\hat{B}(\hat{x}) = \hat{B}_\theta(\hat{x}) \hat{e}_\theta + \hat{B}_\zeta(\hat{x}) \hat{e}_\zeta, \quad (\text{A3})$$

hence

$$\begin{aligned} B \circ \Phi &= (\det D\Phi)^{-1} D\Phi \hat{B} \\ &= (\det D\Phi)^{-1} (\hat{B}_\theta \partial_\theta \Phi + \hat{B}_\zeta \partial_\zeta \Phi). \end{aligned} \quad (\text{A4})$$

At the same time, the normal vector n is proportional to $\partial_\theta \Phi \times \partial_\zeta \Phi$ and therefore orthogonal to $B \circ \Phi$. $\square$

Proof of Lemma 16. The first identity is Proposition 12 with $v \times B$ replaced by $v \times B - \eta J$, since $\langle J, \eta J \rangle_\Omega = \eta \|J\|_\Omega^2$. For the second, the harmonic component of B is unchanged by admissible variations, so $\delta_v \mathcal{H} = \langle \delta_v A, B \rangle_\Omega + \langle A, \delta_v B \rangle_\Omega$.

Now substitute $\delta_v B$ and $\delta_v A$ by their definitions to reach $\delta_v \mathcal{H} = 2\langle B, v \times B - \eta J \rangle_\Omega = -2\eta \langle J, B \rangle_\Omega$. $\square$

Proof of Lemma 25. Let $B_{\mathfrak{H}} = \Pi_{\mathfrak{H}_0^2} B^n$, constant in n because $\text{curl } E \perp \mathfrak{H}_0^2$. Then $\text{curl}(A^{n+1} - A^n - \delta t E) = 0$ with $A^{n+1} - A^n - \delta t E \in V_0^1$. Hence this expression is a gradient: $A^{n+1} - A^n - \delta t E = \text{grad } \phi$, $\phi \in V_0^0$ (as $\dim \mathfrak{H}_0^1 = 0$ in our case), and $\langle \text{grad } \phi, B^{n+1} + B_{\mathfrak{H}} \rangle_\Omega = 0$ since both fields are divergence-free with vanishing normal trace. It follows that

$$\begin{aligned} \mathcal{H}^{n+1} - \mathcal{H}^n &= \langle A^{n+1} - A^n, B^{n+1} + B_{\mathfrak{H}} \rangle_\Omega \\ &\quad + \langle A^n, B^{n+1} - B^n \rangle_\Omega \\ &= \delta t \langle E, B^{n+1} + B_{\mathfrak{H}} \rangle_\Omega + \delta t \langle \text{curl } A^n, E \rangle_\Omega \\ &= \delta t \langle E, B^{n+1} + B_{\mathfrak{H}} \rangle_\Omega + \delta t \langle B^n - B_{\mathfrak{H}}, E \rangle_\Omega \\ &= \delta t \langle E, B^{n+1} + B^n \rangle_\Omega, \end{aligned} \quad (\text{A5})$$

by partial integration (with both A^n and E in V_0^1) and $\text{curl } A^n = B^n - B_{\mathfrak{H}}$. $\square$

Proof of Proposition 27. From $B^{n+1} = B^n + \delta t \text{curl } E_\sigma$, $\mathcal{E}^{n+1} - \mathcal{E}^n = \delta t \langle B^n, \text{curl } E_\sigma \rangle_\Omega + \frac{\delta t^2}{2} \|\text{curl } E_\sigma\|_\Omega^2$. Since $E_\sigma \in V_0^1$, the definition of the weak curl operator gives $\langle B^n, \text{curl } E_\sigma \rangle_\Omega = \langle J, E_\sigma \rangle_\Omega$. As $J \in V_0^1$ and Π_0^1 is L^2 -orthogonal,

$$\begin{aligned} \langle J, E_\sigma \rangle_\Omega &= \langle J, v \times B^n \rangle_\Omega - \sigma \langle J, B^n \rangle_\Omega \\ &= -\langle v, J \times B^n \rangle_\Omega - \sigma \langle J, B^n \rangle_\Omega \\ &= -\langle v, J \times B^n - \widetilde{\text{grad } p} \rangle_\Omega - \sigma \langle J, B^n \rangle_\Omega, \end{aligned} \quad (\text{A6})$$

since v is divergence-free. Substitute $(\text{Id} + \varepsilon \text{curl } \widetilde{\text{curl}}) v = J \times B^n - \widetilde{\text{grad } p}$ and the claim follows. $\square$

Appendix B: Additional numerical results

1. Further numerical results: L-BFGS

This section discusses further numerical results and hyperparameter sweeps for the L-BFGS approach.

a. L-BFGS memory The impact of varying the L-BFGS memory parameter m is shown in Fig. 15. Steepest descent ($m = 0$) decreases the force monotonically but reaches only a force residual of 10^{-7} . Non-linear conjugate gradient ($m = 1$) reaches 10^{-8} at step 1481 and $m = 5$ at step 1142, both with a floor of 2×10^{-9} . The cost per step is essentially the same, 0.30, 0.32 and 0.33 s for $m = 0, 1, 5$.

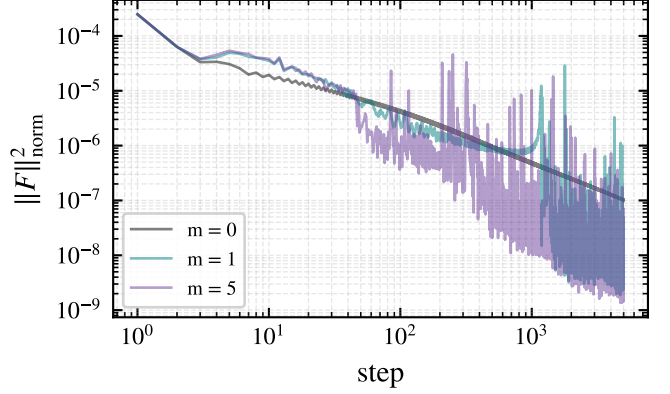


FIG. 15: Effect of the L-BFGS memory.

b. Velocity smoothing Table VI shows the hyperparameter sweep for the velocity smoothing strength. For some values of c , the L-BFGS memory is dropped on most steps (Powell's restart, last column), the method degenerates to steepest descent and stalls between 5×10^{-8} and 1.4×10^{-7} .

TABLE VI: Sweeping the smoothing constant c in $\varepsilon = c h_r^2$. Smoothing too little breaks the L-BFGS method (row one), while smoothing too much makes the method unable to resolve the small features needed to drop the force residual below 10^{-6} .

c	s/step	$\ F\ _{\text{norm}}^2$ floor $\times 10^{-8}$	steps to $\ F\ _{\text{norm}}^2$			restarts	$\Delta \mathcal{H} / \mathcal{H}$
			10^{-6}	10^{-7}	10^{-8}	[%]	$\times 10^{-6}$
0.004	0.29	14.4	1013	—	—	100	-7.4
0.02	0.32	0.19	320	1190	1481	17	-10.6
0.1	0.36	5.13	157	2626	—	62	-8.0
0.5	0.56	13.9	214	—	—	93	-8.6

c. Velocity formulation Table VII compares the two velocity formulations from Section 5.3 on the NCSX case at the reference configuration. The residual traces agree at every step to within the round-off that separates the two trajectories: the same steps to 10^{-6} and 10^{-7} (320 and about 1190), 10^{-8} at step 1481 against 1542, floors of 1.9 and 4.0×10^{-9} , and the same helicity change, as expected from the discrete sequence properties.

The potential route costs 27% less per step (0.32 against 0.44 s), replacing the $k = 3$ Laplacian solve with a $k = 1$ one and the ε -shifted $k = 2$ solve with a $k = 1$ one.

2. Further numerical results: Newton

We give some further numerical results for Newton's descent method in this section.

a. Precision Solver precision limits the accuracy of the descent direction and consequently the force residual

TABLE VII: Velocity formulations: potential and Leray projection. Performance is essentially identical, but the potential route is faster and extends to Newton’s method.

velocity	s/step	$\ F\ _{\text{norm}}^2$	steps to $\ F\ _{\text{norm}}^2$			$\Delta\mathcal{H}/\mathcal{H}$ $\times 10^{-6}$
		floor $\times 10^{-9}$	10^{-6}	10^{-7}	10^{-8}	
potential	0.32	1.9	320	1190	1481	-10.6
Leray	0.44	4.0	320	1189	1542	-10.3

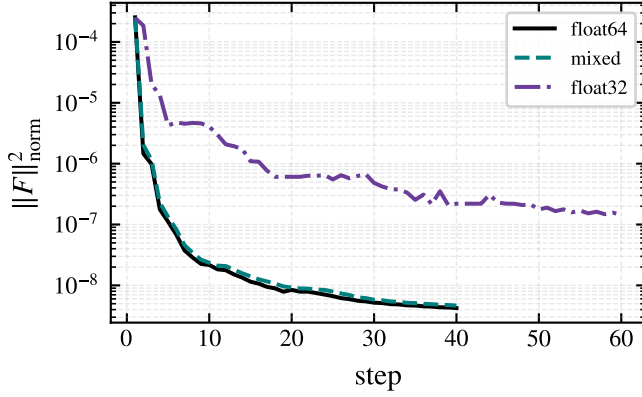


FIG. 16: Newton’s method in single, mixed (single-precision solves, double-precision residual) and double precision.

that can be reached. Fig. 16 shows that single precision floors at $\|F\|_{\text{norm}}^2 \approx 1.4 \times 10^{-7}$, while mixed and double precision agree. The tolerance of the inner solves does not enter: at 10^{-6} , 10^{-8} and 10^{-10} the three trajectories coincide.

Per step, mixed precision is approximately $1.5\times$ faster than float64. Single precision float32 is approximately $3.9\times$ faster than float64, but at the cost of residual quality.

b. Order See Fig. 17. Increasing the order past $p = 2$ does not pay off. All resolutions share a residual floor: the limit field is not regular enough to benefit from higher-order approximation. Linear/constant splines ($p = 1$, not shown), perform very badly in contrast and stall at $\|F\|_{\text{norm}}^2 > 10^{-4}$.

c. Hyperparameters The preconditioner of the Newton solve is built from the first term of the hessian (65) with the field replaced by its vacuum part, $(u, v) \mapsto \langle \text{curl}(u \times \text{ch}), \text{curl}(v \times \text{ch}) \rangle_{\Omega}$, which is fixed for the whole relaxation. This operator has a large nullspace, containing every displacement parallel to $\mathfrak{h}$, and it is small on the resonant modes, the displacements nearly constant along the field lines of $\mathfrak{h}$. The preconditioner therefore approximately inverts the shifted operator $(u, v) \mapsto \langle \text{curl}(u \times \text{ch}), \text{curl}(v \times \text{ch}) \rangle_{\Omega} + \kappa (2\pi|\mathfrak{h}|)^2 \langle u, v \rangle_{\Omega}$. The Newton system itself is not modified. SIESTA linearizes the same force operator but assembles it explicitly and factorizes it [37, Section IX.A]. Following, our matrix-free

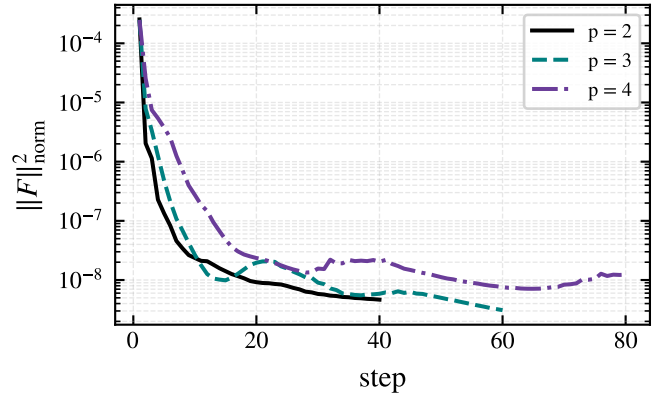


FIG. 17: Newton’s method for $p = 2, 3, 4$.

approach, in MRX it is only ever applied.

There are three hyperparameters that need to be set in the inexact Newton method: The preconditioner floor κ , the line search regularization C , and the number of MINRES iterations to approximately invert the hessian. Their optimal values have been determined by sweeps, shown in Table VIII.

The impact of the line search regularization parameter C is shown in Fig. 18.

For the right value of C , the line-search converges to very small steps at the discretization floor, as the term $\|\text{curl} B\|_{\Omega}^2$ along the step direction grows in the presence of numerous current sheets. The result is that the relaxation dynamics come to a gentle halt ($C \geq 0.1$, yellow, pink, and gray lines). For values of C that are too small, the accepted time-step stays near the Newton’s suggested $\delta t = 1$, and the force residual rises as reconnection on grid-scale levels takes place.

d. Resolution Fig. 19 shows the Poincaré sections of the relaxed states at three resolutions. The grid-scale island size and features are resolution-dependent, as the lowest resolution seems to open up a very small island chain and the highest resolution appears to roughly keep the nested surfaces. Nonetheless, the final pressure profiles, rotational transform profiles, and force balances are in very strong agreement across resolution.

e. Island seeding Table IX shows runtime and relaxation steps to the residual floor as well as energy and helicity change for the seeded relaxation runs. We observe that all runs perform similarly in these metrics despite their markedly different Poincaré cross-sections shown in Fig. 13.

Appendix C: Hyper-parameters

1. Numerical experiments

Table X lists the parameters used for all numerical experiments reported in this work.

TABLE VIII: The Newton configuration on the NCSX case: three one-parameter sweeps, each holding the other two settings at the values in its first column. MINRES iterations determine the speed of descent (6th column), κ determines the quality of the floor (4th column), and C determines how well the solution stays at that floor (Fig. 18).

sweep	level	s/step	$\ F\ _{\text{norm}}^2$ floor $\times 10^{-9}$	at step	1st step $< 10^{-8}$	$\Delta\mathcal{E}/\mathcal{E} \times 10^{-6}$	$\Delta\mathcal{H}/\mathcal{H} \times 10^{-6}$
MINRES it. ($\kappa = 3, C = 0$)	30 it.	1.6	17.7	100	—	-3.81	-8.73
	50 it.	2.3	8.8	70	58	-3.87	-18.21
	100 it.	4.1	4.9	33	19	-4.39	0.74
	200 it.	7.5	4.6	16	8	-5.86	67.56
	300 it.	11.1	4.5	10	5	-9.20	207.04
κ (300 it., $C = 0$)	$\kappa = 10^{-2}$	11.1	14.6	14	—	-7.45	91.44
	$\kappa = 0.1$	11.1	11.9	9	—	-9.86	125.35
	$\kappa = 1$	10.8	5.8	7	5	-10.84	216.23
	$\kappa = 3$	11.1	4.5	10	5	-9.20	207.04
	$\kappa = 10$	11.0	3.9	17	7	-5.48	46.93
C ($\kappa = 3, 100$ it.)	$C = 0$	4.1	4.9	33	19	-4.39	0.74
	$C = 0.01$	4.0	4.1	39	19	-4.31	-3.72
	$C = 0.03$	4.0	3.5	47	19	-4.04	-14.68
	$C = 0.1$	4.1	3.8	65	19	-3.95	-10.97
	$C = 0.3$	4.1	4.6	83	19	-3.92	-10.31
	$C = 1$	4.1	5.0	90	19	-3.92	-10.22

seed at (ι, r)	steps to the floor	time [min]	$\ F\ _{\text{norm}}^2 \times 10^8$	$\Delta\mathcal{E}/\mathcal{E} \times 10^6$	$\Delta\mathcal{H}/\mathcal{H} \times 10^6$	final chain width
—	39	4.4	0.46	-3.89	-10.2	—
(1/2, 0.544)	39	4.4	0.52	-4.18	-10.0	$2.7 h_r$
(3/5, 0.794)	41	4.4	1.28	-6.77	+6.4	$2.1 h_r$

TABLE IX: Island seeds on the NCSX case, Newton's method.

2. Code settings

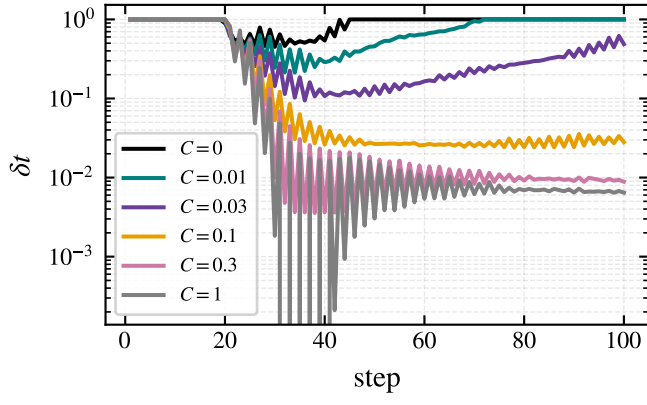
Finally, we provide a complete list of all user-set MRX parameters and their default values in Table XI.

TABLE X: The runs behind the paper’s figures and tables. The reference configuration is Newton’s method on the NCSX li383 case at $16 \times 32 \times 32$, $p = 2$, mixed precision, solver tolerance 10^{-8} , the harmonic atom at $\kappa = 3$ with a budget of 100 MINRES iterations, the regularized step with $C = 0.1$, a CFL cap of 0.5, and a budget of 200 steps from the VMEC field in chunks of 20, stopping at the floor (a chunk’s mean accepted step below 0.1 or its mean squared residual below 10^{-8}); no seed, no reconnection. For L-BFGS, the defaults are $m = 1$ with Powell’s restart, velocity smoothing with $\gamma = 1$ and $c = 0.02$, the potential velocity, $C = 0$ and 5,000 steps in chunks of 500 without a floor stop. Each row names what the experiment sweeps and what else deviates from the reference. A seed (m, n) has amplitude $a = 10^{-2}$ and width 0.1 at the resonant surface. A reconnection event is one resistive solve every 60 steps spending the stated share of the helicity.

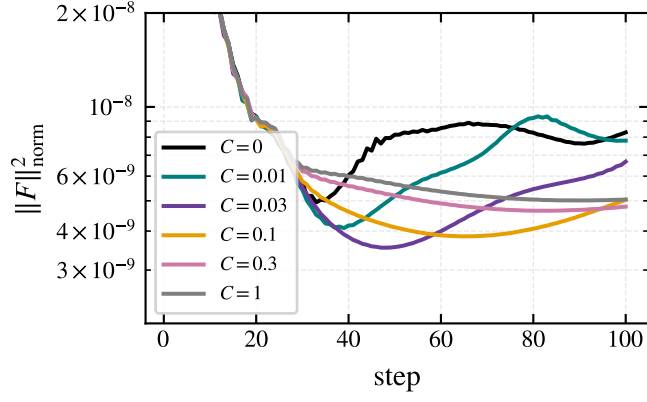
experiment	shown in	swept	otherwise
helicity preservation	Tab. II	precision mixed / float64, $\sigma = 0$ / $\sigma \neq 0$	L-BFGS, $\gamma = 0$, steps = 1000
L-BFGS memory	Fig. 15	$m \in \{0, 1, 5\}$	L-BFGS
Newton against L-BFGS, wall time	Fig. 9	Newton / L-BFGS	–
Newton convergence in n_r	Tab. III, Fig. 19	Newton / L-BFGS, $n_r \in \{12, 16, 24, 32, 48\}$	–
Newton: MINRES budget, atom, C	Tab. VIII, Fig. 18	$it \in \{30, 50, 100, 200, 300\}$, $\kappa \in \{10^{-2}, 0.1, 1, 3, 10\}$, $C \in \{0, 0.01, 0.03, 0.1, 0.3, 1\}$	steps = 100, no floor stop; $C = 0$ in the it and κ sweeps
Newton: resolution, degree	Fig. 10, Fig. 17	$n_r \in \{12, 16, 24, 32, 48\}$, $p \in \{2, 3, 4\}$	–
Newton: solver tolerance, precision	Fig. 16	precision mixed / float64 / float32, tol $\in \{10^{-6}, 10^{-8}, 10^{-10}\}$	–
reconnection ladder	Fig. 11, Fig. 12	–	steps = 300, $C = 0$, no floor stop, 4 reconnections at 2.5%
reconnection ladder, L-BFGS	Fig. 1	–	L-BFGS, Leray velocity, steps = 40000, 4 reconnections at 2.5%, every 8000 steps
seeded island chains	Tab. IX, Fig. 13	seed none / (6, 1) / (5, 1), steps $\in \{100, 200\}$	–
seeded island chains with reconnection	Tab. IV–V, Fig. 14	seed none / (6, 1) / (5, 1)	steps = 300, $C = 0$, no floor stop, 4 reconnections at 3%
velocity formulation	Tab. VII	velocity potential / Leray	L-BFGS
velocity smoothing	Fig. 8	$\gamma \in \{0, 1\}$, steps $\in \{5000, 20000\}$	L-BFGS
velocity smoothing constant	Tab. VI	$c \in \{0.004, 0.02, 0.1, 0.5\}$	L-BFGS

TABLE XI: MRX relaxation parameter choices and default values.

Parameter	Default	Notes
Initial condition: equilibrium file (<code>.wout/.dat</code>)	–	Also used to load the geometry.
Number of splines (n_r, n_θ, n_ζ)	(16, 32, 32)	Fig. 10.
Spline degree and quadrature points per cell (p, q)	(2, $p + 1$)	Fig. 17. Computational cost $\sim p^3$.
Custom knot vectors for mesh refinement	none	Fig. 2.
Precision {float32, float64, mixed}	mixed	Fig. 16.
Solver tolerance	10^{-8}	Fig. 16; 10^{-10} in float64, 3.5×10^{-4} in float32.
Island seed (m, n), cf. Sec. 7.2	off	Add: location r , width w , amplitude a , Eq. (72).
Descent method: Newton, L-BFGS(m)	Newton	Sec. 5.9. Default $m = 1$.
Newton (MINRES iterations, $\delta t_{\max}$)	(100, 1)	Table VIII. More iterations: past the floor sooner.
Newton preconditioner floor κ	3	Table VIII; $\kappa \rightarrow \infty$: the curl–curl preconditioner.
Smoothing (γ, ε)	(1, $0.02 h_r^2$)	Fig. 8, Table VI.
Velocity formulation {Leray, potential}	potential	Table VII. Newton always uses the potential v .
CFL cap on the line-search step	0.5	Eq. (62).
Helicity correction σ_*	off	Eq. (59), Table II; +3–5% per step.
Regularized line search C (Newton)	0.1	Eq. (69), $\varepsilon = C h_r^2$; off for L-BFGS.
Stop on the accepted step δt (Newton)	0.1	Sec. 6.4; the run ends at its floor. Off for L-BFGS.
L-BFGS restart (Powell)	0.2	Memory dropped when $ \langle F_k, F_{k-1} \rangle_\Omega > 0.2 \ F_k\ _\Omega^2$.
Maximum steps	100 (Newton)	The reference runs of Table X use 200 and 5000.
Steps per compiled chunk	20 (Newton)	Diagnostics and check-pointing happens between chunks.
Squared force residual stopping criterion	10^{-8}	Tested on the mean over a chunk.
Resistive reconnection every K steps	off	Fig. 11; helicity share per solve, 1% when on.
Restart: continue from previous checkpoint	none	–



(a)



(b)

FIG. 18: The step-regularization constant C : (a) accepted step δt , (b) squared force residual.

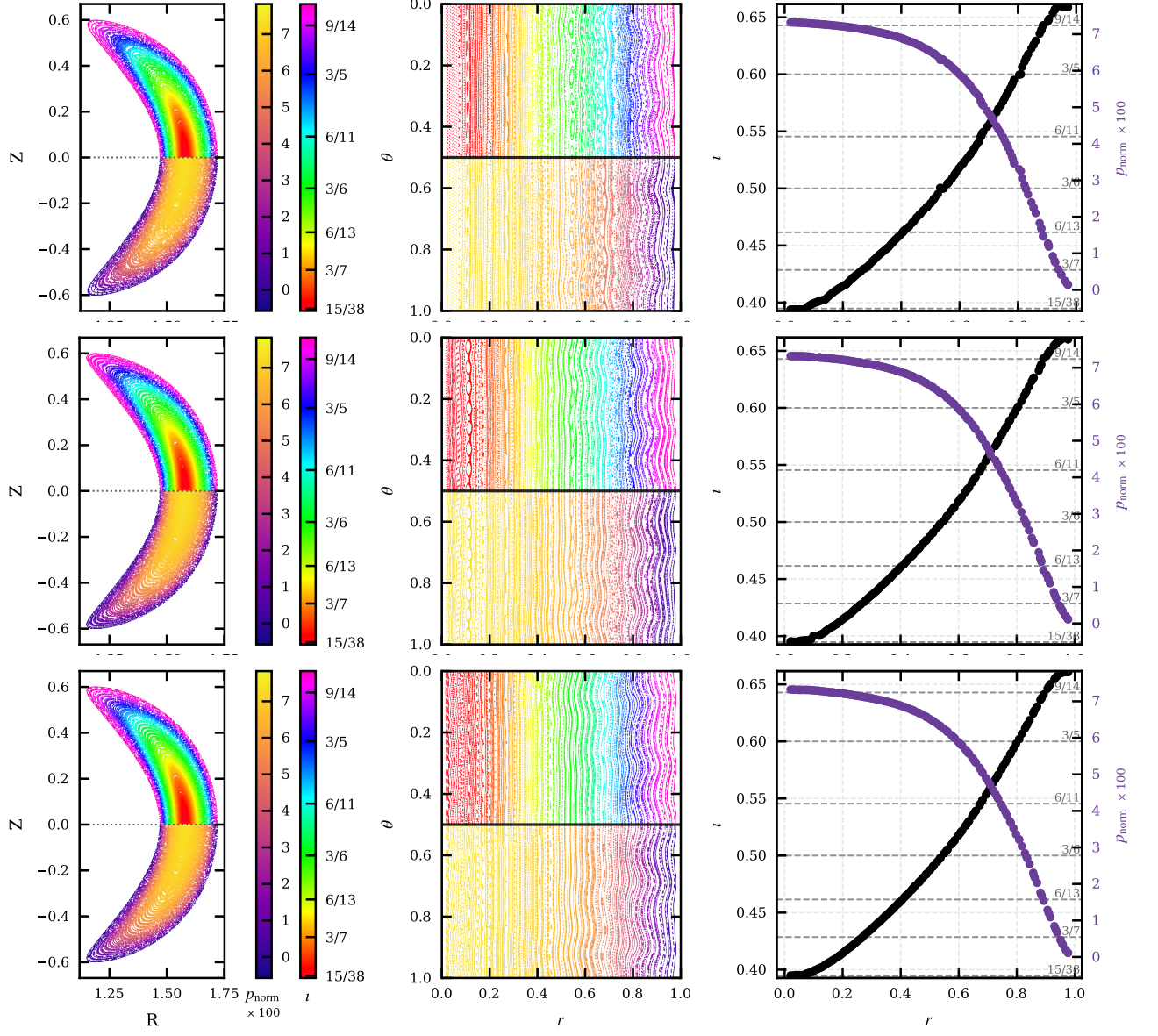


FIG. 19: Poincaré sections at $\zeta = 0$ for the Newton descent runs in Table III. Resolutions are $n = 16 \times 32 \times 32$ (top), $24 \times 48 \times 48$ (middle), and $32 \times 64 \times 64$ (bottom). The shape of grid-scale island features is, as expected, resolution-dependent; the $3/5$ island chain does not form at the highest resolution.

- [1] L. Lao, H. St. John, R. Stambaugh, A. Kellman, W. Pfeiffer, Reconstruction of current profile parameters and plasma shapes in tokamaks, *Nuclear Fusion* 25 (11) (1985) 1611. [doi:10.1088/0029-5515/25/11/007](https://doi.org/10.1088/0029-5515/25/11/007).
- [2] R. Sanchez, S. P. Hirshman, J. C. Whitson, A. S. Ware, COBRA: An Optimized Code for Fast Analysis of Ideal Ballooning Stability of Three-Dimensional Magnetic Equilibria, *Journal of Computational Physics* 161 (2) (2000) 576–588. [doi:10.1006/jcph.2000.6514](https://doi.org/10.1006/jcph.2000.6514).
- [3] W. I. van Rij, S. P. Hirshman, Variational bounds for transport coefficients in three-dimensional toroidal plasmas, *Physics of Fluids B: Plasma Physics* 1 (3) (1989) 563–569. [doi:10.1063/1.859116](https://doi.org/10.1063/1.859116).
- [4] N. R. Mandell, W. Dorland, I. Abel, R. Gaur, P. Kim, M. Martin, T. Qian, GX: a GPU-native gyrokinetic turbulence code for tokamak and stellarator design, *Journal of Plasma Physics* 90 (4) (2024) 905900402. [doi:10.1017/S0022377824000631](https://doi.org/10.1017/S0022377824000631).
- [5] N. Nikulsin, R. Ramasamy, M. Hoelzl, F. Hindenlang, E. Strumberger, K. Lackner, S. Günter, JOEK3D: An extension of the JOEK nonlinear MHD code to stellarators, *Physics of Plasmas* 29 (6) (2022) 063901. [doi:10.1063/5.0087104](https://doi.org/10.1063/5.0087104).
- [6] S. C. Jardin, N. Ferraro, J. Breslau, J. Chen, Multiple timescale calculations of sawteeth and other global macroscopic dynamics of tokamak plasmas, *Computational Science & Discovery* 5 (1) (2012) 014002. [doi:10.1088/1749-4699/5/1/014002](https://doi.org/10.1088/1749-4699/5/1/014002).
- [7] C. Sovinec, Modeling of Macroscopic Dynamics in Three-Dimensional Plasma Configurations (Final Technical Report), Tech. Rep. DOE-UW-Madison-SC0018642, Univ. of Wisconsin, Madison, WI (United States) (Jun. 2024). [doi:10.2172/2376548](https://doi.org/10.2172/2376548).
- [8] B. Gonçalves, J. Sousa, C. Varandas, Real-time control of fusion reactors, *Energy Conversion and Management* 51 (9) (2010) 1751–1757. [doi:10.1016/j.enconman.2010.02.004](https://doi.org/10.1016/j.enconman.2010.02.004).
- [9] H. Grad, Toroidal Containment of a Plasma, *The Physics of Fluids* 10 (1) (1967) 137–154. [doi:10.1063/1.1761965](https://doi.org/10.1063/1.1761965).
- [10] P. Constantin, T. D. Drivas, D. Ginsberg, Flexibility and Rigidity in Steady Fluid Motion, *Communications in Mathematical Physics* 385 (1) (2021) 521–563. [doi:10.1007/s00220-021-04048-4](https://doi.org/10.1007/s00220-021-04048-4).
- [11] C. Hansen, I. Stewart, D. Burgess, M. Pharr, S. Guizzo, F. Logak, A. Nelson, C. Paz-Soldan, TokaMaker: An open-source time-dependent Grad-Shafranov tool for the design and modeling of axisymmetric fusion devices, *Computer Physics Communications* 298 (2024) 109111. [doi:10.1016/j.cpc.2024.109111](https://doi.org/10.1016/j.cpc.2024.109111).
- [12] S. P. Hirshman, J. C. Whitson, Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria, *The Physics of Fluids* 26 (12) (1983) 3553–3568. [doi:10.1063/1.864116](https://doi.org/10.1063/1.864116).
- [13] J. Schilling, The numerics of vmec++ (2025). [arXiv:2502.04374](https://arxiv.org/abs/2502.04374).
- [14] R. Jorge, VMEX: a differentiable JAX implementation of VMEC with certified force-balance polishing (Sep. 2026).
- [15] F. Hindenlang, G. G. Plunk, O. Maj, Computing MHD equilibria of stellarators with a flexible coordinate frame, *Plasma Physics and Controlled Fusion* 67 (4) (2025) 045002. [doi:10.1088/1361-6587/adba11](https://doi.org/10.1088/1361-6587/adba11).
- [16] F. Hindenlang, O. Maj, R. Babin, R. Köberl, D. Muir, T. T. Ribeiro, M. Rampp, E. Sonnendrücker, Gvec: A flexible 3d mhd equilibrium solver, *Journal of Open Source Software* 11 (120) (2026) 9670. [doi:10.21105/joss.09670](https://doi.org/10.21105/joss.09670).
- [17] D. W. Dudt, E. Kolemen, DESC: A stellarator equilibrium solver, *Physics of Plasmas* 27 (10) (2020) 102513. [doi:10.1063/5.0020743](https://doi.org/10.1063/5.0020743).
- [18] J. Loizu, S. Hudson, A. Bhattacharjee, P. Helander, Magnetic islands and singular currents at rational surfaces in three-dimensional magnetohydrodynamic equilibria, *Physics of Plasmas* 22 (2) (2015) 022501. [doi:10.1063/1.4906888](https://doi.org/10.1063/1.4906888).
- [19] S. A. Lazerson, J. Loizu, S. Hirshman, S. R. Hudson, Verification of the ideal magnetohydrodynamic response at rational surfaces in the VMEC code, *Physics of Plasmas* 23 (1) (2016) 012507. [doi:10.1063/1.4939881](https://doi.org/10.1063/1.4939881).
- [20] A. Coelho, J. Loizu, P. Ricci, M. Giacomini, Global fluid simulation of plasma turbulence in a stellarator with an island divertor, *Nuclear Fusion* 62 (7) (2022) 074004, publisher: IOP Publishing. [doi:10.1088/1741-4326/ac6ad2](https://doi.org/10.1088/1741-4326/ac6ad2).
- [21] T. S. Pedersen, M. Otte, S. Lazerson, P. Helander, S. Bozhnikov, C. Biedermann, T. Klinger, R. C. Wolf, H. S. Bosch, The Wendelstein 7-X Team, Confirmation of the topology of the Wendelstein 7-X magnetic field to better than 1:100,000, *Nature Communications* 7 (1) (2016) 13493. [doi:10.1038/ncomms13493](https://doi.org/10.1038/ncomms13493).
- [22] Y. Feng, M. Jakubowski, R. König, M. Krychowiak, M. Otte, F. Reimold, D. Reiter, O. Schmitz, D. Zhang, C. Beidler, C. Biedermann, S. Bozhnikov, K. Brunner, A. Dinklage, P. Drewelow, F. Effenberg, M. Endler, G. Fuchert, Y. Gao, J. Geiger, K. Hammond, P. Helander, C. Killer, J. Knauer, T. Kremeyer, E. Pasch, L. Rudischhauser, G. Schlisio, T. Sunn Pedersen, U. Wenzel, V. Winters, W.-X. team, Understanding detachment of the W7-X island divertor, *Nuclear Fusion* 61 (8) (2021) 086012, publisher: IOP Publishing. [doi:10.1088/1741-4326/ac0772](https://doi.org/10.1088/1741-4326/ac0772).
- [23] A. H. Boozer, A. Punjabi, Simulation of stellarator divertors, *Physics of Plasmas* 25 (9) (2018) 092505. [doi:10.1063/1.5042666](https://doi.org/10.1063/1.5042666).
- [24] K. Garcia, A. Bader, H. Frerichs, G. Hartwell, J. Schmitt, N. Allen, O. Schmitz, Exploration of non-resonant divertor features on the Compact Toroidal Hybrid, *Nuclear Fusion* 63 (12) (2023) 126043, publisher: IOP Publishing. [doi:10.1088/1741-4326/ad0160](https://doi.org/10.1088/1741-4326/ad0160).
- [25] G. Janeschitz, K. Borrass, G. Federici, Y. Igitkhanov, A. Kukushkin, H. D. Pacher, G. W. Pacher, M. Sugihara, The ITER divertor concept, *Journal of Nuclear Materials* 220-222 (1995) 73–88. [doi:10.1016/0022-3115\(94\)00447-1](https://doi.org/10.1016/0022-3115(94)00447-1).
- [26] C. E. Kessel, J. P. Blanchard, A. Davis, L. El-Guebaly, L. M. Garrison, N. M. Ghoniem, P. W. Humrickhouse, Y. Huang, Y. Katoh, A. Khodak, E. P. Marriott, S. Malang, N. B. Morley, G. H. Neilson, J. Rapp, M. E. Rensink, T. D. Rognlien, A. F. Rowcliffe, S. Smo-

- lentsev, L. L. Snead, M. S. Tillack, P. Titus, L. M. Waganer, G. M. Wallace, S. J. Wukitch, A. Ying, K. Young, Y. Zhai, Overview of the fusion nuclear science facility, a credible break-in step on the path to fusion energy, *Fusion Engineering and Design* 135 (2018) 236–270. doi:10.1016/j.fusengdes.2017.05.081.
- [27] W. A. Cooper, J. P. Graves, A. Pochelon, O. Sauter, L. Villard, Tokamak magnetohydrodynamic equilibrium states with axisymmetric boundary and a 3d helical core, *Physical Review Letters* 105 (3) (2010) 035003. doi:10.1103/PhysRevLett.105.035003.
- [28] D. Panici, B. Jang, R. Conlin, D. W. Dudt, Y. G. Elmacioglu, E. Kolemen, *Deflation techniques for stellarator equilibrium and optimization*, preprint (2026). arXiv:2602.09957, doi:10.48550/arXiv.2602.09957. URL <https://arxiv.org/abs/2602.09957>
- [29] T. Thun, R. Conlin, D. G. Panici, D. Böckenhoff, Narrow operator models of stellarator equilibria in fourier-zernike basis, *Journal of Plasma Physics* 92 (2) (2026) E50. doi:10.1017/S0022377826101512.
- [30] S. R. Hudson, R. L. Dewar, G. Dennis, M. J. Hole, M. McGann, G. von Nessi, S. Lazerson, Computation of multi-region relaxed magnetohydrodynamic equilibria, *Physics of Plasmas* 19 (11) (2012) 112502. doi:10.1063/1.4765691.
- [31] Y. Narushima, S. Sakakibara, S. Ohdachi, Y. Suzuki, K. Y. Watanabe, S. Nishimura, S. Satake, B. Huang, M. Furukawa, Y. Takemura, K. Ida, M. Yoshinuma, I. Yamada, LHD Experiment Group, Experimental observation of response to resonant magnetic perturbation and its hysteresis in LHD, *Nuclear Fusion* 55 (7) (2015) 073004. doi:10.1088/0029-5515/55/7/073004.
- [32] J. Loizu, S. R. Hudson, C. Nührenberg, Verification of the SPEC code in stellarator geometries, *Physics of Plasmas* 23 (11) (2016) 112505. doi:10.1063/1.4967709.
- [33] E. Balkovic, J. Loizu, E. Lanti, C. Smiet, C. Lazzati, R. Ramasamy, A. Goodman, J. Geiger, J. P. Graves, Spectre: A robust solver for 3d equilibria with arbitrary topology, arXiv preprint arXiv:2607.27135 (2026).
- [34] M. Drevlak, D. Monticello, A. Reiman, PIES free boundary stellarator equilibria with improved initial conditions, *Nuclear Fusion* 45 (7) (2005) 731. doi:10.1088/0029-5515/45/7/022.
- [35] Y. Suzuki, N. Nakajima, K. Watanabe, Y. Nakamura, T. Hayashi, Development and application of HINT2 to helical system plasmas, *Nuclear Fusion* 46 (11) (2006) L19. doi:10.1088/0029-5515/46/11/L01.
- [36] Y. Suzuki, HINT modeling of three-dimensional tokamaks with resonant magnetic perturbation, *Plasma Physics and Controlled Fusion* 59 (5) (2017) 054008, publisher: IOP Publishing. doi:10.1088/1361-6587/aa5adc.
- [37] S. P. Hirshman, R. Sanchez, C. R. Cook, SIESTA: A scalable iterative equilibrium solver for toroidal applications, *Physics of Plasmas* 18 (6) (2011) 062504. doi:10.1063/1.3597155.
- [38] S. Candelaresi, D. Pontin, G. Hornig, Mimetic Methods for Lagrangian Relaxation of Magnetic Fields, *SIAM Journal on Scientific Computing* 36 (6) (2014) B952–B968. doi:10.1137/140967404.
- [39] C. B. Smiet, S. Candelaresi, D. Bouwmeester, Ideal relaxation of the Hopf fibration, *Physics of Plasmas* 24 (7) (2017) 072110. doi:10.1063/1.4990076.
- [40] M. He, P. E. Farrell, K. Hu, B. D. Andrews, Topology-preserving discretization for the magnetofrictional equations arising in the Parker conjecture, arXiv:2501.11654 [math] (Jan. 2025). doi:10.48550/arXiv.2501.11654.
- [41] S. Mao, R. Xi, A structure-preserving, helicity-conserving high-order ale finite element method for compressible ideal mhd systems, *SIAM Journal on Scientific Computing* 48 (3) (2026) C1–C25. doi:10.1137/25M1796230.
- [42] F. Bauer, O. Betancourt, P. Garabedian, *A Computational Method in Plasma Physics*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1978. doi:10.1007/978-3-642-85470-5.
- [43] F. Bauer, O. Betancourt, P. Garabedian, *Magnetohydrodynamic Equilibrium and Stability of Stellarators*, Springer New York, New York, NY, 1984. doi:10.1007/978-1-4612-5240-5.
- [44] O. Betancourt, BETAS, a spectral code for three-dimensional magnetohydrodynamic equilibrium and nonlinear stability calculations, *Communications on Pure and Applied Mathematics* 41 (5) (1988) 551–568. doi:10.1002/cpa.3160410504.
- [45] O. Betancourt, F. Herrnegger, P. Merkel, J. Nührenberg, R. Gruber, F. Troyon, Comparison of MHD stability results obtained with the BETA 3D and HERA 2D codes, *Journal of Computational Physics* 52 (1) (1983) 187–197. doi:10.1016/0021-9991(83)90023-2.
- [46] M. Taylor, A high performance spectral code for nonlinear MHD stability, Tech. Rep. DOE/ER/53223–193, 10113306, Courant Institute of Mathematical Sciences (Sep. 1992). doi:10.2172/10113306.
- [47] R. Chodura, A. Schlüter, A 3D code for MHD equilibrium and stability, *Journal of Computational Physics* 41 (1) (1981) 68–88. doi:10.1016/0021-9991(81)90080-2.
- [48] T. Hender, B. Carreras, L. Garcia, J. Rome, V. Lynch, The calculation of stellarator equilibria in vacuum flux surface coordinates, *Journal of Computational Physics* 60 (1) (1985) 76–96. doi:10.1016/0021-9991(85)90018-X.
- [49] S. Hirshman, O. Betancourt, Preconditioned descent algorithm for rapid calculations of magnetohydrodynamic equilibria, *Journal of Computational Physics* 96 (1) (1991) 99–109. doi:10.1016/0021-9991(91)90267-0.
- [50] A. Reiman, H. Greenside, Calculation of three-dimensional MHD equilibria with islands and stochastic regions, *Computer Physics Communications* 43 (1) (1986) 157–167. doi:10.1016/0010-4655(86)90059-7.
- [51] W. Park, D. A. Monticello, H. Strauss, J. Manickam, Three-dimensional stellarator equilibrium as an Ohmic steady state, *The Physics of Fluids* 29 (4) (1986) 1171–1175. doi:10.1063/1.865864.
- [52] K. Harafuji, T. Hayashi, T. Sato, Computational study of three-dimensional magnetohydrodynamic equilibria in toroidal helical systems, *Journal of Computational Physics* 81 (1) (1989) 169–192. doi:10.1016/0021-9991(89)90069-7.
- [53] S. Hudson, R. L. Dewar, M. J. Hole, G. Dennis, M. McGann, G. von Nessi, S. A. Lazerson, Computation of multi-region relaxed magnetohydrodynamic equilibria, *Phys. Plasmas* 19 (11) (2012). doi:10.1063/1.4765691.
- [54] D. Panici, R. Conlin, D. Dudt, K. Unalmis, E. Kolemen, The DESC stellarator code suite. Part 1. Quick and

- accurate equilibria computations, *Journal of Plasma Physics* 89 (3) (2023) 955890303. doi:[10.1017/S0022377823000272](https://doi.org/10.1017/S0022377823000272).
- [55] A. Reiman, H. Greenside, Computation of zero β three-dimensional equilibria with magnetic islands, *Journal of Computational Physics* 87 (2) (1990) 349–365. doi:[10.1016/0021-9991\(90\)90257-2](https://doi.org/10.1016/0021-9991(90)90257-2).
- [56] H. Oliver, A. Reiman, D. Monticello, Solving the 3D MHD equilibrium equations in toroidal geometry by Newton’s method, *Journal of Computational Physics* 211 (1) (2006) 99–128. doi:[10.1016/j.jcp.2005.05.007](https://doi.org/10.1016/j.jcp.2005.05.007).
- [57] C. Bressan, Metriplectic relaxation for calculating equilibria: theory and structure-preserving discretization, Ph.D. thesis, Technische Universität München (2023).
- [58] H. Moffatt, Some topological aspects of fluid dynamics, *Journal of Fluid Mechanics* 914 (2021) P1. doi:[10.1017/jfm.2020.230](https://doi.org/10.1017/jfm.2020.230).
- [59] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, *Acta Numerica* 15 (2006) 1–155. doi:[10.1017/S0962492906210018](https://doi.org/10.1017/S0962492906210018).
- [60] D. Boffi, F. Brezzi, M. Fortin, *Mixed Finite Element Methods and Applications*, Vol. 44 of Springer Series in Computational Mathematics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. doi:[10.1007/978-3-642-36519-5](https://doi.org/10.1007/978-3-642-36519-5).
- [61] L. D. Landau, E. M. Lifšic, L. D. Landau, *Statistical physics*, third edition, revised and enlarged Edition, no. volume 5 in *Course of theoretical physics*, Butterworth-Heinemann, Amsterdam Boston Heidelberg, 1996.
- [62] K. Hu, Y. Ma, J. Xu, Stable finite element methods preserving $\text{div } \mathbf{b} = 0$ exactly for MHD models, *Numerische Mathematik* 135 (2) (2017) 371–396. doi:[10.1007/s00211-016-0803-4](https://doi.org/10.1007/s00211-016-0803-4).
- [63] E. S. Gawlik, F. Gay-Balmaz, A structure-preserving finite element method for compressible ideal and resistive magnetohydrodynamics, *Journal of Plasma Physics* 87 (5) (2021) 835870501. doi:[10.1017/S0022377821000842](https://doi.org/10.1017/S0022377821000842).
- [64] K. Hu, Y.-J. Lee, J. Xu, Helicity-conservative finite element discretization for incompressible MHD systems, *Journal of Computational Physics* 436 (2021) 110284. doi:[10.1016/j.jcp.2021.110284](https://doi.org/10.1016/j.jcp.2021.110284).
- [65] E. S. Gawlik, F. Gay-Balmaz, A finite element method for MHD that preserves energy, cross-helicity, magnetic helicity, incompressibility, and $\text{div } \mathbf{B} = 0$, *Journal of Computational Physics* 450 (2022) 110847. doi:[10.1016/j.jcp.2021.110847](https://doi.org/10.1016/j.jcp.2021.110847).
- [66] F. Holderied, S. Possanner, Magneto-hydrodynamic eigenvalue solver for axisymmetric equilibria based on smooth polar splines, *Journal of Computational Physics* 464 (2022) 111329. doi:[10.1016/j.jcp.2022.111329](https://doi.org/10.1016/j.jcp.2022.111329).
- [67] V. Carlier, M. Campos-Pinto, Variational discretizations of ideal magnetohydrodynamics in smooth regime using structure-preserving finite elements, *Journal of Computational Physics* 523 (2025) 113647. doi:[10.1016/j.jcp.2024.113647](https://doi.org/10.1016/j.jcp.2024.113647).
- [68] R. Frostig, M. Johnson, C. Leary, *Compiling machine learning programs via high-level tracing*, 2018.
- [69] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).
- [70] R. Babin, F. Hindenlang, O. Maj, R. Köberl, Construction of an invertible mapping to boundary conforming coordinates for arbitrarily shaped toroidal domains, *Plasma Physics and Controlled Fusion* 67 (3) (2025) 035005.
- [71] R. Abraham, J. E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Vol. 75 of Applied Mathematical Sciences, Springer New York, New York, NY, 1988. doi:[10.1007/978-1-4612-1029-0](https://doi.org/10.1007/978-1-4612-1029-0).
- [72] D. Arnold, R. Falk, R. Winther, Finite element exterior calculus: from Hodge theory to numerical stability, *Bulletin of the American Mathematical Society* 47 (2) (2010) 281–354. doi:[10.1090/S0273-0979-10-01278-4](https://doi.org/10.1090/S0273-0979-10-01278-4).
- [73] A. Buffa, G. Sangalli, R. Vázquez, Isogeometric analysis in electromagnetics: B-splines approximation, *Computer Methods in Applied Mechanics and Engineering* 199 (17–20) (2010) 1143–1152. doi:[10.1016/j.cma.2009.12.002](https://doi.org/10.1016/j.cma.2009.12.002).
- [74] M. Campos-Pinto, Y. Güçlü, Broken-FEEC discretizations and Hodge Laplace problems, arXiv:2109.02553 [math] (Feb. 2024). doi:[10.48550/arXiv.2109.02553](https://doi.org/10.48550/arXiv.2109.02553).
- [75] F. Holderied, STRUPHY: a structure-preserving hybrid MHD-kinetic code for the interaction of energetic particles with Alfvén waves in magnetized plasmas, PhD Thesis, Technische Universität München (2022).
- [76] D. Toshniwal, H. Speleers, R. R. Hiemstra, T. J. Hughes, Multi-degree smooth polar splines: A framework for geometric modeling and isogeometric analysis, *Computer Methods in Applied Mechanics and Engineering* 316 (2017) 1005–1061. doi:[10.1016/j.cma.2016.11.009](https://doi.org/10.1016/j.cma.2016.11.009).
- [77] D. Toshniwal, T. J. Hughes, Isogeometric discrete differential forms: Non-uniform degrees, Bézier extraction, polar splines and flows on surfaces, *Computer Methods in Applied Mechanics and Engineering* 376 (2021) 113576. doi:[10.1016/j.cma.2020.113576](https://doi.org/10.1016/j.cma.2020.113576).
- [78] F. Patrizi, D. Toshniwal, Isogeometric de Rham complex discretization in solid toroidal domains, arXiv:2106.10470 [math] (Jan. 2025). doi:[10.48550/arXiv.2106.10470](https://doi.org/10.48550/arXiv.2106.10470).
- [79] M. Landreman, E. Paul, Magnetic fields with precise quasisymmetry for plasma confinement, *Phys. Rev. Lett.* 128 (2022) 035001. doi:[10.1103/PhysRevLett.128.035001](https://doi.org/10.1103/PhysRevLett.128.035001).
- [80] H. Grad, H. Rubin, Hydromagnetic equilibria and force-free fields, in: *Proceedings of the second United Nations international conference on the peaceful uses of atomic energy*. V. 31. Theoretical and experimental aspects of controlled nuclear fusion, United Nations, Geneva, Switzerland, 1958, pp. 190 – 197.
- [81] I. B. Bernstein, E. A. Frieman, M. D. Kruskal, R. M. Kulsrud, An energy principle for hydromagnetic stability problems, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 244 (1236) (1958) 17–40. doi:[10.1098/rspa.1958.0023](https://doi.org/10.1098/rspa.1958.0023).
- [82] S. Lundquist, On the Stability of Magneto-Hydrostatic Fields, *Physical Review* 83 (2) (1951) 307–311. doi:[10.1103/PhysRev.83.307](https://doi.org/10.1103/PhysRev.83.307).
- [83] K. Hain, R. Lust, A. Schlüter, Zur Stabilität eines Plasmas, *Zeitschrift für Naturforschung A* 12 (10) (1957) 833–841. doi:[10.1515/zna-1957-1011](https://doi.org/10.1515/zna-1957-1011).

- [84] P. Merkel, A. Schlüter, Das holonome Energieprinzip der Magnetohydrodynamik, Bayerische Akademie der Wissenschaften 1976, 7 (1977) 117–123, place: München Publisher: Verl.d.Bayer.Akad.d.Wiss.
- [85] L. Woltjer, A theorem on force-free magnetic fields, Proceedings of the National Academy of Sciences 44 (6) (1958) 489–491. doi:10.1073/pnas.44.6.489.
- [86] H. K. Moffatt, The degree of knottedness of tangled vortex lines, Journal of Fluid Mechanics 35 (1) (1969) 117–129. doi:10.1017/S0022112069000991.
- [87] H. K. Moffatt, Magnetostatic equilibria and analogous Euler flows of arbitrarily complex topology. Part 1. Fundamentals, Journal of Fluid Mechanics 159 (-1) (1985) 359. doi:10.1017/S0022112085003251.
- [88] H. K. Moffatt, Magnetostatic equilibria and analogous Euler flows of arbitrarily complex topology. Part 2. Stability considerations, Journal of Fluid Mechanics 166 (-1) (1986) 359. doi:10.1017/S0022112086000198.
- [89] V. I. Arnold, B. A. Khesin, Topological Methods in Hydrodynamics, Vol. 125 of Applied Mathematical Sciences, Springer International Publishing, Cham, 2021. doi:10.1007/978-3-030-74278-2.
- [90] D. I. Pontin, G. Hornig, The Parker problem: existence of smooth force-free fields and coronal heating, Living Reviews in Solar Physics 17 (1) (2020) 5. doi:10.1007/s41116-020-00026-5.
- [91] D. I. Pontin, G. Hornig, A. L. Wilmot-Smith, I. J. D. Craig, Lagrangian relaxation schemes for calculating force-free magnetic fields, and their limitations, The Astrophysical Journal 700 (2) (2009) 1449–1455. doi:10.1088/0004-637X/700/2/1449.
- [92] V. I. Arnold, The asymptotic Hopf invariant and its applications, in: A. B. Givental, B. A. Khesin, A. N. Varchenko, V. A. Vassiliev, O. Y. Viro (Eds.), Vladimir I. Arnold - Collected Works, Springer Berlin Heidelberg, Berlin, Heidelberg, 1974, pp. 357–375. doi:10.1007/978-3-642-31031-7_32.
- [93] A. Ern, J. Guzmán, P. Potu, M. Vohralík, Discrete Poincaré inequalities: a review on proofs, equivalent formulations, and behavior of constants, version Number: 1 (2024). doi:10.48550/ARXIV.2412.11796.
- [94] D. MacTaggart, A. Valli, Magnetic helicity in multiply connected domains, Journal of Plasma Physics 85 (5) (2019) 775850501. doi:10.1017/S0022377819000576.
- [95] A. Enciso, D. Peralta-Salas, Obstructions to Topological Relaxation for Generic Magnetic Fields, Archive for Rational Mechanics and Analysis 249 (1) (2025) 6. doi:10.1007/s00205-024-02078-5.
- [96] A. Enciso, D. Peralta-Salas, F. T. De Lizaaur, Helicity is the only integral invariant of volume-preserving transformations, Proceedings of the National Academy of Sciences 113 (8) (2016) 2035–2040. doi:10.1073/pnas.1516213113.
- [97] R. Beekie, S. Friedlander, V. Vicol, On Moffatt’s Magnetic Relaxation Equations, Communications in Mathematical Physics 390 (3) (2022) 1311–1339. doi:10.1007/s00220-021-04289-3.
- [98] M. J. D. Powell, Restart procedures for the conjugate gradient method, Mathematical Programming 12 (1) (1977) 241–254. doi:10.1007/BF01593790.
- [99] P. Kidger, On Neural Differential Equations, arXiv:2202.02435 [cs] (Feb. 2022). doi:10.48550/arXiv.2202.02435.
- [100] E. N. Parker, Spontaneous Current Sheets In Magnetic Fields: With Applications to Stellar X-rays, Oxford University Press New York, NY, 1994. doi:10.1093/oso/9780195073713.001.0001.
- [101] H. Bae, H. Kwon, J. Shin, Local and global solutions to Stokes-Magneto equations with fractional dissipations, Partial Differential Equations and Applications 6 (4) (2025) 32. doi:10.1007/s42985-025-00341-2.
- [102] K. Cieliebak, E. Volkov, A note on the stationary Euler equations of hydrodynamics, Ergodic Theory and Dynamical Systems 37 (2) (2017) 454–480. doi:10.1017/etds.2015.50.
- [103] J. Nocedal, S. J. Wright, Numerical optimization, 2nd Edition, Springer series in operations research, Springer, New York, 2006.
- [104] S. Hirshman, K. Perumalla, V. Lynch, R. Sanchez, BCYCLIC: A parallel block tridiagonal matrix cyclic solver, Journal of Computational Physics 229 (18) (2010) 6392–6404. doi:10.1016/j.jcp.2010.04.049.
- [105] M. C. Zarnstorff, L. A. Berry, A. i. Brooks, J. Ehrhardt, M. Galloway, D. A. Gates, S. P. Hirshman, S. Hudson, L. P. Ku, E. Lazarus, H. i. Mynick, B. i. Nelson, C. Nuehrenberg, A. Reiman, G. Rewoldt, J. Roach, J. Schmidt, D. A. Spong, D. J. Strickler, A. Sykes, R. Taylor, P. Valanju, P. Ward, H. Weitzner, D. Williamson, S. Zweben, et al., Physics of the compact advanced stellarator ncsx, Plasma Physics and Controlled Fusion 43 (12A) (2001) A237. doi:10.1088/0741-3325/43/12A/318.
- [106] A. R. Yeates, On the limitations of magneto-frictional relaxation, Geophysical & Astrophysical Fluid Dynamics 116 (4) (2022) 305–320. doi:10.1080/03091929.2021.2021197.
- [107] P. Helander, Theory of plasma confinement in non-axisymmetric magnetic fields, Reports on Progress in Physics 77 (8) (2014) 087001. doi:10.1088/0034-4885/77/8/087001.
- [108] J. Liu, M. Sato, A. Ishizawa, Pressure gradient-driven plasmoid formation and ejection from magnetic island, Physics of Plasmas 33 (5) (2026) 052506. doi:10.1063/5.0314607.
URL <https://pubs.aip.org/pop/article/33/5/052506/3391242/Pressure-gradient-driven-plasmoid-formation-and>
- [109] J. R. Cary, J. D. Hanson, Simple method for calculating island widths, Physics of Fluids B: Plasma Physics 3 (4) (1991) 1006–1014. doi:10.1063/1.859829.
- [110] M. Landreman, B. Medasani, F. Wechsung, A. Giuliani, R. Jorge, C. Zhu, SIMSOPT: A flexible framework for stellarator optimization, Journal of Open Source Software 6 (65) (2021) 3525. doi:10.21105/joss.03525.
- [111] A. Kaptanoglu, T. Blickhan, Computational boundary specification in 3D fixed-boundary magnetohydrodynamic equilibrium modeling, arXiv:2605.01652 [physics.plasm-ph] (May 2026). doi:10.48550/arXiv.2605.01652.