J. Plasma Phys. (2026), *vol*. 92, E50 © The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.
doi:10.1017/S0022377826101512

Narrow operator models of stellarator equilibria in Fourier Zernike basis

Timo Thun[1], Rory Conlin[2], Dario Giovanni Panici[3] and Daniel Böckenhoff[1]

[1]Max-Planck-Institute for Plasma Physics, 17491 Greifswald, Germany
[2]Institute for Research in Electronics & Applied Physics, University of Maryland, College Park, MD 20742, USA
[3]Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA
Corresponding author: Timo Thun, timo.thun@ipp.mpg.de

(Received 30 October 2025; revision received 4 February 2026; accepted 26 February 2026)

Numerical computation of the ideal magnetohydrodynamic (MHD) equilibrium magnetic field is at the base of stellarator optimisation and provides the starting point for solving more sophisticated partial differential equations like transport or turbulence models. Conventional approaches solve for a single stationary point of the ideal MHD equations, which is fully defined by three invariants and the numerical scheme employed by the solver. We present the first numerical approach that can solve for a continuous distribution of equilibria with fixed boundary and rotational transform, varying only the pressure invariant. This approach minimises the force residual by optimising parameters of multilayer perceptrons that map from a scalar pressure multiplier to the Fourier Zernike basis as implemented in the modern stellarator equilibrium solver `DESC`.

Key words: fusion plasma, plasma confinement, plasma simulation

1. Introduction

Stellarators are inherently steady-state plasma confinement devices, which is among the key reasons behind their renaissance as promising candidates for fusion power plants. Ideal magnetohydrodynamic (MHD) equilibria are a central part of optimising the complex, three-dimensional plasma shapes which is a necessary condition for steady-state operation of such devices. The equilibrium magnetic field is required not only in optimisation, but also plays a role in future real-time control algorithms and simulation frameworks (Schissel, Nazikian & Gibbs 2025).

Solving the three-dimensional MHD equations requires numerical approaches, because no analytical solution throughout the full volume of strongly shaped toroidal ideal MHD equilibria with nested magnetic topology exists yet (Bruno & Laurence 1996). Recent work advanced analytical models for Fourier components of the equilibrium magnetic field in a subset of reactor-relevant magnetic fields and analytical expansions close to the magnetic axis are used extensively in research (Nikulsin *et al.*

2024; Sengupta *et al.* 2024). These analytical solutions and the following numerical solvers assume a nested magnetic topology, or integrability throughout the volume, and computation of chaotic regions or magnetic islands takes considerably more effort (Hudson *et al.* 2012).

Accuracy of numerical partial differential equation (PDE) solutions is inherently connected to the representation which defines gradients, and commonly used ideal MHD equilibrium solvers with nested magnetic field topology can be differentiated accordingly: a widely used finite-difference solver employed in the design of currently operating stellarator devices is VMEC (Hirshman & Whitson 1983), another pseudo-spectral solver is DESC (Dudt & Kolemen 2020) and a third example is GVEC (Hindenlang, Plunk & Maj 2025), that abstracts the notion of basis functions, which enabled computation of plasmas with a figure-8 shape (Plunk *et al.* 2025).

Active control of stellarator plasmas is much less required than active control of tokamaks, which are prone to disruptive events that can damage the machine because confinement in tokamaks is dependent on large toroidal plasma currents (Schissel *et al.* 2025). Modern control policies enabled accurate tracking of the location, current and shape of axisymmetric plasmas realisable within the Tokamak à Configuration Variable device (Degrave *et al.* 2022). This shows that digital twins and real-time control can also be helpful tools in future stellarator devices, especially regarding control of transport and turbulence and possibly accessing novel plasma states by careful search in a device's configuration space. Many transport and turbulence codes, and accordingly their surrogate models, rely on either a coordinate system in which the magnetic field lines are straight (Mandell *et al.* 2024) or the equilibrium magnetic field (Landreman *et al.* 2014). Computation of straight-field line coordinate systems requires the equilibrium magnetic field and models with very rapid inference of the equilibrium field will be helpful in sophisticated stellarator control strategies. Furthermore, real-time interpretation of diagnostic data is facilitated if the inference time of magnetic equilibria is reduced as much as possible (Merlo *et al.* 2023*b*).

Artificial neural networks (NNs) enable quick inference by transferring the bulk of the computation to training the NN, which is then composed of simple nonlinearities and parallelisable matrix multiplications.

We introduce simple NN-based models with low residuals over parametrised spaces of equilibria within fixed boundaries and with fixed rotational transform. These models are parametrised by a unit interval scalar multiplier of the pressure coefficients and achieve volume-average force residuals very close to DESC's force residual over the whole interval, from near-vacuum conditions to full pressure.

1.1. *Motivation*

This work takes the next step on the path to precise operator models of a subset of fusion relevant ideal MHD equilibria by integrating NNs into DESC. Previous work presented advantages of small multilayer perceptrons (MLPs) which output Fourier decomposed equilibrium magnetic fields (Thun *et al.* 2026) and we test the same approach in DESC's Fourier Zernike basis.

DESC can solve equilibria prescribed with current or rotational transform, includes many features such as omnigenous field optimisation (Dudt *et al.* 2024) and Mercier stability (Panici *et al.* 2023) and has implemented interfaces to gyrokinetic turbulence codes (Kim *et al.* 2024) – all this is immediately available to evaluate operator models parametrised by NNs in future work. The implementation of DESC allows for easy integration of NNs and DESC's optimisation subspace, in which linear constraints

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

are satisfied by construction, and reduces the dimensionality of the minimisation problem. We train narrow operator models in DESC's optimisation subspace ($\boldsymbol{y}$ in (2.17)) using only the force residual evaluated on typical concentric grids at discrete multipliers of the pressure coefficients.

Operator models that parametrise equilibria with low normalised force error are the scaffolding for digital twins, real-time control algorithms and rapid interpretation of diagnostic data. Furthermore, precise equilibrium operator models of the configuration space of a machine are necessary to create sophisticated real-time capable simulation frameworks, for example, including transport and turbulence operator models which use deviations from the equilibrium magnetic field (Schissel *et al.* 2025).

Another application for the presented operator models is in optimisation: parametrised operator models ensure low sensitivity of stellarator optimisation targets towards uncertainty in the prescribed pressure profile. The presented models are a first step towards parametrisation of flexible configurations that preserve low optimisation metrics throughout a device's operational limits and map out the landscape of said metrics.

In terms of flight simulators or digital twins, control is likely to benefit from such models: once trained, they can better inform control algorithms by rapidly propagating aleatoric uncertainties through the magnetic field topology to the control algorithm. Models with rapid inference of plasma evolution are expected to play an important role in sophisticated control strategies of advanced fusion experiments (Schissel *et al.* 2025).

Prior research determined that the minimum size, or complexity, of simple MLPs which parametrise a single ideal MHD equilibrium in Fourier space necessitate two hidden layers, each with a nonlinear activation function (Thun *et al.* 2026). In the following, we will use this result to answer whether similar two-layer MLPs are also capable of reproducing narrow regions of the ideal MHD PDE operator defined by a scaling factor of the pressure, while keeping the plasma boundary and the rotational transform profile constant. Because we only vary the pressure in this work, we will call these models narrow operator models in the sense that they parametrise a narrow subspace of the ideal MHD PDE.

2. Three-dimensional ideal magnetohydrostatic problem

Stationary points of the ideal MHD PDE with isotropic pressure p describe plasmas as fluids with one species only in the limit of long wavelengths, low frequencies and no electric resistivity (Freidberg 2014)

$$\boldsymbol{J} \times \boldsymbol{B} = \nabla \boldsymbol{p}, \tag{2.1}$$

$$\mu_0 \boldsymbol{J} = \nabla \times \boldsymbol{B}, \tag{2.2}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{2.3}$$

with the magnetic field $\boldsymbol{B}$, currents $\boldsymbol{J}$, pressure $\boldsymbol{p}$ and vacuum permeability μ_0. Inserting Ampère's law (2.2) into the momentum (2.1) removes currents from this system of equations, yielding the residual force $\boldsymbol{F}$

$$\begin{aligned}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} &= \mu_0 \nabla \boldsymbol{p} \\ \Leftrightarrow \boldsymbol{F} &= (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} - \mu_0 \nabla \boldsymbol{p} \\ \Leftrightarrow \boldsymbol{F} &= F_\rho \nabla \rho + F_\theta \nabla \theta + F_\theta \nabla \zeta. \end{aligned} \tag{2.4}$$

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

Equilibrium states are defined by the magnetic field, which has toroidal, or ring-shaped, form for magnetically confined plasmas in tokamaks and stellarators.

Under the assumption of a nested, or integrable, structure of this magnetic field, the component in radial direction ρ of the magnetic field $B^\rho = \boldsymbol{B} \cdot \nabla\rho$ is 0, and we can write the magnetic field as

$$\boldsymbol{B} = \nabla\zeta \times \nabla\chi + \nabla\psi \times \nabla\theta^\star = B^\theta \boldsymbol{e}_\theta + B^\zeta \boldsymbol{e}_\zeta, \tag{2.5}$$

with toroidal magnetic flux $2\pi\psi$ and poloidal magnetic flux $2\pi\chi$. The radial magnetic coordinate in this work is the same as DESC's $\rho = \sqrt{s} = \sqrt{\psi/\psi_\mathrm{b}} \in \mathbb{R} \cap [0, 1)$, $\psi_\mathrm{b} := \psi(\rho = 1)$, $\theta^\star \in [0, 2\pi]$ is a poloidal angle which straightens magnetic field lines and the magnetic toroidal angle $\zeta \in [0, 2\pi]$ is equal to the cylindrical toroidal angle (Helander 2014) where s is the fraction of ψ and ψ_b.

Nestedness of the magnetic topology implies constant toroidal and poloidal magnetic flux on isobaric flux surfaces. Assuming nested flux surfaces, the ideal MHD equilibrium equations can be solved in an inverse manner, i.e. they are fully defined by the map from independent to cylindrical laboratory coordinates $[\rho, \theta, \zeta]^\mathsf{T} \to [R, \lambda, Z]^\mathsf{T}$ and three invariants (Hirshman & Whitson 1983).

Under Gauss's law for magnetism, the contravariant $\boldsymbol{B}$-field reduces to

$$\boldsymbol{B} = \frac{\partial_\rho\psi}{\sqrt{g}} \left((\iota(\rho) - \partial_\zeta\lambda)\boldsymbol{e}_\theta + (1 + \partial_\theta\lambda)\boldsymbol{e}_\zeta \right). \tag{2.6}$$

Because the poloidal angle $\theta \in [0, 2\pi]$ is arbitrary, as long as it is periodic and the Jacobian of the inverse map stays finite and does not switch sign, λ is introduced as a periodic renormalisation function which straightens magnetic field lines: $\theta^\star = \theta + \lambda(\rho, \theta, \zeta)$.

Hirshman & Whitson (1983) defined the covariant basis vectors of the inverse map as

$$\boldsymbol{e}_\rho = \begin{bmatrix} \partial_\rho R \\ 0 \\ \partial_\rho Z \end{bmatrix} \quad \boldsymbol{e}_\theta = \begin{bmatrix} \partial_\theta R \\ 0 \\ \partial_\theta Z \end{bmatrix} \quad \boldsymbol{e}_\zeta = \begin{bmatrix} \partial_\zeta R \\ R \\ \partial_\zeta Z \end{bmatrix}, \tag{2.7}$$

in conjunction with the Jacobian

$$\sqrt{g} = \boldsymbol{e}_s \cdot \boldsymbol{e}_\theta \times \boldsymbol{e}_\zeta = (\boldsymbol{e}^s \cdot \boldsymbol{e}^\theta \times \boldsymbol{e}^\zeta)^{-1}. \tag{2.8}$$

And the contravariant basis vectors are $\boldsymbol{e}^i = \nabla i = \boldsymbol{e}_j \times \boldsymbol{e}_k / \sqrt{g}$ with (i, j, k) a cyclic permutation in $\{\rho, \theta, \zeta\}$.

The last components required for solving (2.1) to (2.3) are three invariants: any equilibrium needs some prescribed (isotropic) pressure profile $p(\rho)$, a rotational transform profile $\iota(\rho)$ or some current profile $c(\rho)$ (Hirshman & Meier 1985), and optionally, the plasma boundary can be enforced via $R(\rho = 1) = R_\mathrm{b}$ and $Z(\rho = 1) = Z_\mathrm{b}$, in which case the equilibrium is called fixed boundary (Kruskal & Kulsrud 1958). Additionally, the total toroidal flux through the torus ψ_b can be either specified or set to 1 Wb with later rescaling of other values.

Finally, the output is an equilibrium magnetic field $\boldsymbol{B}$, determined by a balance between plasma pressure gradient and Lorentz force under the assumption of nested magnetic surfaces inside a fixed plasma boundary. The relative strength of both forces is commonly described by a ratio, the plasma beta

$$\langle \boldsymbol{\beta} \rangle_\mathrm{vol} = \frac{\langle p \rangle_\mathrm{vol} \, 2\mu_0}{\langle \boldsymbol{B}^2 \rangle_\mathrm{vol}}, \tag{2.9}$$

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

with brackets denoting the volume average of some quantity $(\cdot)$

$$\langle(\cdot)\rangle_{\text{vol}} = \frac{1}{V}\int_\rho\int_\theta\int_\zeta (\cdot)\sqrt{g}\,d\rho\,d\theta\,d\zeta. \tag{2.10}$$

The plasma volume is computed by integrating the Jacobian (2.8) over the triplet (ρ, θ, ζ).

Minimisation of (2.4) is simplified by inserting the contravariant $\boldsymbol{B}$-field (2.6), revealing two independent directions of the covariant force: on the one hand $F_\rho = \sqrt{g}(J^\zeta B^\theta - J^\theta B^\zeta) + \partial_\rho p(\rho)$ in the radial direction $\nabla\rho$ and, on the other hand, $F_\beta = \sqrt{g}J^\rho$ in the helical direction $\boldsymbol{\beta}_{\text{DESC}} = B^\zeta\nabla\theta - B^\theta\nabla\zeta$ (Panici *et al.* 2023)

$$\boldsymbol{F} = \left(\sqrt{g}(J^\zeta B^\theta - J^\theta B^\zeta) + \partial_\rho p(\rho)\right)\nabla\rho + \sqrt{g}J^\rho(B^\zeta\nabla\theta - B^\theta\nabla\zeta). \tag{2.11}$$

The currents in this expression are given by $\mu_0 J^i = \nabla\cdot(\boldsymbol{B}\times\nabla i)$ with i a cyclic permutation of $\{\rho, \theta, \zeta\}$.

Numerical solutions to (2.1) to (2.3) with different characteristics can be compared using the normalised force

$$\boldsymbol{F}_{\text{norm}} = \frac{|(\nabla\times\boldsymbol{B})\times\boldsymbol{B} - \mu_0\nabla p(\rho)|}{\langle|\nabla|B|^2/(2\mu_0)|\rangle_{\text{vol}}}. \tag{2.12}$$

The denominator in this equation is the volume average of the magnetic pressure gradient with $\nabla|B|^2 = 2(|B|\nabla|B|)$. We will denote the scalar volume average of $\boldsymbol{F}_{\text{norm}}$ as $\langle\boldsymbol{F}\rangle_{\text{vol,norm}}$ in the following (see e.g. figure 2).

Due to physics and engineering reasons, stellarators are commonly split into N_{FP} self-similar parts, each occupying $2\pi/\text{N}_{\text{FP}}$ of the full toroidal angle $\zeta\in[0, 2\pi)$.

2.1. *DESC solver*

DESC is a pseudo spectral code that not only efficiently solves (2.1) to (2.3), but also includes other important stellarator minimisation problems. It can solve free-boundary equilibria where, instead of fixing the plasma's boundary, a current is specified some distance from the plasma boundary which is then split into discrete coils (Conlin *et al.* 2024). DESC implements stellarator optimisation targets, such as the direct minimisation of omnigenous field errors, Mercier- or the infinite-N ideal ballooning stability, and is coupled to other codes like the turbulence code GX or the gyrokinetic solver GS2 (Gaur *et al.* 2024; Kim *et al.* 2024). The previously mentioned analytic near-axis expansion can be used as a starting point for optimisation in DESC, which then computes solutions valid throughout the volume. A comparison between DESC, VMEC and a code which can resolve magnetic islands, SPEC, agreed well on the magnetic axis position of a Heliotron-like equilibrium (Hudson *et al.* 2025).

DESC minimises (2.11) by weighting F_ρ and F_β with the occupied volume of each collocation point

$$f_\rho = F_\rho||\nabla\rho||_2\sqrt{g}\Delta\rho\Delta\theta\Delta\zeta, \tag{2.13}$$

$$f_\beta = F_\beta||\boldsymbol{\beta}_{\text{DESC}}||_2\sqrt{g}\Delta\rho\Delta\theta\Delta\zeta. \tag{2.14}$$

A nonlinear system of equations is then solved by least-squares optimisers

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} f_{\rho,j}(\boldsymbol{x}) \\ f_{\beta,k}(\boldsymbol{x}) \end{bmatrix} = \boldsymbol{0} \tag{2.15}$$

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

Parameter	DIII-D	Heliotron	W7-X	Quasi-helical
NN nodes per layer	1, 8, 16, 182	1, 64, 256, 351	1, 32, 64, 265	1, 32, 128, 661
M	12	9	6	8
N	0	3	6	8
$x_{\text{scale},j}$	1	$\frac{1}{l_{ZP,j}+\lvert m_j\rvert+\lvert n_j\rvert+1}$	$\frac{1}{l_{ZP,j}+\lvert m_j\rvert+\lvert n_j\rvert+1}$	$\frac{1}{l_{ZP,j}+\lvert m_j\rvert+\lvert n_j\rvert+1}$
Current prescribed	Yes	No	No	Yes
$\langle\boldsymbol{\beta}\rangle_{\text{vol}}$ at $\eta_{\text{p}}=1$	0.645 %	10.18 %	4.43 %	5.08 %
Pressure function type	$\sum_l \rho^{2l} p_l$	$\sum_l \rho^{2l} p_l$	$\sum_l \rho^{2l} p_l$	$\sum_l p_l B_l^3(\rho^2)$
$\max_{\eta_{\text{p,train}}}(\epsilon_{\text{X}}(\eta_{\text{p}}))$ [m]	$\epsilon_{\text{X}}(1.0) = 0.0062$	$\epsilon_{\text{X}}(1.0) = 0.0079$	$\epsilon_{\text{X}}(0.1) = 0.0134$	$\epsilon_{\text{X}}(0.4) = 0.06$

TABLE 1. Summary of model and equilibrium parameters. The last number of nodes per layer is the size of $\boldsymbol{y}$ and B_l^3 is a cubic B-spline basis.

for $\boldsymbol{x} = [\boldsymbol{R}_{l_{ZP}mn}(\rho,\theta,\zeta), \boldsymbol{\lambda}_{l_{ZP}mn}(\rho,\theta,\zeta), \boldsymbol{Z}_{l_{ZP}mn}(\rho,\theta,\zeta)]^{\mathsf{T}}$, with $j = 0,\ldots,J$ and $k = 0,\ldots,K$ indexing collocation points on possibly two different grids. $\boldsymbol{x}$ are the coefficients for each dependent coordinate decomposed into a Fourier Zernike basis with Zernike polynomial of finite radial order $l_{ZP} = 0, 1, \ldots, L_{ZP}$ and finite poloidal mode numbers $m = -M, -M+1, \ldots, 0, \ldots, M-1, M$, and a toroidal decomposition into Fourier modes with finite mode numbers $n = -N, -N+1, \ldots, 0, \ldots, N-1, N$. The radial function of the Zernike polynomials is a shifted Jacobi polynomial, which fulfils the mathematical condition for physical scalars on the unit disc (Lewis & Bellan 1990). Due to brevity, we omit the full basis description and refer interested readers to Dudt & Kolemen (2020) and Panici *et al.* (2023).

Each equilibrium solved by DESC is defined by the result of the nonlinear least-squares minimisation

$$\begin{aligned} &\text{minimise} && ||(\nabla\times\boldsymbol{B})\times\boldsymbol{B} - \mu_0\nabla\boldsymbol{p}||_2^2 \\ &\text{subject to} && \boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{b}, \end{aligned} \tag{2.16}$$

where $\boldsymbol{b}$ is the target for the constraints in $\bar{\boldsymbol{x}} = [\boldsymbol{x}, \boldsymbol{c}]$. In this work, the constraint vector includes the ideal MHD invariants for fixed-boundary equilibria, namely $\boldsymbol{c} = [\boldsymbol{R}_{b,mn}, \boldsymbol{Z}_{b,mn}, \boldsymbol{p}_l, \boldsymbol{\iota}_l, \psi_{\text{b}}]^{\mathsf{T}}$ (Conlin *et al.* 2022). $\boldsymbol{\iota}_l$ are the coefficients of the prescribed rotational transform profile $\iota(\rho)$, or toroidal current profile $c(\rho)$, and $\boldsymbol{p}_l$ are the coefficients of some isotropic pressure function $\boldsymbol{p} = p(\rho)$ (see table 1) with the indices l depending on each prescribed function.

The constrained minimisation problem defined by (2.16) is then transformed into an unconstrained problem by splitting $\bar{\boldsymbol{x}}$ into a particular solution $\boldsymbol{x}_{\text{p}}$ and a vector $\boldsymbol{y}$ on a hyperplane defined by the constraint manifold $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

$$\begin{aligned} &\boldsymbol{A}\boldsymbol{Z} = 0 \\ &\boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x}_{\text{p}} + \boldsymbol{Z}\boldsymbol{y}) = \boldsymbol{b} \\ \iff &\boldsymbol{A}(\boldsymbol{x}_{\text{p}} + \boldsymbol{Z}\boldsymbol{y}) = \boldsymbol{b}. \end{aligned} \tag{2.17}$$

The nullspace $\boldsymbol{Z}$ then only needs to be computed once before the start of optimisation via singular value decomposition, yielding an unconstrained optimisation problem over the projected parameter vector $\boldsymbol{y} = \boldsymbol{Z}^{\mathsf{T}}(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\text{p}})$.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

DESC commonly uses an incremental minimisation, or automatic continuation, parametrised by two multipliers: $\eta_b \in [0, 1]$ and $\eta_p \in [0, 1]$. Starting with a circular tokamak, each incremental step either changes η_b, which deforms the initial circular plasma boundary into the desired stellarator shape, or η_p, which ramps up the pressure.

The NN based parametrisation of this work sets $\boldsymbol{y}$ as the output layer of simple MLPs and then minimises the sum of force residuals of equilibria defined by a set $\eta_{p,\text{train}}$.

3. Physics informed neural networks in DESC

Instead of directly optimising over $\boldsymbol{y}$ (2.17) as done in DESC, we will set $\boldsymbol{y}$ as the output of two-layer MLPs and optimise over the MLP parameters $\boldsymbol{\nu}$ in a physics informed neural network (PINN) approach (Raissi, Perdikaris & Karniadakis 2019). To this end, we reformulate the optimisation problem as

$$\boldsymbol{\nu}^\star = \arg\min_{\boldsymbol{\nu}} \quad \mathcal{L}_{\text{op}} \tag{3.1}$$

and solve it with the L-BFGS optimiser for some loss function $\mathcal{L}_{\text{op}}$. Paluzo-Hidalgo, Gonzalez-Diaz & Gutiérrez-Naranjo (2020) showed that MLPs with two hidden layers and nonlinearities consisting of rectified linear units can approximate arbitrary functions, and in Fourier space, optimising over parameters of MLPs with two hidden layers is sufficient to solve single, fixed-boundary and finite-$\langle\boldsymbol{\beta}\rangle_{\text{vol}}$ MHD equilibria with the lowest $\langle\boldsymbol{F}\rangle_{\text{vol,norm}}$ over the tested spectral resolutions and NNs (Thun *et al.* 2026). Setting the number of hidden layers or their node numbers too large in the presented approach yields unusable results because minimisation stagnates in local optima.

The plasma boundary is fixed via projection into $\boldsymbol{y}$ (see (2.17)), which reduces the number of parameters and, in our tests, optimising NN with the Fourier Zernike modes in the output layer did not work. As initial guess for all presented narrow operator models we use the default DESC initial guess. If it fails to produce nested flux surfaces, for example for the quasi-helical equilibrium (figure 3), an invertible mapping to boundary conforming coordinates introduced by Babin *et al.* (2025) calculates the $N = 2$ axis initial guess. This axis guess is then interpolated towards the boundary, ensuring a well-posed initial guess throughout the volume. The initial guess in Fourier Zernike space is projected into the tangent space as $\boldsymbol{y}_{\text{init}}$, using the same $\boldsymbol{A}$ and $\boldsymbol{Z}$ matrices, and then added to the MLP prediction (see (3.7)).

We show that it is possible to minimise the sum of ideal MHD force residuals defined by equilibria evenly distributed in $\eta_{p,\text{train}}$ over the parameters of MLPs with two hidden layers. Each operator MLP parametrises the function $\text{MLP} : \eta_{p,i} \to \boldsymbol{y}_i$ for $\eta_{p,\text{train}} = \{\eta_{p,i=0}, \ldots, \eta_{p,i=I-1}\}$. The input can be easily modified to include, for example, the boundary Fourier modes or rotational transform coefficients, but results in this work only use the scalars $\eta_{p,i}$ as input.

Each full training step consists of the model predicting $\boldsymbol{y}_{\text{train}}$ for all $\eta_{p,i}$ and the sum of all residuals for all i as target function, scaled by $\alpha_{\text{MHD}} = 10^7$ to avoid optimisation problems caused by the residual approaching machine precision

$$\mathcal{L}_{\text{op}} = \alpha_{\text{MHD}} \sum_{i=0}^{I-1} |\boldsymbol{f}(\boldsymbol{x})|_i^2 = \alpha_{\text{MHD}} \sum_{i=0}^{I-1} |\hat{\boldsymbol{f}}(\boldsymbol{y})|_i^2, \tag{3.2}$$

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

where $\hat{\boldsymbol{f}}$ is the composition of $\boldsymbol{f}$ and the inverse of the linear projection $\bar{\boldsymbol{x}} = \boldsymbol{x}_\mathrm{p} + \boldsymbol{Z}\boldsymbol{y}$.

We use $I = 10$ equispaced $\eta_{\mathrm{p,train},i}$ points to train all presented narrow operator models. The MLPs use the same activation function as self-normalising NNs (Klambauer *et al.* 2017), which Merlo *et al.* (2021) also deemed optimal through hyperparameter search

$$\sigma(x) = \mathrm{selu}(x) = \lambda_s \begin{cases} x, & x > 0, \\ \alpha_s e^x - \alpha_s, & x \leqslant 0, \end{cases} \tag{3.3}$$

with $\lambda_s = 1.0507\ldots$ and $\alpha_s = 1.6732\ldots$.

Each MLP has the following functional form:

$$\hat{\boldsymbol{y}}_\mathrm{mlp}(\eta_\mathrm{p,train}) = \boldsymbol{W}_2(\sigma \boldsymbol{z}_1(\eta_\mathrm{p,train})) + \boldsymbol{b}_2, \tag{3.4}$$

$$\boldsymbol{z}_1(\eta_\mathrm{p,train}) = \boldsymbol{W}_1(\sigma \boldsymbol{z}_0(\eta_\mathrm{p,train})) + \boldsymbol{b}_1, \tag{3.5}$$

$$\boldsymbol{z}_0(\eta_\mathrm{p,train}) = \boldsymbol{W}_0(\eta_\mathrm{p,train}) + \boldsymbol{b}_0. \tag{3.6}$$

Weights of the MLPs $\boldsymbol{W}_k$ for $k = 0, 1, 2$ are initialised with a normal distribution $\mathcal{N}(0, 0.01^2)$, while the bias vectors $\boldsymbol{b}_k$ for $k = 0, 1, 2$ are initialised with $\boldsymbol{0}$.

The MLP output is scaled and added to the linear projection of the initial guess, which is necessary for convergence for all non-axisymmetric equilibria we tested

$$\boldsymbol{y} = \boldsymbol{y}_\mathrm{init} + \boldsymbol{y}_\mathrm{scale}\hat{\boldsymbol{y}}_\mathrm{mlp}. \tag{3.7}$$

$\boldsymbol{y}_\mathrm{init}$ is the projected initial guess, and the scaling vector $\boldsymbol{y}_\mathrm{scale}$ is the projection of the inverse of the sum of absolute mode numbers l_{ZP}, m and n (see table 1 for the non-projected scales). All $\boldsymbol{y}$ in (3.7) are projected with the same $\boldsymbol{A}$ and $\boldsymbol{Z}$ operators (see (2.17)).

Optimisation of each MLP is split into two stages: first, the loss function is modified to only include the outliers, i.e. $i = 0$ and $i = I - 1$, and in a second minimisation all I equilibria are included. The trained models are then tested on $\eta_\mathrm{p,test}$, which oversamples I by a factor of 10, staying within the interval $[\eta_{\mathrm{p},0}, \eta_{\mathrm{p},I-1}]$ (see figure 2) and an extrapolation of each model is plotted in figure 6. We provide detailed hyperparameters for models and optimisation in table 1 and code which reproduces the plots in the supplementary data.

4. Results

This section compares `DESC`'s `lsq-exact` optimiser with an `L-BFGS` optimiser applied to the free parameters $\boldsymbol{v}$ of MLPs that parametrise the linear projection of the Fourier Zernike basis over $\eta_\mathrm{p,train}$. All equilibria we show are fixed-boundary equilibria with $\langle \boldsymbol{\beta} \rangle_\mathrm{vol} > 0$ and constant rotational transform or current profile. Furthermore, all results presented in this section do not use continuation methods or iterative refinement of the grid and compute $\mathcal{L}_\mathrm{op}$ (3.2) on concentric grids commonly used by `DESC` (Conlin *et al.* 2022). Volume- or surface-averaged quantities are calculated on quadrature grids.

Each `DESC` equilibrium in this comparison is solved in two stages: first, the equilibrium is optimised with automatic continuation and moderate tolerances, and in a second optimisation the tolerances of the resulting equilibrium are reduced to zero with a maximum of 100 iterations. If automatic continuation yields intermediate equilibria that `DESC` cannot solve, we instead solve the equilibrium without

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

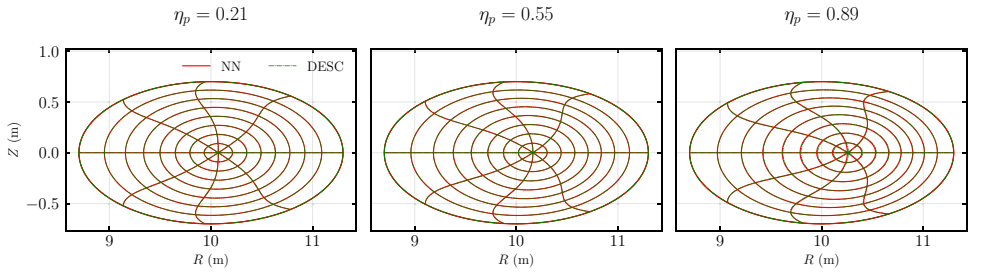


FIGURE 1. Operator MLP solution of Heliotron-like equilibria at $\zeta = 0$ with elliptical boundary for different pressure scaling factors η_p. The DESC solution in green and MLP solution in red match qualitatively for the plotted flux surfaces.

automatic continuation. This is only the case for some $\eta_{\mathrm{p},i}$ in the quasi-helical configuration. An automatic continuation where η_p is increased first can fail due to intermediate equilibria having unrealistic pressure, and this will be remedied in a future version of DESC by performing continuation in η_b first and then η_p. Lastly, we run DESC with the same spectral resolutions M and N as prescribed in the input files which are included in the supplementary material, and the Zernike polynomials are of order $L_{ZP} = M$. Except for the Wendelstein 7-X (W7-X) equilibrium where we use $L_{ZP} = M + 1 = 7$ as $L_{ZP} = 6$ did not resolve the Shafranov shift properly.

Comparing the MLP operator model to DESC solutions is only possible at discrete points $\eta_{\mathrm{p,train}}$, because DESC solves single equilibria rather than spaces of equilibria. DESC solutions in figure 2 are marked with a plus sign while the training points of the operator models are marked by a cross.

We present operator models for a $\mathrm{N_{FP}} = 5$ W7-X-like equilibrium in standard configuration, a $\mathrm{N_{FP}} = 19$ Heliotron-like equilibrium, a $\mathrm{N_{FP}} = 4$ quasi-helical equilibrium and an axisymmetric, but not stellarator symmetric, equilibrium akin to the experimental device DIII-D. The W7-X equilibrium is a good example for a three-dimensional plasma in an experimental device, and the quasi-helical equilibrium is representative for optimised quasi-symmetric stellarators with self-consistent bootstrap current (Landreman, Buller & Drevlak 2022).

Out of all three stellarator equilibria, the Heliotron-like equilibrium has the highest sensitivity of its axis position with respect to the plasma $\langle\boldsymbol{\beta}\rangle_\mathrm{vol}$, moving its axis by 25.7cm between $\langle\boldsymbol{\beta}\rangle_\mathrm{vol}(\eta_{\mathrm{p},0}) = 1.018\,\%$ and $\langle\boldsymbol{\beta}\rangle_\mathrm{vol}(\eta_{\mathrm{p},9}) = 10.18\,\%$. The operator model is able to resolve this change along $\eta_{\mathrm{p},i}$, as seen in figure 1. Additionally, we plot the flux surfaces of operator NN similar to DIII-D and W7-X in standard configuration in Appendix A.

Table 1 provides the optimisation parameters of each narrow operator model. The last line of table 1 shows the maximum discrepancy for $\eta_{\mathrm{p,train}}$ between the DESC solution and the model solution

$$\epsilon_\mathrm{X} = \frac{1}{N}\sum_i \sqrt{(R_{NN,i} - R_{\mathrm{DESC},i})^2 + (Z_{NN,i} - Z_{\mathrm{DESC},i})^2} \tag{4.1}$$

evaluated on an equal grid with collocation points indexed by $i = 0, 1, \ldots, N$.

Figure 1 illustrates the $\boldsymbol{B}$-field topology at $\zeta = 0$ of the Heliotron equilibrium for $\eta_\mathrm{p} = \{0.21, 0.55, 0.89\}$, which are all points on which the model was not trained, but which lie within the training interval $\eta_{\mathrm{p,train}}$. The MLP parametrised solution is plotted in red while the solution of DESC is plotted in green and both agree well for all $\eta_\mathrm{p} \in [0.1, 1]$.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

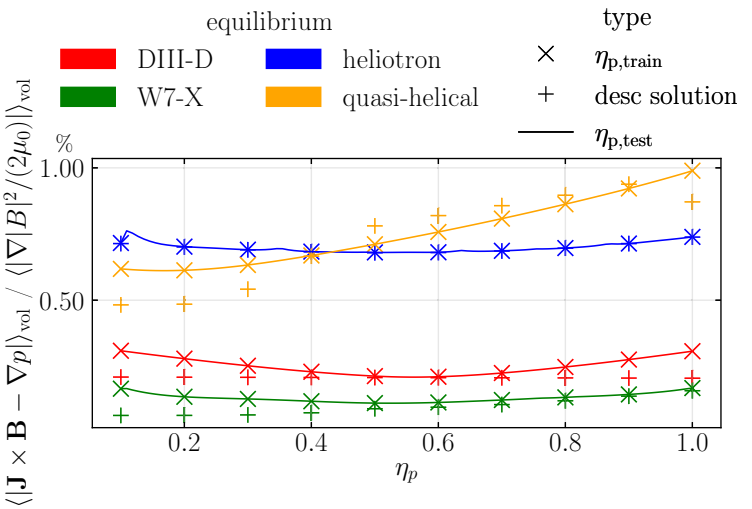


FIGURE 2. Operator MLP solutions for equilibria presented in this work and trained on $I = 10$ equispaced $\eta_{\mathrm{p,train}}$ (cross signs) compared with their DESC solution (plus signs) at the training points in terms of $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$.

Figure 2 plots the scalar quantity $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ over $\eta_{\mathrm{p}} \in [0.1, 1]$ for all tested equilibria. The models and DESC's force error of the Heliotron equilibrium match well for all $\eta_{\mathrm{p}} \in [0.1, 1]$ with the largest discrepancy around $\eta_p = 0.11$, where the operator model shows a small spike in force error. Removing this spike requires increasing the number of training points I. All NN models show good agreement with DESC and stay below $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}} < 1\,\%$. In contrast, DESC achieves comparable $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ in the quasi-helical and lower $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ for the DIII-D and W7-X test cases for all $\eta_{\mathrm{p,train}}$.

To illustrate quantities of interest that depend on higher-order derivatives, we showcase quasi-symmetry in Boozer coordinates for the quasi-helical equilibrium with self-consistent bootstrap current computed at $\eta_{\mathrm{p,train}} = 1$. Throughout $\eta_{\mathrm{p,test}}$, the quasi-helical operator MLP shows good quasi-symmetry at radial position $s = \rho^2 = 0.75$ plotted for $\eta_{\mathrm{p}} = \{0.21, 0.55, 0.89\}$ in figure 3. Only the maxima of $|\boldsymbol{B}|$ change slightly with decreasing η_{p} at θ_{Boozer} close to 0. The topology of the magnetic well is qualitatively preserved.

4.1. *Discussion*

Optimisation of the presented narrow operator learning models was stopped at an arbitrary number of iterations, and further optimisation could yield models that close existing gaps between DESC's and the model's $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ (see figure 2).

Including $\eta_{\mathrm{p}} < 0.1$ in the training set resulted in slight differences of the $\boldsymbol{B}$-field topology between the model and DESC solutions at low η_{p}. This could be caused by large-scale differences in $\boldsymbol{y}$ and whether such low beta regions are relevant for flight simulators remains an open question. Rigid start-up sequences of experimental plasmas (Grulke *et al.* 2024) and control at high densities increase the importance of operator models closer to $\eta_{\mathrm{p}} = 1$. Evaluating such models at $\eta_{\mathrm{p,test}} > 1$, i.e. outside $\eta_{\mathrm{p,train}}$, shows a monotonic increase in $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ (see Appendix B). Extrapolation with this approach to unseen equilibria seems unlikely, but extending $\eta_{\mathrm{p,train}}$ to relevant regions is straightforward.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

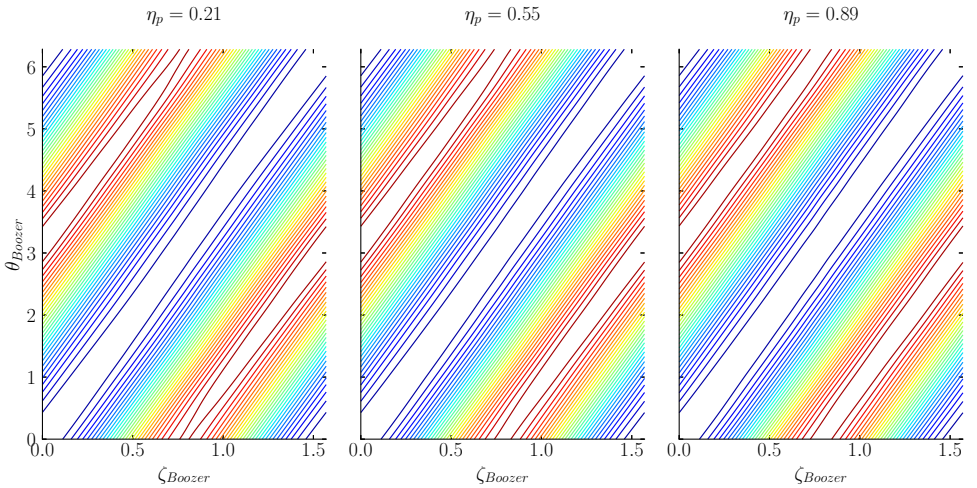


FIGURE 3. Good quasi-symmetry for the $N_{FP} = 4$ quasi-helical equilibrium at $s = \rho^2 = 0.75$ for $\eta_p = \{0.21, 0.55, 0.89\}$ with a constant current that was optimised for $\eta_p = 1$. θ_{Boozer} and ζ_{Boozer} are straight-field line coordinates in which transport equations are nearly isomorphic to axisymmetric equilibria (Pytte & Boozer 1981).

Increasing the MLP layer size or depth, or increasing the spectral resolution L_{ZP}, M or N too much, forces the minimisation to settle in local minima, far away from DESC's optima. Automatic continuation methods similar to DESC could avoid these local minima.

The presented optimisation of narrow operator models yields models that capture the equilibria in $\eta_{p,test}$ as good as DESC and can even achieve lower $\langle \boldsymbol{F} \rangle_{vol,norm}$ in some regions. In all the tested cases, we see that optimising over an ensemble of (2.16), parametrised by the set $\eta_{p,train,i}$, can yield a continuous and precise model of the narrow PDE operator. To verify this, we optimised $\boldsymbol{y}$ with DESC again until the change in parameters was below machine precision (i.e. with a ceiling of 10^6 `lsq-exact` iterations) and arrived at the same conclusion and qualitatively equal results as plotted in figure 2.

Quantities that depend on higher-order derivatives of the equilibrium magnetic field such as the quasi-symmetry evaluated in Boozer coordinates and the magnetic well are preserved for the quasi-helical test case (figure 3).

Training these narrow operator models incurs higher computational costs compared with verifying the model with I DESC solutions (see Appendix C), however, the increased cost must be weighted against the advantages of continuously parametrised models over $\eta_{p,test}$.

One common approach to training ideal MHD operator models is to construct a dataset of equilibrium magnetic fields with a conventional solver and then train a model on this dataset, and possibly an additional physics-based addition to the loss function. For the quasi-helical equilibrium in figure 2, this training scheme would not improve upon DESC's force error for $\eta_p < 0.4$, hinting that additional training of operator models directly on the physics yields more precise models, or optimisation potential in DESC's minimisation for this specific equilibrium at $\eta_p < 0.4$.

Merlo *et al.* (2023*a*) also present improvements in optimisation with operator models trained on a surrogate for the force residual (2.16) that assumes the helical force F_β to be zero.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

4.2. *Outlook*

To improve the applicability and training efficiency of the presented models, future work should explore the sampling granularity in $\eta_{\mathrm{p,train}}$ required to achieve good force error over the parameter range and increase the NN complexity while introducing more parameters like $\iota(\rho)$ coefficients. A solution for the optimisation stagnating in local minima must be found when increasing the number of parameters.

In the Heliotron case, the model shows a spike between training points at $\eta_{\mathrm{p}} \in [0.1, 0.2]$, whereas the other operator models follow a continuous trend, raising the question of how many training points I are required for the latter without degradation of $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ in $\eta_{\mathrm{p,test}}$. Investigating the optimal ratio of data from a solver and direct force residual in training sets of operator models could further reduce computational cost and help avoid local minima.

Finding commonalities in the parameters of the presented narrow operator models in terms of $\boldsymbol{x}$ instead of $\boldsymbol{y}$ could yield more efficient optimisation, also named transfer learning in PINN research (Goswami *et al.* 2022). Care must be taken with regard to differently shaped $\boldsymbol{y}$, but modification of the linear operators $\boldsymbol{A}$ and $\boldsymbol{Z}$ to include all equilibria under consideration can alleviate those issues.

Modification of the presented approach to include different inputs, for example the boundary coefficients, can improve sensitivity analysis because first-order gradients of the force residual to the input space are easily computed by automatic differentiation. Using one of the presented narrow operator models delivers continuous gradients of dependent to independent variables over the space parametrised by $\eta_{\mathrm{p,test}}$.

In quasi-isodynamic stellarator optimisation the pressure profiles are usually fixed *a priori* (Sánchez *et al.* 2023; Gaur *et al.* 2024; Goodman *et al.* 2024) and more diverse profiles could yield lower multi-objective targets or more flexible configurations. Optimisation for flexible configurations could also yield more robust optimised stellarators.

Extending the narrow operator models to free-boundary equilibria is not straightforward: the DESC suite already includes numerical free-boundary computation, but in our preliminary research we found that continuation methods are indispensable to solve free-boundary problems in DESC and these incur a change in the shape and encoded information of $\boldsymbol{y}$. However, re-evaluating free-boundary operator models in DESC is more promising with the mentioned improvements in this section, especially transfer learning in terms of $\boldsymbol{x}$ and customised linear matrices $\boldsymbol{A}$ and $\boldsymbol{Z}$.

5. Conclusion

We presented narrow operator models in the form of MLPs with two hidden layers parametrised by a scaling factor of the pressure coefficients. These MLPs reduce $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ of various equilibria types to comparable levels computed by the modern solver DESC. All introduced models precisely interpolate through an oversampled test set $\eta_{\mathrm{p,test}}$ with all points still inside the training set $\eta_{\mathrm{p,train}}$ for all tested equilibria (see figure 2). Importantly, they are trained only on a physics-based residual (2.16), that is without using any equilibrium information pre-computed by existing solver like DESC.

For the DIII-D like and W7-X equilibria, DESC computes equilibria with marginally lower $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ compared with the MLP parametrisation, while for the quasi-helical equilibrium the MLP approaches slightly lower $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ than DESC for $\eta_{\mathrm{p}} \in [0.4, 0.9]$ (see figure 2). This is interesting for future operator model optimisation because it hints at a benefit when training on the force residual: if the training

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

set only consisted of pre-computed DESC solutions, the operator model's $\langle \boldsymbol{F} \rangle_{\text{vol,norm}}$ would be bounded by DESC's $\langle \boldsymbol{F} \rangle_{\text{vol,norm}}$.

The narrow operator model of the quasi-helical equilibrium with self-consistent bootstrap current preserves good higher-order metrics such as quasi-symmetry and magnetic well throughout $\eta_{\text{p,test}}$ (see figure 3).

Extrapolation of the model to unseen $\eta_{\text{p}} > 1$ incurs a monotonically increasing $\langle \boldsymbol{F} \rangle_{\text{vol,norm}}$ (see Appendix B), but including $\eta_{\text{p,train}} > 1$ in the training set is straightforward.

Our training scheme was purposedly kept minimal, in the sense that no advanced enhancements from the PINN literature were included, to evaluate simple MLPs trained solely on the force residual as a baseline for future work. We expect significant improvements to the presented method if recent advances in PINN research are included (Luo *et al.* 2025).

Acknowledgements

The authors thank M. Landreman for valueable feedback and provision of the quasi-helical equilibrium, and we thank J. Geiger for providing the W7-X equilibrium. Our gratitude extends to E. Kolemen for guidance throughout this work and international visiting opportunities. Furthermore, we appreciate R. Gaur's feedback during the project. Last but not least, we are indebted to the various maintainers of the open source codes used in this work: `jax`, `optax`, `jaxopt`, `flax`, `orbax`, `desc` and `matplotlib`. T.T. is supported by a grant from the Simons Foundation (Grant No. 601966). D.P. is funded through the SciDAC program by the US Department of Energy, Office of Fusion Energy Science and Office of Advanced Scientific Computing Research under contract number DE-AC02-09CH11466, DE-SC0022005, and by the Simons Foundation/SFARI (560651). This work was supported by a DOE Distinguished Scientist Award via DOE Contract DE-AC02-09CH11466 at the Princeton Plasma Physics Laboratory. This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No. 101052200 -EUROfusion). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

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

Funding Statement

Open access funding provided by Max Planck Society.

Declaration of interests

The authors report no conflicts of interest.

Data availability

DESC version 13.0 was used for the results in this work. Result Data is provided at https://doi.org/10.5281/zenodo.17360265, including scripts that use trained MLP parameters $\boldsymbol{\nu}$ to regenerate the plots.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

Appendix A. The W7-X and DIII-D equilibria

We show Poincaré sections for the DIII-D like and W7-X equilibria in figures 4 and 5. The slight gap in $\langle \boldsymbol{F} \rangle_{\text{vol,norm}}$ between DESC and the NN for low η_{p} leads to a visible shift in the axis position for W7-X in standard configuration (figure 5, $\eta_{\text{p}} = 0.21$).

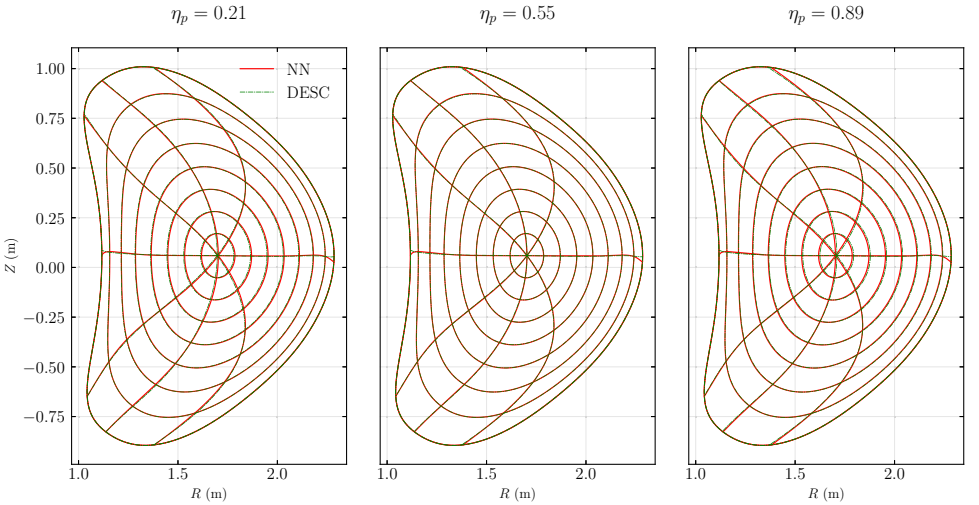


FIGURE 4. Operator MLP solution of an axisymmetric equilibria akin to DIII-D and parametrised by pressure scaling factor η_{p}. The DESC solution in green and MLP solution in red match qualitatively for the plotted flux surfaces.

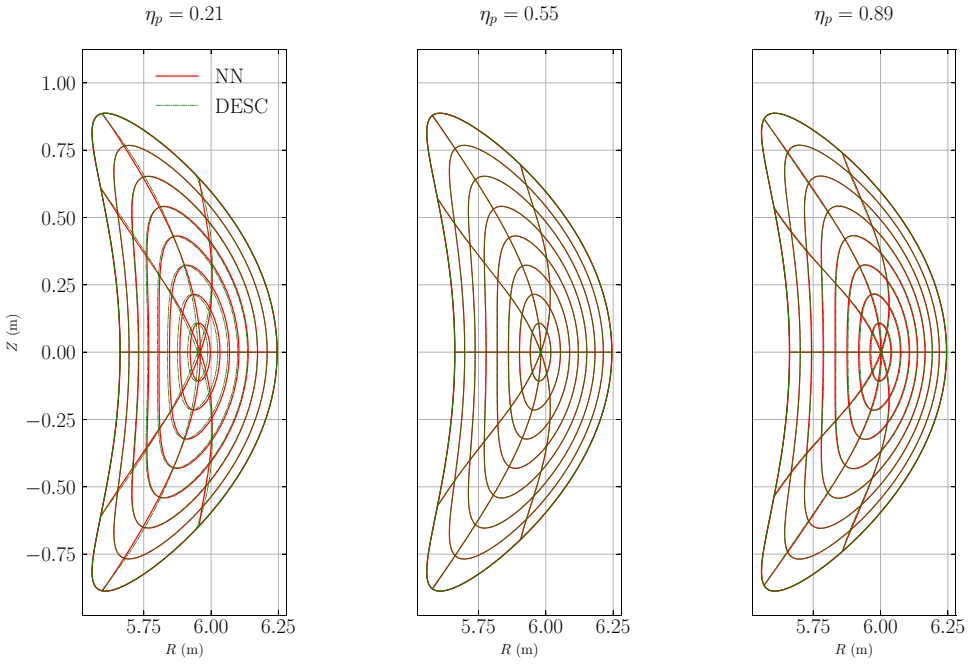


FIGURE 5. Operator MLP solution of W7-X equilibria in standard configuration parametrised by pressure scaling factor η_{p}. The DESC solution in green and MLP solution in red match qualitatively for the plotted flux surfaces, except at $\eta_{\text{p}} = 0.21$.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

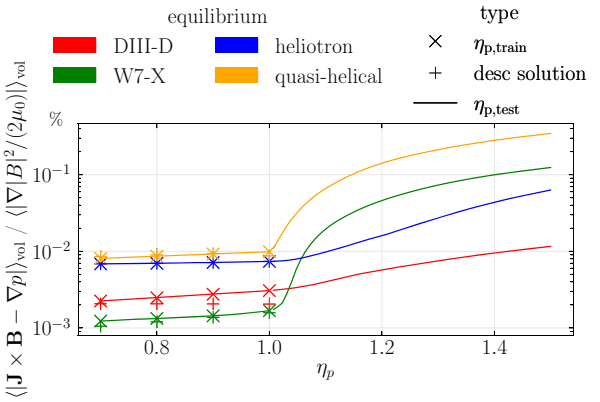


FIGURE 6. Extrapolation of all presented narrow operator models to $\eta_{\mathrm{p,test}} > 1$, outside of $\eta_{\mathrm{p,train}} \in [0.1, 1]$, shows a monotonic increase in $\langle \boldsymbol{F} \rangle_{\mathrm{vol,norm}}$ with increasing $\eta_{\mathrm{p,test}}$.

Appendix B. Extrapolations

Extrapolation of a model trained on $\eta_{\mathrm{p,train}} \in [0.1, 1]$ to $\eta_{\mathrm{p,test}} > 1$ does incur a monotonically increasing force error, as seen in figure 6.

Appendix C. Run time

All DESC and MLP results were computed on the same machine. The operator models for non-axisymmetric equilibria take roughly one to two orders of magnitude more compute resources to train compared with the 10 DESC solutions, while the axisymmetric DIII-D like case was comparable in terms of compute resources to the 10 DESC solutions. We expect improvements to the MLP approach if enhancements from the PINN literature or automatic continuation are added to the current MLP minimisation.

REFERENCES

BABIN, R., HINDENLANG, F., MAJ, O. & KÖBERL, R. 2025 Construction of an invertible mapping to boundary conforming coordinates for arbitrarily shaped toroidal domains. *Plasma Phys. Control. Fusion* **67**, 035005.

BRUNO, O.P. & LAURENCE, P. 1996 Existence of three-dimensional toroidal MHD equilibria with nonconstant pressure. *Commun. Pure Appl. Math.* **49**, 717–764.

CONLIN, R., DUDT, D.W., PANICI, D. & KOLEMEN, E. 2022 The DESC stellarator code suite part II: perturbation and continuation methods. *J. Plasma Phys.* arXiv: 2203.15927.

CONLIN, R., SCHILLING, J., DUDT, D.W., PANICI, D., JORGE, R. & KOLEMEN, E. 2024 High order free boundary MHD equilibria in DESC. arXiv: 2412.05680.

DEGRAVE, J. *et al.* 2022 Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature* **602**, 414–419.

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

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

FREIDBERG, J.P. 2014 *Ideal MHD*. Cambridge University Press.

GAUR, R. *et al.* 2024 Omnigenous stellarator equilibria with enhanced stability. arXiv: 2410.04576.

GOODMAN, A.G. *et al.* 2024 Quasi-isodynamic stellarators with low turbulence as fusion reactor candidates. *PRX Energy* **3**, 023010.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

Goswami, S., Kontolati, K., Shields, M.D. & Karniadakis, G.E. 2022 Deep transfer operator learning for partial differential equations under conditional shift. *Nat. Mach. Intell.* **4**, 1155–1164.

Grulke, O. *et al.* 2024 Overview of the first Wendelstein 7-X long pulse campaign with fully water-cooled plasma facing components. *Nucl. Fusion* **64**, 112002.

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

Hindenlang, F.J., Plunk, G.G. & Maj, O. 2025 Computing MHD equilibria of stellarators with a flexible coordinate frame. arXiv: 2410.17595.

Hirshman, S.P. & Meier, H.K. 1985 Optimized Fourier representations for three-dimensional magnetic surfaces. *Phys. Fluids* **28**, 1387–1391.

Hirshman, S.P. & Whitson, J.C. 1983 Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria. *Phys. Fluids* **26**, 3553–3568.

Hudson, S.R., Dewar, R.L., Dennis, G., Hole, M.J., McGann, M., Von Nessi, G. & Lazerson, S. 2012 Computation of multi-region relaxed magnetohydrodynamic equilibria. *Phys. Plasmas* **19**, 112502.

Hudson, S.R., Panici, D., Zhu, C., Woodbury Saudeau, G.S., Baillod, A., Cianciosa, M. & Ware, A.S. 2025 Verification of the Shafranov shift in free-boundary VMEC, DESC, and SPEC calculations with elliptical geometry. *Phys. Plasmas* **32**, 043906.

Kim, P. *et al.* 2024 Optimization of nonlinear turbulence in stellarators. *J. Plasma Phys.* **90**, 905900210.

Klambauer, G., Unterthiner, T., Mayr, A. & Hochreiter, S. 2017 Self-normalizing neural networks. In *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc.

Kruskal, M.D. & Kulsrud, R.M. 1958 Equilibrium of a magnetically confined plasma in a toroid. *Phys. Fluids* **1**, 265–274.

Landreman, M., Buller, S. & Drevlak, M. 2022 Optimization of quasi-symmetric stellarators with self-consistent bootstrap current and energetic particle confinement. *Phys. Plasmas* **29**, 082501.

Landreman, M., Smith, H.M., Mollén, A. & Helander, P. 2014 Comparison of particle trajectories and collision operators for collisional transport in nonaxisymmetric plasmas. *Phys. Plasmas* **21**, 042503.

Lewis, H.R. & Bellan, P.M. 1990 Physical constraints on the coefficients of Fourier expansions in cylindrical coordinates. *J. Math. Phys.* **31**, 2592–2596.

Luo, K., Zhao, J., Wang, Y., Li, J., Wen, J., Liang, J., Soekmadji, H. & Liao, S. 2025 Physics-informed neural networks for PDE problems: a comprehensive review. *Artif. Intell. Rev.* **58**, 323.

Mandell, N.R., Dorland, W., Abel, I., Gaur, R., Kim, P., Martin, M. & Qian, T. 2024 GX: a GPU-native gyrokinetic turbulence code for tokamak and stellarator design. *J. Plasma Phys.* **90**, 905900402.

Merlo, A., Böckenhoff, D., Schilling, J., Lazerson, S.A., Pedersen, T.S. 2023*a* Physics-regul*a*rized neural network of the ideal-MHD solution operator in Wendelstein 7-X configurations. *Nucl. Fusion* **63**, 066020.

Merlo, A. *et al.* 2023*b* Accelerated Bayesian inference of plasma profiles with self-consistent MHD equilibria at W7-X via neural networks. *J. Instrum.* **18**, 11012.

Merlo, A., Böckenhoff, D., Schilling, J., Höfel, U., Kwak, S., Svensson, J., Pavone, A., Lazerson, S.A. & Pedersen, T.S. 2021 Proof of concept of a fast surrogate model of the VMEC code via neural networks in Wendelstein 7-X scenarios. *Nucl. Fusion* **61**, 096039.

Nikulsin, N., Sengupta, W., Jorge, R. & Bhattacharjee, A. 2024 An asymptotic Grad–Shafranov equation for quasisymmetric stellarators. *J. Plasma Phys.* **90**, 905900608.

Paluzo-Hidalgo, E., Gonzalez-Diaz, R. & Gutiérrez-Naranjo, M.A. 2020 Two-hidden-layer feed-forward networks are universal approximators: a constructive approach. *Neural Netw.* **131**, 29–36.

Panici, D., Conlin, R., Dudt, D.W., Unalmis, K. & Kolemen, E. 2023 The DESC stellarator code suite. Part 1. Quick and accurate equilibria computations. *J. Plasma Phys.* **89**, 955890303.

Plunk, G.G., Drevlak, M., Rodríguez, E., Babin, R., Goodman, A. & Hindenlang, F. 2025 Back to the figure-8 stellarator. *Plasma Phys. Control. Fusion* **67**, 035025.

Pytte, A. & Boozer, A.H. 1981 Neoclassical transport in helically symmetric plasmas. *Phys. Fluids* **24**, 88–92.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press

RAISSI, M., PERDIKARIS, P. & KARNIADAKIS, G.E. 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707.

SCHISSEL, D.P., NAZIKIAN, R.M. & GIBBS, T. 2025 Summary report from the mini-conference on Digital Twins for Fusion Research. *Phys. Plasmas* **32**, 050601.

SENGUPTA, W., RODRIGUEZ, E., JORGE, R., LANDREMAN, M. & BHATTACHARJEE, A. 2024 Stellarator equilibrium axis-expansion to all orders in distance from the axis for arbitrary plasma beta. *J. Plasma Phys.* **90**, 905900407.

SÁNCHEZ, E., VELASCO, J.L., CALVO, I. & MULAS, S. 2023 A quasi-isodynamic configuration with good confinement of fast ions at low plasma β. *Nucl. Fusion* **63**, 066037.

THUN, T., MERLO, A., CONLIN, R., PANICI, D. & BÖCKENHOFF, D. 2026 Improving ideal MHD equilibrium accuracy with physics-informed neural networks. *Nucl. Fusion* **66**, 026029.

https://doi.org/10.1017/S0022377826101512 Published online by Cambridge University Press