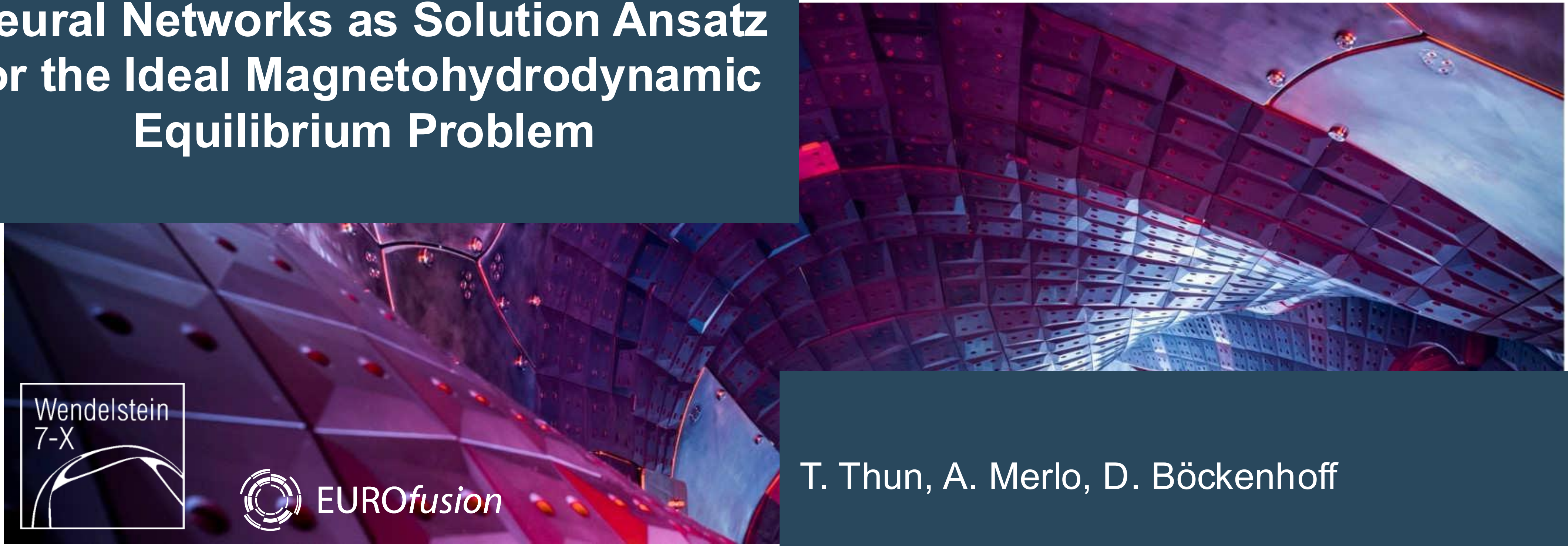
MAX-PLANCK-INSTITUT
FÜR PLASMAPHYSIK

Neural Networks as Solution Ansatz for the Ideal Magnetohydrodynamic Equilibrium Problem



T. Thun, A. Merlo, D. Böckenhoff

THIS WORK WAS FUNDED BY THE SIMONS FOUNDATION

Motivation

- **Different parametrisations (VMEC, DESC, GVEC, …) provide different ideal MHD equilibrium solution** (nested flux surfaces, fixed boundary) **with differing residuals and resource requirements.**

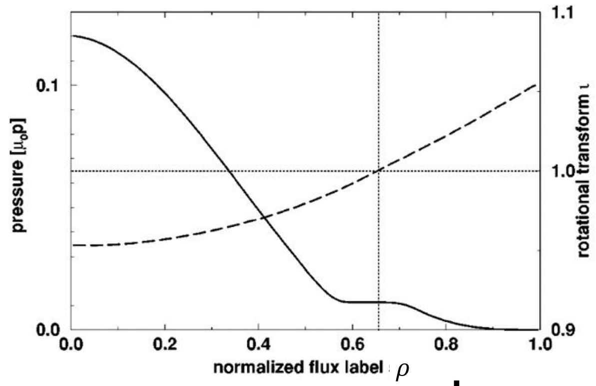
- **Fast solutions of ideal MHD required for real-time inference, plasma flight-simulators, real-time control and many other data-intensive algorithm.**

- **Solution Operator of W7-X configuration space accelerated Bayesian Inference & showed promise for optimisation** [Merlo 2023]**, but is constrained by VMEC parametrization.**

- **How does a Neural Network as parametrization (i.e., without a dataset) of one single equilibrium compares to the "classical" solvers in terms of compactness, efficiency and accuracy?**

NN training on 3D MHD force-residual

$\boldsymbol{BC} \equiv$ Invariants $p(\rho), \iota(\rho)$ and boundary shape

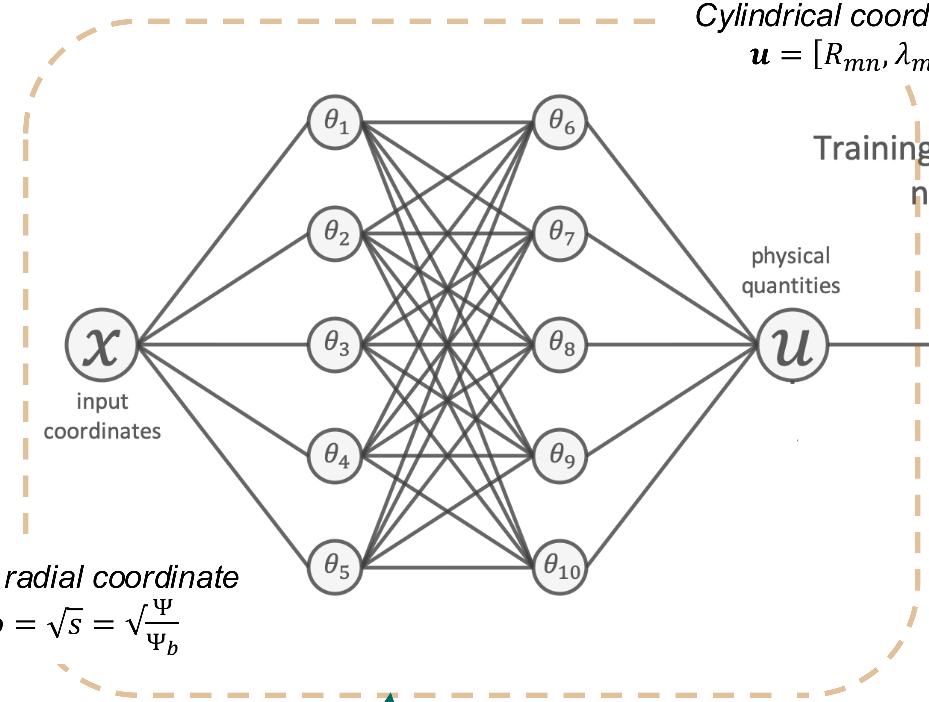


[Nührenberg 2003]

1st step: function learning $[\boldsymbol{R}_{mn}, \boldsymbol{\lambda}_{mn}, \boldsymbol{Z}_{mn}]^T = \boldsymbol{NN}(\rho; \boldsymbol{BC})$
(one equilibrium)

2nd step: operator learning $[\boldsymbol{R}_{mn}, \boldsymbol{\lambda}_{mn}, \boldsymbol{Z}_{mn}]^T = \boldsymbol{NN}(\rho, \boldsymbol{BC})$
(space of equilibria)

Cylindrical coordinate Modes
$\boldsymbol{u} = [R_{mn}, \lambda_{mn}, Z_{mn}]^T$



Training a **physics-informed** neural network

$\frac{du}{dx}, \frac{d^2u}{dx^2}, \ldots$

Compute derivatives and minimise underlying equation residual

$\boldsymbol{\mathcal{F}^{-1}}$

$$\vec{F} = -\vec{j} \times \vec{B} + \vec{\nabla} p = 0,$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j},$$

$$\vec{\nabla} \cdot \vec{B} = 0$$

$$\mu_0 \vec{F} = (\vec{\nabla} \times \vec{B}) \times \vec{B} - \mu_0 \vec{\nabla} p = \boldsymbol{F}_\mu$$

Magnetic radial coordinate
$x = \rho = \sqrt{s} = \sqrt{\frac{\Psi}{\Psi_b}}$

$$L_{PDE} = \frac{1}{2} \left\langle \left|\left| \boldsymbol{F}_\mu \right|\right| \right\rangle^2_{vol}$$

"train" Δ

AdamW & BFGS

$L_{PDE}\ [+L_{spec\_cond}]$

"Loss"

"physics-informed"

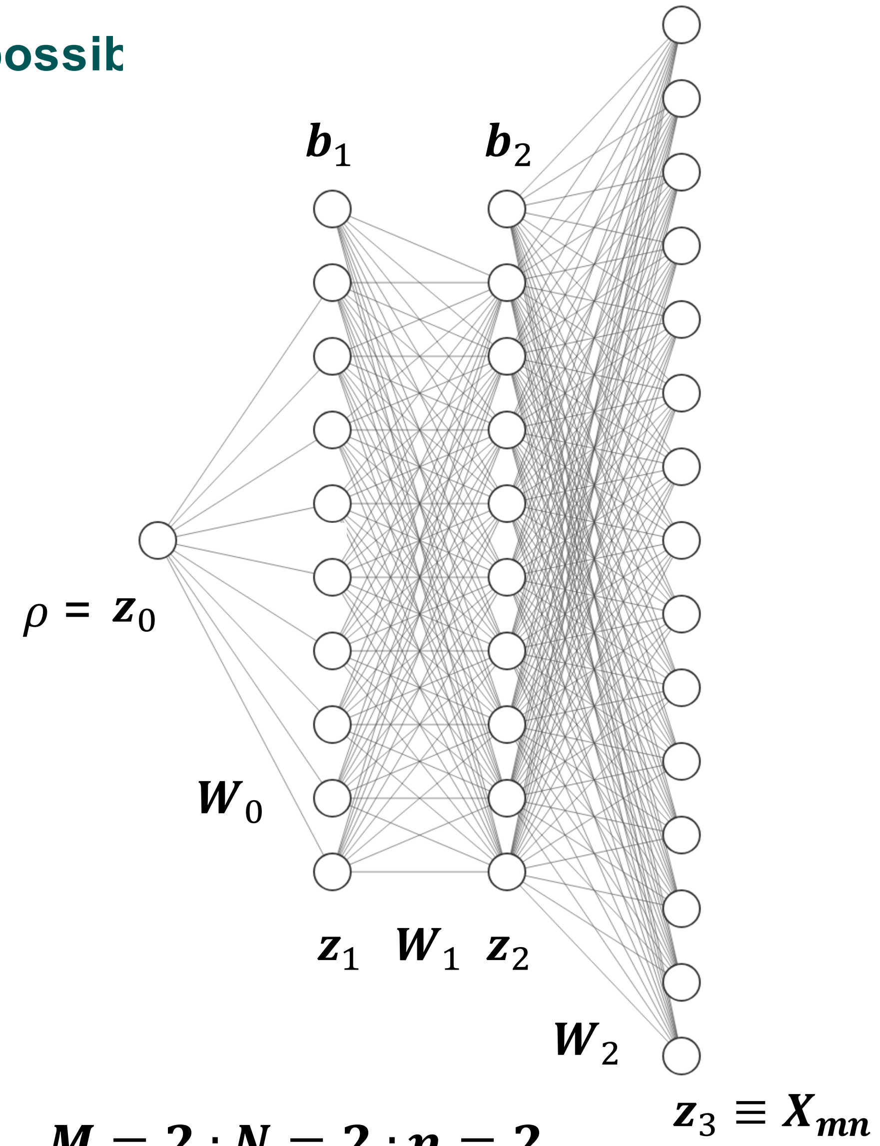
Neural Network (NN) used in this work

- **Task is to learn a single function → NN Topology as simple as possib**
  → 2 hidden layers with n nodes

- **For each dependent mode-vector**
  $\boldsymbol{X}_{mn,i}(\rho) = \{\boldsymbol{R}_{mn}(\rho), \boldsymbol{\lambda}_{mn}(\rho), \boldsymbol{Z}_{mn}(\rho)\}$; $\boldsymbol{X}_{mn,i}(\rho) \in \mathbb{R}^{(M+1)*(2N+1)-N}$
  we use a NN, for a total of 3 NNs.

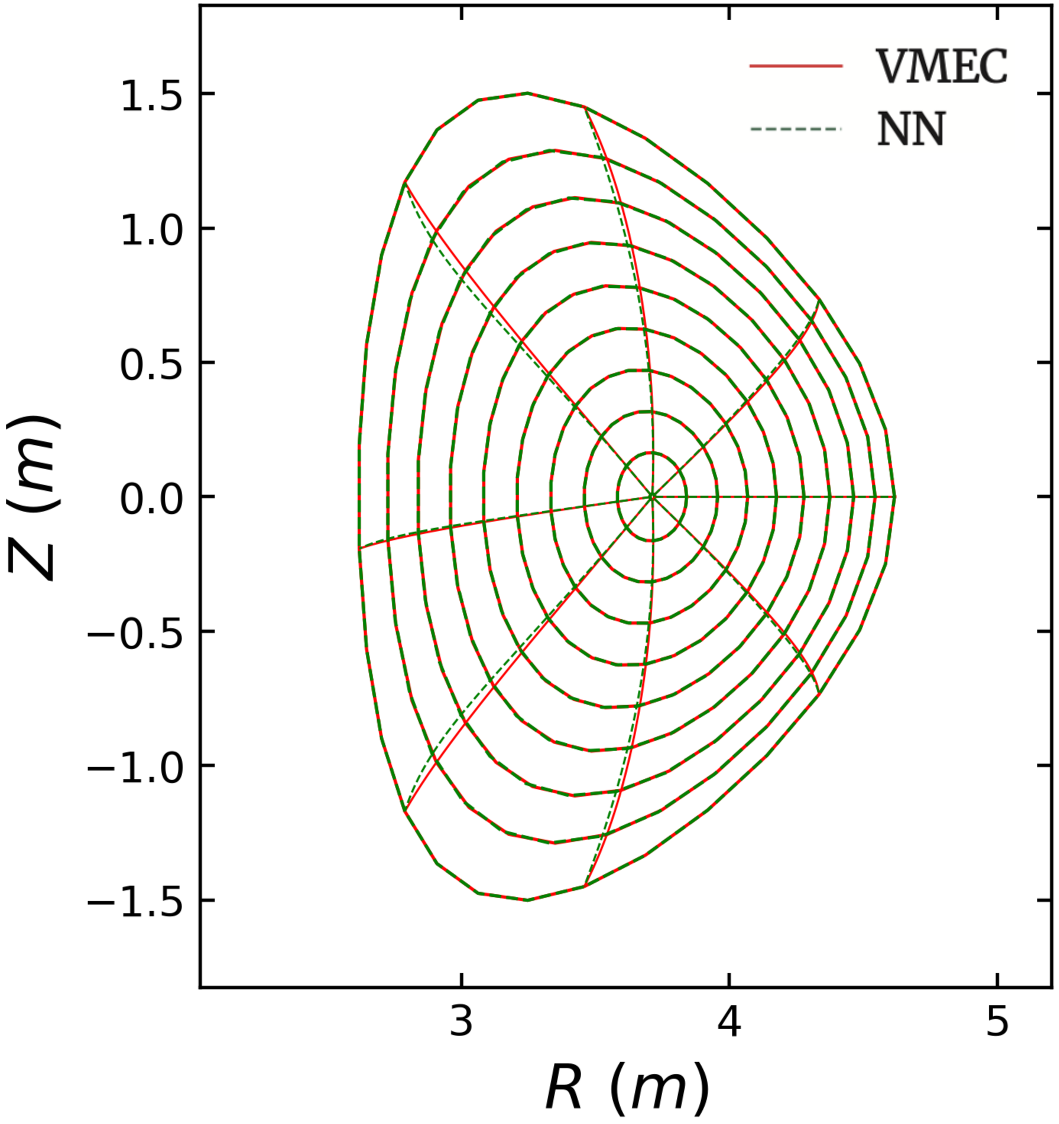
- **Hyperbolic tangent activation function enforces continuous functions**
  $\sigma_i(x) = \ tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

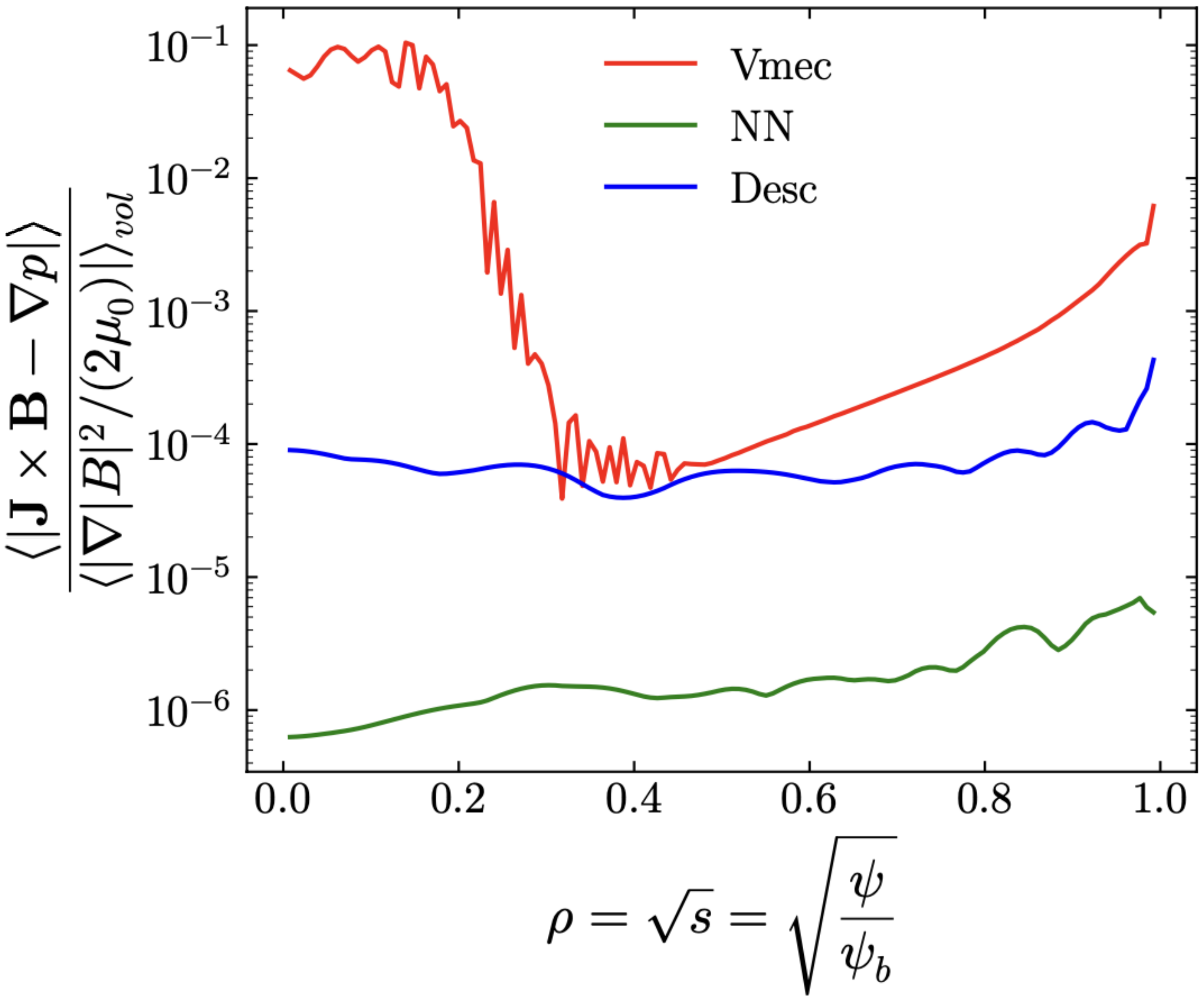
- **No continuation-methods (DESC) or mesh-refinement (VMEC) during training**



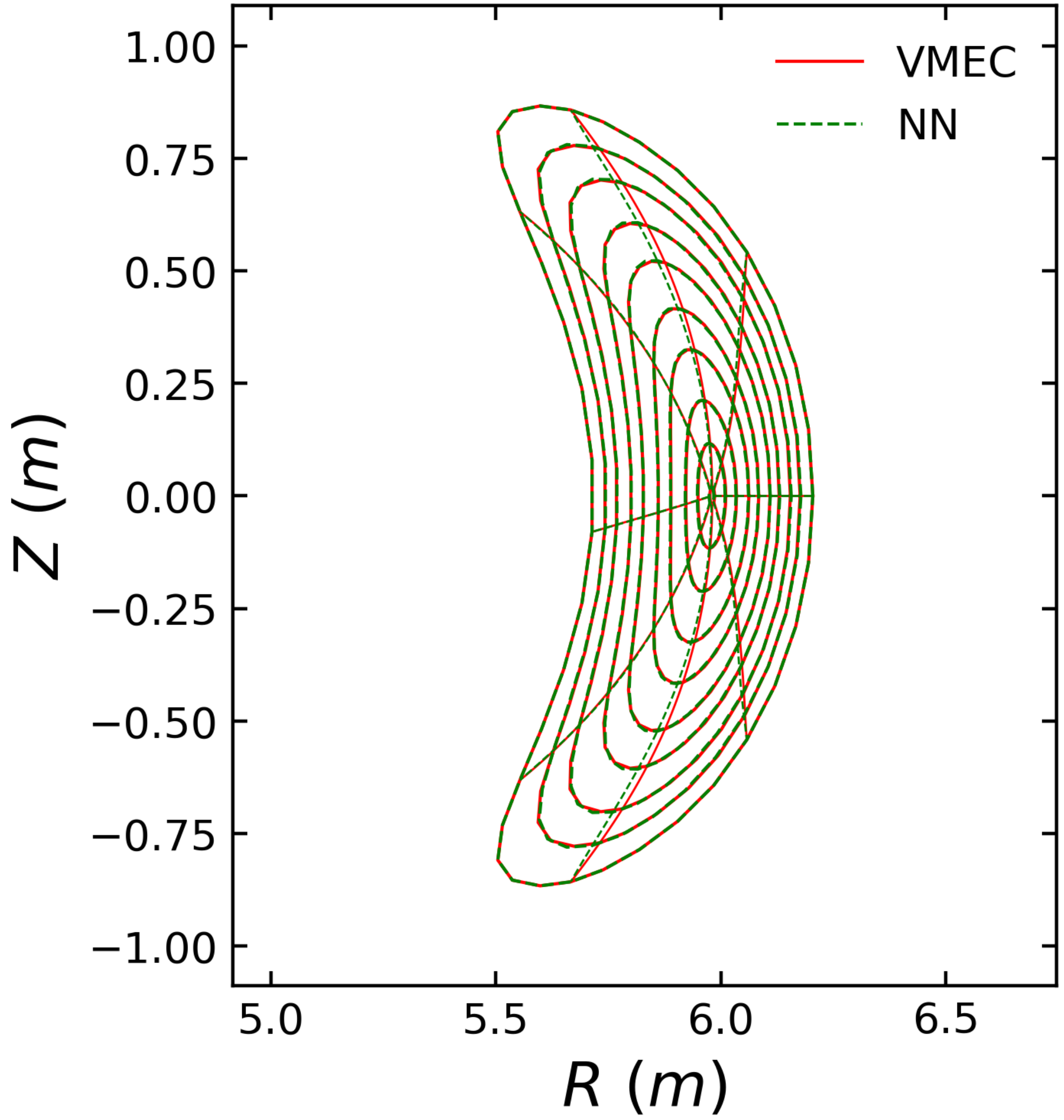
$\boldsymbol{M} = \boldsymbol{2}$; $\boldsymbol{N} = \boldsymbol{2}$; $\boldsymbol{n} = \boldsymbol{2}$

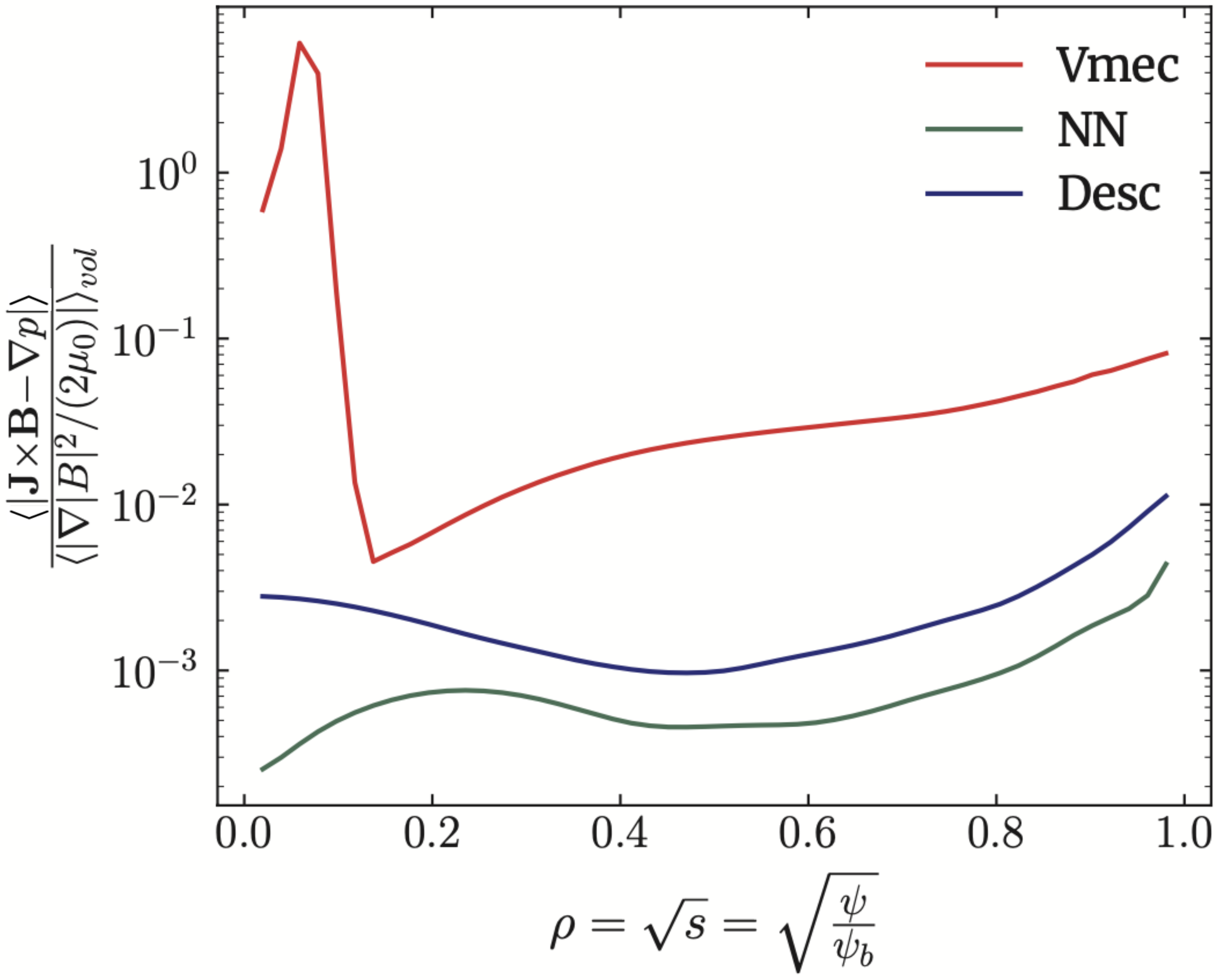
Dshape [Hirshman 1983] $\beta = 2\%$



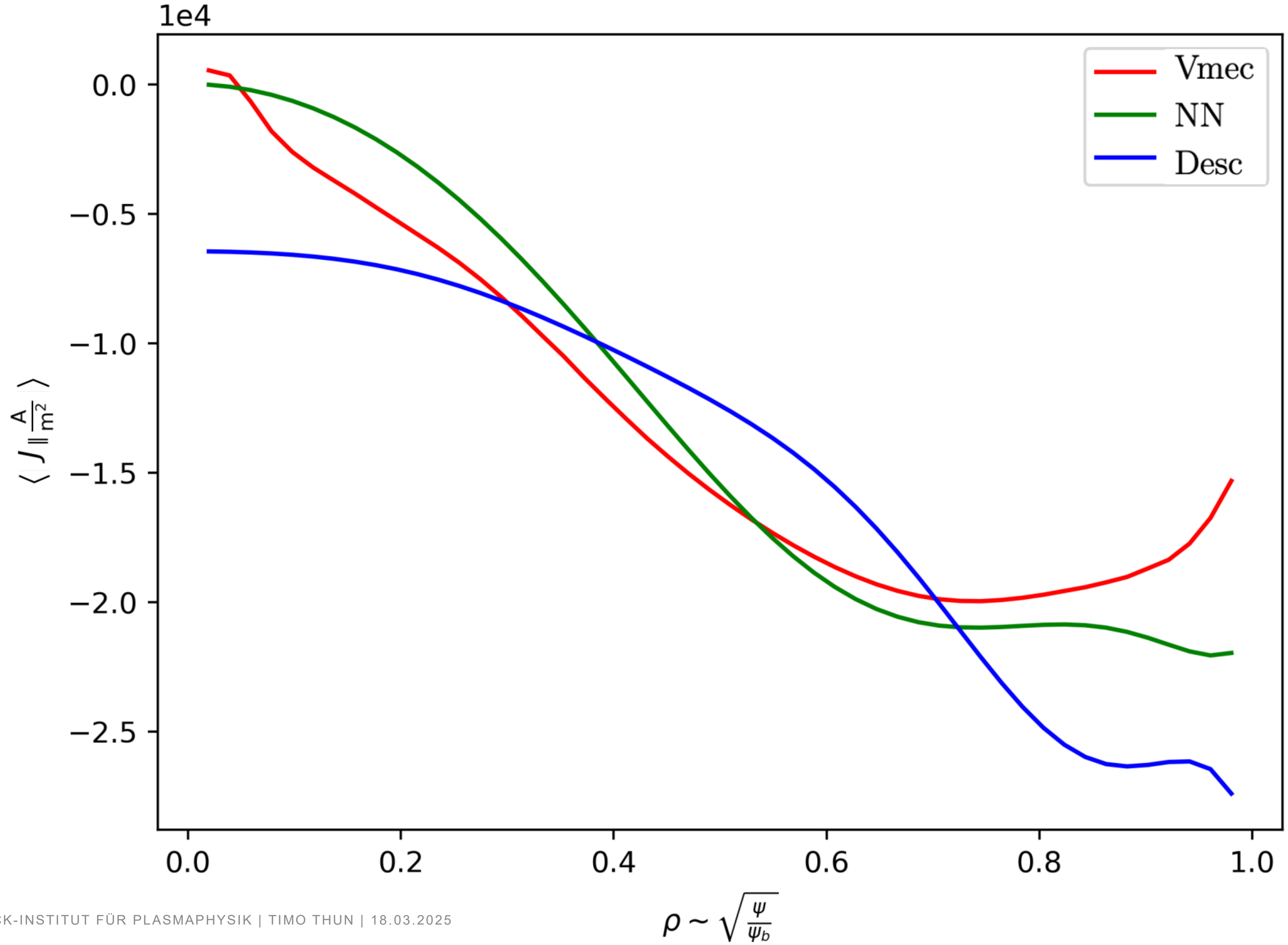


W7X M12 N12 [Panici 2023] $\beta = 2.5\%$

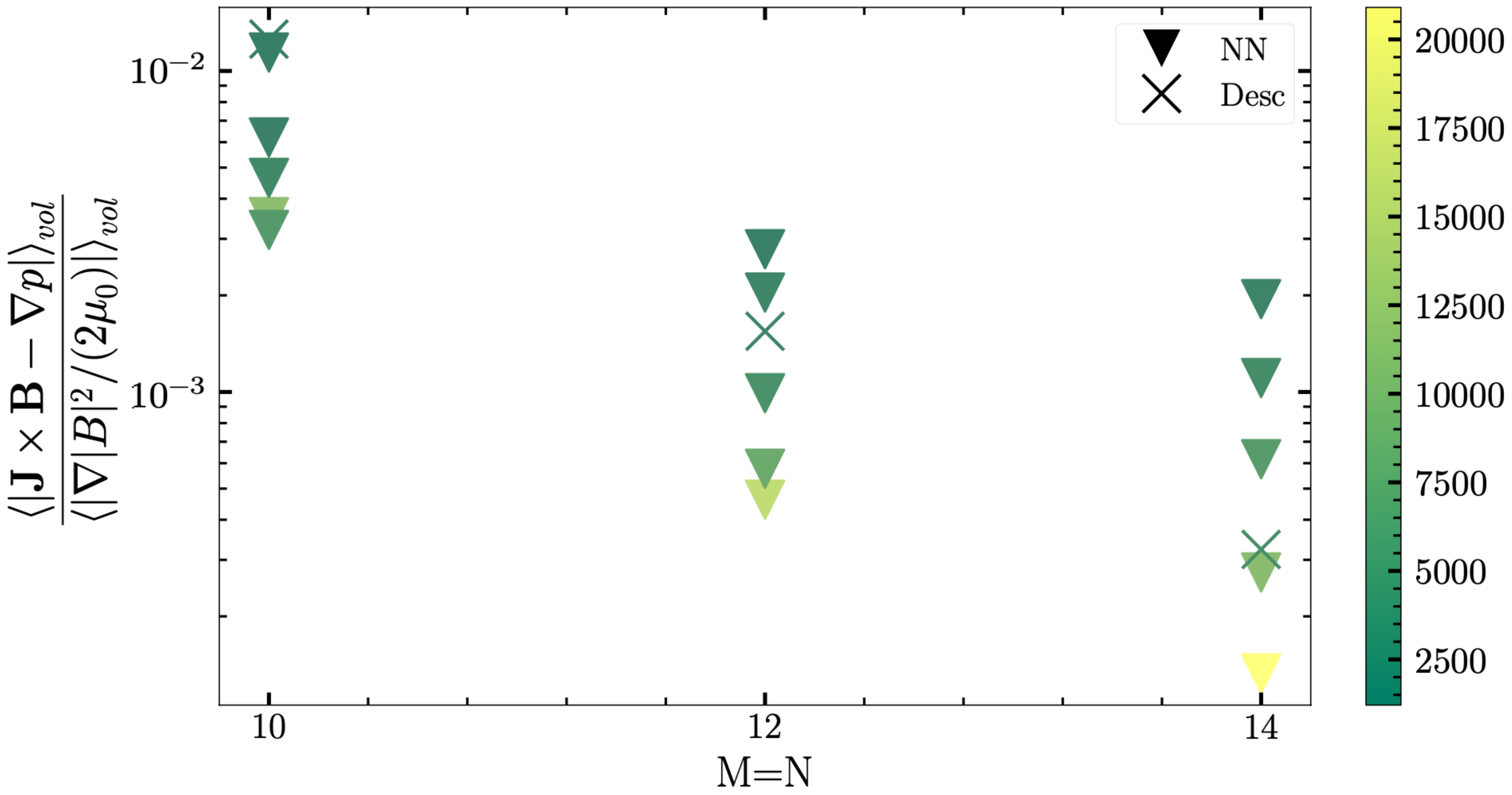




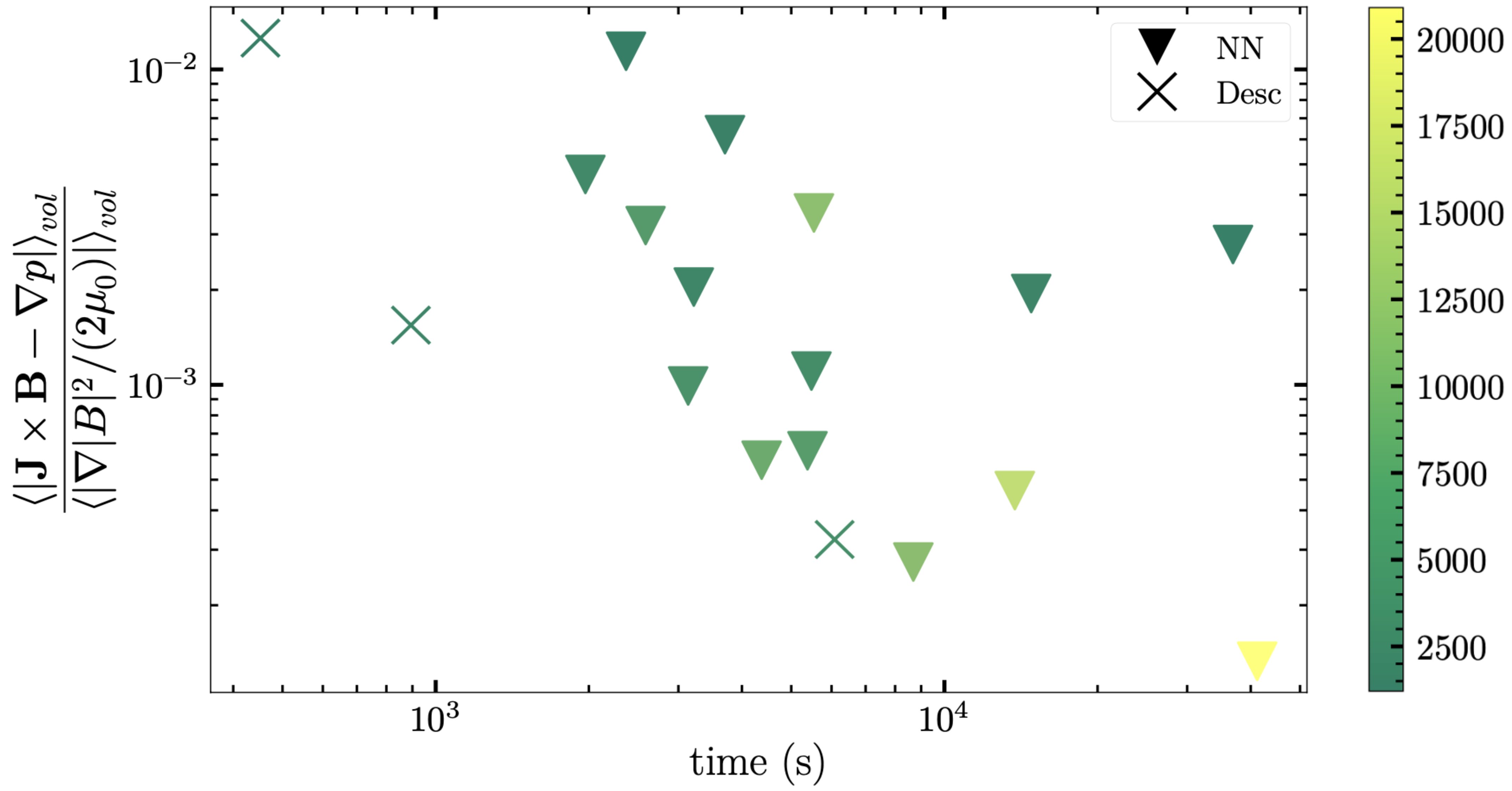
W7X M12N12 [Panici 2023] $\beta = 2.5\%$; parallel current profile



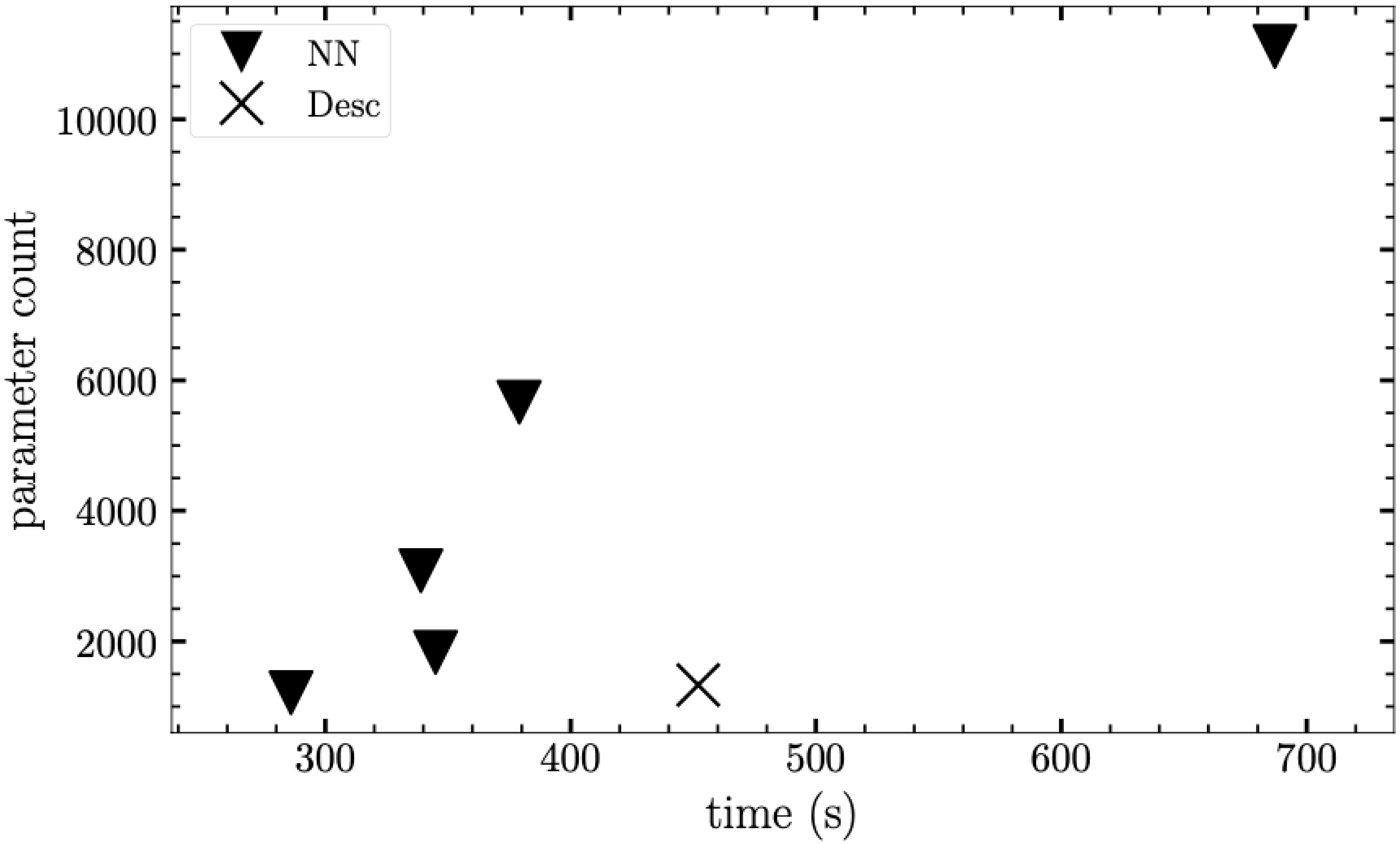
W7X M10 N10 $\beta = 2.5\%$; Scan of Fourier resolution



W7X M10 N10 $\beta = 2.5\%$; Run-time Comparison on RAVEN GPU



W7X M10 N10 $\beta = 2.5\%$; Runtime until DESC's $F_{vol,norm}$ is reached by NN



Conclusion and Outlook

- **Simple PiNNs can achieve lower $< ||\boldsymbol{F}|| >_{vol}$ than DESC/VMEC.**

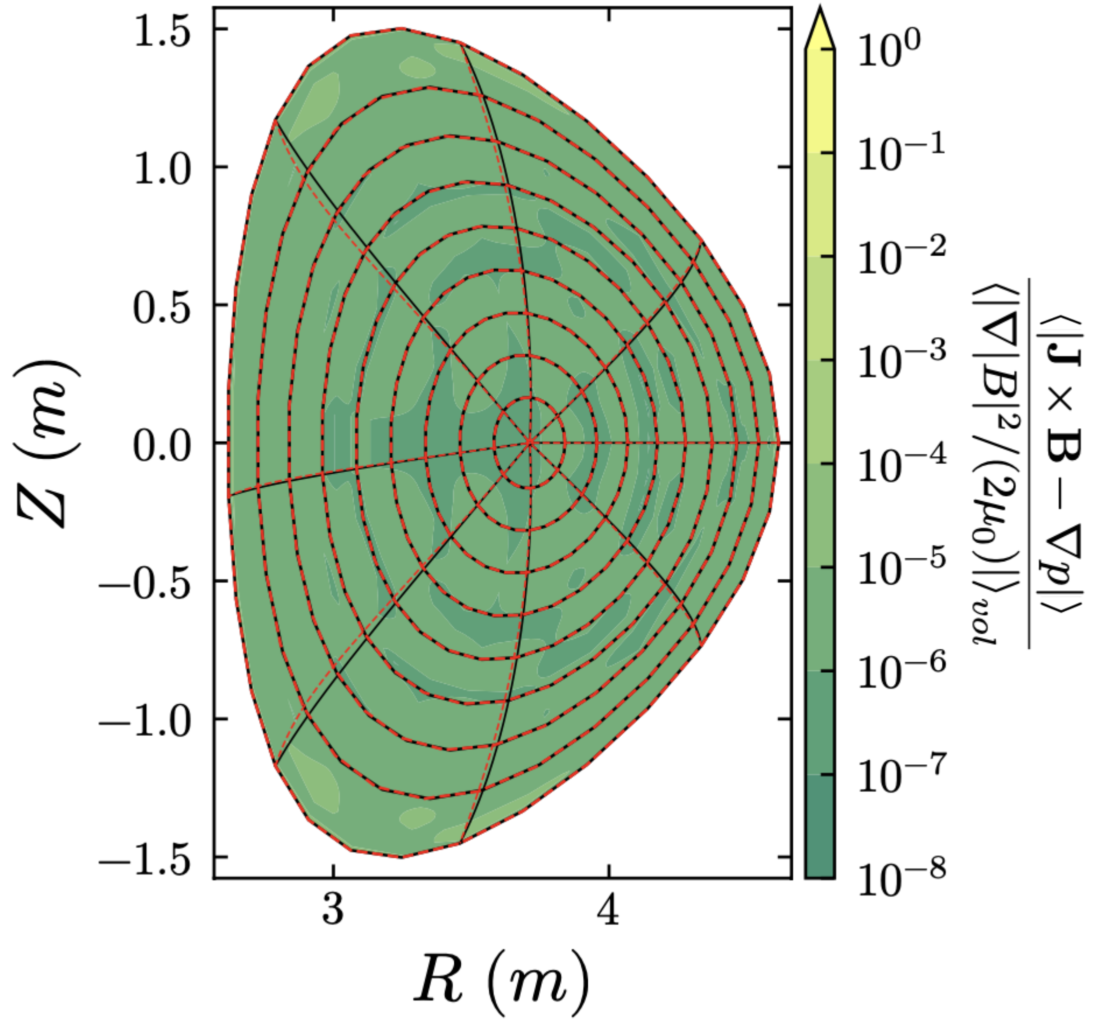
- **PiNNs can cost more resources to train, but:**
  - Residual-based Attention
  - Perturbation methods
  - Adaptive Grid-Refinement

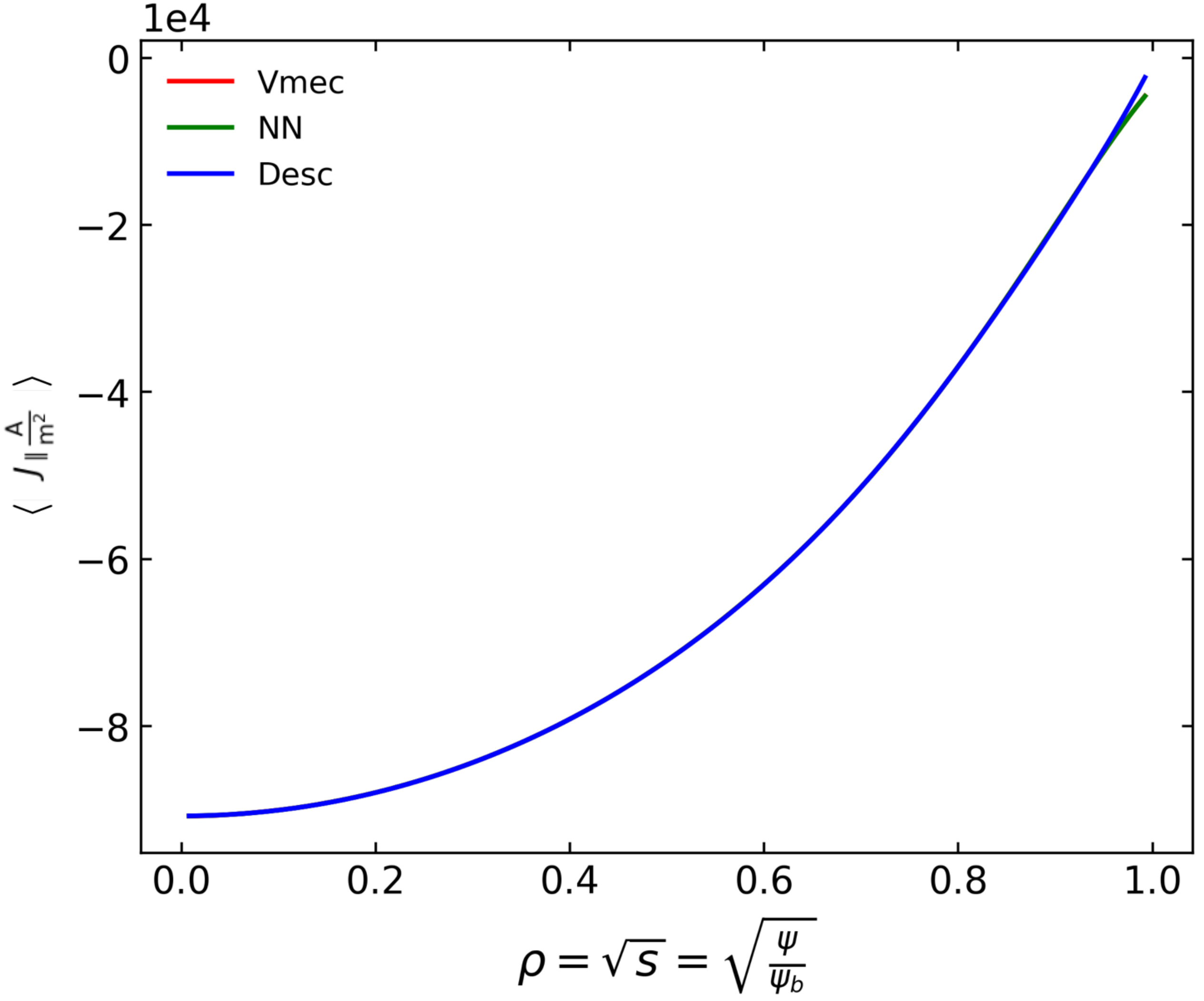
- **With enhancements, PiNNs could become more compact and cost less to train.**

- **Outlook:**
  - Operator learning → $\boldsymbol{NN}(\rho, \boldsymbol{BC})$
  - Investigate effect of lower $< ||\boldsymbol{F}|| >_{vol}$ in stability codes and stellarator optimization metrics

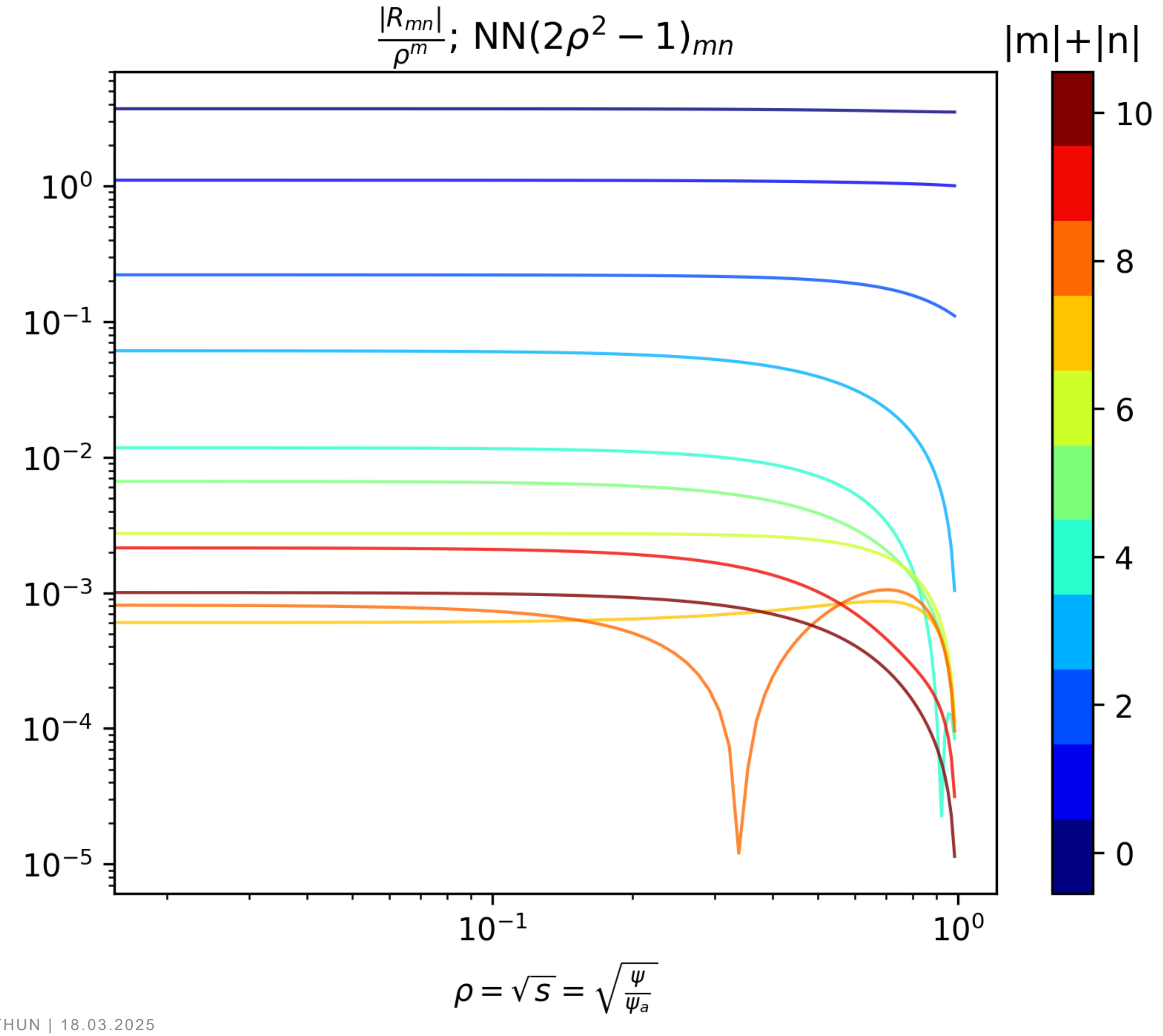
[Merlo 2023] - Physics-regularized neural network of the ideal-MHD solution operator in Wendelstein 7-X configurations
[Hirshman 1984] - Steepest descent moment method for three-dimensional magnetohydrodynamic equilibria
[Nührenberg 2003] - Magnetic islands and perturbed plasma equilibria
[Panici 2023] - The DESC stellarator code suite. Part 1. Quick and accurate equilibria computations

Appendix – Dshape solved with [8, 8] MLP's





Appendix - Dshape solved with [8, 8] MLP's, Analyticity @ axis



$\boldsymbol{NN}(\rho; \boldsymbol{BC})$

NN training on 3D MHD force-residual – implemented code

Find $\boldsymbol{X}_{mn}(\rho)$ s.t. min $||F||$

m_{pol} n_{tor} → $\mathcal{F}$ →

$\rho\ u\ v \in \mathbb{R}^{3 \times n_\rho n_u n_v}$ → $\mathcal{F}$

$p(\rho), \iota(\rho), R_{bmn}, Z_{bmn}$ →

$\rho = \sqrt{s} \in (0,1)$

$\theta \in [0, 2\pi)$

$\zeta \in [0, \frac{2\pi}{N_{FP}}]$

$$\boldsymbol{X}_{mn}(\rho) = \rho^m * [\, \boldsymbol{X}_{bmn} + (1-\rho^\alpha) * NN_X([\rho^2, \rho^4, \dots]^T)]$$

$\boldsymbol{X}_{mn} \in \{\boldsymbol{R}_{mn,c}, \boldsymbol{\lambda}_{mn,s}, \boldsymbol{Z}_{mn,s}\}$ $\quad \boldsymbol{\lambda}_{bmn} = \boldsymbol{0}$

$$NN_X([\rho^2, \rho^4, \dots]^T) = \boldsymbol{W}_{X,1} * \tanh\left(\boldsymbol{W}_{X,0} * \begin{bmatrix} \rho^2 \\ \rho^4 \\ \vdots \end{bmatrix} + \boldsymbol{b}_{X,0}\right) + \boldsymbol{b}_{X,1}$$

→ $\mathcal{F}^{-1}$ ↓

$R\ \lambda\ Z \in \mathbb{R}^{3 \times n_\rho n_u n_v}$

↓ AD

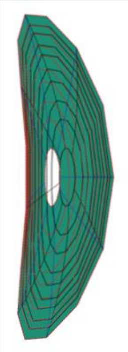
$$\frac{\partial(R, \lambda, Z)}{\partial(\rho, u, v)}$$

$$\frac{\partial^2(R, \lambda, Z)}{\partial(\rho^2, u^2, v^2)}$$

←

$$\frac{\partial \iiint ||J \times B - \nabla p||_2 \sqrt{g}\, d\rho\, du\, dv}{\partial(NN_X)}$$

↑ **Optimizer** ↑



VMEC-like Initial condition

W7-X standard configuration

Black: VMEC
Red+Blue: NN