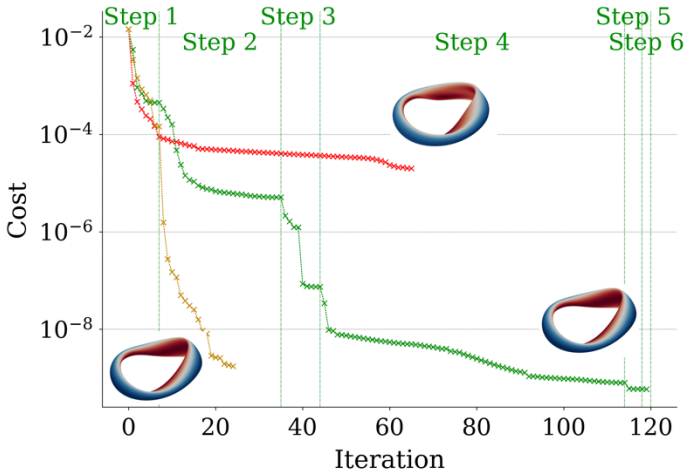
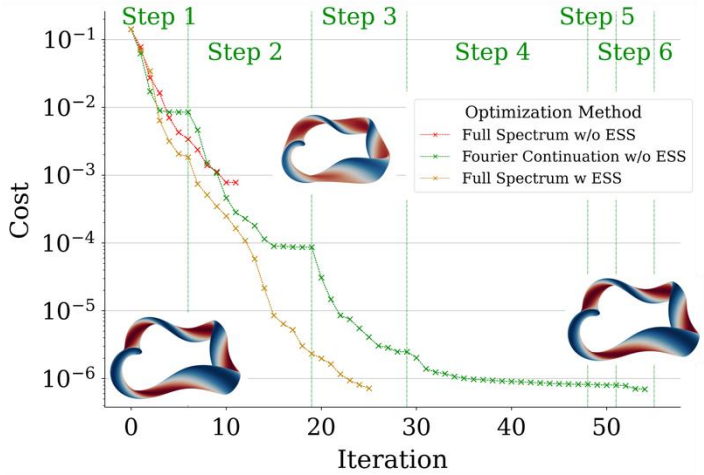
Exponential Spectral Scaling: Robust and Efficient Stellarator Boundary Optimization via Mode-Dependent Scaling

Byoungchan Jang[1], Rory Conlin[1], Matt Landreman[1]
[1]*University of Maryland, College Park, MD, USA*



Simons/HiFiStell Hour
12/18/2025



U.S. DEPARTMENT OF ENERGY | Office of Science

Jang *et al.*, Journal of Plasma Physics (2025)
arXiv:2509.16320

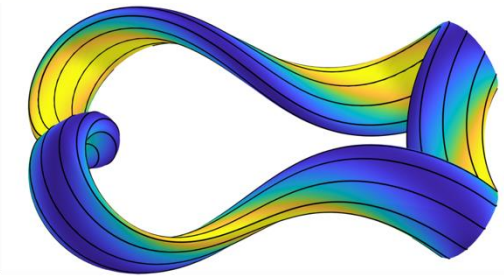
Motivation

- Stellarators: 3D magnetic configurations enabling steady-state fusion.
- In stellarator optimization, gradient-based algorithms (optimizers) can get stuck in local minima, and the solution depends on the initial condition.
- Stellarator optimization pipeline:

Motivation

- Stellarators: 3D magnetic configurations enabling steady-state fusion.
- In stellarator optimization, gradient-based algorithms (optimizers) can get stuck in local minima, and the solution depends on the initial condition.
- Stellarator optimization pipeline:

  Stage 1: optimize plasma boundary

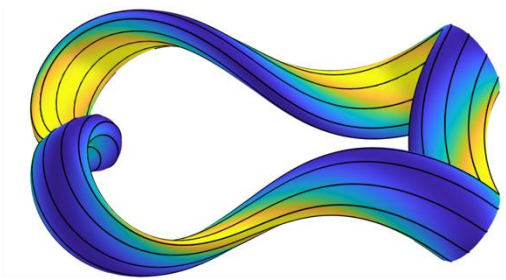


Landreman, Paul (2021)

Motivation

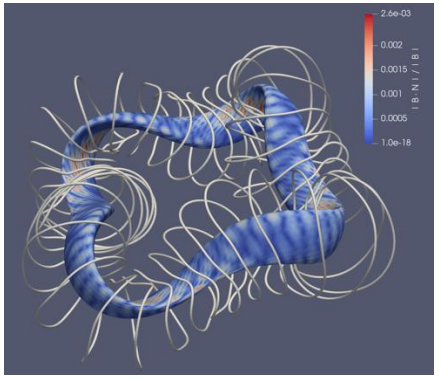
- Stellarators: 3D magnetic configurations enabling steady-state fusion.
- In stellarator optimization, gradient-based algorithms (optimizers) can get stuck in local minima, and the solution depends on the initial condition.
- Stellarator optimization pipeline:

Stage 1: optimize plasma boundary



Landreman, Paul (2021)

Stage 2: design coils to realize that boundary

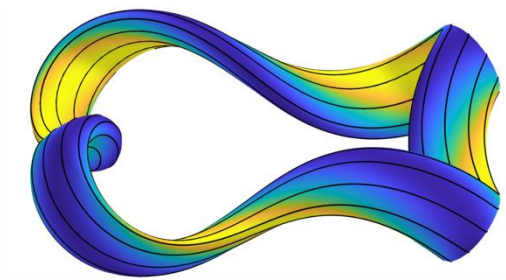


Wiedman *et al.* (2024)

Motivation

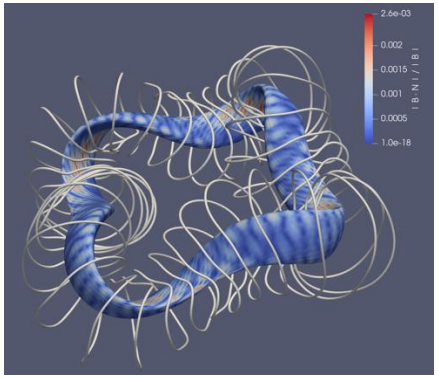
- Stellarators: 3D magnetic configurations enabling steady-state fusion.
- In stellarator optimization, gradient-based algorithms (optimizers) can get stuck in local minima, and the solution depends on the initial condition.
- Stellarator optimization pipeline:

Stage 1: optimize plasma boundary



Landreman, Paul (2021)

Stage 2: design coils to realize that boundary



Wiedman *et al.* (2024)

This talk: **stage-1 boundary optimization**, and how to make it more robust and efficient

Stage-1 Stellarator Optimization

Given a "cost function" $f: \mathbb{R}^n \to \mathbb{R}$, minimize $f(\boldsymbol{x})$

(aka "loss function", "objective function")

Stage-1 Stellarator Optimization

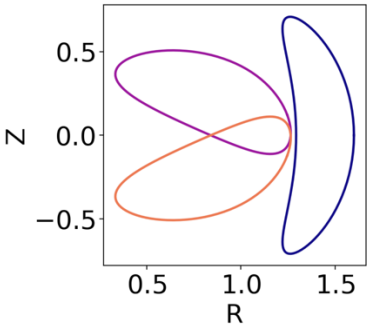
Given a "cost function" $f: \mathbb{R}^n \to \mathbb{R}$, minimize $f(\boldsymbol{x})$

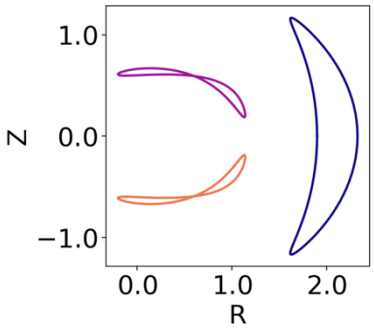
(aka "loss function", "objective function")

Parameter space: $\boldsymbol{x}$

$\boldsymbol{x}$ = shape of toroidal surface (~100 dofs)

- Part of the parameter space corresponds to unphysical self-intersecting shapes





Stage-1 Stellarator Optimization

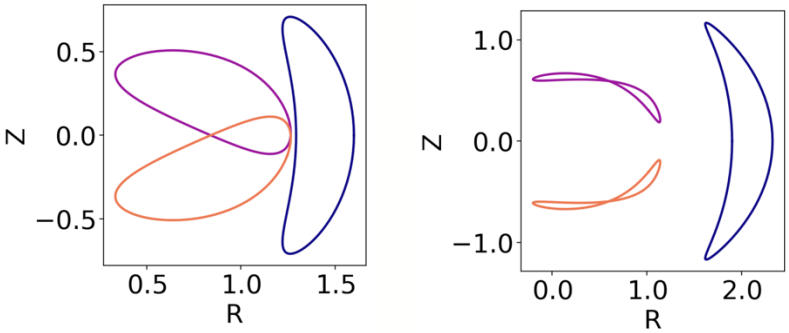
Given a "cost function" $f: \mathbb{R}^n \to \mathbb{R}$, minimize $f(\boldsymbol{x})$

(aka "loss function", "objective function")

Parameter space: $\boldsymbol{x}$

$\boldsymbol{x}$ = shape of toroidal surface (~100 dofs)

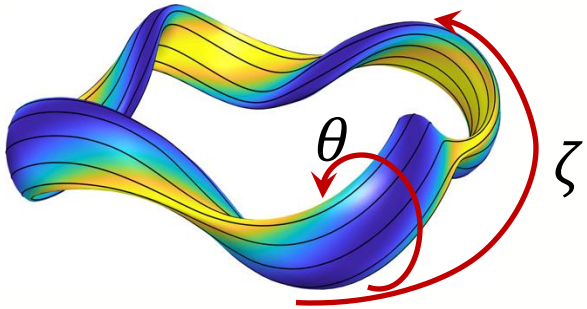
- Part of the parameter space corresponds to unphysical self-intersecting shapes



Objective functions: $\boldsymbol{f}$

- Large volume of good magnetic surfaces (no islands and chaos)
- Rotational transform (i.e., inverse safety factor)
- Good confinement of particle trajectories
- Low neoclassical transport
- Low turbulent transport
- Magnetohydrodynamic (MHD) stability
- Coil Buildability ($L_{\nabla B}$, etc.)

Parameter space

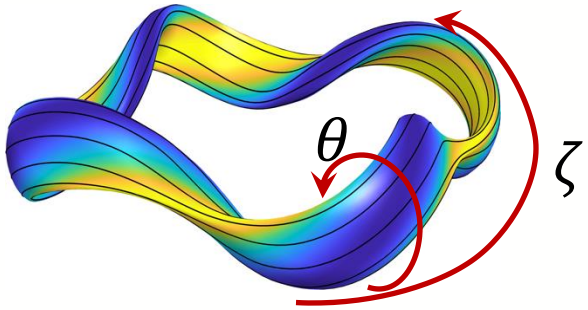


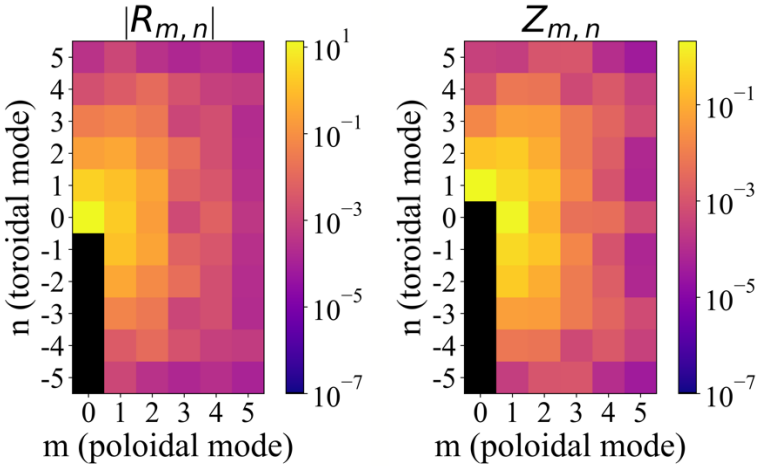
Parametrization of boundary(i.e. LCFS) surface:

$$R(\theta,\zeta) = \sum_{m,n} R_{m,n}\cos(m\theta - n_{fp}n\zeta) \qquad Z(\theta,\zeta) = \sum_{m,n} Z_{m,n}\sin(m\theta - n_{fp}n\zeta)$$

ζ = toroidal angle, θ = poloidal angle, n_{fp} = number of field periods

Parameter space



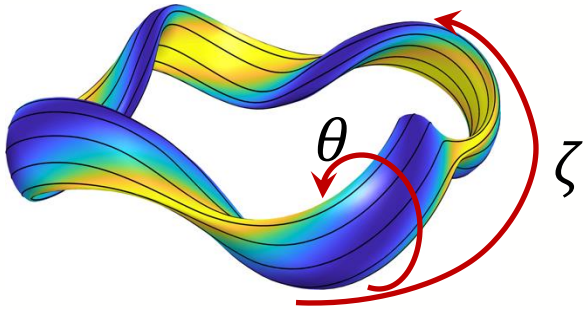


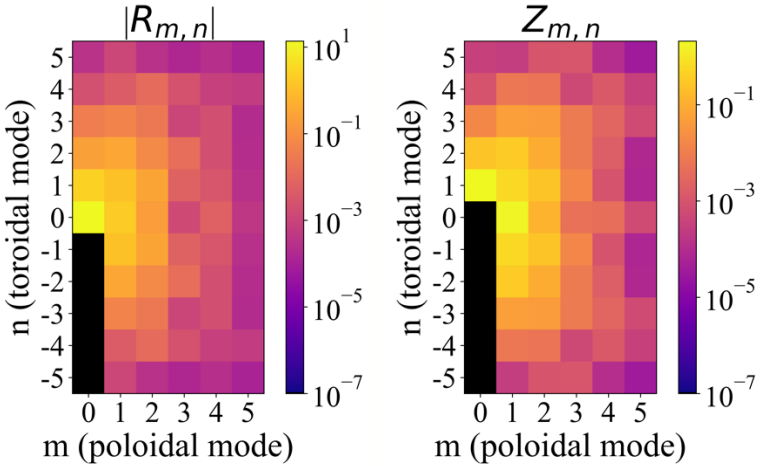
Parametrization of boundary(i.e. LCFS) surface:

$$R(\theta,\zeta) = \sum_{m,n} R_{m,n}\cos(m\theta - n_{fp}n\zeta) \qquad Z(\theta,\zeta) = \sum_{m,n} Z_{m,n}\sin(m\theta - n_{fp}n\zeta)$$

ζ = toroidal angle, θ = poloidal angle, n_{fp} = number of field periods

Parameter space





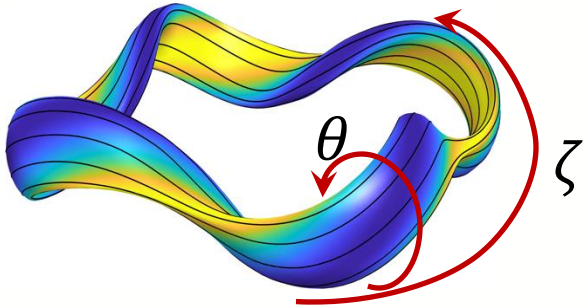
Parametrization of boundary(i.e. LCFS) surface:

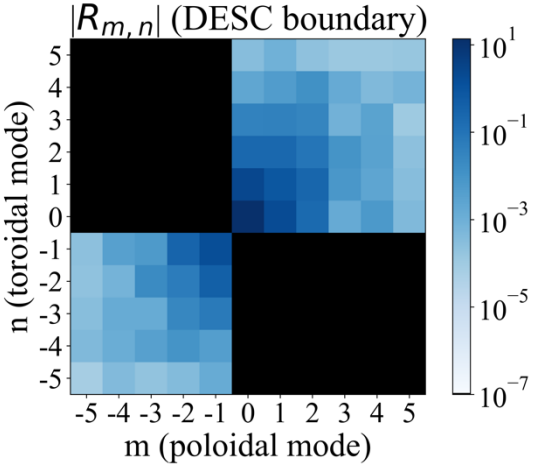
$$R(\theta,\zeta) = \sum_{m,n} R_{m,n}\cos(m\theta - n_{fp}n\zeta) \qquad Z(\theta,\zeta) = \sum_{m,n} Z_{m,n}\sin(m\theta - n_{fp}n\zeta)$$

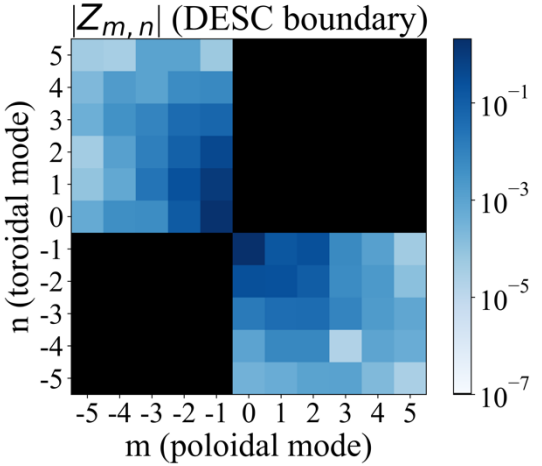
ζ = toroidal angle, θ = poloidal angle, n_{fp} = number of field periods

Parameter space for optimization: $x = [R_{m,n}, Z_{m,n}]$

Parameter space







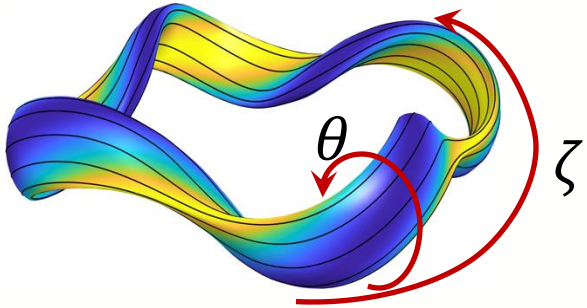
DESC Analogue

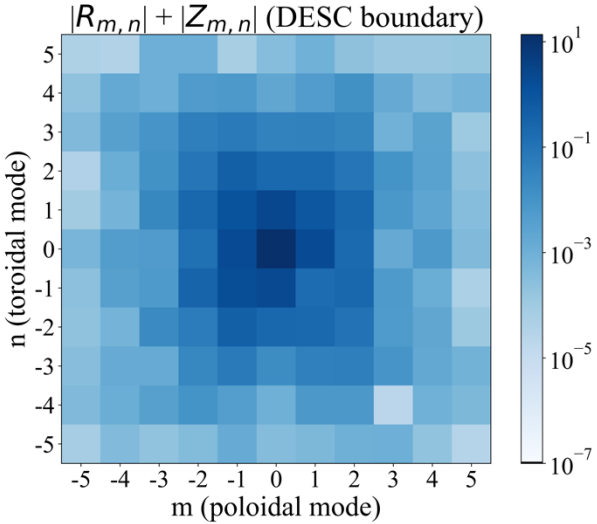
$$\cos(m\theta - n_{fp}n\zeta) = \left[\cos(m\theta)\cos(n_{fp}n\zeta) + \sin(m\theta)\sin(n_{fp}n\zeta)\right]$$

$$\sin(m\theta - n_{fp}n\zeta) = \left[\sin(m\theta)\cos(n_{fp}n\zeta) - \cos(m\theta)\sin(n_{fp}n\zeta)\right]$$

Parameter space for optimization: $x = \left[R_{m,n}, Z_{m,n}\right]$

Parameter space





DESC Analogue

$$\cos(m\theta - n_{fp}n\zeta) = \left[\cos(m\theta)\cos(n_{fp}n\zeta) + \sin(m\theta)\sin(n_{fp}n\zeta)\right]$$

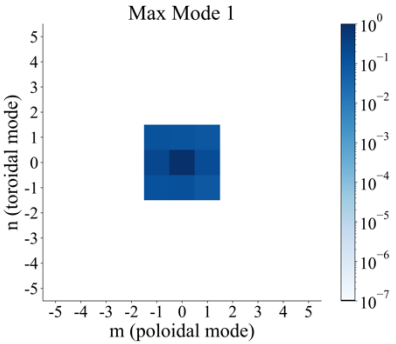
$$\sin(m\theta - n_{fp}n\zeta) = \left[\sin(m\theta)\cos(n_{fp}n\zeta) - \cos(m\theta)\sin(n_{fp}n\zeta)\right]$$

Parameter space for optimization: $x = \left[R_{m,n}, Z_{m,n}\right]$

Problem

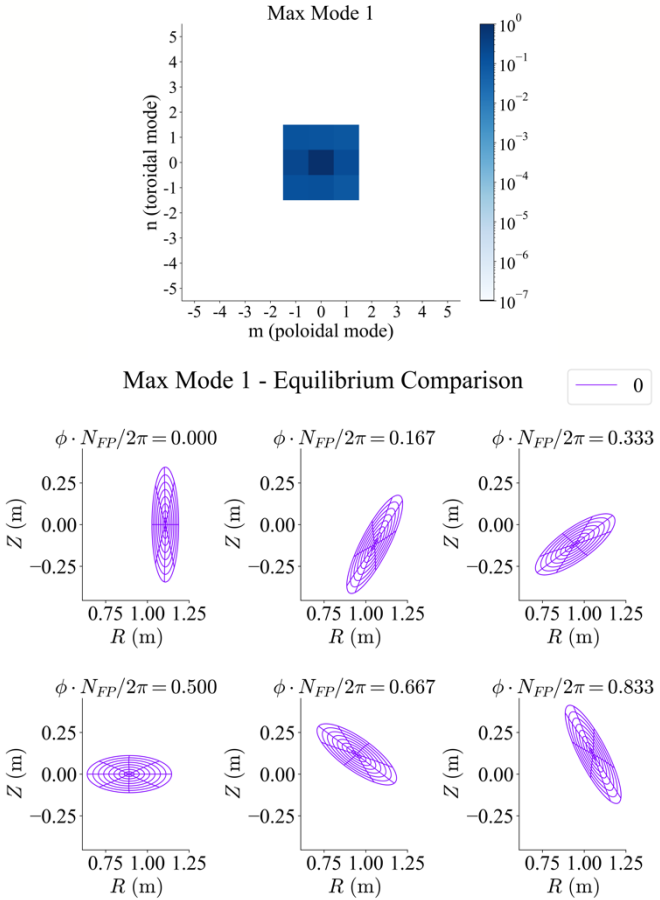
Problem: Full Spectrum Optimization Stalls

Problem: Full Spectrum Optimization Stalls

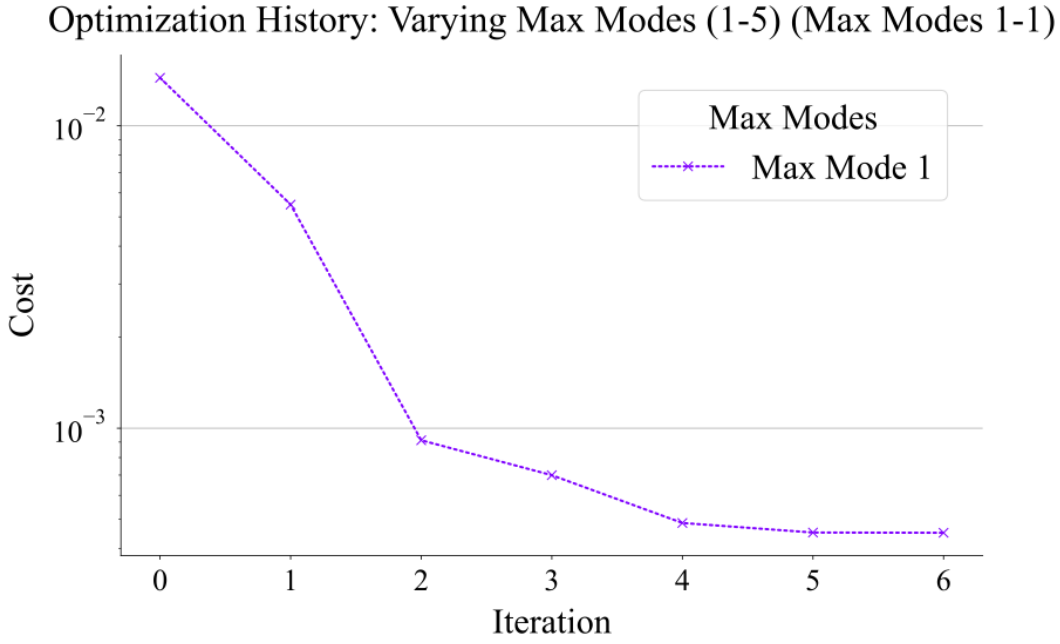


$$\text{vac QA} \quad \min_{x} f(x)$$

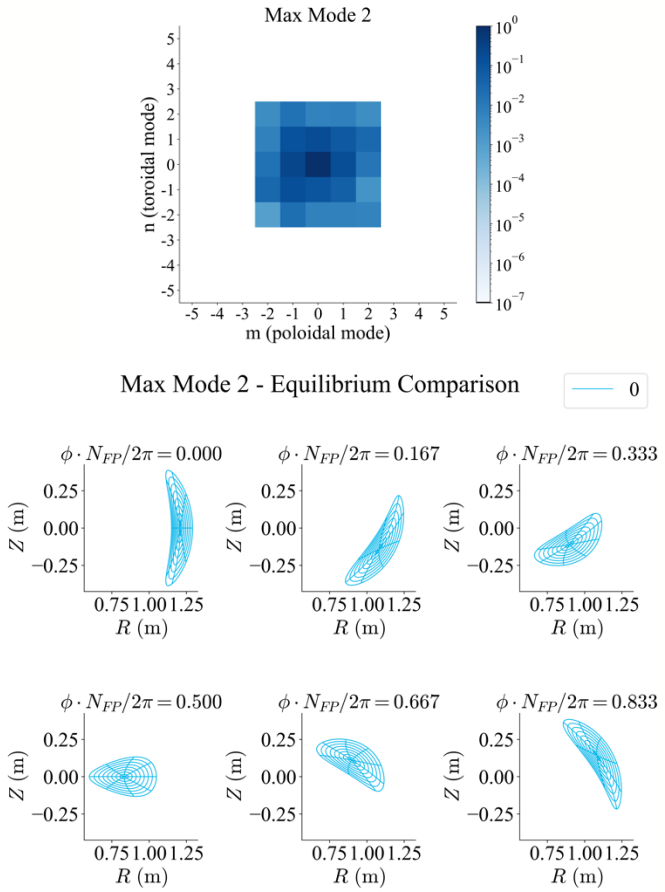
Problem: Full Spectrum Optimization Stalls



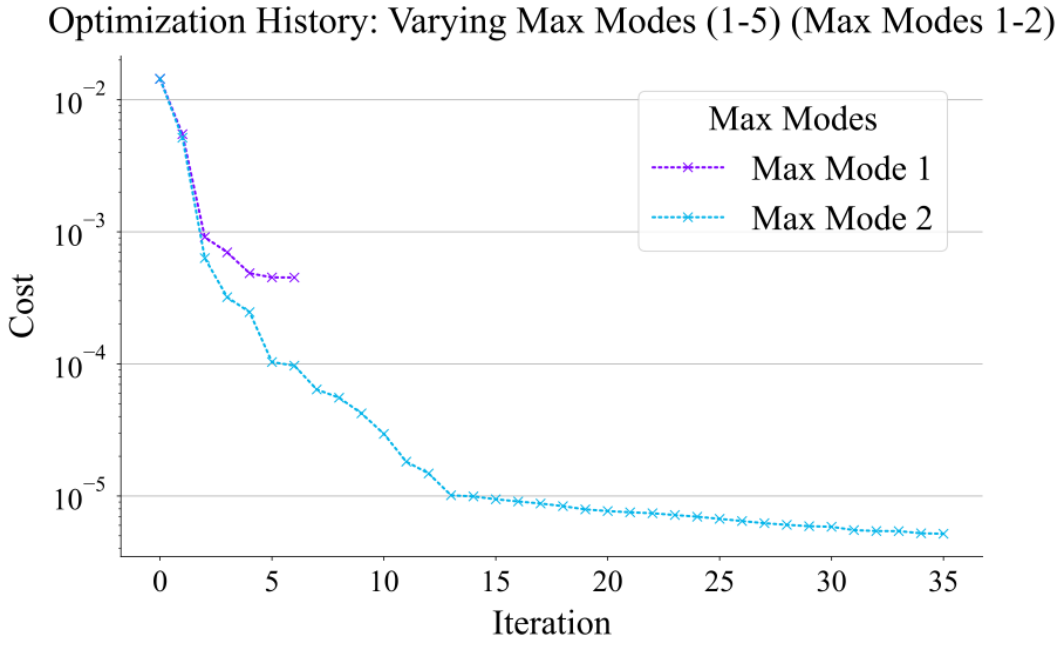
$$\text{vac QA} \quad \min_{x} f(x)$$



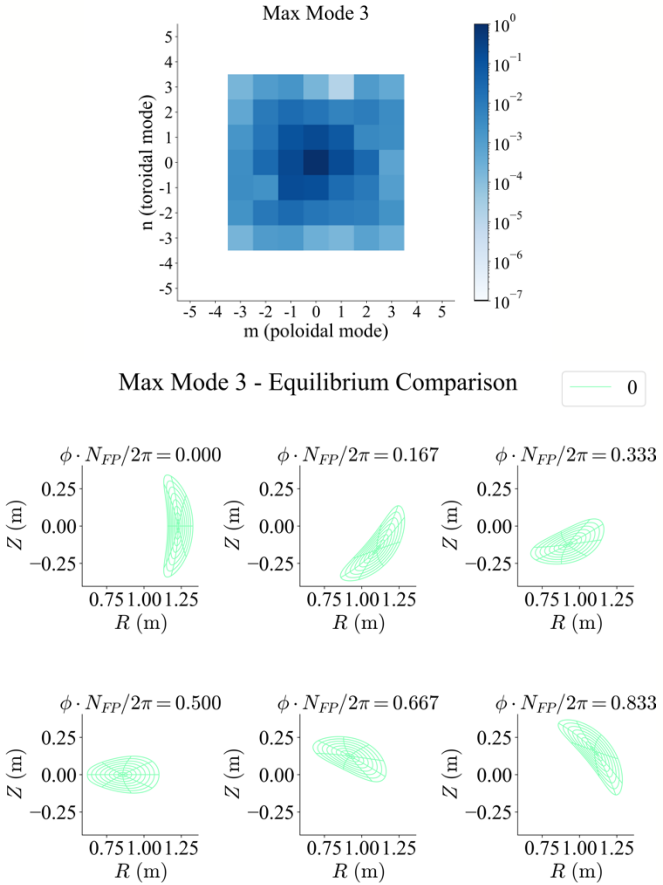
Problem: Full Spectrum Optimization Stalls



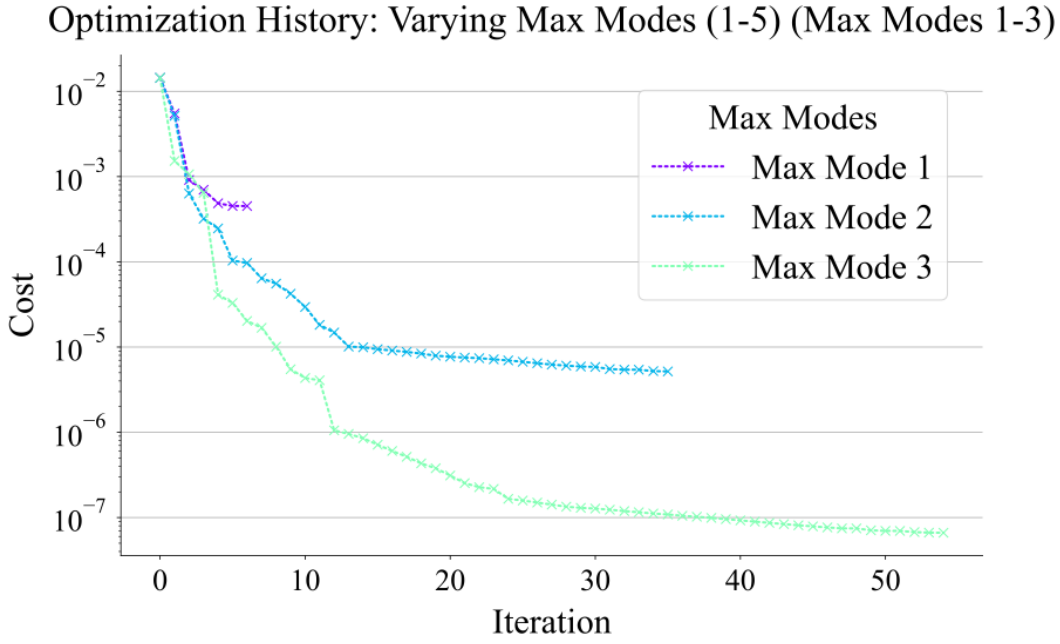
vac QA $\min_x f(x)$



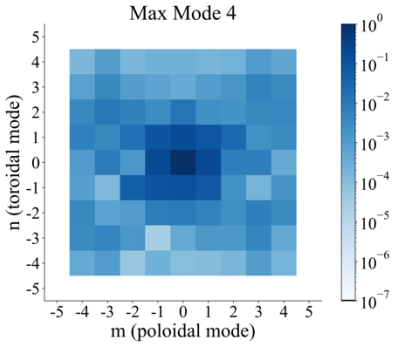
Problem: Full Spectrum Optimization Stalls

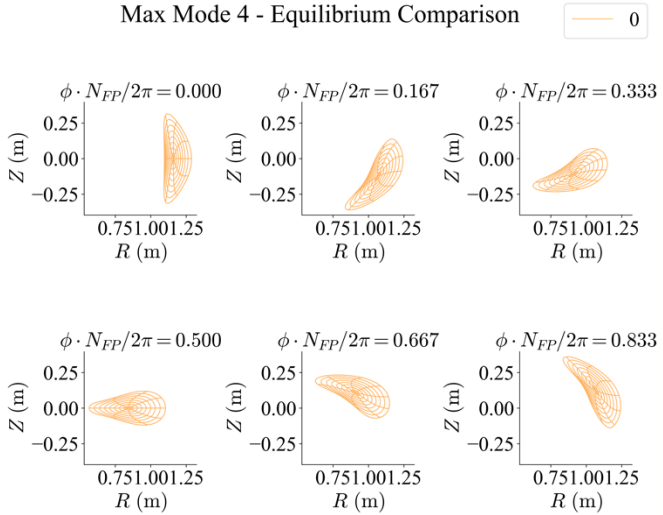


$$\text{vac QA} \quad \min_{x} f(x)$$

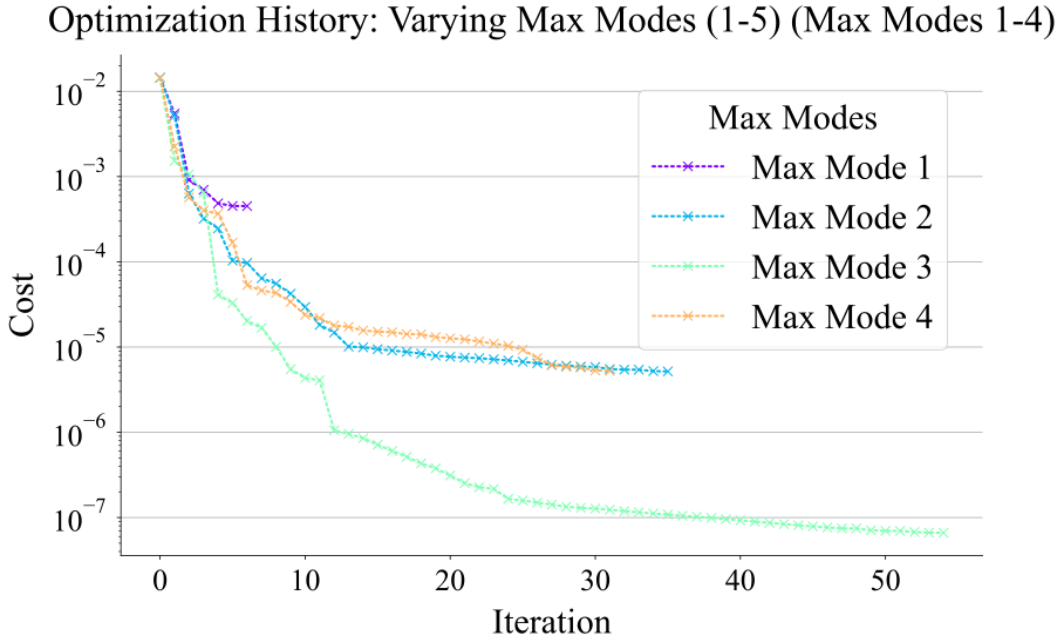


Problem: Full Spectrum Optimization Stalls

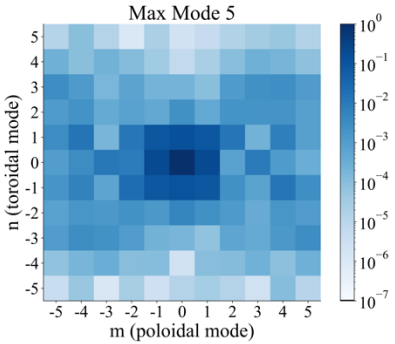


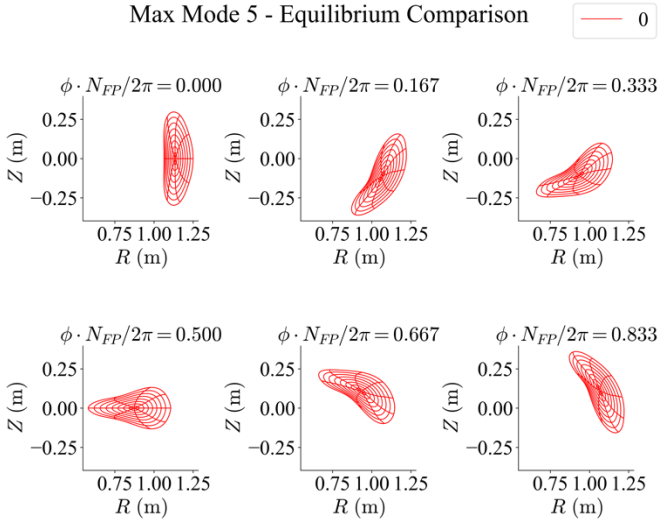


vac QA $\min_x f(x)$

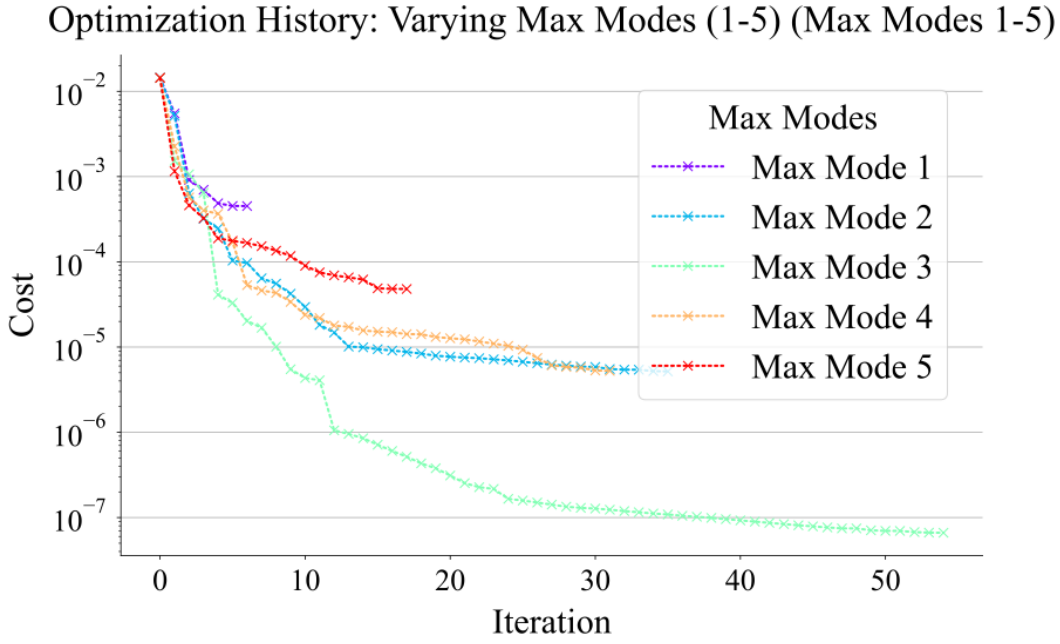


Problem: Full Spectrum Optimization Stalls



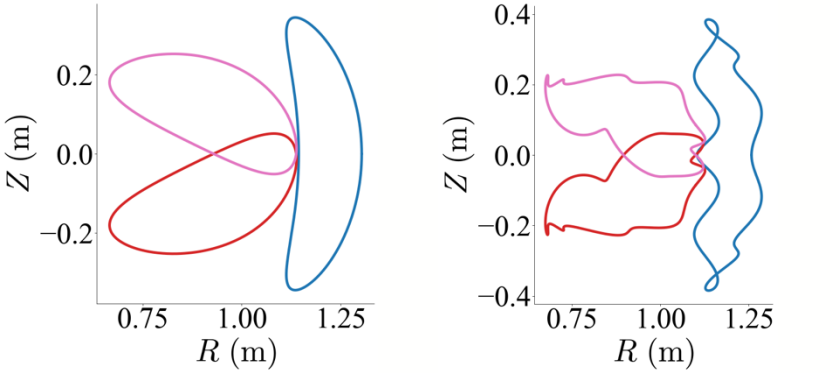


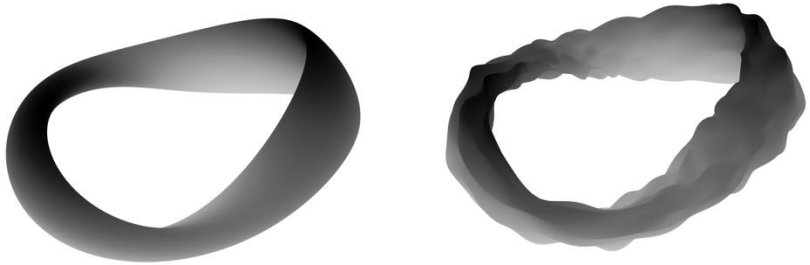
vac QA $\min_x f(x)$



Why does this happen?

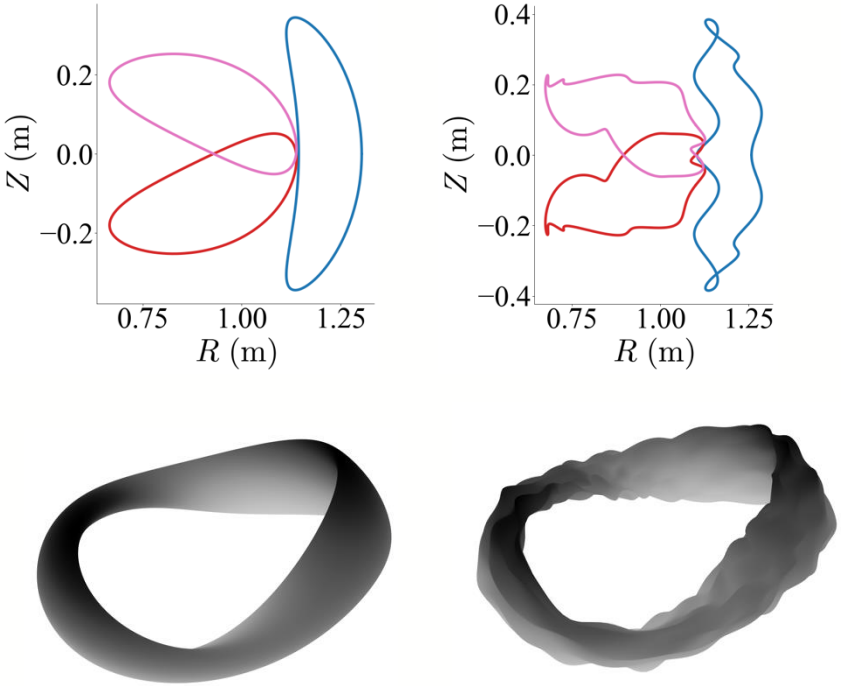
- Part of the parameter space corresponds to unphysical self-intersecting shapes





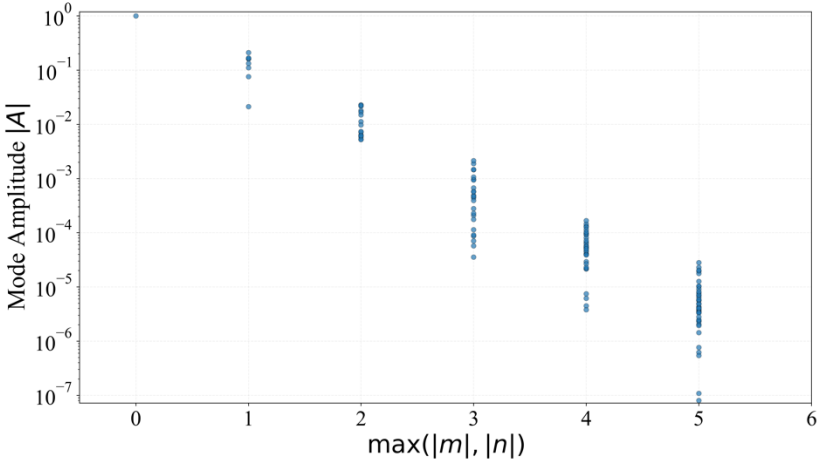
Why does this happen?

- Part of the parameter space corresponds to unphysical self-intersecting shapes



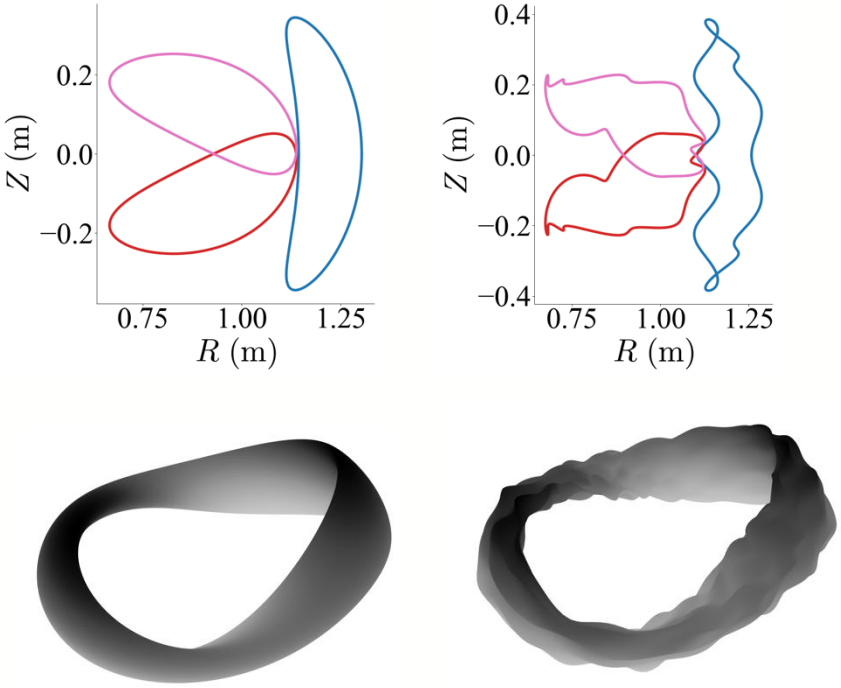
- Extreme Mode Amplitude Disparity

$$\mathrm{MAD} = \frac{\max_{(m,n)} |A_{m,n}|}{\min_{(m,n)\,:\,A_{m,n}\neq 0} |A_{m,n}|}$$



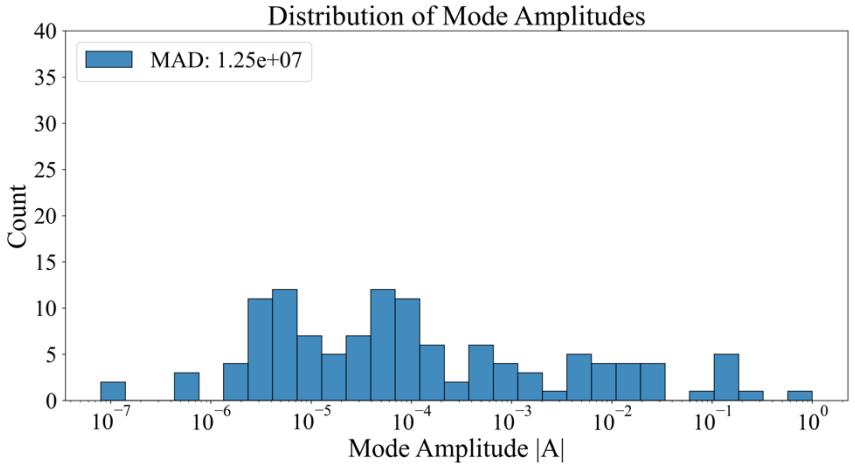
Why does this happen?

- Part of the parameter space corresponds to unphysical self-intersecting shapes



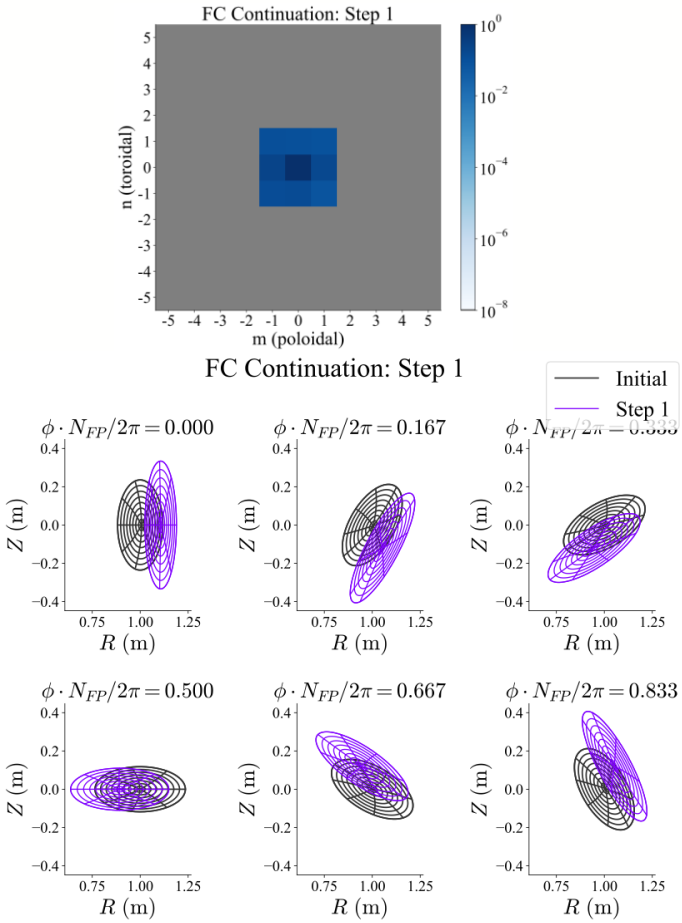
- Extreme Mode Amplitude Disparity

$$\mathrm{MAD} = \frac{\max_{(m,n)} |A_{m,n}|}{\min_{(m,n)\,:\,A_{m,n}\neq 0} |A_{m,n}|}$$

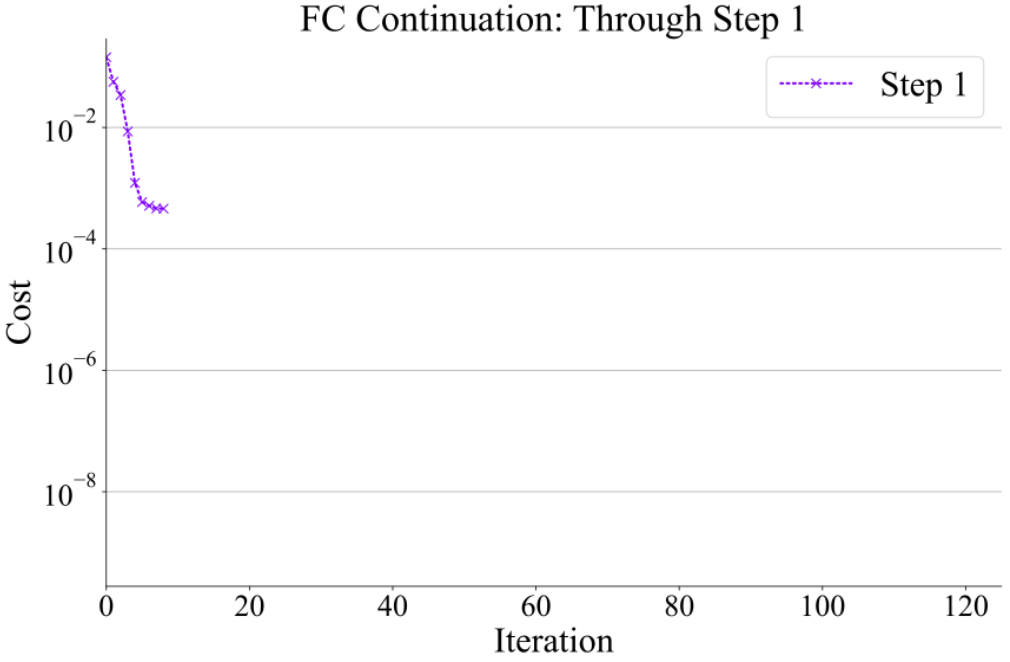


Fourier Continuation as a Workaround

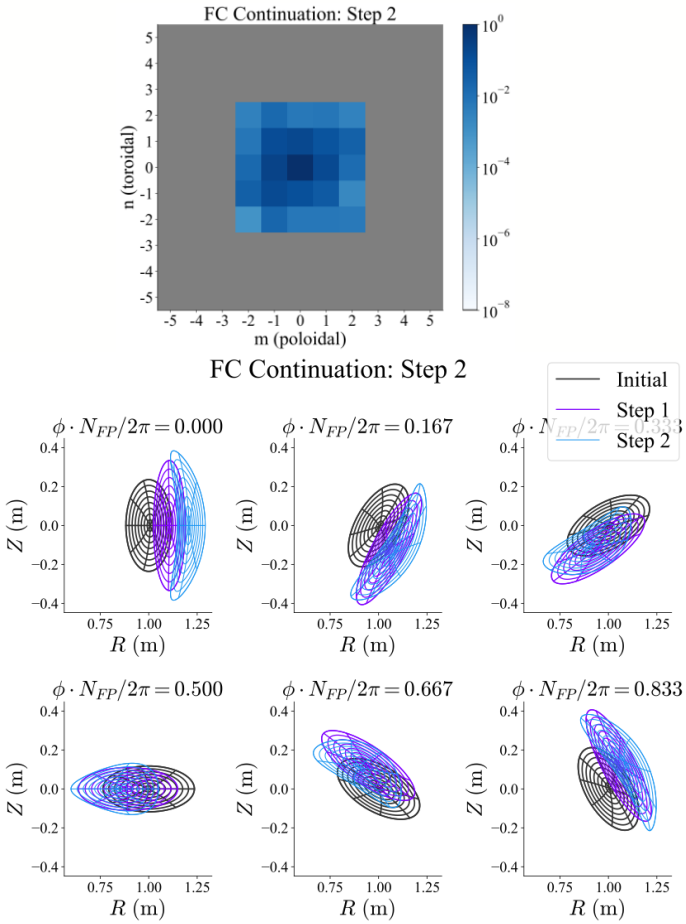
Fourier Continuation as a Workaround



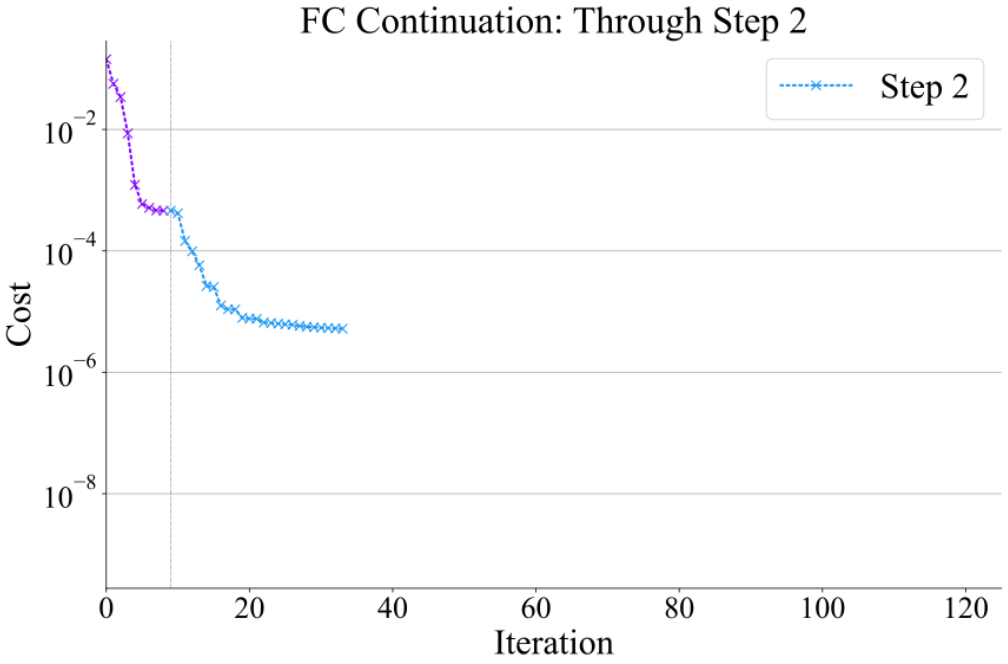
$$\text{vac QA} \quad \min_{x} f(x)$$



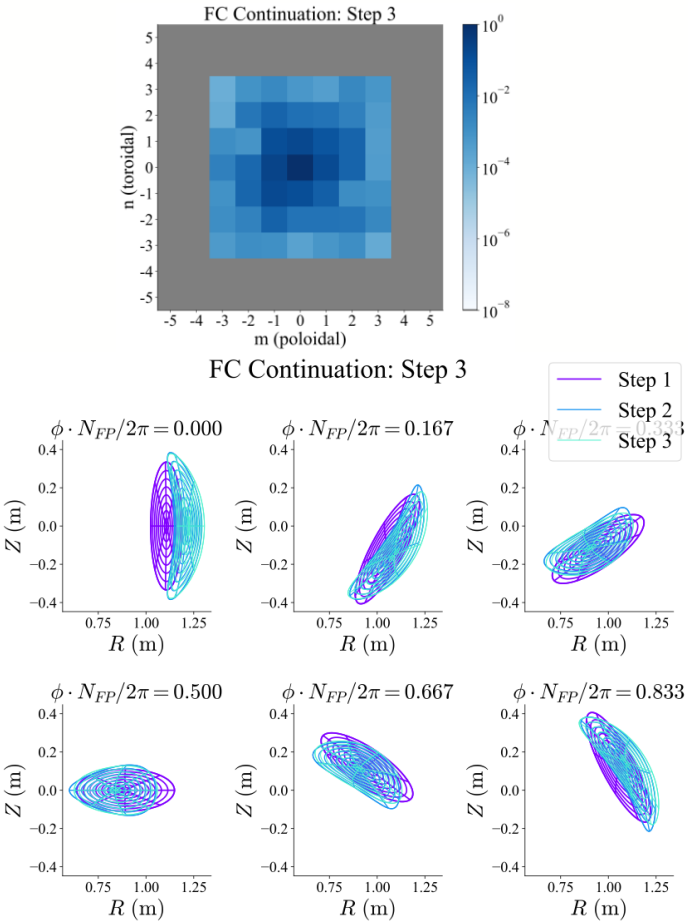
Fourier Continuation as a Workaround



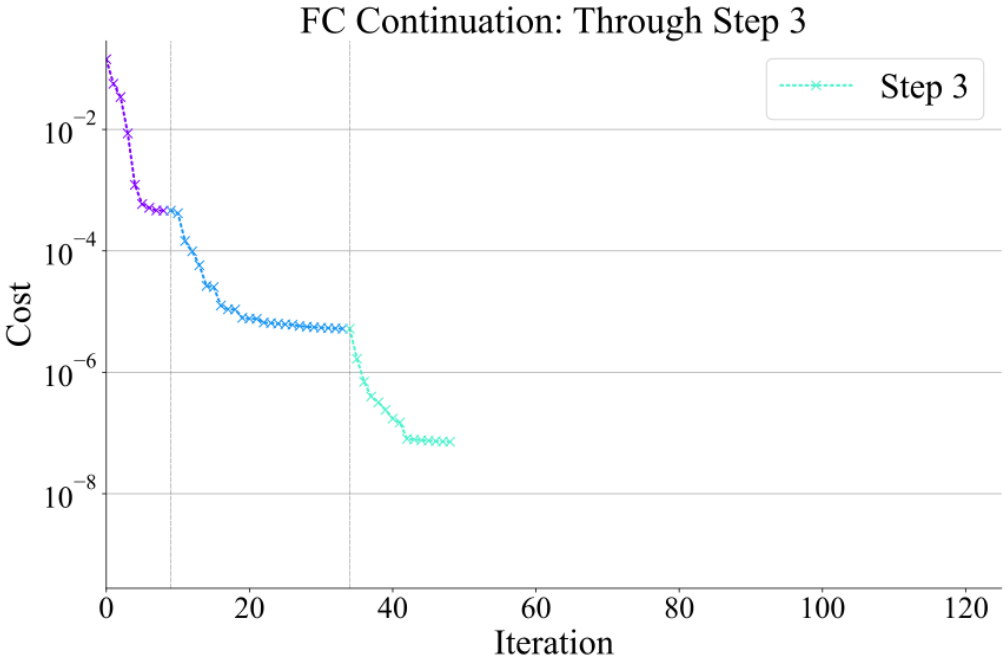
$$\text{vac QA} \quad \min_{x} f(x)$$



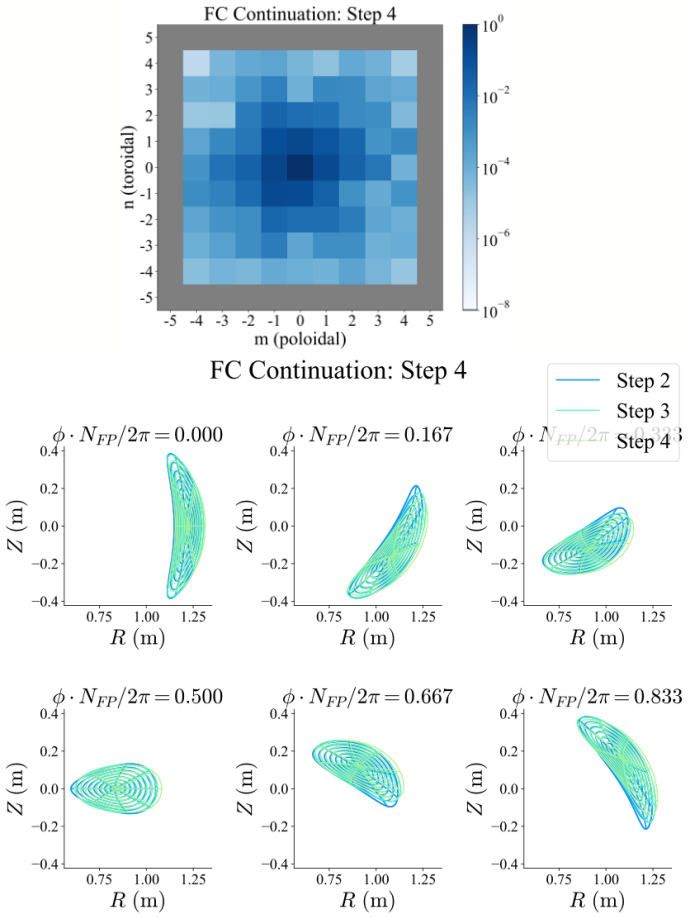
Fourier Continuation as a Workaround



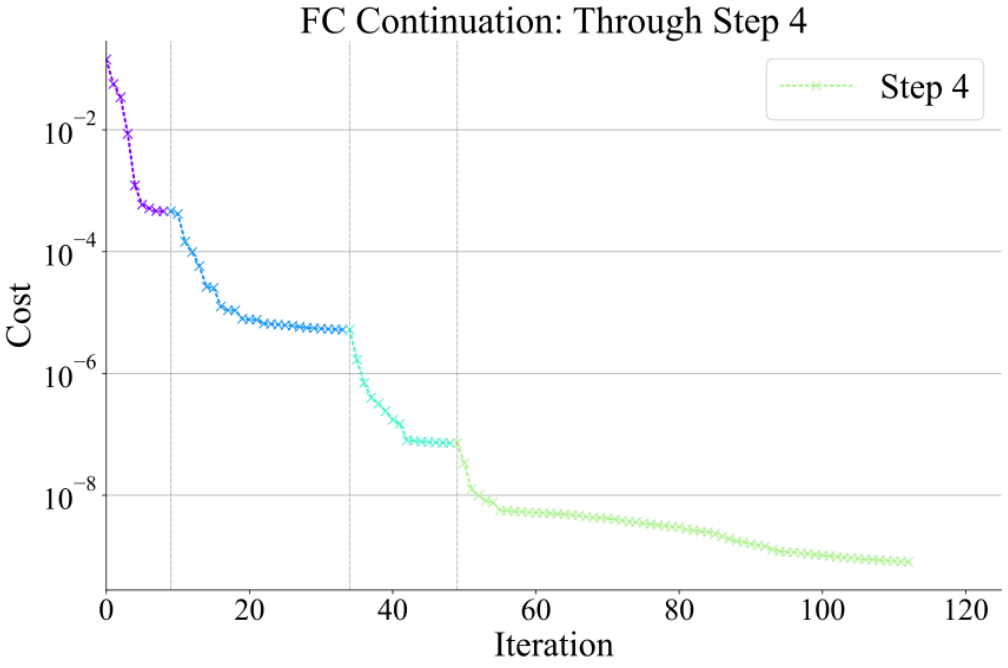
vac QA $\min_x f(x)$



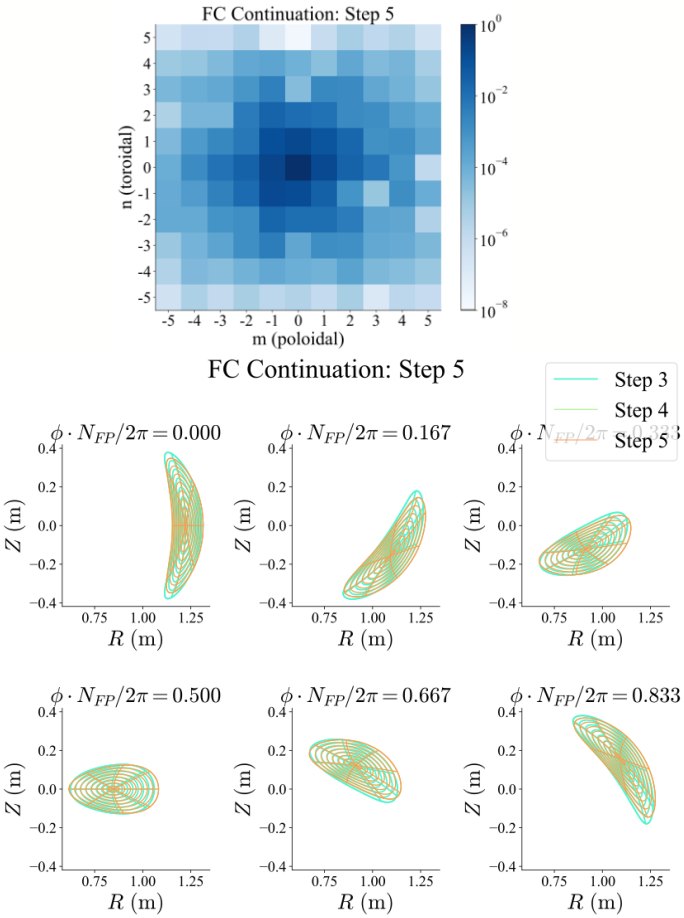
Fourier Continuation as a Workaround



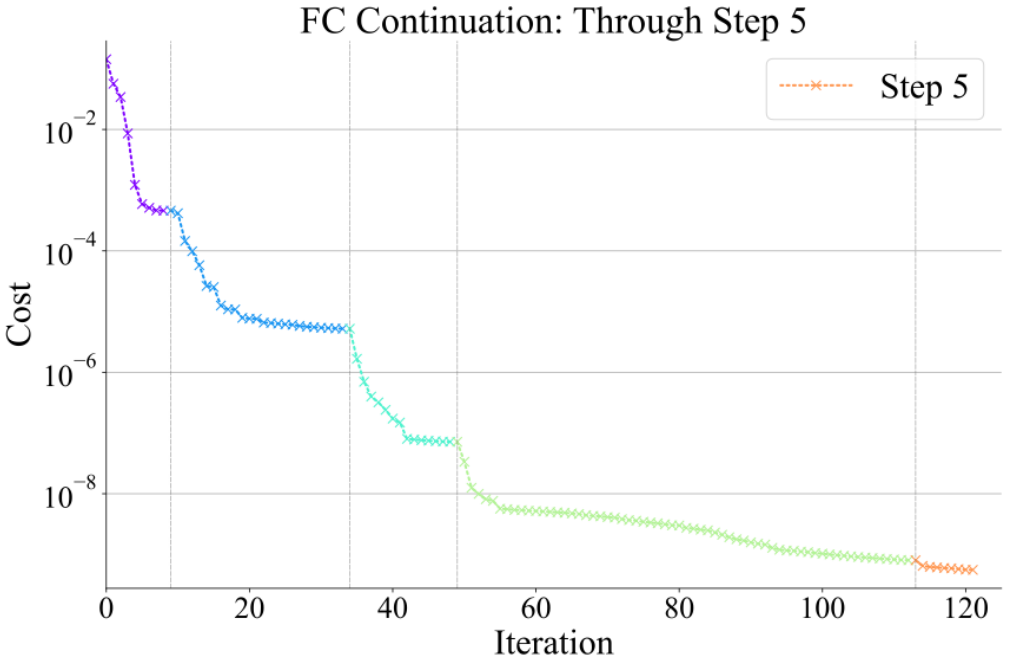
$$\text{vac QA} \quad \min_{x} f(x)$$



Fourier Continuation as a Workaround



vac QA $\min_x f(x)$

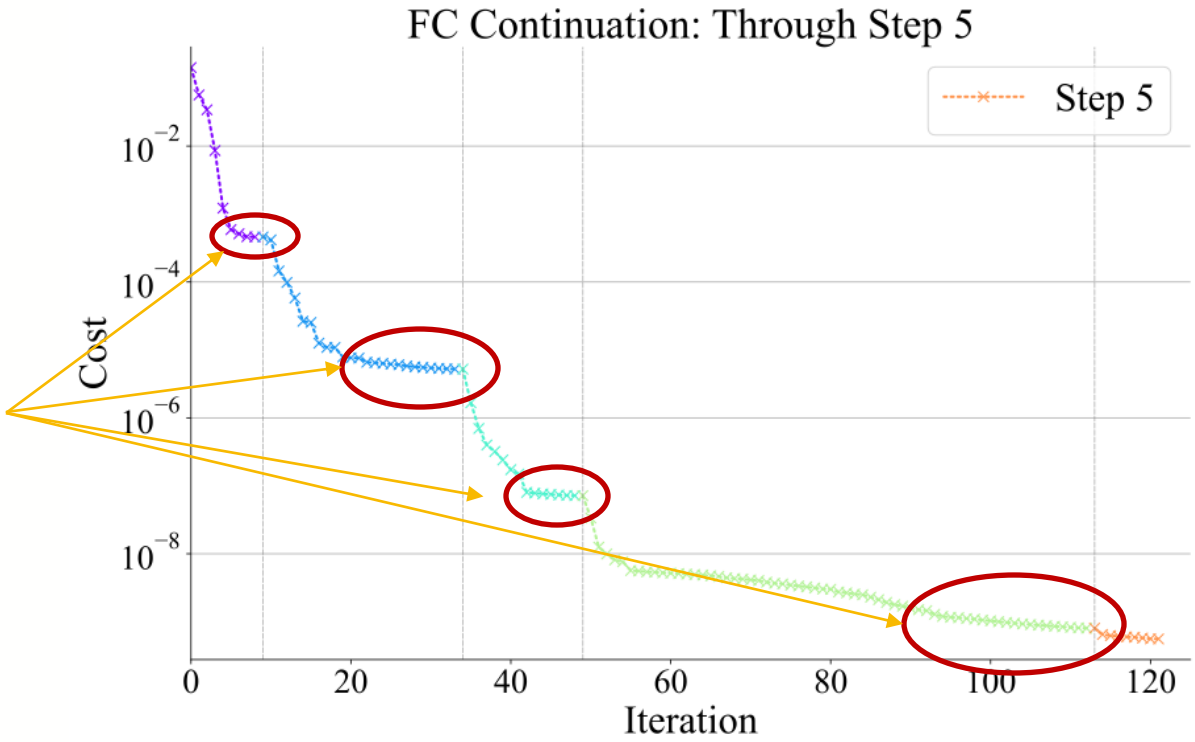


Is there a better way?

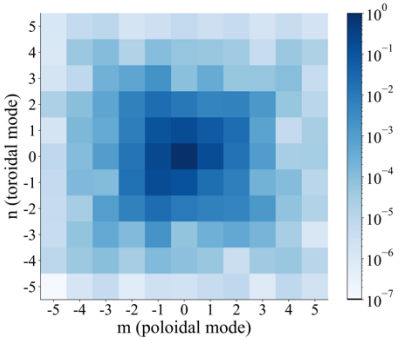
Over-solving => Inefficiency

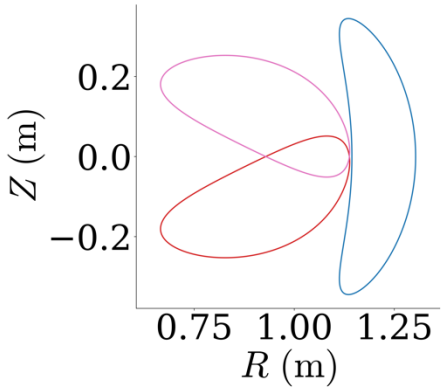
smooth solution

decrease MAD



Smoothness and Fourier Decay

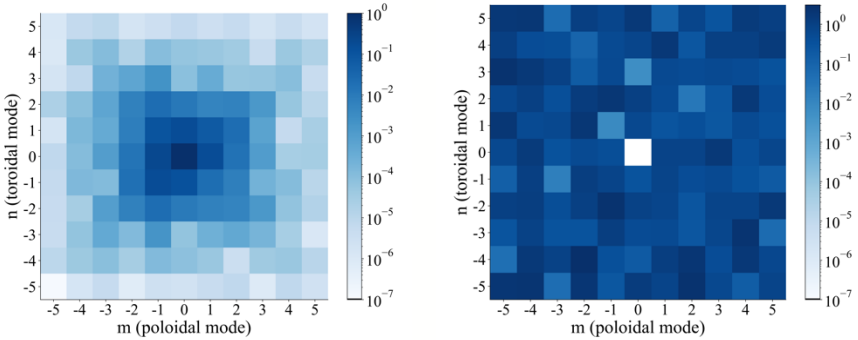


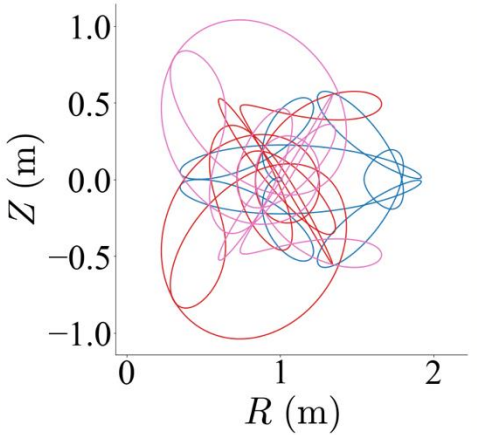


smooth solution

Smoothness and Fourier Decay

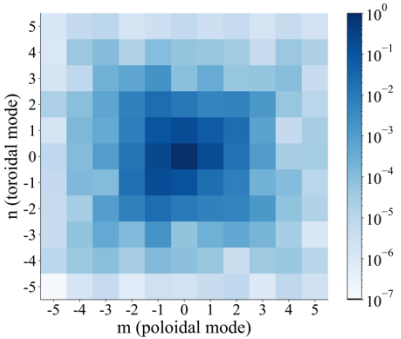
smooth solution

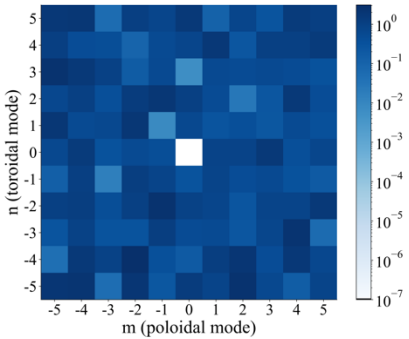


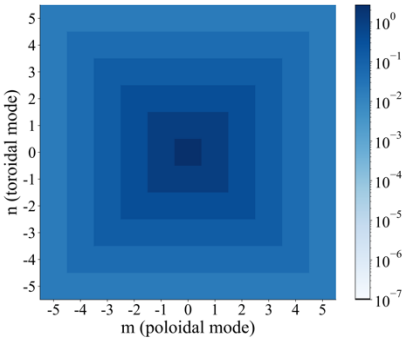


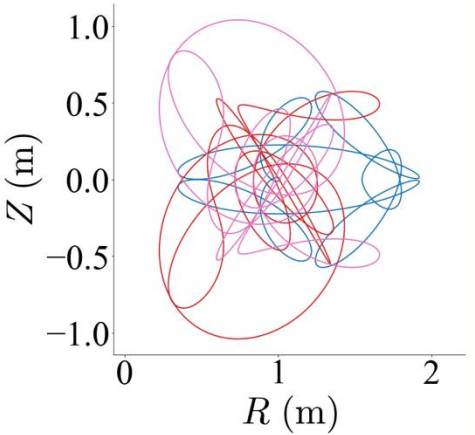
Smoothness and Fourier Decay

smooth solution









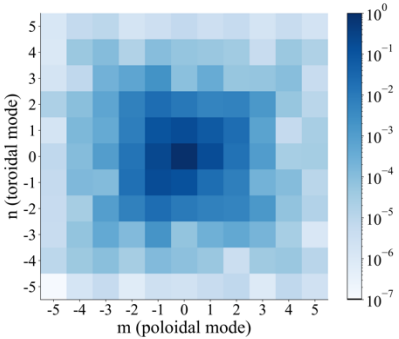
$$|\hat{f}_n| \leq Ce^{-\alpha|n|}$$

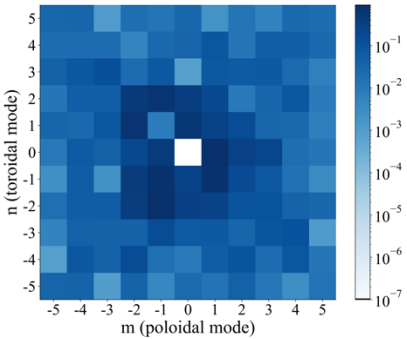
$$C, \alpha > 0$$

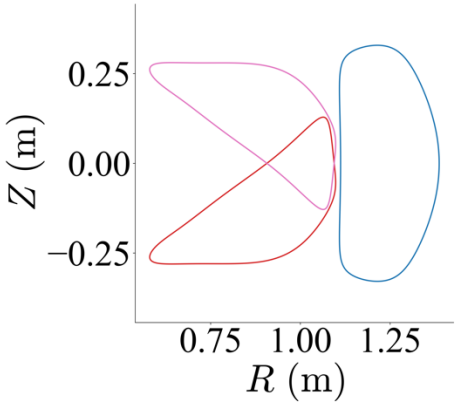
Paley–Wiener (1934), Thm. XII, p. 16

Smoothness and Fourier Decay

smooth solution



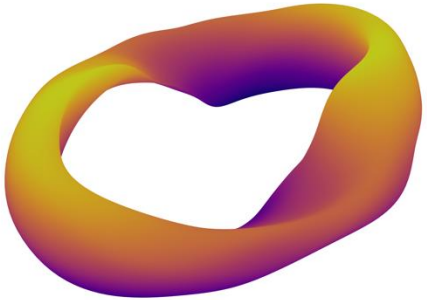




$$|\hat{f}_n| \leq Ce^{-\alpha|n|}$$

$$C, \alpha > 0$$

Paley–Wiener (1934), Thm. XII, p. 16



Exponential Spectral Scaling

$$|\hat{f}_n| \leq Ce^{-\alpha|n|}$$

$$C, \alpha > 0$$

Paley–Wiener (1934), Thm. XII, p. 16

Exponential Spectral Scaling

$$|\hat{f}_n| \leq Ce^{-\alpha|n|}$$

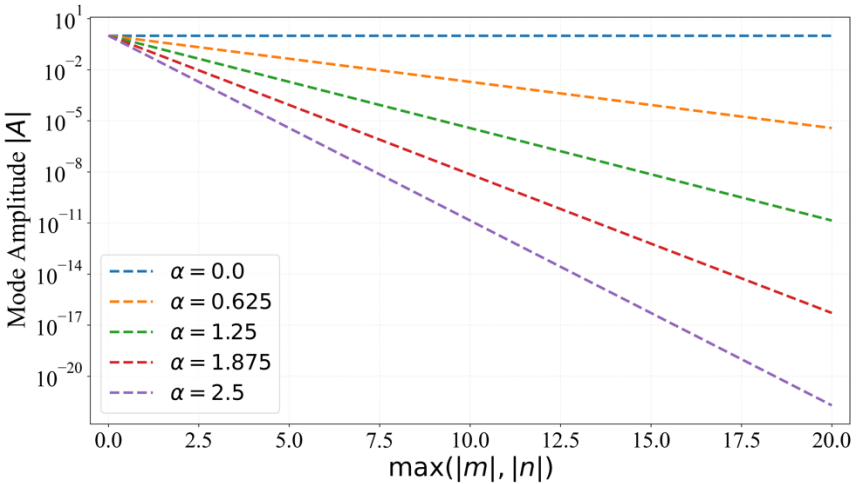
$$C, \alpha > 0$$

Paley–Wiener (1934), Thm. XII, p. 16

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

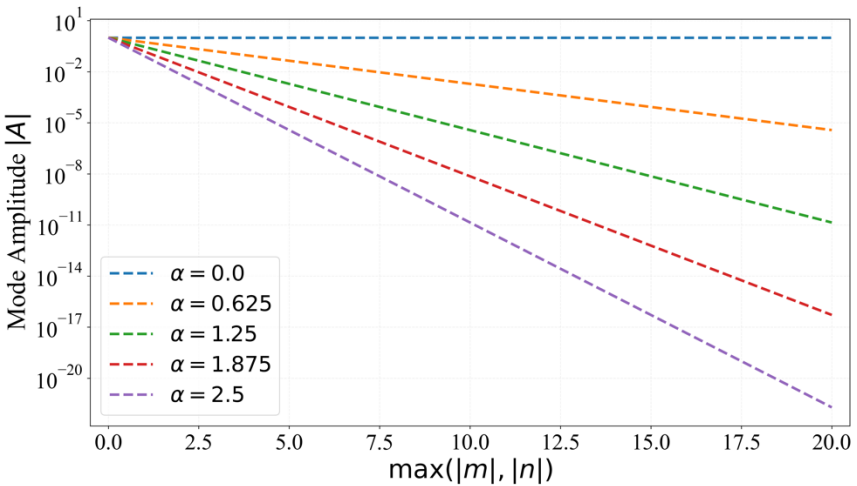
Exponential Spectral Scaling

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

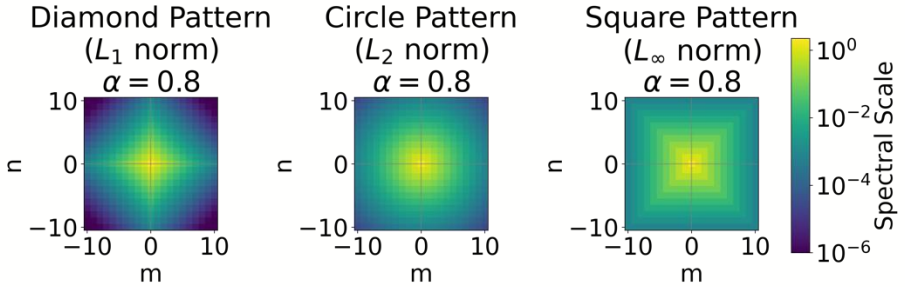


Exponential Spectral Scaling

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

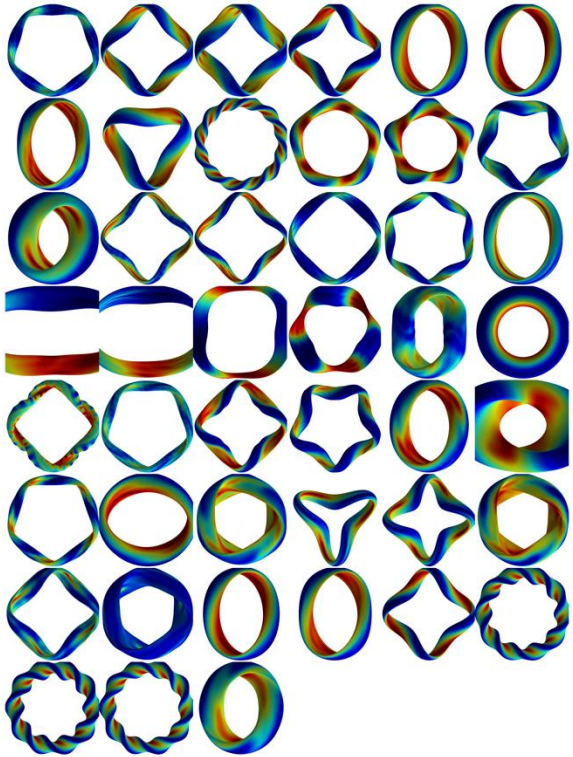


$$g_1 = |m| + |n|, \quad g_2 = \sqrt{m^2 + n^2}, \quad g_\infty = \max\{|m|, |n|\}.$$

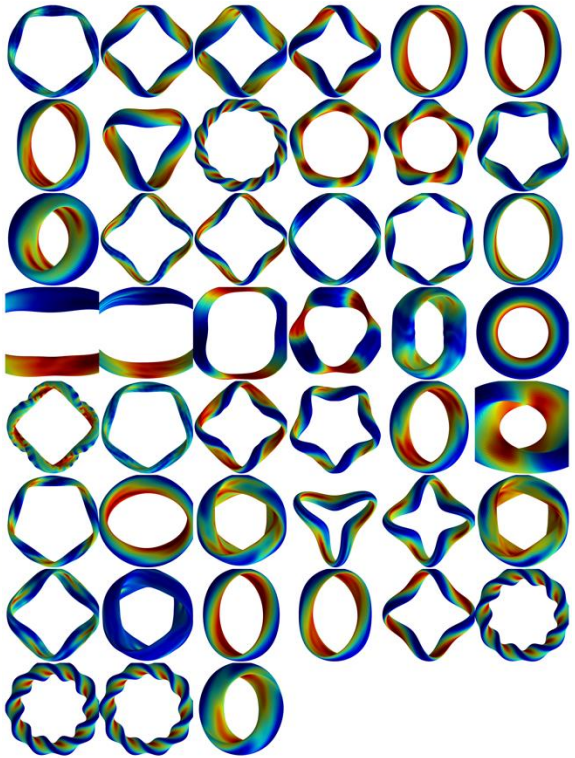


Decay patterns for L_1, L_2, and L_∞ norms.

What would be a good value for the decay/norm type?

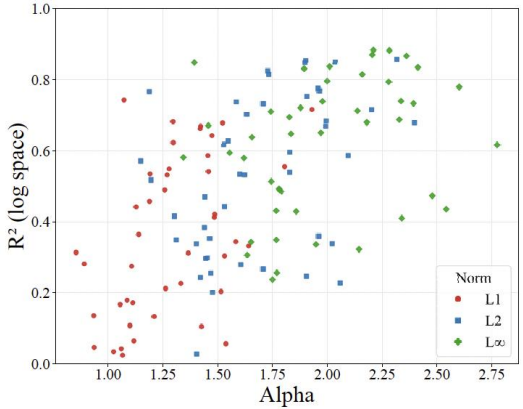


What would be a good value for the decay/norm type?



Fit $\log|A_n^m| \approx -\alpha\, g(m,n)$ across a library of configurations.
For L_2 and L_∞:
- α clusters in ~ 1.1–2.8.
- Good R^2 values.

For L_1:
- Narrower range.
- Systematically worse fits.



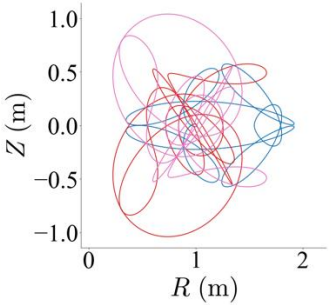
Fitted α vs R^2 across many stellarator configurations.

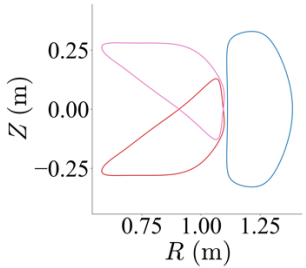
Exponential Spectral Scaling

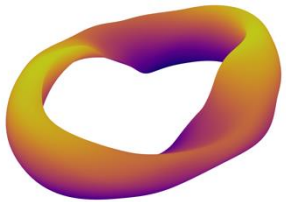
smooth solution

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

decrease MAD





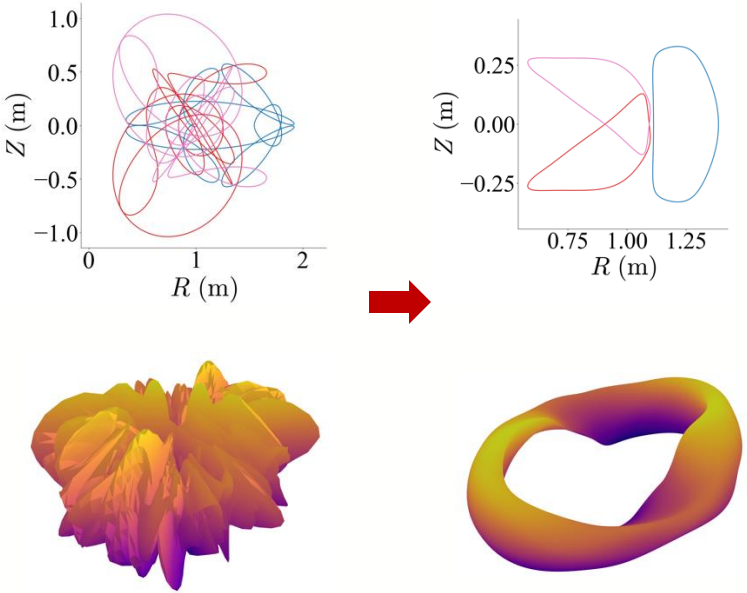


Exponential Spectral Scaling

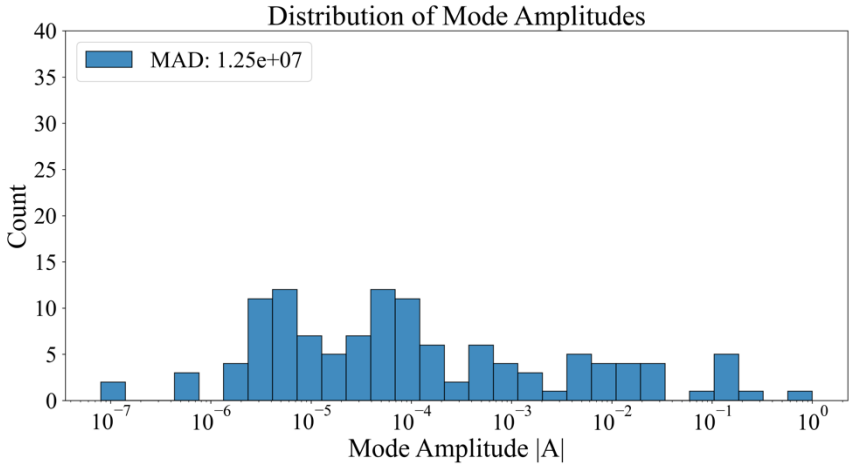
smooth solution

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

decrease MAD



$$\text{MAD} = \frac{\max_{(m,n)} |A_{m,n}|}{\min_{(m,n)\,:\,A_{m,n}\neq 0} |A_{m,n}|}$$

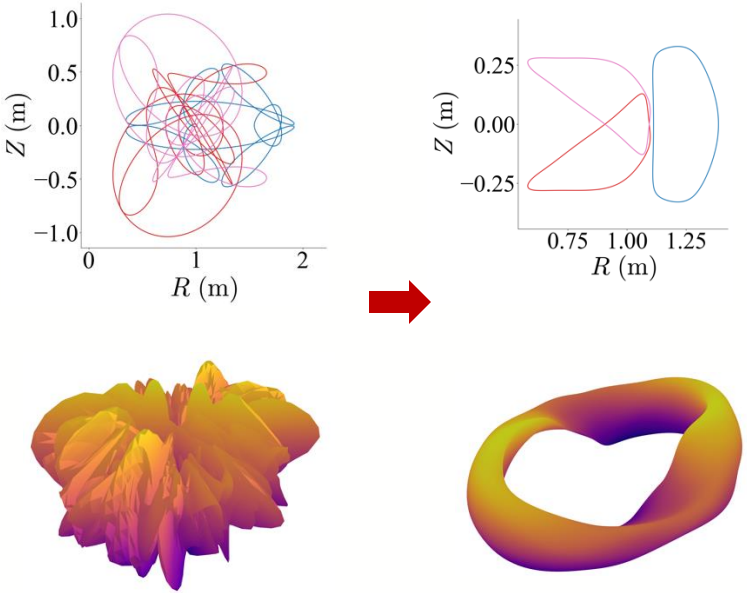


Exponential Spectral Scaling

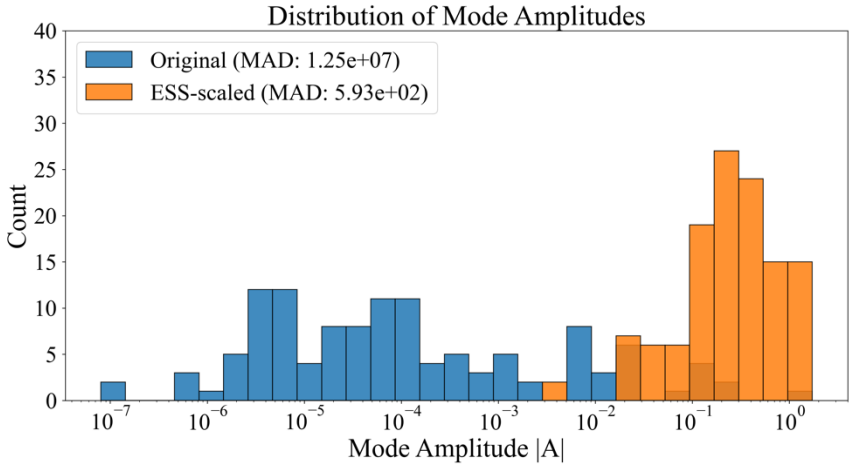
smooth solution

$$\mathbf{x}_{\text{scaled}} = \frac{\mathbf{x}_{\text{original}}}{\exp(-\alpha\, g(m,n))}$$

decrease MAD



$$\text{MAD} = \frac{\max_{(m,n)} |A_{m,n}|}{\min_{(m,n)\,:\,A_{m,n}\neq 0} |A_{m,n}|}$$



Numerical Results

Optimization Objective: QA/QH Problem

vac QS $\min_x f(\boldsymbol{x})$

Objective: weighted sum of three terms

$$f(\boldsymbol{x}) = \omega_{QS} f_{QS} + \omega_{\bar{\iota}} f_{\bar{\iota}} + \omega_A f_A, \quad \omega_\circ = 1.$$

f_{QS}: volume integral of quasisymmetry error.

$f_{\bar{\iota}}$: penalizes deviation of mean rotational transform

f_A: penalizes deviation of aspect ratio.

Benchmarks:

QA: $N_{FP} = 2, (\bar{\iota}_{tar}, A_{tar}) = (0.42, 6)$

QH: $N_{FP} = 4, (\bar{\iota}_{tar}, A_{tar}) = (1.24, 8)$

Numerical Setup

Codes:

- `DESC`: spectral equilibrium + audodiff derivatives
- `SIMSOPT` + `VMEC`: modular optimization with finite-difference

Optimization:

- Trust-region reflective (`method='trf'`)
- Also tested Levenberg-Marquardt
- Initial condition: rotating ellipse

Variables: all Fourier coefficients with $|m|, |n| \leq 5$

Default Scaling (Moré)

Nonlinear least squares problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}\sum_{i=1}^{m} f_i(\mathbf{x})^2, \quad J(\mathbf{x}) \in \mathbb{R}^{m\times n}.$$

Trust-region step with scaled norm:

$$\|Dp\|_2 \le \Delta,$$

$p \in \mathbb{R}^n$ is the step,
$D = \mathrm{diag}(d_1, \ldots, d_n)$ is a diagonal scaling matrix,
$\Delta > 0$ is the trust-region radius in the *scaled* space.

$$d_j = \|J_{:,j}\|_2, \quad (\text{default } d_j = 1 \text{ if } \|J_{:,j}\|_2 = 0)$$
$$d_j \leftarrow \max\{d_j, \|J_{:,j}\|_2\} \quad (\text{monotone update}).$$

Default Scaling (Moré)

Nonlinear least squares problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}\sum_{i=1}^{m} f_i(\mathbf{x})^2, \quad J(\mathbf{x}) \in \mathbb{R}^{m\times n}.$$

Trust-region step with scaled norm:

$$\|Dp\|_2 \le \Delta,$$

$p \in \mathbb{R}^n$ is the step,
$D = \mathrm{diag}(d_1, \ldots, d_n)$ is a diagonal scaling matrix,
$\Delta > 0$ is the trust-region radius in the *scaled* space.

$$d_j = \|J_{:,j}\|_2, \quad (\text{default } d_j = 1 \text{ if } \|J_{:,j}\|_2 = 0)$$
$$d_j \leftarrow \max\{d_j, \|J_{:,j}\|_2\} \quad (\text{monotone update}).$$

What it achieves

Before Moré Scaling

Unscaled variables differ greatly:

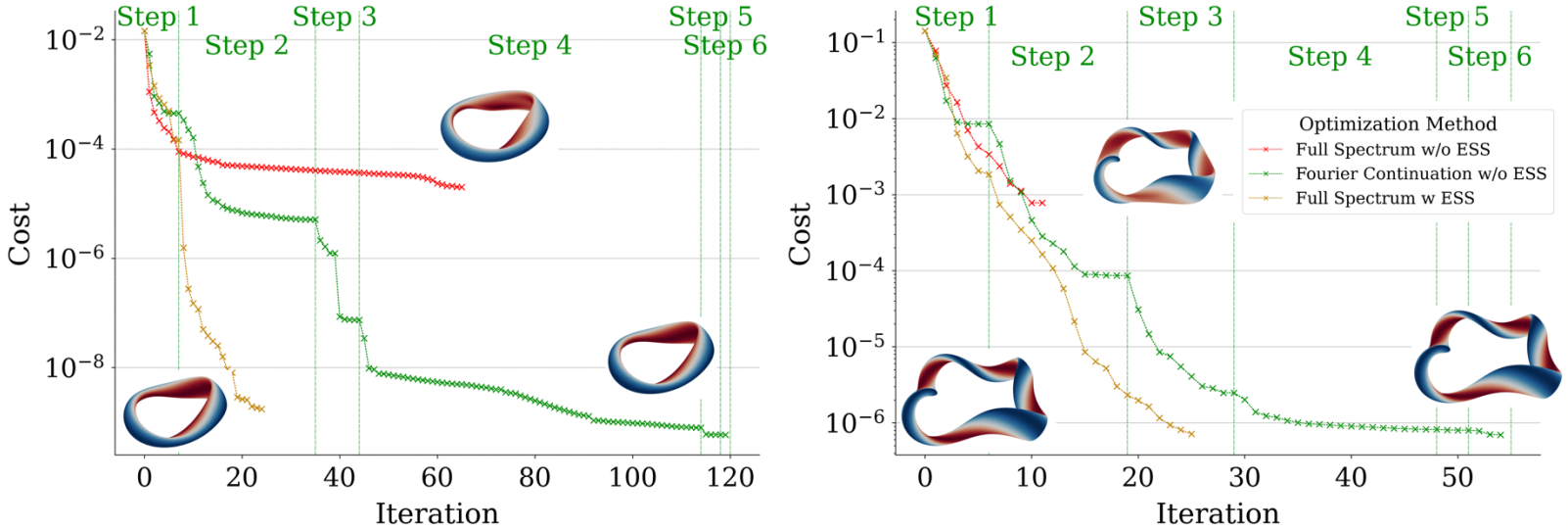
- $x_1 \in [0,1]$ vs $x_2 \in [0,1000]$
- Same step $\Delta = 0.1 \rightarrow$ different effects

After Moré Scaling

- Normalized directions
- Equal step → equal effect

Equal step along any scaled variable has similar effect on cost.

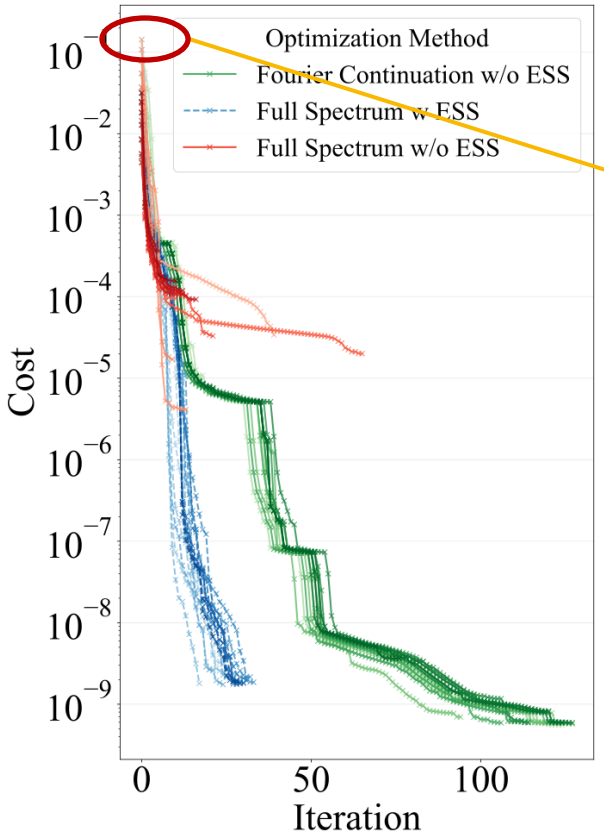
ESS reaches low cost values faster, without staging

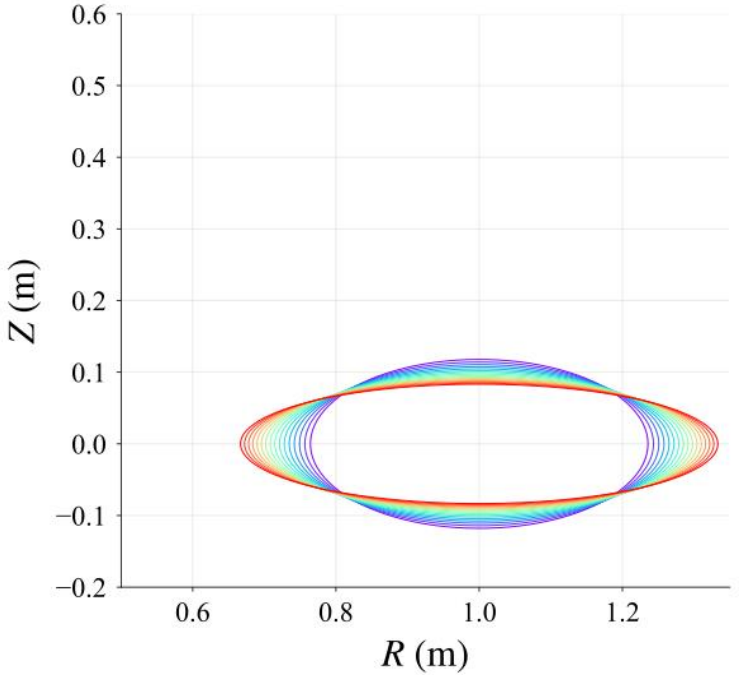


QA and QH cost histories: Default full-spectrum (red), FC (green), and ESS full-spectrum (gold).

Avoids distorted shapes and self-intersections that appear with poorly scaled full-spectrum runs

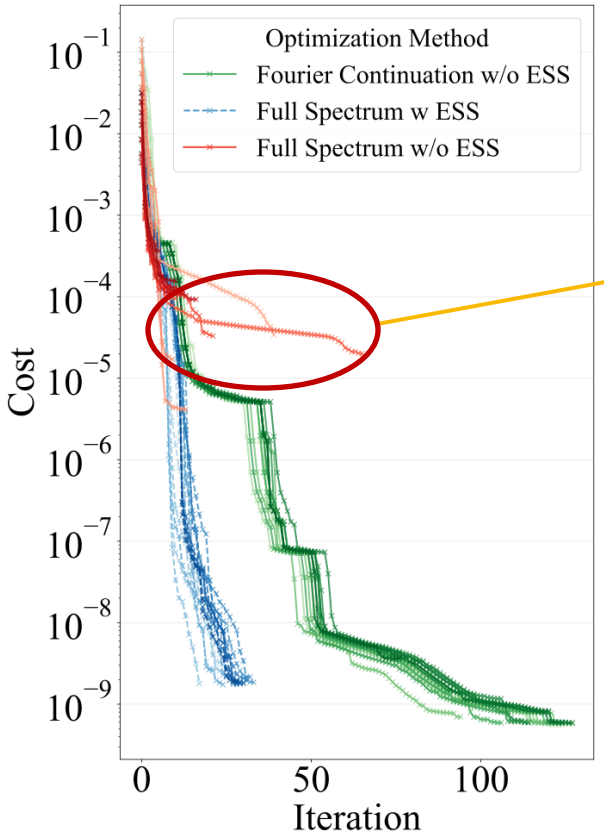
ESS Delivers Consistent Convergence and Shape Quality

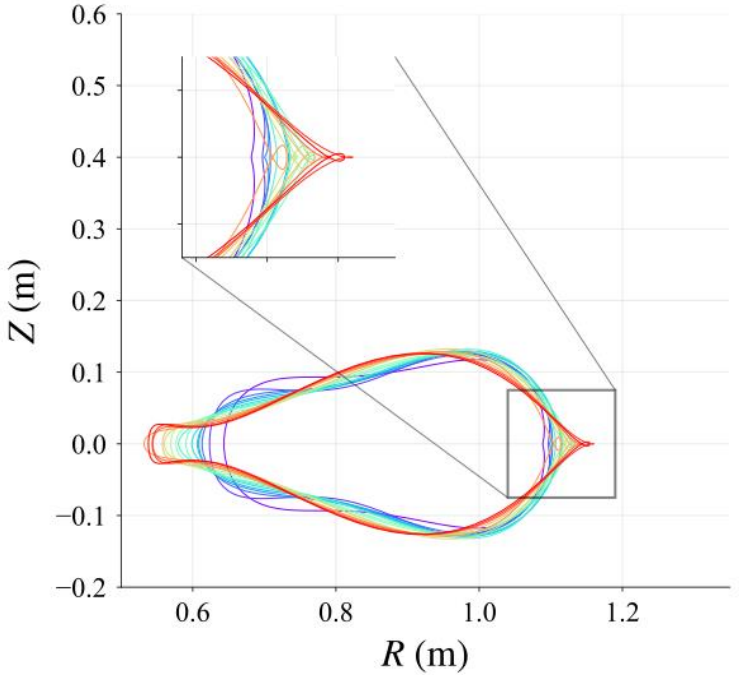




Initial Conditions

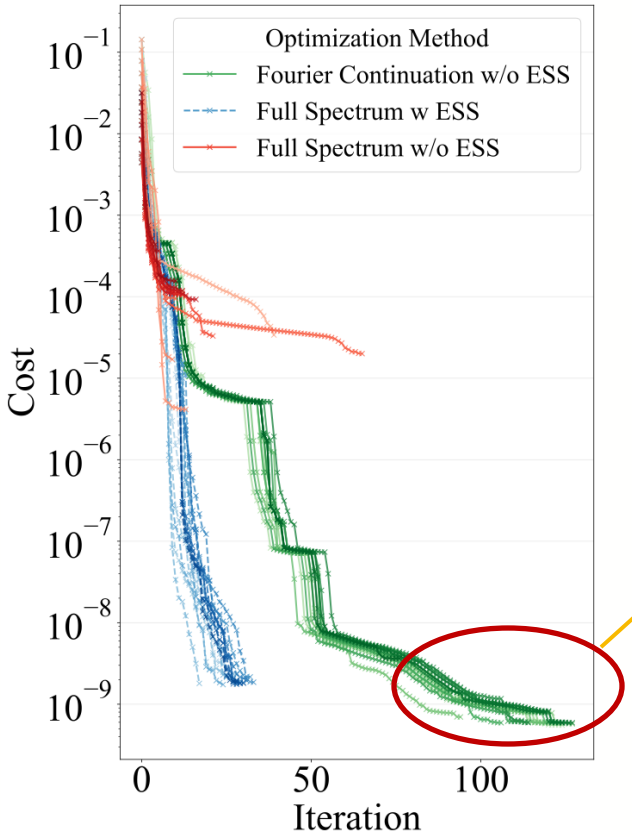
ESS Delivers Consistent Convergence and Shape Quality

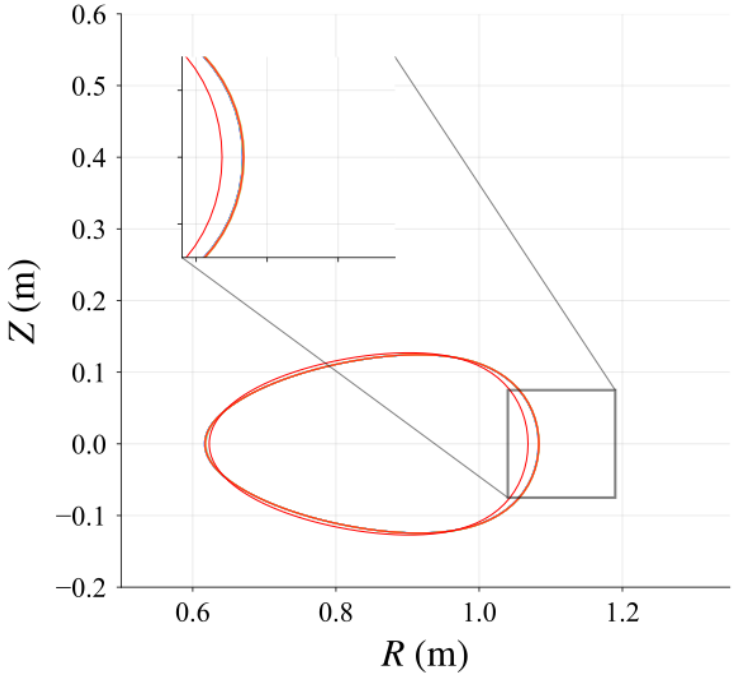




Stalled Cases

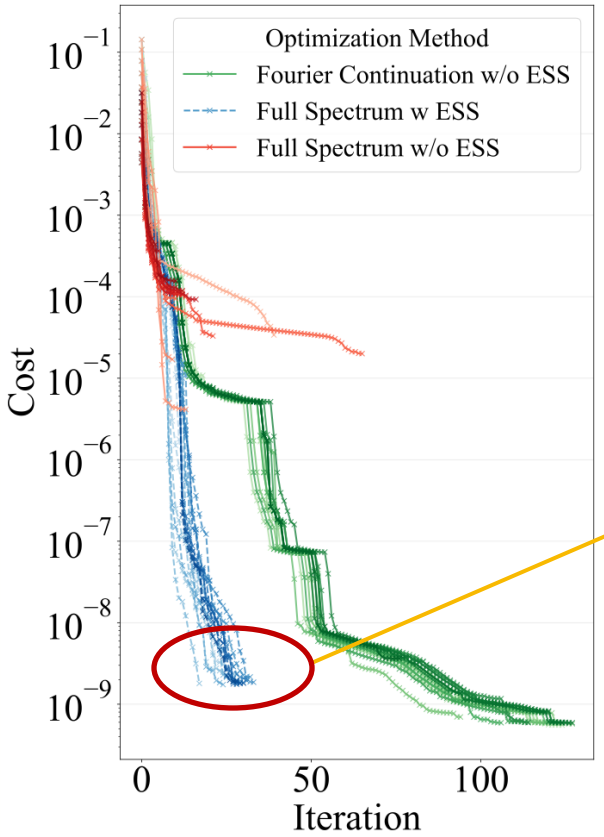
ESS Delivers Consistent Convergence and Shape Quality

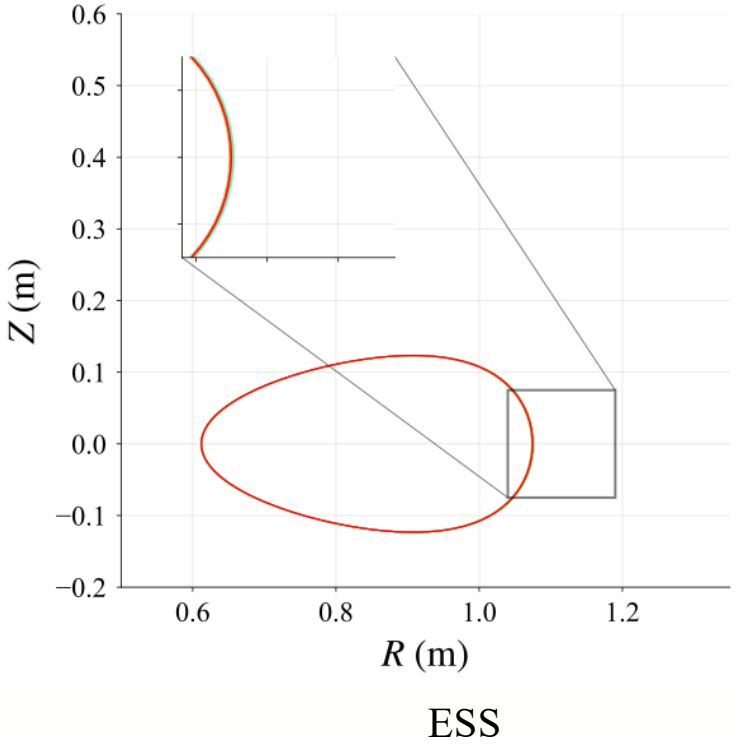




Fourier Continuation

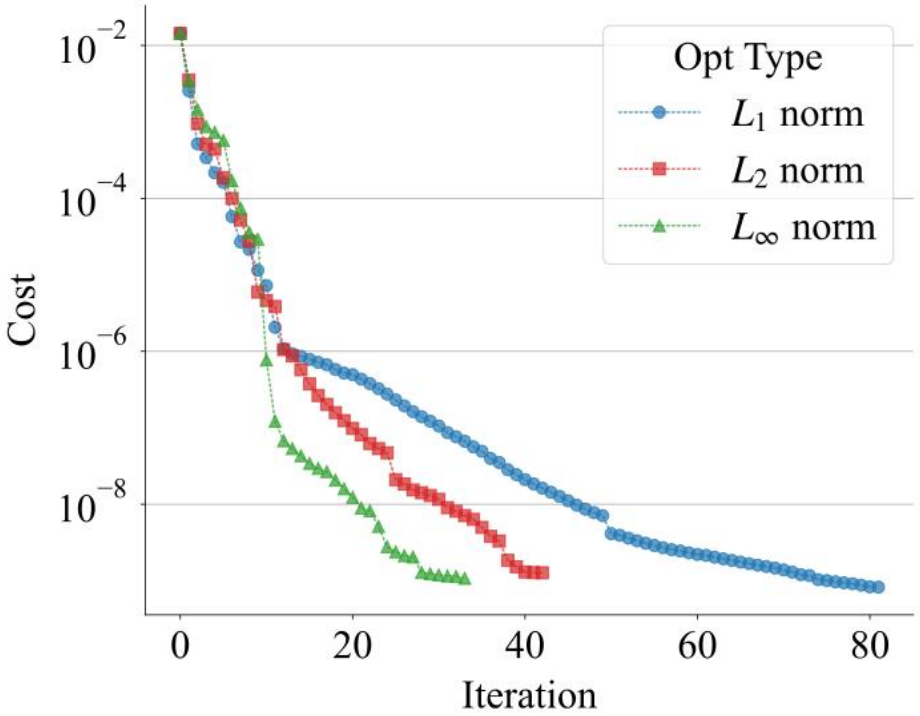
ESS Delivers Consistent Convergence and Shape Quality





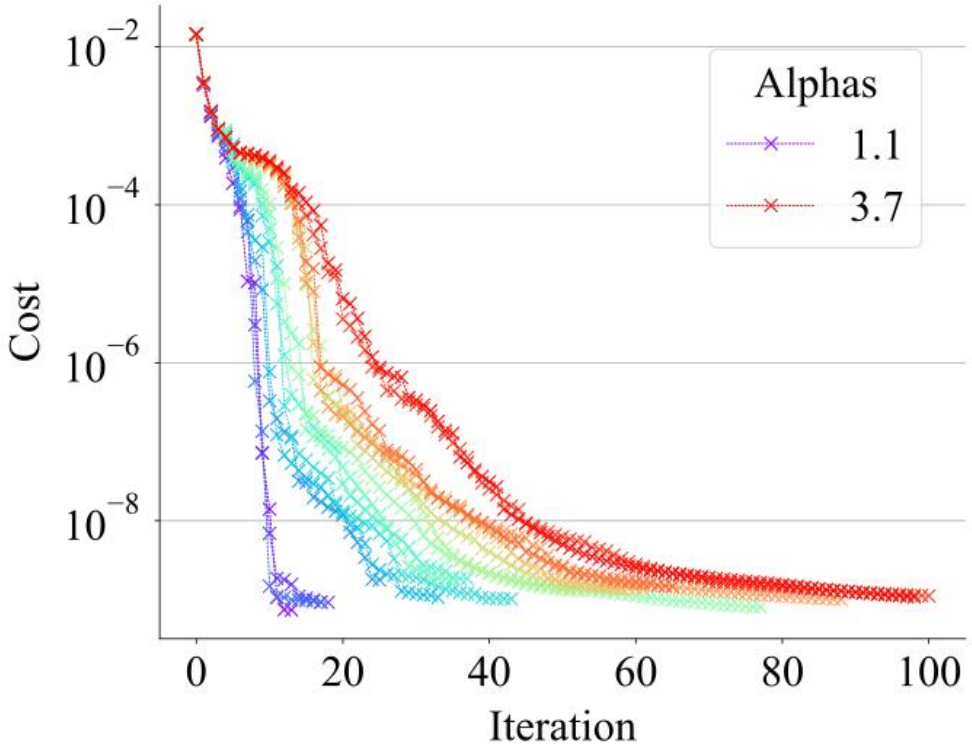
ESS

Norm Choice: L_1 vs L_2 vs L_∞



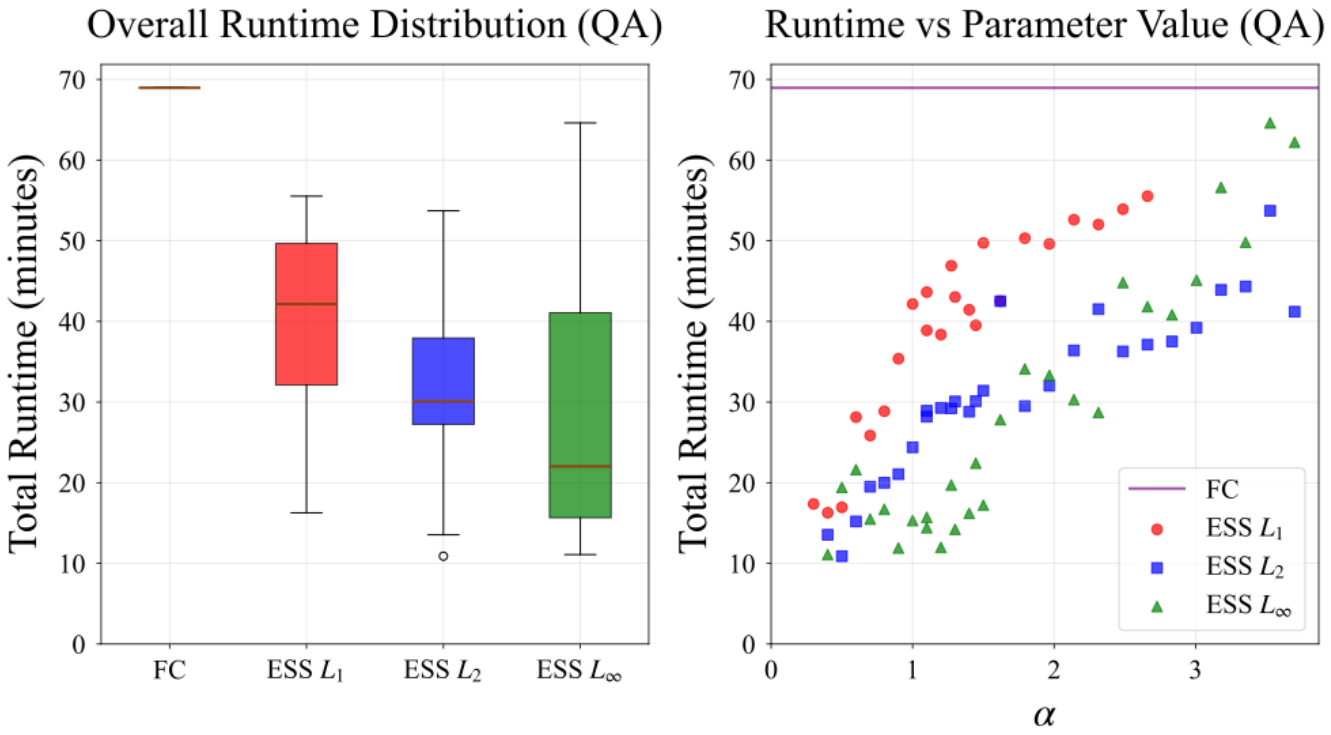
Cost histories for different ESS norm choices (α fixed)

Robustness in α for L_∞



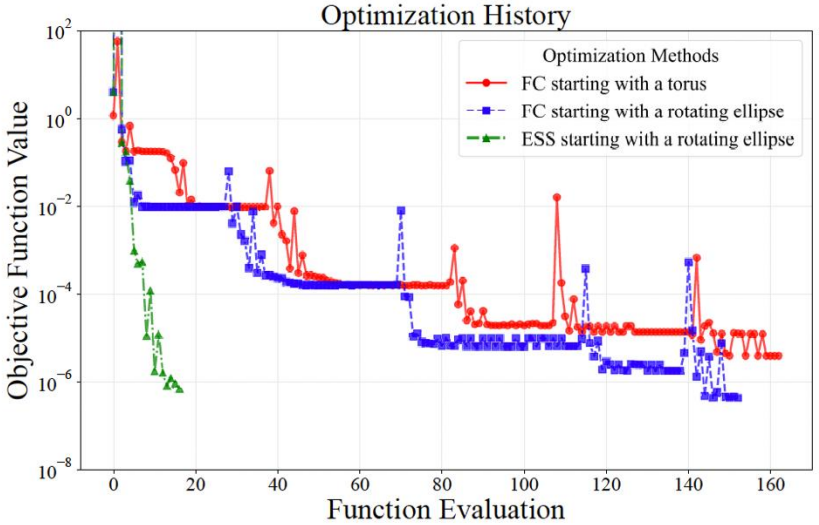
ESS convergence for different α with L_∞ norm

Wall-Clock Performance vs α and L_*



Time to reach f $\leq 10^{-8}$ for various (α, L_*); FC shown for comparison

ESS is Solver-Agnostic: Results with SIMSOPT



Cost histories for three QA runs in SIMSOPT with ESS ($L_\infty,\ \alpha = 1.2$).

Same optimization problem and initial conditions as in DESC.

Similar convergence behavior despite different backend (VMEC+ finite differences).

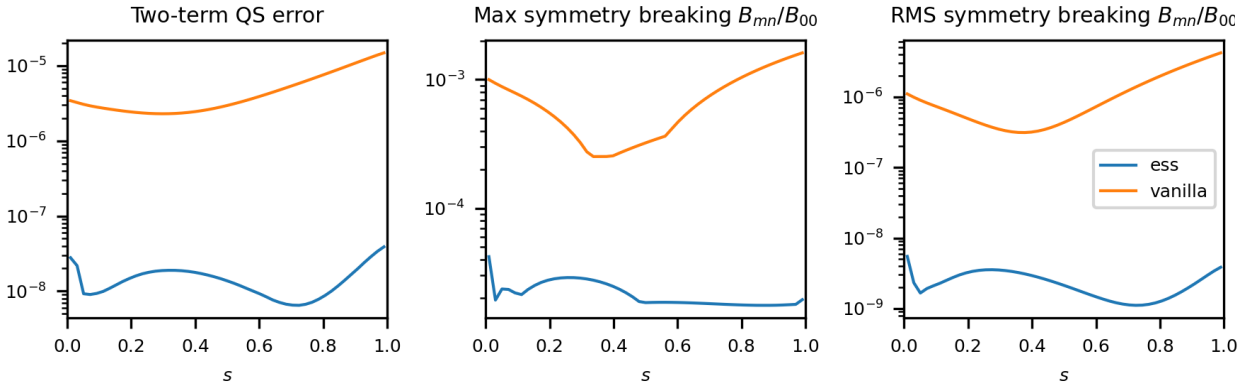
ESS acts purely in coefficient space => portable to other toolchains

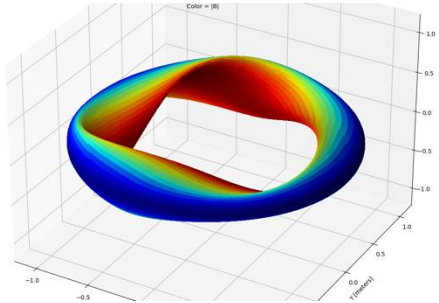
Other Results

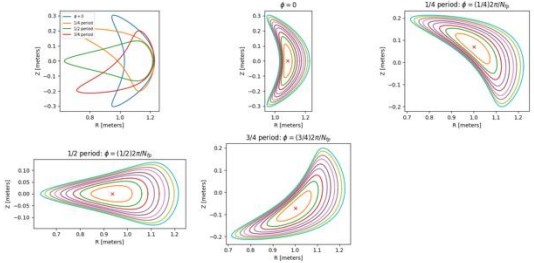
Different rotational transform objective - Stefan

vac QA $\min_{x} f(x)$

$f_{(\iota_{target}>0.55)}$



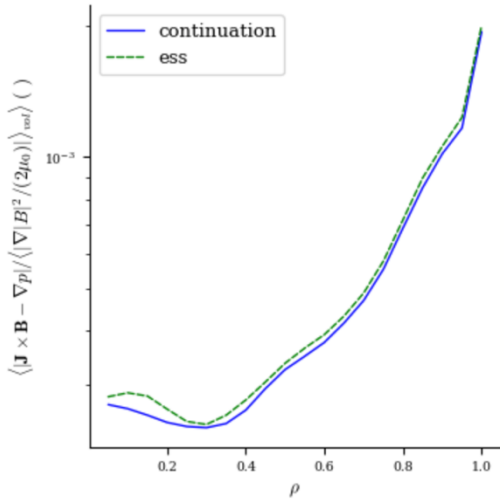




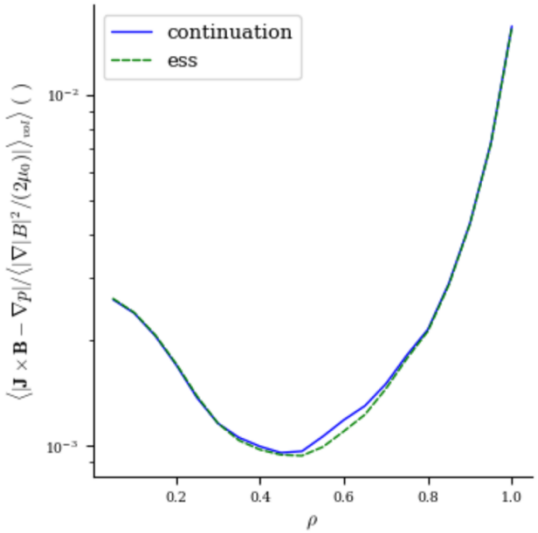
Inner Modes for Force Balance - Rory

fixed boundary force balance $\min_{x} f(\boldsymbol{x})$

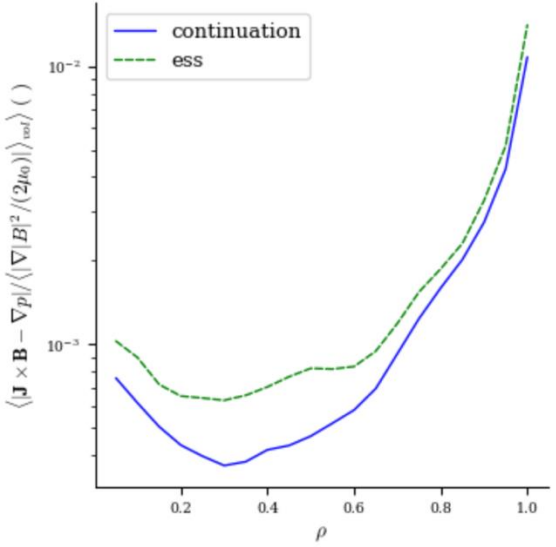
HSX



W7-X

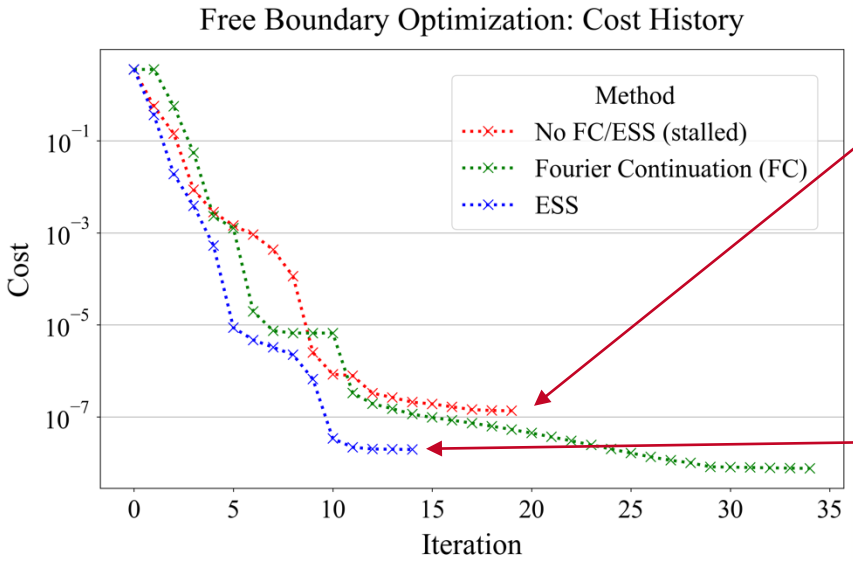


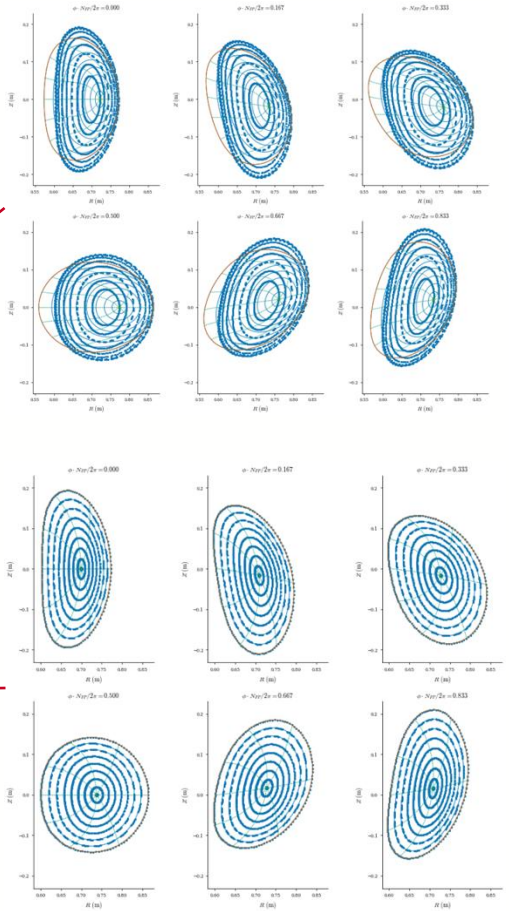
NCSX



Free Boundary - Dario

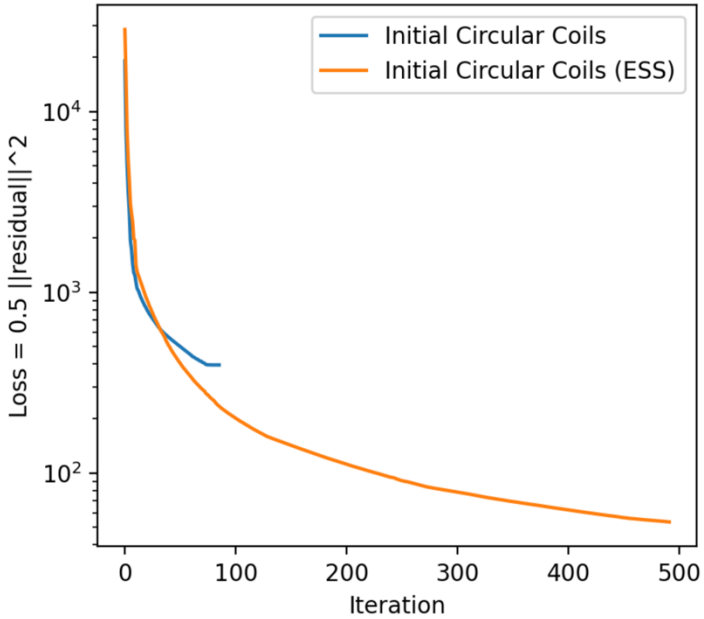
free boundary QA $\min_{x} f(x)$





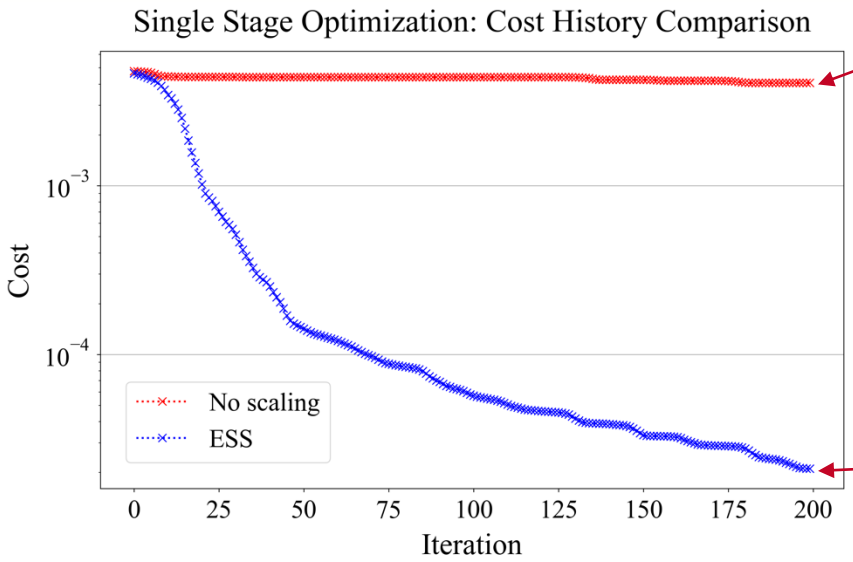
Stage-2 - Rogerio

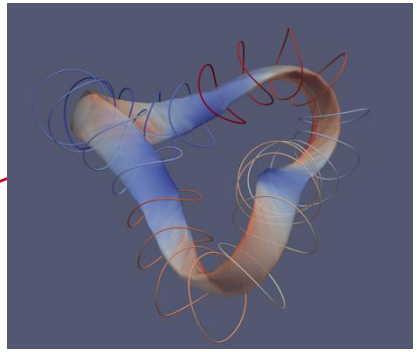
Stage-2 cold start $\min_x J(\boldsymbol{x})$

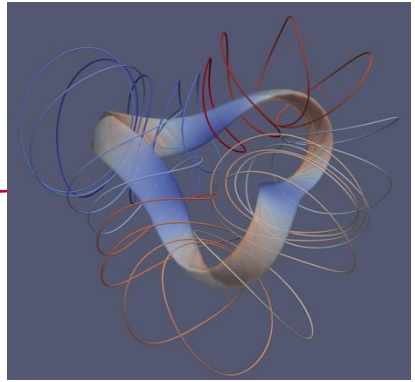


NAE - Misha

finite β single-stage $\min_x J(x)$

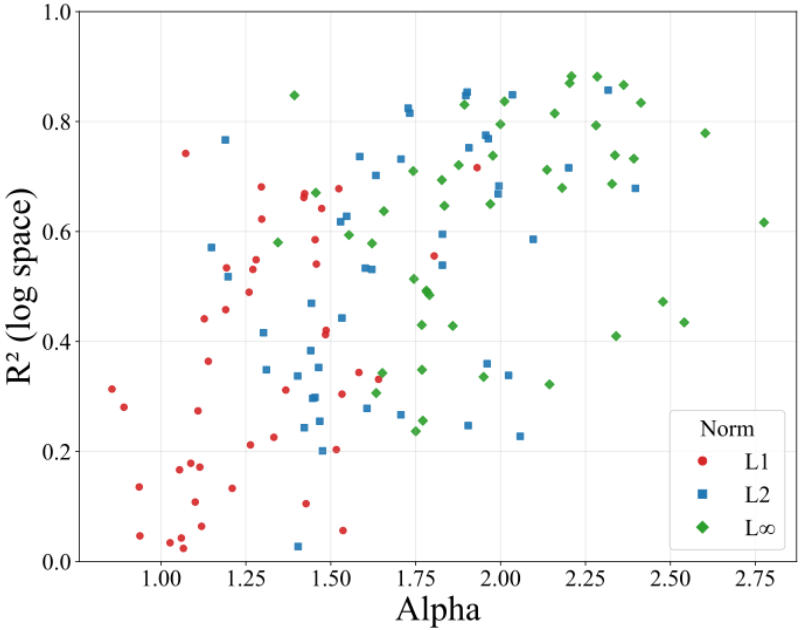


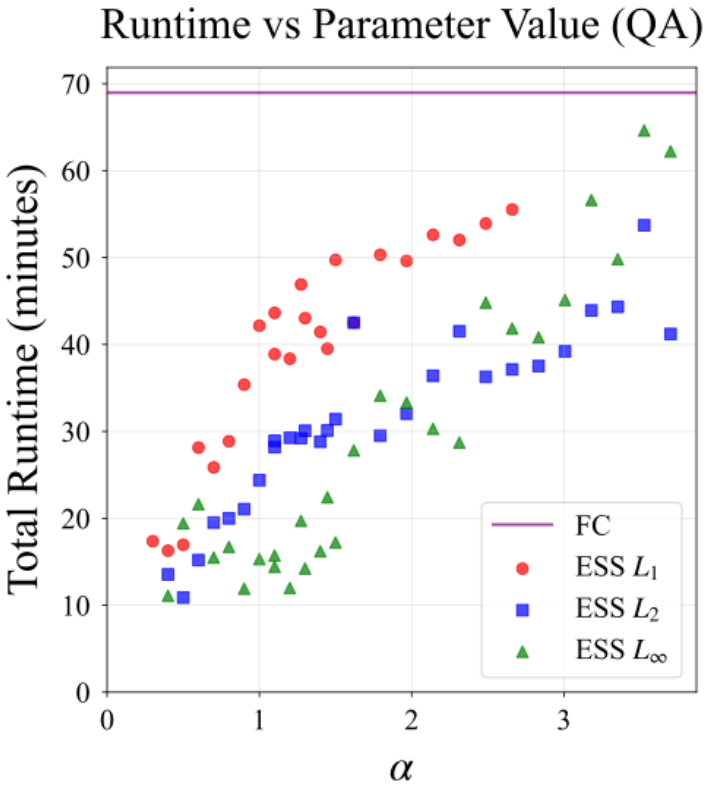




Discussions

Why is the best performing ESS parameter different?



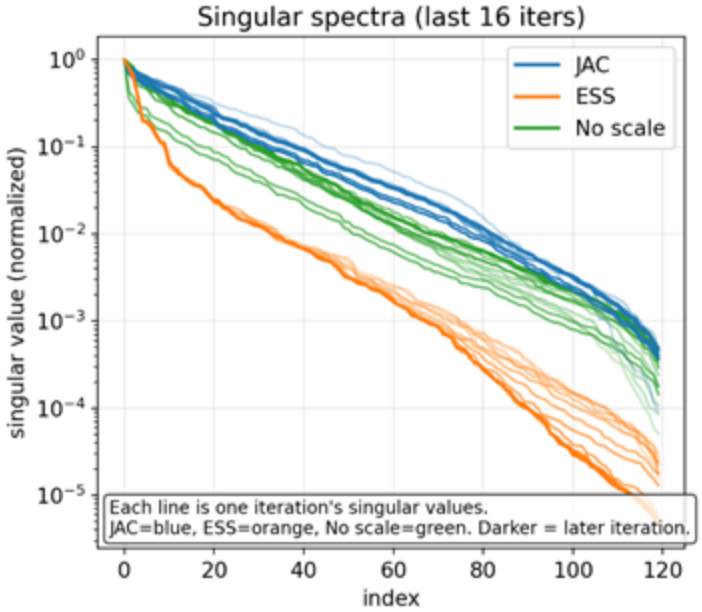


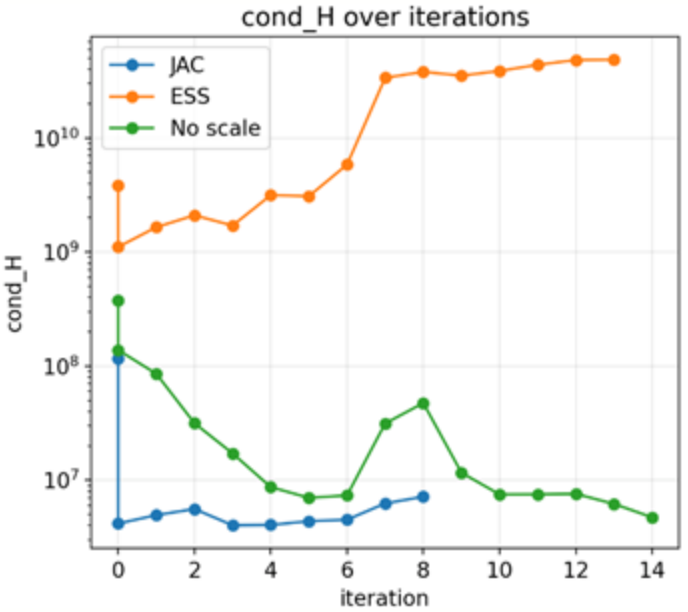
Condition number doesn't seem to tell the full story

It works but why? It's not a preconditioner as the condition number increases

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}\sum_{i=1}^{m} f_i(\mathbf{x})^2$$

$$J(\mathbf{x}) \in \mathbb{R}^{m\times n}$$





Future Work

Apply ESS to other symmetry classes:
- Quasi-iso => anisotropic scaling in (m,n) if needed

- Combine with global optimization (PR in stellopt)

Deeper numerical questions

- Why Jacobian-based Moré scaling underperforms in stage-1

- Why ESS/FC are less critical in stage-2?

Code & Preprint

Code (Zenodo)

DOI: 10.5281/zenodo.1714501

Preprint (arXiv)

arXiv: 2509.16320

Accepted for publication in the Journal of Plasma Physics on November 18th, 2025.

Thank You!

byoungj@umd.edu