Accelerating Transport Solvers with Surrogate Modeling and Uncertainty Quantification

P. Rodriguez-Fernandez[1*], N. T. Howard[1], A. Martin-Sanabria[1], A. Ho[1], J. Pimentel-Aldaz[1], J. Hall[1], A. Saltzman[1] and C. Holland[2]

[1]MIT Plasma Science and Fusion Center, Cambridge, MA, United States
[2]University of California San Diego, San Diego, CA, United States
pablorf@mit.edu

2025 Sherwood Fusion Theory Conference
New York City (NY), April 7th-10th 2025

This work was funded by US DoE DE-SC0024399

Abstract

Core transport modeling of tokamak plasmas has become a critical tool to scope fusion power plants and plan experiments in upcoming burning plasmas. However, increasing fidelity of the turbulent transport models often comes with an associated increase in computational cost and additional challenges to reach convergence or steady-state conditions. Critical gradient, stiff transport and the uncertainty derived from either a limited time average in initial value solvers or discontinuities in eigenvalue solvers are usually the culprits for the high computational expense of transport simulations. This work aims to tackle these issues, with the goal of making physics-based transport simulations accessible to predict performance and enable their routine use for reactor design, as opposed to empirical scaling law extrapolations. To this end, refinements to the PORTALS framework [1] —performed under the Surrogate Models for Accurate and Rapid Transport Solutions (SMARTS) SciDAC partnership— are presented. Surrogate model selection, quantification of evaluation uncertainty, physics-guided optimization methods and custom convergence metrics are implemented in PORTALS, demonstrating the workflow can produce steady-state solutions using the TGLF quasilinear transport model [2] efficiently with minimal user input. PORTALS-TGLF has been used to produce databases of thousands of converged, steady-state transport simulations for numerical studies in experimentally-relevant plasma regimes, enabling studies that were only possible with neural network versions of turbulent transport models or with manual tuning and convergence checks.

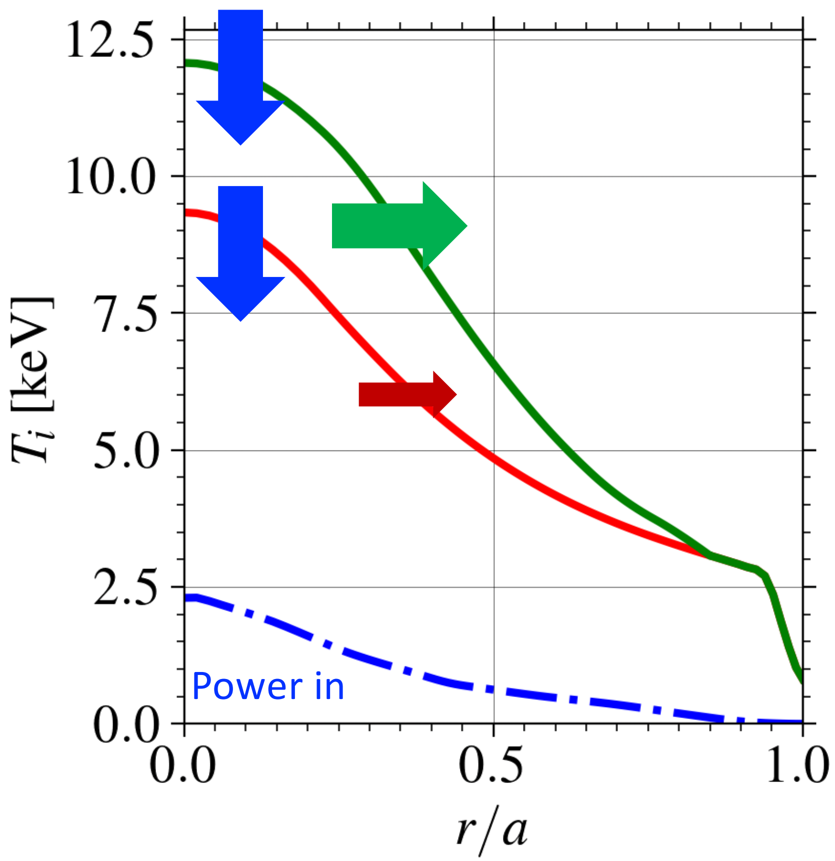
This work was funded by US DoE DE-SC0024399.

[1] P. Rodriguez-Fernandez NF 2024
[2] G. M. Staebler PoP 2007

Practical predictions require an inverse solver for δf turbulence codes

- Solving for steady-state with local transport models becomes an inverse problem.

$$\frac{1}{V'}\left[\frac{\partial}{\partial t}\left(\frac{3}{2}n_i T_i V'\right) + n_i T_i \frac{\partial V'}{\partial t}\right] = -\frac{1}{V'}\frac{\partial}{\partial r}\left(\langle q_i \cdot \nabla r\rangle V'\right) + \langle S_i\rangle$$

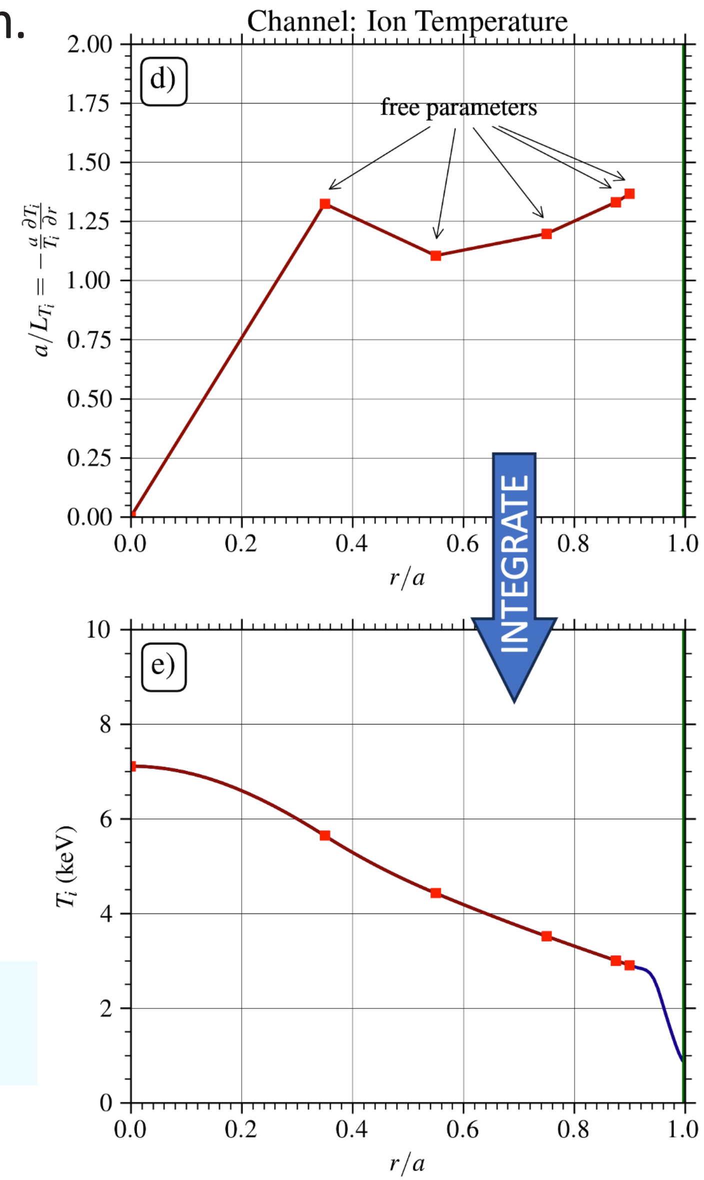


Steady-state:

$$\langle q_i \cdot \nabla r\rangle = \frac{1}{V'}\int_0^r \langle S_i\rangle V' dr$$

$$Q_i^{TRANSPORT} = Q_i^{TARGET}$$

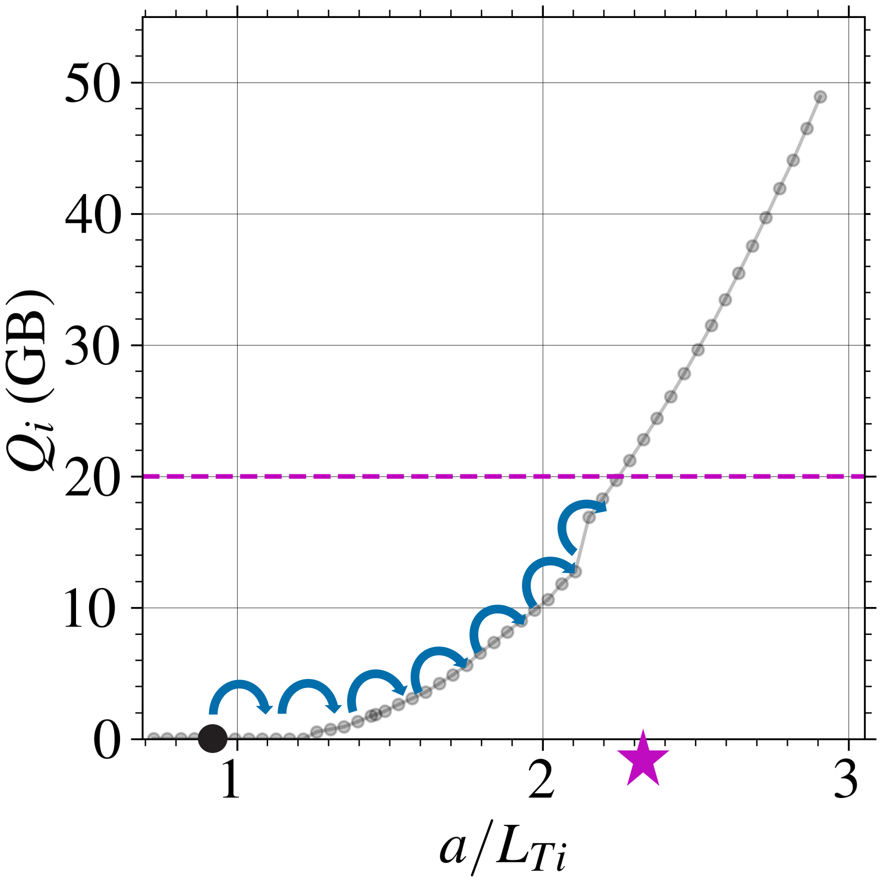
$$a/L_X^* = \text{argmin}\,|Q_i^{TRANSPORT} - Q_i^{TARGET}|$$

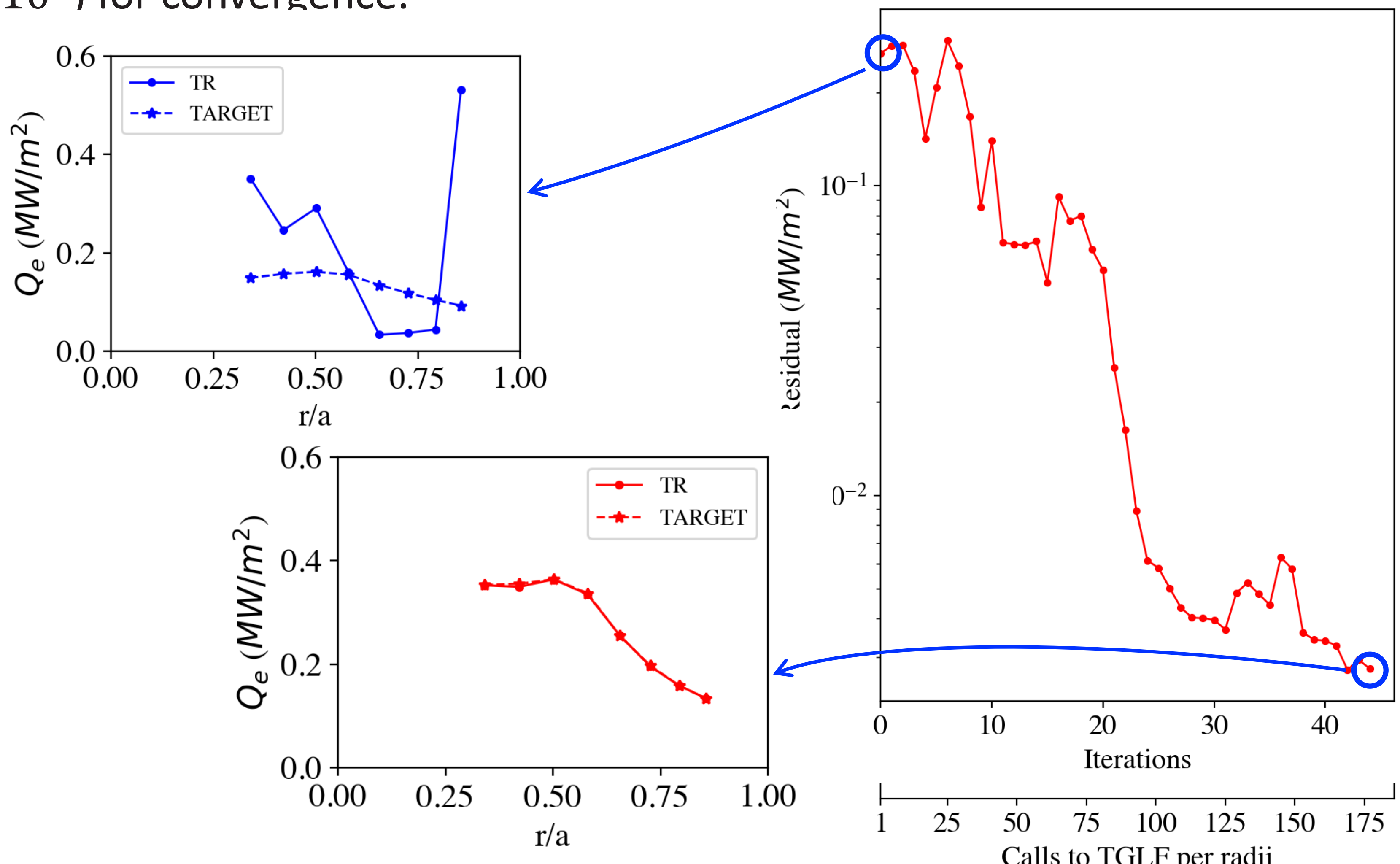


Standard solvers require too many iterations → *expensive predictions*

MIT PSFC

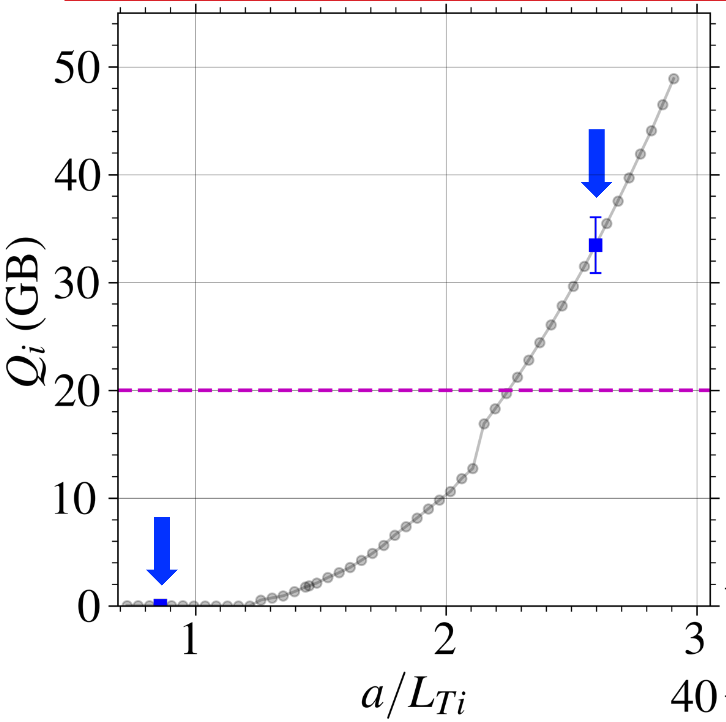
- Standard solvers find the steady-state via small “steps” to reducing a residual or dW/dt.
- Jacobian calculation, use of small steps or ad-hoc modifications of flux response to achieve convergence often leads to many evaluations ($\sim 10^2 - 10^3$) for convergence.



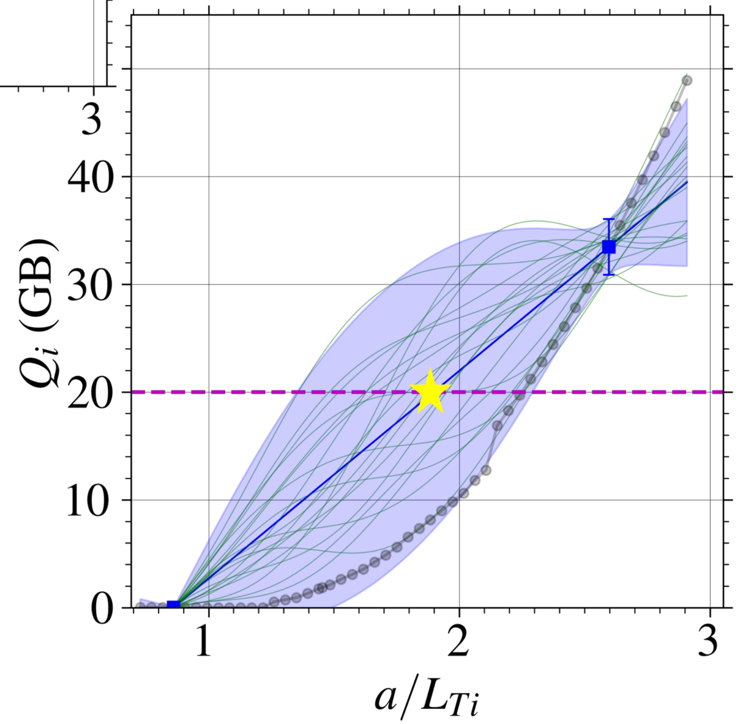


PORTALS leverages surrogate-based optimization to reach convergence

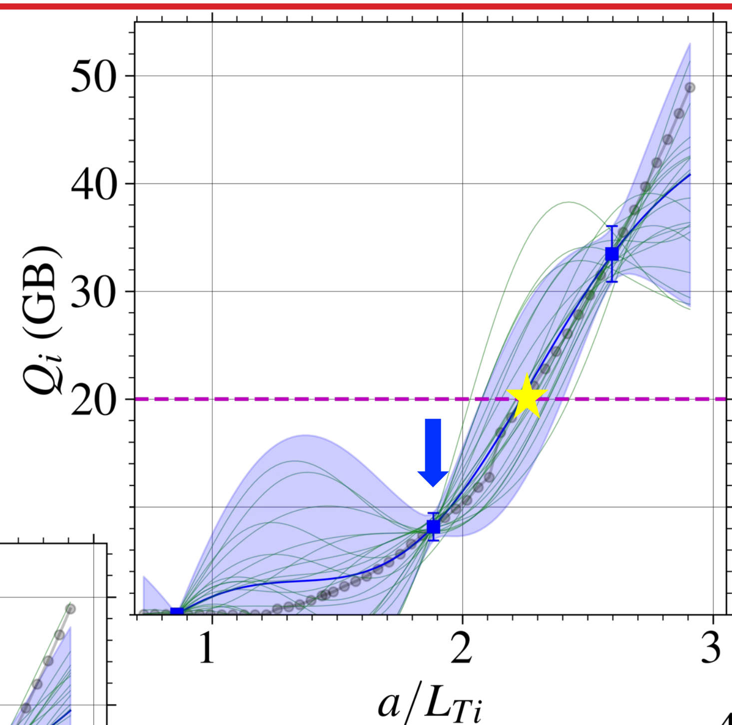
MIT
PSFC



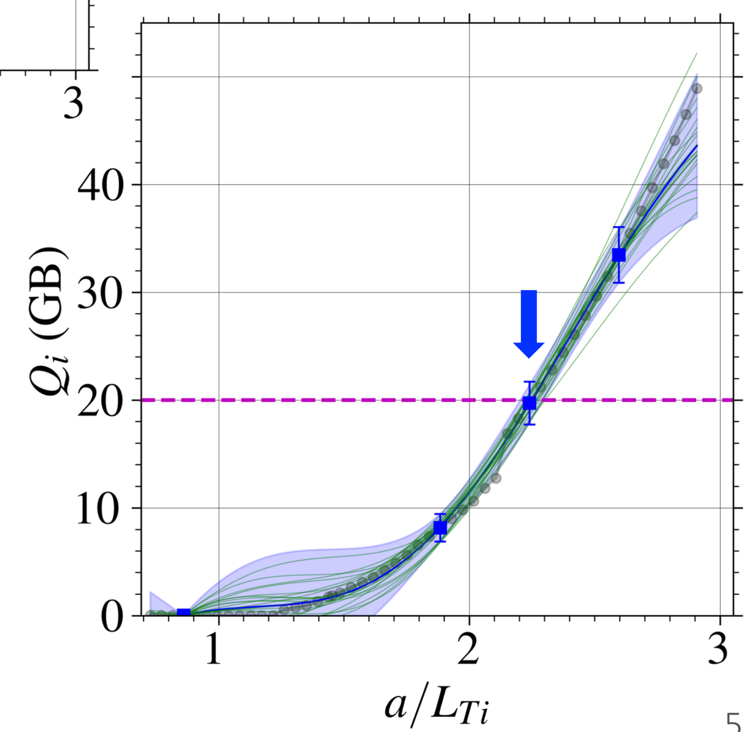
Surrogate model trained on a few simulations



Point predicted by surrogate to be flux-matched is simulated and added to database



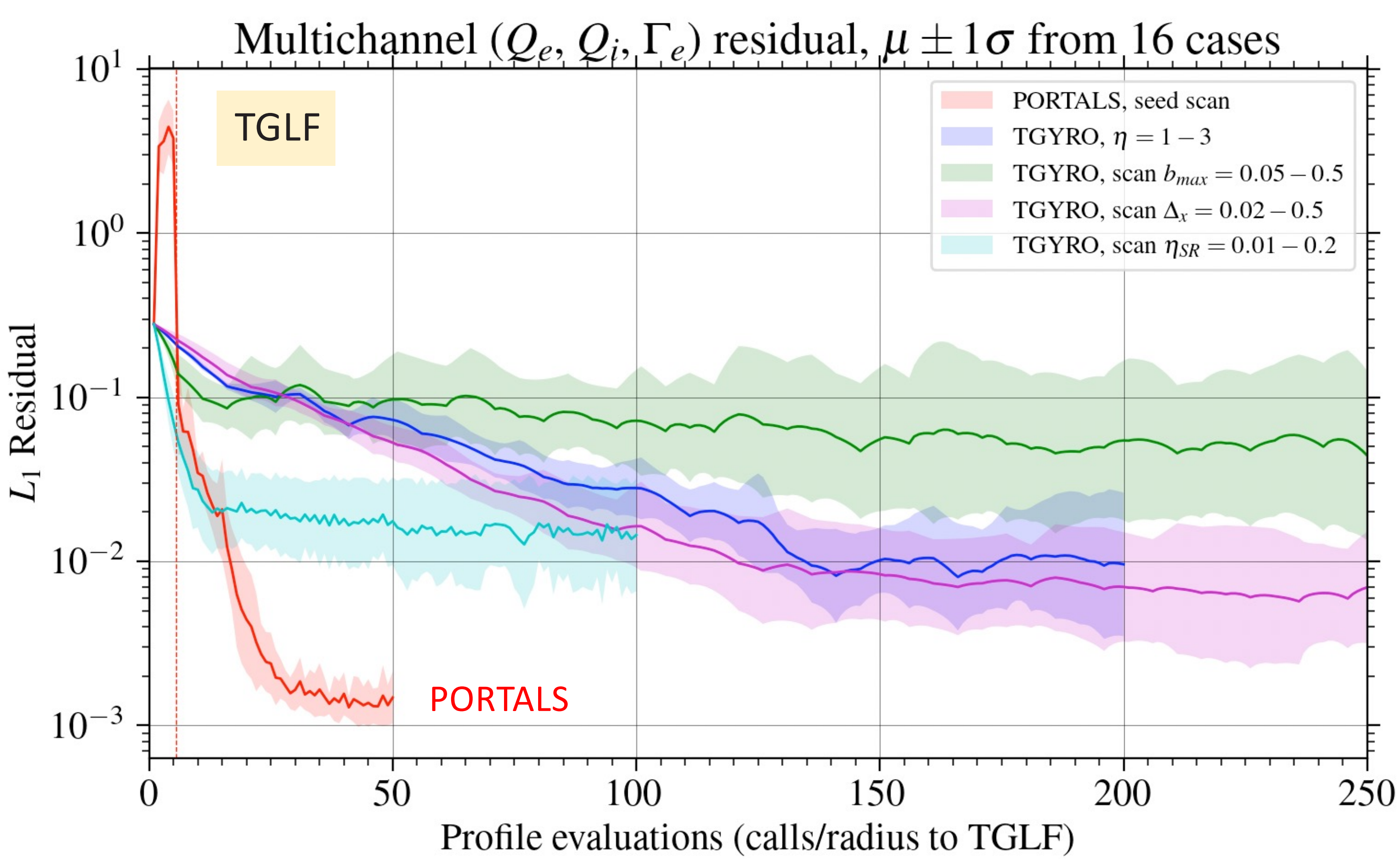
Surrogate becomes more accurate as solution is approached

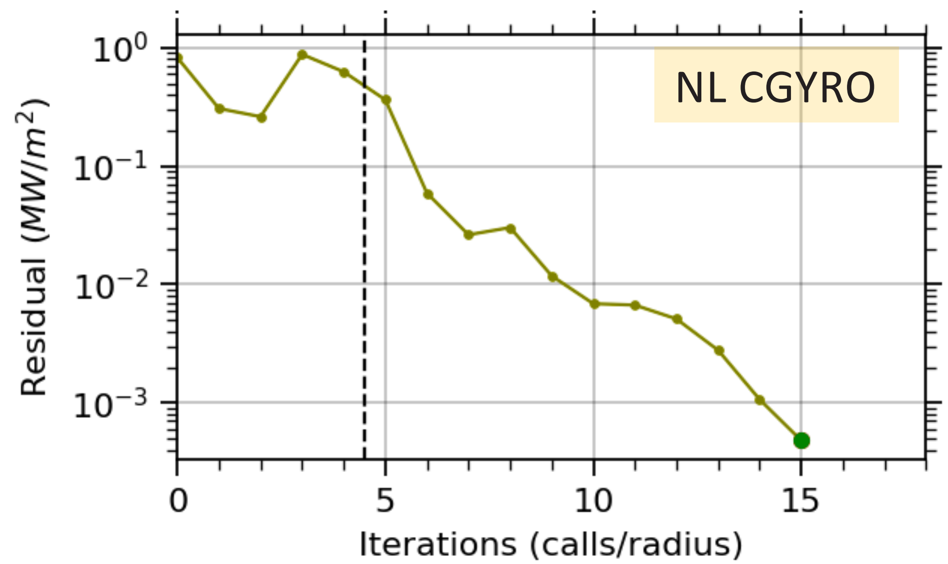


Surrogate techniques proved to be very effective at flux-matching problem

MIT PSFC

- We find that $\sim\mathbf{15-20}$ **evaluations** (simulations/radius) are typically enough for convergence with NL CGYRO[1].
- Benchmarks with quasilinear TGLF[2] indicate **faster convergence** than with standard numerical solvers.

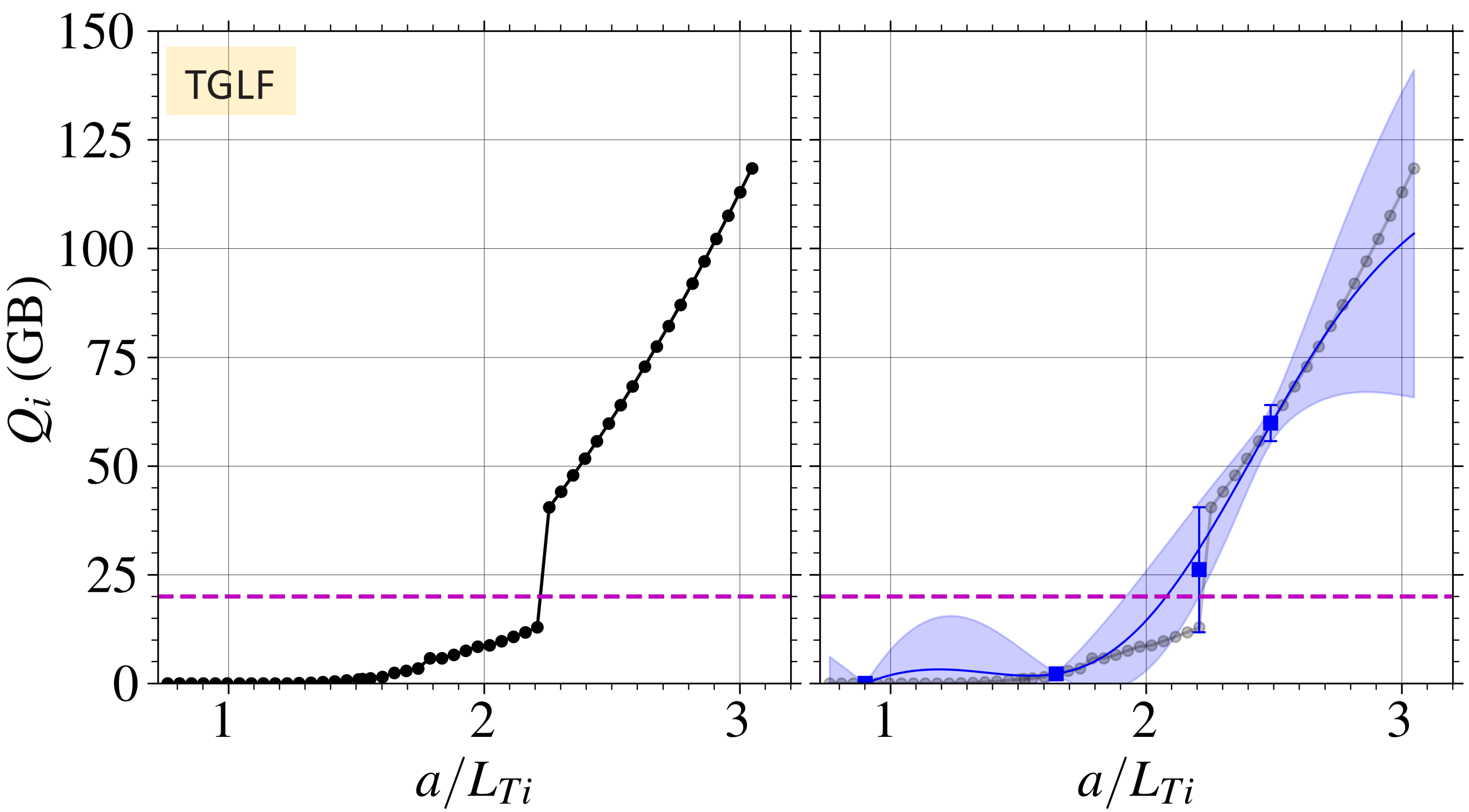


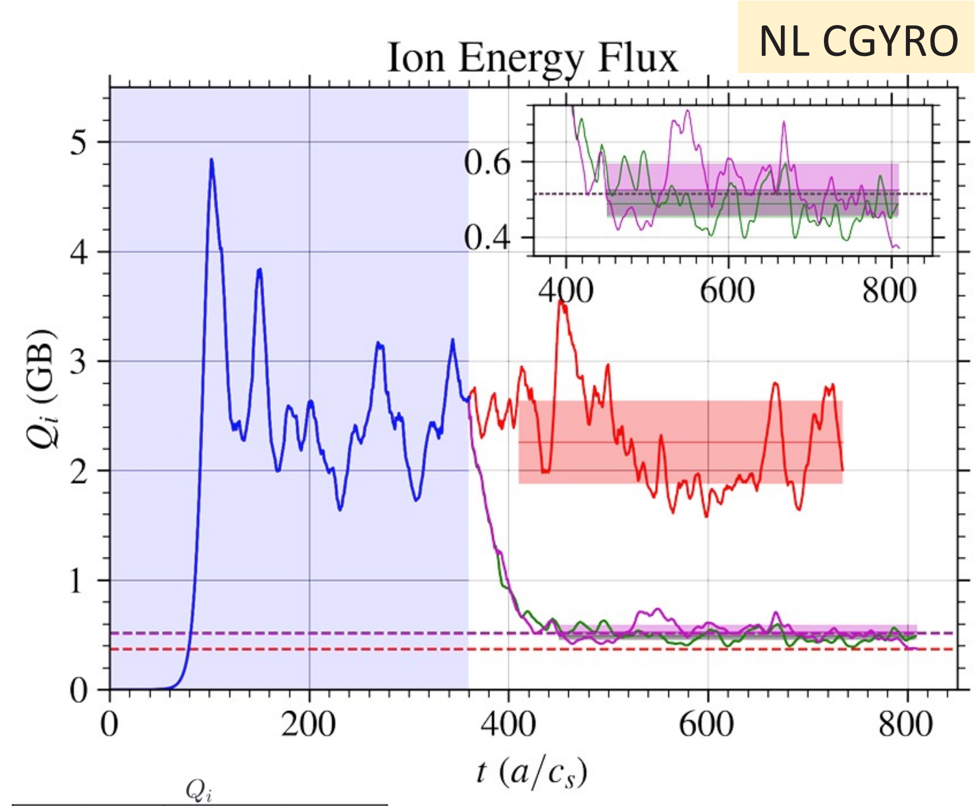


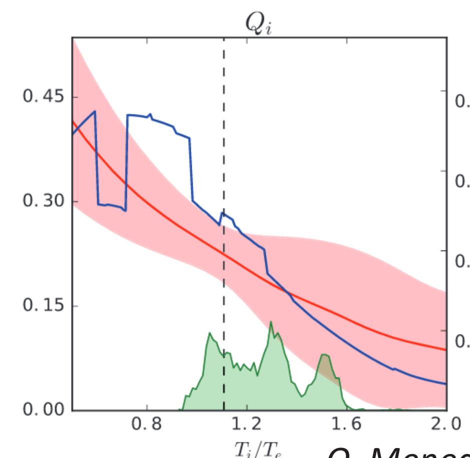
[1] J. Candy JPC 2016
[2] G.M. Staebler PoP 2007

Uncertainty quantification identified as key contributor to high efficiency

- Use of **uncertainties** from limited time averaging (initial value solvers) or discontinuities (reduced quasilinear models) is identified as key contributor to PORTALS efficiency.





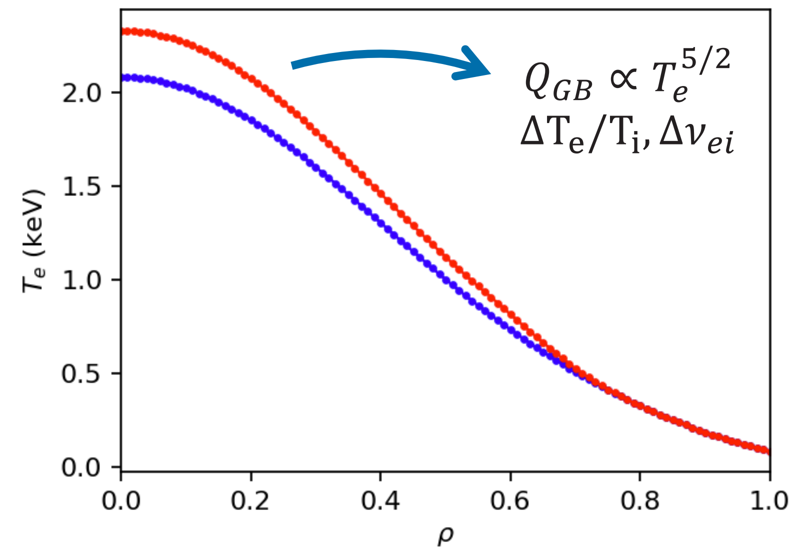


- Easier convergence with smooth models

O. Meneghini NF 2017

Dimensionality reduction becomes very important to build surrogates

- Surrogate training simplified by **dimensionality reduction** (local simulations) and de-coupled from high-dimensional flux-matching problem.



Flux-matching problem

$$a/L_{Te}^{r_1},\quad a/L_{Te}^{r_2},\quad a/L_{Te}^{r_3}\ldots$$
$$a/L_{Ti}^{r_1},\quad a/L_{Ti}^{r_2},\quad a/L_{Ti}^{r_3}\ldots$$
$$a/L_{ne}^{r_1},\quad a/L_{ne}^{r_2},\quad a/L_{ne}^{r_3}\ldots$$

$N_c \times N_r$

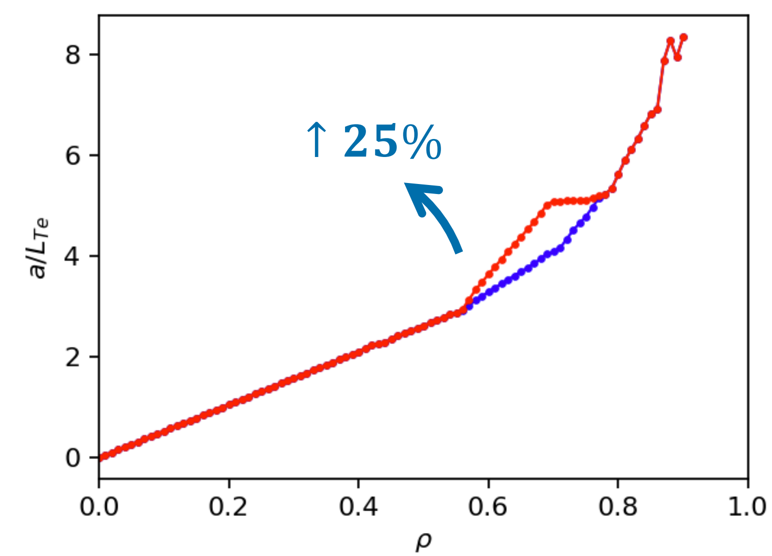
Surrogate training (at r_1)

$$a/L_{Te}^{r_1}, T_e^{r_1}$$
$$a/L_{Ti}^{r_1}, T_i^{r_1}$$
$$a/L_{ne}^{r_1}, n_e^{r_1}$$

$N_c \times 2$

Turbulence-guided variables

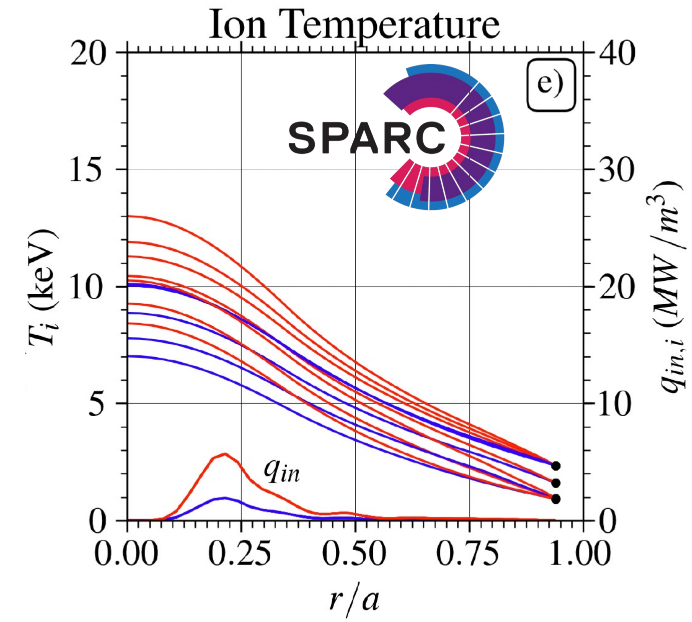
$$a/L_{Te}^{r_1}, \beta_e^{r_1}$$
$$a/L_{Ti}^{r_1}, (T_i/T_e)^{r_1}$$
$$a/L_{ne}^{r_1}, \nu_{ei}^{r_1}$$

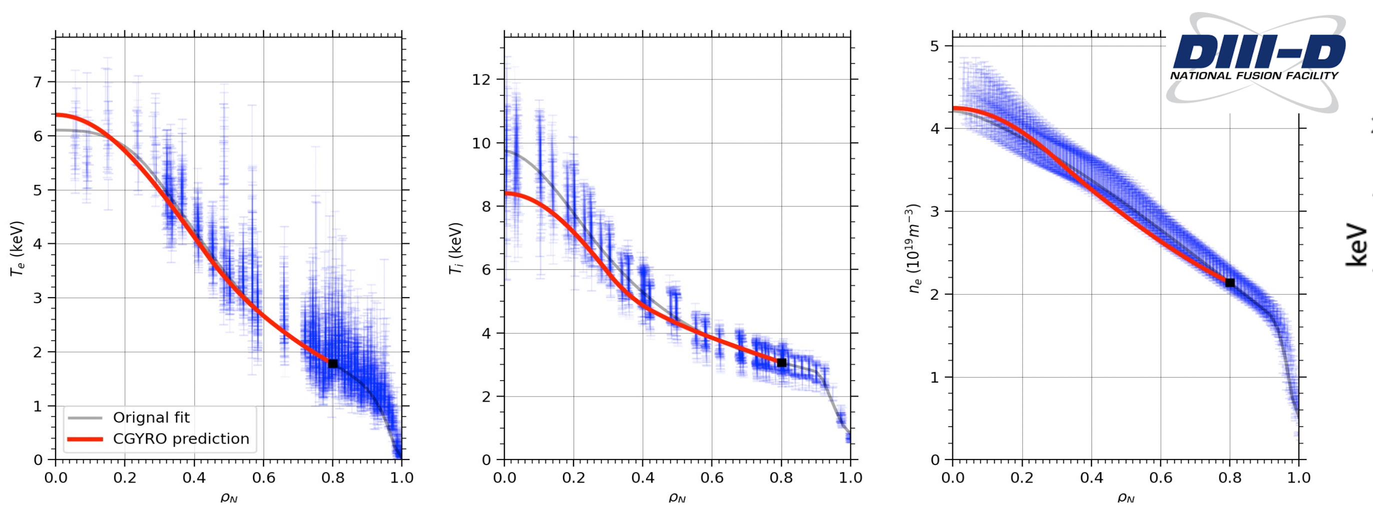


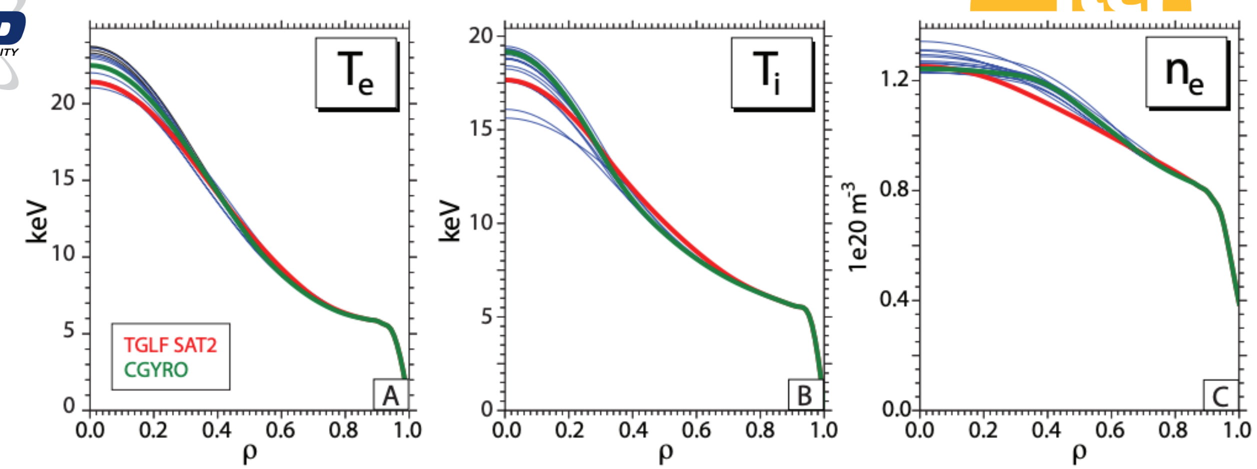
PORTALS-CGYRO has been used to simulate many plasmas and tokamaks

Methods + SPARC H-mode: P. Rodriguez-Fernandez et al Nucl. Fusion 62 076036 (2022).
JET: P. Rodriguez-Fernandez et al invited talk at EPS (2023).
ARC: C. Holland et al invited talk at APS (2023).
DIII-D: N. T. Howard et al Phys. Plasmas 31, 032501 (2024).
Methods + DIII-D: P. Rodriguez-Fernandez et al Nucl. Fusion 64 076034 (2024).
SPARC L-mode: P. Rodriguez-Fernandez et al Phys. Plasmas 31, 062501 (2024).
ITER: N. T. Howard et al Nucl. Fusion 65 016002 (2025).
AUG: R. Bielajew et al submitted (2025).

(+ simulations/papers in progress to study density peaking,
ITER W-wall, impurity effects in SPARC, impurity transport in DIII-D)



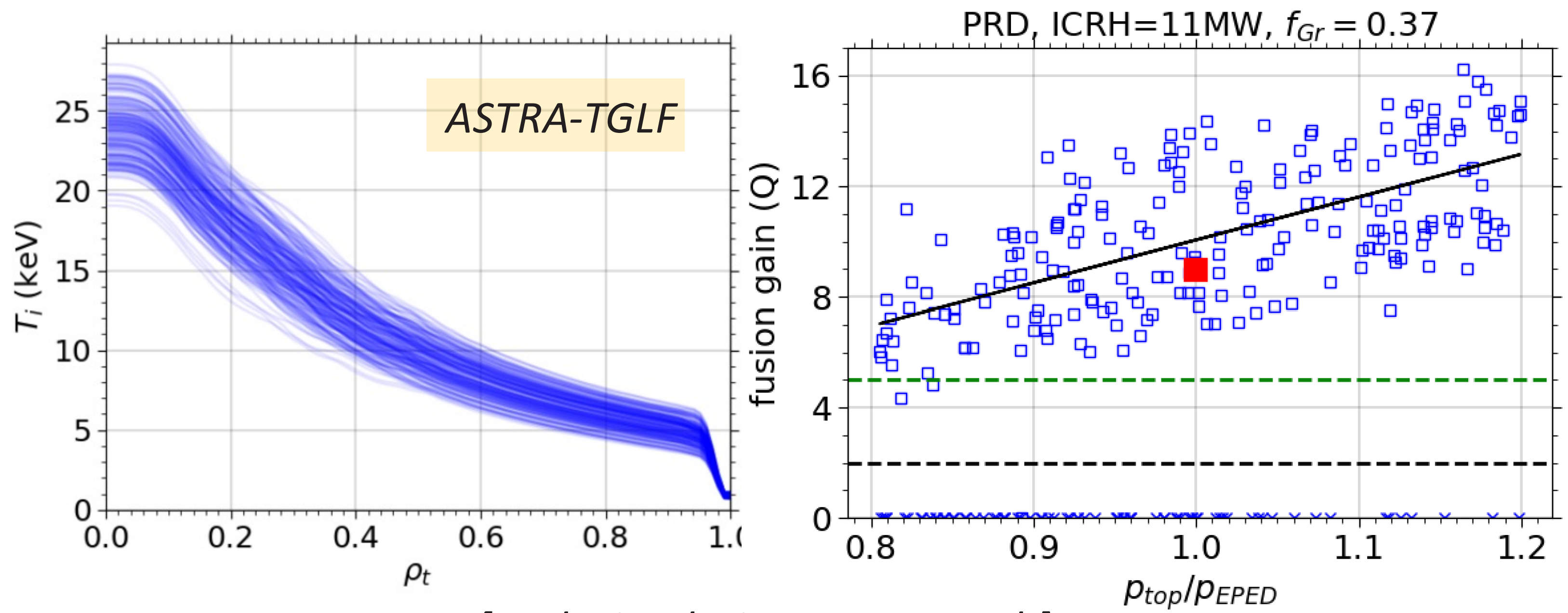




Can PORTALS become a go-to solution for quasilinear models?

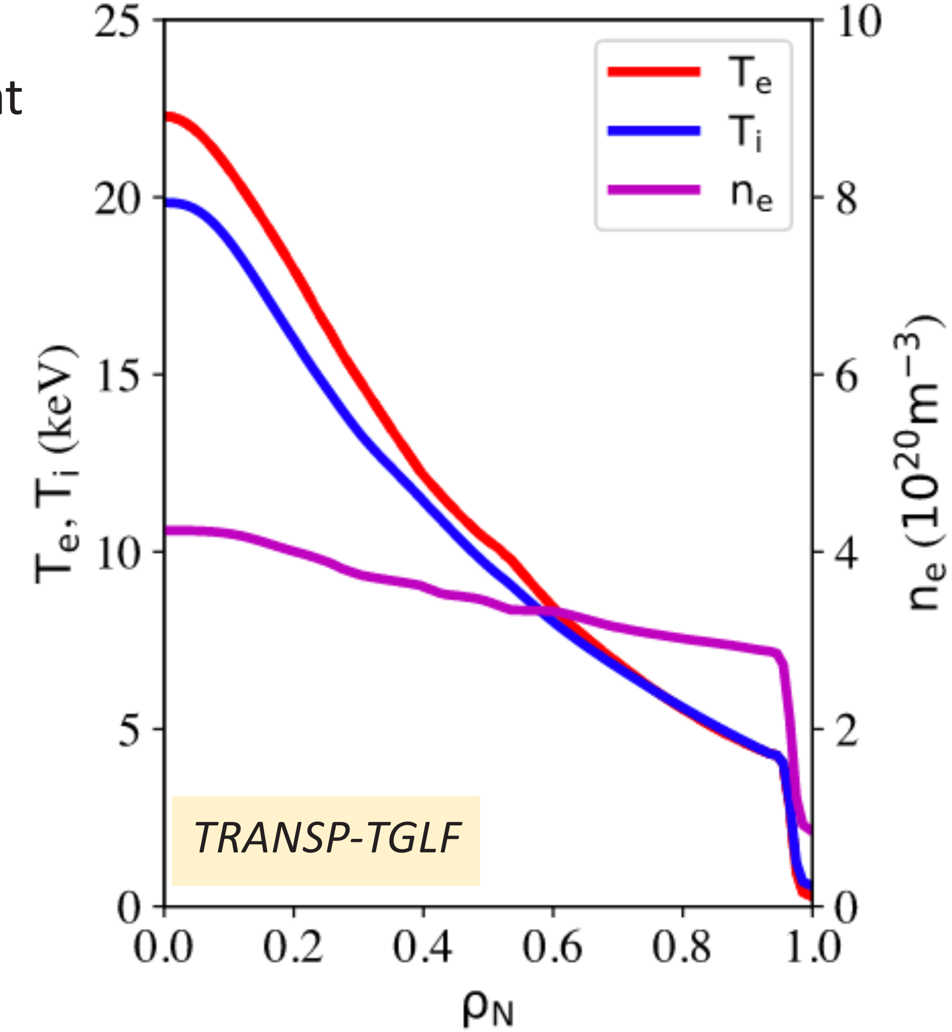
- PORTALS was designed to be efficient at flux-matching with expensive solvers (nonlinear gyrokinetics).
- Solvers with quasilinear models (e.g. TGLF, QuaLiKiz) remain somewhat expensive and require fine tuning.

M. Muraca arXiv 2025



[Each simulation ~100 CPU h]

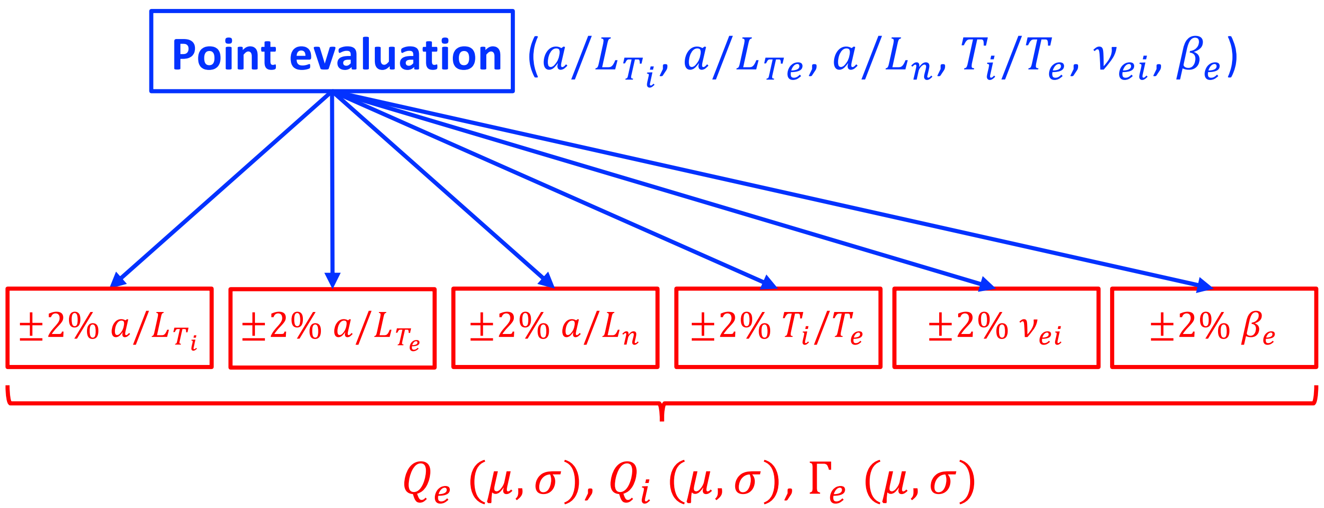
P. Rodriguez-Fernandez JPP 2020

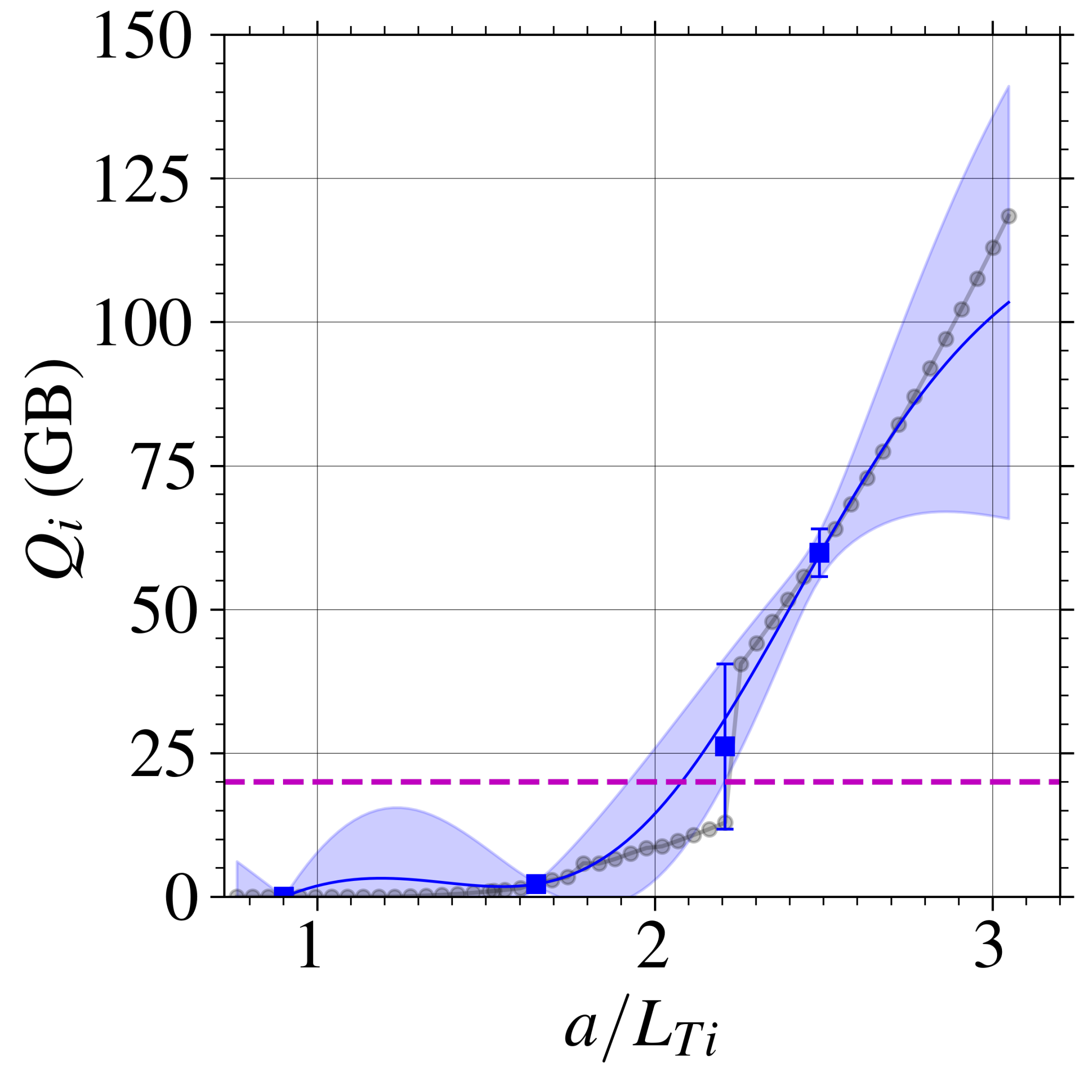


[Each simulation ~2k CPU h]

Development #1: Uncertainty quantification around discontinuities

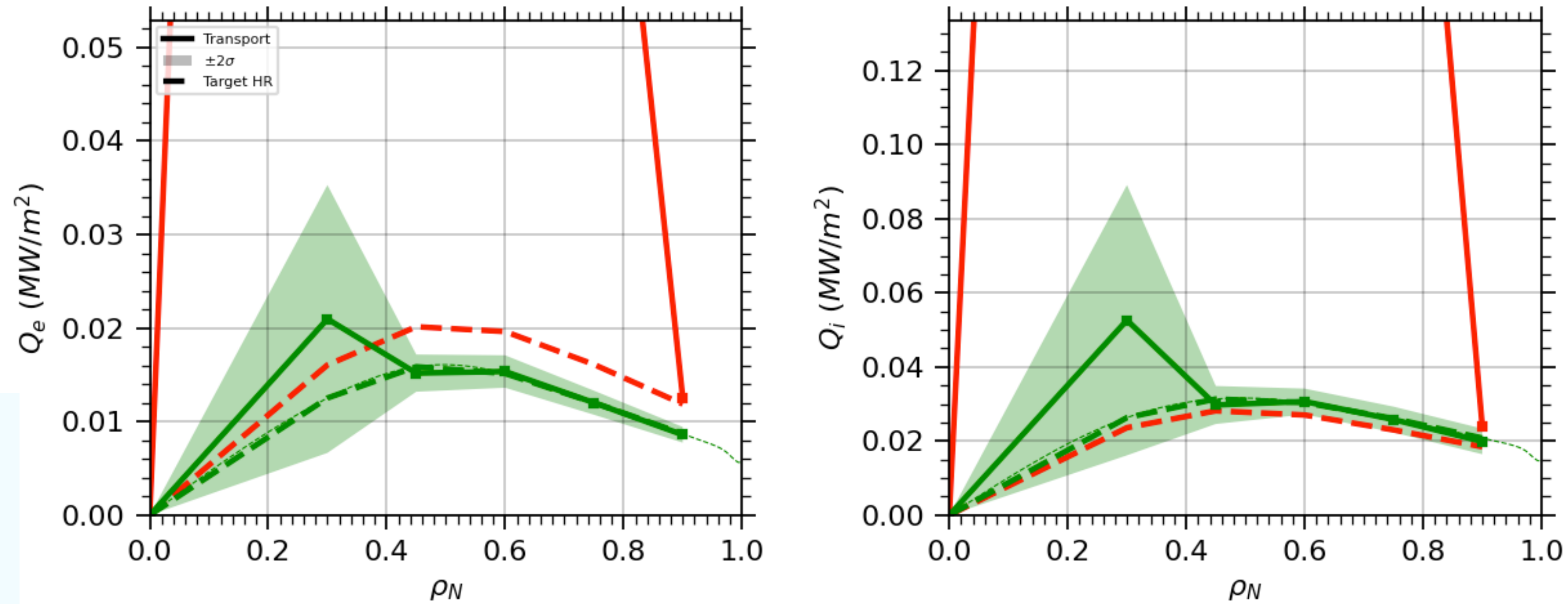
- Using **uncertainty quantification** in surrogate fitting can lead to smoother behavior around discontinuities, marginal stability and stiff transport.
- For each transport flux evaluation, evaluate fluxes response to small variations in inputs (1%-2%) and define mean and standard deviation of evaluation from the flux distribution.





Development #2: Convergence metrics based on flux variance

- Defining proper **convergence metrics** based on flux variance (instead of residuals or gradient variations) can help stop simulations earlier, reducing computational waste.



Ricci validation metric:

$$\chi_R = \frac{\sum R_j S_j}{\sum S_j}$$

$$R_j = \frac{1}{2}\left(1 + \tanh\frac{d_j - d_0}{\lambda}\right), \text{ with } d_j = \sqrt{\frac{(x_j - y_j)^2}{\Delta x_j^2 + \Delta y_j^2}}$$

$$S_j = \exp\left(-\frac{\Delta x_j + \Delta y_j}{|x_j| + |y_j|}\right)$$

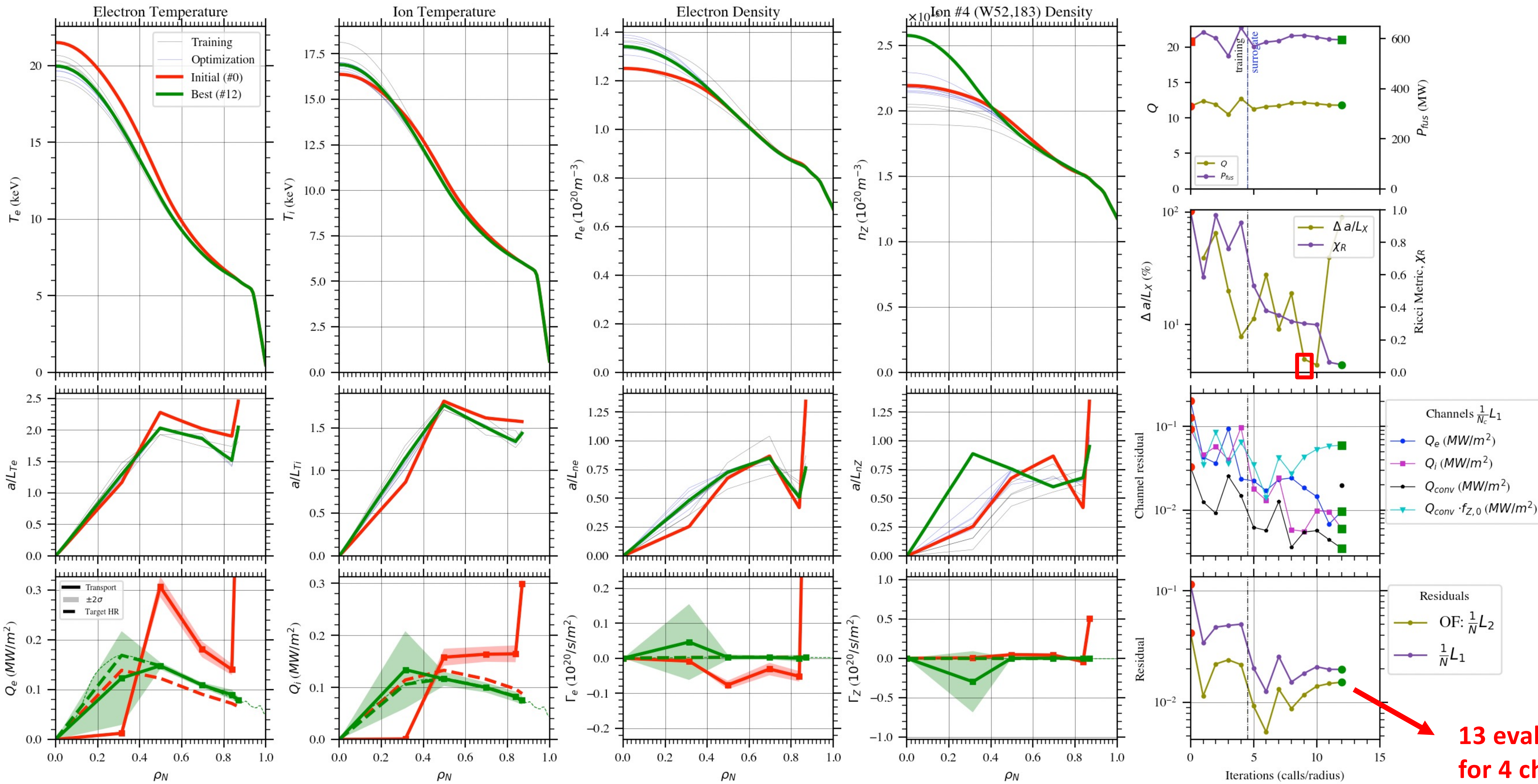
$x_j \rightarrow$ transport flux per channel/radius

$y_j \rightarrow$ target flux per channel/radius

$\Delta \rightarrow$ uncertainty in evaluations

P. Ricci PoP 2011

Development #2: Convergence metrics based on flux variance (example)



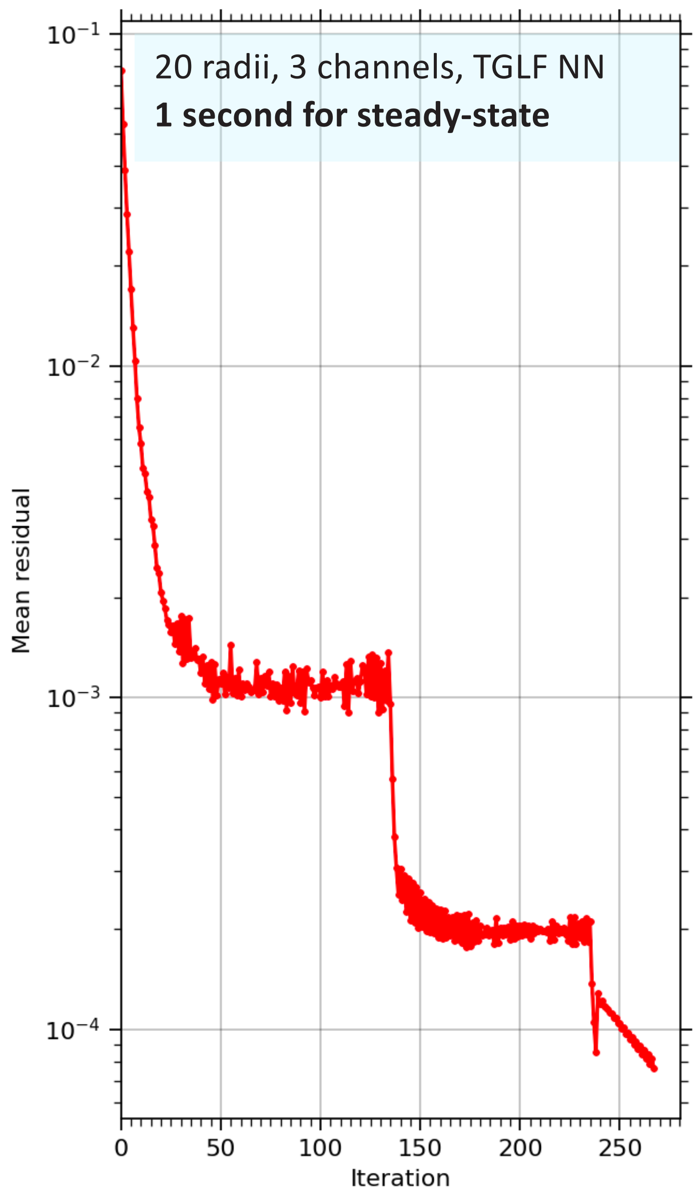
13 evaluations for 4 channels!

Development #3: Physics-guided solvers of acquisition optimization

- Each iteration is expensive ($\mathcal{O}(min)$) because finding the next point requires solving a $\boldsymbol{N_c \times N_r}$ **inputs** with $\boldsymbol{N_c \times N_r}$ **outputs** minimization problem, evaluating $\boldsymbol{N_c \times N_r \times N_K}$ surrogates. *c = channels, r = radii, k = models*
- Standard ML methods to optimize acquisition are too slow, so custom optimization methods are under development.
- Example: A **dynamic simple relaxation** (DSR) method* is proven efficient!

$$z_{j,c}^{(i+1)} = z_{j,c}^{(i)} + \boxed{\eta_{j,c}} \frac{F_{j,c}^{\mathrm{target}} - F_{j,c}^{\mathrm{tr}}}{\sqrt{\left(F_{j,c}^{\mathrm{target}}\right)^2 + \left(F_{j,c}^{\mathrm{tr}}\right)^2}} \cdot \left|z_{j,c}^{(i)}\right|$$

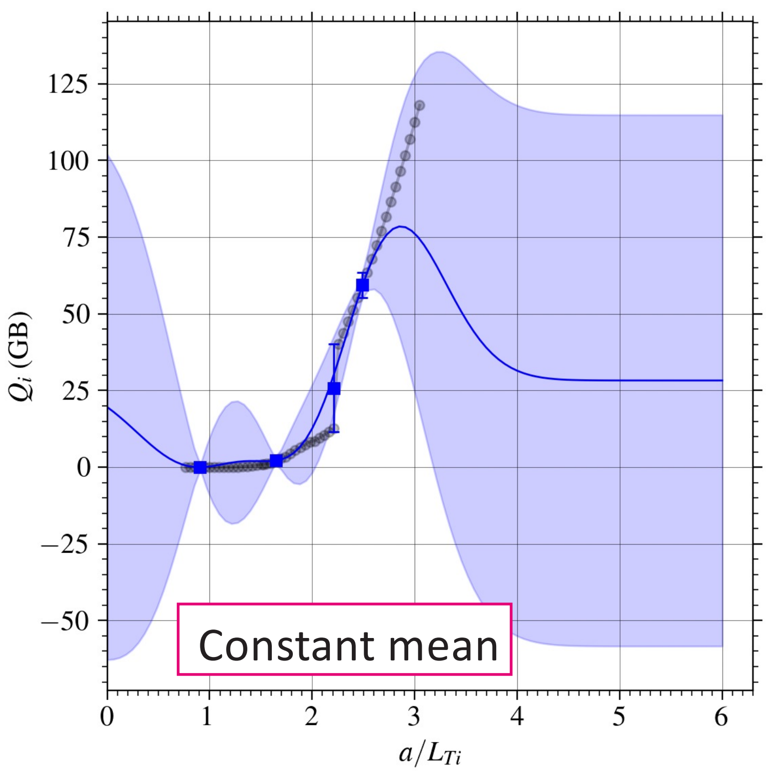
Reduced if residual doesn't improve by custom amount

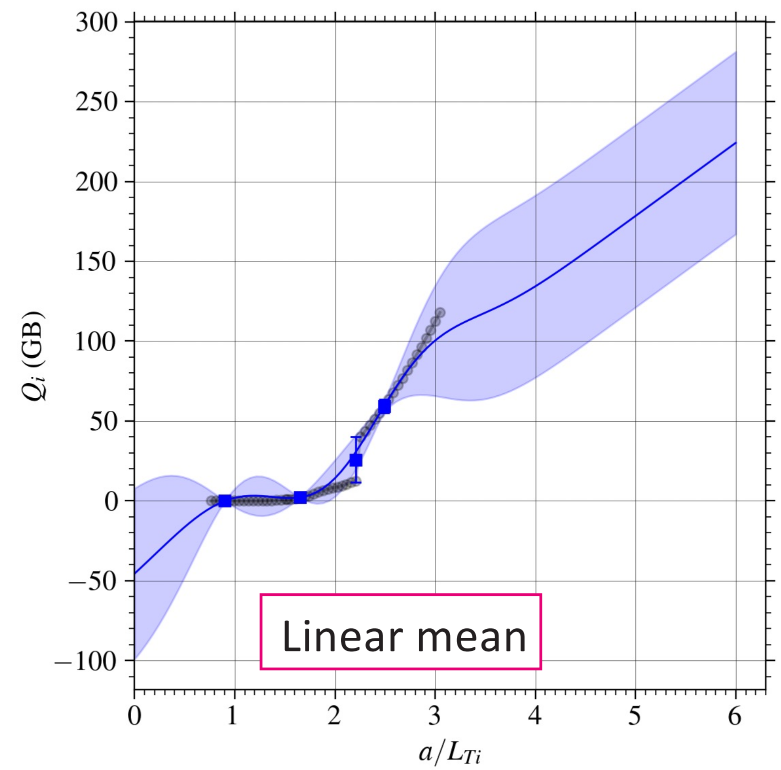


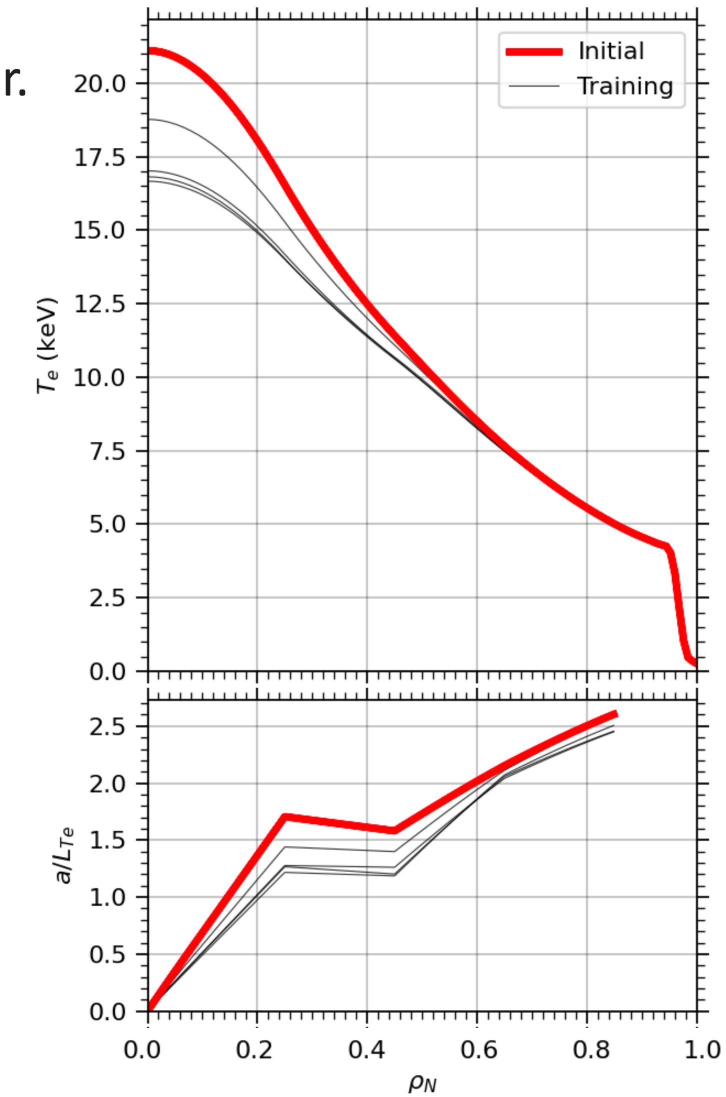
*Inspired by TGYRO method-6 [J. Candy PoP 2009]

Development #3b: Curation of surrogate models is critical

- To help robustness, we enforce a **positive-slope linear-mean** Gaussian Process for diagonal terms of transport matrix → enforcing $Q_i \rightarrow \beta_{>0} \cdot a/L_{Ti}$ far from training ranges.
- Initial random training changed for simple relaxation, leading to better behavior.

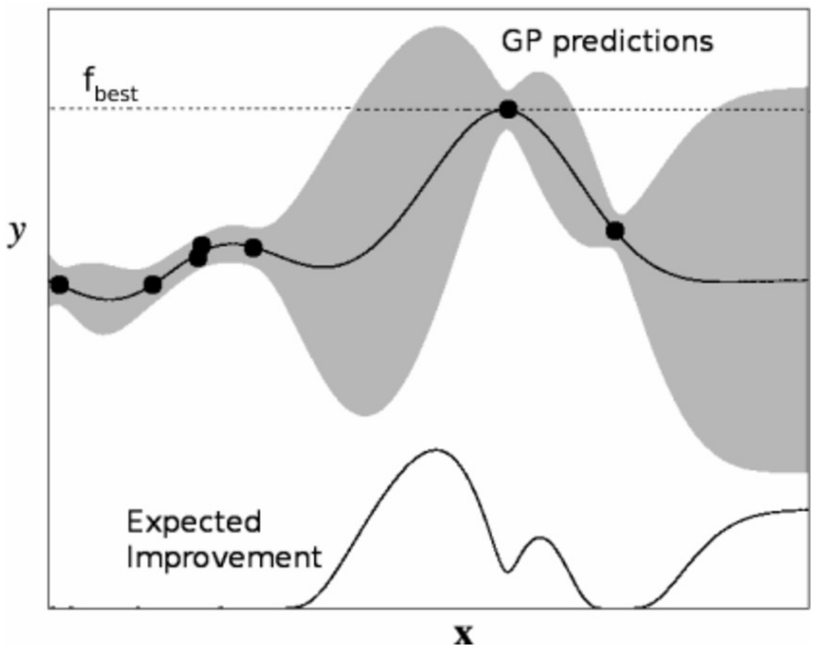




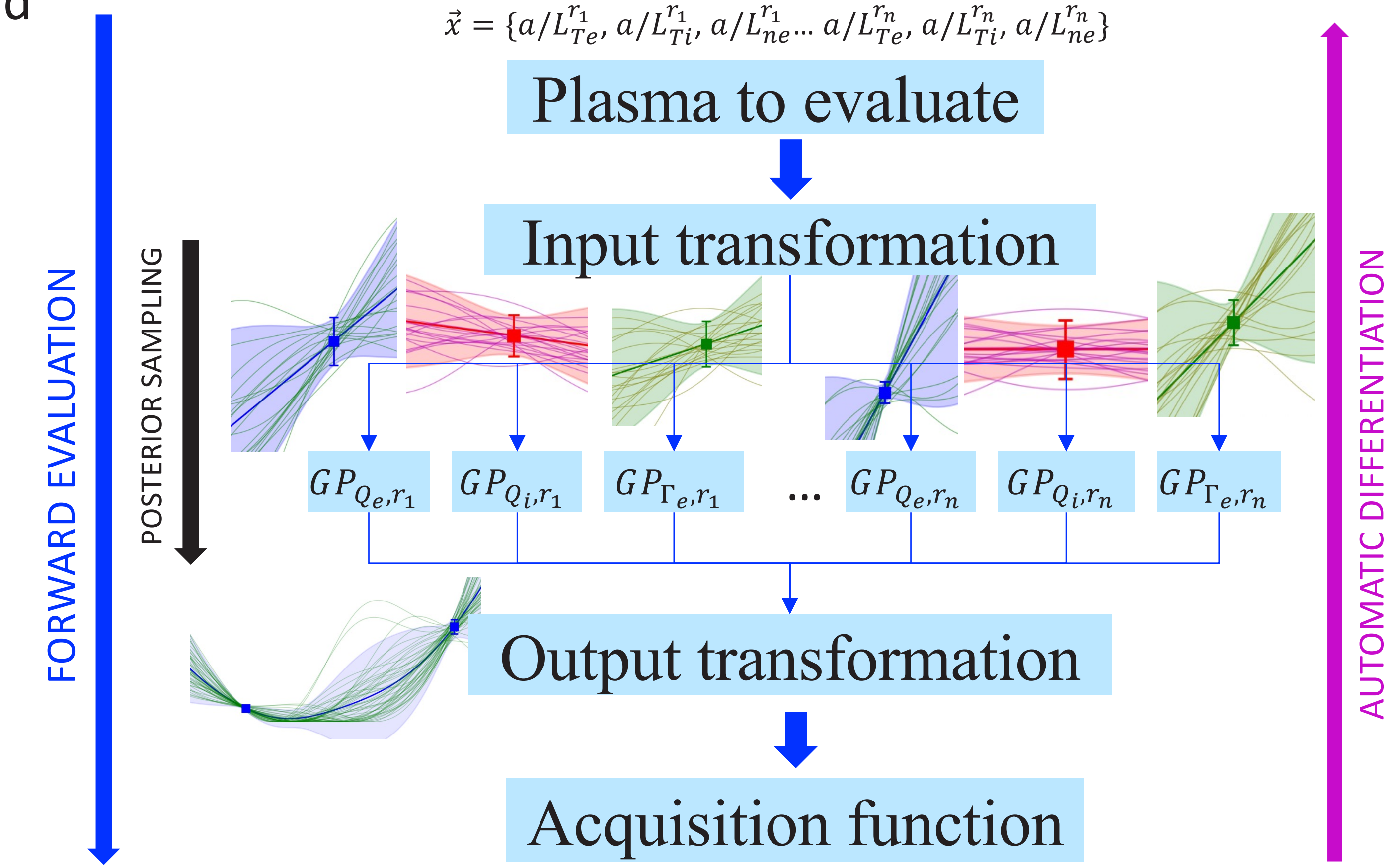


Development #4: Bayesian optimization with exploratory acquisitions

- According to Bayesian Optimization studies, adding **exploration in acquisition function** should reduce number of evaluations for convergence. Methods under investigation.



M. Frean AAI 2008



- Framework fully in PyTorch, enabling automatic differentiation through surrogates!

New frontiers: the SMARTS project

SMARTS
Surrogate Models for Accurate and Rapid Transport Solutions

MIT PSFC

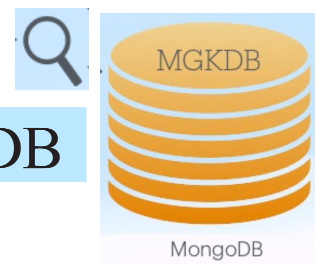
- SMARTS is one of DoE's SciDAC-5 partnerships between FES and ASCR: UCSD (lead), MIT, GA, ORNL, SNL, SapientAI and Kalling. Two examples of PORTALS extensions in SMARTS:

  - Coupling of PORTALS with **gyrokinetic simulation databases** (MGKDB) to help inform new simulations to run.
  - Development of **multi-fidelity** Bayesian optimization approaches to leverage massive cost savings between TGLF and NL CGYRO.

PORTALS

Flux evaluation

GK simulation → MGKDB



HPC

PORTALS

Flux evaluation

arbitrary low cost → BgB model

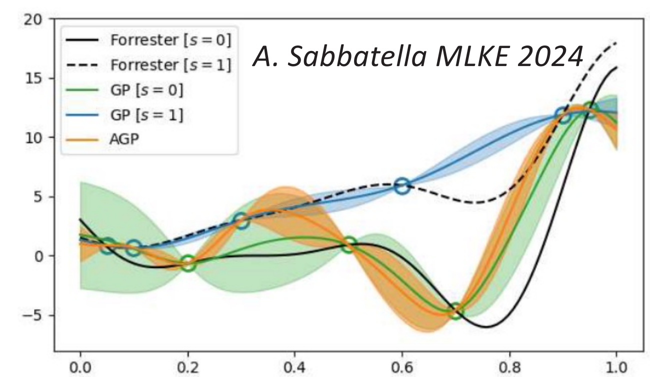
M. Erba NF 1998

~10 CPU sec → TGLF

G. Staebler PoP 2007

~10^2 GPU h → GK simulation

HPC



Thanks to SMARTS team and specially to C. Holland, A. Ho , V. Ayyar, P. Robbe

New frontiers: MAESTRO integrated modeling workflow

MIT PSFC

- MAESTRO (**M**odular and **A**ccelerated **E**ngine for **S**imulation of **T**ransport and **R**eactor **O**ptimization) is an MIT in-house multi-scale integrated modeling workflow inspired by the STEP workflow in OMFIT [O. Meneghini NF 2015].

Engineering parameters

FreeGS → **EPED-NN** / **TRANSP** → **TRANSP** → **EPED-NN** → **PORTALS** → **TRANSP** → **EPED-NN** ↔ **PORTALS** → Self-consistent* core solution

+ $\beta_N = 1.0$
+ fixed core a/LT

EPED-NN

TRANSP

Current diffusion
Grad-Shafranov
Sawtooth

TRANSP

Current diffusion
Grad-Shafranov
Sawtooth
TORIC-FPP
NUBEAM

Effect of new β_N

EPED-NN

PORTALS

Fixed-targets prediction
Soft convergence criterion

Effect of new kinetic profiles on q-profile, ICRF deposition, fast ions

TRANSP

Effect of new β_N (including fast!)

EPED-NN

Last β_N NL interactions

PORTALS

Self-consistent* core solution

*Including:

- Turbulent transport (full TGLF)
- Neoclassical transport (full NEO)
- ICRF deposition (TORIC-FPP)
- Alpha distribution (NUBEAM)
- Power balance, current diffusion
- Pedestal physics (EPED-NN)
- Grad-Shafranov (TEQ)
- Sawtooth trigger and mixing (Porcelli)
- Bootstrap current (Hager)
- Radiation (Bremms, line, sync)

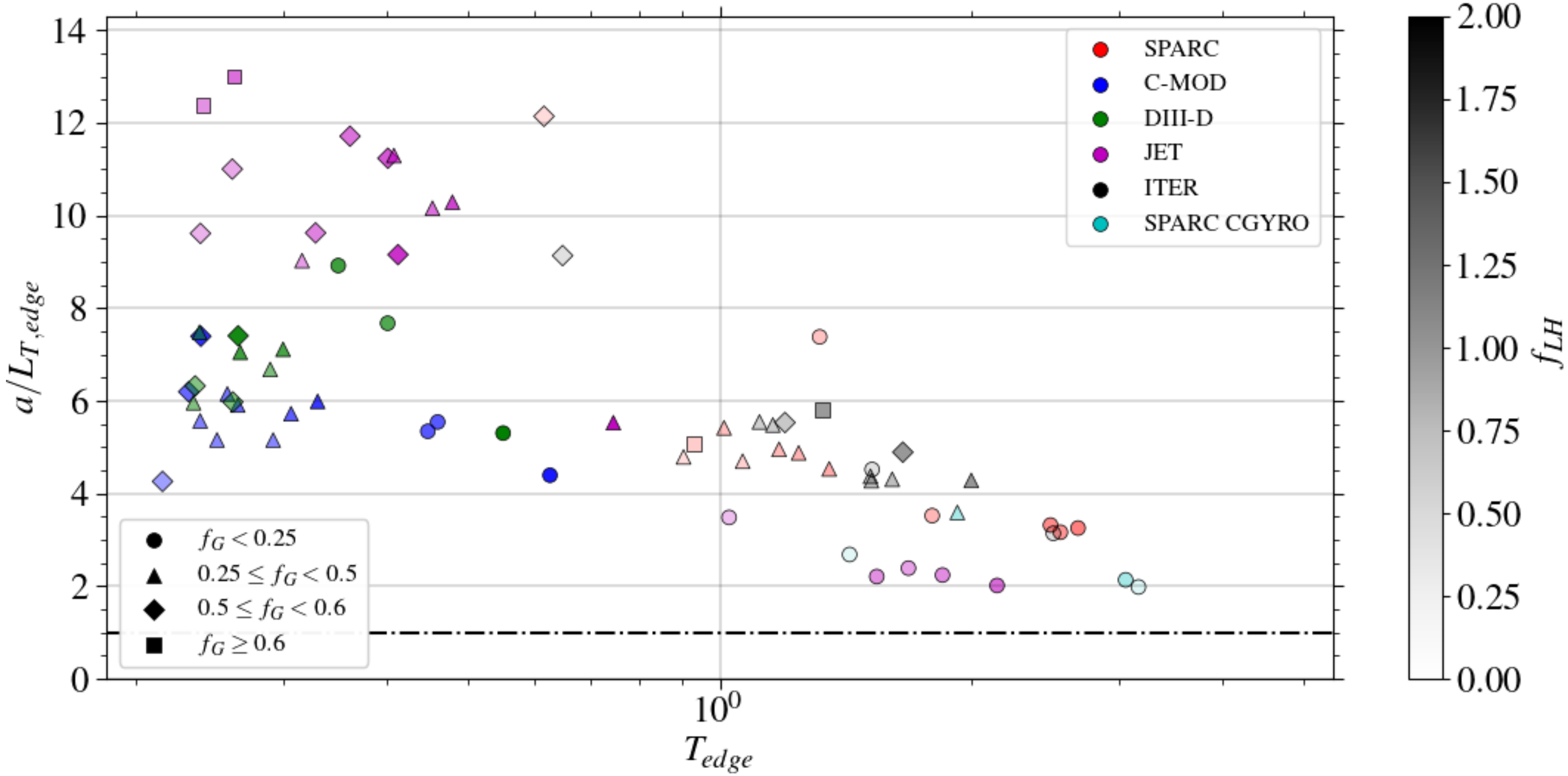
Total cost: ~40 CPU h (& a lot of it is I/O bc of the weak coupling)

ChatGPT 4o (DALL.E) view of a maestro conducting a tokamak

Example of MAESTRO deployment to study parametric dependencies

MIT PSFC

- MAESTRO enables studies with large number of TGLF simulations from engineering parameters to full physics.
- Example of the study of edge boundary condition effect on edge normalized gradients, at $H_{89} = 1$:



Conclusions and next steps

MIT PSFC

- Surrogate-based optimization (BO) designed for expensive black-boxes very suitable for flux-matching problem with both nonlinear gyrokinetics and reduced quasilinear models.
- PORTALS opens new pathways to *1) predict performance in future reactors and 2) study turbulence in current experiments* **with first-principles simulations**.
- Development of uncertainty quantification techniques, physics-guided acquisition optimization solvers and stopping criteria can make PORTALS the *go-to* tool to predict steady-state scenarios.

Want to use PORTALS?
Get MITIM-fusion repo!

bitly

- Next steps:
  - **Improve BO aspects**: exploratory acquisition functions, multi-fidelity approaches, reduction of training/optimization overhead.
  - **Extend physics**: multi-channel, multi-species, fast ion transport, ease implementation of other codes.
  - **Study relevant physics**: fast ions, negative triangularity, density peaking, extension to stellarators.