NSTX-U is sponsored by the
U.S. Department of Energy Office of Science
Fusion Energy Sciences

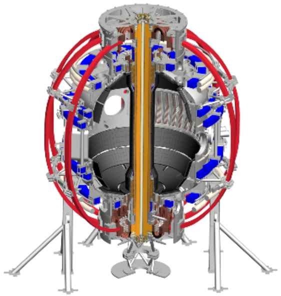
Machine learning accelerated models for scenario optimization on NSTX-U

Dan Boyer[1], Stan Kaye[1], Jason Chadwick[2]

28th IAEA Fusion Energy Conference
Remote e-conference
10-15 May, 2021

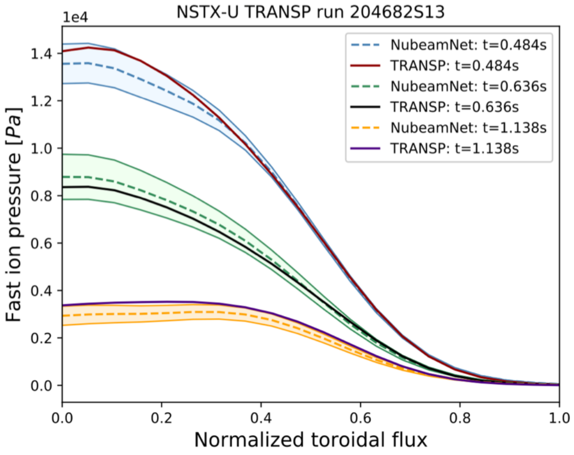
1) PPPL

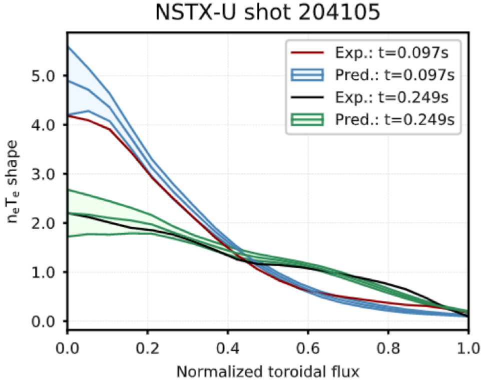
2) Carnegie Mellon University



Understanding plasma at reactor relevant conditions requires multi-billion dollar devices - *how do we make the most of this investment?*

- Experimental time is limited and operators must ensure machine and personnel safety
  - Shot development, control commissioning, parameter scans, between shot decisions by operators: expensive and limited
- Predict first: use integrated models (e.g., TRANSP) to develop experiments.
  - Predictive models are steadily improving but not complete
  - Can take hours/days to run predictive models for whole discharges
    - Need to have rapid optimization capabilities to respond to situations as they occur between/during shots
- **Proposed solution:** Scenario optimization using machine learning models
  1. Accelerate well-validated models with ML
  2. ML empirical models where physics models are lacking
  3. Genetic algorithms for optimization of actuator trajectories





PPPL's TRANSP code integrates numerous physics modules to enable scenario prediction

- Used for both analysis of experiments and prediction
- Ongoing development to improve physics fidelity and validate modules
- NSTX-U predictions have been used to propose flattop scenarios and ramp-up trajectories for NSTX-U

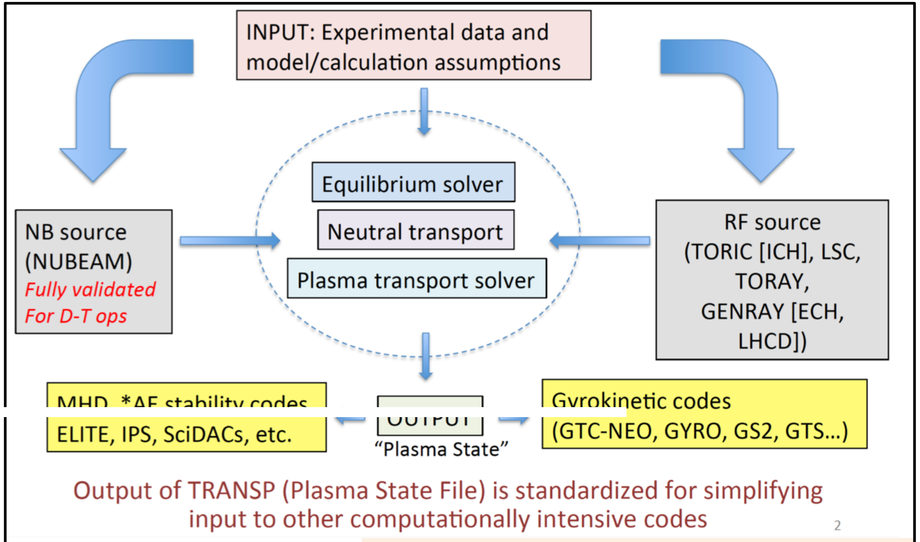
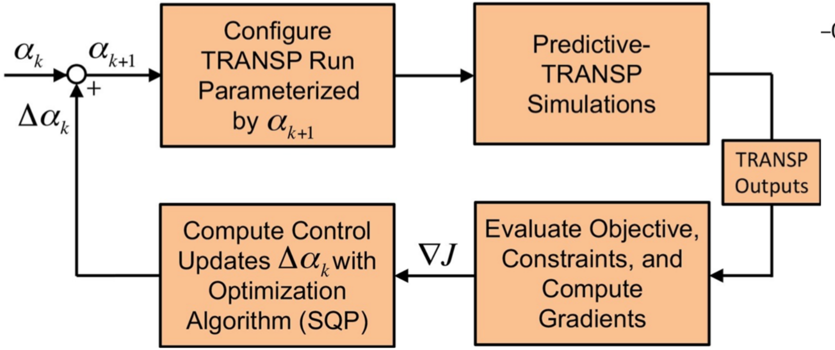
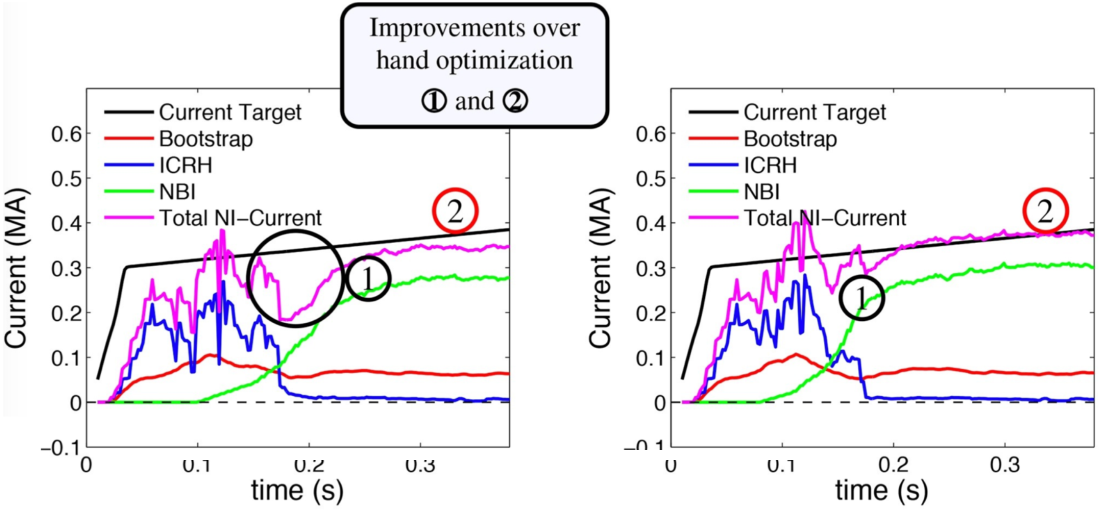


Figure from S. Kaye, TRANSP User Group Meeting 2015

Numerical optimization algorithms have been used to automate trajectory design with TRANSP

- Reduces manual time required by physicists
- Enables improved trajectory design
- Still, expensive in terms of computation time: ~1 week to converged optimal trajectory





(a) By-hand optimization.

(b) TRANSP-based optimization.

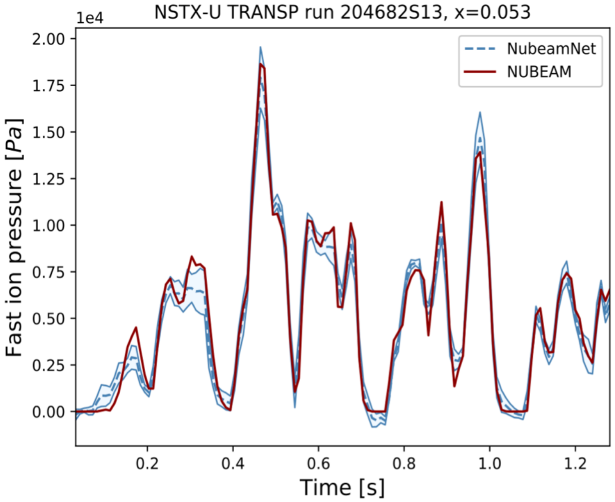
Wehner, et al., *Fusion Eng. Des.* 146, 2019

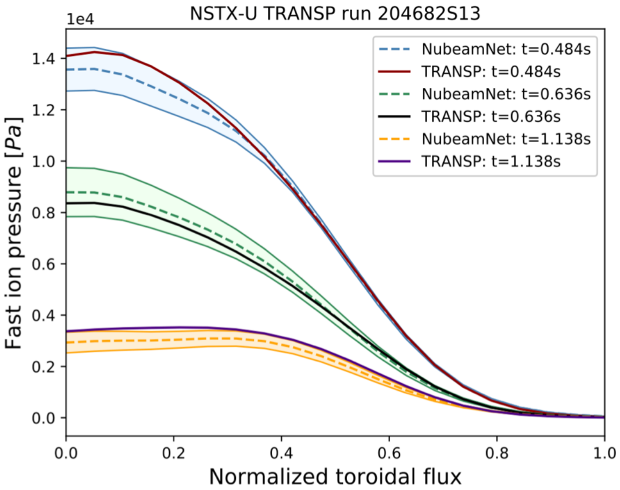
(1) Machine learning enables **faster versions of physics models** for optimization and control

- Predictive TRANSP simulations can take hours per simulation second
- **NUBEAM** is a Monte Carlo code that calculates the effect of neutral beams on the plasma (heating, current drive, torque)
  - Often takes >30% of calculation time
- Basic machine learning approaches enable the development of **NubeamNet**

Boyer et al., *Nuclear Fusion* 2019

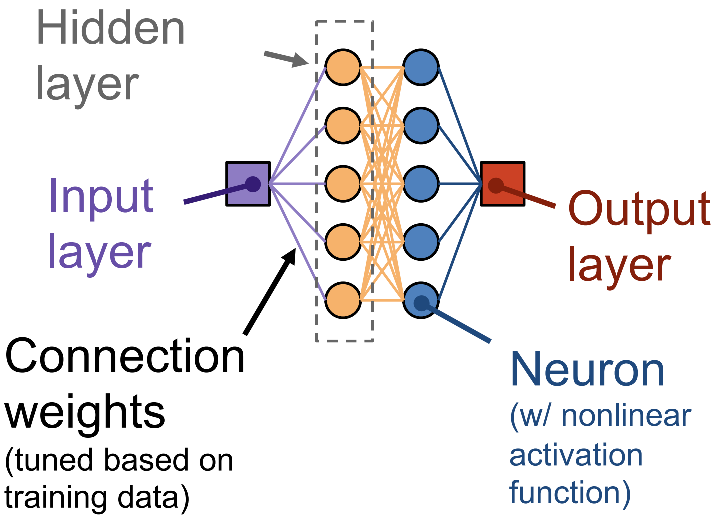
Calculation of beam effects (at 5ms intervals) took **less than 50ms for the entire shot**, compared to minutes to hours for NUBEAM





Orders of magnitude speed increase enabled by neural networks trained on database of NUBEAM results

Neural network - a universal approximator



Neural network model development

1. **Database generation:** NSTX-U TRANSP runs (~2000, ~100 samples per run), including scans of important parameters
2. **Data reduction:** Make the data manageable. e.g., reduce profile data, time history

Dataset

3. Training
Tune connection weights

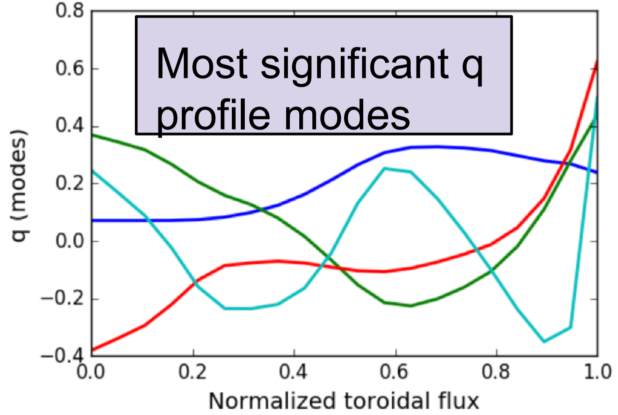
4. Validation
Select model topology

5. Testing
Does the model generalize?

Two of the challenges to machine learning for NUBEAM: spatially distributed data and time history dependence

Profiles projected onto only the most significant modesl reduced number of coefficients used for neural network training.

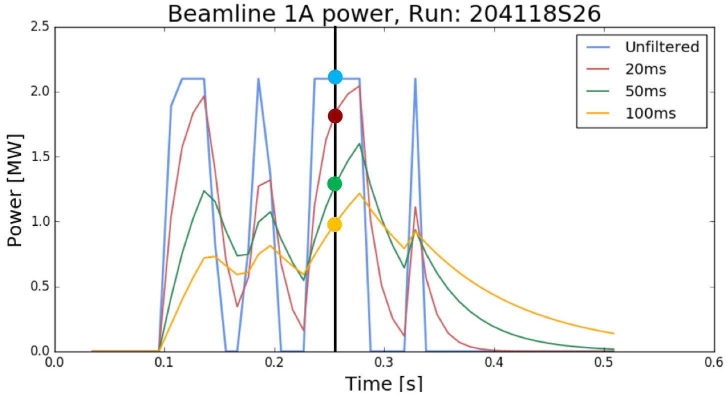
Coef. = Modes · profile
Recon. profile = Modes^T · Coef.



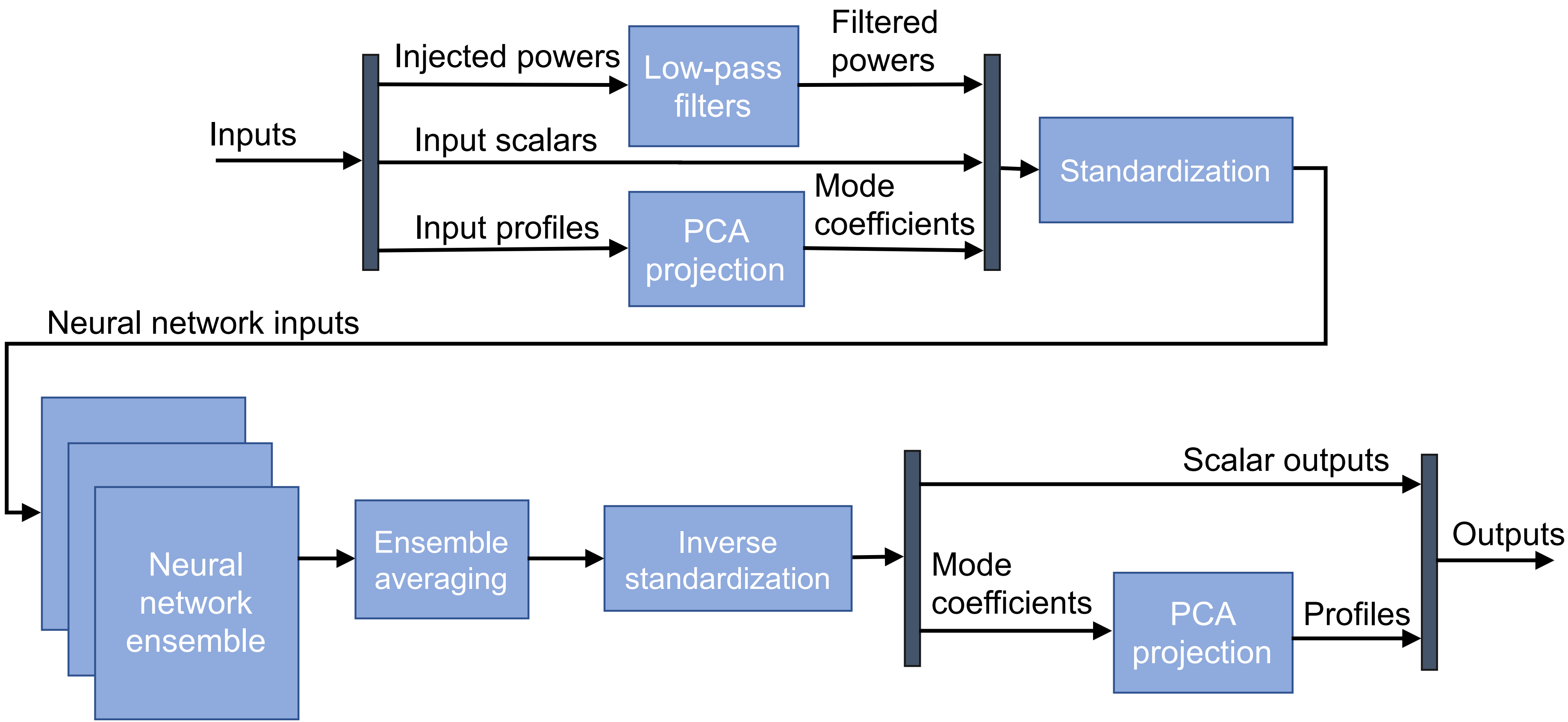
Most significant q profile modes

- Blackbox machine learning would address with **convolutional and/or recurrent neural networks**
- Much simpler approach used here:
  - **Principal component analysis** compresses spatial data and **low-pass filtering** encodes time-history

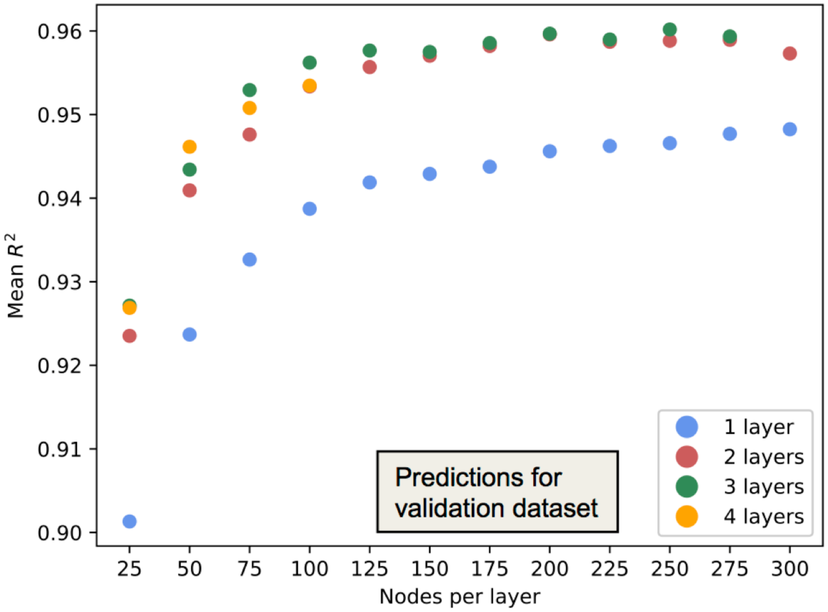
Beam power augmented with low pass filtered versions representing slowing down



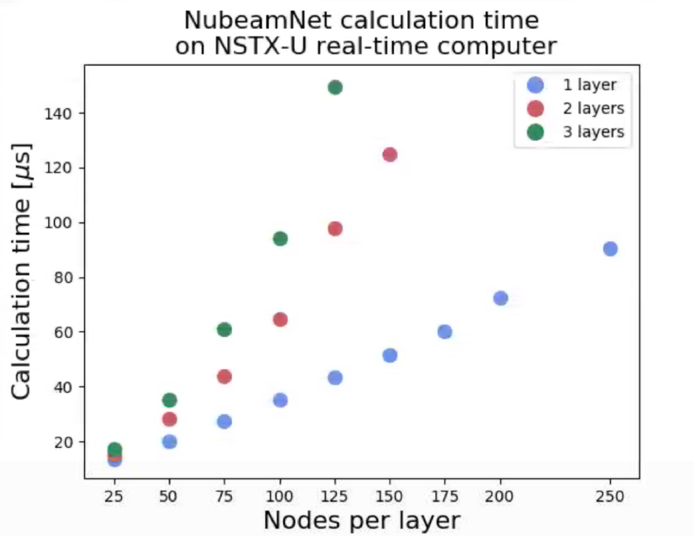
Overview of NUBEAM Neural Network



Validation: The topology of the model must be selected to optimize quality of fit and evaluation time



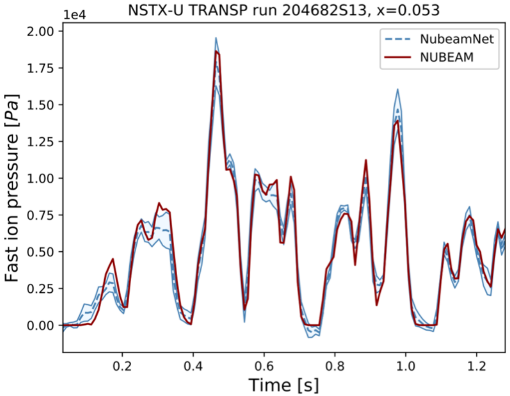
- Neural net code implemented on NSTX-U real-time computer
- Calculation times well within the 200 microsecond control system cycle time, much faster than time scales of interest

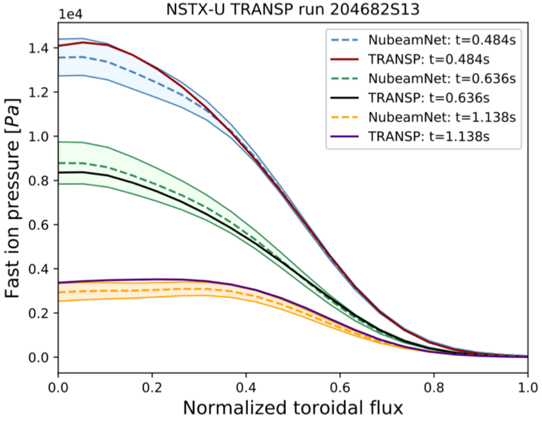


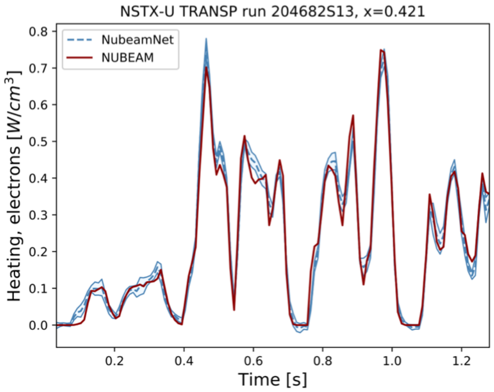
- Significant improvement with more than one layer, but not much benefit in going deeper
- Adding nodes improves fit, but improvement slows or rolls over around 100-125 nodes per layer

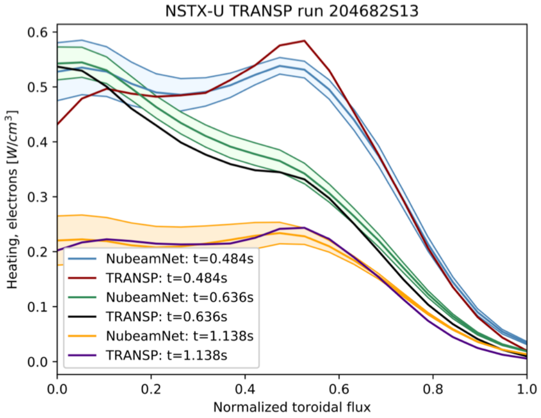
Trained neural network is able to accurately reproduce **time history** and **profiles** in **testing dataset**

- Successfully accelerated a computationally intensive code
- Accuracy and timing indicate the model is **well-suited for real-time applications**
- Promising approach that could be applied to accelerate other physics modules









(2) For phenomena not well described by physics models: **machine learning models from empirical data**

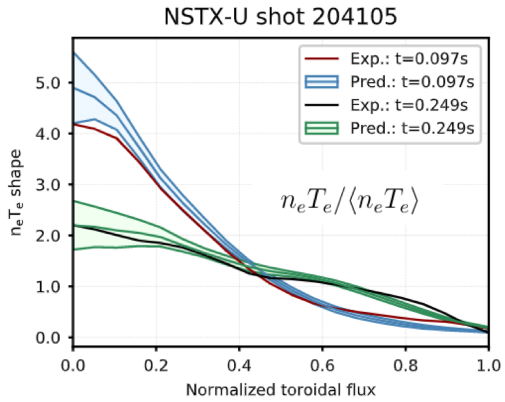
- While ions were found to behave neoclassically on NSTX high collisionality H-modes, electron transport was found to be anomalous and dominant
- Multi-point Thomson scattering data is available for thousands of discharges from NSTX and NSTX-U
- **Goal:** Develop prediction of electron density and pressure profiles suitable for accelerated optimization and real-time control applications
- Considerations:
  - Inputs readily available in real-time and predictable by models
  - Decouple from details of sources, profiles, and time history as much as possible (simplifies training and integration with other models)

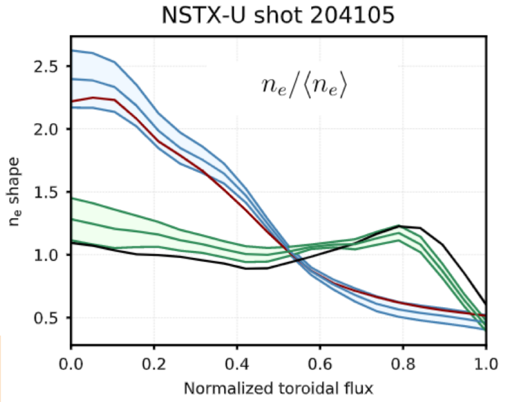
Electron density and pressure profile shape can be well predicted from a small number of scalar parameters

Inputs	
Symbol	**Name**
R_0	Major radius
κ	Elongation
I_p	Plasma current
a	Minor radius
$B_{\phi,v}R$	Vacuum toroidal field
δ_u	Upper triangularity
δ_l	Lower triangularity
$\langle n_e \rangle$	Volume-averaged electron density
$\langle n_e T_e \rangle$	Volume-averaged electron pressure
Outputs	
$n_e/\langle n_e \rangle$	Electron density profile shape
$n_e T_e/\langle n_e T_e \rangle$	Electron pressure profile shape

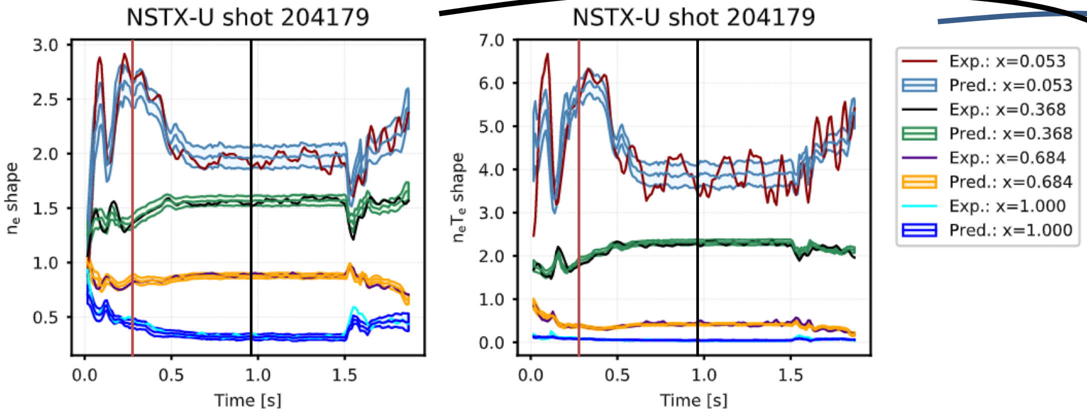
- Prediction of **profiles shapes** (profiles normalized by volume avg.)
  - Volume averages considered measured or predicted by particle balance models
  - No information on sources included
  - **Prediction still quite successful**: indicates stiff, self-organized profiles
- Thomson scattering profile data reduced through principal component analysis

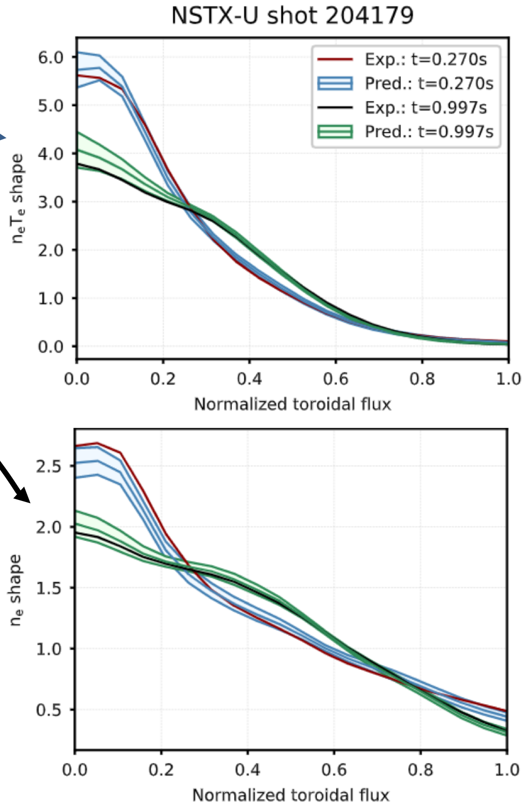
Boyer et al., *Nuclear Fusion* 2021





Test results show time history and profiles are accurately modeled

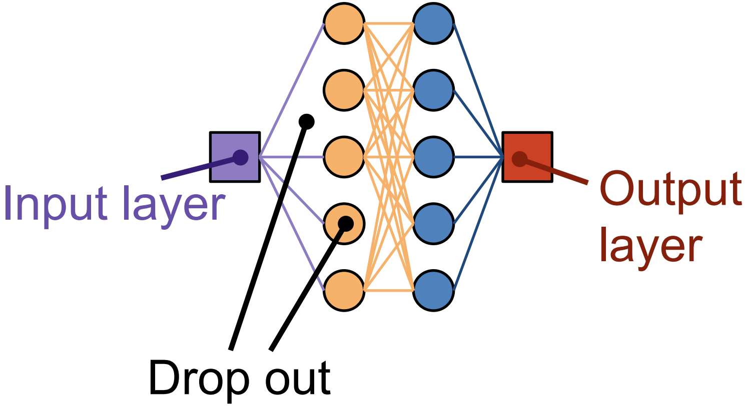


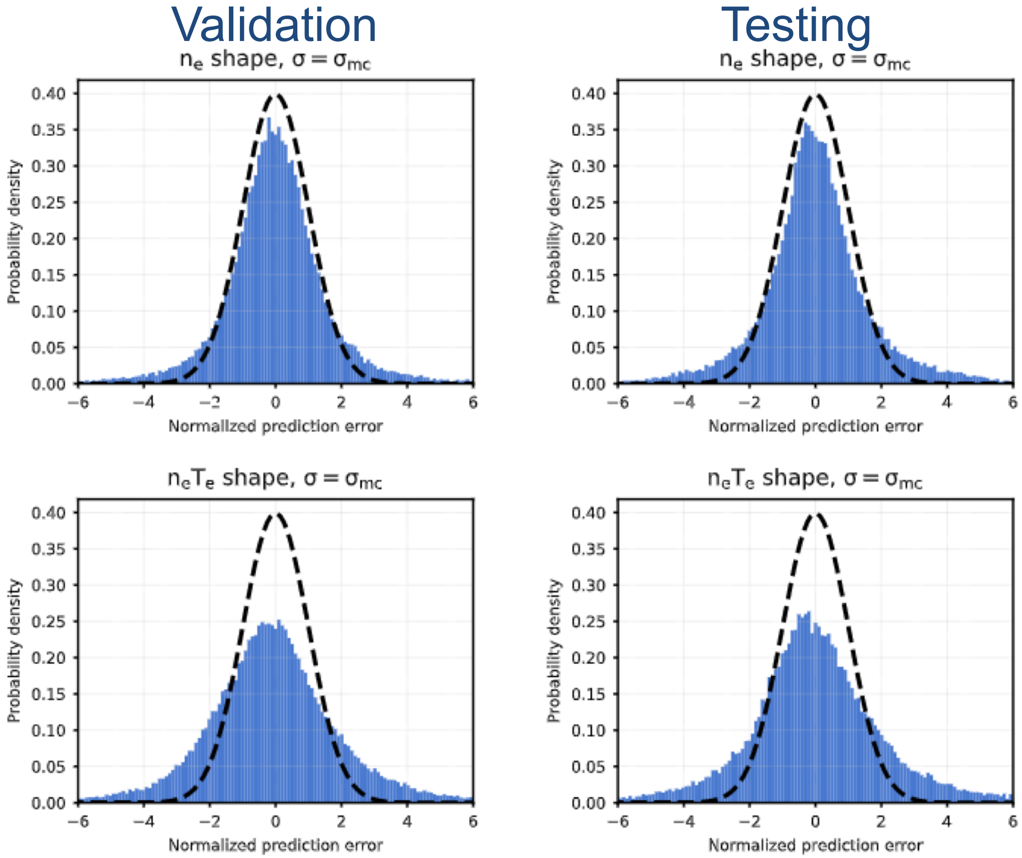


- Time history of changes in profile peaking are captured by model, as well as changes in edge density.
- The model generally averages over noise and fast transients in the core

Uncertainty estimated from standard deviation of ensemble of models. Ensemble created from Monte Carlo dropout works well.

- Generating an ensemble of perturbed models through **Monte Carlo dropout** provides a useful estimate of uncertainty that generalizes well

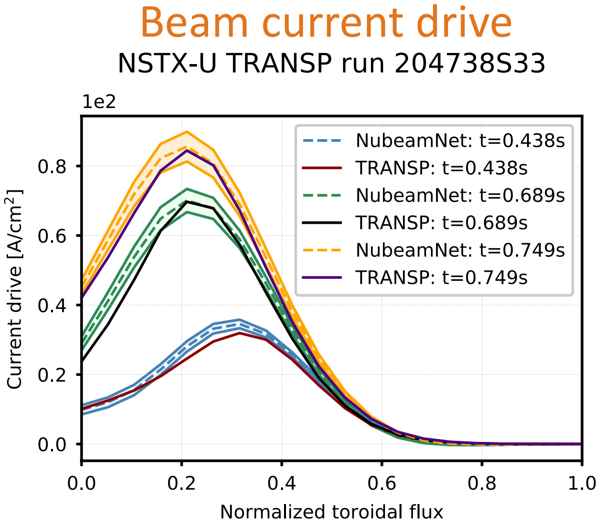


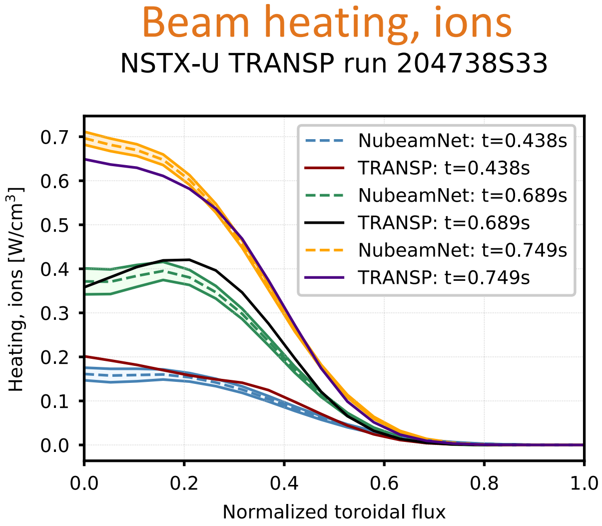


(1+2) Initial combination of machine learning & reduced physics models for faster prediction

- Neutral beams
- Density
- Temperature
- Current profile
- Rotation

- NubeamNet prediction of heating, current drive, torque, neutron rate, etc.



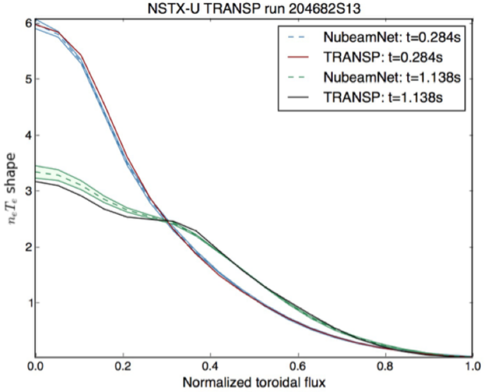


(1+2) Initial combination of machine learning & reduced physics models for faster prediction

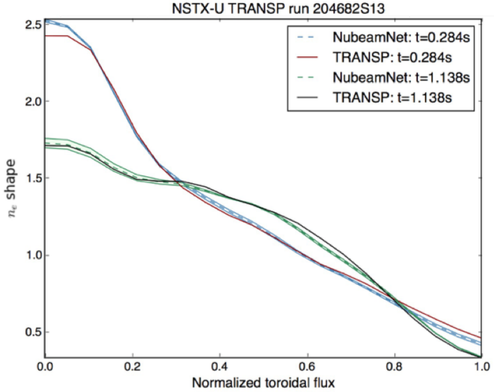
- Neutral beams
- Density
- Temperature
- Current profile
- Rotation

- 0D particle and energy balance, ITER confinement scaling
  - Free parameters: **Greenwald fraction, H98y2**
- Electron temperature and density profile shapes from neural network
  - Trained on empirical data (electron transport anomalous on NSTX)
- Ion density from quasi-neutrality, assumed Zeff
  - Free parameters: **Impurity species, Zeff**
- Ion temperature assumed to be multiple of electron temperature
  - Free parameter: **scale factor**

Electron pressure (shape)



Electron density (shape)

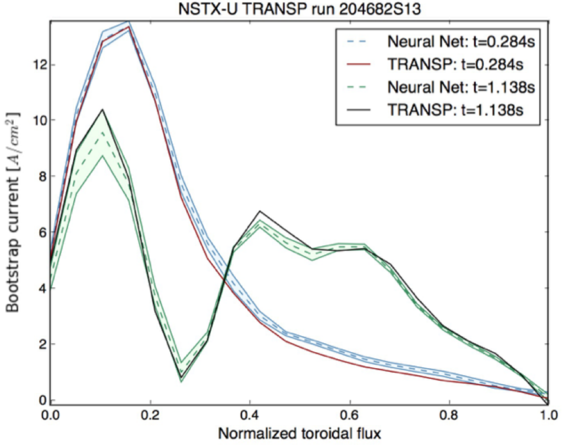


(1+2) Initial combination of machine learning & reduced physics models for faster prediction

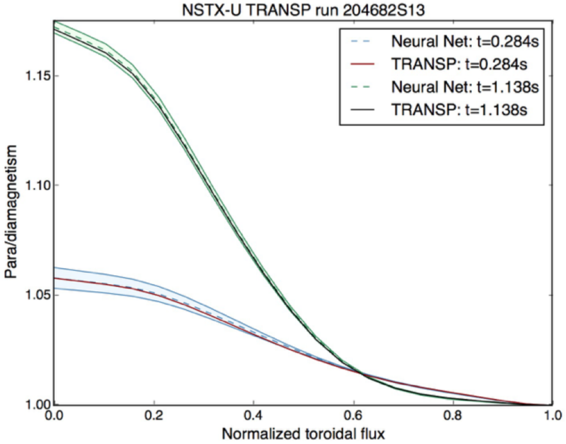
- Neutral beams
- Density
- Temperature
- Current profile
- Rotation

- Magnetic/momentum diffusion equation, semi-implicit solver
  - Extends simplified version used in previous control modeling efforts
- Neural networks trained on TRANSP runs provide PDE parameters
  - Equilibrium dependent geometric parameters
  - Bootstrap current and resistivity profiles

Bootstrap current

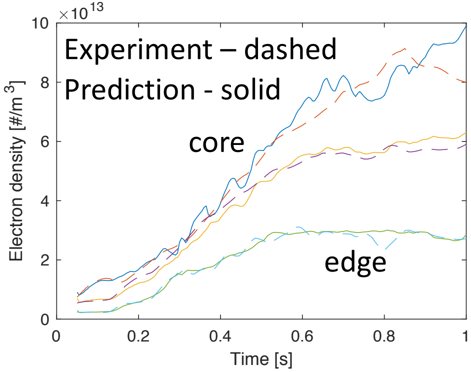


Para/diamagnetism

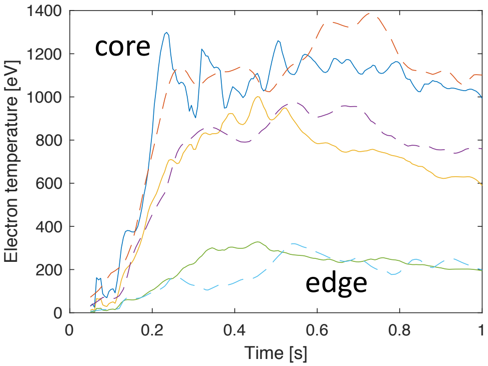


(1+2) Initial combination of machine learning & reduced physics models for faster prediction

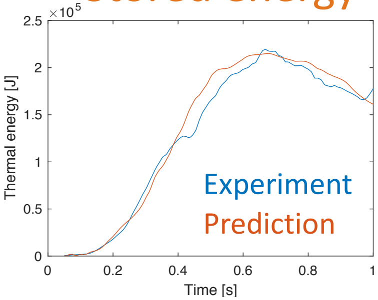
Electron density



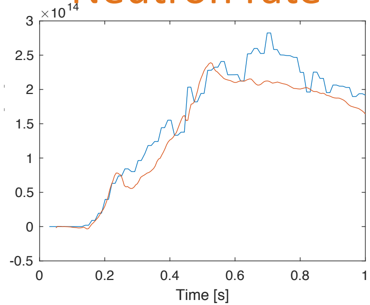
Electron temperature

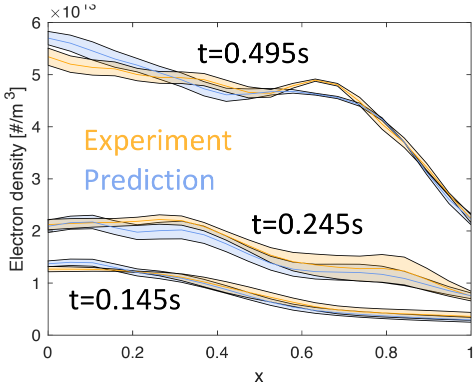


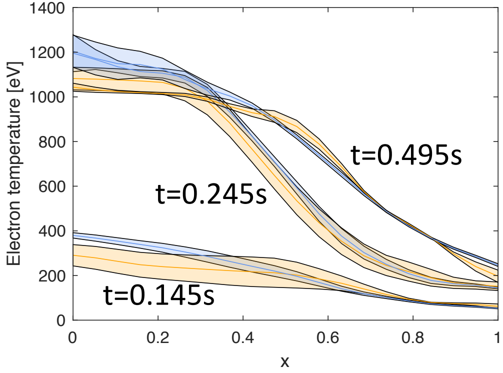
Stored energy

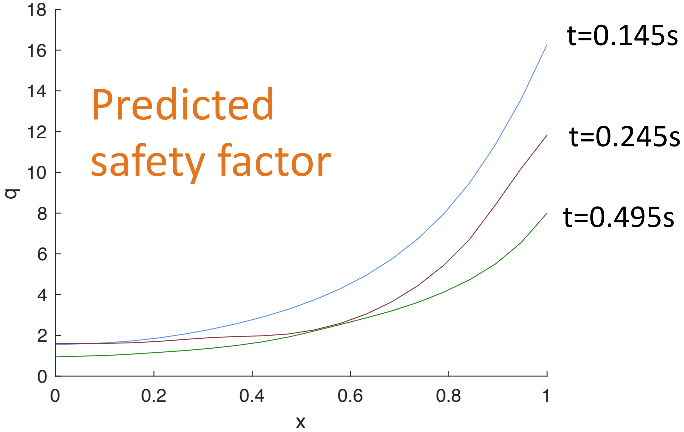


Neutron rate



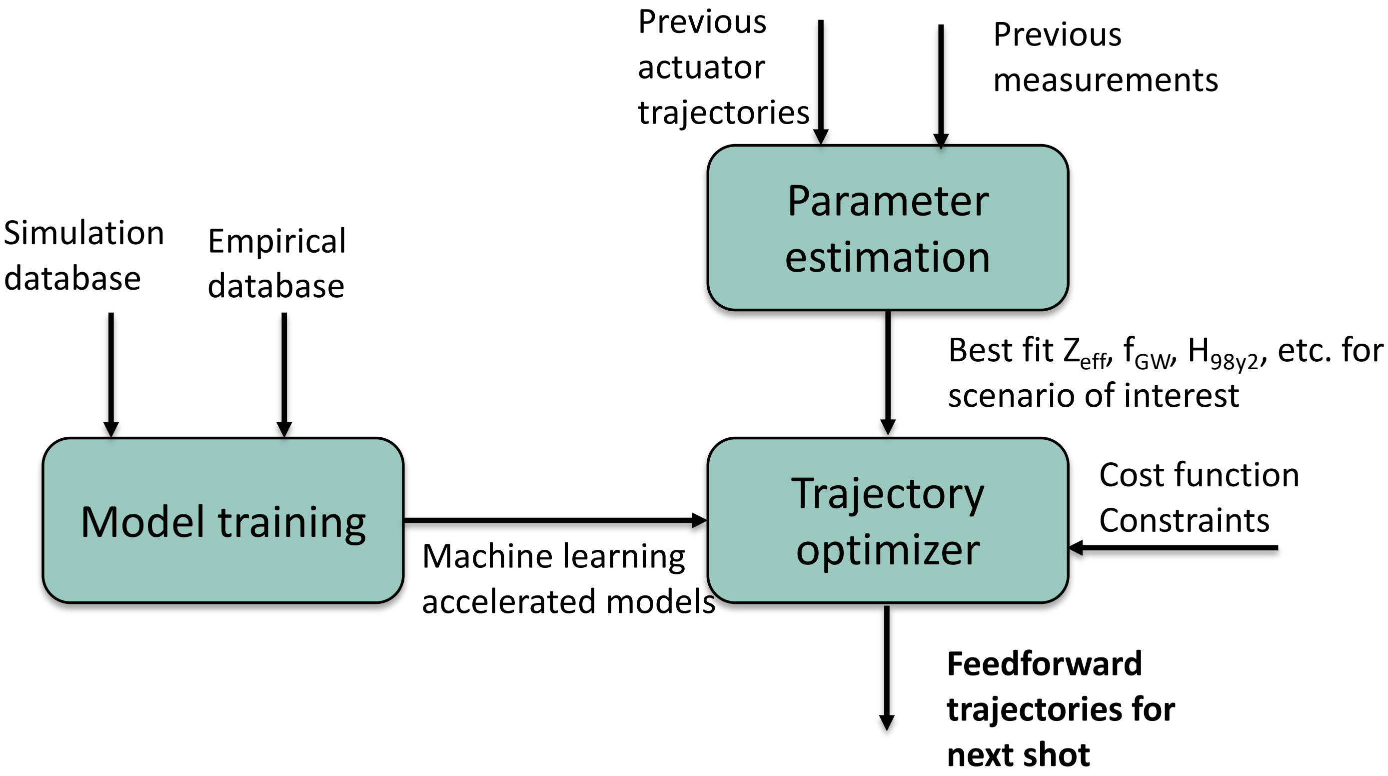






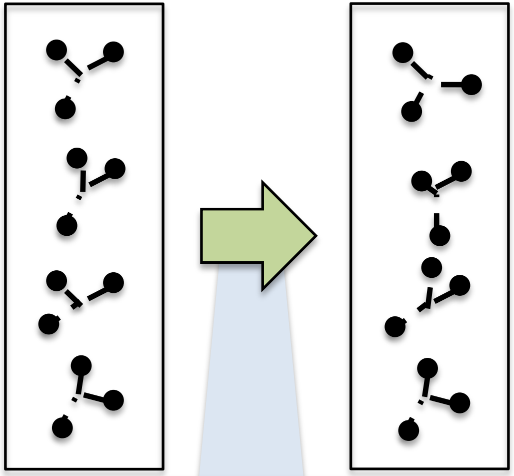
<1s wall clock time

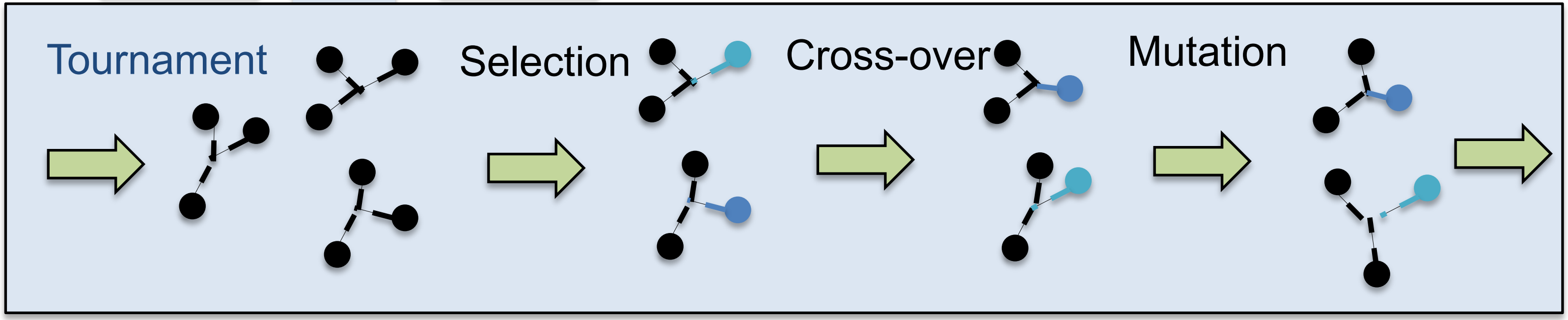
(3) Rapid calculation time facilitates scenario/trajectory optimization for experiment planning and control



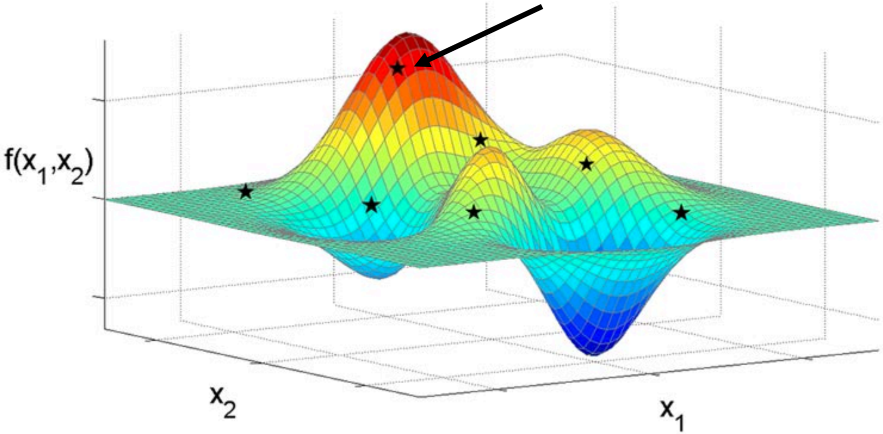
Genetic algorithm enables global optimization of actuator trajectories, gradient-based methods enable refinement

Initial population New population





Genetic algorithm finds approximate global optimum



Note: Individuals in this case are arrays of actuator magnitudes and associated times

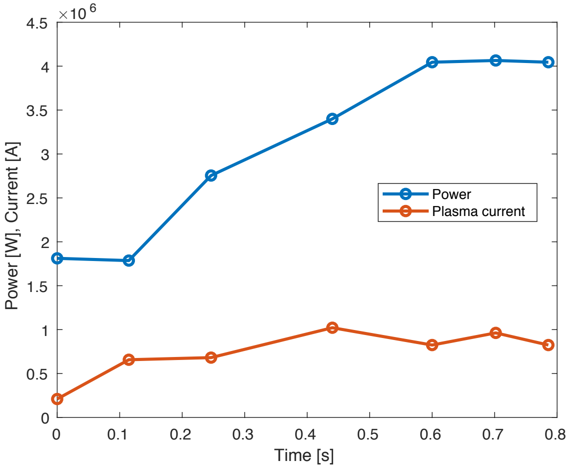
Different tasks can be optimized with the same model and codebase with change in cost function and actuators

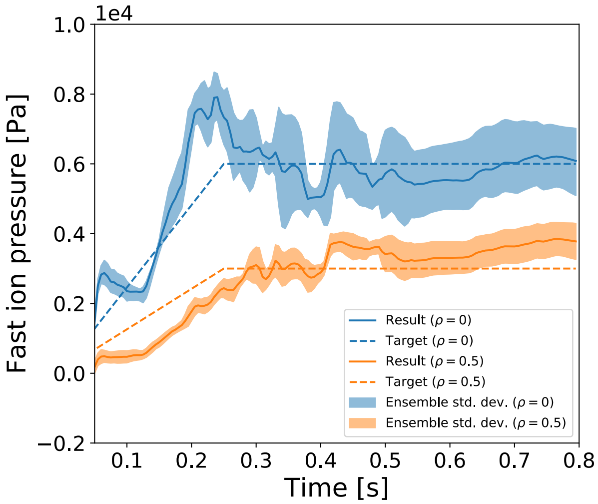
$$J_0 = \int_{t_i}^{t_f} \left[(p_e(0) - p_{e,t}(0))^2 + (p_e(0.5) - p_{e,t}(0.5))^2 + (P_{fi}(0) - P_{fi,t}(0))^2 + (P_{fi}(0.5) - P_{fi,t}(0.5))^2 \right] dt,$$

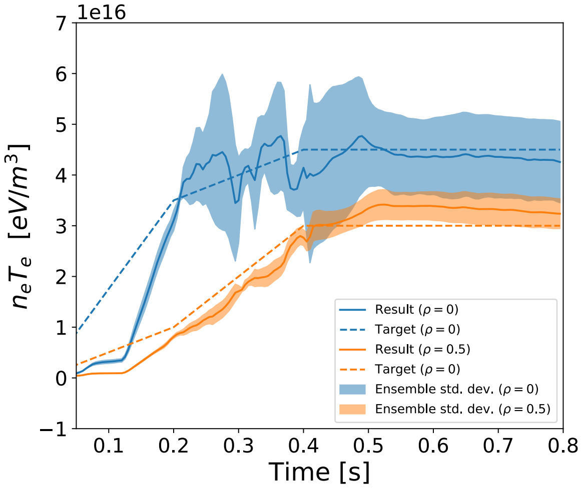
$$J_1 = J_0 + \int_{t_i}^{t_f} [\lambda_{p_e}(\sigma^2_{p_e}(0) + \sigma^2_{p_e}(0.5) + \lambda_{P_{fi}}(\sigma^2_{P_{fi}}(0) + \sigma^2_{P_{fi}}(0.5)]dt$$

Weights fast ion and electron pressure profile tracking at two points

Penalty on model uncertainty keeps solution in confidence region







Machine learning enables accelerated integrated modeling for scenario optimization and control

- Accelerate well-validated physics models
  - Example: NubeamNet

Boyer et al., *Nuclear Fusion 59, 5,* 2019

- Generate empirical models for phenomena not well described by physics models
  - Example: Electron transport on NSTX

Boyer et al., *Nuclear Fusion 61, 4,* 2021

- Combined models for accelerated discharge prediction
  - Initial version developed, on-going work to include more models
- Genetic algorithms show promise for global optimization of actuator trajectories

Boyer, Proceedings of the 2nd Conference on Learning for Dynamics and Control, PMLR 120:698-707, 2020.

- Future/ongoing work
  - Experimentally validate trajectory design approach
  - Integrate accelerated models into feedback control algorithms
  - Extend approach to more tokamaks

DIII-D NubeamNet: Morosohk, et al., FED, 163, 112125, 2021.