Stabilizing Tokamak Plasmas with ML-based Feedback Control

A. Rothstein[1],

J. Abbate[2], J. Seo[3], H. Farre[2], I. Char[4], R. Shousha[2], A. Jalalvand[1], S.K. Kim[2], and E. Kolemen[1,2]

[1] Princeton University

[2] Princeton Plasma Physics Lab

[3] Chung-Ang University

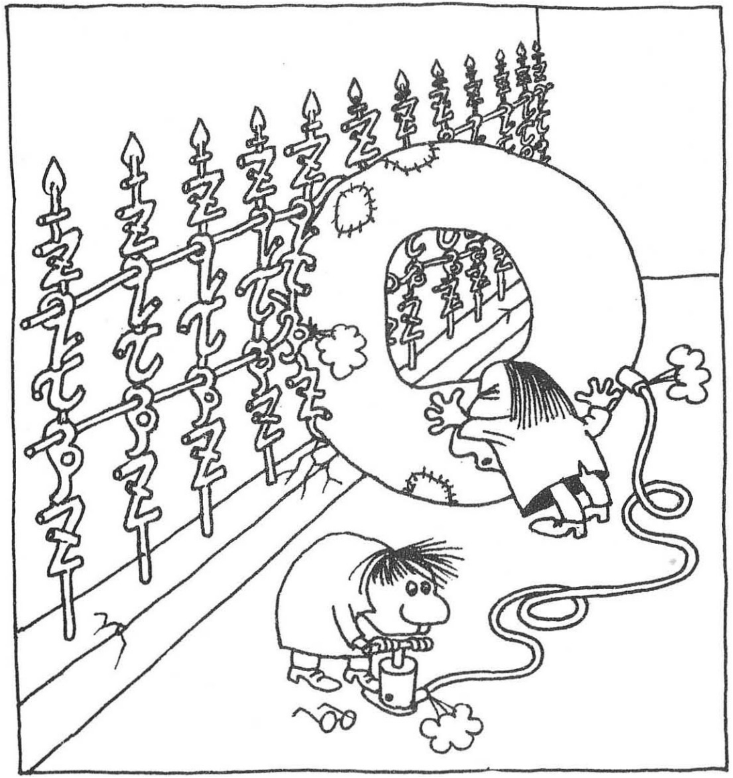
[4] Carnegie Mellon University

Plasma Control

Mechanical & Aerospace Engineering

DIII-D NATIONAL FUSION FACILITY

PPPL



Cartoon: EPS 1981

DIII-D Experiment Status, San Diego CA 2021-06-16 08:51:16

Next Shot	Ip Request	Shot Setup Time	Ready to Start
187065	1.00 MA	01:28	Beams · PCS · Data Acq
	Bt Request	Time Since Last Power Shot	Ready to Fire
ReferenceShot	1.99 T	0:21:47	Beams · ECH · Cryo · MPRB
			Requested for Shot: N/A · N/A

One Minute to Countdown

First Fault	B Cooldown	F Cooldown	Glow Left	PCS Status	Pit Run
None	00:00	00:00	00:00	In Lockout	CLEAR

EPS Settings

Reverse	105.0 KA	Forward	130.0 KA
Off	6.504 s	EBTime	0.840 s

BPS Settings

+	118.0 KA		1.99 Tesla	IB Control From Ops	
Flattop	-0.155 s	Off	5.226 s	Cool Down	9.940 m

ECH Valve/Ready Status

Gyro	Tin Man	Leia	Luke	Vader	Nasa	R2D2
Load	Local	Local	Local	Local	Local	Local

Density Limit

$5.6 \times 10^{13} cm^3$

Neutral Beams

	30L	30R	150L	150R	210L	210R	330L	330R
Voltage (KV)	81.0	75.0	62.0	62.1	70.0	75.7	72.0	76.0
TIV	Open		Open		Open		Open	

Vessel Pressure

8.8862e-08 Torr

150 Beamline

Tilt	Source Angle Left	Source Angle Right
16.08 degree	17.46 '	-0.02 '

Arcminutes

Plasma Control

Improving plasmas by trial-and-error

- “[after a previous full-day experiment we were] able to reproduce 133103 in 180636, 180643 and 180644.
- Many shots had MHD modes at 3 s… to try to improve that we changed **Electron Cyclotron Heating** deposition (180639-180642), and go to lower (180643-180646) and higher (180647) **plasma current**… none of which were successful.
- We also **tried lowering the voltage on the off-axis beams** (180645) to get rid of the bursty modes and **moving the BetaN ramp earlier** (180646.)”
- Ultimately, got “good reproduction of 133103, but no significant improvement”

Human operators combine simulations, heuristics, and experience to achieve desired state by trial-and-error

Stabilizing Tokamak Plasmas with ML-based Feedback Control

A. Rothstein[1],

J. Abbate[2], J. Seo[3], H. Farre[2],
I. Char[4], R. Shousha[2], A. Jalalvand[1],
S.K. Kim[2], and E. Kolemen[1,2]

[1] Princeton University

[2] Princeton Plasma Physics Lab

[3] Chung-Ang University

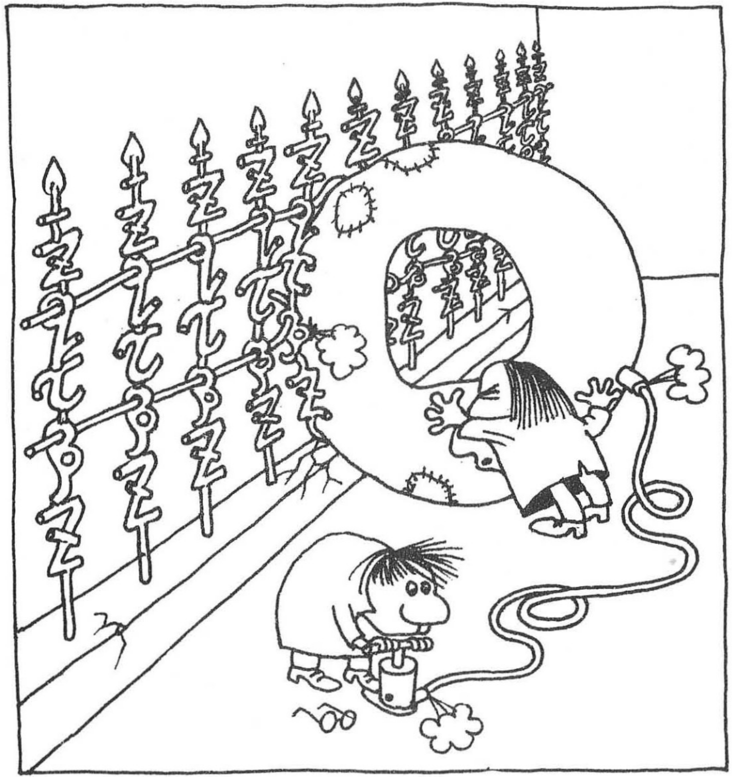
[4] Carnegie Mellon University

Plasma Control

Mechanical & Aerospace Engineering

DIII-D NATIONAL FUSION FACILITY

PPPL



Cartoon: EPS 1981

Stabilizing Tokamak Plasmas with ML-based Feedback Control

A. Rothstein[1],

J. Abbate[2], J. Seo[3], H. Farre[2], I. Char[4], R. Shousha[2], A. Jalalvand[1], S.K. Kim[2], and E. Kolemen[1,2]

[1] Princeton University

[2] Princeton Plasma Physics Lab

[3] Chung-Ang University

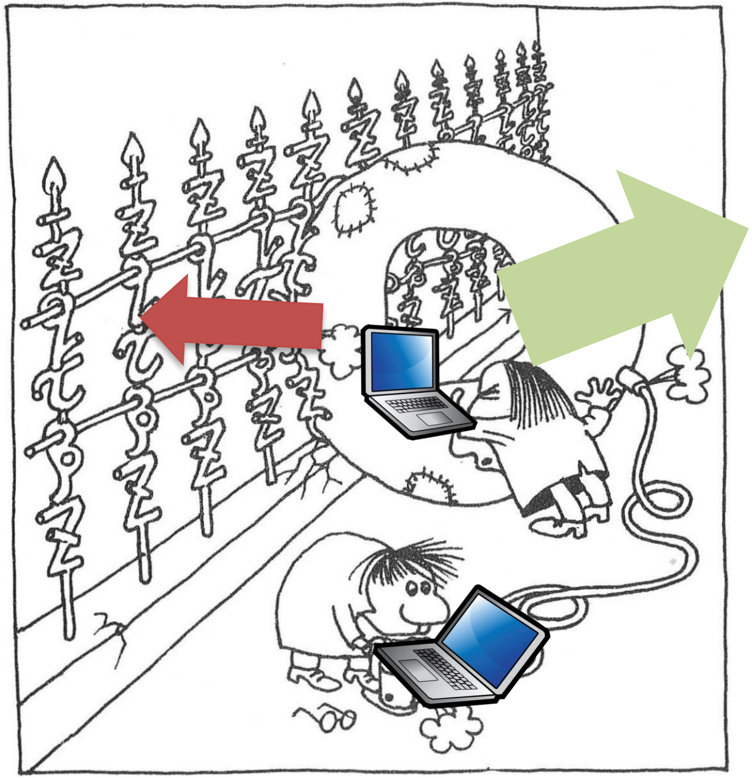
[4] Carnegie Mellon University

Plasma Control

Mechanical & Aerospace Engineering

DIII-D NATIONAL FUSION FACILITY

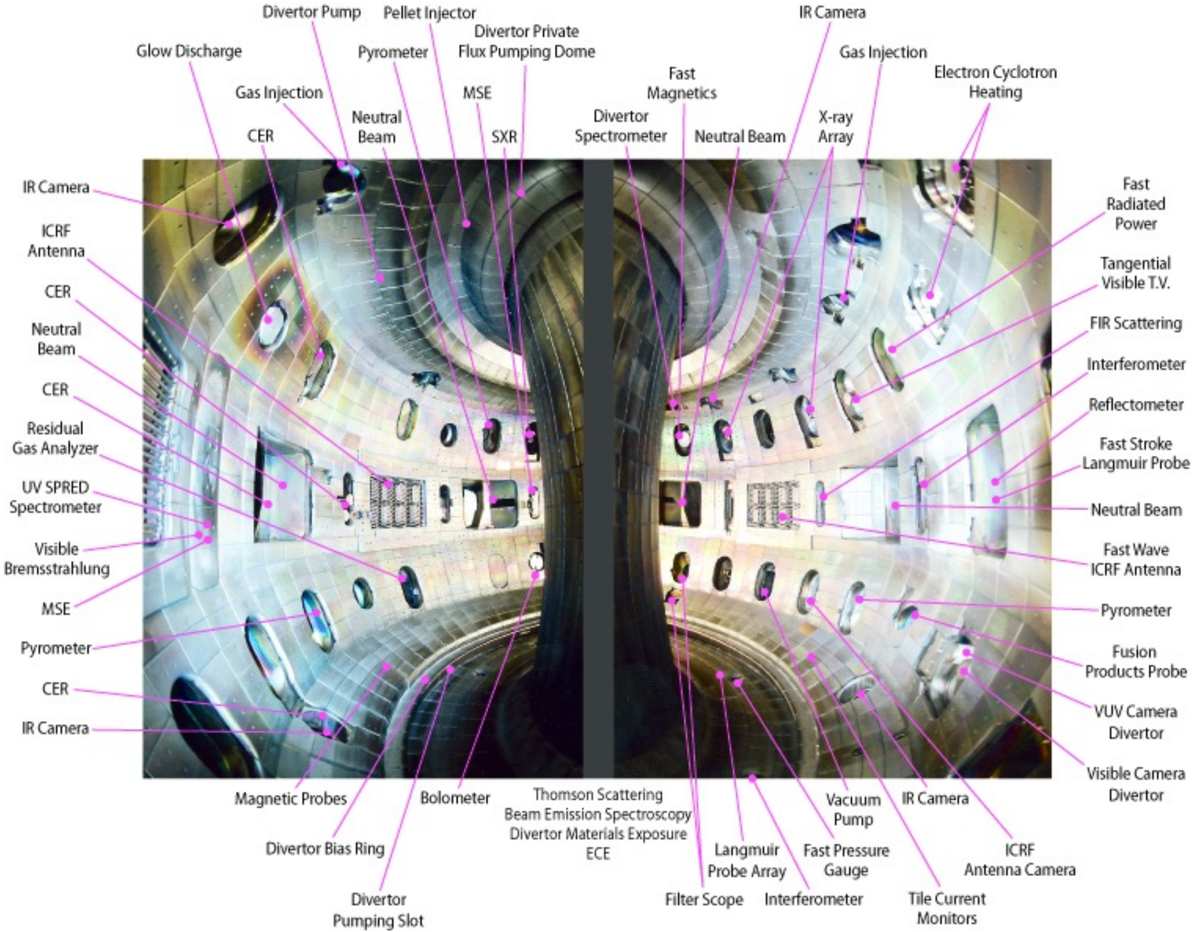
PPPL



Cartoon: EPS 1981

Crash Course to Tokamak Experiments

Observing the plasma state

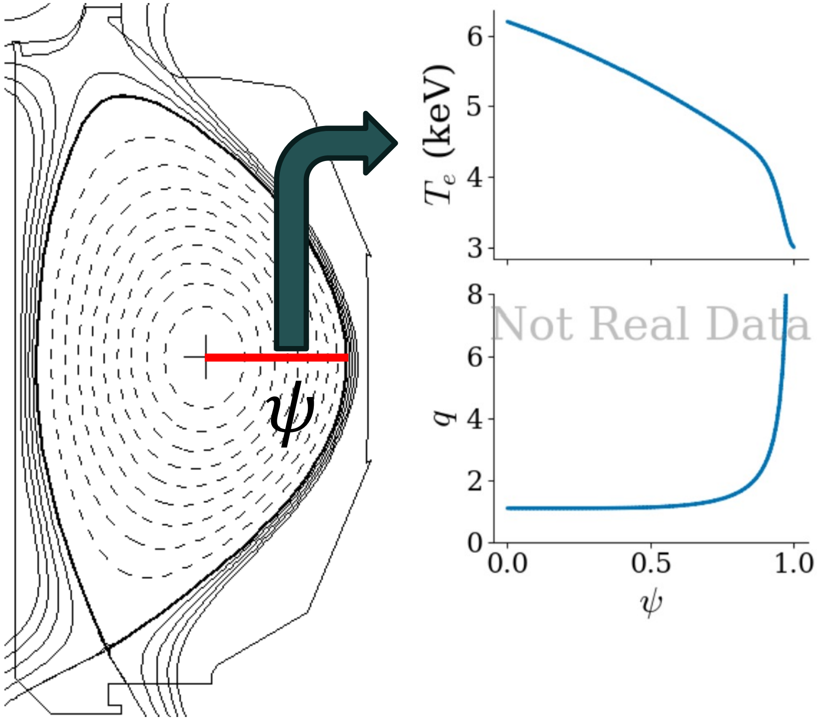


Reconstructing the plasma state

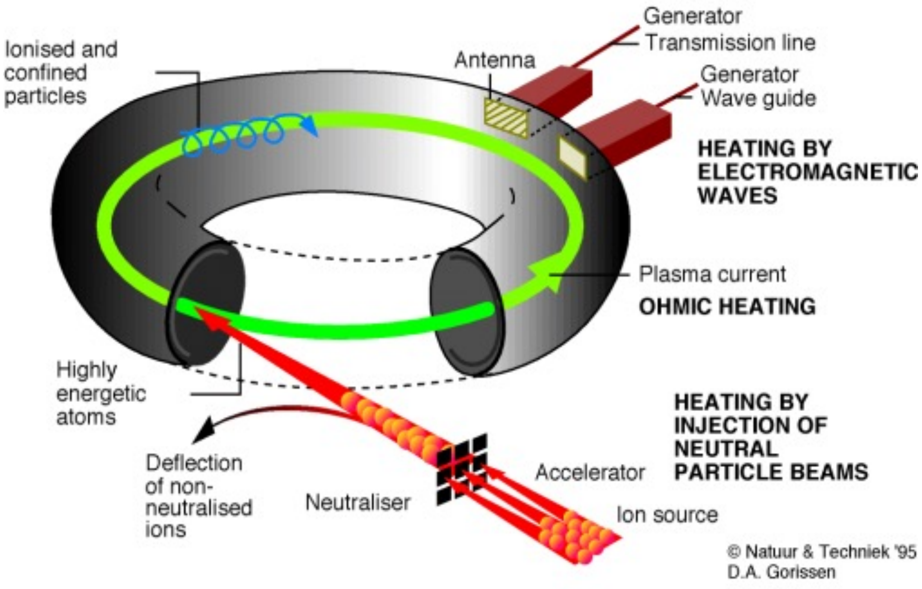
- Use diagnostic data to reconstruct plasma equilibrium

Scalar Parameters
- Plasma shape and boundary ($\kappa, \delta_{u,l}$ etc)
- Normalized pressure (β_N)
- Plasma current (I_P)
- Magnetic field (B_T)

1D Profiles
- Pressure (P)
- Safety factor (q)
- Electron temperature and density (T_e, n_e)
- Ion temperature and density ($\mathrm{T_i}, \mathrm{n_i}$)
- Rotation (Ω)



Actuators that affect plasma state



Heating Sources
- Neutral Beams
- Electron Cyclotron Heating
- Other RF Waves (Helicon + Lower Hybrid)

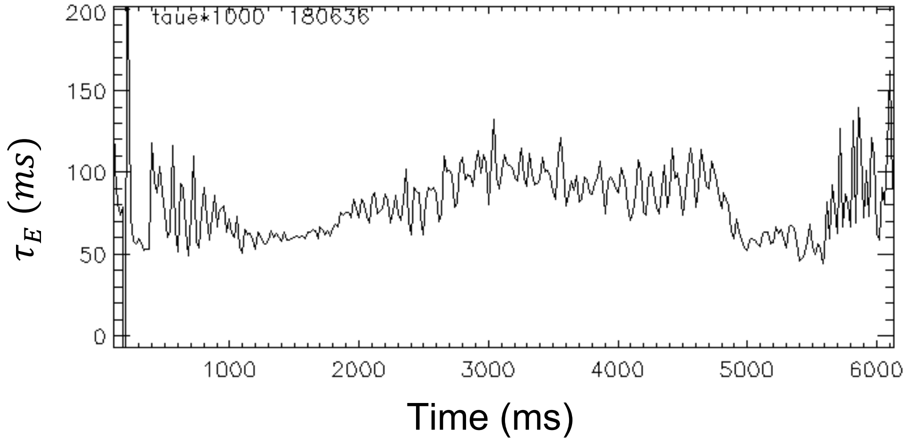
Magnetic Coils
- Central solenoid ramp rate
- Toroidal field coils
- Poloidal field coils
- 3D field coils to perturb toroidal symmetry

Gases
- Gas valves
- Pellet injection

Experimental Timescales

Profile Evolution
- τ_E: 50-100ms
- τ_R: $\approx$ 1s

Instabilties
- Tearing Modes: 1-10ms
- VDEs: μs scale
- Disruptions: $\approx$ 1ms



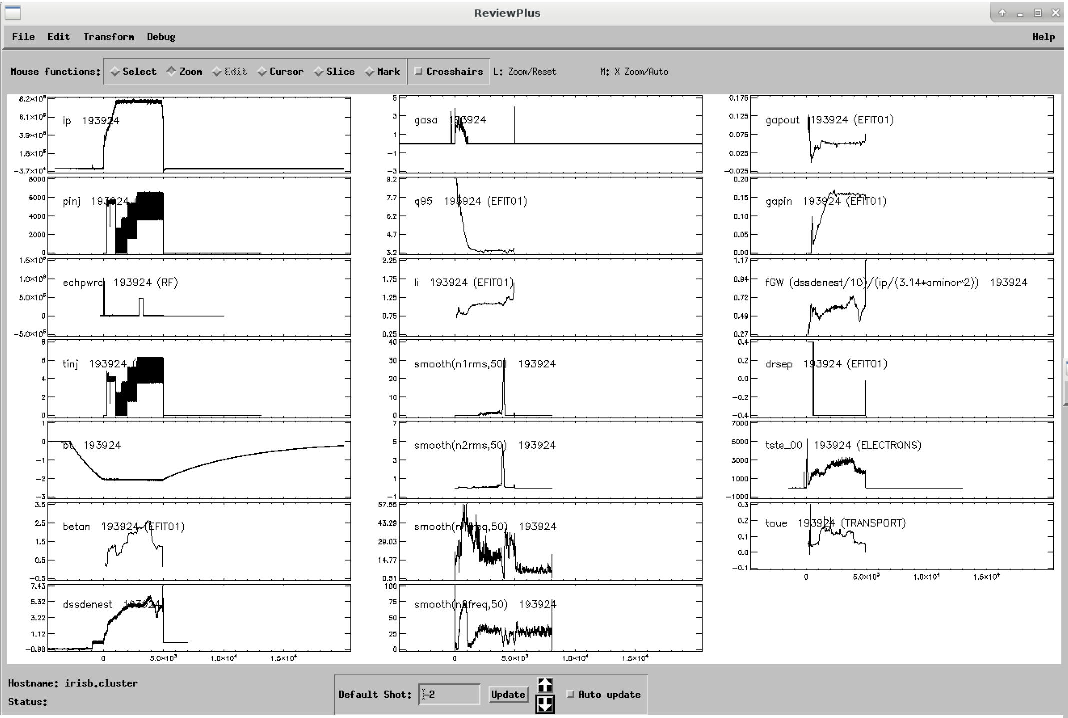
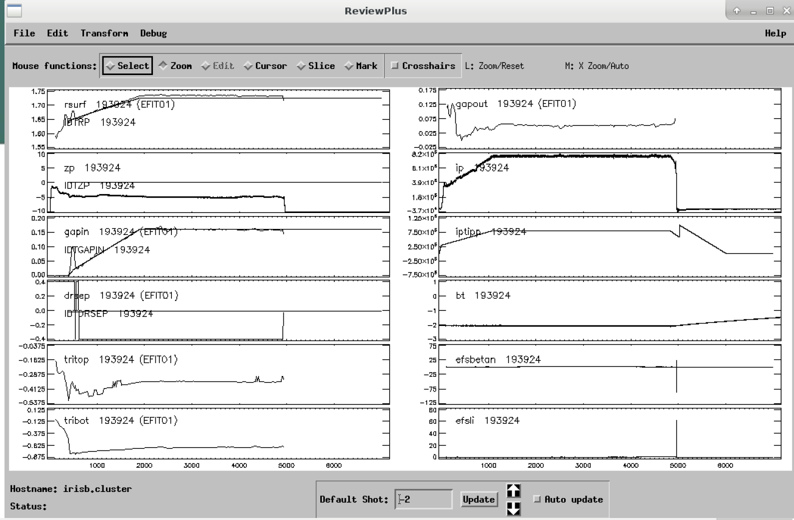
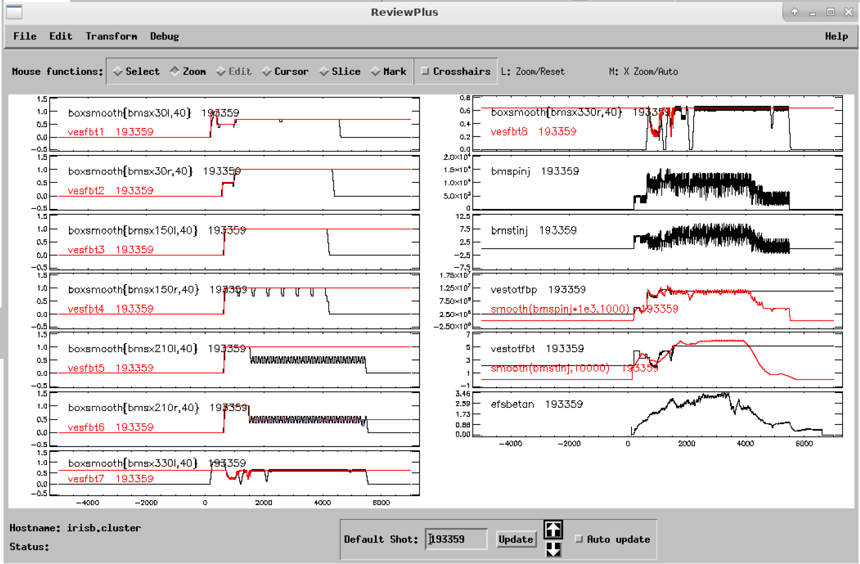
Real-time control system
- Shape control: <1ms
- NBI heating: 50ms
- ECH heating: 50ms
- ML models: 1-10ms
- Magnetic diagnostics: <1ms
- Profile diagnostics: $\approx$ 20ms

Plasma Control

Why machine learning?

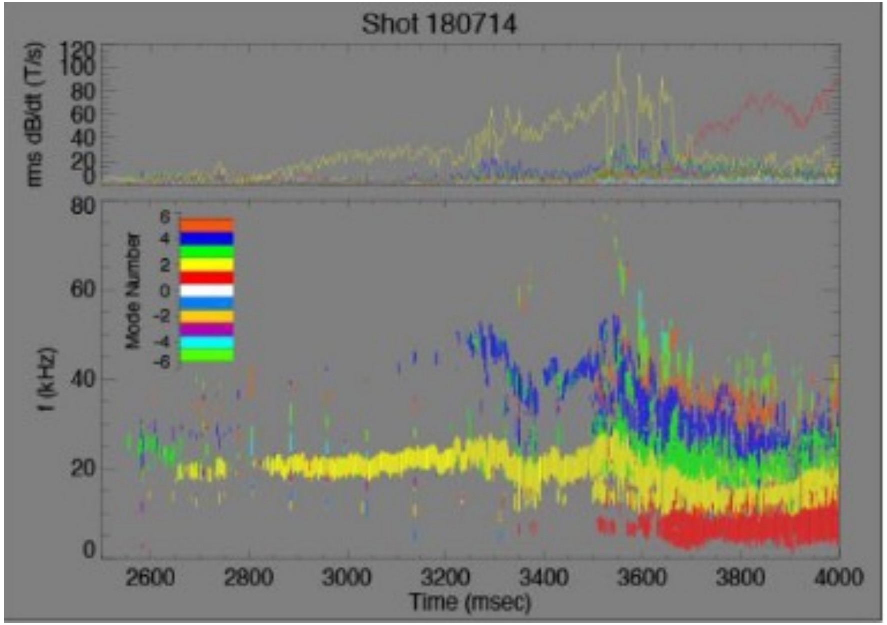
- Lots of data

Plasma Control

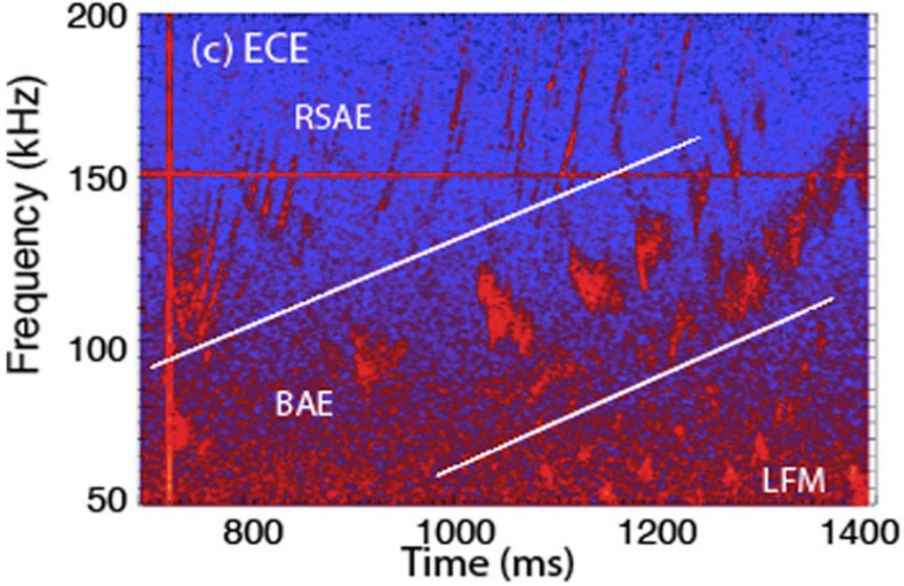

ReviewPlus
File Edit Transform Debug Help
Mouse functions: Select Zoom Edit Cursor Slice Mark Crosshairs L: Zoom/Reset M: X Zoom/Auto
ip 193924
pinj 193924
echpwrc 193924 (RF)
tinj 193924
bt 193924
betan 193924 (EFIT01)
dssdenest 193924
gasa 193924
q95 193924 (EFIT01)
li 193924 (EFIT01)
smooth(n1rms,50) 193924
smooth(n2rms,50) 193924
smooth(n1freq,50) 193924
smooth(n2freq,50) 193924
gapout 193924 (EFIT01)
gapin 193924 (EFIT01)
fGW (dssdenest/10)/(ip/(3.14*aminor^2)) 193924
drsep 193924 (EFIT01)
tste_00 193924 (ELECTRONS)
taue 193924 (TRANSPORT)
Hostname: irisb.cluster
Status:
Default Shot: -2
Update
Auto update

ReviewPlus
File Edit Transform Debug Help
Mouse functions: Select Zoom Edit Cursor Slice Mark Crosshairs L: Zoom/Reset M: X Zoom/Auto
rsurf 193924 (EFIT01)
IDTRP 193924
zp 193924
IDTZP 193924
gapin 193924 (EFIT01)
IDTGAPIN 193924
drsep 193924 (EFIT01)
IDTDRSEP 193924
tritop 193924 (EFIT01)
tribot 193924 (EFIT01)
gapout 193924 (EFIT01)
ip 193924
iptipp 193924
bt 193924
efsbetan 193924
efsli 193924
Hostname: irisb.cluster
Status:
Default Shot: -2
Update
Auto update

ReviewPlus
File Edit Transform Debug Help
Mouse functions: Select Zoom Edit Cursor Slice Mark Crosshairs L: Zoom/Reset M: X Zoom/Auto
boxsmooth(bmsx30l,40) 193359
vesfbt1 193359
boxsmooth(bmsx30r,40) 193359
vesfbt2 193359
boxsmooth(bmsx150l,40) 193359
vesfbt3 193359
boxsmooth(bmsx150r,40) 193359
vesfbt4 193359
boxsmooth(bmsx210l,40) 193359
vesfbt5 193359
boxsmooth(bmsx210r,40) 193359
vesfbt6 193359
boxsmooth(bmsx330l,40) 193359
vesfbt7 193359
boxsmooth(bmsx330r,40) 193359
vesfbt8 193359
bmspinj 193359
bmstinj 193359
vestotfbp 193359
smooth(bmspinj*1e3,1000) 193359
vestotfbt 193359
smooth(bmstinj,10000) 193359
efsbetan 193359
Hostname: irisb.cluster
Status:
Default Shot: 193359
Update
Auto update

Why machine learning?

- Lots of data
- Models can be run real-time (ms time-scale)
- ML can find patterns to predict instabilities



[Victor IAEA 2020]



[Heidbrink NF 2021]

Why machine learning?

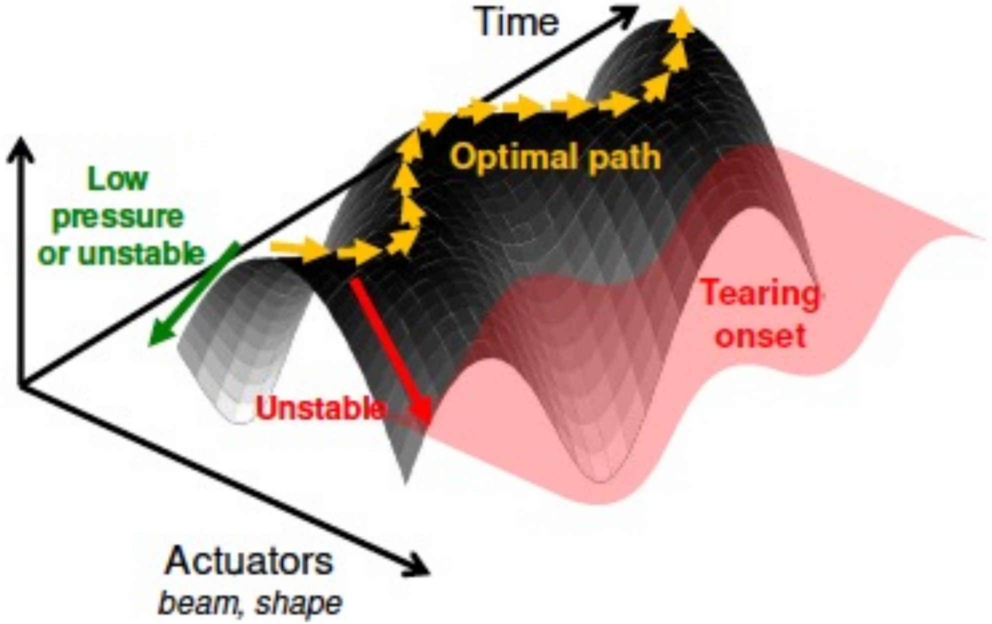
- Lots of data
- Models can be run real-time (ms time-scale)
- ML can find patterns to predict instabilities
- Physics models use artificial actuators

- Minimal experimental time to test my modes and policies
- Shots not reproducible
- Non-linearities also make learning challenging

What should a good ML controller do?

What ML is not good for:
- Developing new scenarios
- Extrapolating to new regimes

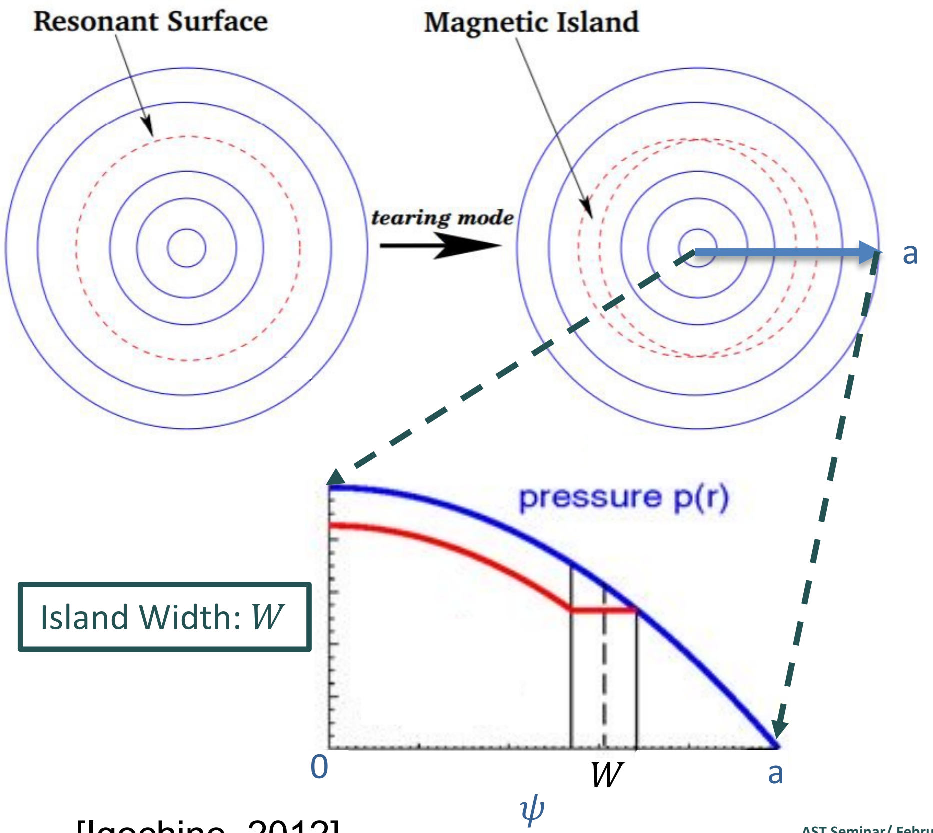
What ML is good for:
- Maintaining stability in previously explored spaces
- Recovering from small deviations to optimized scenario



[Seo Nature 2024]

Tearing Mode Prediction and Control

What are tearing modes?



Island Width: W

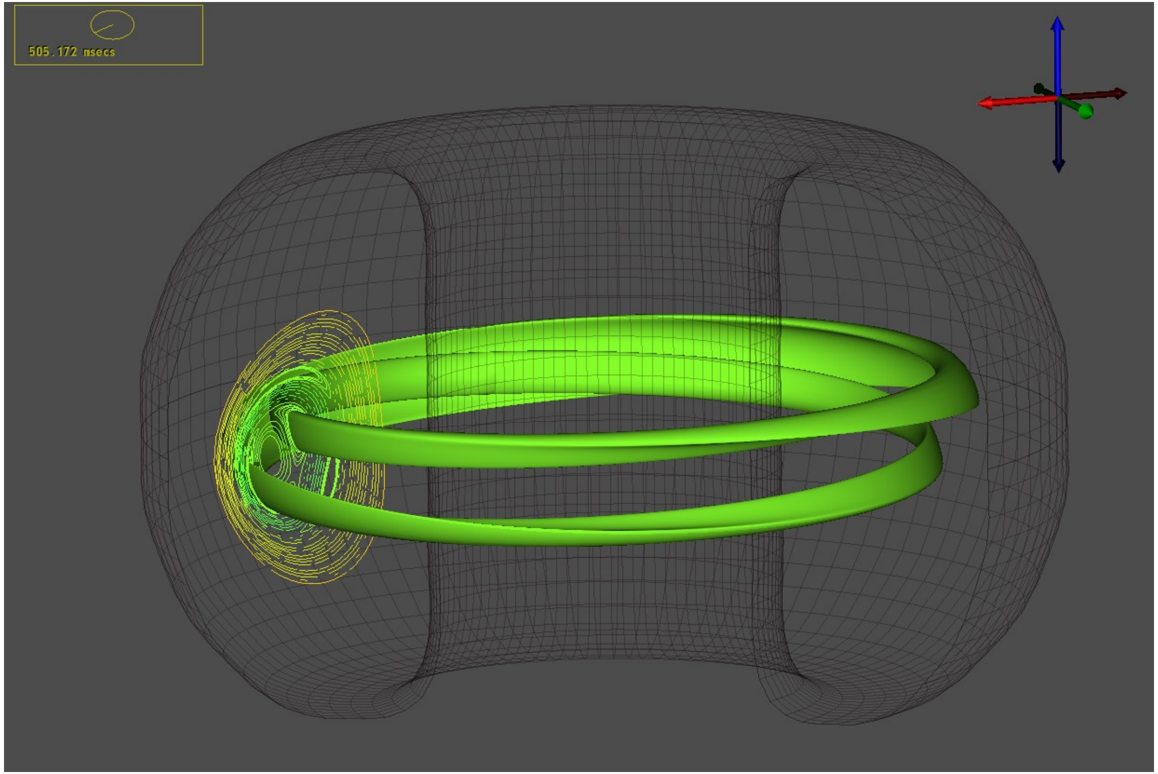
[Igochine, 2012]

- Magnetic field reconfigures to lowest energy
- Occurs at rational surfaces
- Breaks nicely nested flux surfaces

So why do we care?

- “Short circuits” transport
- Modes can lock to wall → disrupts plasma

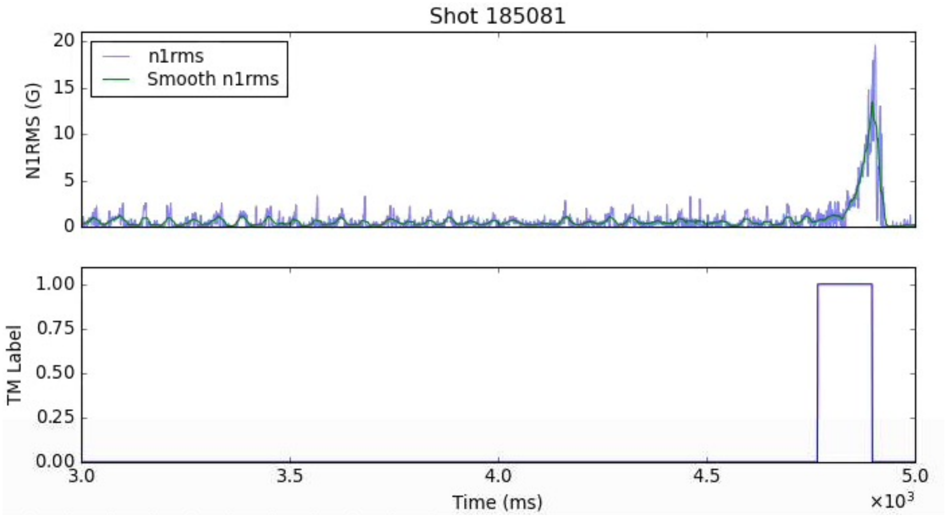
Tearing Modes are 3D Structures



[VACET]

Tearing Mode Database

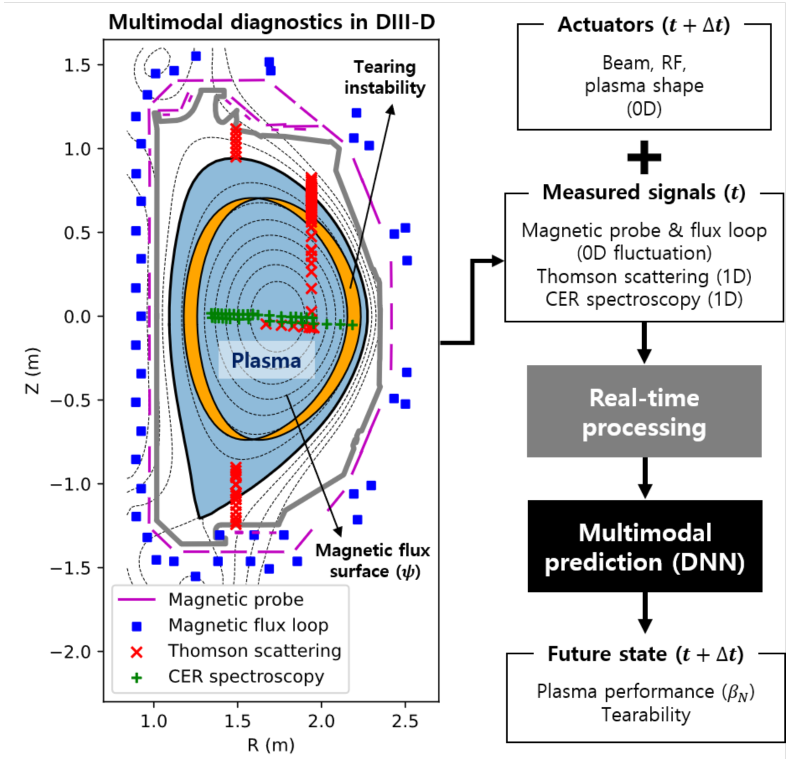
- Conditions for TM label:
  - n1rms > 12G
  - Duration of TM > 50ms
  - $H_{98} > 0.7$
  - q_{95} at TM onset $< 1.5 * \min(q_{95})$
- Includes 8,505 shots from 2011-2022 campaigns with 639,555 time slices
  - $\approx 8\%$ of time slices have TMs



Tearing Mode Predictor

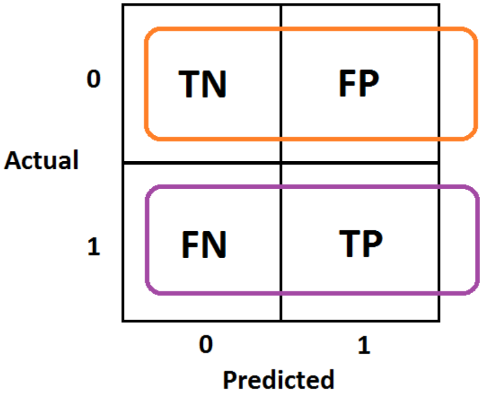
- Uses current time-step profiles + future actuators:
  - This models: "What can a controller do?"
- $\Delta t = 25$ms chosen to capture profile variation

Profiles at $T = t$	Actuators at $T = t + \Delta t$	Outputs $T = t + \Delta t$
rtEFIT: q, p Thomson: n_e, T_e CER: v_{tor}	rtEFIT: $B_T, I_p, R_0, \kappa, \delta_u, \delta_l, gap_{in}$ $P_{NBI}, T_{NBI}, P_{ECH}$	Tearability β_N



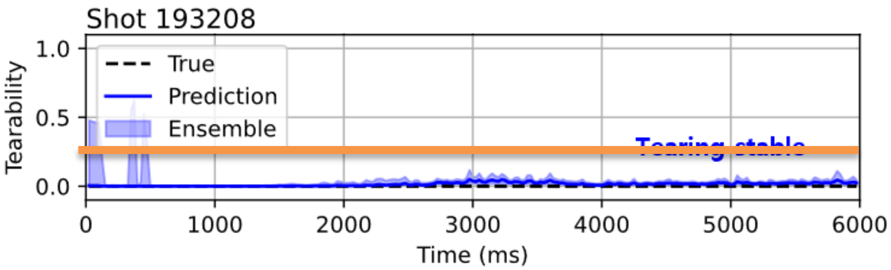
Now stir the pile, but how should we evaluate models?

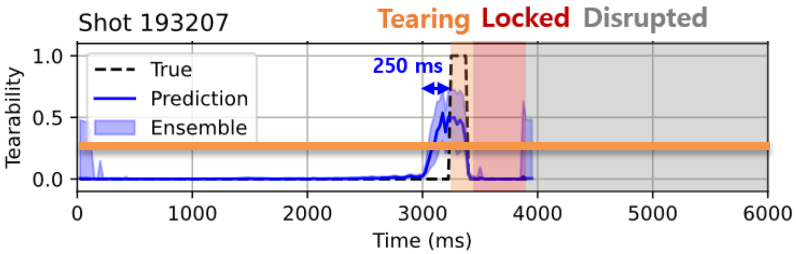
Assessing Event Prediction Models



$$\textit{False Positive Rate (FPR)} = \frac{FP}{(FP + TN)}$$

$$\textit{True Positive Rate(TPR)} = \frac{TP}{(TP + FN)}$$

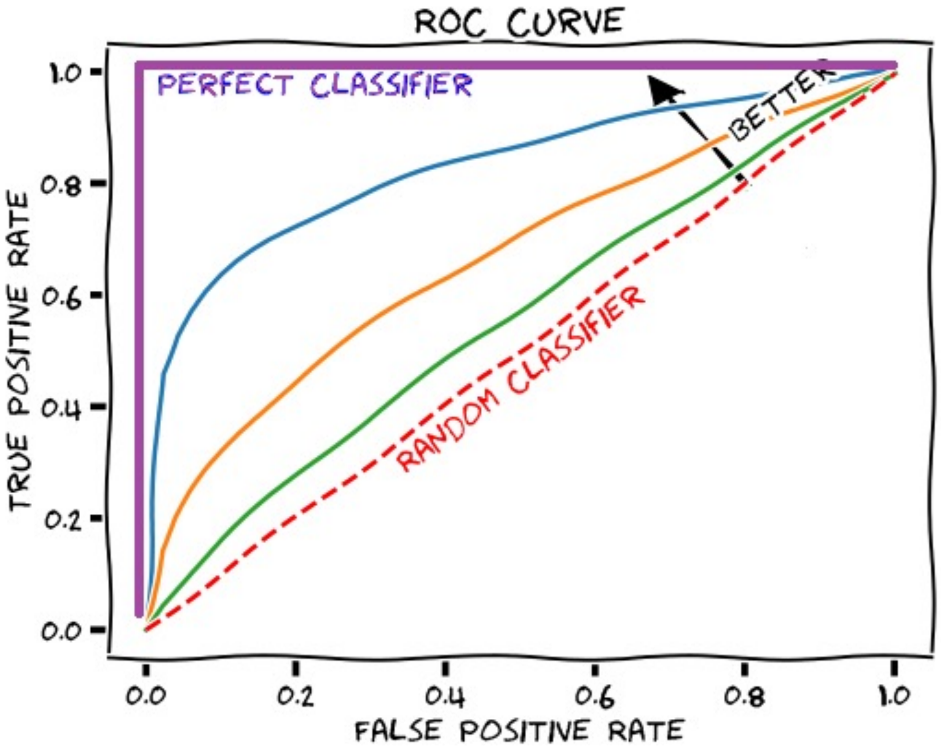




Plasma Control

AUC Metric

- AUC metric integrates TPR by sweeping threshold from $1 \to 0$
  - FPR sweeps from $0 \to 1$
- AUC values:
  - Perfect classifier = 1
  - Random classifier = 0.5

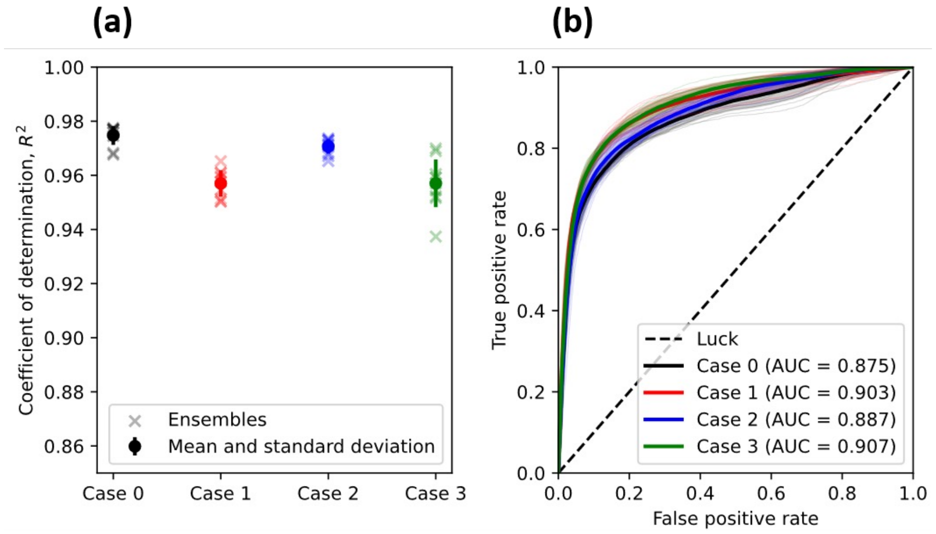


Plasma Control

TM Model Selection

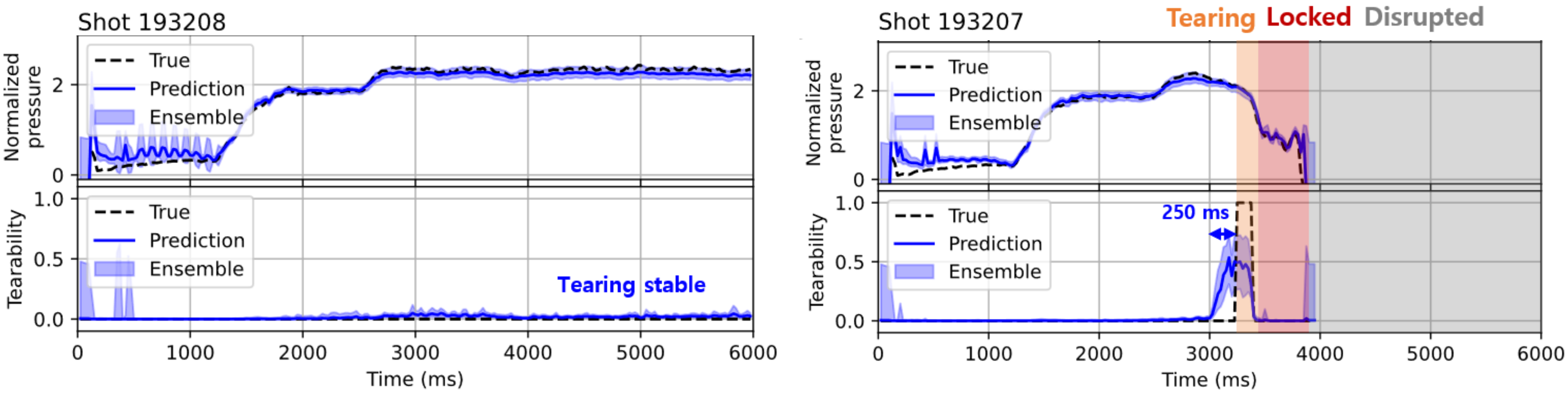
TABLE II
THE INFLUENCE OF THE TYPE OF LOSS FUNCTION AND OVERSAMPLING

Case number	Loss for β_N	Loss for tearability	Over-sampling	R^2 for β_N	AUC for tearability
0	MSE	MSE	No	0.975	0.875
1	MSE	MSE	Yes	0.957	0.903
2	MSE	BCE	No	0.971	0.887
3	MSE	BCE	Yes	0.957	0.907



TM Predictor Results

- Ensemble of 10 models gives uncertainty estimate

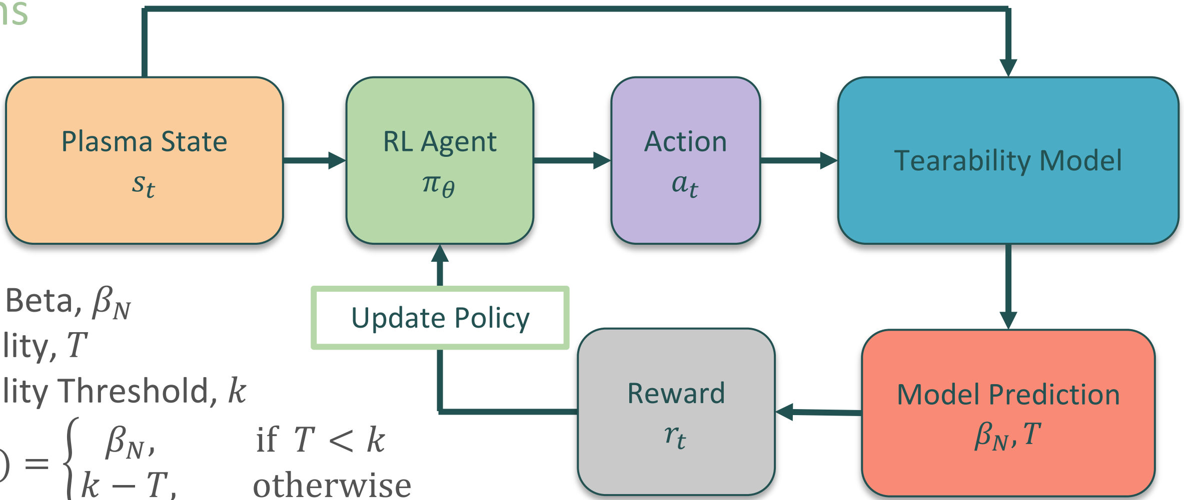


J. Seo et al, IJCNN, 2023.

Plasma Control

RL Training Loop

1. Random, real experimental data, s_t, selected from database
2. RL Agent observes plasma and decided on action to take, a_t
3. Plasma state and RL agent's actions are fed to Tearability model
4. Tearability Model predicts Tearability (T) and β_N
5. Using reward function, the RL agent updates its policy to perform better in future iterations

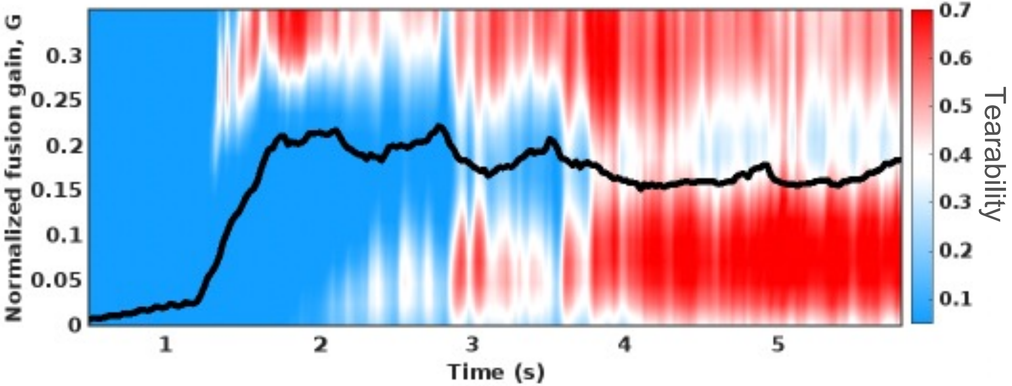


- Plasma Beta, β_N
- Tearability, T
- Tearability Threshold, k

$$r_t(\beta_N, T; k) = \begin{cases} \beta_N, & \text{if } T < k \\ k - T, & \text{otherwise} \end{cases}$$

Why is RL control worth exploring?

- Fusion Gain: $G \propto \beta_N / q_{95}^2$
- RL feedback control can "thread the needle" to do better than other controllers



[Seo Nature 2024]

Summary of TM and instability control

- Build a model that identifies and predicts instabilities:
  - Other instabilities: VDEs, AEs, ELMs, Density limits, etc
- If we understand how to control: can do direct control based on event predictions
  - Example: predict disruptions → safe ramp down
- If we don't understand how to control: RL is a possible option that can learn solutions based on experimental data
  - RL controller can turn multiple actuators to find solutions classical control cannot
  - Improvements to instability predictor → improvements to RL agent

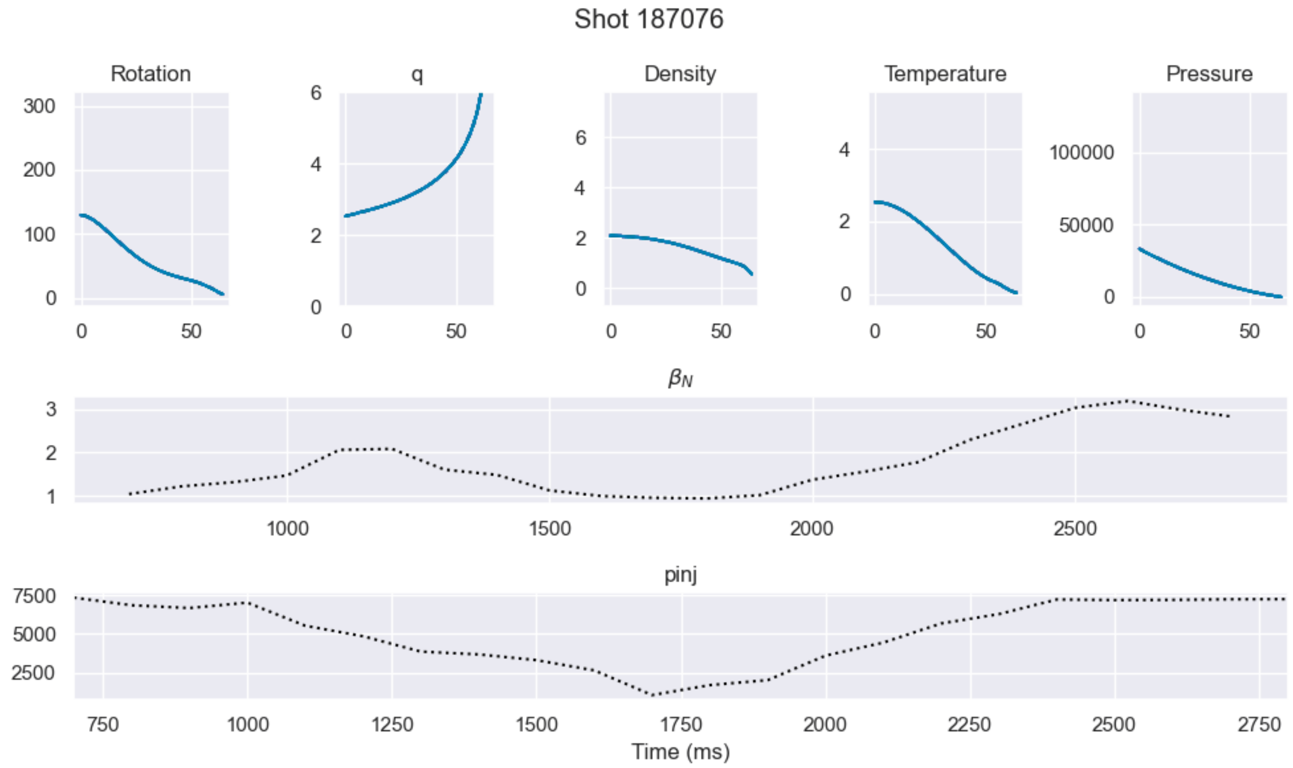
Full Shot Profile Prediction and Control

Plasma Control

Predict full shots with actuator trajectories

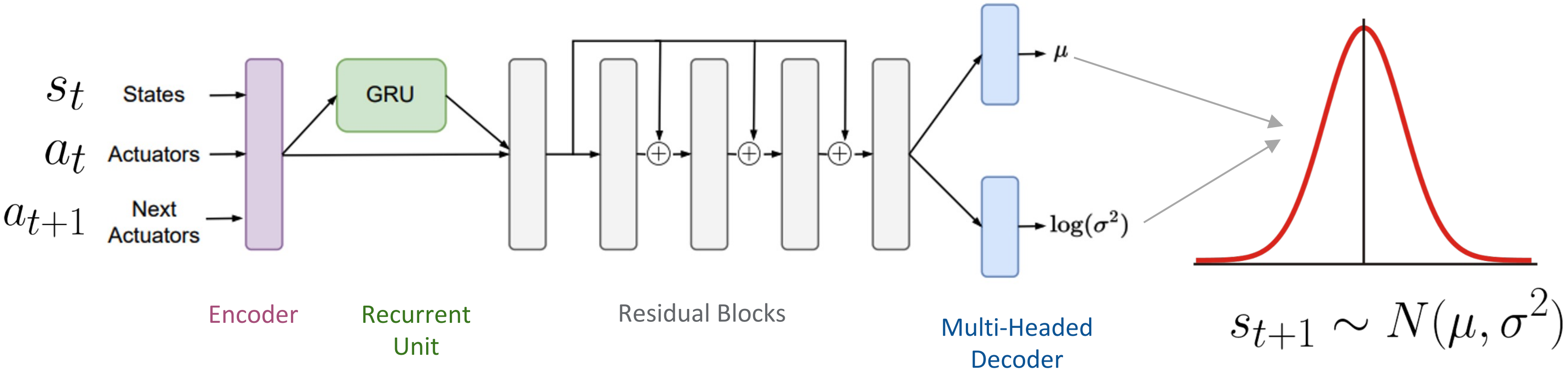
Ground Truth

Predictions



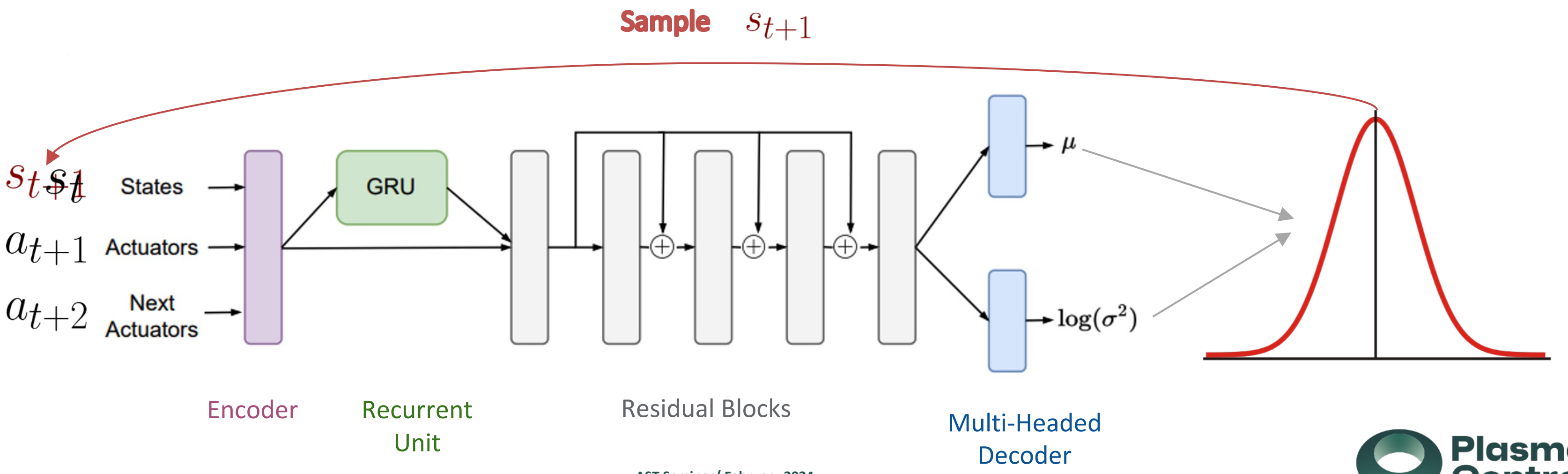
Model Architecture

- Predictions are made 25ms into future
- Model predicts a Gaussian distribution of the next state



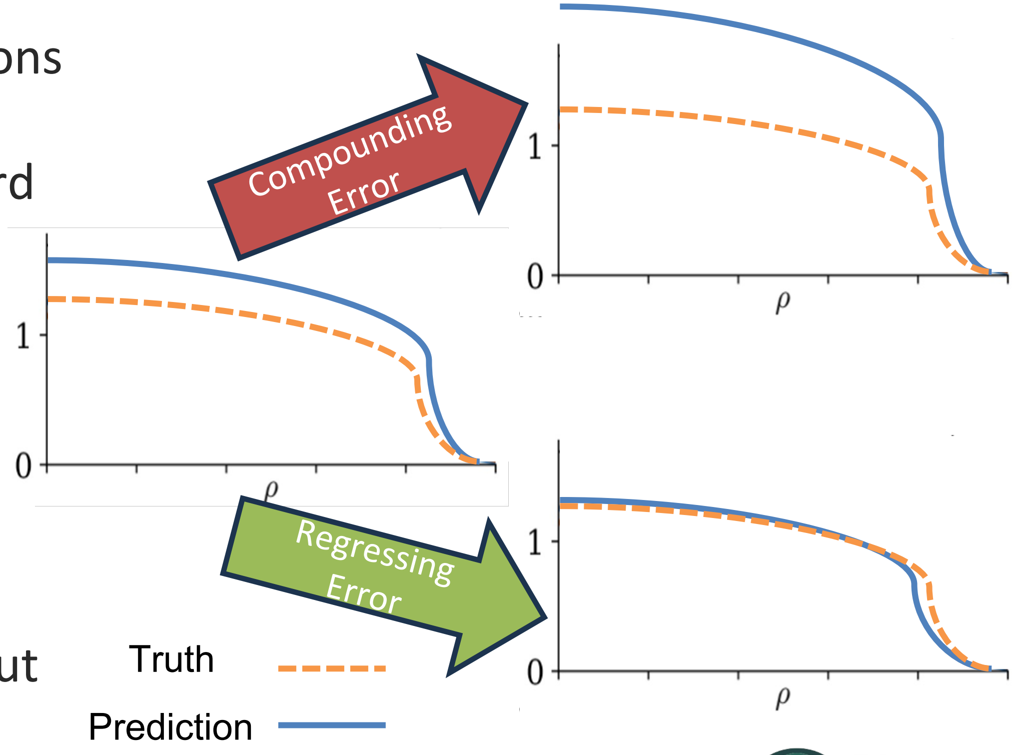
Predicting Full Shots

- Next state sampled from Gaussian and fed back into model
- Actuators can be taken from historical data ("replaying a shot") or provided by some optimization algorithm



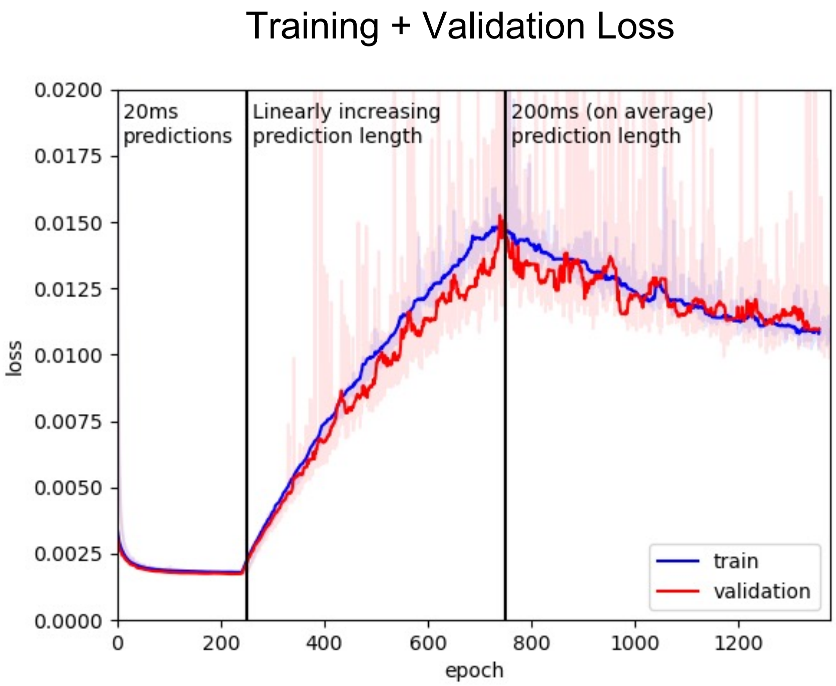
How to keep long-term predictions stable

- Predicting off previous predictions causes errors to compound
- Need to have ”regression toward the mean”
- Solution 1: Uncertainty predictions
  - Predict (μ, σ)
- Solution 2: model ensembling
  - Multiple models = further averaging
- Solution 3: autoregressive rollout



Curriculum Learning

- Start by having model predict $\mu = 1$ time steps into future
  - Use time t to predict $t + 1$
- Ramp prediction horizon from $\mu = 1$ to $\mu = 10$
- Continue training at $\mu = 10$



Model Predictive Control

$$\boldsymbol{x_{t+1}} = \boldsymbol{Ax_t} + \boldsymbol{Bu_t}$$

Predicted state

Current state

Control Actuators

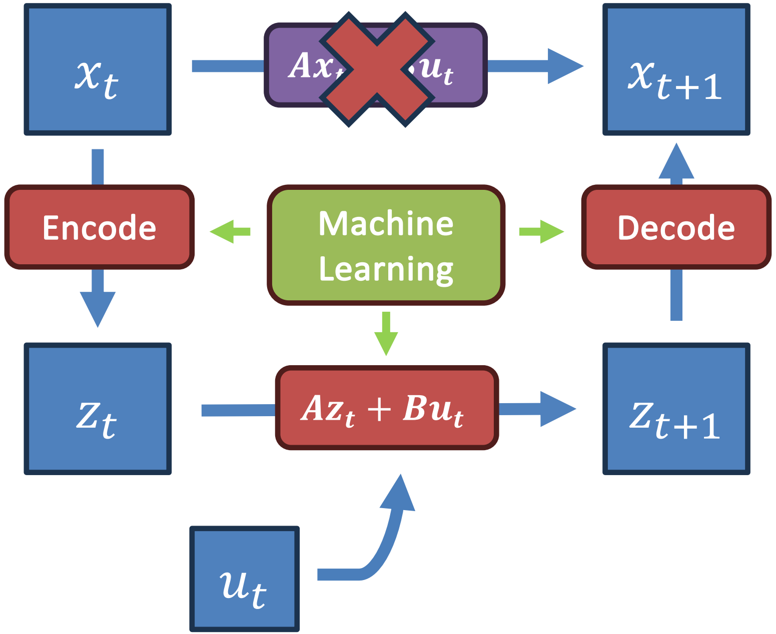
$$\boldsymbol{Cost} = \sum_t \underbrace{(x_{target} - x_t)^T Q (x_{target} - x_t)}_{\text{Tracking Error}} + \underbrace{u_t^T R u_t}_{\text{Control Effort}}$$

- MPC efficiently finds the optimal (cheapest) actuator trajectory to reach a desired state
- Requires linearized dynamics model of the plasma, but we know plasmas are strongly nonlinear!
- How can we control in real-time?

Plasma Control

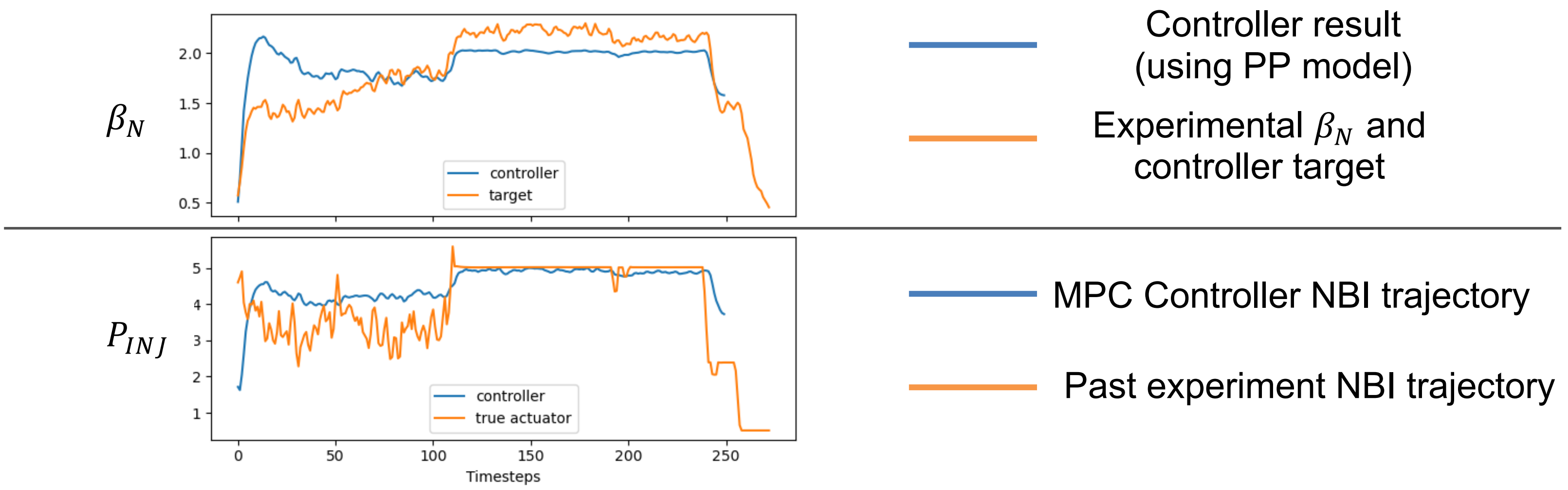
ML Linear Projection

- Nonlinear plasma behavior can be approximately mapped to a larger linear space
- The encoder, decoder, A, and B matrices are learned from DIII-D data
- MPC can be applied to this linear model to find optimal actuator trajectories



[M. Watter, 2015]

Testing out MPC Controller



- Proof of concept: control $\boldsymbol{\beta_N}$ with NBI heating
- The controller finds similar actuator trajectory to experiment.
- Working on full profile controller given a broader set of actuators
  - NBI power and torque, ECH heating, Ip, Bt, shaping and gas injection

Experimental validation hopefully this year…

Wanna bet who will run first?

NSTX-U

Plasma Control

Conclusions and Future Work

- ML can be good!
  - When used in correct situations
- Instability event predictors can be used to develop controllers or be integrated into control systems as safety alarms
- Profile predictor is an offline tool that can simulate full shots based on true machine actuators
  - Physics simulations require mapping artificial diagnostics and actuators
- Can we learn physics from the linear mapping learned by ML model?
  - Perhaps…

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Award DE-FC02- 04ER54698. In addition this material was supported by the U.S. Department of Energy, under Awards DE-SC0015480.

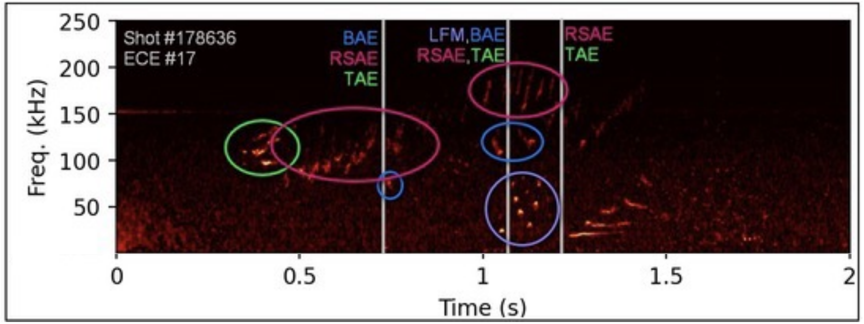
Backup Slides

Alfvén Eigenmode Detection and Control

Plasma Control

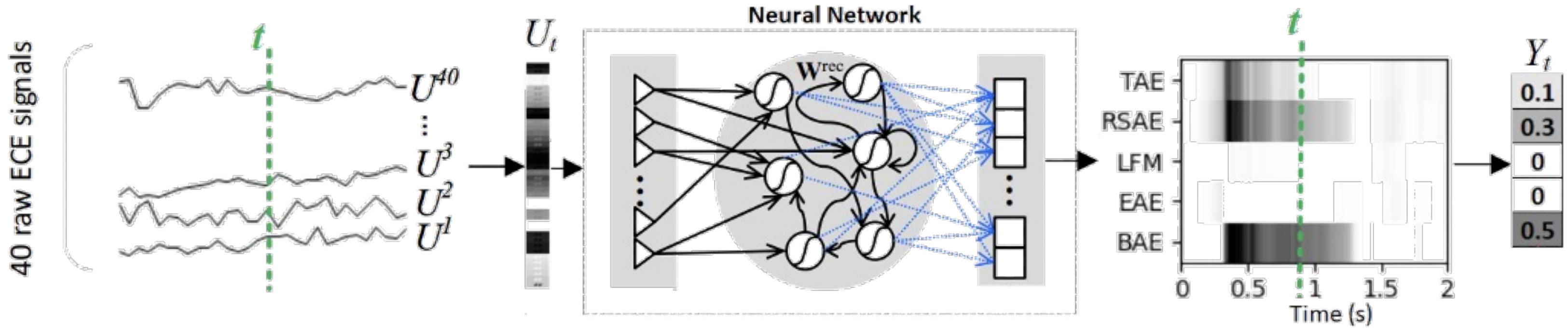
AE Control Introduction

- AEs occur at undamped resonances in the Alfvén continuum
- AEs degrade confinement and have potential to release enough energy to damage vessel walls
- Can be most easily identified in spectrograms of fluctuation diagnostics like ECE or CO_2
- Most straightforward form of control is adjusting P_{NBI}
  - Can be controlled by anything that will adjust Energetic Particle distribution



[Jalalvand NF 2021]

ECE-Based AE Detection

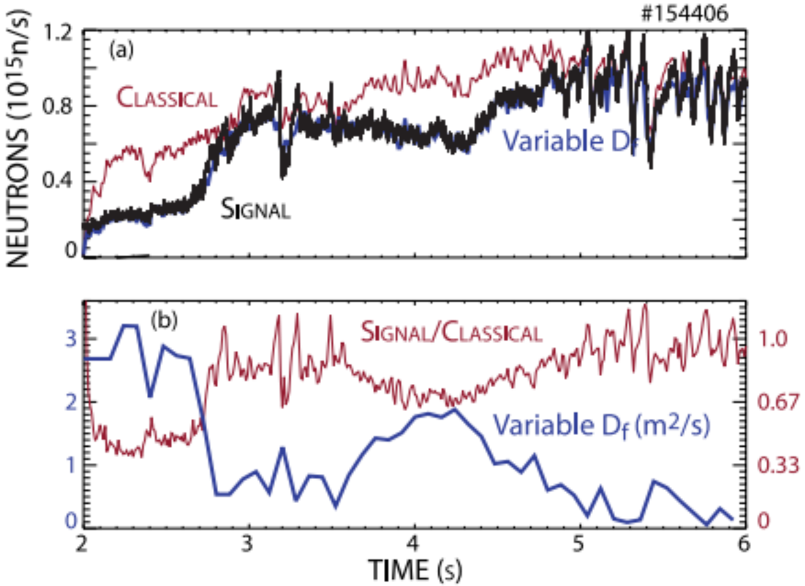


- Input: 40 high frequency ECE channels
- Output: Labels for 5 flavors of AEs
- Hand-produced labels provide 450 training and 150 validation shots
  - Skewed to RSAE and TAE activity
- Achieves >90% true positive rate with <10% false positive rate
- Runs real-time in ≈ 0.5ms

[Jalalvand NF 2021]

AE Detection by Neutron Rate

- When AEs present, classical neutron rate will be much larger than measured neutron rate
- Fastest physics calculation from RABBIT [Weiland NF 2018] still takes hundreds of ms
- Need a faster model to produce classical neutron rates



[Heidbrink PPCF 2014]

Classical Neutron Rate Prediction

- Uses NUBEAM as part of TRANSP to produce *Classical Neutron Rate* along with other NBI-related data
- Uses shape information and profiles from rtEFIT and rtThomson
- Profiles are reduced by PCA to 4 components each
- Runs real-time in ≈ 0.25ms

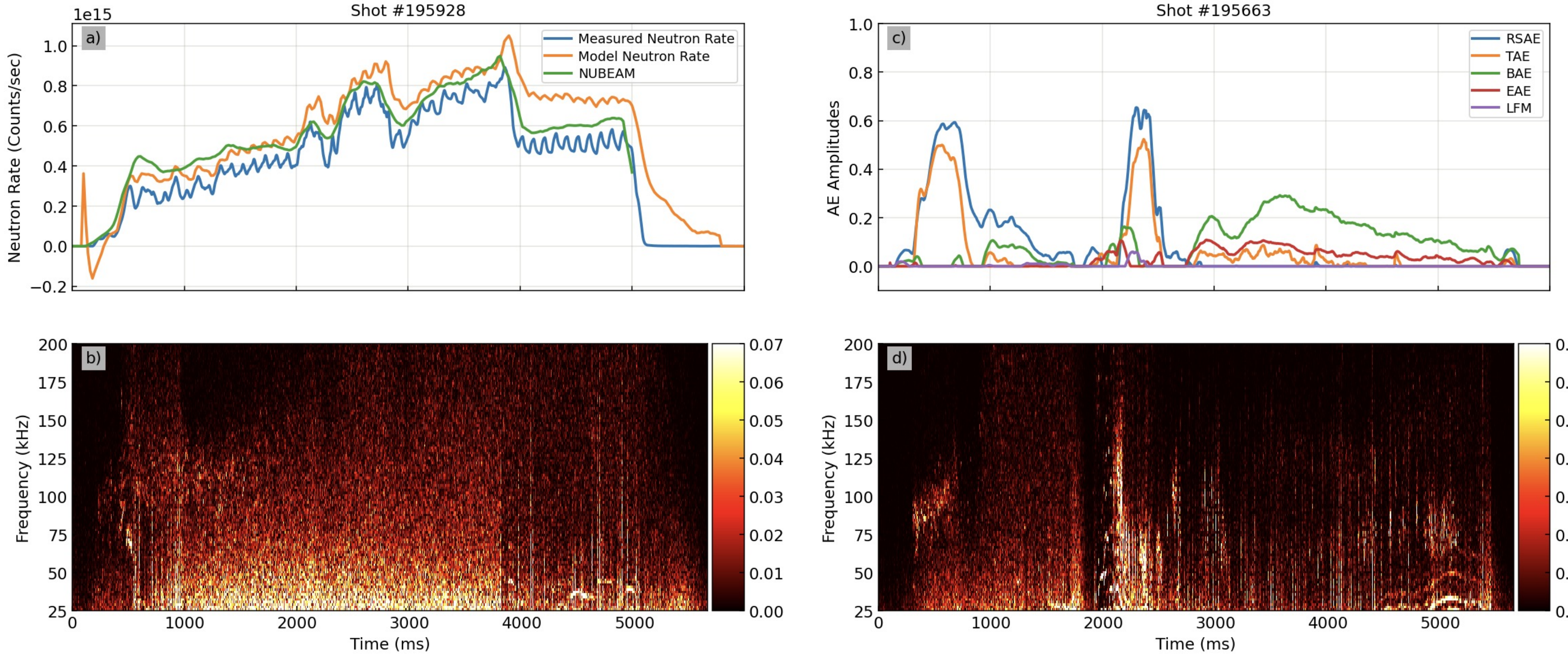
Inputs	Description	Outputs	Description
R_0	Major radius	$s_{neutron}$	Classical neutron rate
κ	Elongation	P_{shine}	Shine through power
I_p	Plasma current	P_{cx}	Charge-exchange loss power
a	Minor radius	P_{orb}	Orbit loss power
B_T	Vacuum toroidal field	$P_{b,e}$	Beam heating electron profile
δ_u	Upper triangularity	$P_{b,i}$	Beam heating ion profile
δ_l	Lower triangularity	$T_{b,e}$	Beam torque electron profile
P_{inj}	Injected power for each beam	$T_{b,i}$	Beam torque ion profile
T_{inj}	Injected torque per beam	n_b	Beam ion density profile
V_{inj}	NBI voltages per beam	j_b	Beam current drive profile
n_e	Electron density profile	p_{fast}	Fast ion pressure
T_e	Electron temperature profile		
q	Safety factor profile		
p	Plasma pressure profile		

[Rothstein in review]

Plasma Control

Real-time Model Results

CO_2 Interferometer Crosspower

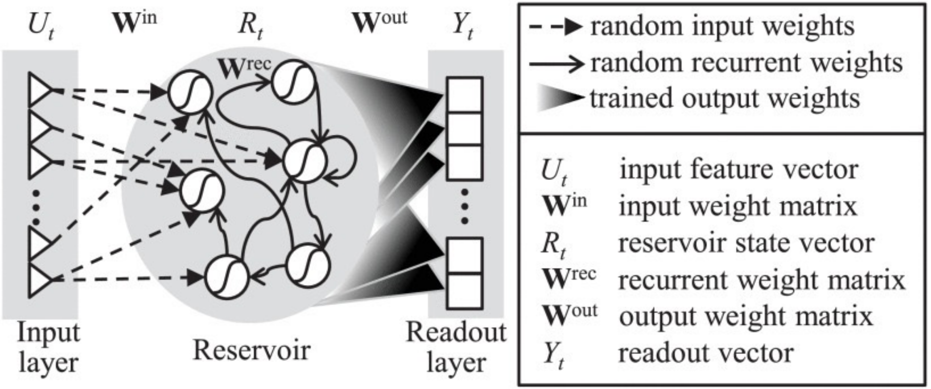


AST Seminar/ February 2024

Plasma Control

Benefit of RCN Architecture

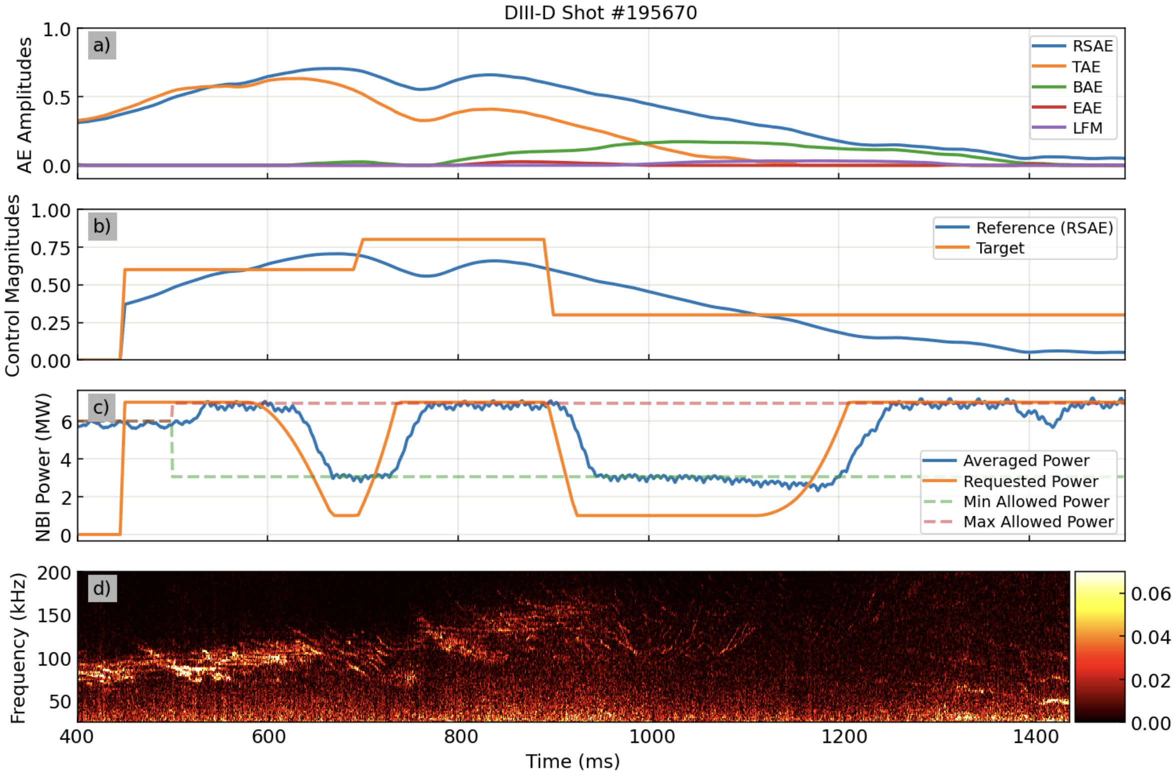
- Lightweight models that show comparable performance to large deep learning models
  - Recurrent connections provide 'memory'
- Basic structure easy and fast to implement on PCS
  - $<$ 1ms to run both ECE and Neutron rate models
- Since only W_{out} needs to be retrained, models can be changed without PCS changes



[Jalalvand NF 2021]

Experimental Control Results

- Single actuator proportional control using P_{NBI}
- Targeted AE activity in ramp-up so limited time for control
- AE amplitude follows target, but highly delayed



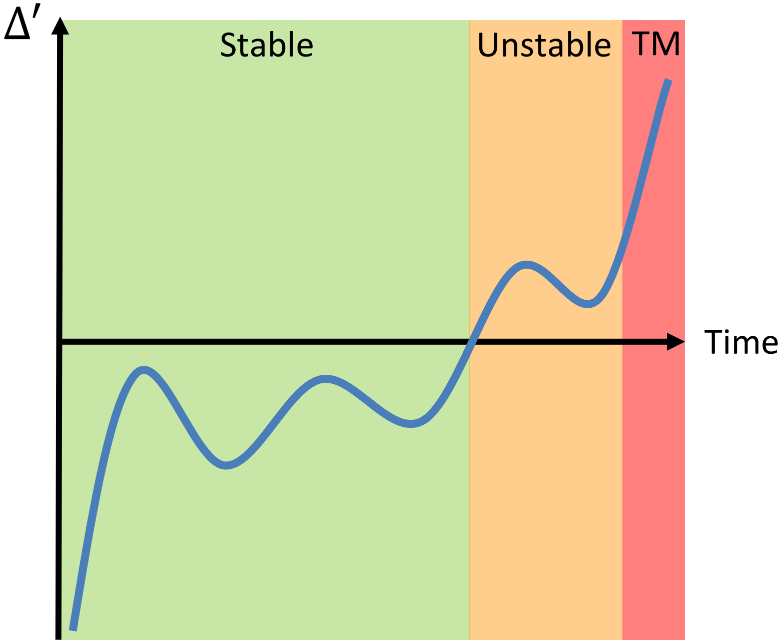
Present Limitations

- ECE detector is limited by cutoff and can be unreliable as viewing positions change
- Neutron Model requires rtEFIT and rtThomson inputs
  - Unreliable before ~1sec
- Single actuator control not the most realistic when there are other objectives
- Need to explore other ways to affect EP distribution:
  - NBI voltage modulation
  - ECH
  - Shaping parameters

STRIDE Development

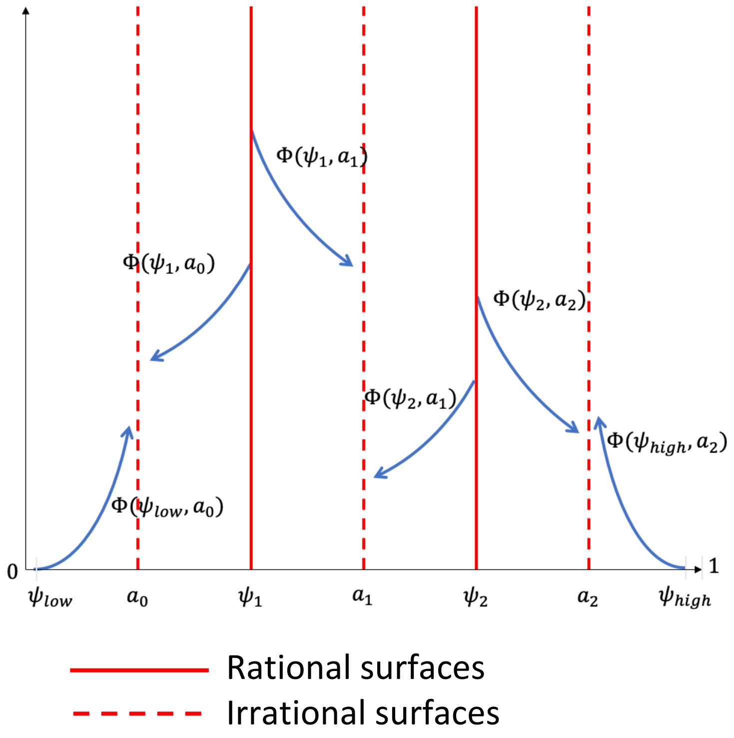
Goal for STRIDE

- Want a metric that describes TM stability – Δ'
- Must be robustly correlated with stability
- Calculation must be reproducible and reliable on database of shots



STRIDE Δ' Calculation

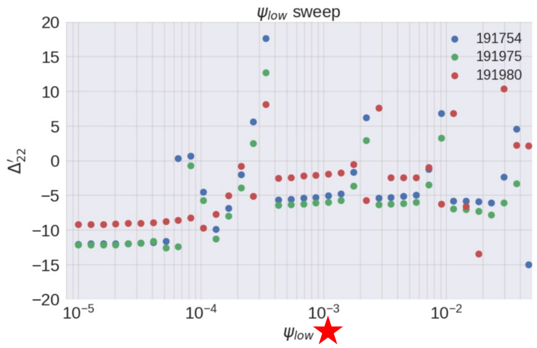
- $\boldsymbol{\Delta}'$ calculated from integration of state transition matrix Φ
  - Subintervals split at rational surfaces and at locations in between rational surfaces
  - Allows for parallelization of integration
- Shooting method integrates away from rational surfaces
  - Matching condition at rational surfaces



A. S. Glasser *et al.* 2018 *Phys. Plasmas* **25** 082502

Current Problems with STRIDE

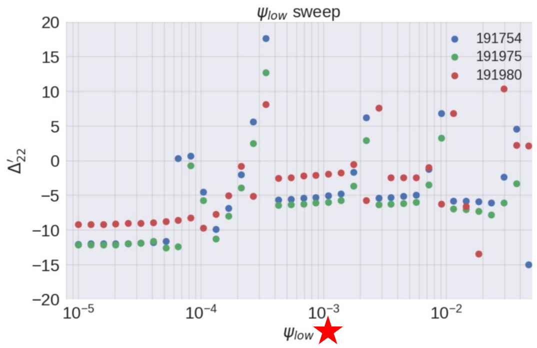
- Bounds of integration are mildly problematic
  - Should be fixable by adjusting the grid packing algorithm
  - ★ Typical parameter value

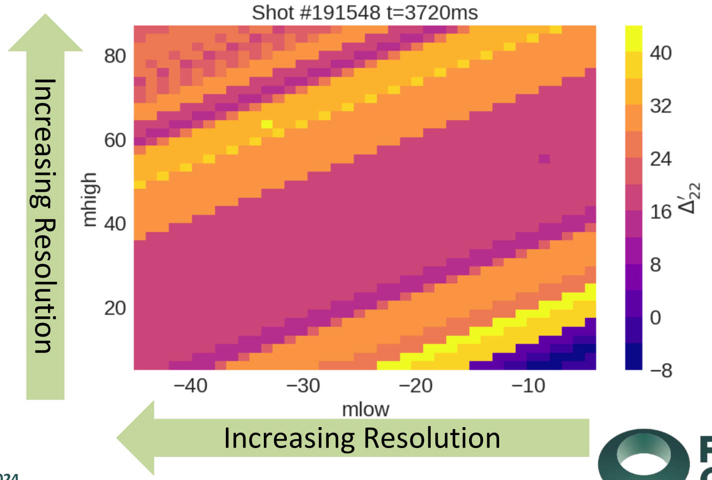


Plasma Control

Current Problems with STRIDE

- Bounds of integration are mildly problematic
  - – Should be fixable by adjusting the grid packing algorithm
  - ★ Typical parameter value
- Increasing number of Fourier modes changes result
  - – Not present in RDCON

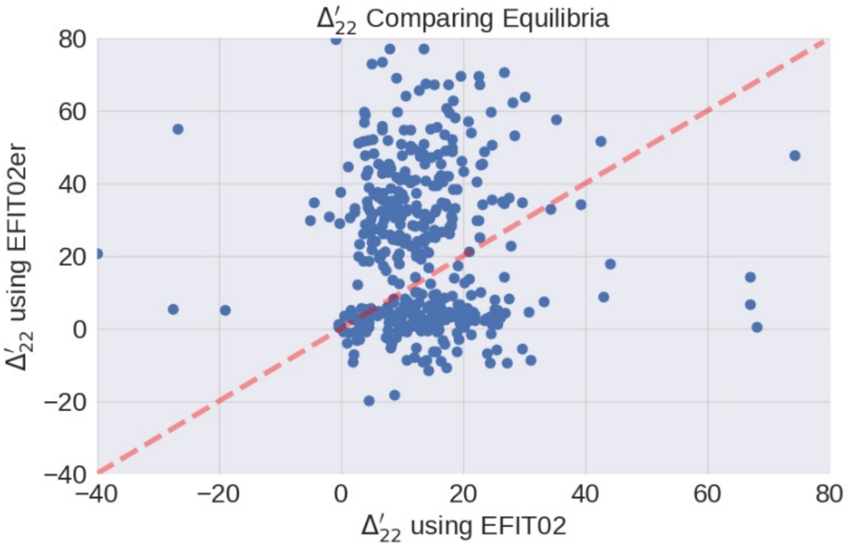


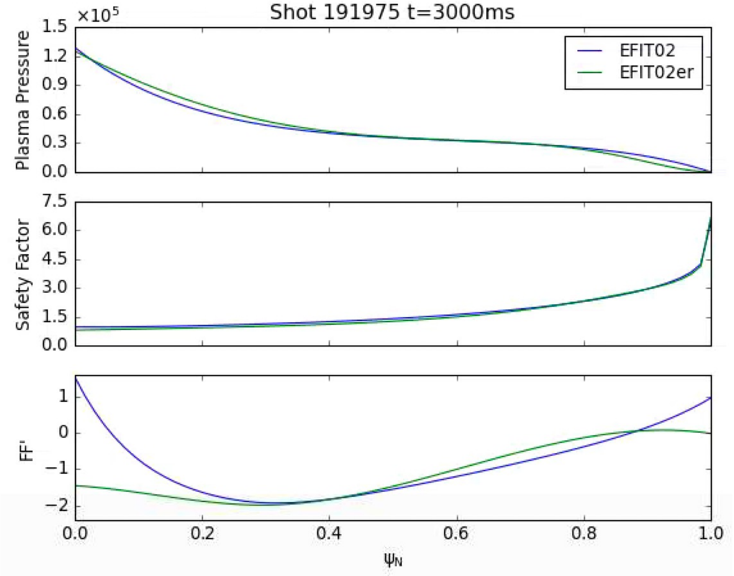


Plasma Control

STRIDE Profiles Dependence

- Profile quality matters a lot $\boldsymbol{\Delta}'$ and real-time profiles need to be closer to offline, kinetic-constrained equilibria



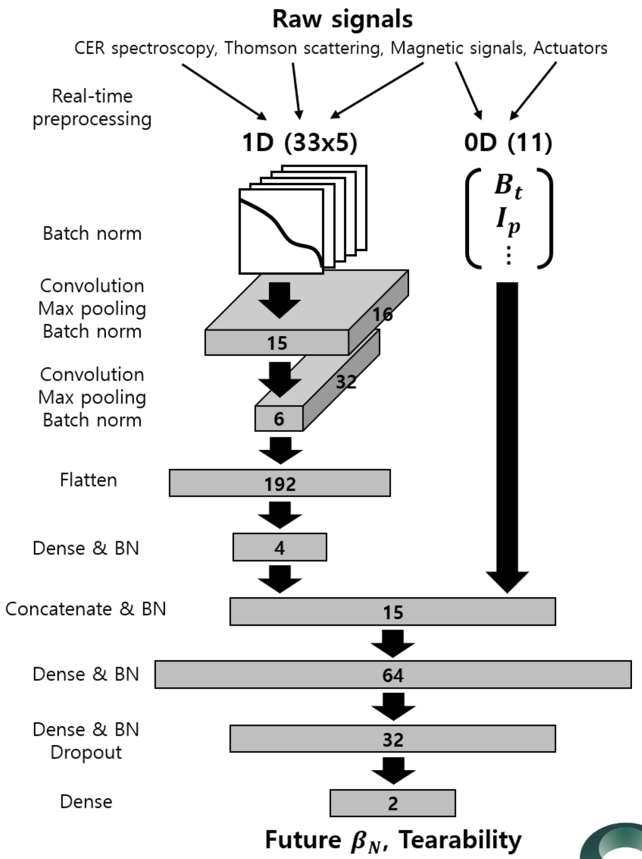


Plasma Control

TM/RL Back-up Slides

Tearing Mode Predictor

- 12,086 parameters
  - $\approx 50\times$ more time slices than parameters
- Ensemble of 10 models provided uncertainty estimates
- Memory-free model
  - No LSTMs or other recurrent layers
  - No sense of how the profiles are evolving, adding memory could be promising



TM Model Selection

- Mean-squared error (MSE) loss

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}[(y_{1,i} - \hat{y}_{1,i})^2 + (y_{2,i} + \hat{y}_{2,i})^2]$$

- Binary cross entropy (BCE) loss

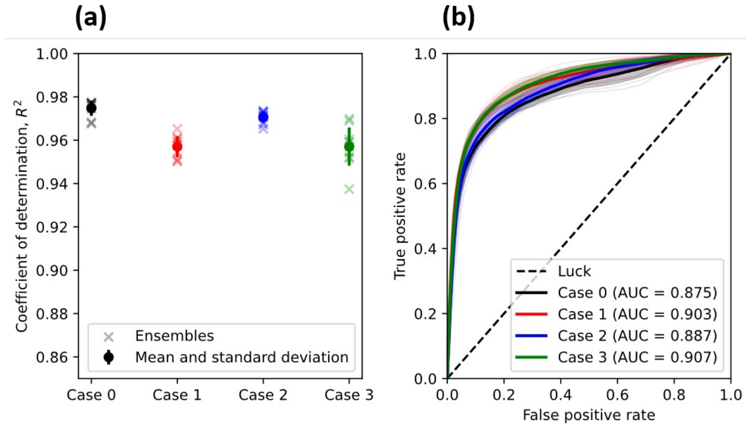
$$L_{BCE} = \frac{1}{N}\sum_{i=1}^{N}[(y_{1,i} - \hat{y}_{1,i})^2$$

$$\boxed{-w_{BCE}(y_{2,i}\log\hat{y}_{2,i} + (1 - y_{2,i})\log\hat{y}_{2i})^2]}$$

  – Generally better for binary prediction tasks

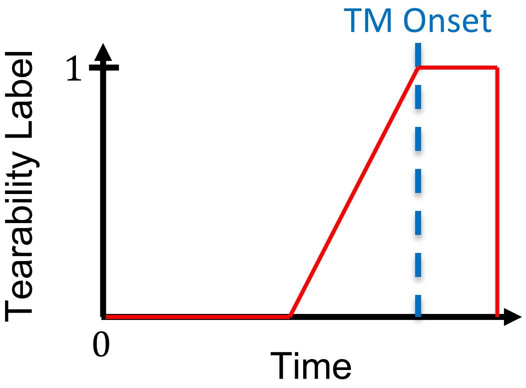
TABLE II
THE INFLUENCE OF THE TYPE OF LOSS FUNCTION AND OVERSAMPLING

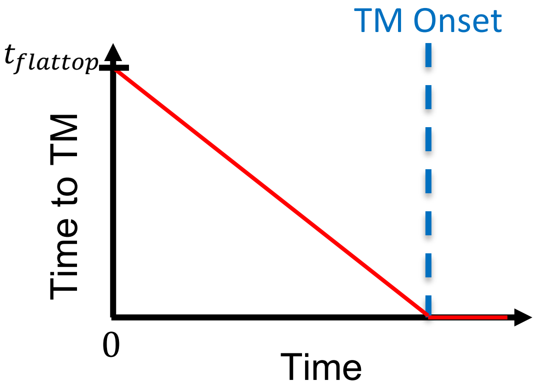
Case number	Loss for β_N	Loss for tearability	Over-sampling	R^2 for β_N	AUC for tearability
0	MSE	MSE	No	0.975	0.875
1	MSE	MSE	Yes	0.957	0.903
2	MSE	BCE	No	0.971	0.887
3	MSE	BCE	Yes	0.957	0.907

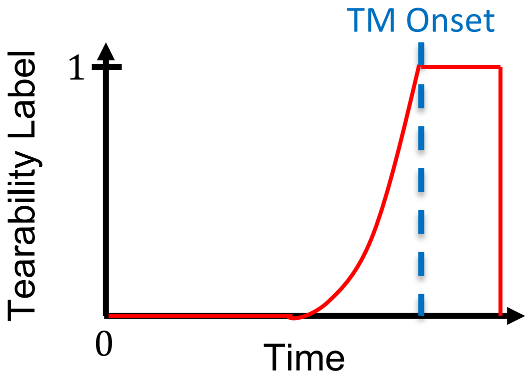


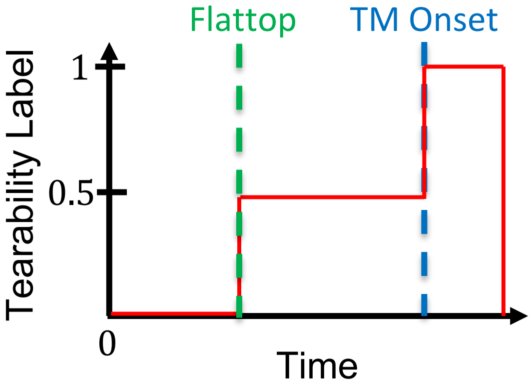
Plasma Control

How to incorporate stability physics into labels?





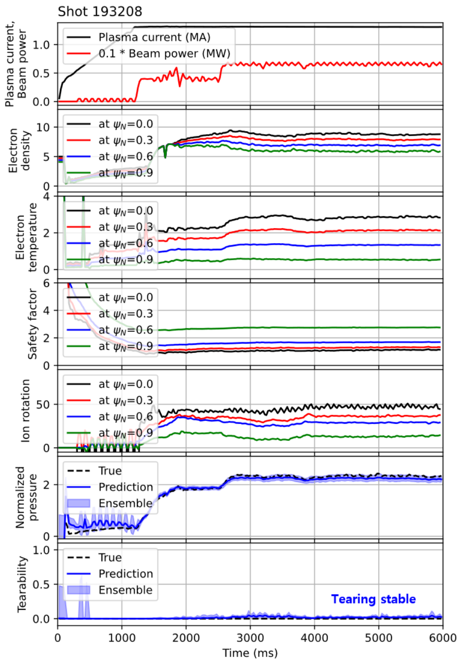


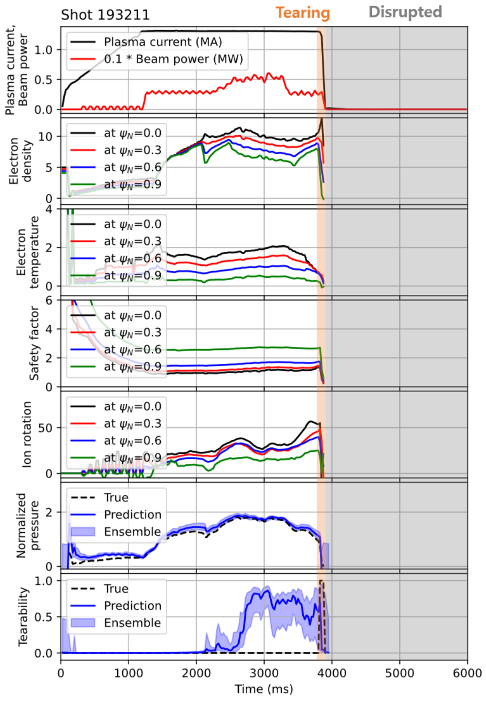


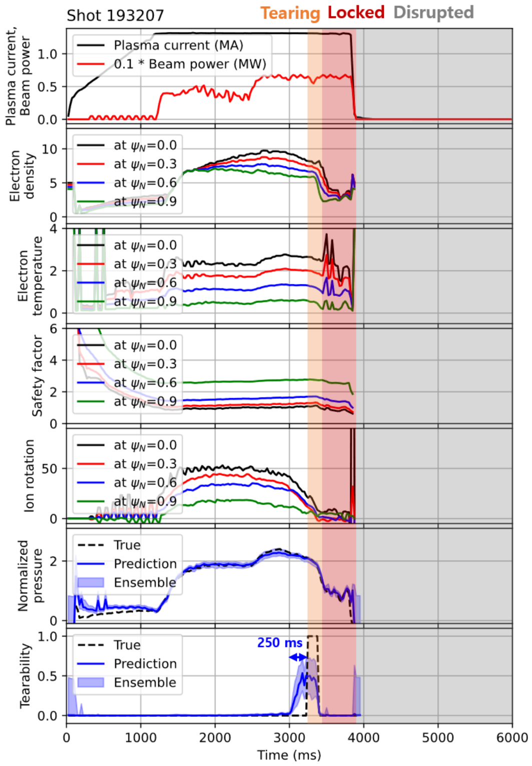
Discussion on Labels and Predictor

- Even with caveats to TM labels, model has good performance at predicting TMs
  - At present, no sense of marginal stability
- Incorporating physics insights to improve labels seems like it would provide better performance
- All of this starts incorporating biases, so need to be very careful with changes made

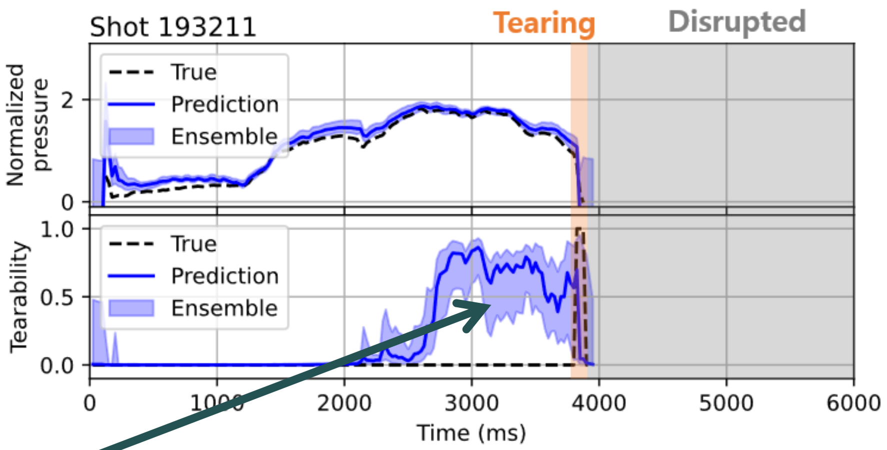
TM Predictor Results



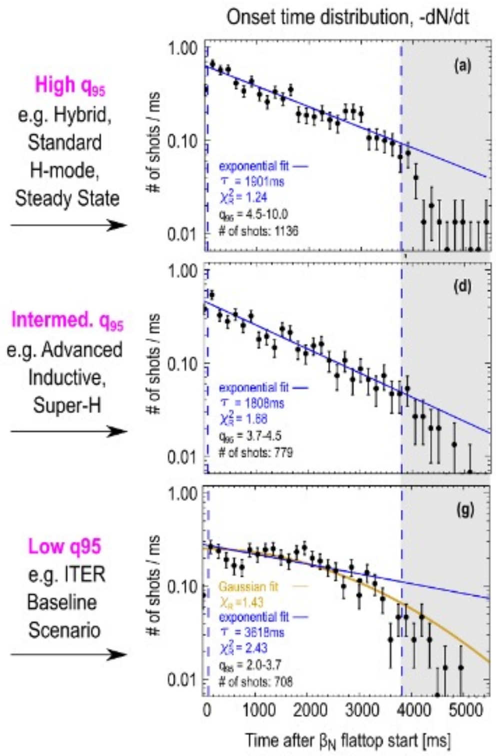




TM Predictor Results



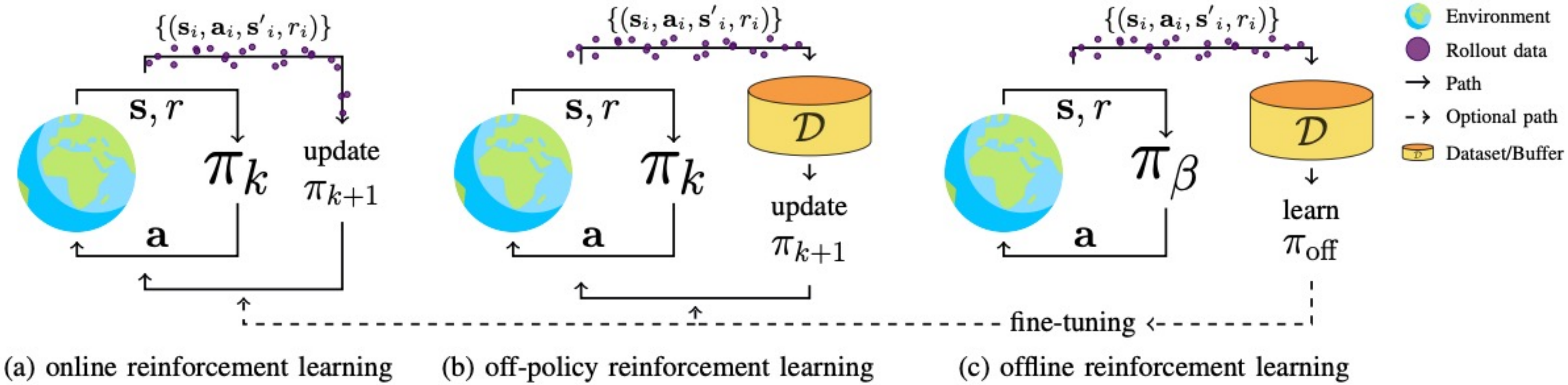
- Unstable but waiting for seed event?
- Losses and AUC metric punish this behavior, but we want this early warning



[J. Seo et al, IJCNN, 2023]

[Bardóczi NF 2023]

Reinforcement Learning Overview



- Cannot test new policies on environment
- Restricted to offline RL

- Train ML simulator for environment – learn a single instability
- Environment models gives intuition to plasma behavior

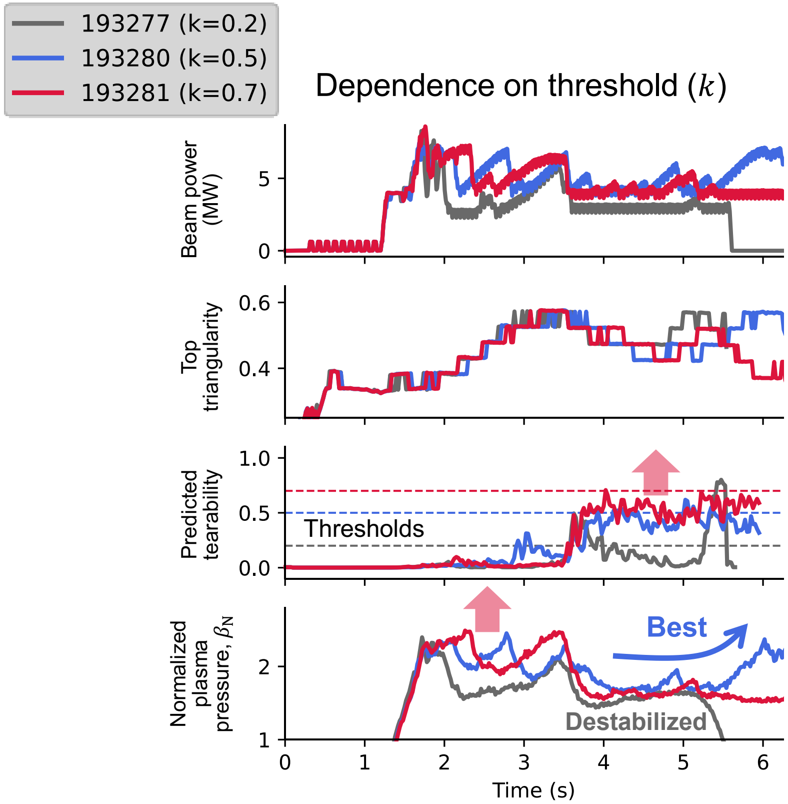
[Prudencio, 2022]

RL Background

- Deep Deterministic Policy Gradient
  - Actor-critic type of RL
  - Learns $Q^\star$ (value function) and $a^\star$ (optimal policy) concurrently
- Uses off-policy data to learn Q function
- Uses Q function to learn policy

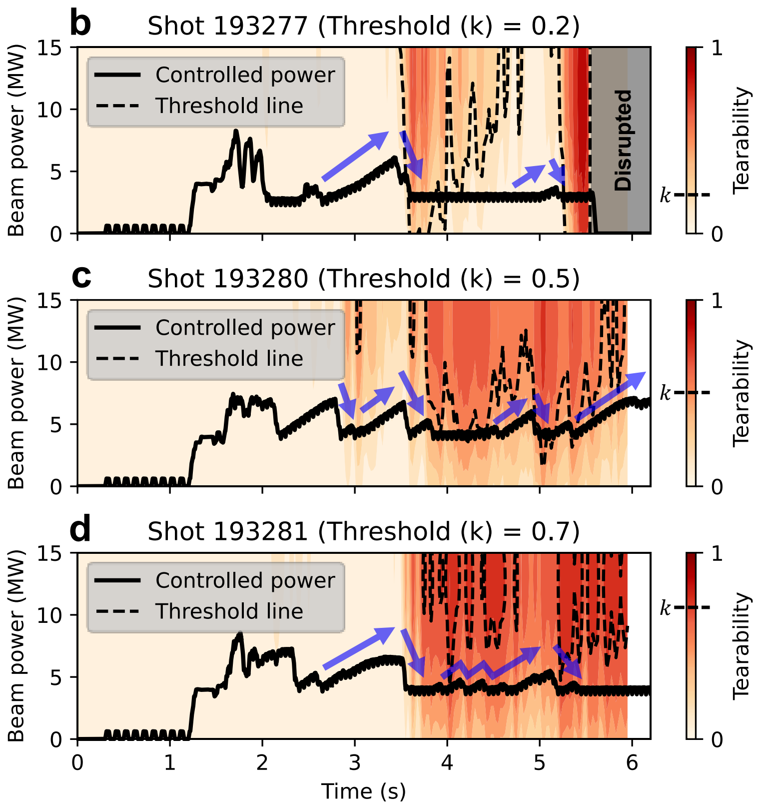
RL Threshold Dependence

- Lower threshold = safer controller
  - Less likely to cause TMs but at cost of lower β_N
- Higher threshold = riskier controller
  - Higher β_N but more likely to cause TMs
- Moderate threshold found most effective in experimental shots on DIII-D

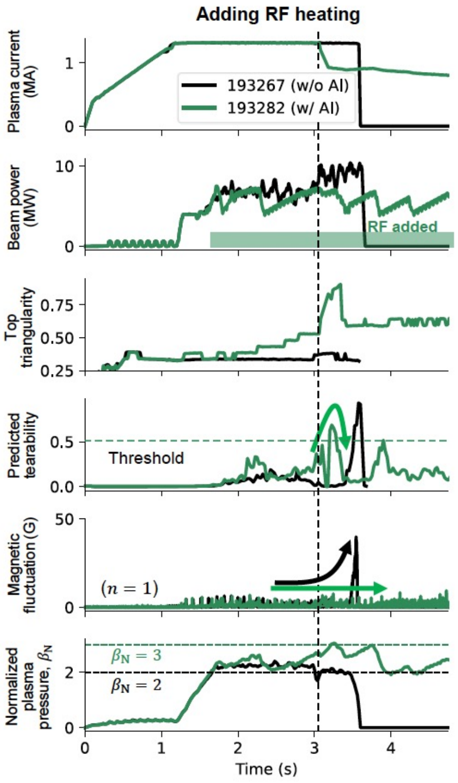


[Seo Nature 2024]

RL Results



b
Shot 193277 (Threshold (k) = 0.2)
Controlled power
Threshold line
Disrupted
c
Shot 193280 (Threshold (k) = 0.5)
Controlled power
Threshold line
d
Shot 193281 (Threshold (k) = 0.7)
Controlled power
Threshold line
Beam power (MW)
Tearability
k
Time (s)



Adding RF heating
Plasma current (MA)
193267 (w/o AI)
193282 (w/ AI)
Beam power (MW)
RF added
Top triangularity
Predicted tearability
Threshold
Magnetic fluctuation (G)
$(n = 1)$
Normalized plasma pressure, β_N
$\beta_N = 3$
$\beta_N = 2$
Time (s)

Advantages of RL Control

- An RL agent can balance multiple actuators to take advantage of nonlinear affects
  - RL controller can find complex trajectories classical control cannot
- Long-term trajectory planning
  - This agent is a ”Greedy bandit”
    - Tries to maximize rewards single step into the future
  - Future version can get long term planning by taking advantage of Q-learning

The Heart of Q-Learning

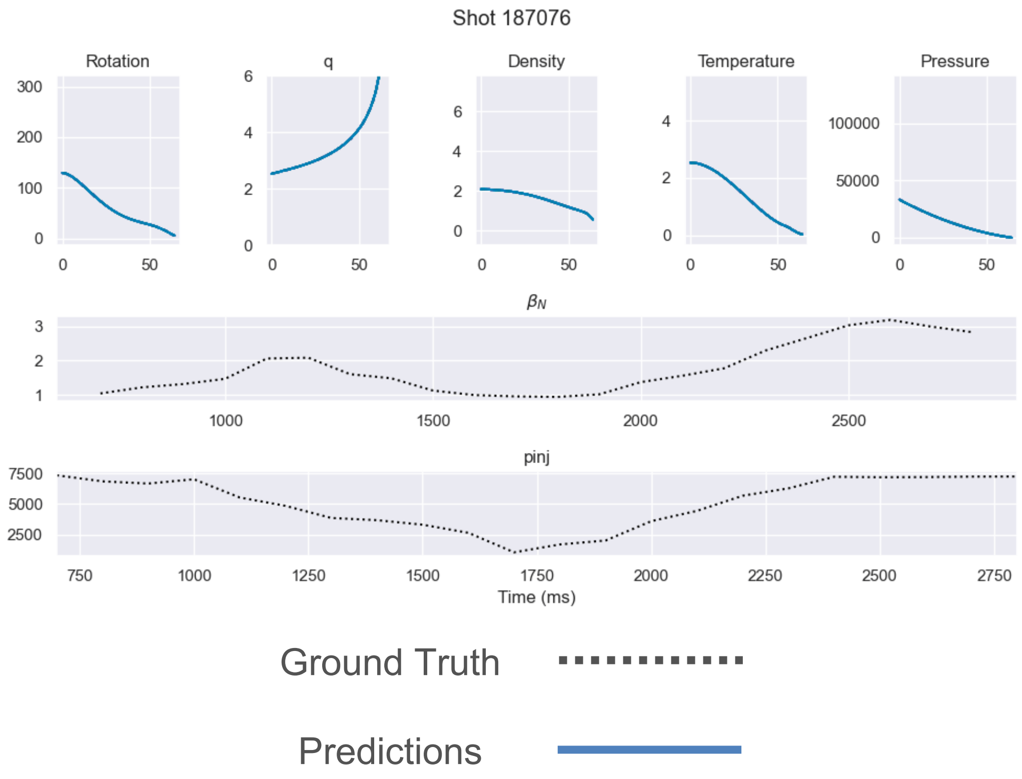
$$\boxed{Q_{new}(s_t, a_t)} \leftarrow \boxed{Q(s_t, a_t)} + \boxed{\alpha}\left[\boxed{r_{t+1}} + \boxed{\gamma}\,\boxed{\cancel{\max Q(s_{t+1}, a')}} - Q(s_t, a_t)\right]$$

- Bellman equation
  - Q – Value function (or just Q-function)
  - s_t, a_t – States and actions at time t
  - α – Learning rate (hyperparameter)
    - How quickly do we change Q
  - r_{t+1} – Reward from doing a_t at s_t
  - γ – Discount rate (hyperparameter)
    - How much do we care/trust the future?

Long Term Planning

$$\max_{a'} Q(s_{t+1}, a')$$

- Need some model that can produce s_{t+1} from (s_t, a_t)
  - This could be a fully physics model!
- In the meantime, more ML



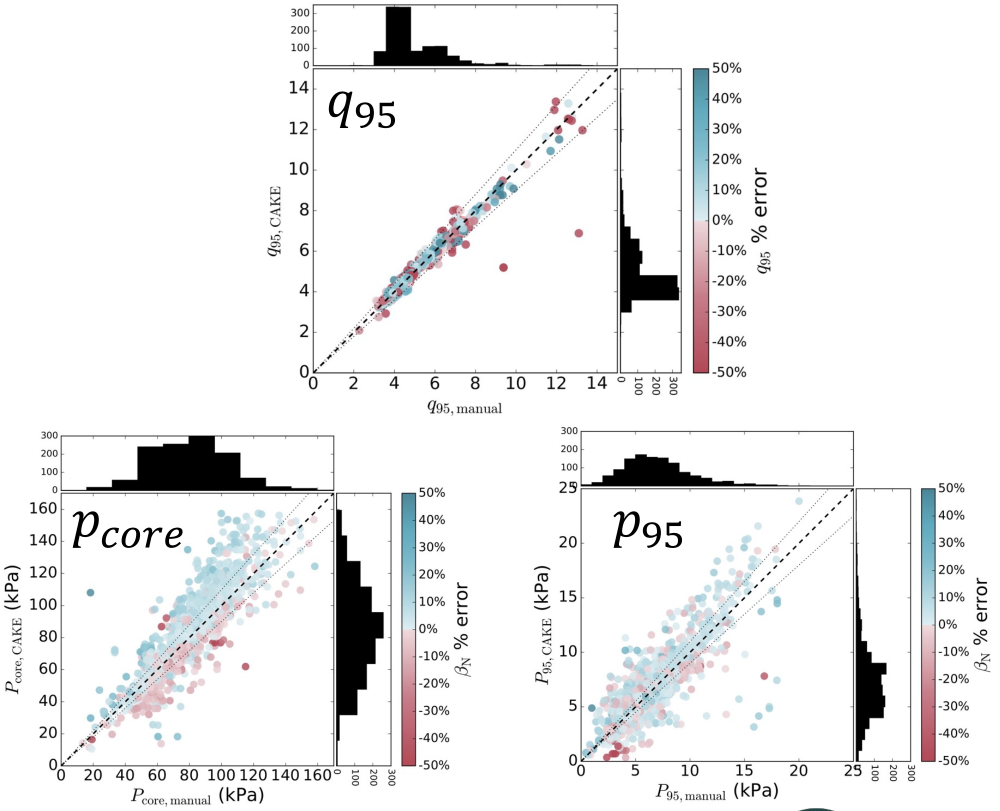
[Courtesy of I. Char]

Plasma Control

Real-time Tools: rtCAKENN and rtGPEC

CAKE Database

- Large database of constrained equilibria
- Reconstructions are more consistent across shots
- Large size + Consistency = Ideal for machine learning!

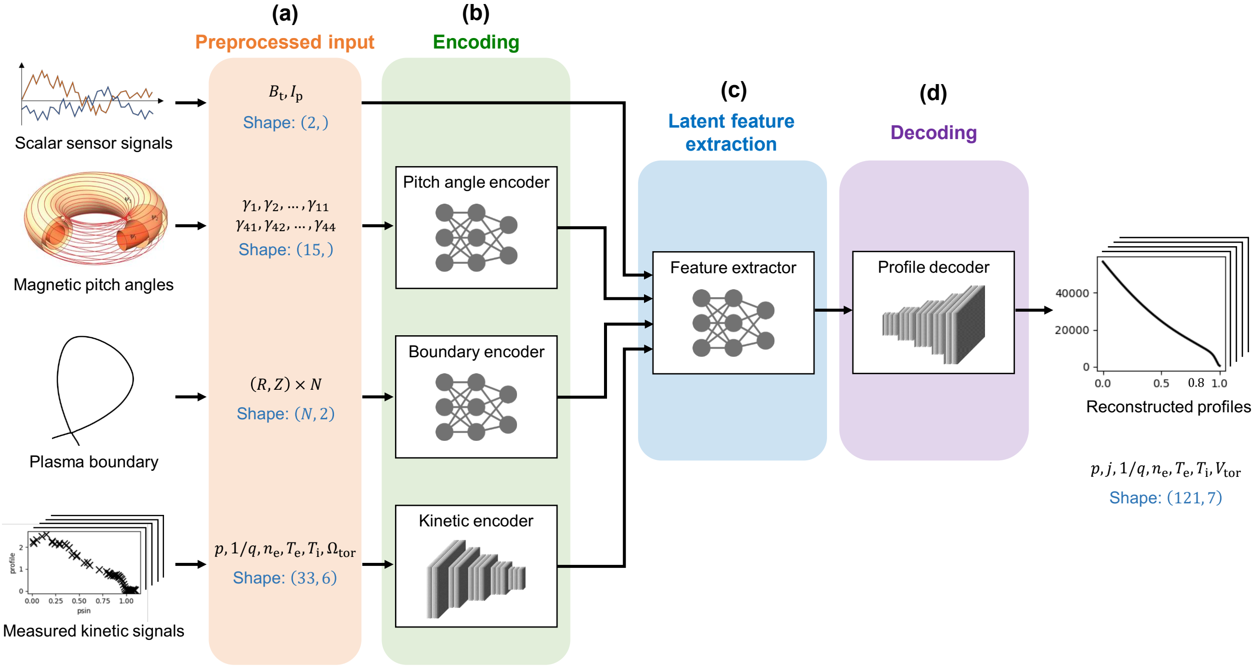


[A.O. Nelson SET Meeting]

Plasma Control

rtCAKENN Architecture

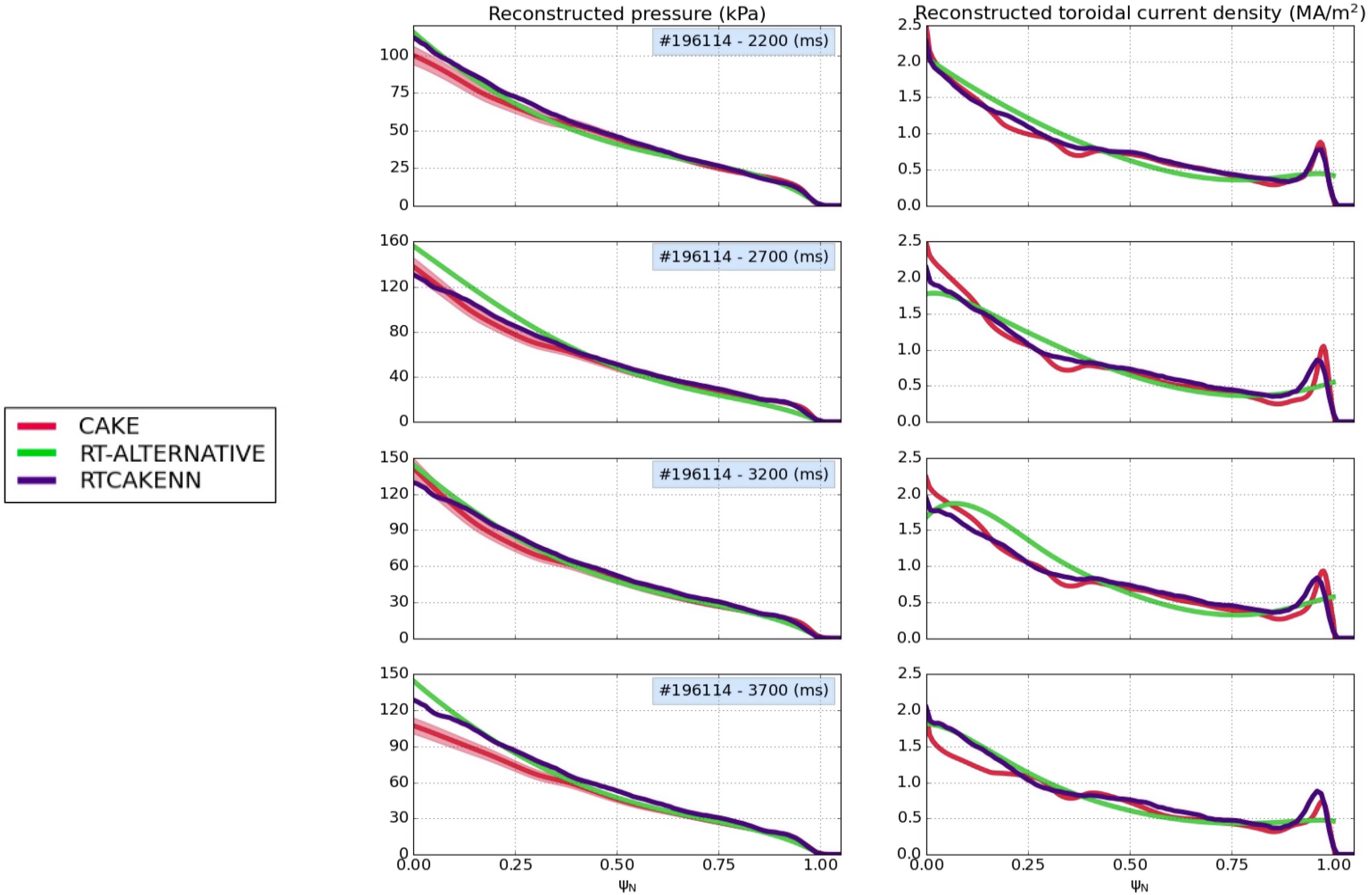
- Ensemble of 10 identical architecture models
  - Improves performance and gives uncertainties
- Produces profiles:
  - $p, J, \frac{1}{q}, n_e, T_e, T_i, v_{tor}$
- Runs in DIII-D PCS in $\approx$ 7.7ms



[Shousha NF 2023]

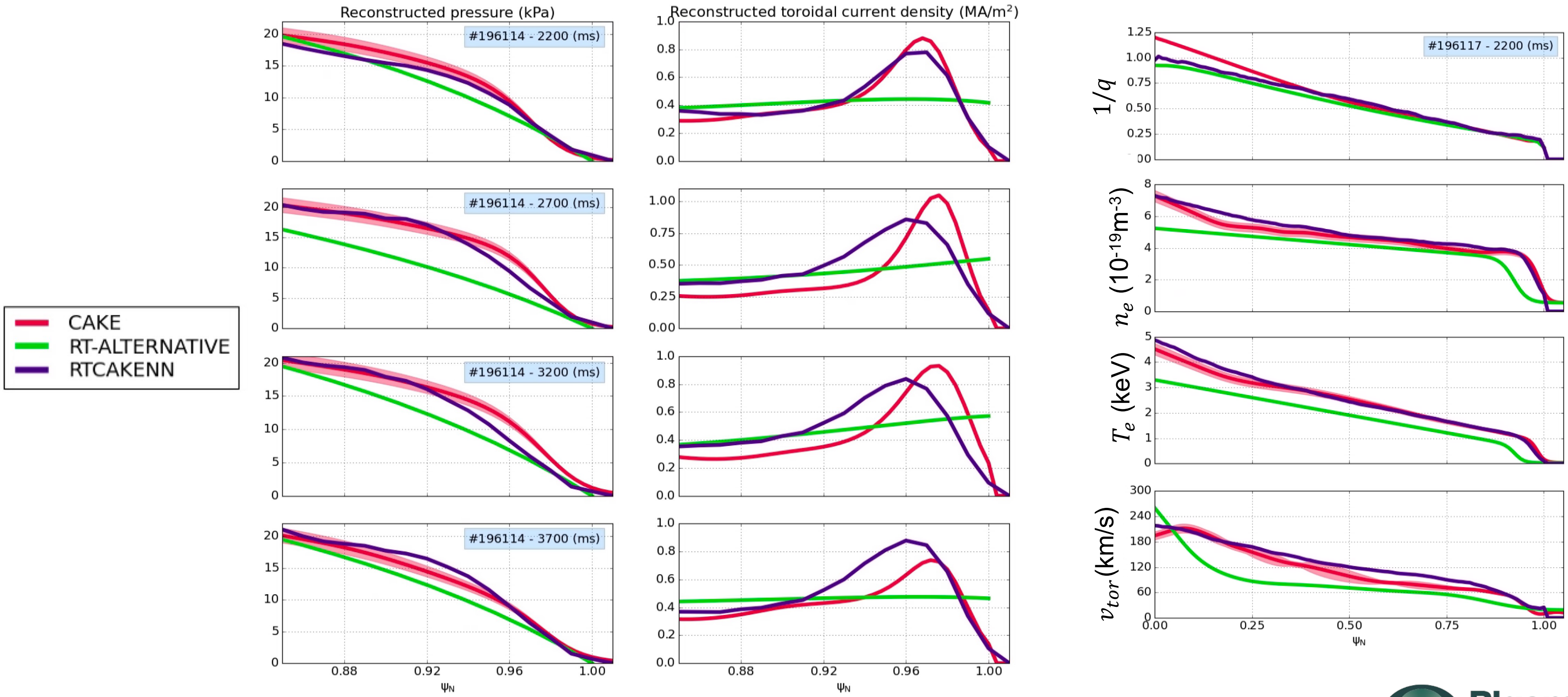
Plasma Control

RTCAKENN agrees with CAKE and is more accurate than RT-Alternatives



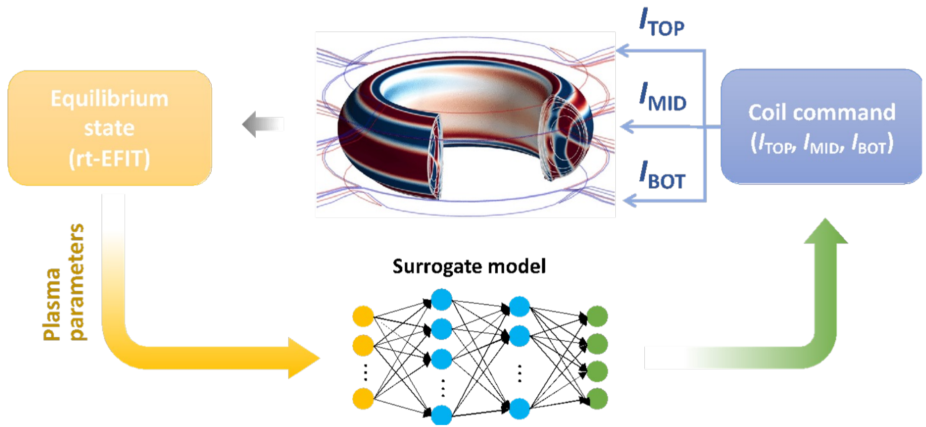
[Shousha NF 2023]

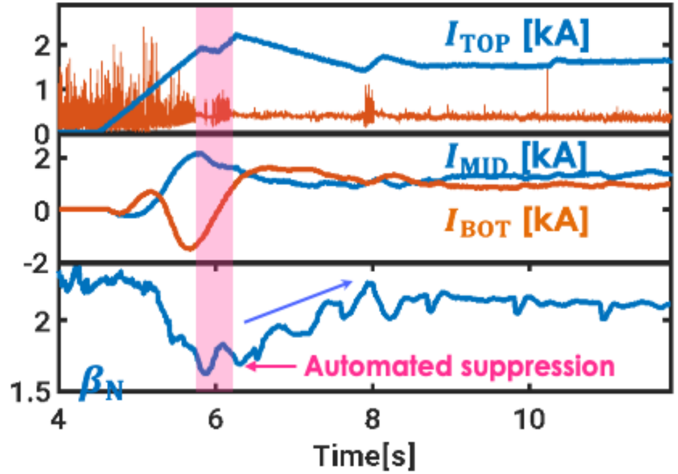
RTCAKENN Captures Pedestal Behavior and J Well



[Shousha NF 2023]

ML-technique is integrated into the adaptive ELM controller for automated safe ELM suppression without empirical approach





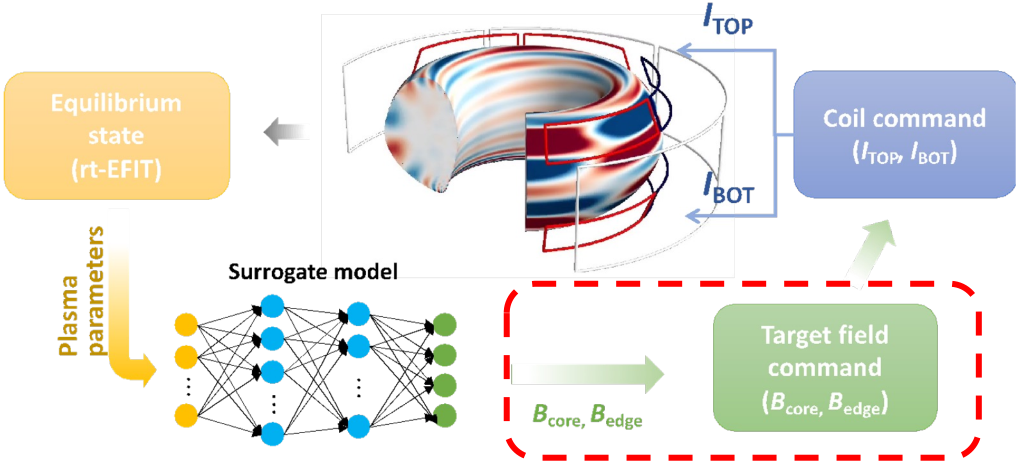
- ➢ An unoptimized spectrum leads to disruption. [C. Paz-Soldan, PRL 15] → Empirical spectrum.
- ➢ An empirical approach is not feasible in ITER.
- ➢ Model-based 3D-coil configuration (ERMP). [J.-K. Park, NP 18, S.M. Yang, NC accepted]
- ➢ ML-surrogate model to accelerate offline derivation (~s → ~ms).
- ➢ Automated & adaptive ELM control in KSTAR without empirical or human decision. → ITER applicable.

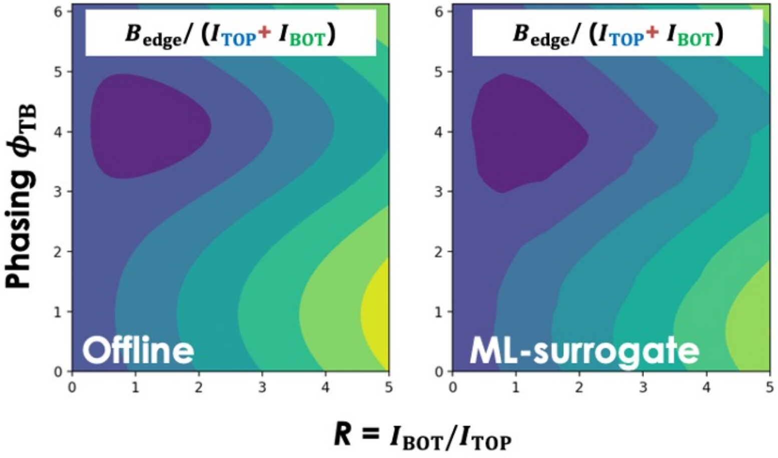
[Courtesy of S.K. Kim]

A more general surrogate model is developed and will be tested in DIII-D

- ➢ **Surrogate model for perturbed 3D-field calculation (rt-GPEC) instead of coil configuration.**
- ➢ **Direct model of core/edge 3D fields ($B_{core/edge}$).**

  - Allows extended and flexible applications (including non-RMP).
- ➢ **Reasonable preliminary results.**





[Courtesy of S.K. Kim]

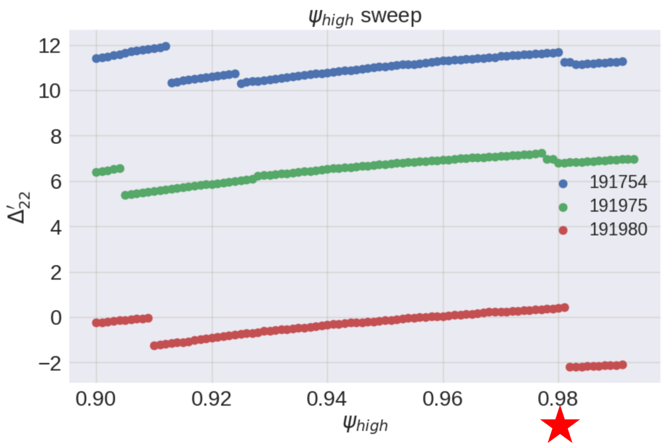
Misc Back-up Slides

Δ' Definition

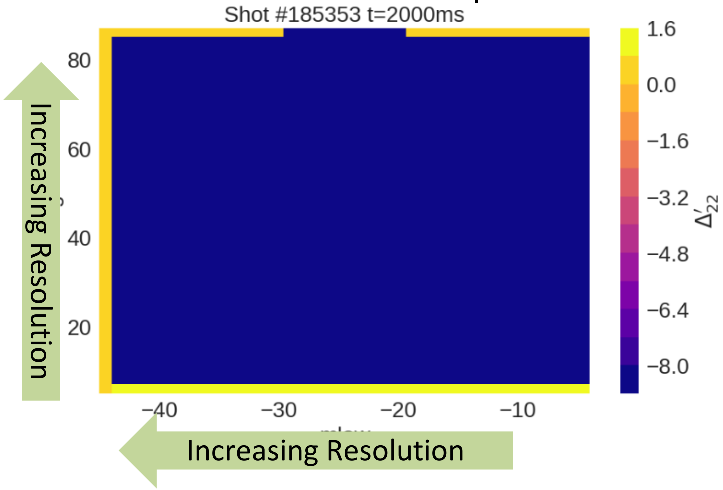
$$\Delta'_{22} = \begin{bmatrix} \Delta'_{LL} & \Delta'_{RL} \\ \Delta'_{LR} & \Delta'_{RR} \end{bmatrix}$$

- For a given rational surface ($q = 2$)
- In slab model definition is a scalar
- In cylindrical geometry matrix is diagonal (no coupling between different m modes)
- In toroidal geometry, toroidicity and noncircularity couple different m modes

STRIDE Parameter Dependence Extras



RDCON Sweep


Shot #185353 t=2000ms
80
60
40
20
−40
−30
−20
−10
1.6
0.0
−1.6
−3.2
−4.8
−6.4
−8.0
Δ'_{22}
Increasing Resolution
Increasing Resolution

STRIDE Equilibria Extras