SciDAC ISEP:
Integrated Simulation of Energetic Particles in Burning Plamsas

Zhihong Lin

University of California, Irvine

SciDAC ISEP Center

UCI, GA, PPPL, ORNL, LBNL, LLNL,
PU, UCSD

ISEP INCITE project

Outlines: ISEP

- Develop EP integrated simulation: ISEP framework
- Study high priority EP physics issues
- Develop EP modules for fusion WDM

[D. Pace, W. Heidbrink, M. Van Zeeland, 2015]

SciDAC ISEP: Integrated Simulation of Energetic Particles

- Confinement of energetic particles (EP) is a critical issue for burning plasma experiments since ignition in ITER relies on self-heating by energetic fusion products (α-particles)
- Plasma confinement properties in the *new* ignition regime of self-heating by α-particles is one of the most uncertain issues when extrapolating from existing fusion devices to ITER
- **EP turbulence and transport**: EP excite meso-scale instabilities and drive large transport, which can degrades overall plasma confinement and threaten machine integrity
- **Interaction between EP and thermal plasmas**: EP can strongly influence microturbulence responsible for turbulent transport of thermal plasmas and macroscopic magnetohydrodynamic (MHD) instabilities potentially leading to disruptions
- SciDAC ISEP: integrated simulations of EP turbulence by treating relevant physical processes from micro to macro scales on same footing

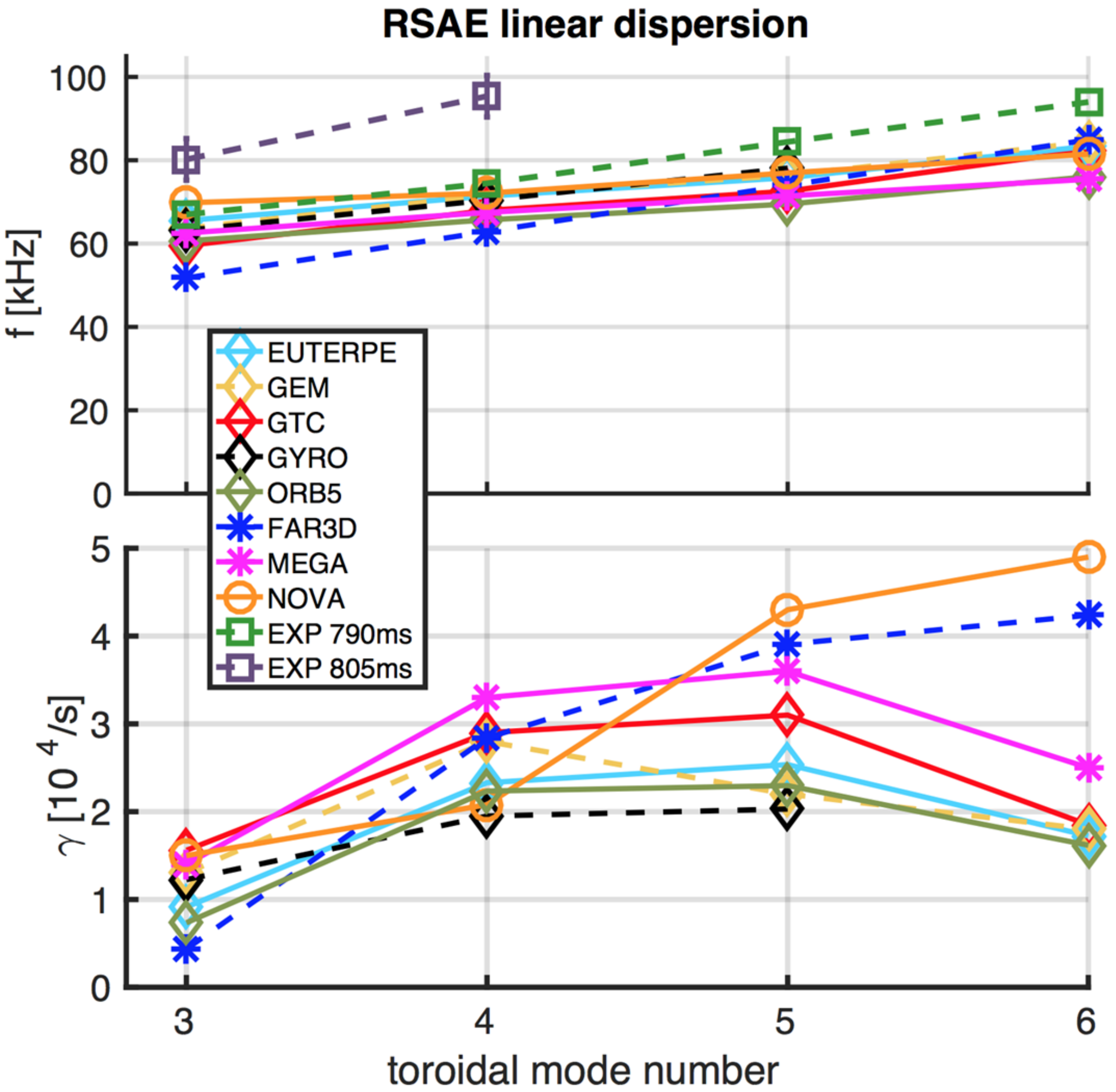
►ISEP Center: UCI, GA, PPPL, ORNL, LBNL, LLNL, PU, UCSD

ISEP Objectives

- Study EP physics needed for predictive capability using 1st-principles GTC, GYRO, FAR3D
  - ► EP transport by mesoscale EP turbulence
  - ► EP coupling with microturbulence and macroscopic MHD modes
- Develop integrated simulation for studying EP physics and verifying reduced transport model
  - ► ISEP framework based on GTC
- Develop EP module with predictive capability for WDM
  - ► Reduced EP transport models (CGM, RBQ, kick model, machine learning)
  - ► First-principles ISEP framework
  - ► V&V
- Computational partnership
  - ► Workflow/data management [*Klasky*]
  - ► Solvers [*Falgout*]
  - ► Optimization & portability [*Williams & Tang*]

Highlights on ISEP Progress

- Developed practical reduced models (CGM, RBQ, Kick) [*1 DPP18 invited talk, 2 IAEA2018 orals*]
- Linear Alfven eigenmode (AE) V&V completed
- Progress on AE NL saturation and coupling with microturbulence
- GTC optimized on GPU [*SC18 workshop oral*]
- ADIOS-2 implemented in GTC

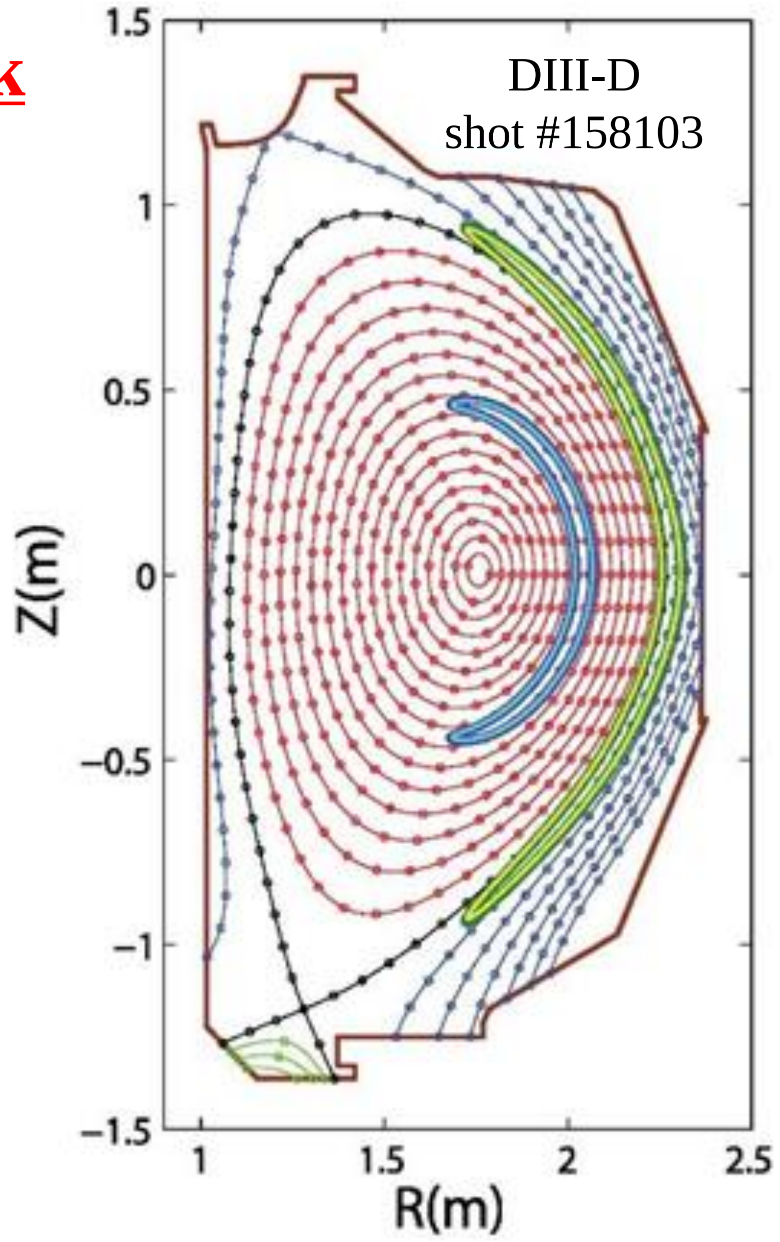


Toroidal mode n=4 RSAE in DIII-D shot 158243 at 805ms

[*Taimourzadeh et al, NF2019*]

Integrated Simulation: 1st-principles ISEP Framework

- ISEP framework based on gyrokinetic toroidal code GTC:
  - ✓ SciDAC GPS (01-11), GSEP (08-17), ISEP (17-)
- First-principles, global, integrated simulation capability for nonlinear interactions of multiple kinetic-MHD processes
- Current capability in the central version
  - ✓ Global 3D toroidal geometry for tokamak, stellarator, FRC
  - ✓ Microturbulence: 5D gyrokinetic ions & electrons, electromagnetic compressible fluctuations, collisionless/collisional tearing modes
  - ✓ MHD and energetic particle (EP): Alfven eigenmodes, kink, resistive tearing modes
  - ✓ Neoclassical transport: Fokker-Planck collision operators
  - ✓ Radio frequency (RF) waves: 6D Vlasov ions
- Large user community (>40 users/developers); Broad impacts to fusion (*12 papers in PRL, Science, Nature Comm*.)



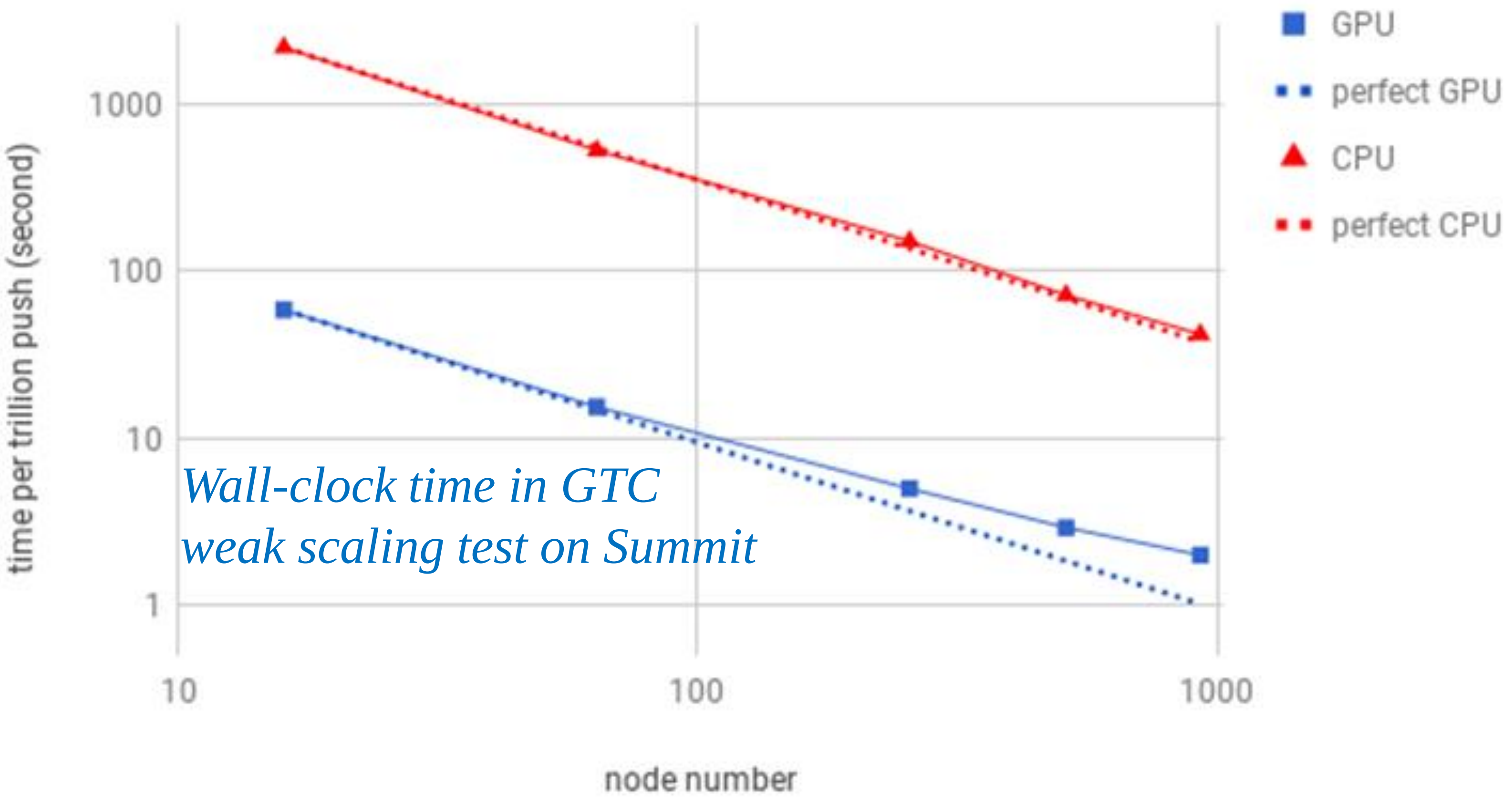
Z. Lin et al, Science 281, 1835 (1998)

Open source: Phoenix.ps.uci.edu/GTC

GTC Performance on Summit GPU

- GPU optimization by CAAR project: UCI, PU, ORNL, NVIDIA, IBM
- GTC speeds up 40x from CPU to GPU on 384 GPUs; speeds up 20x from CPU to GPU on 5556 GPUs (1/5 of SUMMIT)
- INCITE: 2% of Summit time

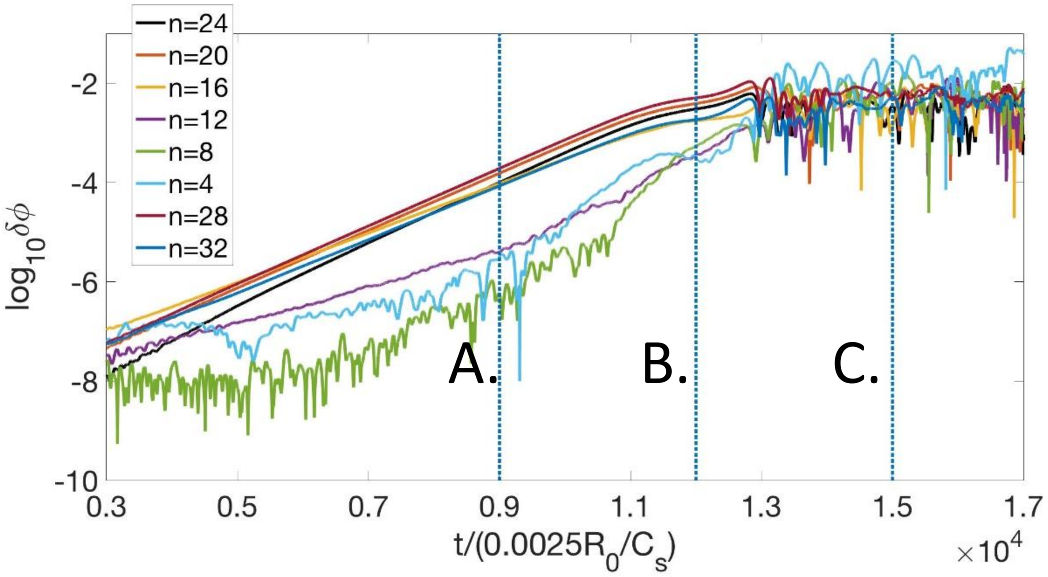
Heterogeneous Programming and Optimization of Gyrokinetic Toroidal Code Using Directives, Wenlu Zhang et al, [*SC2018, WACCPD 2018 Workshop*]



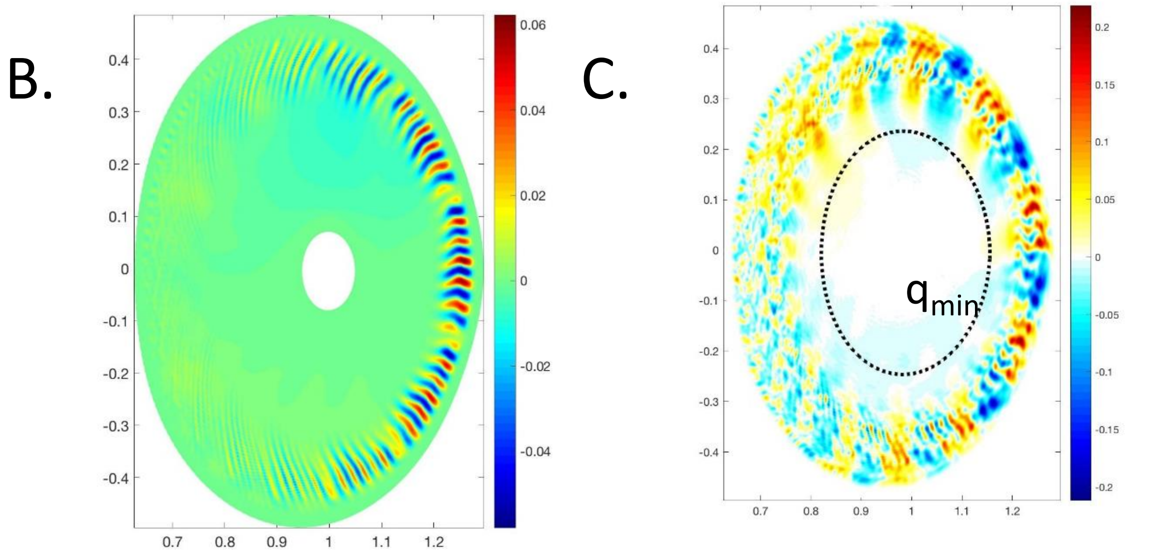
Wall-clock time in GTC weak scaling test on Summit

High Priority Physics Issues: GTC integrated simulation of multiple processes, cross-scale interaction

- Gyrokinetic model for all species: thermal ions & electrons, beam ions
- Linear simulations find unstable low-*n* reversed shear Alfven eigenmode (RSAE), intermediate-*n* toroidal Alfven eigenmode (TAE), and high-*n* ion temperature gradient (ITG)
- Initial nonlinear simulation beyond saturation of RSAE/TAE/ITG



RSAE/TAE/ITG in DIII-D shot 158243 at 805ms



Progress on reduced fidelity models for EP stability and transport is essential for whole device modeling

- **First step: need to rapidly evaluate Alfvén stability and mode structures**
  - Perturbative analysis (NOVA-K, AE3D-K)
  - Non-perturbative gyrofluid closure models (FAR3D, TGLF-EP)
- **Second step: must couple EP stability with energetic particle transport evaluation**
  - Critical gradient models (TGLF-EP & Alpha)
  - Resonance-broadened quasilinear (RBQ) model
  - Perturbative phase space orbits (Kick model)
  - Rapid (GPU-based) fast ion Monte Carlo models with Alfvén mode structures (future versions of AE3D-K)

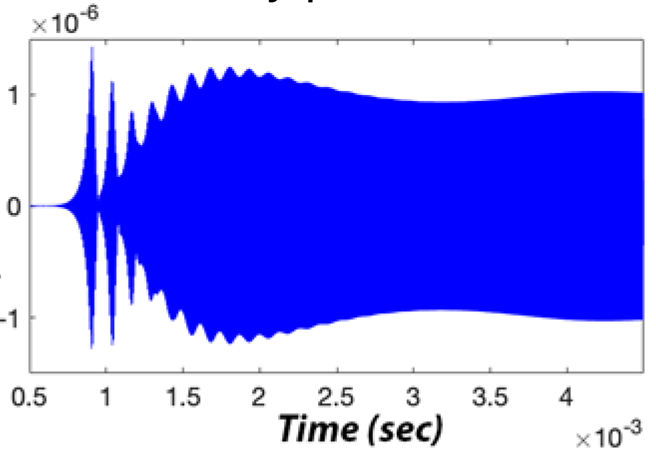
High Priority Issues: long time nonlinear simulations via Computational improvements to the EP gyro-Landau fluid model (FAR3D)

- Source/sink models affect intermittency
- Zonal flows/currents stronger for the fixed EP profile case
- Neoclassical damping of zonal flows recently added
- Will allow first-of-a-kind consistent study of intermittency

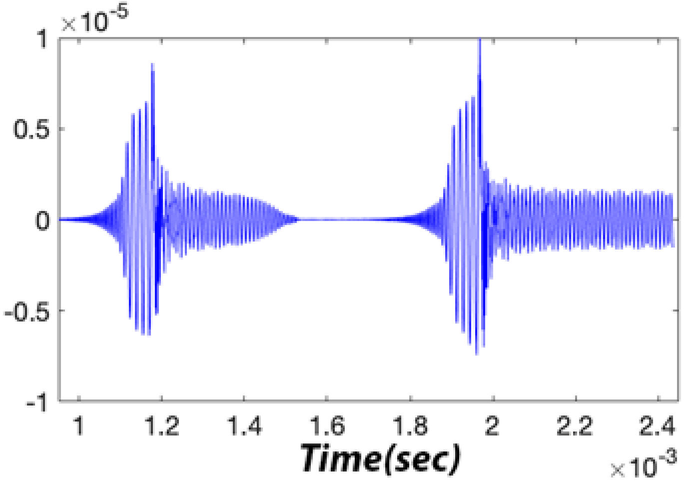
[Poster by Spong]

EP density profile evolves

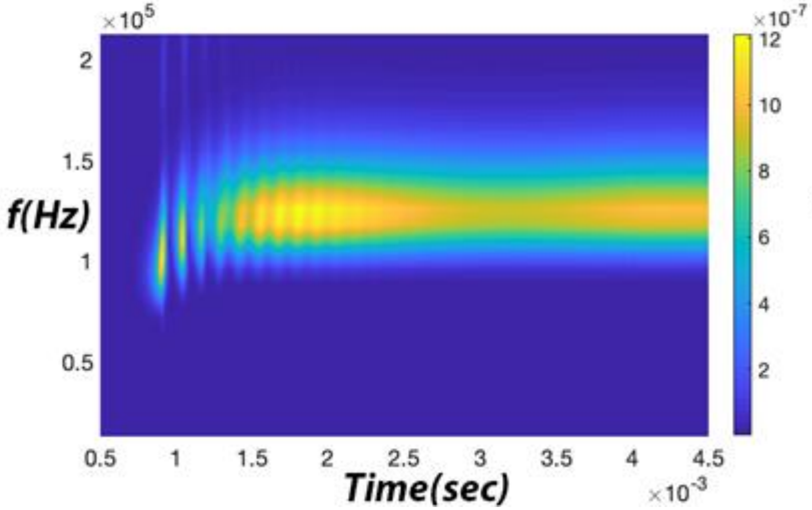
Edge magnetic fluctuations $\delta B_\theta / B$

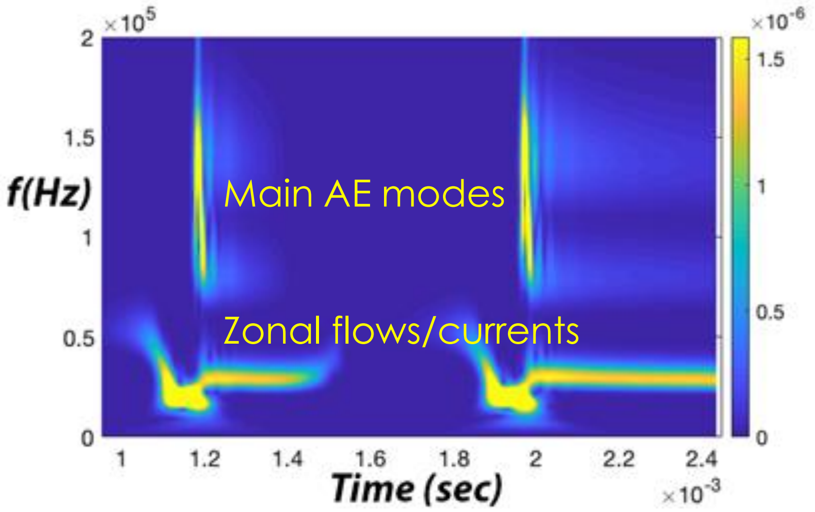


EP density profile fixed



Spectrograms



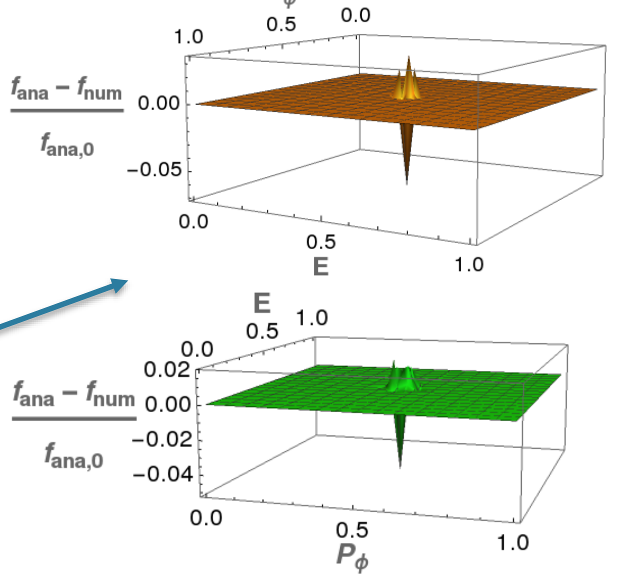


Development of EP modules for WDM: 2D generalization of RBQ

"Formulate RBQ2D approach for full resolution in velocity. Use NOVA-K interface for single modes. Implement slanted phase space diffusion. Apply RBQ1D for VV with velocity space resolution."

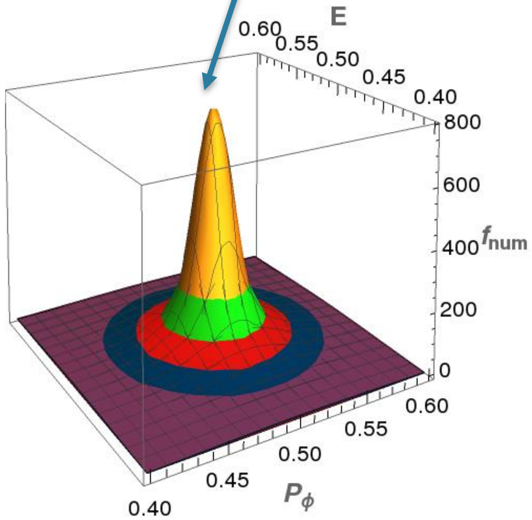
- slanted scheme being implemented via an improved iterative implicit method (McKee *et al*, J. Comp. Phys. 1996)
- Scheme has been verified against known analytical solutions in limiting cases: error accumulation and solution stability were studied.
- Multiple modes introduce a variety of diffusion paths
- Burning plasma modeling will be computationally expensive

Relative error does not accumulate

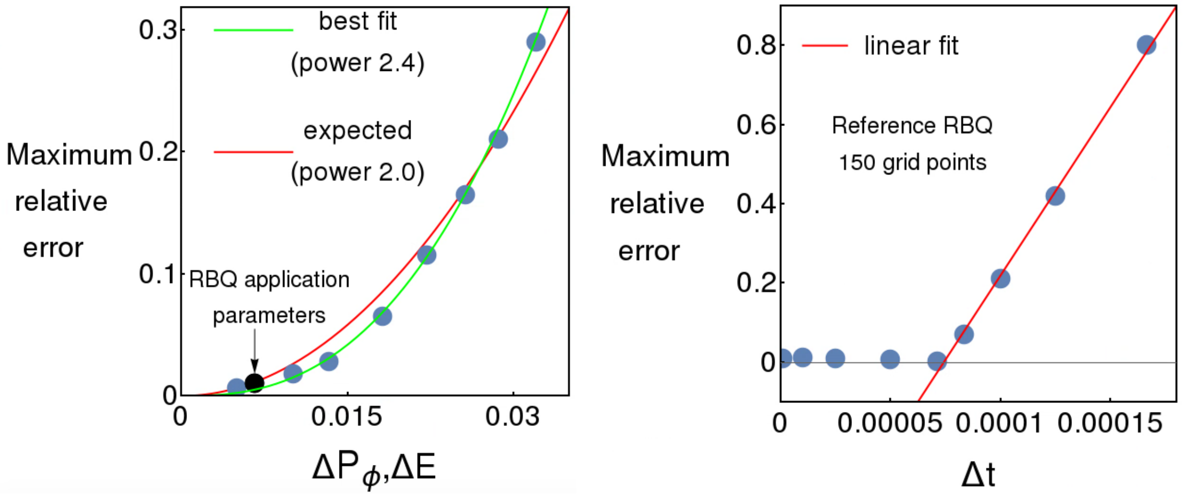


[*Poster by Gorelenkov*]

Numerical distribution function: unconditionally stable



Scaling laws for numerical error: grid spacing and time stepping dependence

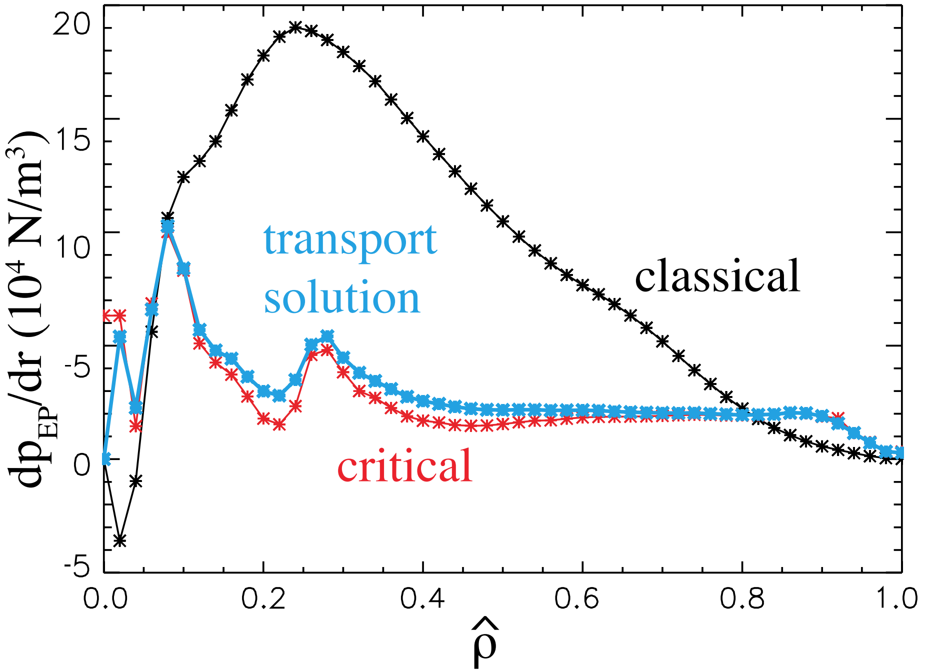


PPPL
PRINCETON PLASMA PHYSICS LABORATORY

Kick model has been extensively tested for NTMs, being extended to other low-f instabilities

- Successful application of kick model to DIII-D experiments with NTMs (from Y3) [Heidbrink NF 2018] [Bardoczi PPCF 2018] [Podestà NF 2019 (in press)]
  - Model reproduces measured EP transport from FIDA, NPA, neutrons based on experimental NTM parameters (island width, frequency, helicity)
- TRANSP + kick simulations being extended to DIII-D scenarios with NTMs + fishbones (from Y4) [Liu IAEA-TCM EP 2019]
- Kick model being updated to deal with broad class of instabilities
  - Fishbones, kinks [Podestà NF 2019 (in press)] [Cecconello IAEA-TCM EP 2019]
  - 3D fields: start with ripple then extend to Resonant Magnetic Perturbations (from Y4)
  - Microturbulence: enable direct comparison with TGLF-EP/ALPHA

TGLF-EP+Alpha is the simplest, fastest EP transport model available → extensive validation possible and necessary



Stiff transport forces the gradient to not (much) exceed a "critical gradient" of AE transport (essentially the linear stability threshold).

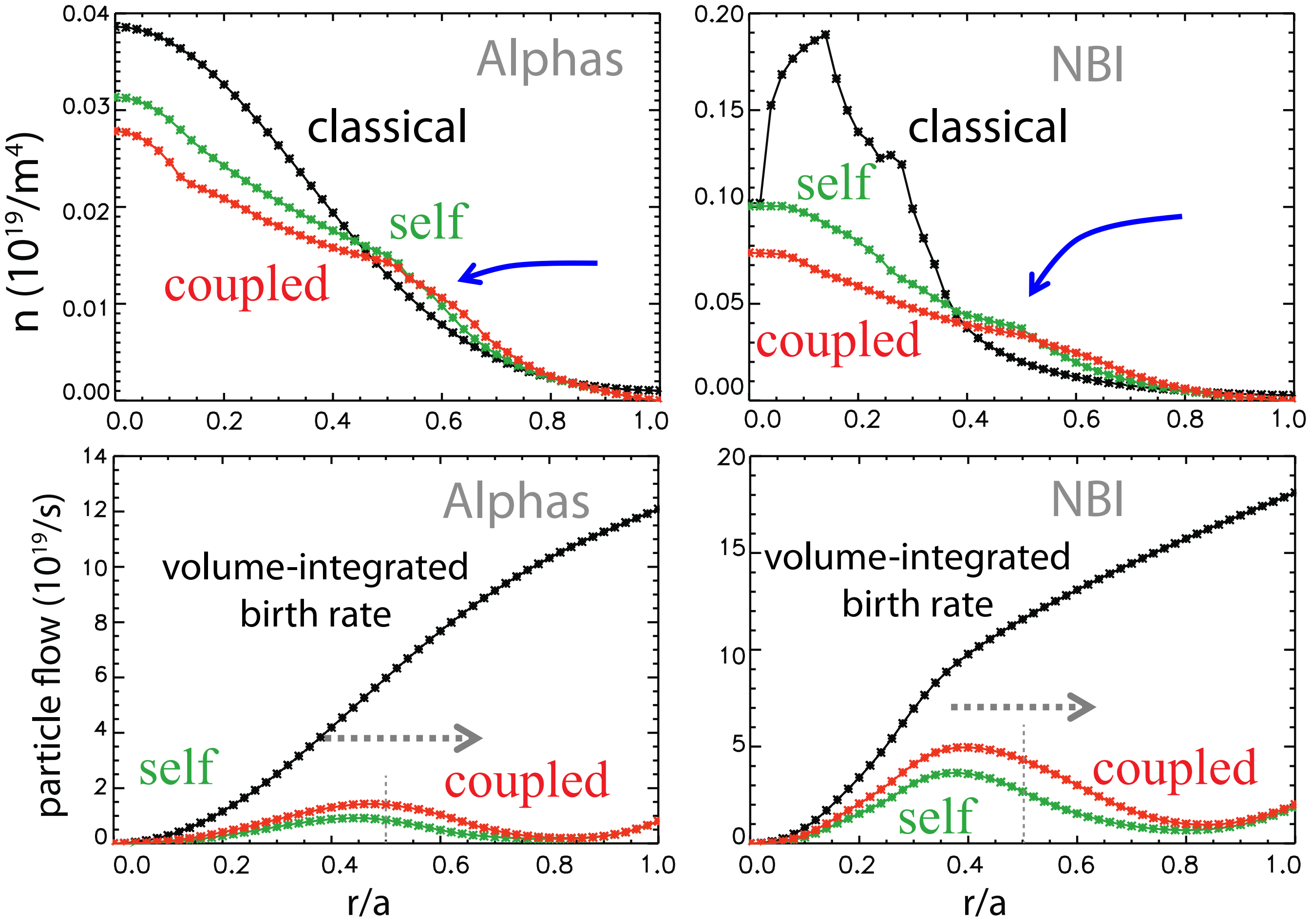
TGLF-EP+Alpha is a 1D critical-gradient model (CGM) using gyro-fluid stability calculations and a stiff AE-EP transport assumption.

Model features:

- Highly reduced → inexpensive
- Increasingly automated, minimal human judgment required
- **Fully physics-based!** No "fudge factors" or AE inputs from experiment.

Simplifying assumptions (Maxwellian EPs; stiff, local transport; no velocity-space dependence; etc.) make **validation** especially necessary to **map applicability**.

In ITER, coupled alpha and NBI drive nearly doubles confinement loss from mid core. Net edge loss is small !



Mid-core AEs redeposit EPs outward

self: Each EP species drives only its own transport

coupled: Simultaneous drive transports both species.

EPs redistributed from **inner core to outer core**

	alphas	NBI ions
self:	14.1%	23.1%
coupled:	23.5%	37.3%

Outside AE-unstable region (center and edge) flux comes from background transport component.

SciDAC ISEP: Integrated Simulation of Energetic Particles

- Develop integrated simulation of EP physics via 1st-principles ISEP framework
- Study high priority EP physics issues:
  - ✓ integrated simulation of EP coupling with thermal plasmas
  - ✓ long time scale simulation of EP confinement [*Poster by Spong*]
- Develop EP modules for WDM: TGLF-EP + Alpha, RBQ & Kick (+TRANSP) [*Poster by Gorelenkov*]
- Convergence of 1st-principles simulation with deep learning? [*Oral by Tang; Poster by Dong*]
- ISEP leads V&V of EP modules in world fusion program

Highlights of AI/Deep Learning FRNN Code

SciDAC-4 ISEP Project

William Tang
Princeton University/Princeton Plasma Physics Laboratory (PPPL)

SciDAC-4 PI Meeting
Rockville, MD

July 16-18, 2019

FRNN Project Team

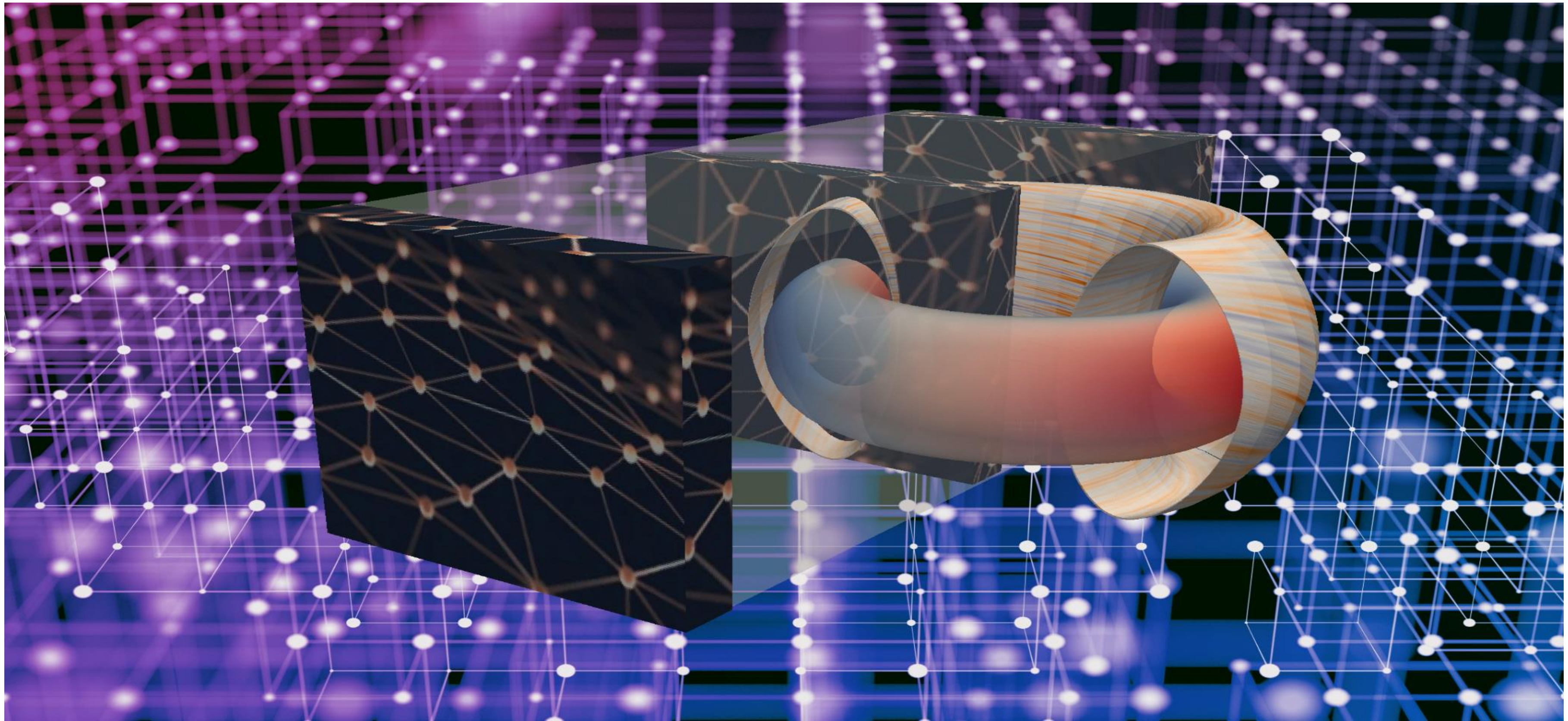
Julian Kates-Harbeck (Harvard U/PPPL), Alexey Svyatkovskiy (Microsoft/PPPL), Eliot Feibush (PPPL/Princeton U), Dan Boyer (PPPL),
Keith Erickson (PPPL), Ge Dong (PPPL), Kyle Felker (ANL/PPPL)

Artificial Intelligence/Deep Learning brings new technology to accelerate progress

"Predicting Disruptive Instabilities in Controlled Fusion Plasmas through Deep Learning"

NATURE: (accepted for publication, Jan. 2019, published, April 17, 2019 –
DOI: 10.1038/s41586-019-1116-4)

Princeton's Fusion Recurrent Neural Network code (FRNN) uses convolutional & recurrent neural network components to integrate both spatial and temporal information for predicting disruptions in tokamak plasmas with unprecedented accuracy and speed on top supercomputers



Machine Learning Workflow

Identify Signals
• Classifiers

Preprocessing and feature extraction

Normalization

Train model, Hyper parameter tuning

Use model for prediction

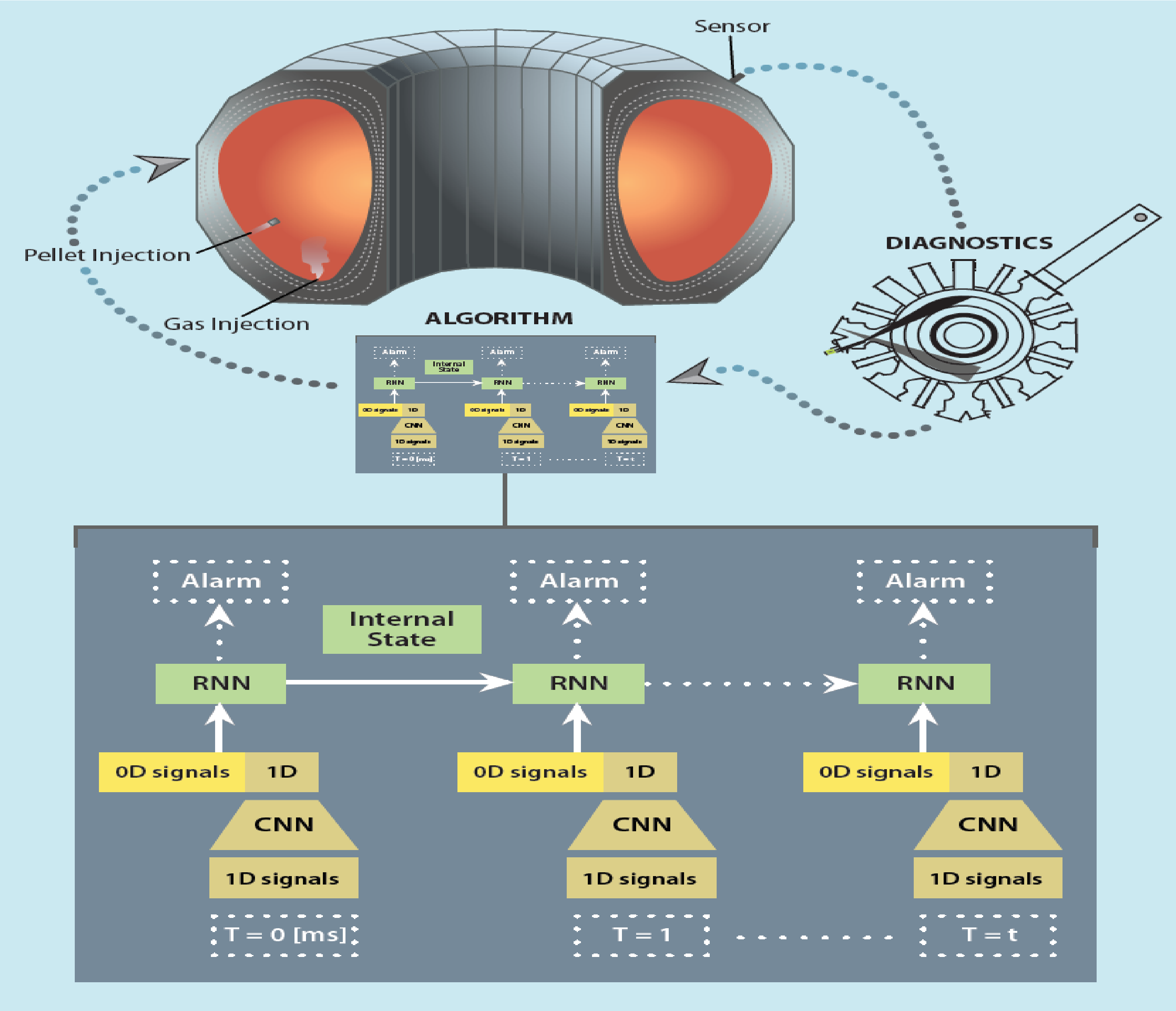
Princeton/PPPL DL predictions now advancing to *multi-D time trace signals* (*beyond zero-D*)

Measured sequential data arranged in patches of equal length for training

All data placed on appropriate numerical scale ~ O(1)
e.g., Data-based with all signals divided by their standard deviation

• All available data analyzed;
• Train LSTM (Long Short Term Memory Network) iteratively;
• Evaluate using ROC (Receiver Operating Characteristics) and cross-validation loss for every epoch (equivalent of entire data set for each iteration)

Apply ML/DL software on new data


Sensor
Pellet Injection
Gas Injection
DIAGNOSTICS
ALGORITHM
Alarm
Alarm
Alarm
Internal State
RNN
RNN
RNN
0D signals
1D
0D signals
1D
0D signals
1D
CNN
CNN
CNN
1D signals
1D signals
1D signals
T = 0 [ms]
T = 1
T = t

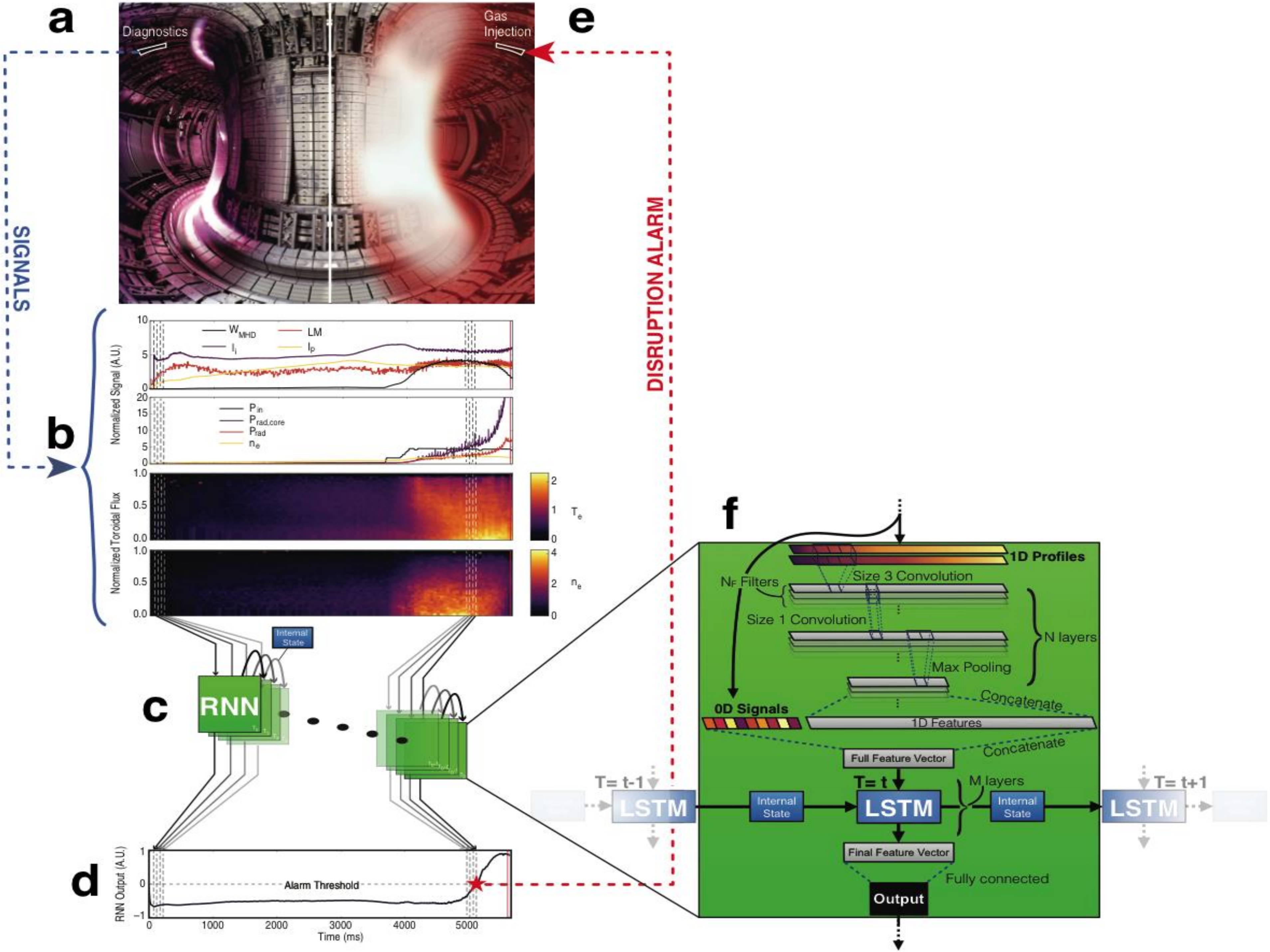

a
Diagnostics
Gas
Injection
e
SIGNALS
DISRUPTION ALARM
b
Normalized Signal (A.U.)
W_MHD
LM
l_i
I_p
P_in
P_rad,core
P_rad
n_e
Normalized Toroidal Flux
T_e
n_e
Internal State
c
RNN
d
RNN Output (A.U.)
Alarm Threshold
Time (ms)
f
1D Profiles
Size 3 Convolution
N_F Filters
Size 1 Convolution
N layers
Max Pooling
Concatenate
0D Signals
1D Features
Concatenate
Full Feature Vector
T= t-1
T= t
T= t+1
M layers
LSTM
Internal State
LSTM
Internal State
LSTM
Final Feature Vector
Fully connected
Output

Figure Caption: System overview and disruption-prediction workflow **(a—e)**

Top Image: interior view of the JET tokamak, with a non-disruptive plasma on the left and a disruptive plasma on the right.

Diagnostics **(a)** provide streams of sensory data **(b)** which are fed to the RNN-based deep learning algorithm **(c)** every 1 ms, producing a corresponding 'disruptivity' output at every time step **(d)**.

If output crosses a preset threshold value (dashed horizontal line), a disruption alarm is called (red star)
Alarm triggers mitigation action, such as gas injection **(e)** into the tokamak, to reduce the deleterious effects of the impending disruption **(f)**.

Detailed schematic of our deep-learning model:
Input data consist of scalar zero-dimensional (0D) signals and 1D profiles.

N layers of convolutional (containing NF filters each) and down-sampling (max-pooling) operations reduce dimensionality of the profile data and extract salient low-dimensional representations (features).

Features are concatenated with the 0D signals and fed into a multi-layer long-short term memory network (LSTM) with M layers, which also receives its internal state from the last time step as input.

The resulting final feature vector ideally contains salient information from the past temporal evolution and the present state of all signals.
This **vector** is fed through a fully connected layer to produce the output.

HIGHLIGHTS OF KEY ACHIEVEMENTS FEATURED IN NATURE PAPER (2019)

- ***Implementation of modern AI/Deep Learning advances enabled key achievements for Fusion Energy Science including:***

(1) Establishing ability to deal with one-dimensional physics signals for the first time – a significant improvement over previous Machine Learning R&D with focus on scalar-only “zero-D” signals.

(2) First demonstration of crucially-needed ability for predictive software trained on one experimental device (e.g., DIII-D tokamak) to make accurate predictions on another (e.g., the much larger, more powerful JET system) –> a key requirement for ITER relevance.

(3) Unique demonstration of AI/DL software capability to efficiently utilize leadership class supercomputers -- e.g., Titan, Summit in US; Tsubame-3 in Japan, etc. – and exciting powerful systems in near future such as AURORA-21 (US), ABCI (Japan), ...

Integration of HPC (using GTC Exascale Code) with Deep Learning Workflows (using FRNN DL Code)

- *"Knowledge & experience" now in place for carrying out path-to-exascale HPC simulations of ITER-relevant burning plasmas with powerfu*
- *Electromagentic GTC code →* ***ESP selection for SUMMIT and 2019 INCITE awardee of 740K SUMMIT Node Hours – 151% above our request !***
- *Neoclassical tearing modes (NTM's) already experimentally observed in JET, but NO realistic models yet developed as improved pre-disruption classifiers in Machine Learning workflows → because of inability to include measured higher-D profiles (only scalars)*

- ***CNN & RNN allow including realistic 1D & higher-D measurements of profiles to enable first-principles-based reduced models of NTM's (supported by exascale GTC code) to be used in FRNN workflows***

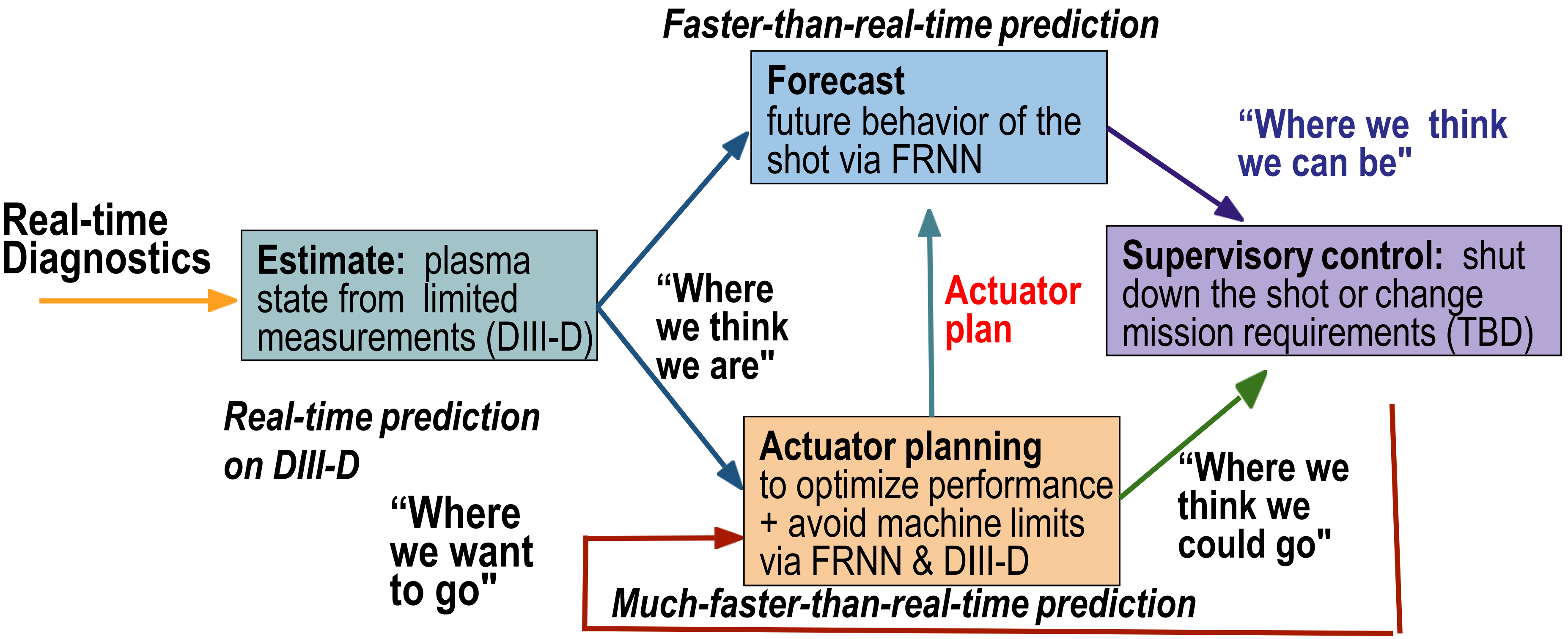
- *Very encouraging recent progress – see Poster by Ge Dong:*

→ FRNN Sensitivity Study: Connection between NTM & Disruption Prediction in DIII-D

Good Example of "integration of HPC with DL" !

Vision for Control Capabilities to Enable Real-Time Experimental Planning

Dan Boyer, Keith Erickson, *… plus experimental/advanced diagnostic expertise*



- **Can we make our models fast & accurate enough?**

--- e.g., via reinforcement learning/inference/ ……

- **Can we make our models realistic enough?**

--- e.g., via focused actuator planning with experimental partners

Initial plans for moving from AI Predictions to Control:

Year 1:

• Design/train/test predictive plasma models for real-time control (e.g., beam Neural network), including gaining experience on modern real-time computer;

• Modify AI/DL FRNN predictor for real-time determination of disruption probability and sensitivity to controllable plasma parameters;

Year 2:

• Initial development of control strategy with variety of decision-making algorithms; & initial deployment and testing of FRNN in real-time hardware.

Year 3:

• Delivery of initial results from an operating integrated tokamak control system equipped with modern Real-time computer hardware needed for the FRNN AI/DL predictor trained on large existing databases.

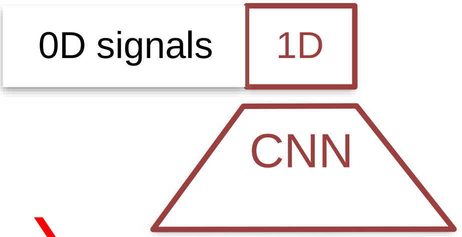
Long Term Impact on Future Devices/ITER: *Successful Development of*

• *Advanced control strategies based on unique AI/DL Predictors needed for optimization of performance and avoidance of disruptions in tokamaks.*

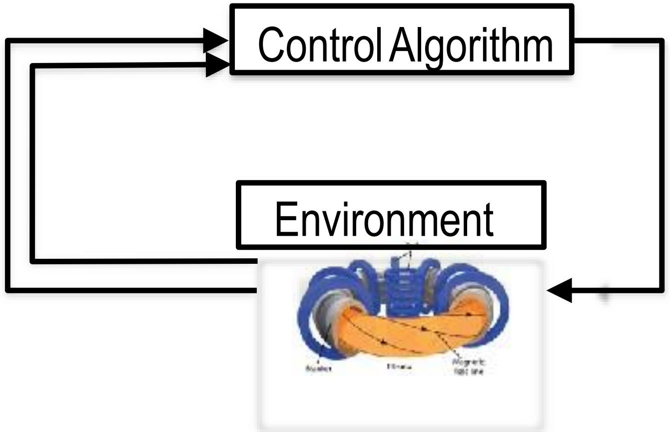
• *Vetting of stable, scalable, portable control systems and associated methodology methodology on existing tokamaks (e.g., DIII-D, KSTAR, …)*

DL/AI Vision Summary in Moving from Prediction to Control

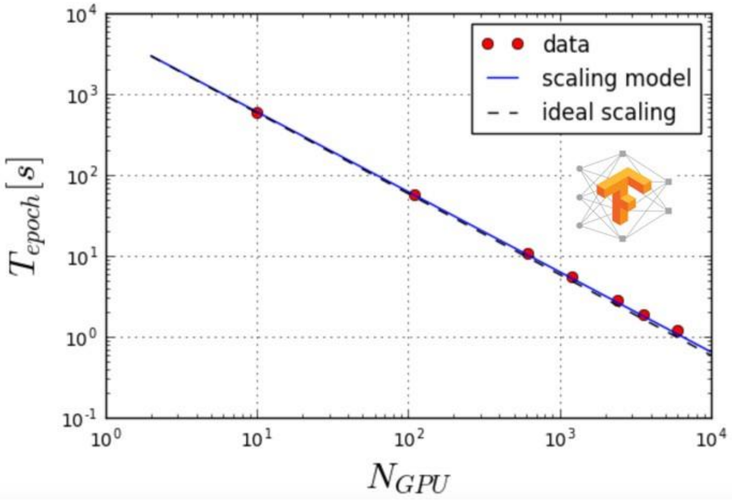
ZERO-D to HIGHER-D SIGNALS via CONVOLUTIONAL NEURAL NETS (CNN)



- *Enables immediate learning of generalizable features (→ helps enable **cross-tokamak portability of DL/AI software**)*



- ***Reinforcement learning enables transition from PREDICTION to CONTROL !***



- *Takes advantage of increasingly powerful **world class HPC (supercomputing) facilities !***