EPFL

Reinforcement Learning for Autonomous Accelerators (RL4AA) Collaboration

Plasma integrated control and trajectory optimization via reinforcement learning: applications in magnetic confinement fusion

Alessandro Pau

Swiss Plasma Center

alessandro.pau@epfl.ch

RL4AA

Swiss Plasma Center

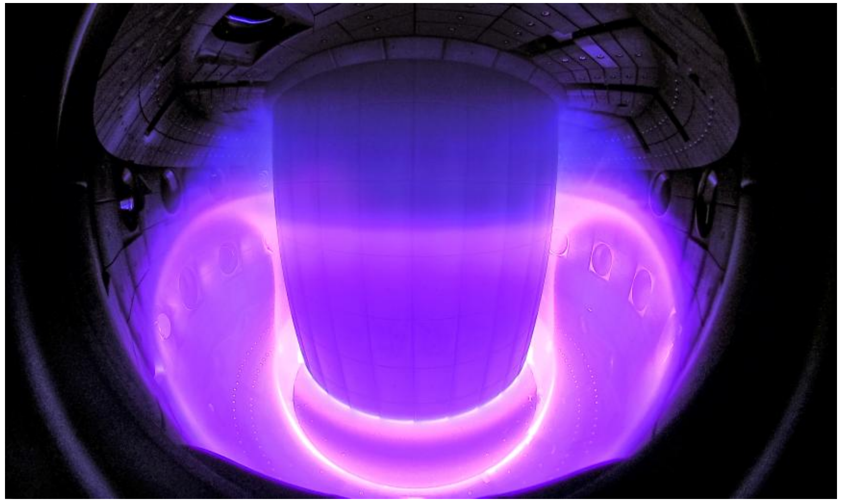
03/04/2025

Swiss Plasma Center - EPFL

Swiss Plasma Center - EPFL

Swiss
Plasma
Center

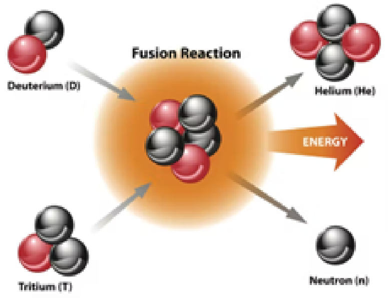
Tokamak a Configuration Variable (TCV)



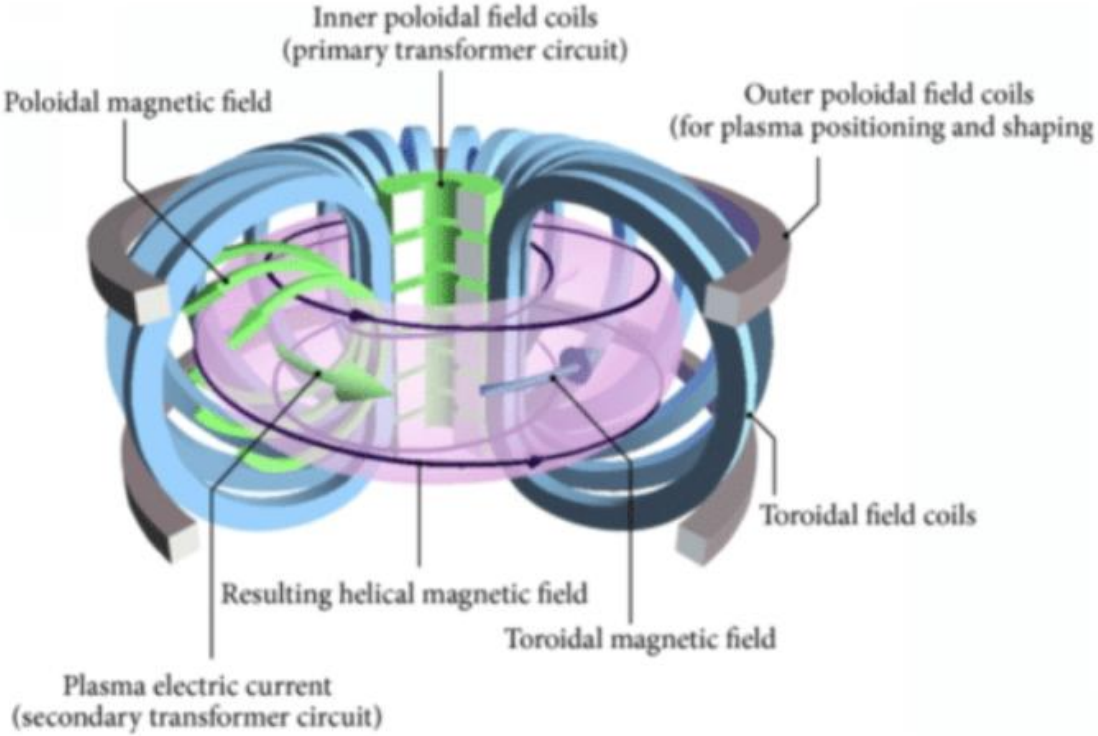
Swiss Plasma Center

Nuclear Fusion to produce electricity

- **Magnetic confinement fusion in brief**
  - For thermonuclear fusion one needs a high product of temperature, density, and confinement. Typically, T>100MK



  - Matter is in the ***plasma state***
    - ions and electrons are dissociated.
  - Magnetic fusion: keep the plasma confined by **magnetic fields**

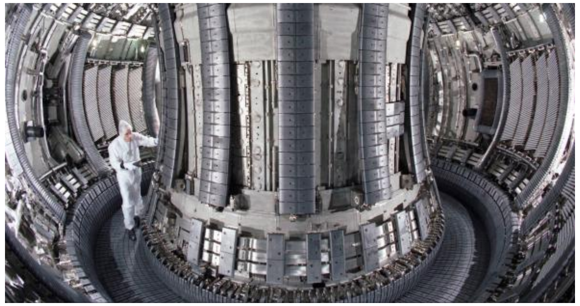


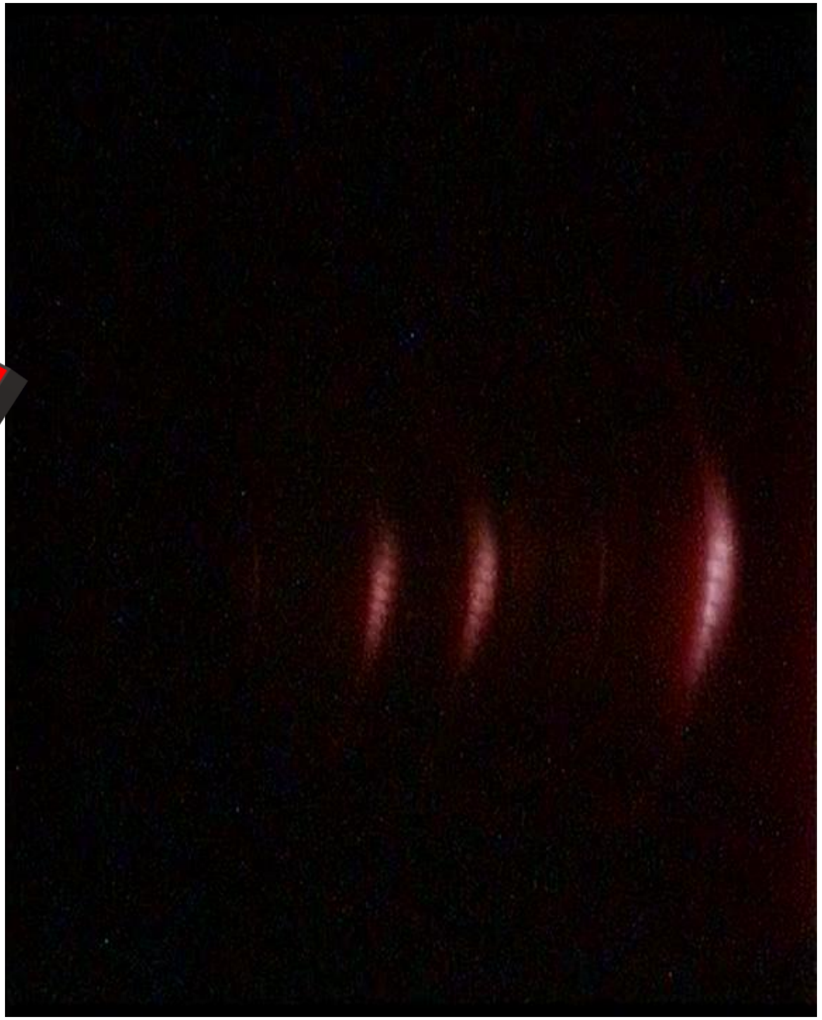
- **Tokamaks**
  - Axisymmetric configuration: **toroidal** (donut) shape with nested flux surfaces.
  - plasma confined by superposition of 2 magnetic fields:
    - **Toroidal** field generated by external coils
    - **Poloidal** field generated by current in plasma

Why Tokamaks? Record of produced energy at JET

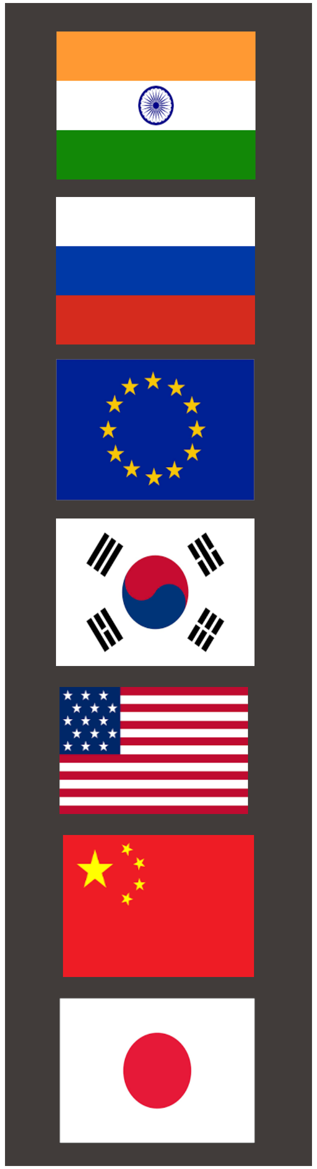
JET, EU, Dec 21st, 2021

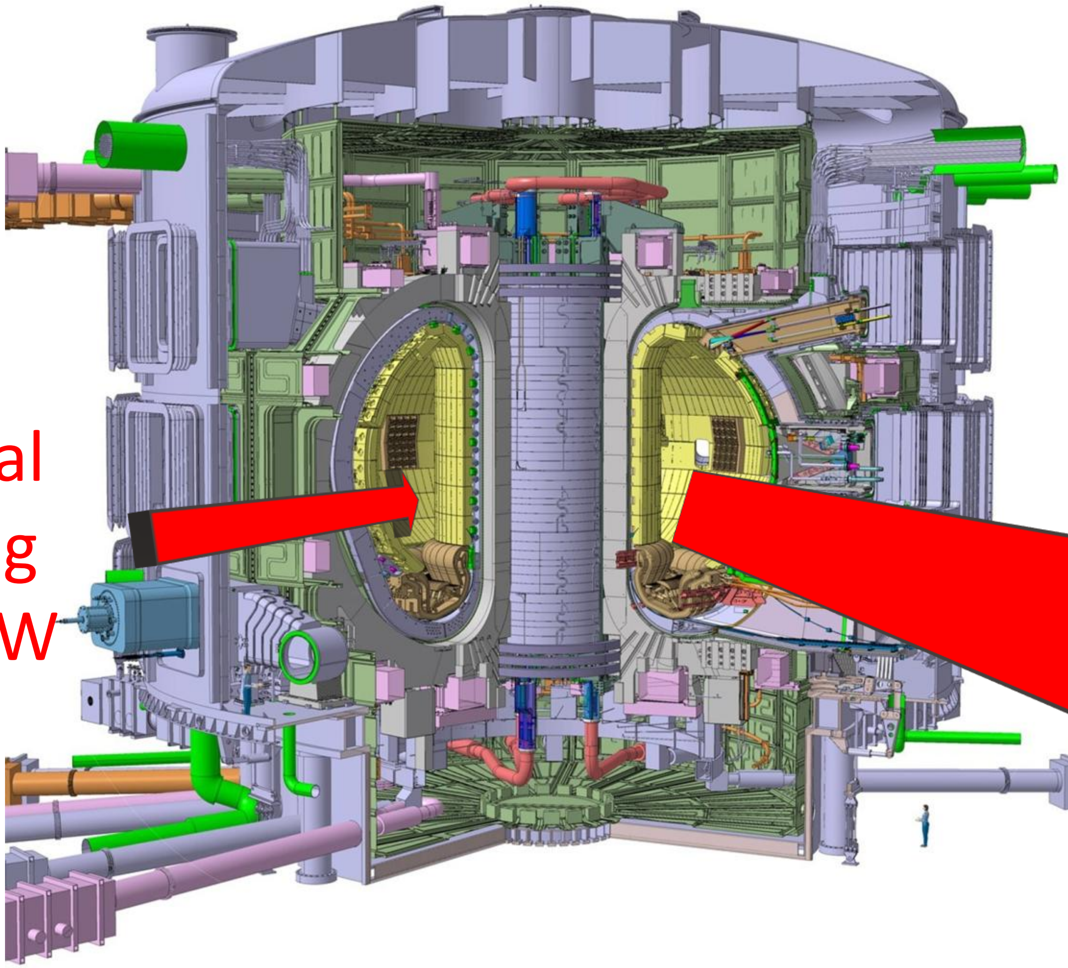
59 MJ for 5s!



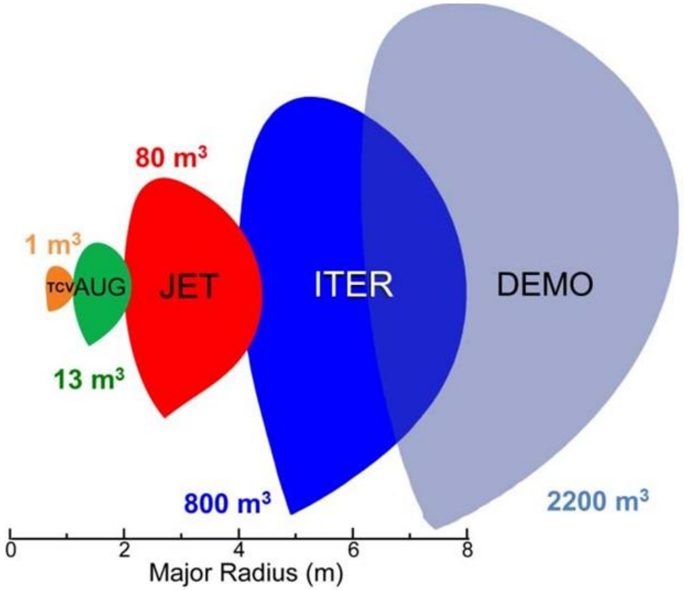


ITER, feasibility of fusion energy





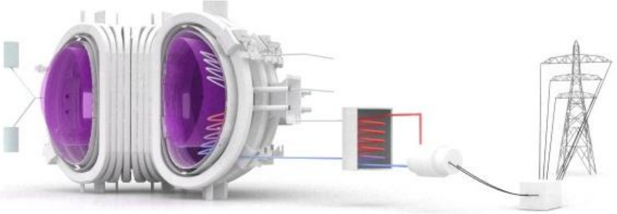
External Heating ~50 MW



~500 MW

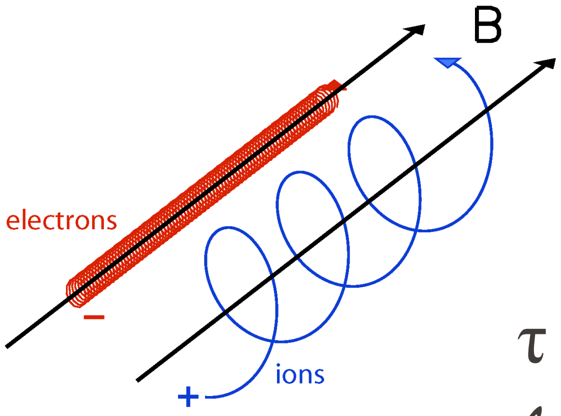
~1g of fuel in 840m³!

The step before DEMO
(demonstration fusion reactor)



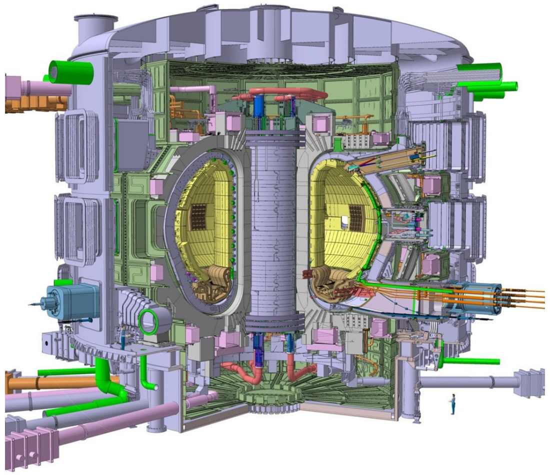
Simulations and control in fusion: a fascinating challenge

- Multiscale
- Multiphysics
- Intrinsic nonlinearity
- Turbulent dynamics
- Extreme anisotropy
- Complex geometry



$\tau \sim 10^{-10}$ s

$\ell \sim 10^{-4}$ m

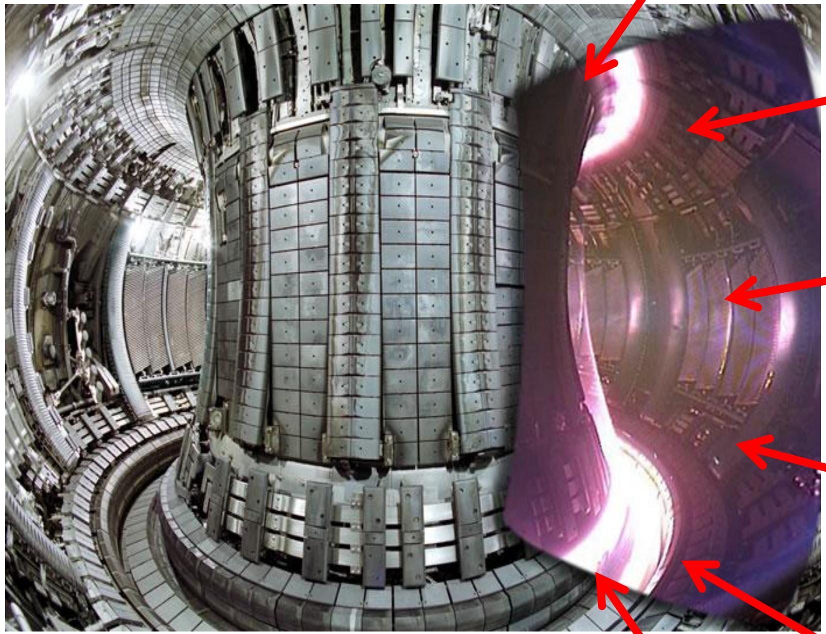


$\tau \sim 10^{3}$ s

$\ell \sim 1$ m

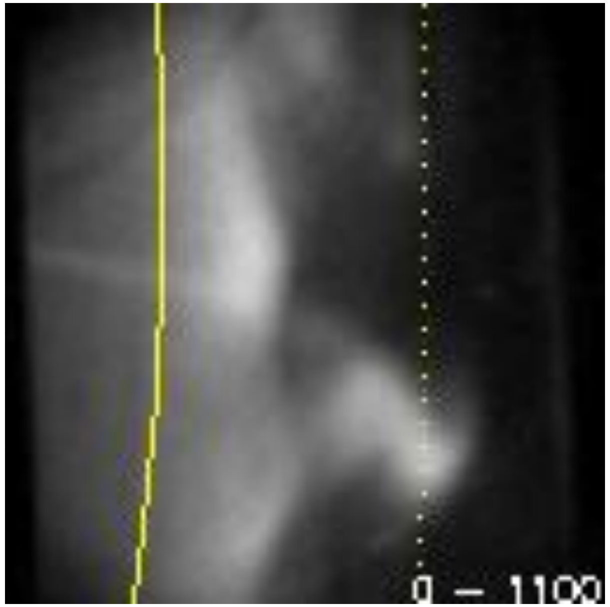
Simulations and control in fusion: a fascinating challenge

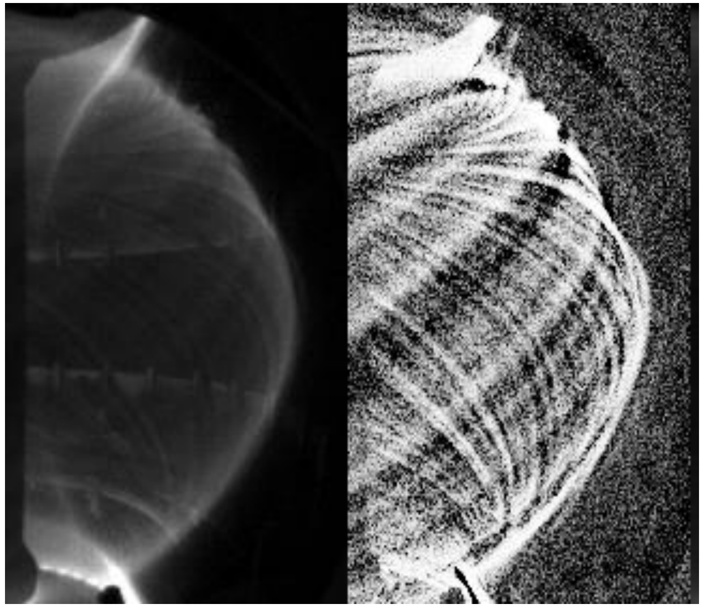
- Multiscale
- Multiphysics
- Intrinsic nonlinearity
- Turbulent dynamics
- Extreme anisotropy
- Complex geometry



Simulations and control in fusion: a fascinating challenge

- Multiscale
- Multiphysics
- Intrinsic nonlinearity
- Turbulent dynamics
- Extreme anisotropy
- Complex geometry

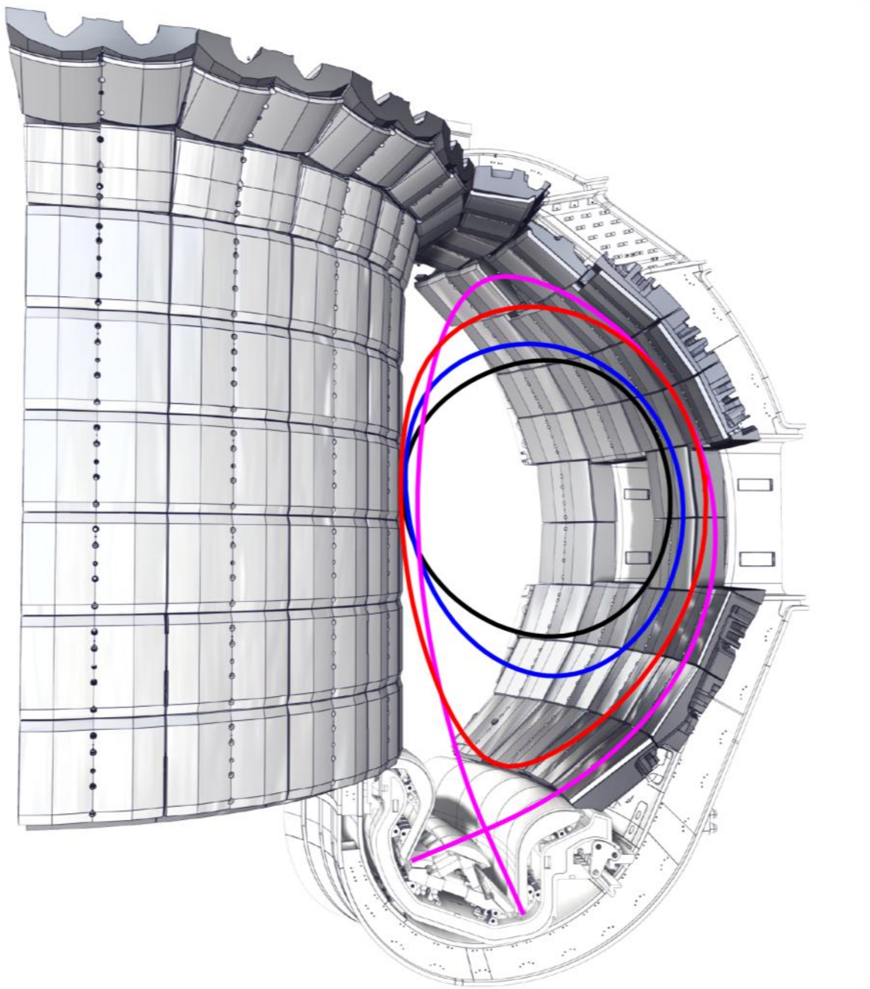




Swiss Plasma Center

Simulations and control in fusion: a fascinating challenge

- Multiscale
- Multiphysics
- Intrinsic nonlinearity
- Turbulent dynamics
- Extreme anisotropy
- Complex geometry

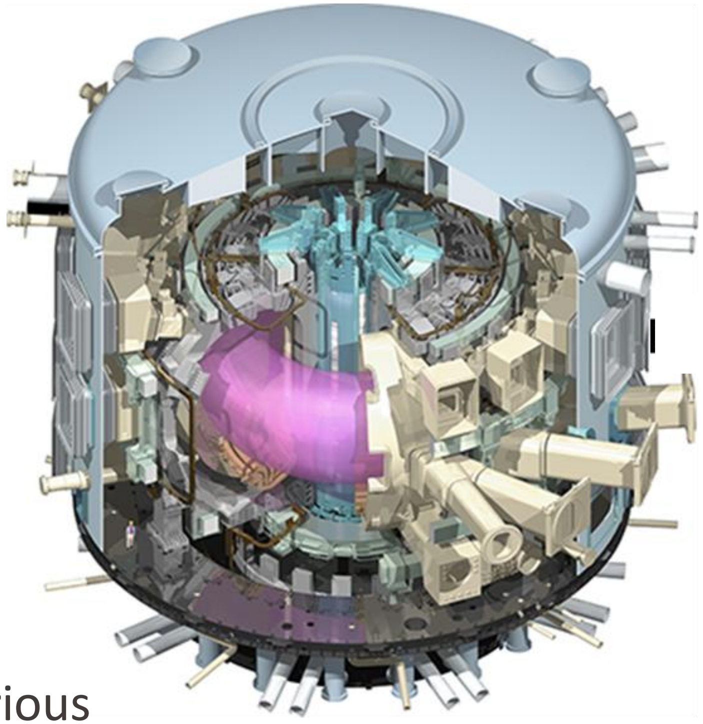


Swiss Plasma Center

Data in fusion: another challenge

- **Tokamaks** are one of the most complex system in nature…

- **Massive amount of data** (Big data – 2PB/day at ITER, high bandwidth diagnostics)



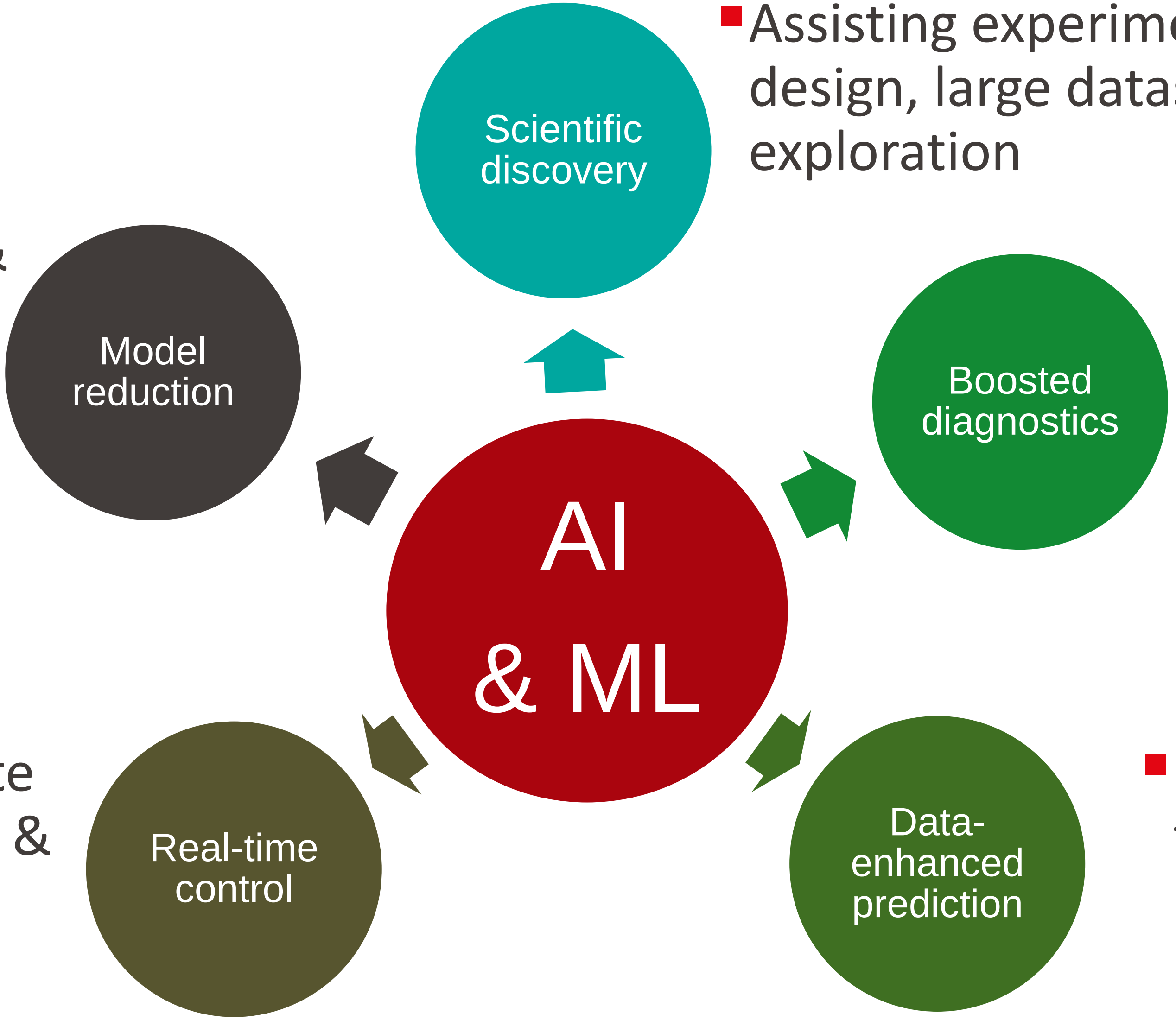
- **High-dimensionality** (many diagnostics measuring various plasma properties)

- **Heterogenous** (various formats, facility dependent data ecosystem)

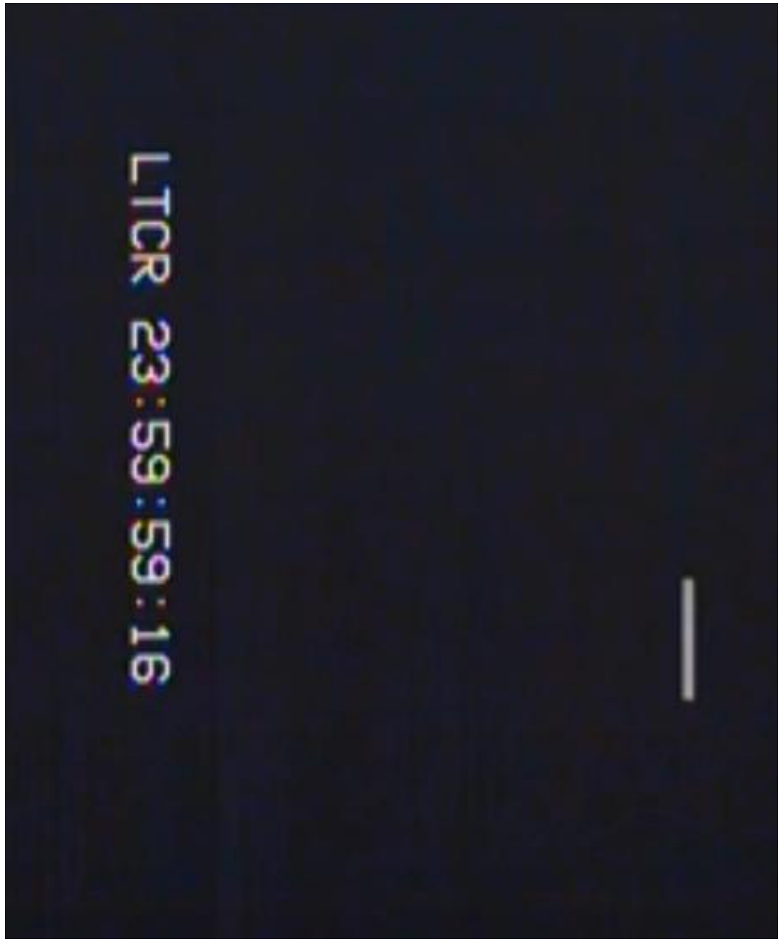
- **Multi-scales, multi-physics** (integrated tokamak modeling)

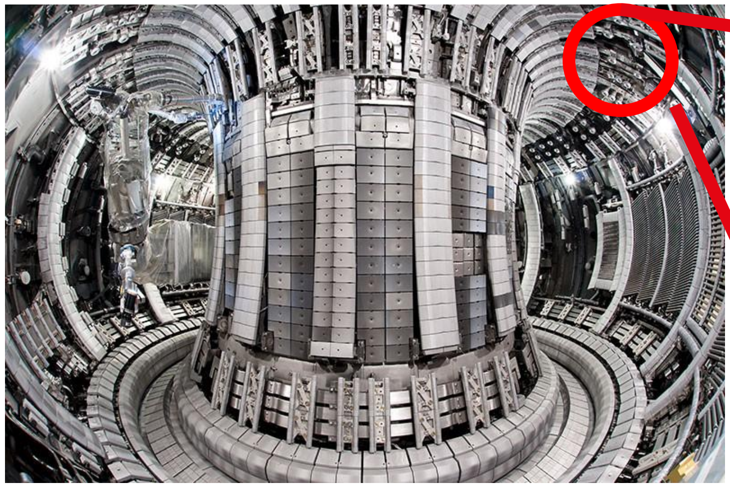
Advancing fusion with ML & AI

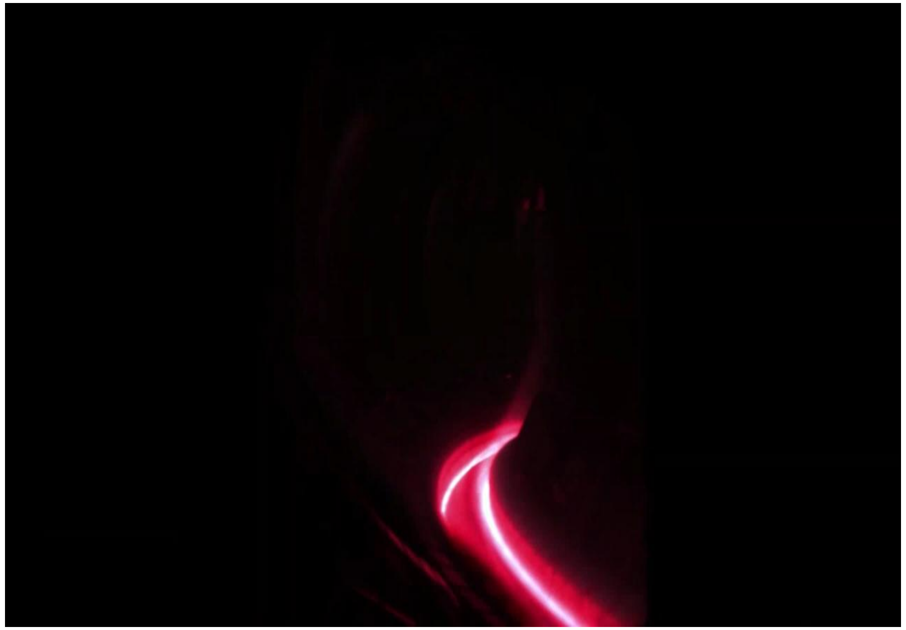
- Reduced & Surrogate models
- Assisting experimental design, large dataset exploration
- Information extraction, data fusion
- Plasma state monitoring & control
- Disruptions, transient events



Thermal loads & catastrophic losses of confinement







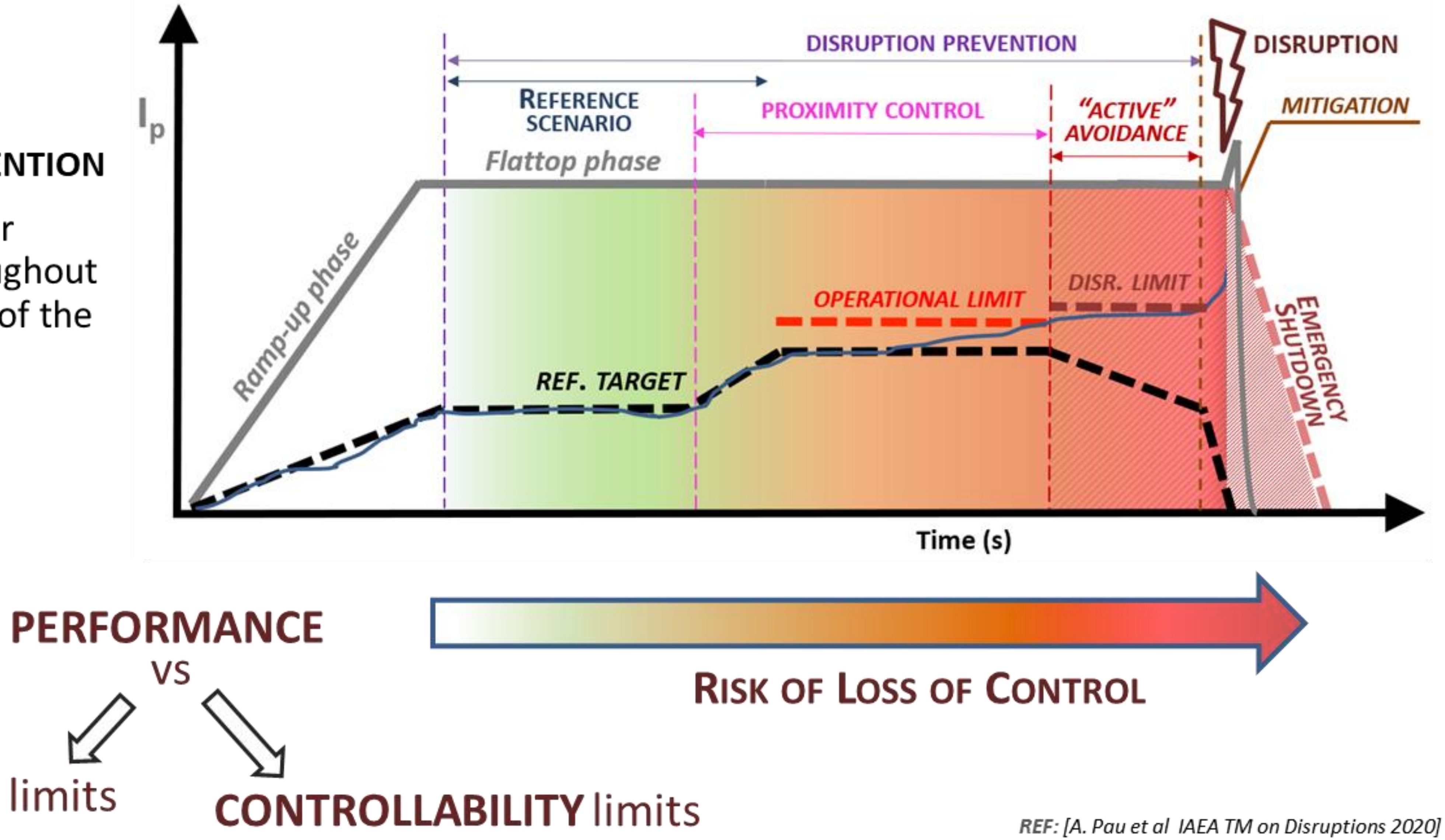
Possible serious damage to ITER structure

Continuous control and disruption prevention

EPFL

- **DISRUPTION PREVENTION**

a key requirement for plasma control throughout the <u>entire evolution</u> of the discharge

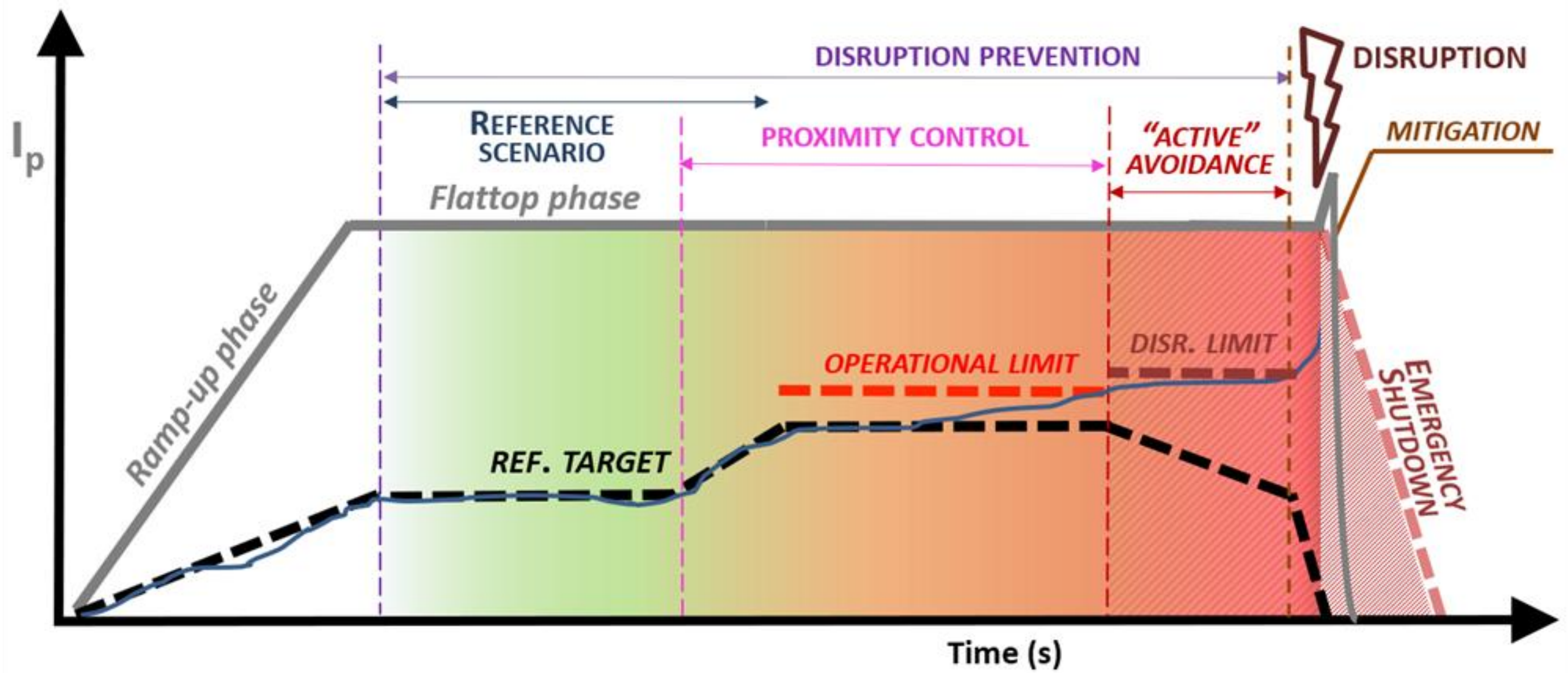


REF: [A. Pau et al IAEA TM on Disruptions 2020]
REF: [J. Barr et al IAEA TM on Disruptions 2020]

Continuous control and disruption prevention

- **DISRUPTION PREVENTION**

Break down in different “control phases”:



REF. SCENARIO

- keep the target scenario stable again disturbances (ST, ELM, MHD modes, etc.)

PROXIMITY CONTROL

- keep stability while pushing performance by regulating proximity to stability & controllability boundaries

ACTIVE AVOIDANCE

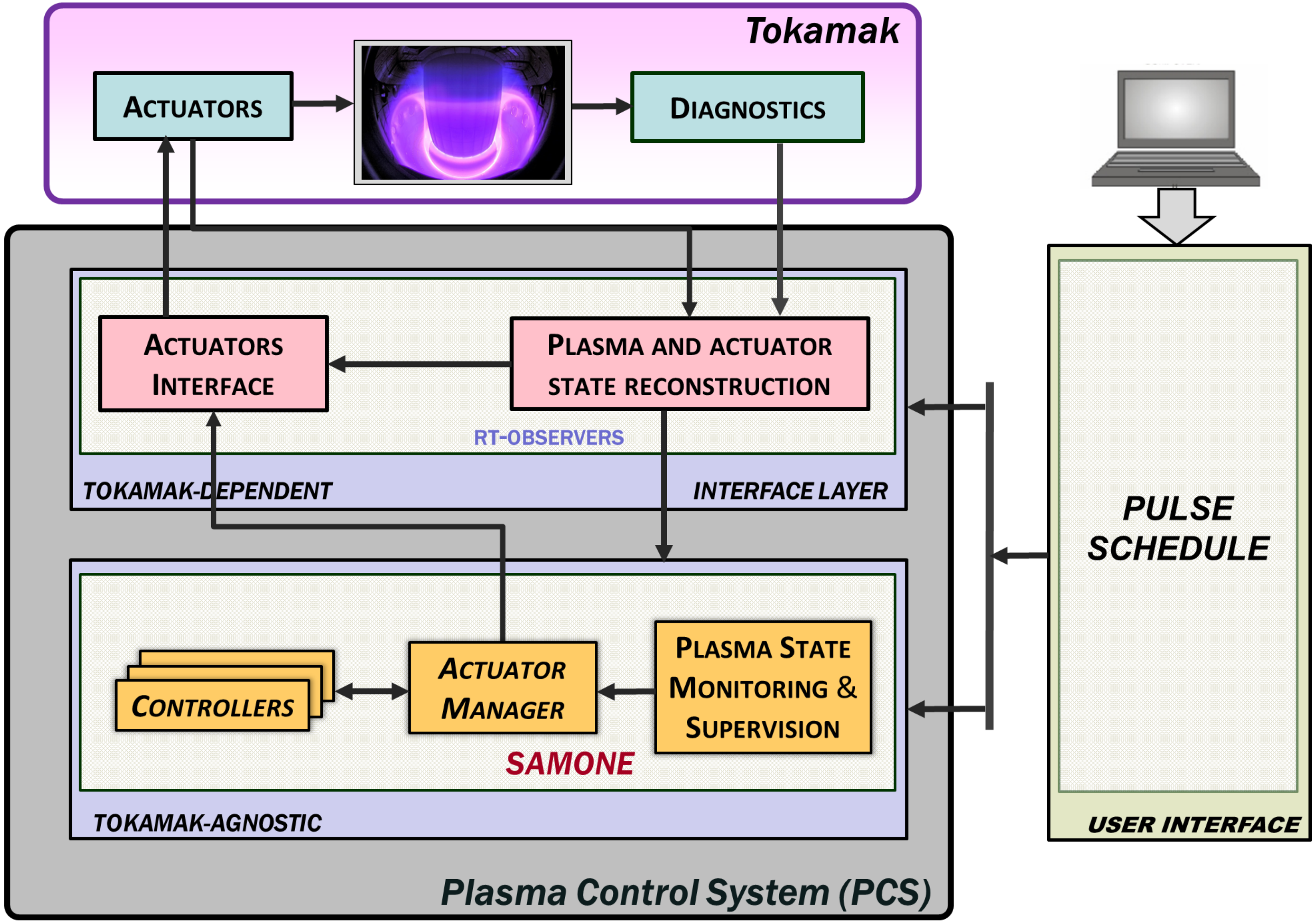
- asynchronous response when crossing operational boundaries (danger levels)

EMERGENCY SHUTDOWN

- Fast controlled shutdown
- mitigation

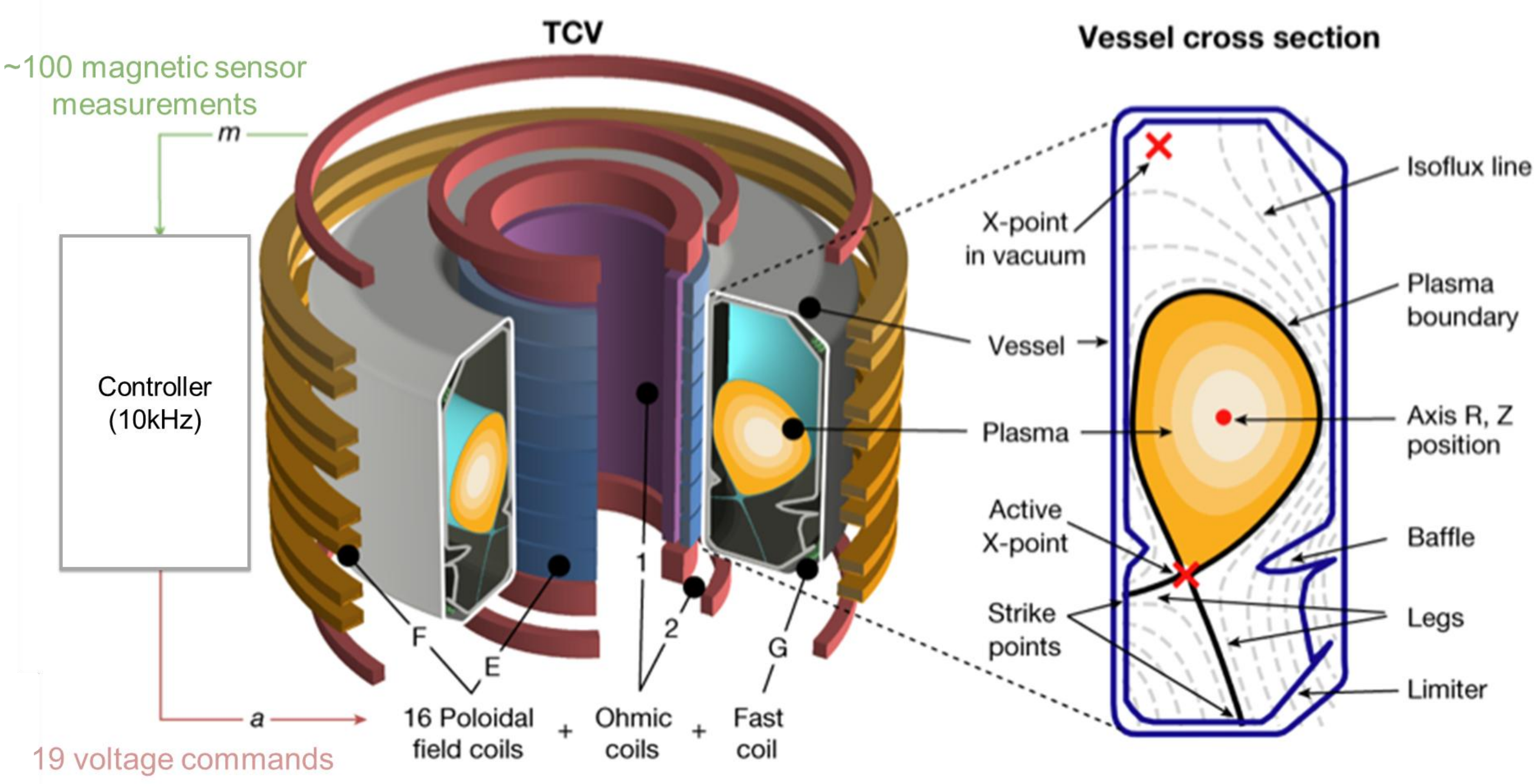
The TCV Plasma Control System (PCS)

- Separation of tokamak **dependent** and **agnostic** layers
- **Generic** implementation
- **Flexible** framework allowing easy **maintenance** and **upgrade**
- Concepts of **integration** and **portability**



REF: [T. Vu et al. IEEE TNS 2021]
REF: [F. Felici et al. IAEA 2021]
REF: [C. Galperti et al IAEA TM on Plasma Control Systems 2021]

Plasma control by deep reinforcement learning

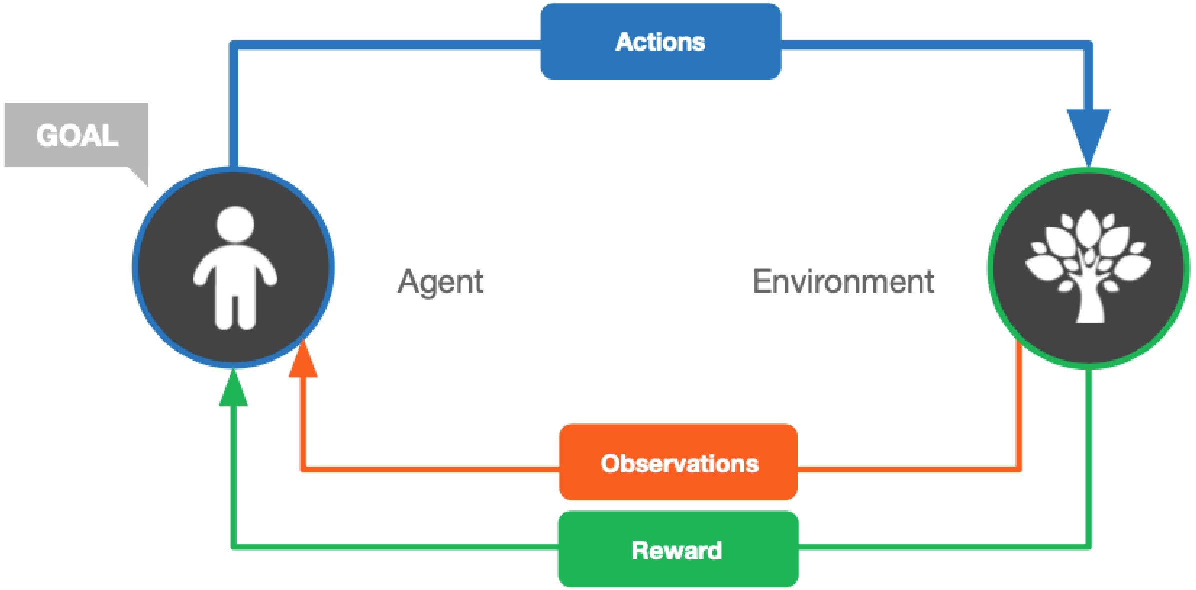


Swiss Plasma Center

DeepMind

REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

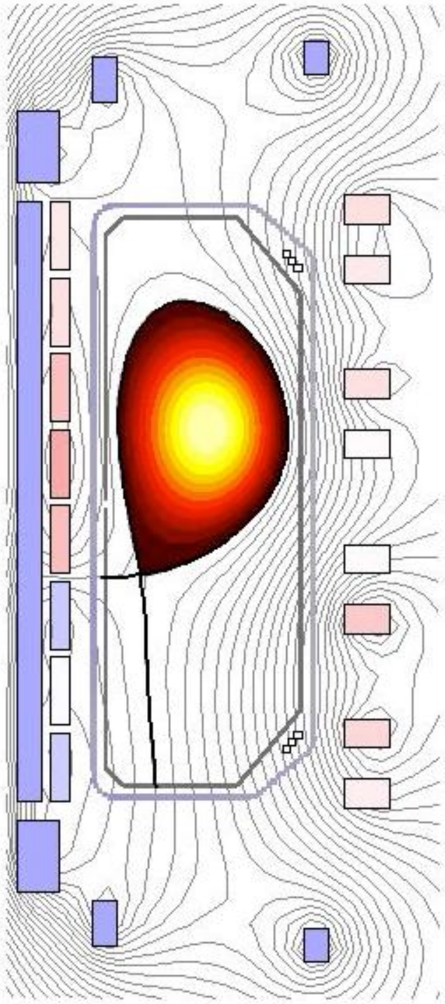
Plasma control by deep reinforcement learning



[Figure and RL slide material from hereon: courtesy A. Abdolmaleki]

- Agent tries to learn a **control policy** mapping states (observations) to actions
- **MPO: Actor-Critic RL** implementation for continuous-valued states
- **Deep RL**: recurrent architecture to efficient model of long-range dependencies (plasma dynamics)

- **TCV Free-boundary simulator**



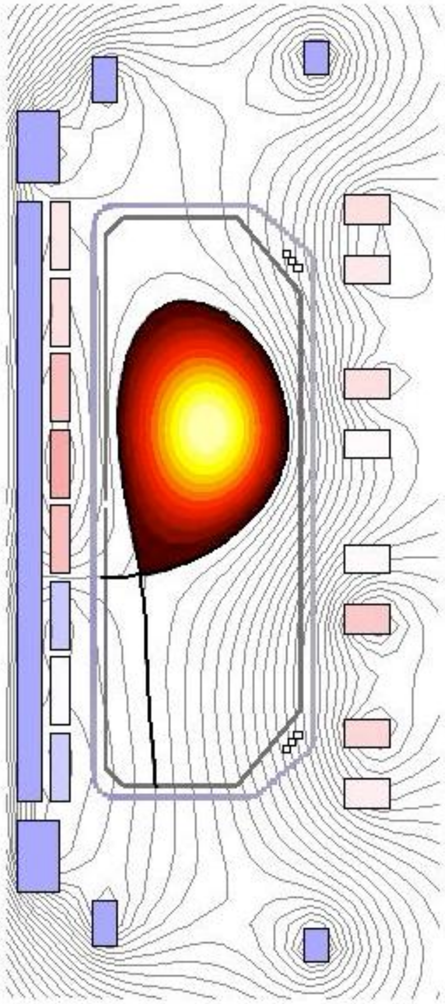
REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

Swiss Plasma Center

Plasma control by deep reinforcement learning

- **Free-boundary simulator** for the Tokamak magnetic equilibrium
  - Solve coupled system of equations:
    - **Circuit equations** for time evolution of currents, coupled equations for *Coils*, *Passive conductors* (vessel), *Plasma*
    - **Grad-Shafranov equation** (ideal MHD force balance: 2D static, elliptic PDE) to determine plasma equilibrium depending on external currents and internal constraints (provided)
  - Typically, **~hours for simulating** a few seconds discharge (50,000 steps / s) of plasma evolution.

- **TCV Free-boundary simulator**

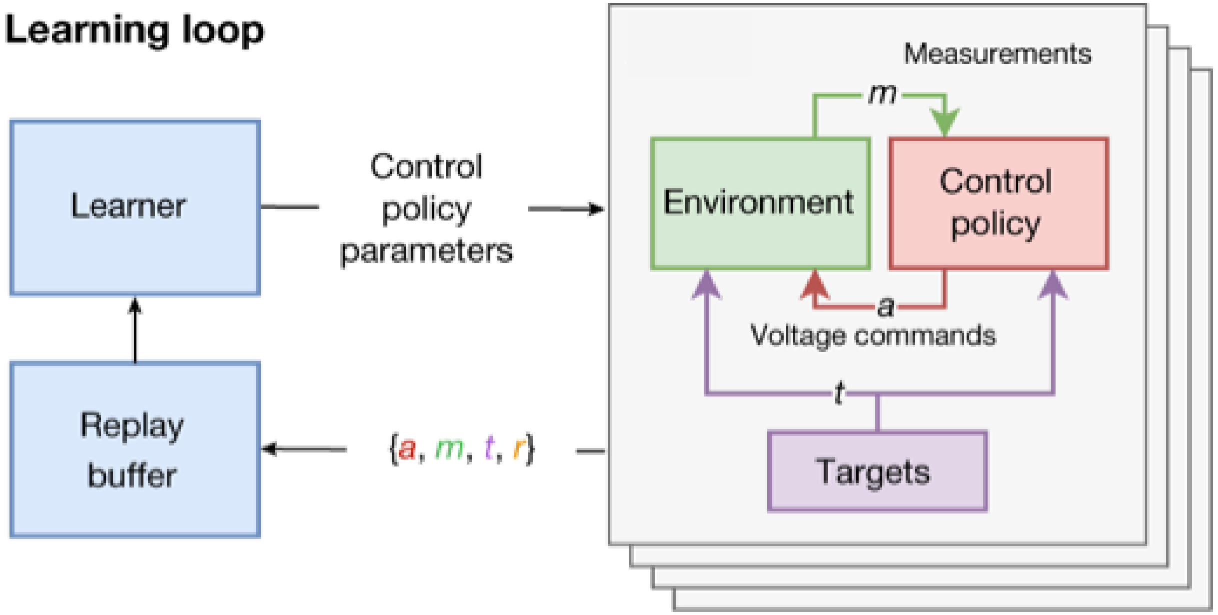


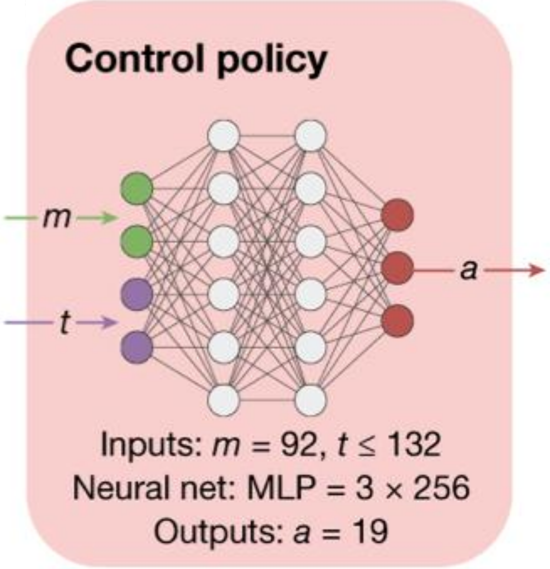
- **What we need to control:**
  - Total plasma current I_p ; Radial position **R**
  - Vertical position Z; Plasma shape (**LCFS,** etc.)

REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

Plasma control by deep reinforcement learning

- **MPO: Actor-Critic RL** implementation for continuous-valued states
  - Stable (constrains deviation from the current policy)
  - efficiently explores **continuous state-action spaces**
  - Efficient data sampling
  - Robust against **uncertainties**
- **Distributed implementation:** many actors run in parallel…
- **Off-policy**: also use data generated using previous policies



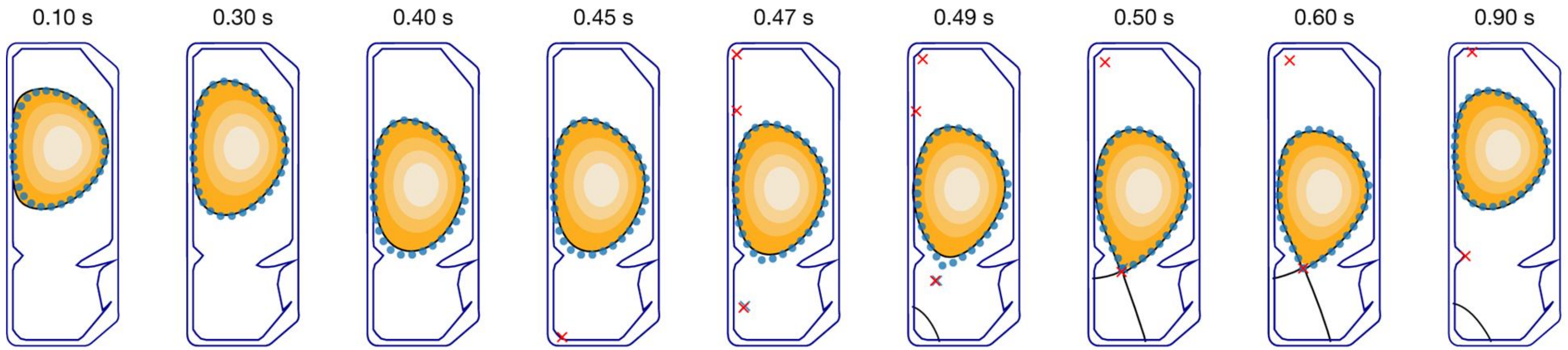


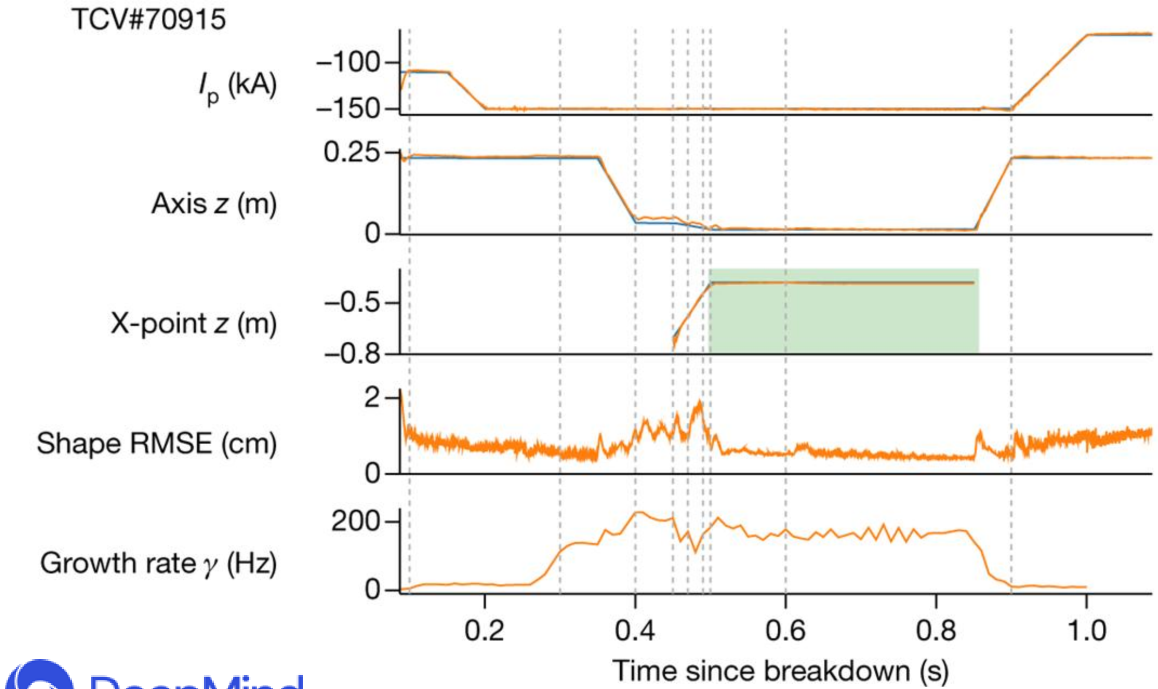
Actor: learns the control **policy** π (taking policy gradients on the *Q* function learned by the critic) -> **"small" feedforward NN**

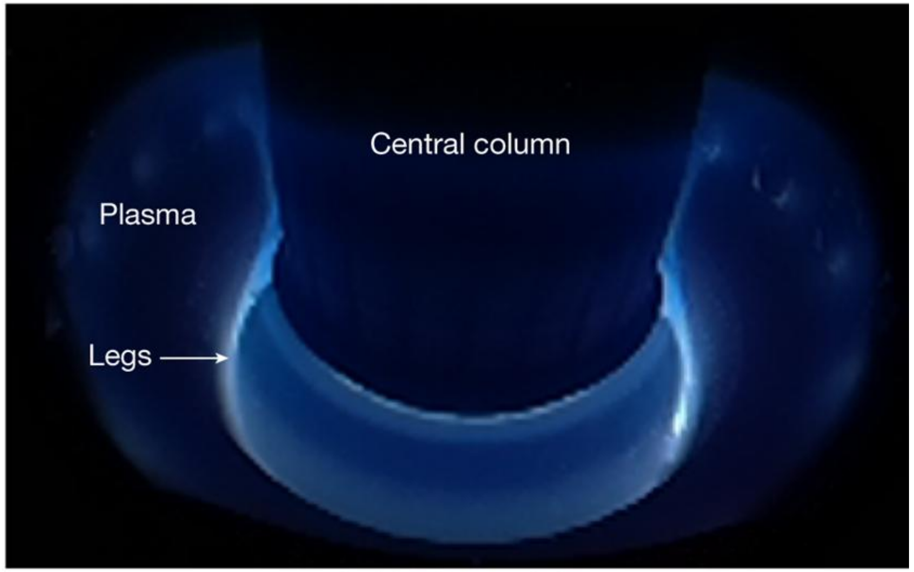
Critic: learns the ***Q* function** from the data generated by the interaction of the actor(s) with the environment (estimates expected cumulative future rewards) -> **deep recurrent NN**

REF: **J. Degrave, F. Felici** et al., *Nature*, 2022

Plasma control by deep reinforcement learning



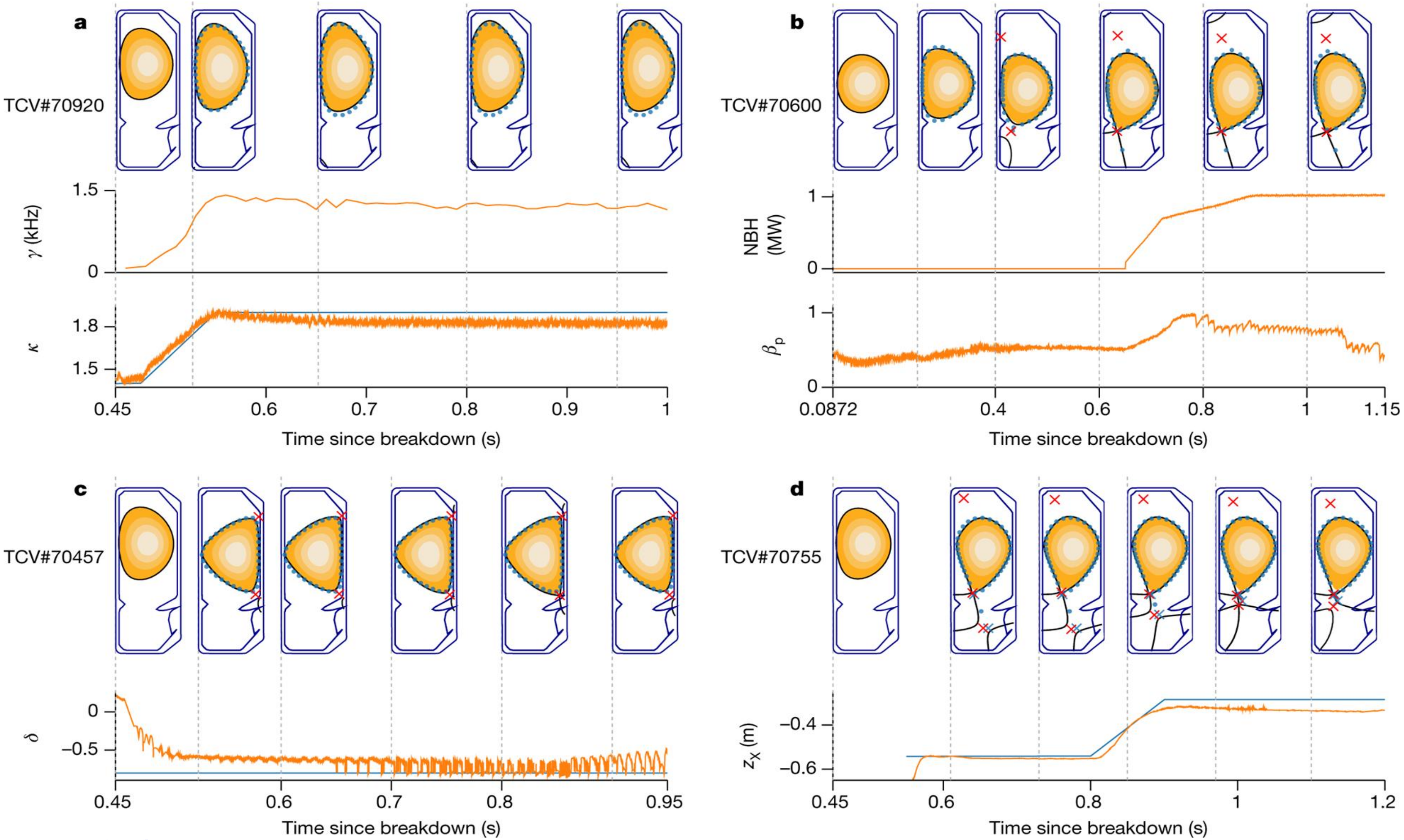




Inside view at 0.6 s

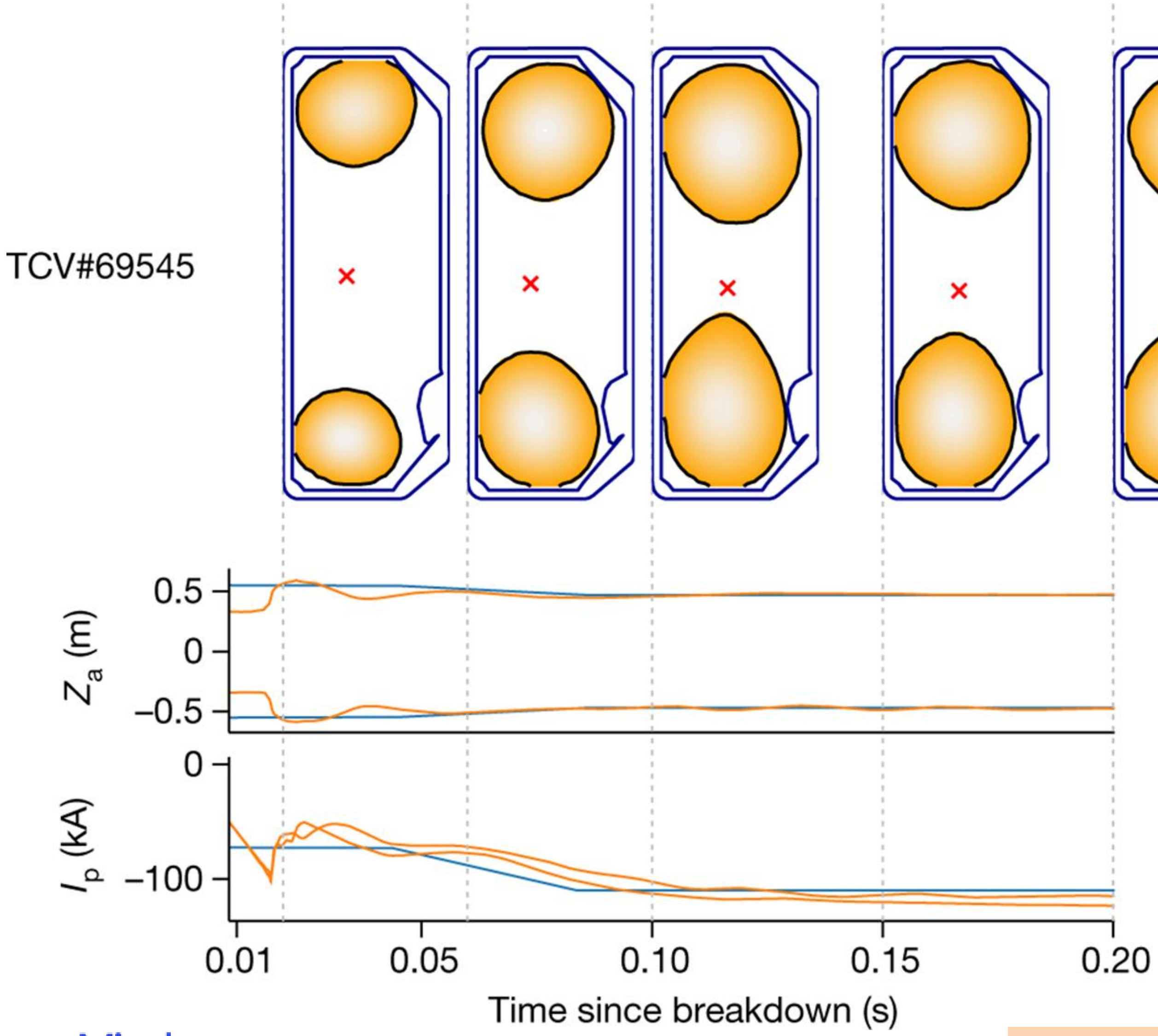
REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

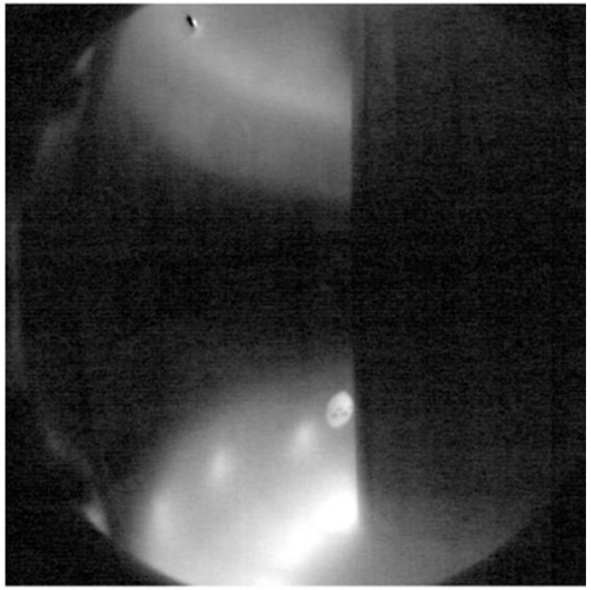
Plasma control by deep reinforcement learning



REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

Plasma control by deep reinforcement learning



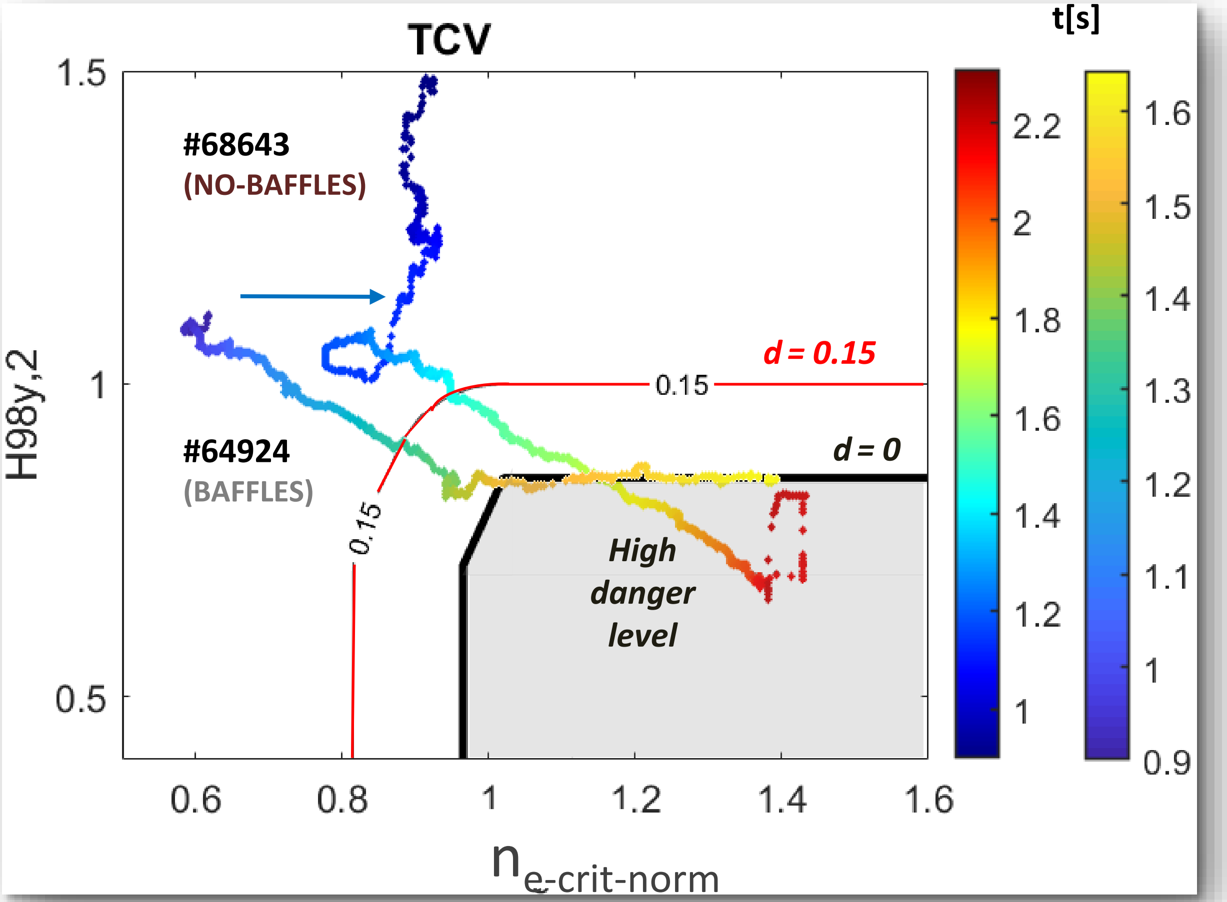


Swiss Plasma Center

DeepMind

REF: **J. Degrave, F. Felici** et al., *Nature*, vol. 602, no. 7897, pp. 414–419, Feb. 2022

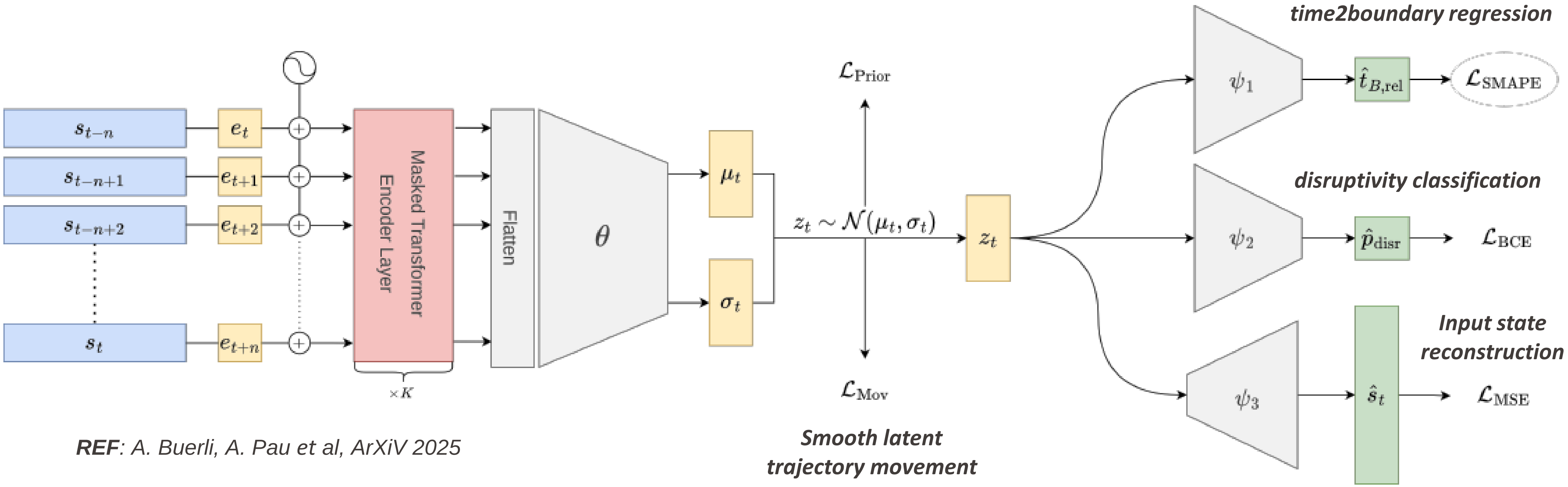
Plasma trajectories: control objectives & optimization



- **Trajectory** allows mapping the distance from operational/disruption boundaries to **probabilities of transition** to specific **states** (occurrence of an **event**)

- **High-performance Density Limit** dynamics described through trajectories in a physics-based **“state space” [H98y,2-$n_{e\text{-crit-norm}}$]**
- **Reduced set of variables** to describe the relevant system dynamics… but sometimes the operational space is complex and high-dimensional (hidden factor, etc.)

Latent variable models for plasma state monitoring



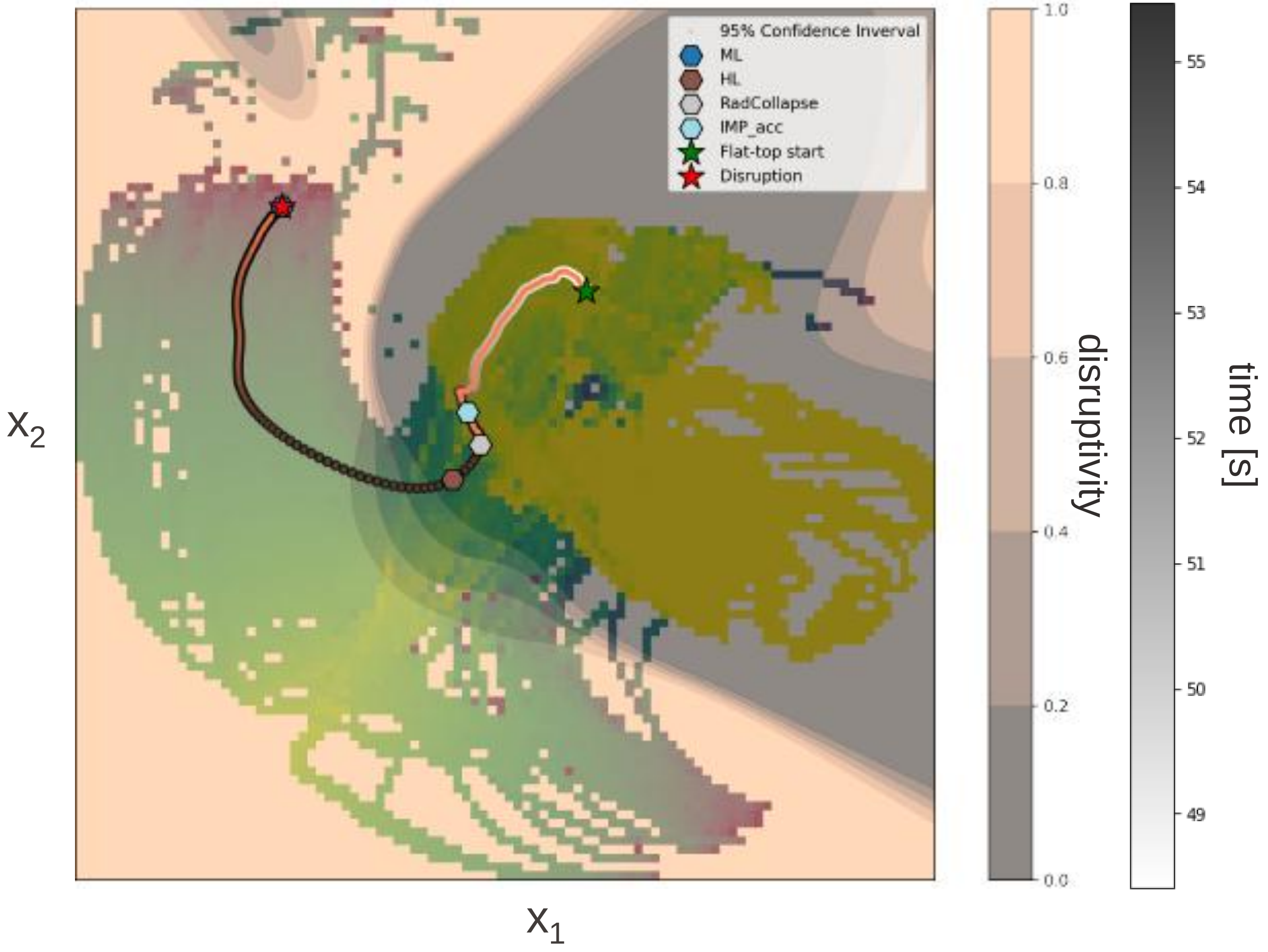
***REF**: A. Buerli, A. Pau et al, ArXiV 2025*

- **Sequence-based** model: a variational autoencoder (transformer, GPT-alike architecture) leveraging **multi-task learning.**
- **Multi-task learning:** by learning tasks jointly(supervised and unsupervised), the model can discover common features or structures across tasks (shared representation).

Plasma State Monitoring with sequence-based DL

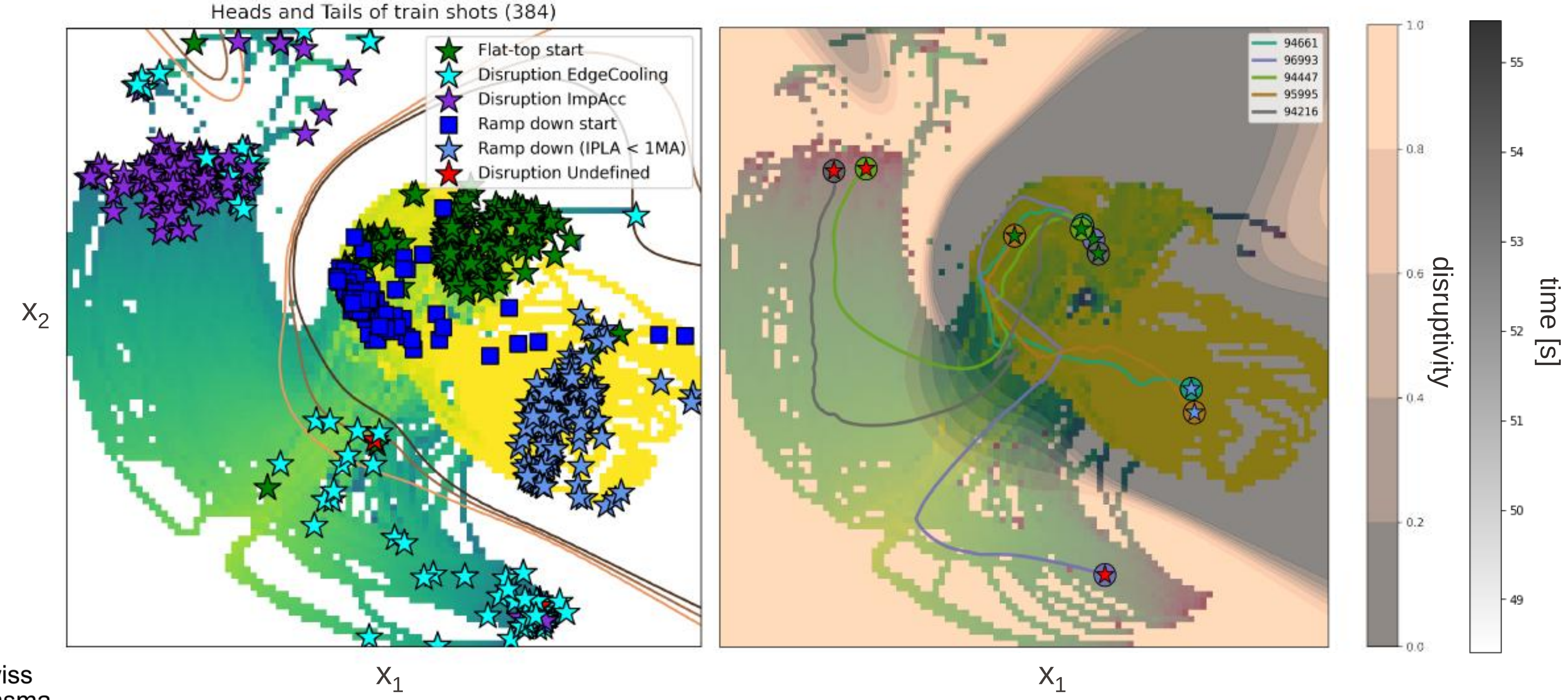
REF: A. Buerli, A. Pau et al, paper sumbitted

- Discharge **plasma state evolution** in time projected on the learned low-D latent space representation (zero-centred and with spherical edges).
- Smoother trajectories with **movement penalty**
- Two main and **diverging paths** along which the state can evolve towards the **boundary condition**, which corresponds to **two different disruptions dynamics at JET**



Plasma State Monitoring with sequence-based DL

REF: *A. Buerli, A. Pau et al, paper submitted*

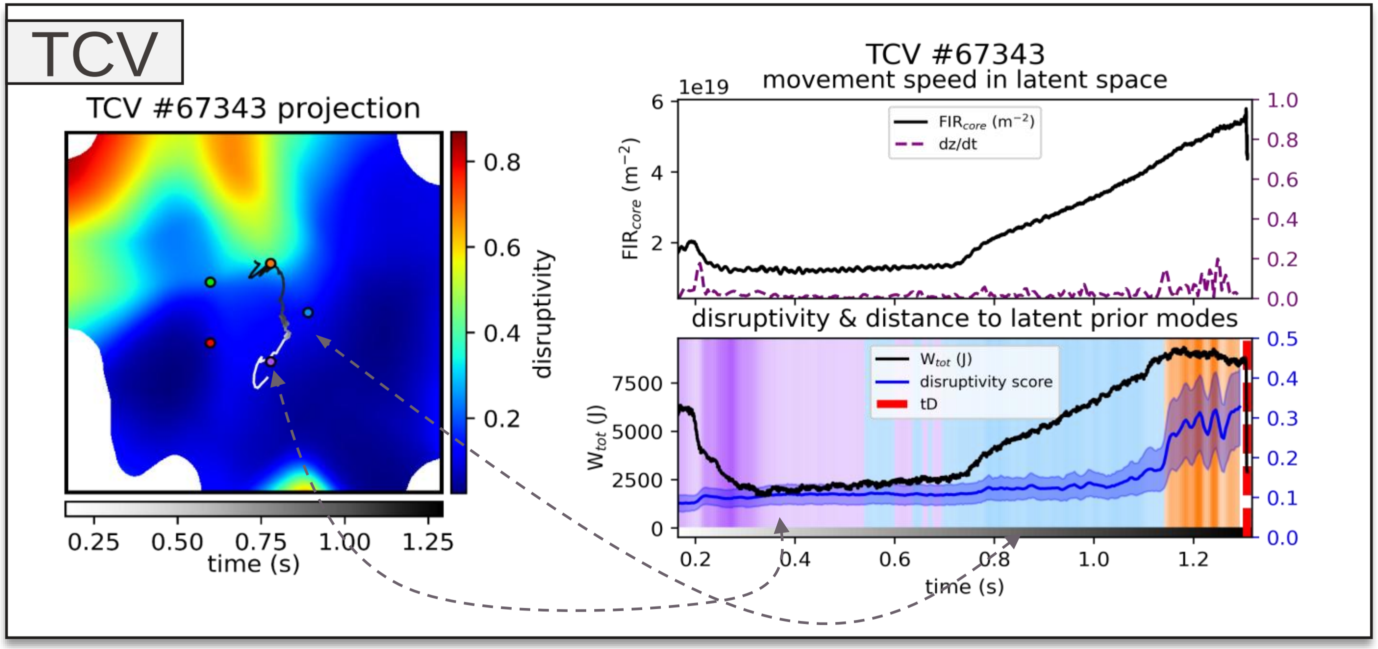


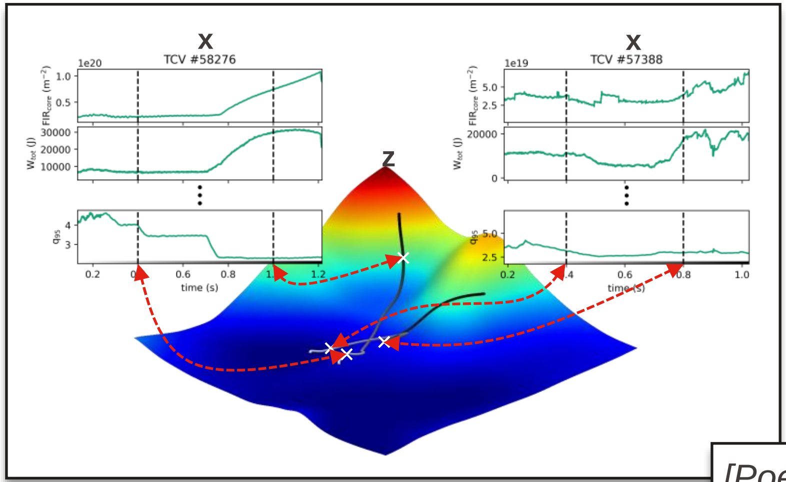
Latent variable models for disruption monitoring

Sequential VAE with multimodal prior

- Project **disruptive boundaries** & physics quantities to inspect connections
- Project full discharges to track proximity to disruption
- *Future: Investigate identified modes in posterior distribution*
- *Future: Discretize projections as sequences of states*

Swiss Data Science Center



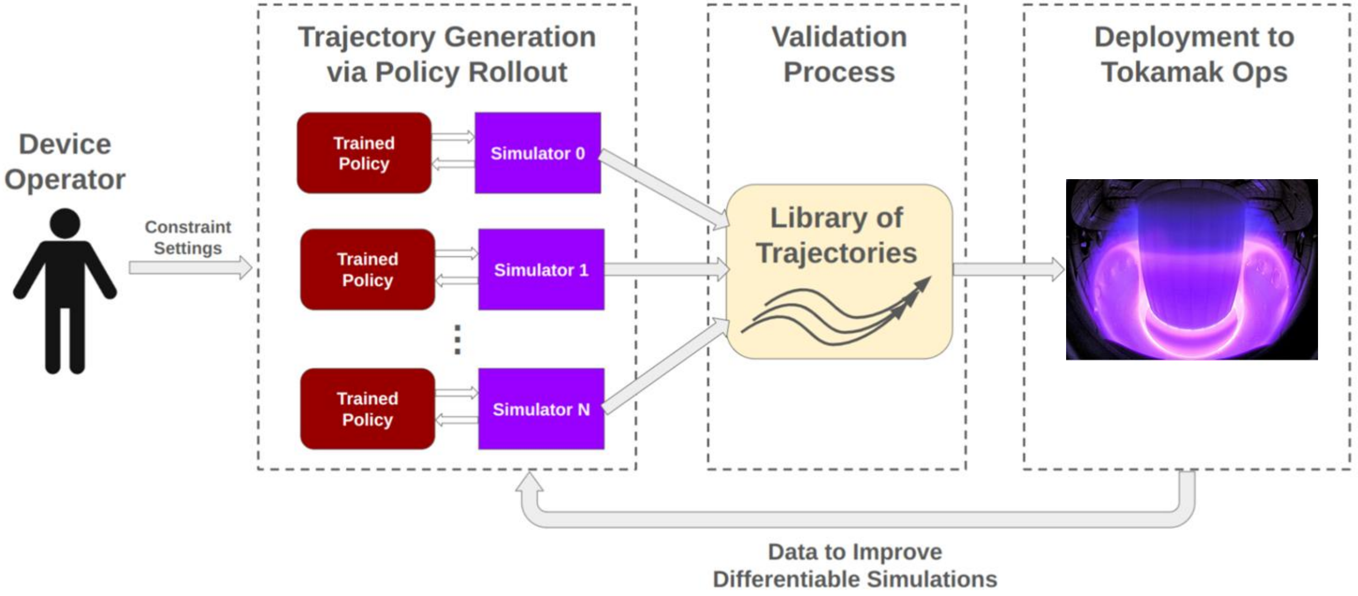


[Poels et al. WIP]

Plasma trajectories optimization with physics constraints

COLLABORATION WITH MIT

- Scientific machine learning for building simulators that **combine physics + machine learning**
- Reinforcement learning to design **trajectories** and **controllers** to meet operator specifications that are robust to **physics uncertainty**



A. Wang, A. Pau, et al. Submitted to Nature Communications

- **Trajectory**: sequence of states, actions, and rewards that an agent experiences as it interacts with the environment.
- **Neural State Space Models** (**NSSM**) to learn the temporal dynamics of some observed quantities in response to actions (physics structure and data-driven models).

Plasma trajectories optimization with physics constraints

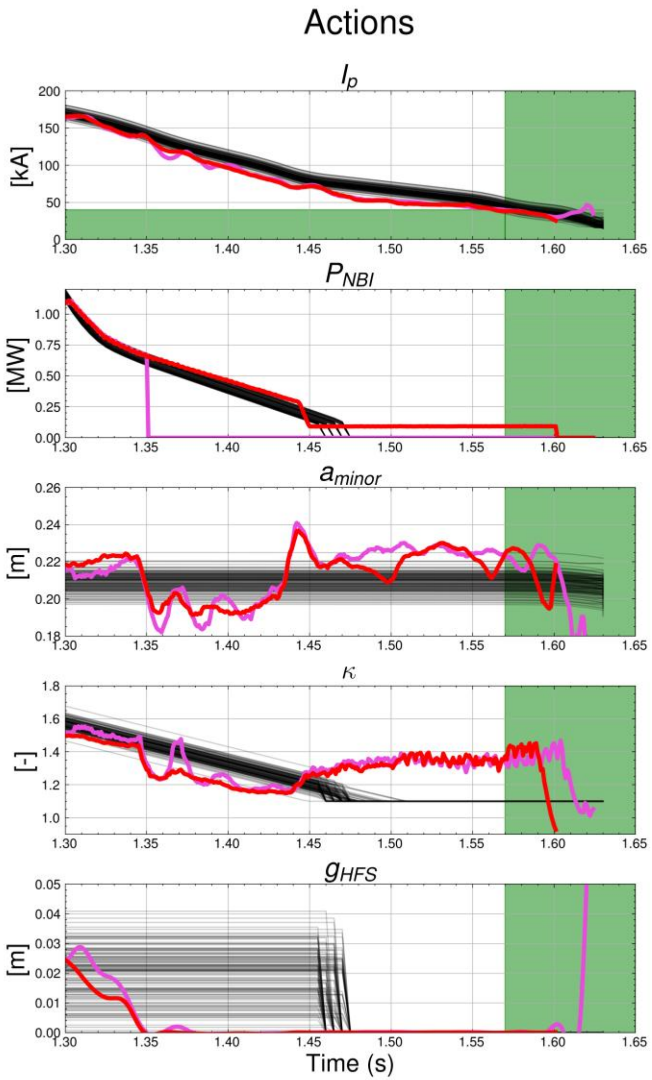
COLLABORATION WITH MIT

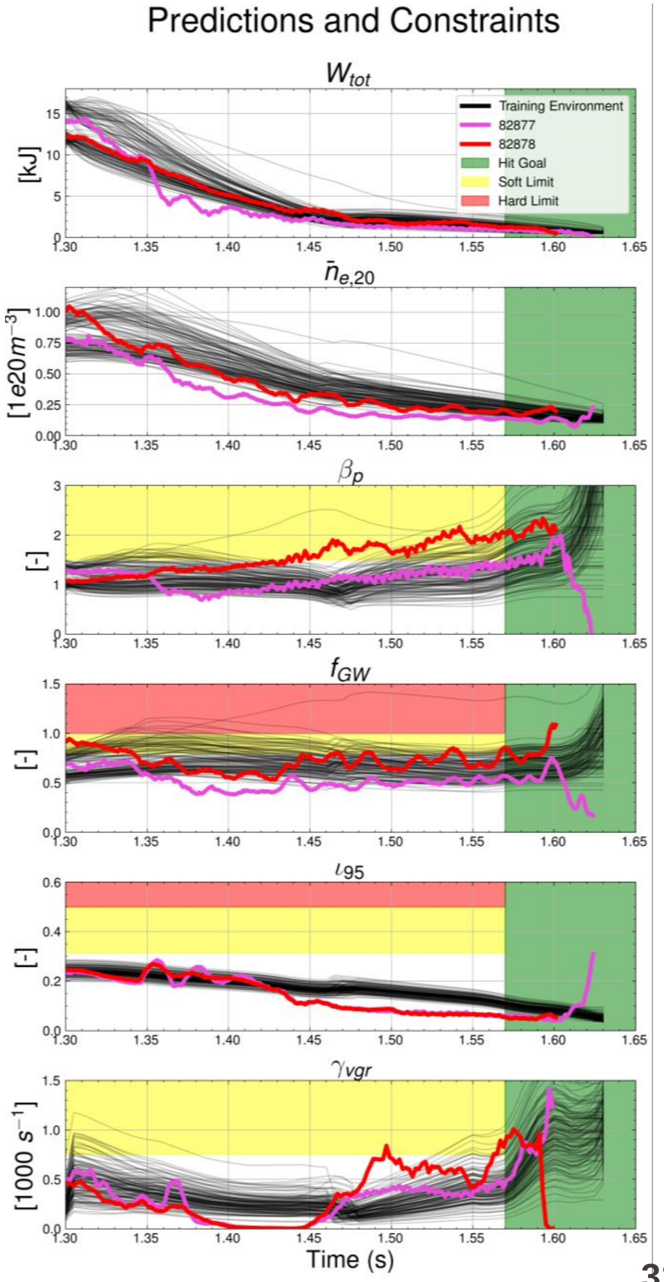
Reward function

$$r(\mathbf{x}(t),\mathbf{a}(t)) = \underbrace{-c_{time}}_{\text{Penalty for time}} - \underbrace{c_W W_{tot}(t) - c_{I_p} I_p(t)}_{\text{Penalty for current and energy}} - \underbrace{\sum_{i=1}^{n_{soft}} c_{soft} s_i(\mathbf{x}(t))}_{\text{Soft chance-constraints}} - \underbrace{\sum_{i=1}^{n_{hard}} c_{hard} h_i(\mathbf{x}(t))}_{\text{Hard chance-constraints}}$$

Reward function parameters

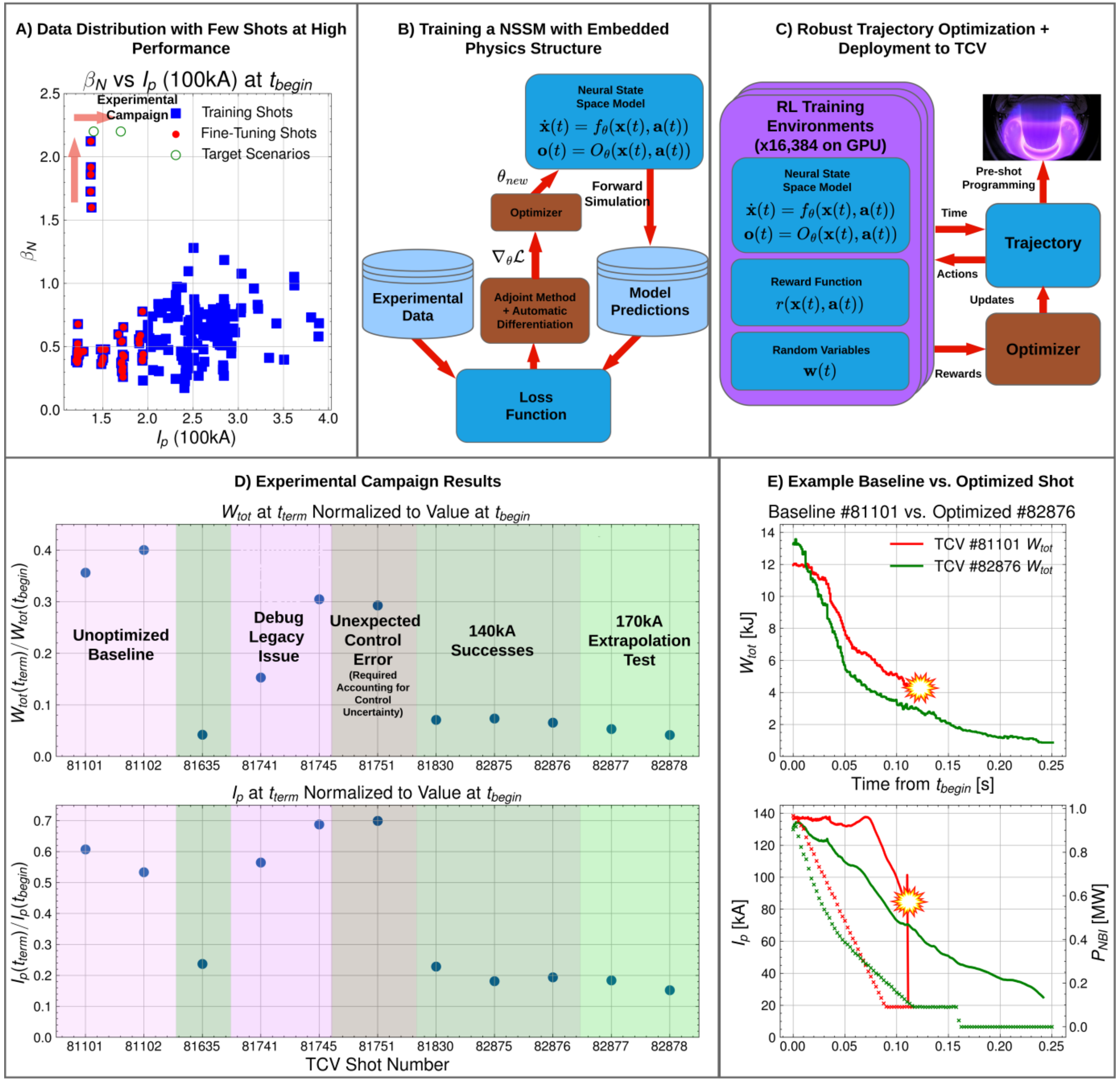
Category	Parameter	Value
Hard Limits	f_{GW}	1.0
	ι_{95}	0.5
Soft Limits	f_{GW}	0.8
	β_p	1.75
	γ_{vgr}	0.75
	ι_{95}	0.313
Parameters	c_{time}	5.0
	c_{I_p}	1.0
	c_W	1.0
	c_{soft}	1.0×10^3
	c_{hard}	5.0×10^4





Swiss Plasma Center

Plasma trajectories optimization with physics constraints

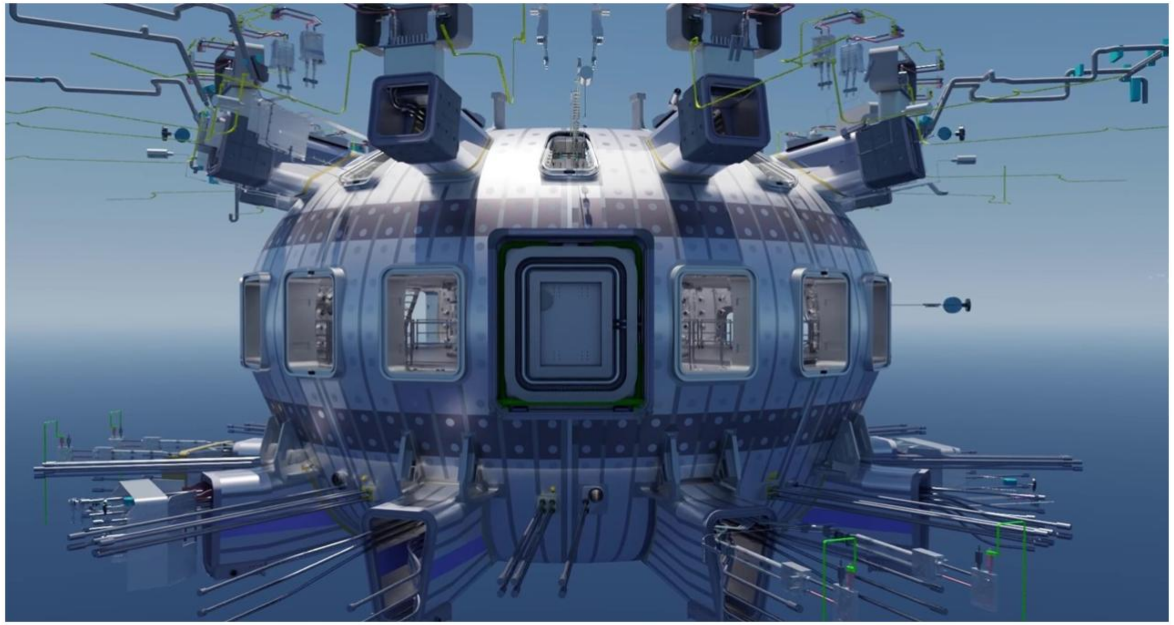


- Initial trial with Proximal Policy Optimization (PPO), before moving to an "**Evolutionary Strategies**" (OpenAI-ES), an algorithm designed for policy optimization.
- ES approaches provide an efficient framework for problems with **long time horizons** and actions that have **long-lasting effects**.

Swiss Plasma Center

Summary and outlook

THE FUTURE: **DIGITAL TWINS**



NVIDIA Omniverse

- AI & ML bridging the gaps between **modeling**, **simulation** and **experiments:**

  - **Integration of AI technologies** to enhance interpretation capabilities, exploration of alternative ideas in virtual simulated environments to enable scientific discovery.

  - **Reinforcement learning** can enable advanced plasma control, allowing the design of optimal control policies in complex environments leveraging modeling, simulation and experimental data

Backup slides

AI & ML challenges in fusion

*REF: **D. Smith**, IAEA AI4Atoms, 2021*

1. Data ecosystems and cross-device datasets
- Challenges: facility-specific data access, format, and schema; database curation; unsupervised or semi-supervised learning for automated labelling
- MFE area: disruptions, ELMs, instabilities

2. Extrapolation to new devices and transfer learning
- Challenges: strategies for transfer or extrapolation; institutional policies for operational certification
- MFE area: extrapolation to ITER; transfer learning to a new device

3. Real-time control and prediction
- Challenges: streaming data, advanced compute hardware (FPGA, ASICs)
- MFE area: disruption mitigation, stability boundaries

4. Near-real-time analysis
- Challenges: multi-channel, high-bandwidth data; automated analysis; advanced compute hardware
- MFE area: fast cameras, multi-channel fluctuation diagnostics, magnetics

5. Surrogate models for simulations
- Challenges: robustness, convergence
- MFE area: plasma-surface interactions, turbulence and transport, RF, stellarator configuration, multi-scale calculations

6. Physics-based models
- Challenges: strategies to blend physics and ML models; Bayesian inference
- MFE area: ELMs and stability boundaries

7. Big data
- Challenges: data product pipeline, automated analysis, data mirroring and provenance
- MFE area: 2 PB/day at ITER, high bandwidth diagnostics

8. Plant operation
- Challenges: real-time monitoring and information synthesis
- MFE area: thermal power generation, nuclear environment, cryogenics, superconducting magnets, vacuum systems, and diagnostic instrumentation

International Atomic Energy Agency (IAEA) Consultancy Meeting on Artificial Intelligence for Nuclear Fusion Research