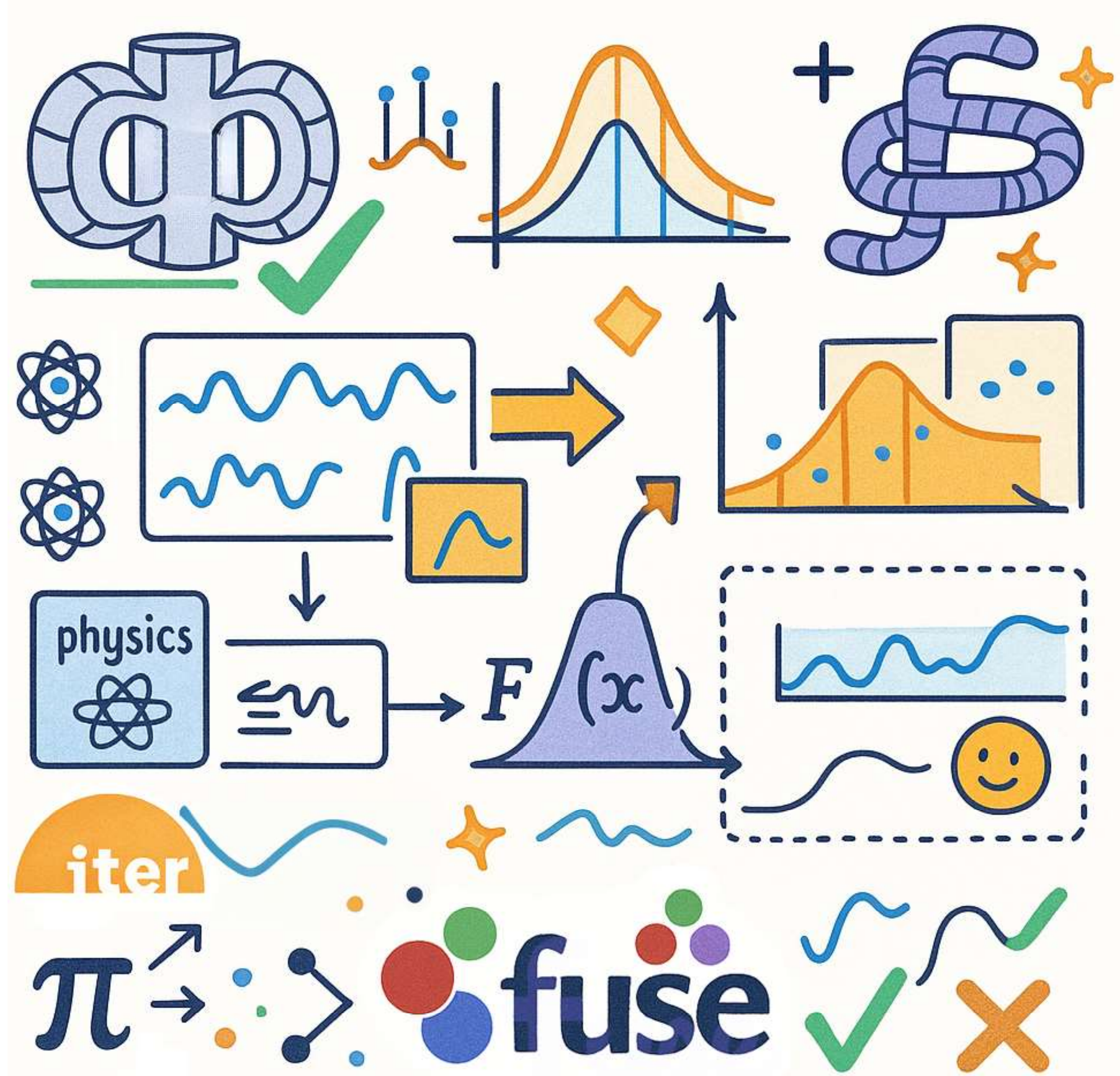
Validation of Integrated Models and Uncertainty Quantification

O. Meneghini

B. C. Lyons, T. F. Neiser, T. Slendebroek,
A. G. Ghiozzi, J. T. McClenaghan, N. Shi,
S. Denk, M. Yoo, G. Avdeeva, G. Dose,
L. Stagner, J. Harvey, J. Guterl,
D. B. Weisberg, T. B. Cote, M. Clark,
A. Zalzali, H. Anand, T. Bechtel, S. P. Smith,
R. M. Nazikian, B. A. Grierson, J. Candy



ITER summer school
30 June – 4 July 2025

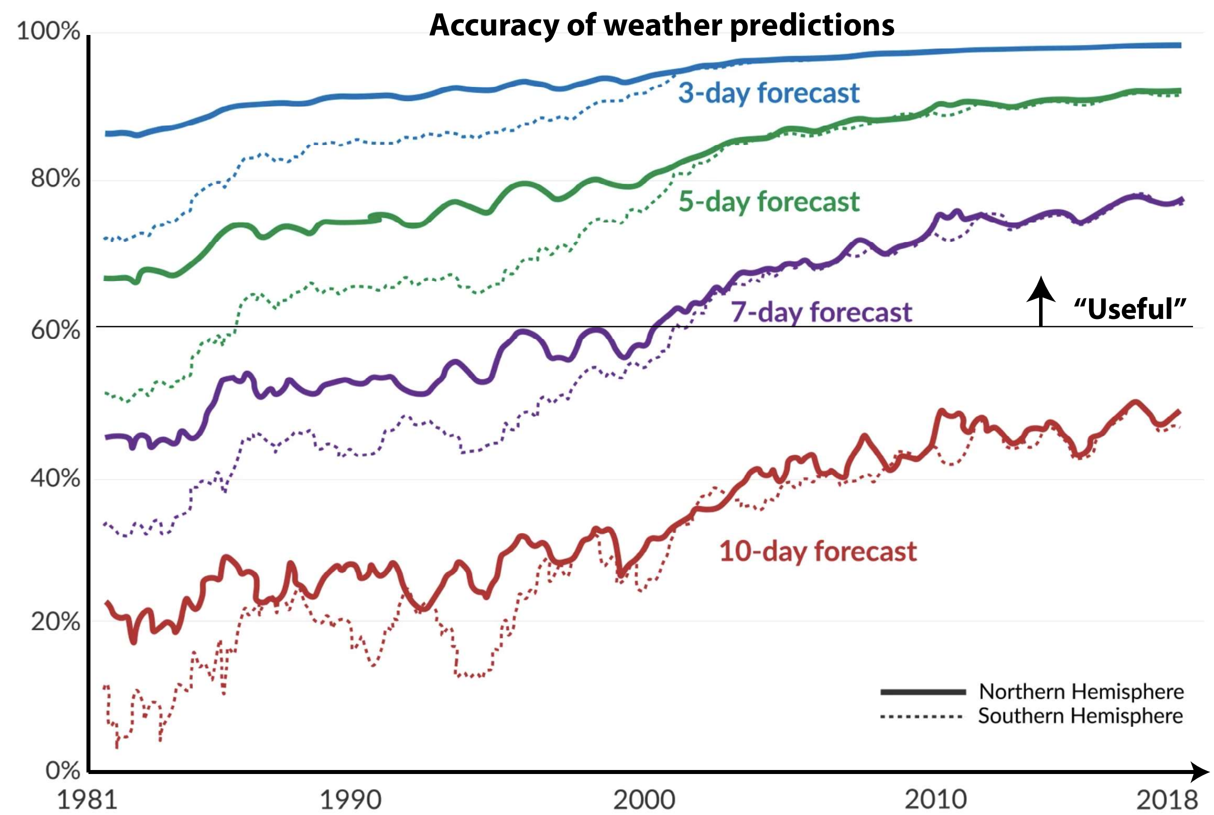
Integrated modelling of magnetic fusion plasmas

GENERAL ATOMICS

“All models are wrong, but some are useful.”
— George Box

Validation is the process of establishing if a model can indeed be useful

- When **applicable**?
- What **uncertainty**?



e.g. It took decades of model development and validation for weather models to became trusted predictive tools

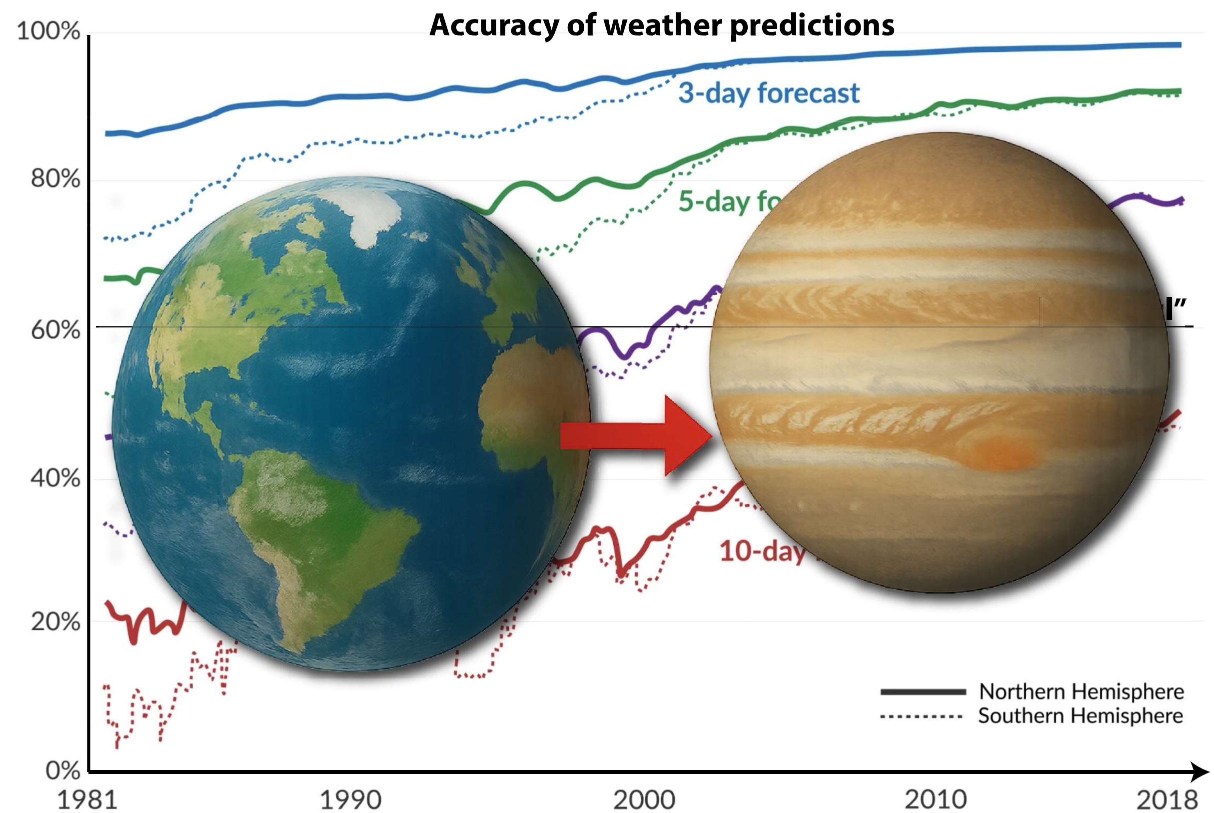
"All models are wrong, but some are useful."
— George Box

Validation is the process of establishing if a model can indeed be useful

- When **applicable**?
- What **uncertainty**?

In fusion, we aim to **predict new operating scenarios and reactors**

- Like predicting weather on Jupiter based on Earth experience
- Extrapolation to unseen regimes requires **theory-based models**

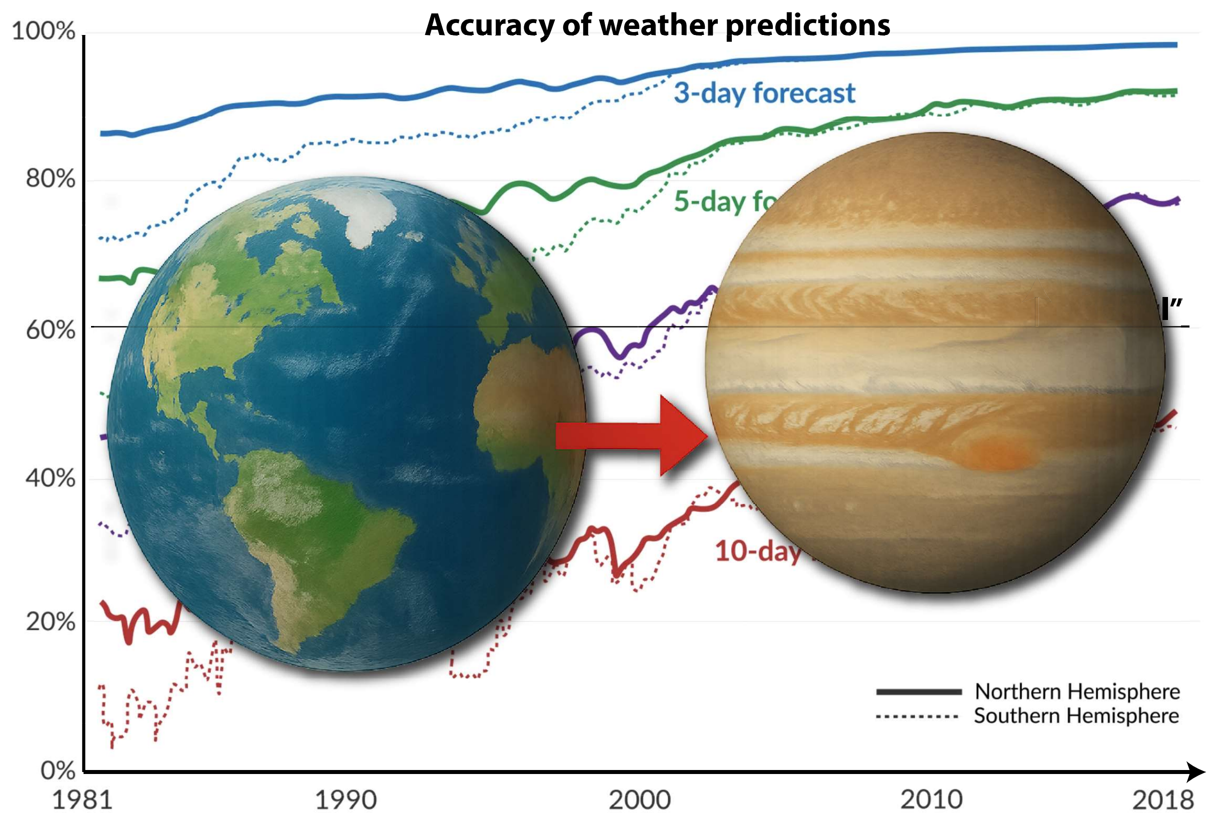


e.g. It took decades of model development and validation for weather models to became trusted predictive tools

“All models are wrong, but some are useful.”
— George Box

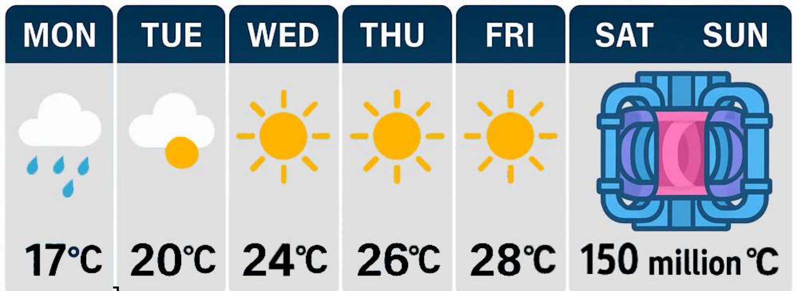
Validation is the process of establishing if a model can indeed be useful

- When **applicable**?
- What **uncertainty**?



In fusion, we aim to **predict new operating scenarios and reactors**

- Like predicting weather on Jupiter based on Earth experience
- Extrapolation to unseen regimes requires **theory-based models**



e.g. It took decades of model development and validation for weather models to became trusted predictive tools

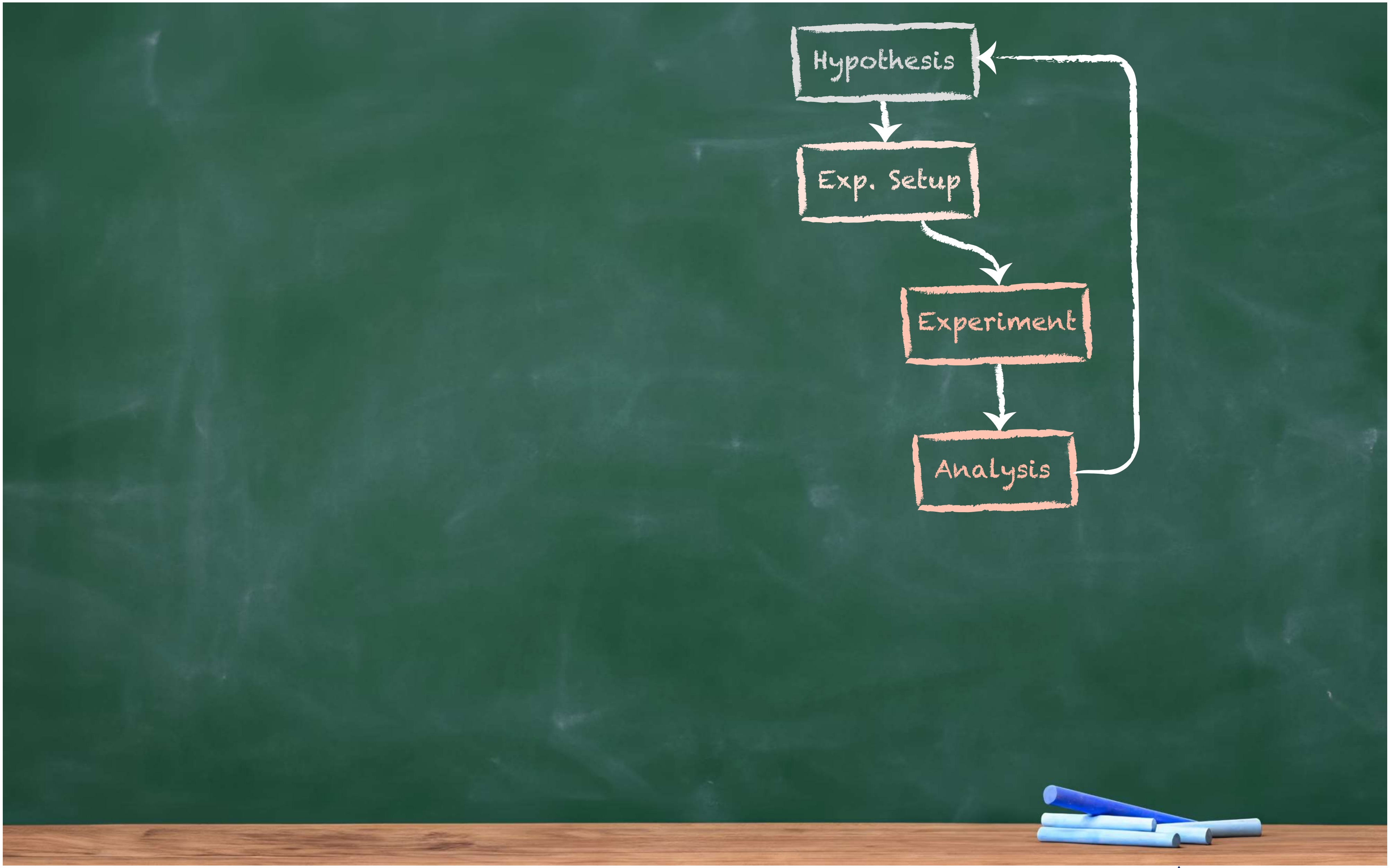
Outline

- Fundamentals of V&V and UQ
  - Validation and Verification
  - Fidelity hierarchy for V&V
  - Statistical validation
  - Uncertainty propagation
- FUSE: Implementing V&V and UQ in practice
  - Machine optimization
  - Integrated data analysis
  - Pulse design
- Key Takeaways

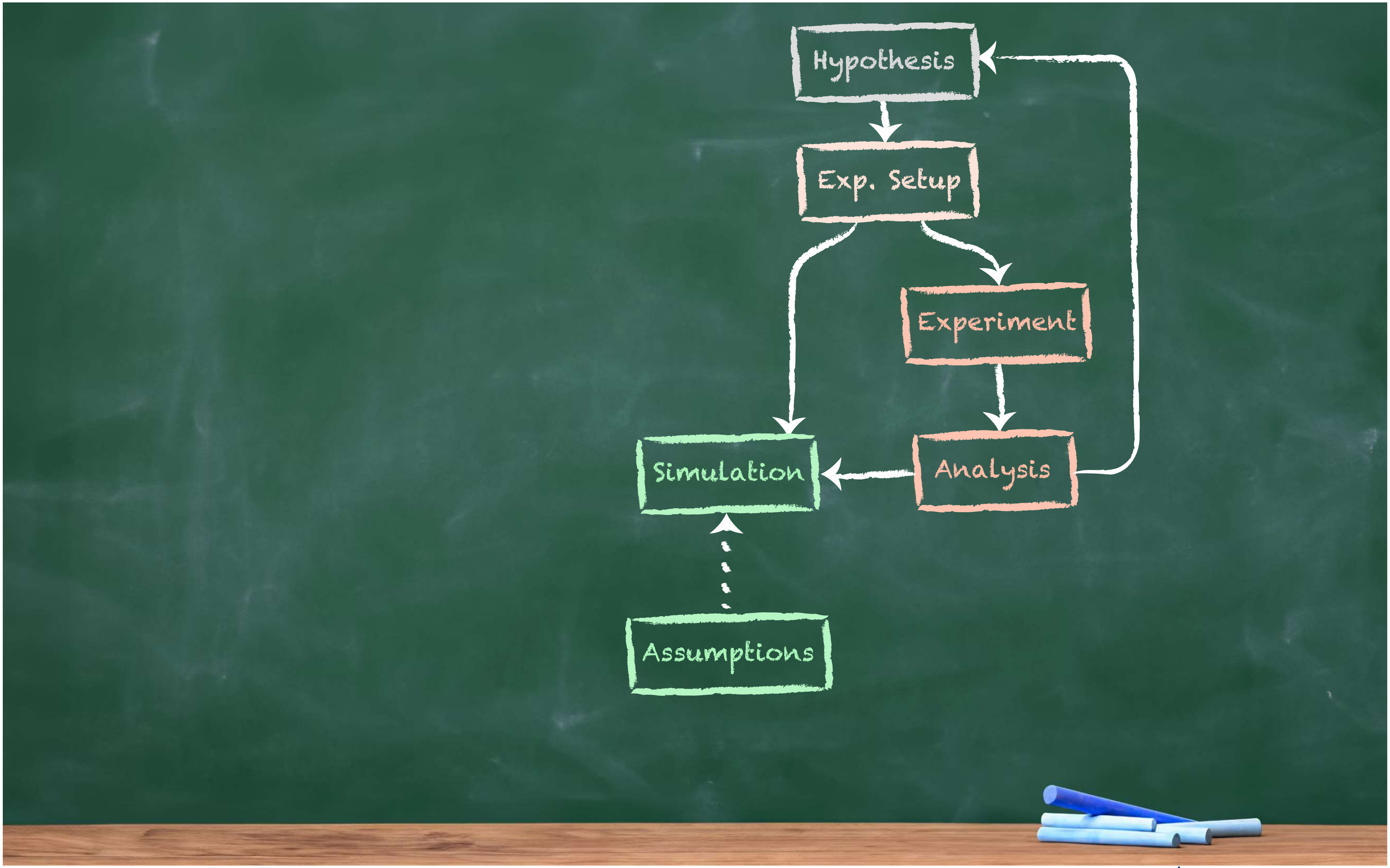
Outline

- Fundamentals of V&V and UQ
    - Validation and Verification
    - Fidelity hierarchy for V&V
    - Statistical validation
    - Uncertainty propagation
- FUSE: Implementing V&V and UQ in practice
    - Machine optimization
    - Integrated data analysis
    - Pulse design
- Key Takeaways

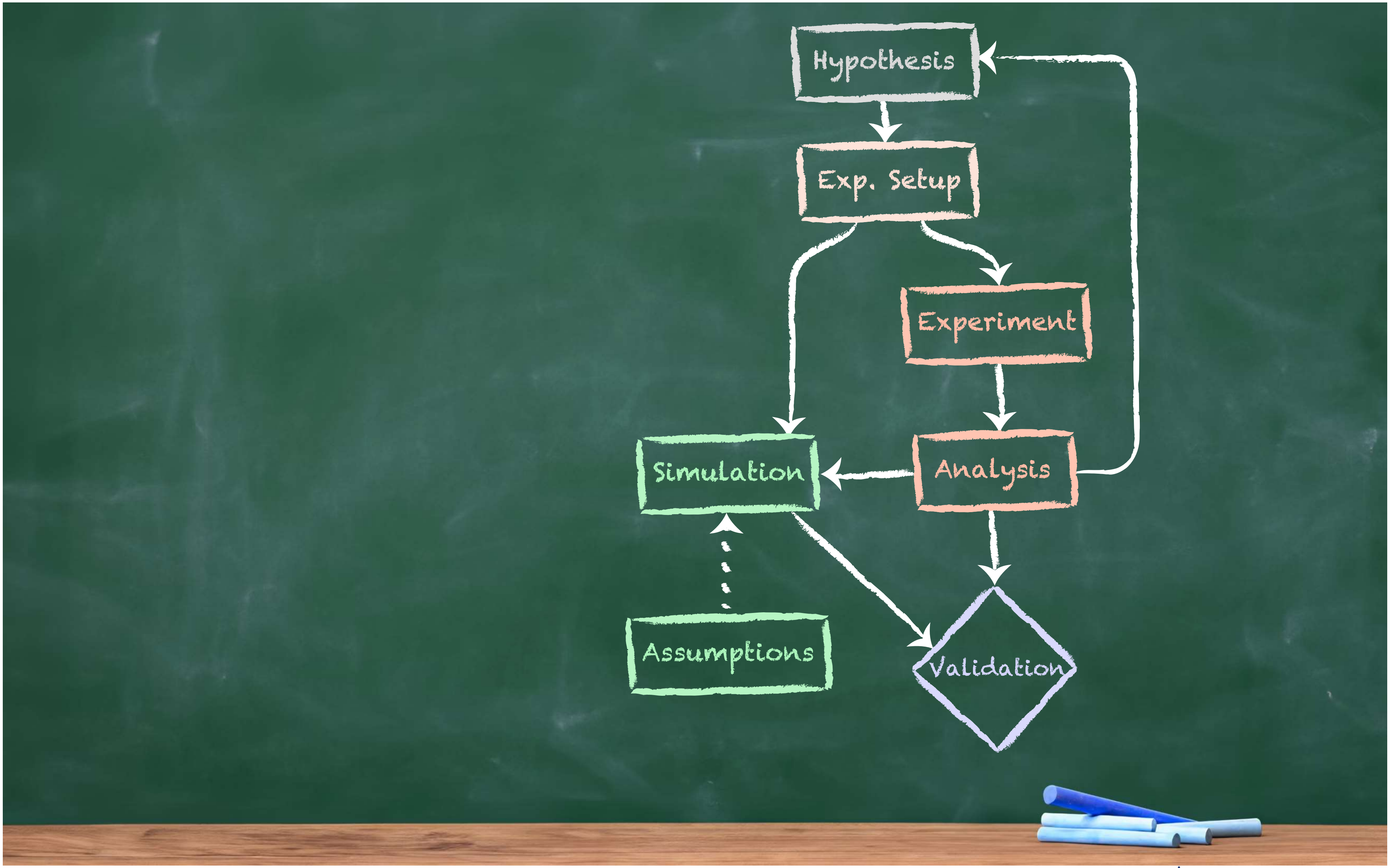
Start with an experiment



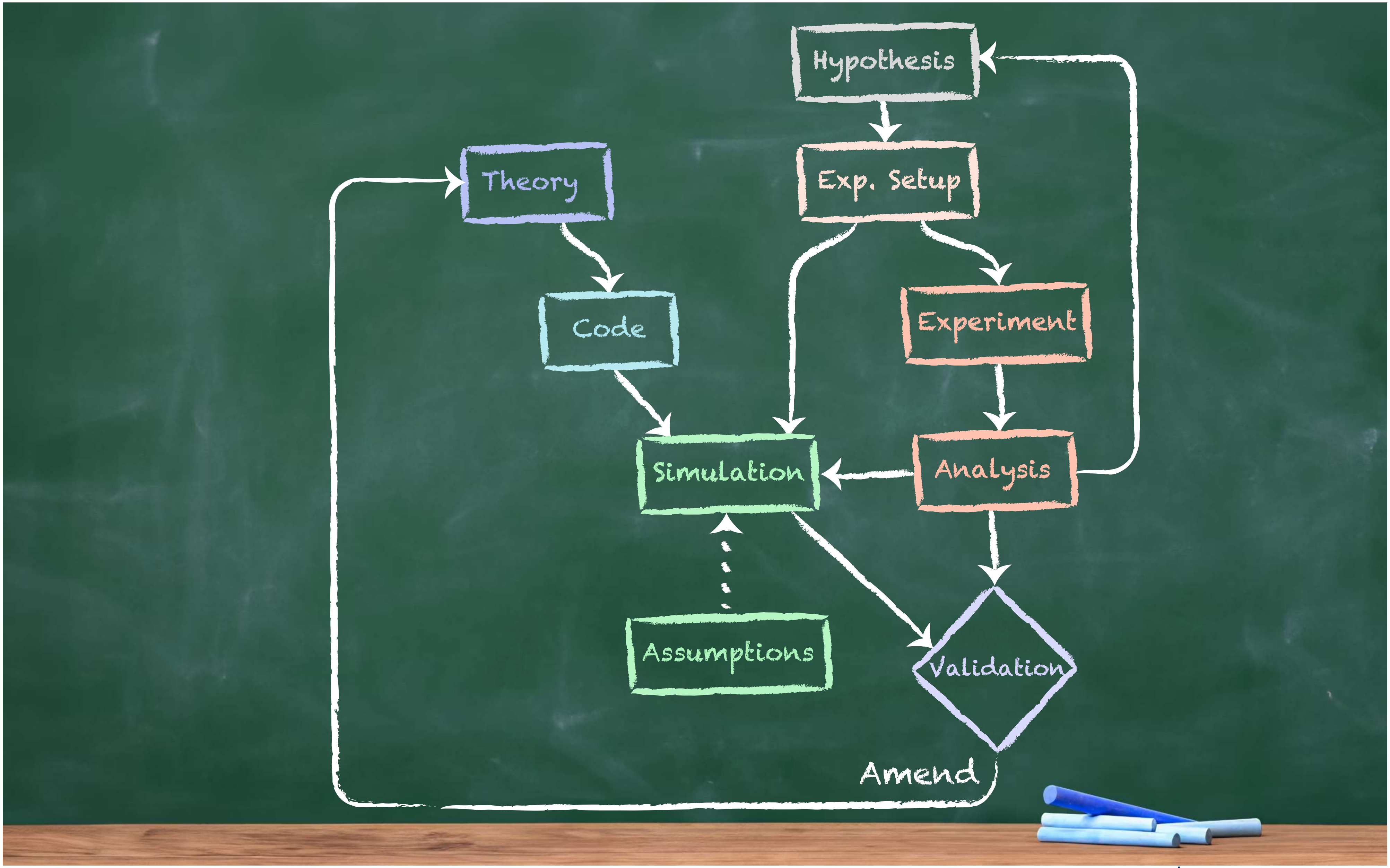
Repeat experiment setup in simulation



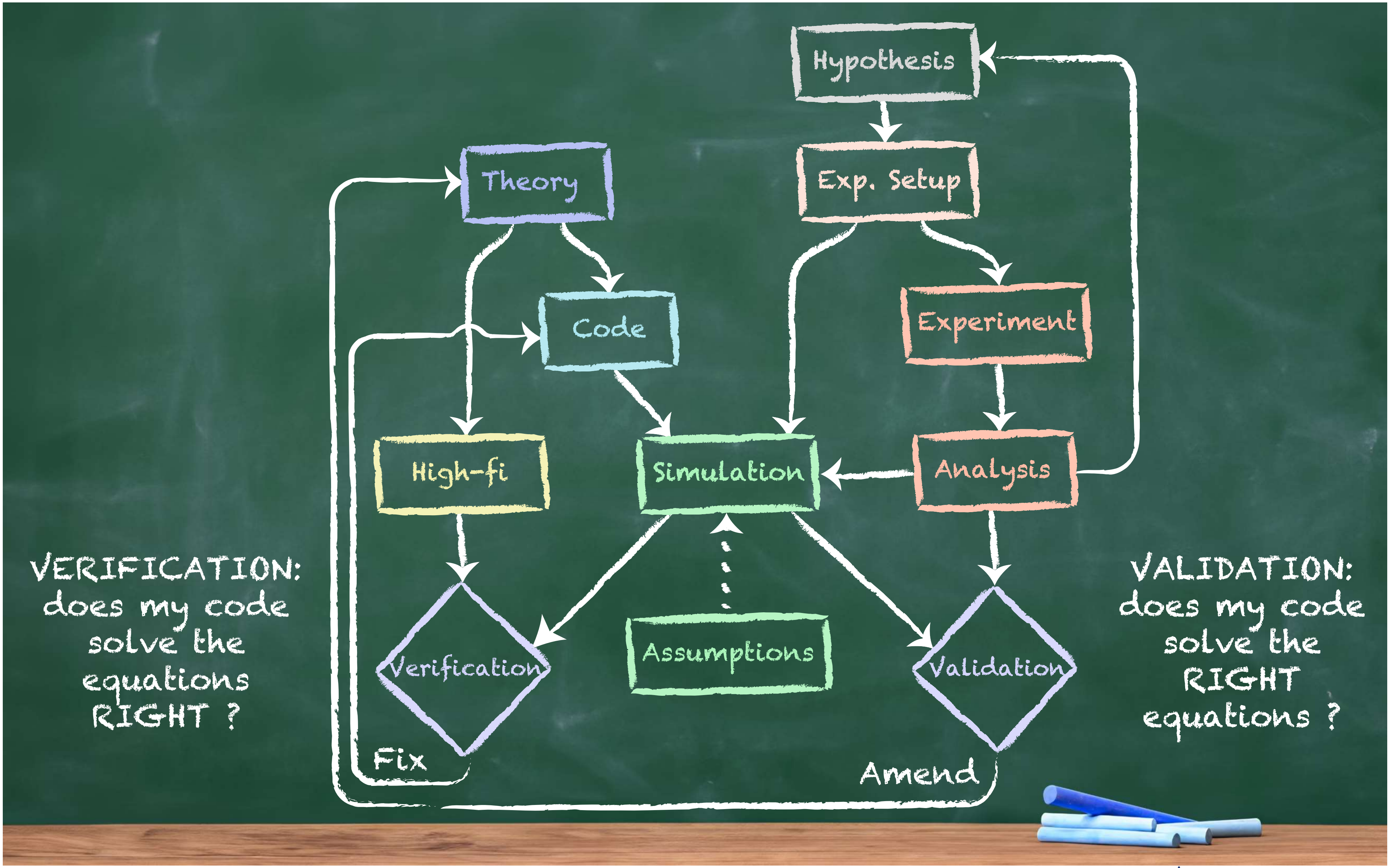
Validation quantifies the agreement between simulation and experiment



If unsatisfactory, we can amend the theory or acknowledge limited region of validity



Validation is not verification, we need both

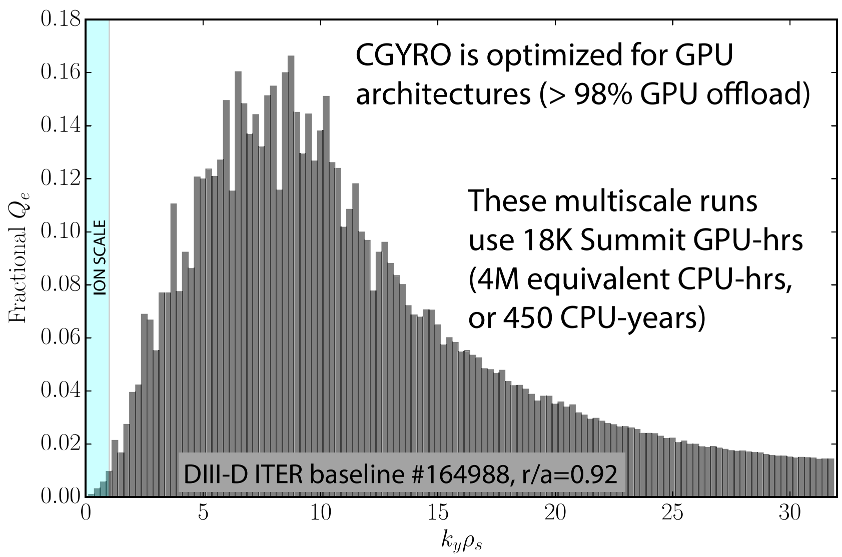


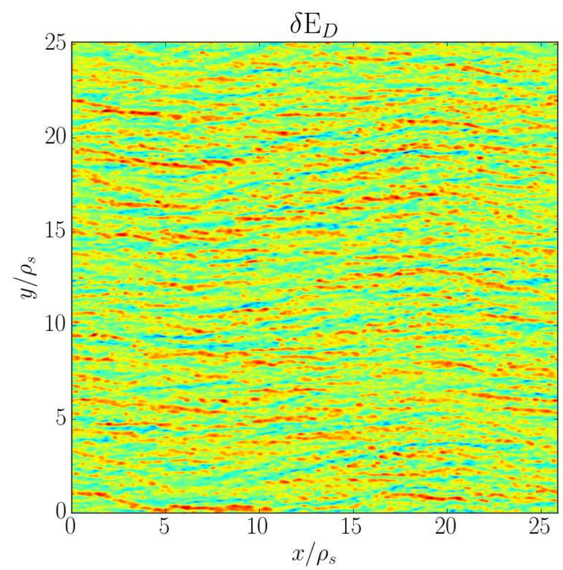
Validation gives confidence in our fundamental understanding of the physics at play

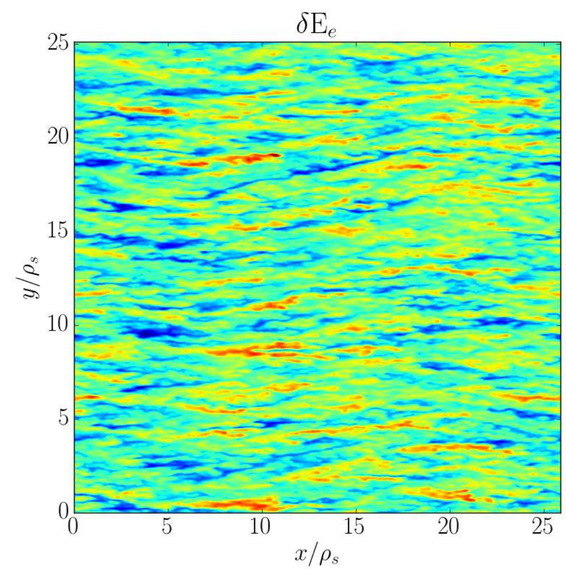
eg. CGYRO+NEO validation with DIII-D data used to establish applicability of multiscale GK in pedestal

	Ions	Electrons
Power balance	2.5	8.2
NEO	2.7	negligible
CGYRO	negligible	8.0

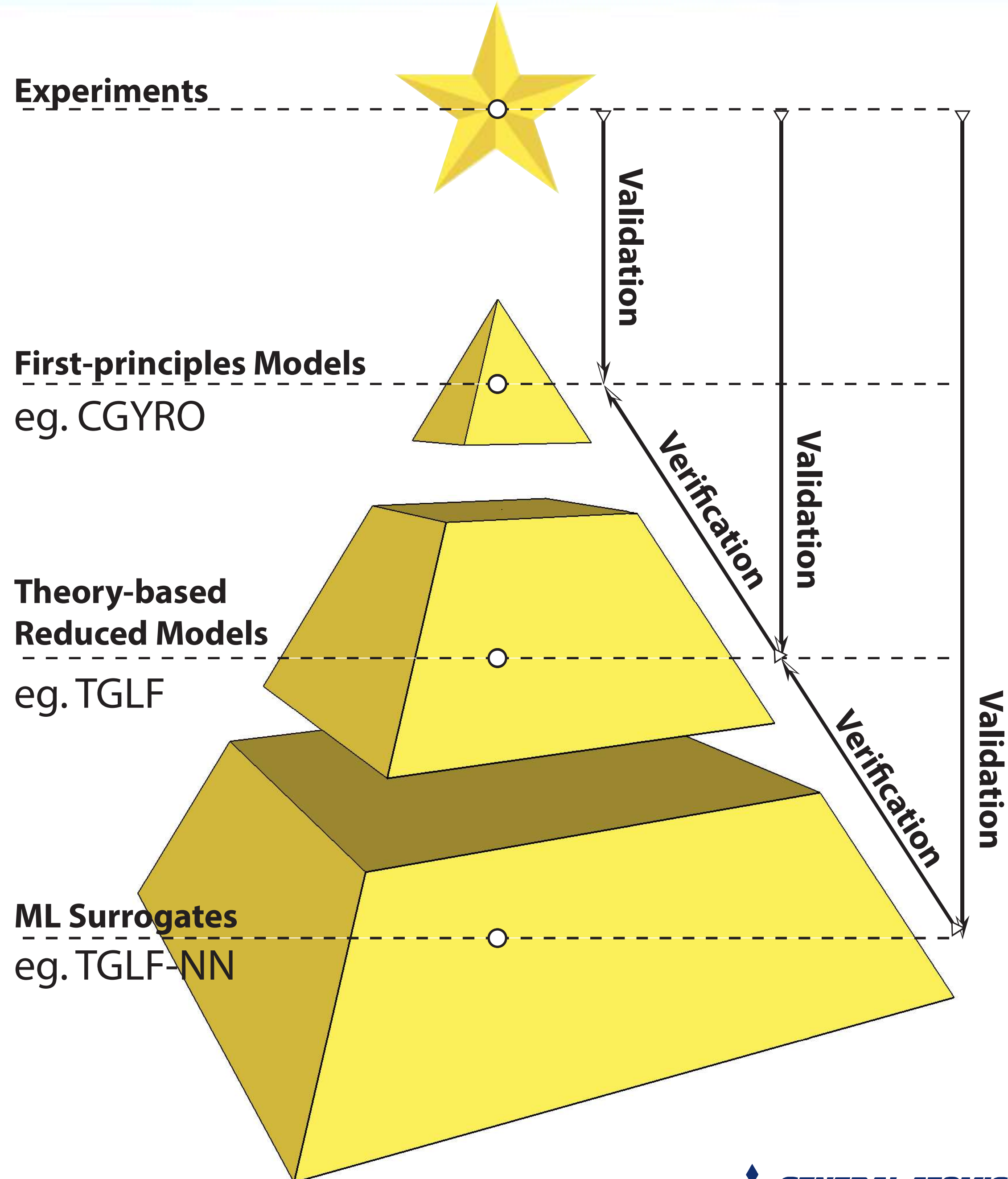
- Have we validated all GK theory? Of course not, we need more points!
- Extensive validation of most computationally expensive first principles models is challenging
- Reduced models are key!



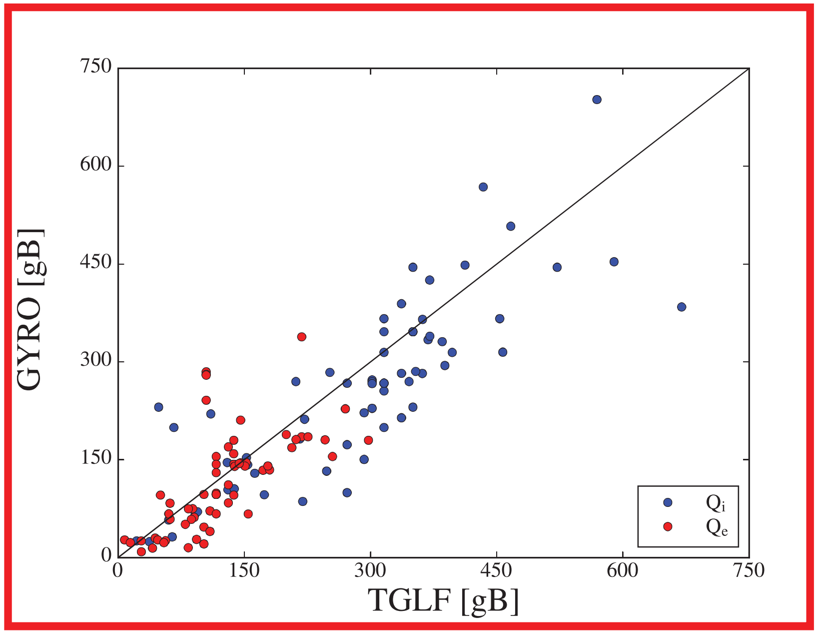


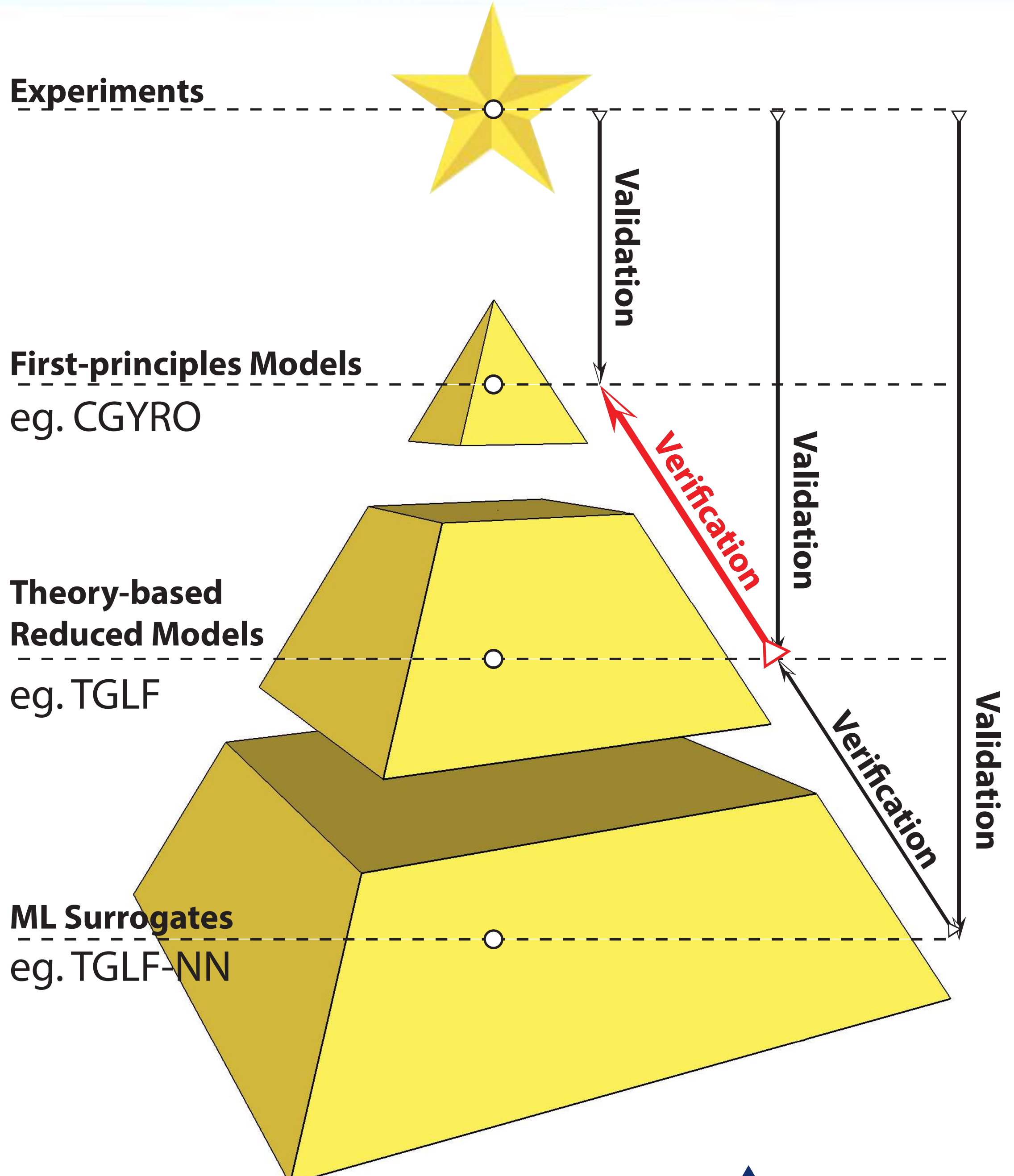


Reduced models and ML surrogates play a key role in providing indirect validation of high-fidelity models

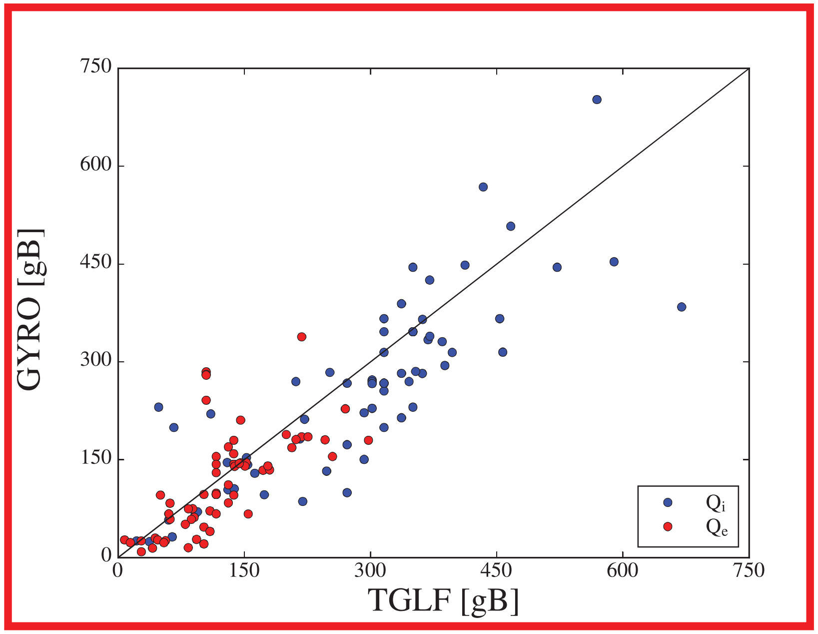


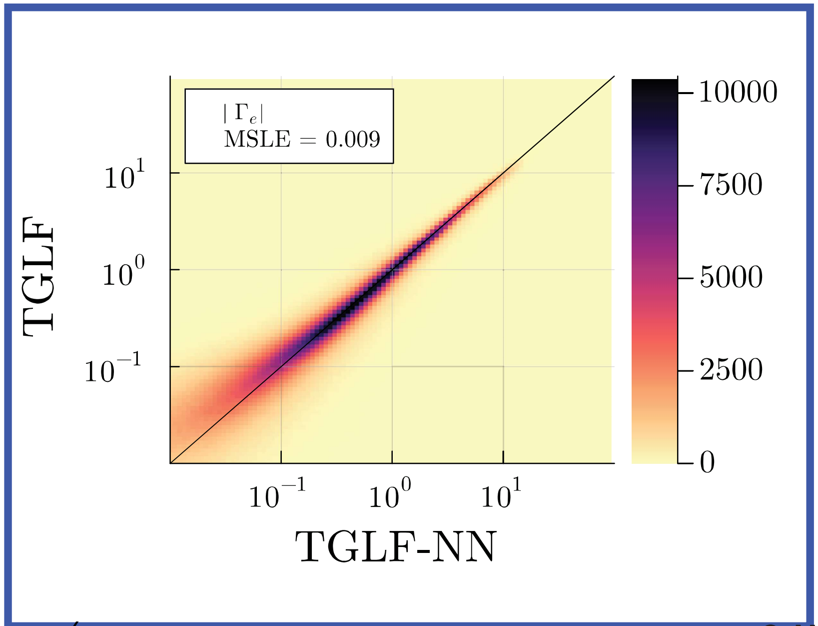
Reduced models and ML surrogates play a key role in providing indirect validation of high-fidelity models

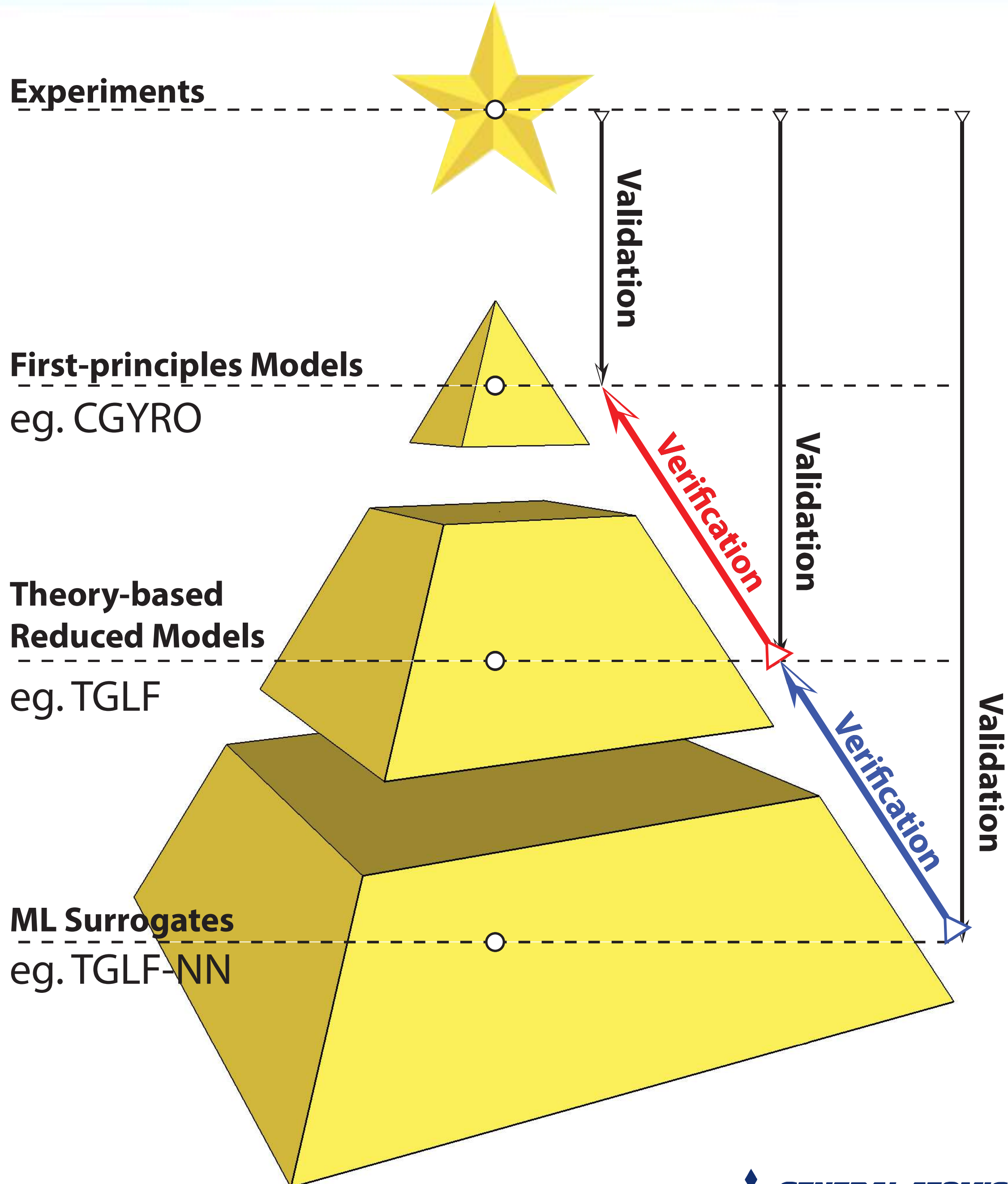




Reduced models and ML surrogates play a key role in providing indirect validation of high-fidelity models

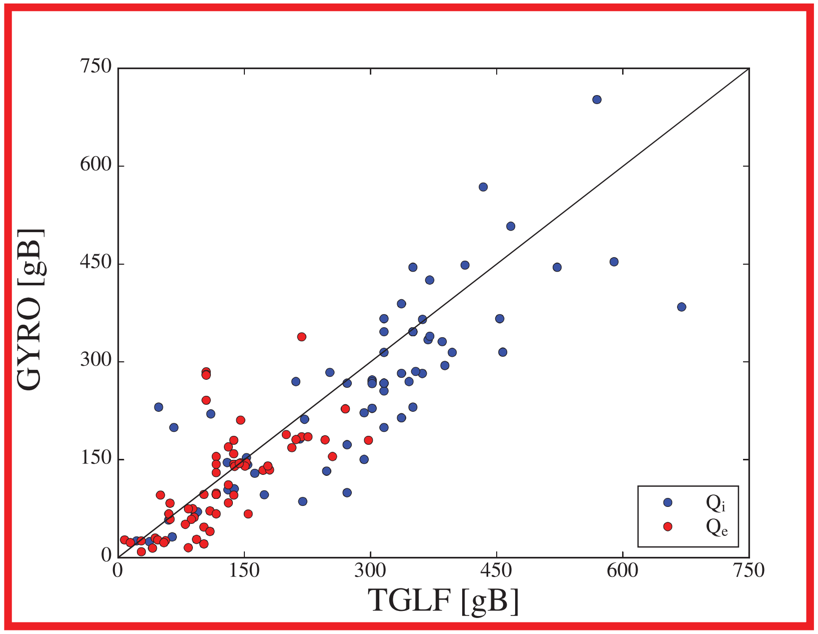


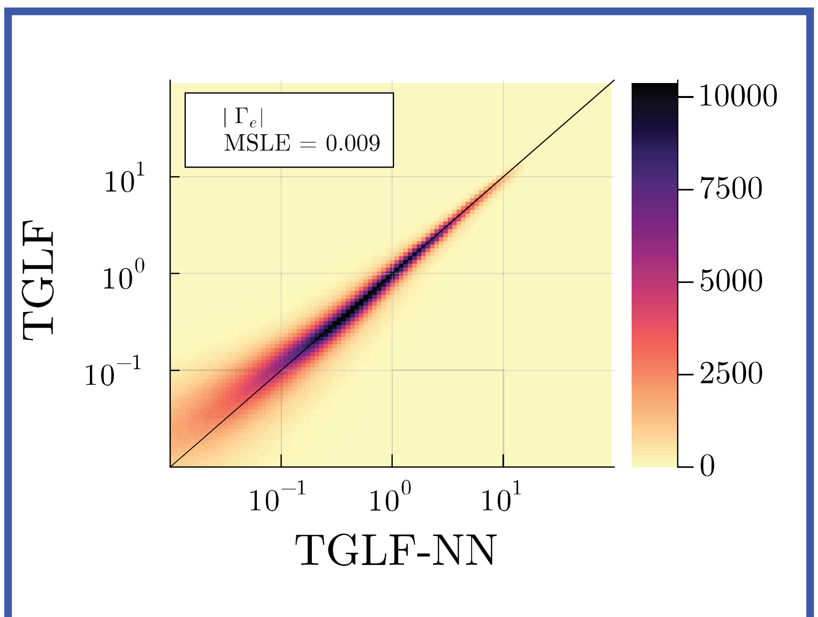


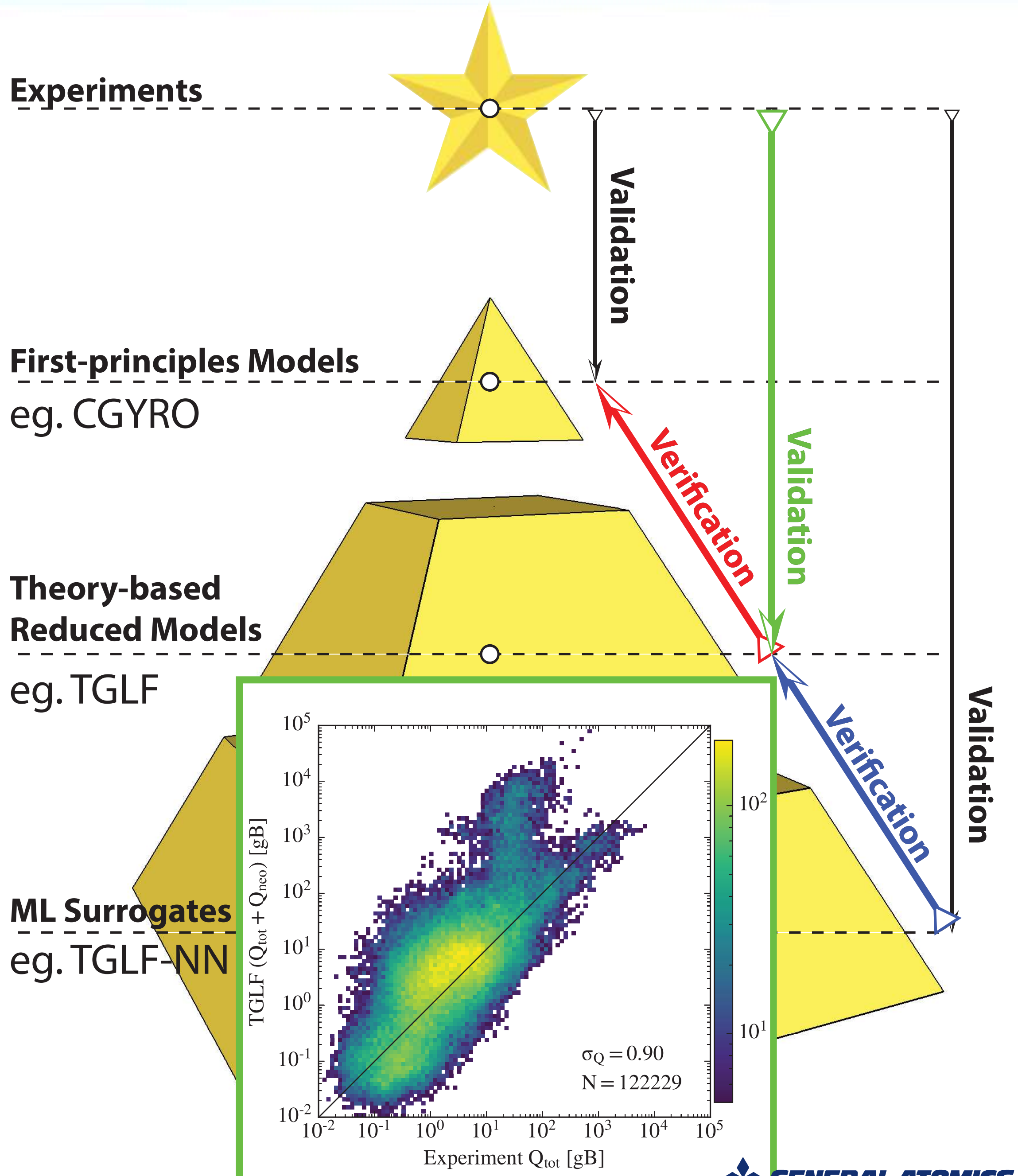


GENERAL ATOMICS

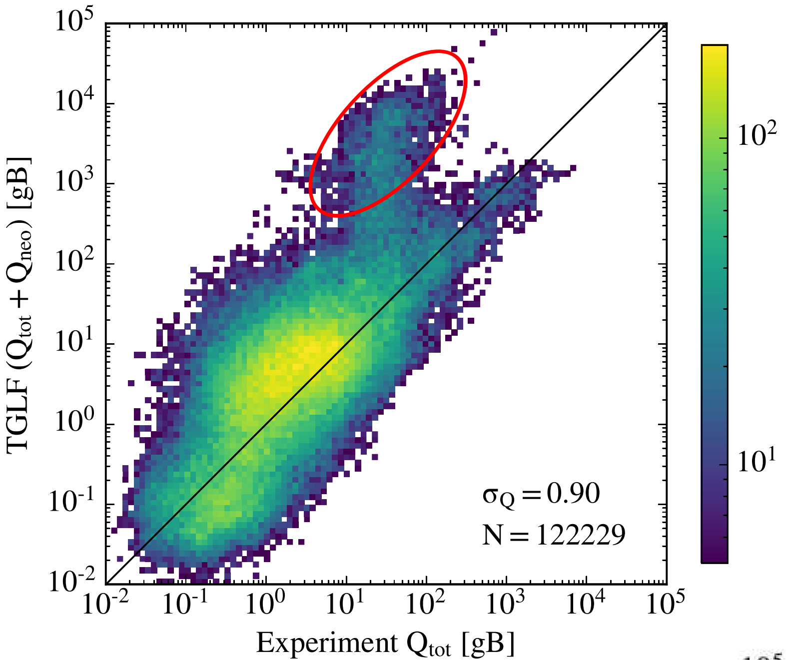
Reduced models and ML surrogates play a key role in providing indirect validation of high-fidelity models







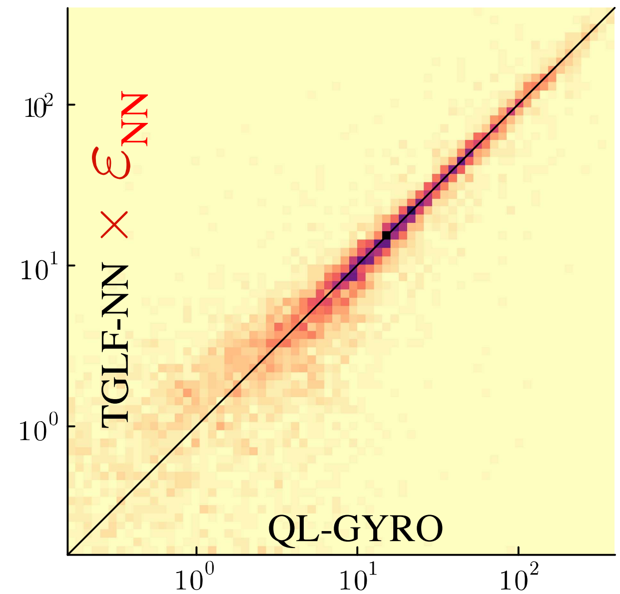
So, validation is not a point, it's a statistical measure spanning large parameter ranges



Verification of TGLF with QLGYRO

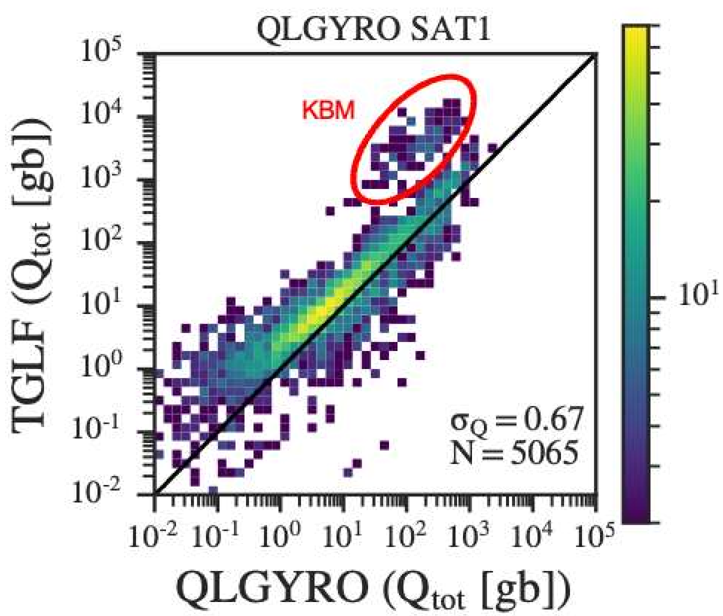
QLGYRO uses:
- linear CGYRO spectra
- TGLF saturation rule

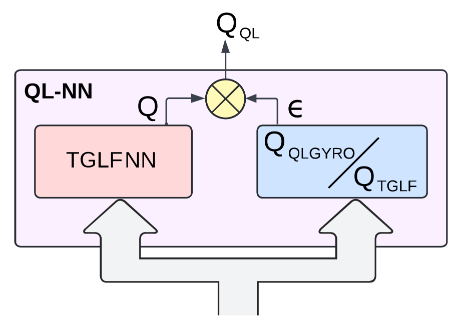
Identifies limitation with the TGLF linear growth rate spectra model



Validation identifies a regime where TGLF is not great.

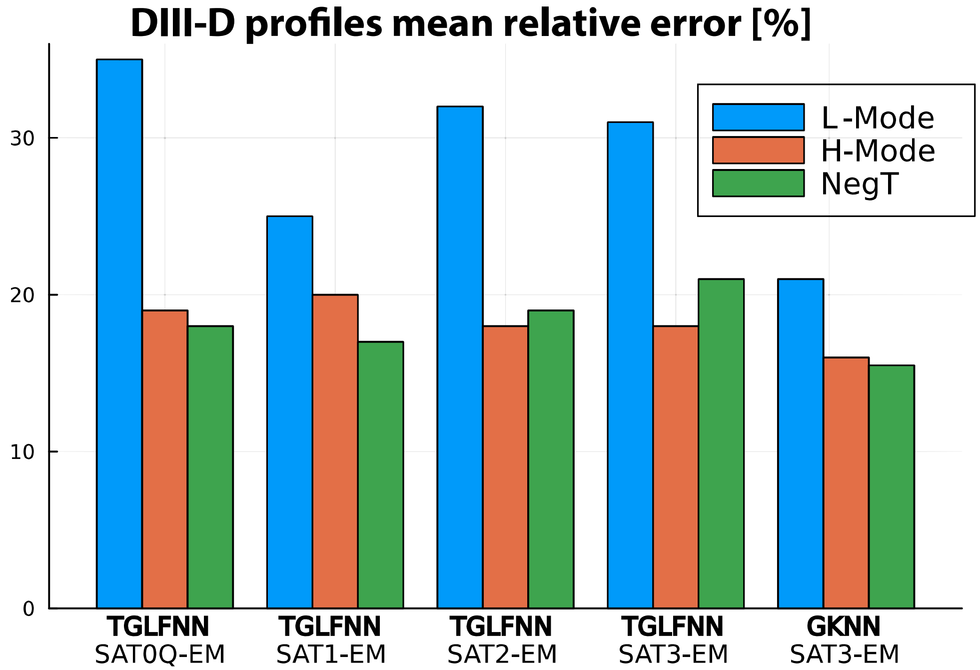
Is it an issue with gyrokinetic theory or a limitation of TGLF?

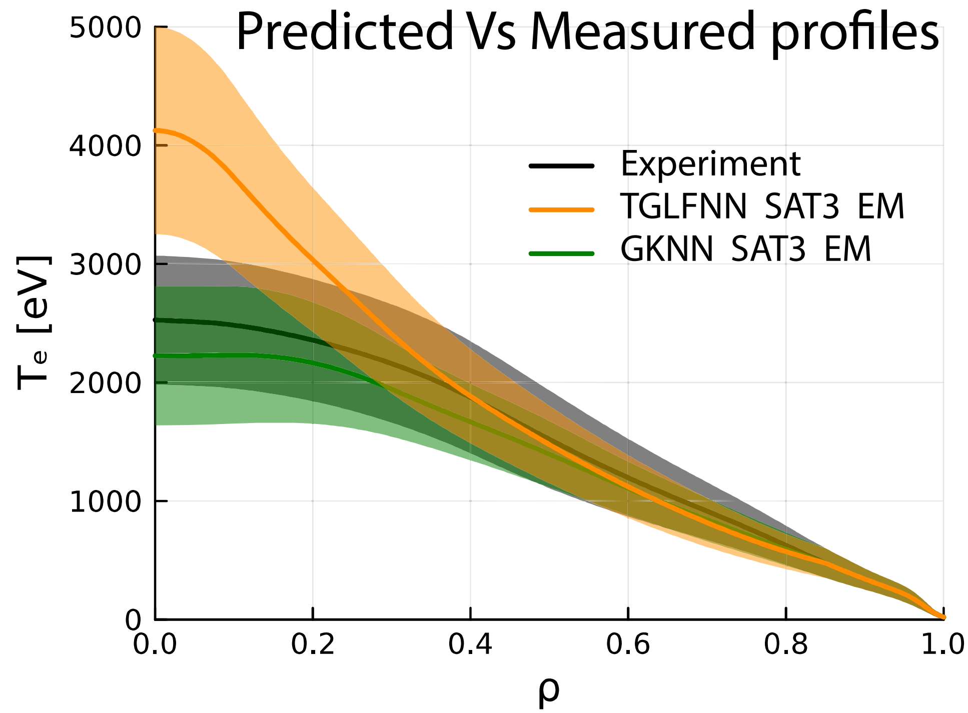




Leads to development of GKNN a new ML surrogate that corrects TGLFNN predictions to match results from QLGYRO

Statistical validation allows systematic tracking of model improvements and confident model selection





- Cross-machine validation of transport simulations
- Different regimes: L / H / Neg-D
- Experiment equilibria, HCD info
- Profiles validation: T_e, T_i, n_e, ω

TGLFNN GKNN	H-mode	L-mode	negD
DIII-D 4750 cases	SAT0Q EM SAT3 EM	SAT1 EM SAT3 EM	SAT1 EM SAT3 EM
MAST-U 1000 cases	SAT0Q EM SAT3 EM	SAT0 ES SAT3 EM	N/A
NSTX 750 cases	SAT3 EM SAT2 EM	SAT3 EM SAT2 EM	N/A

Statistical validation relies on databases of high-quality experimental data analyses

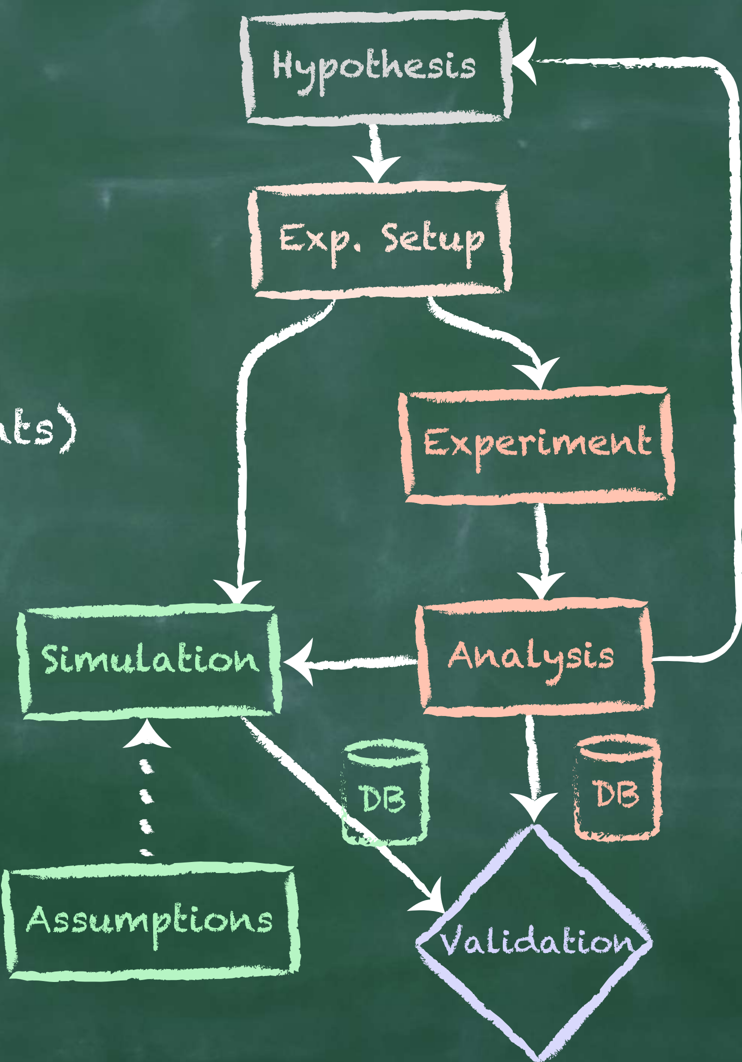
NOTE simulation inputs when doing validations:

1. Exp setup
   (eg. PCS trajectories)

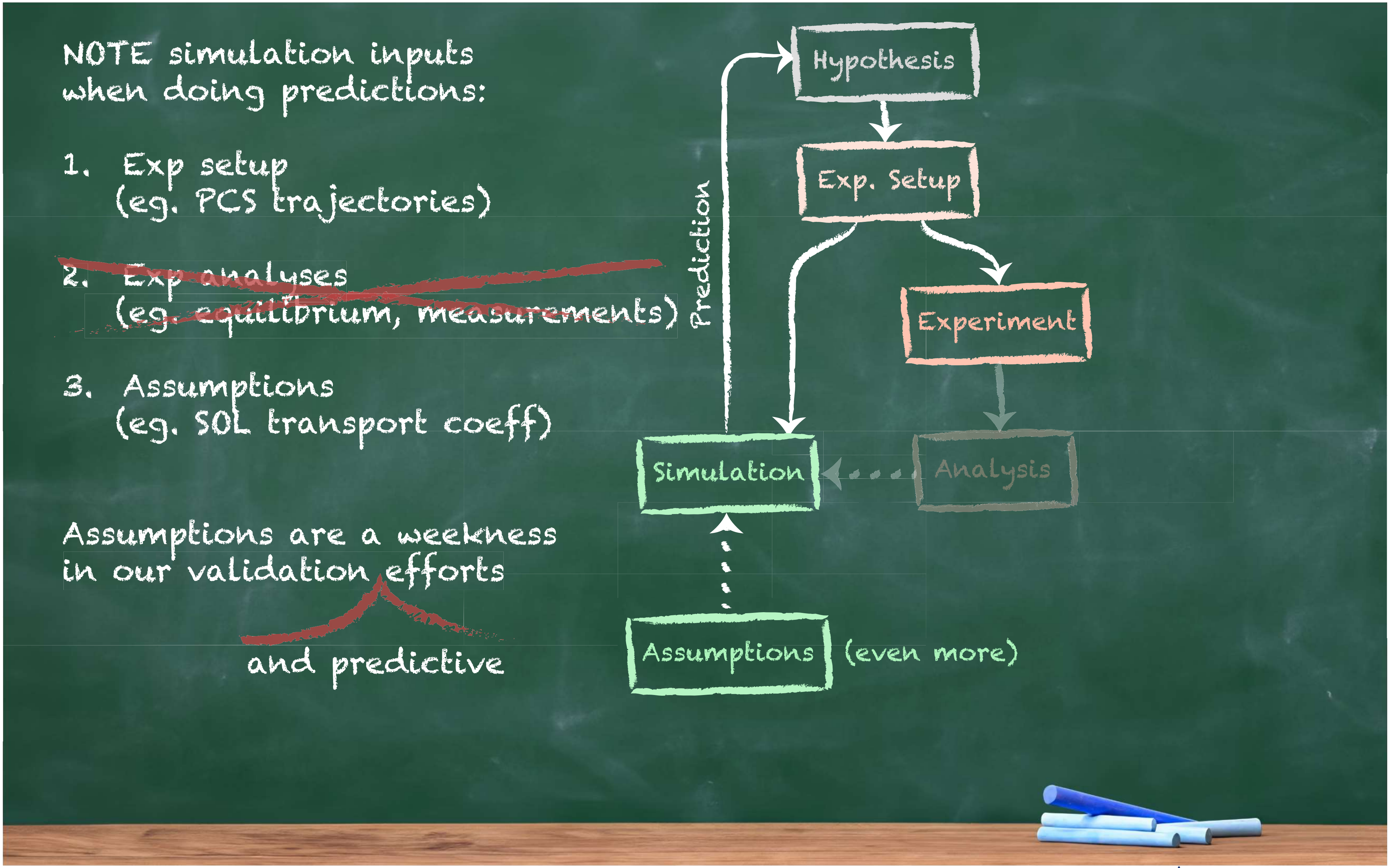
2. Exp analyses
   (eg. equilibrium, measurements)

3. Assumptions
   (eg. SOL transport coeff)

Assumptions are a weekness in our validation efforts

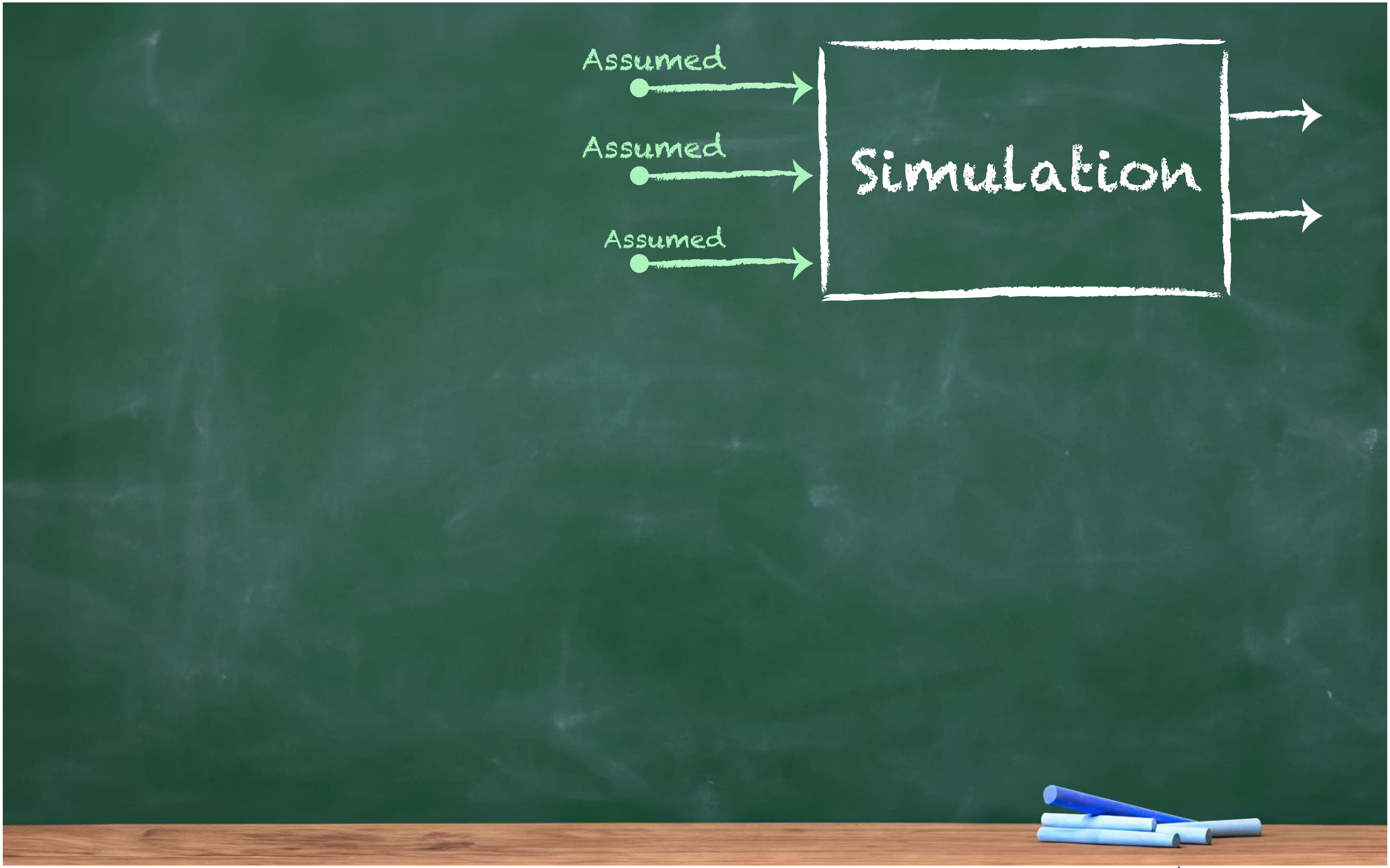


Statistical validation establishes trust needed to apply models predictively, to inform experiments and machine designs

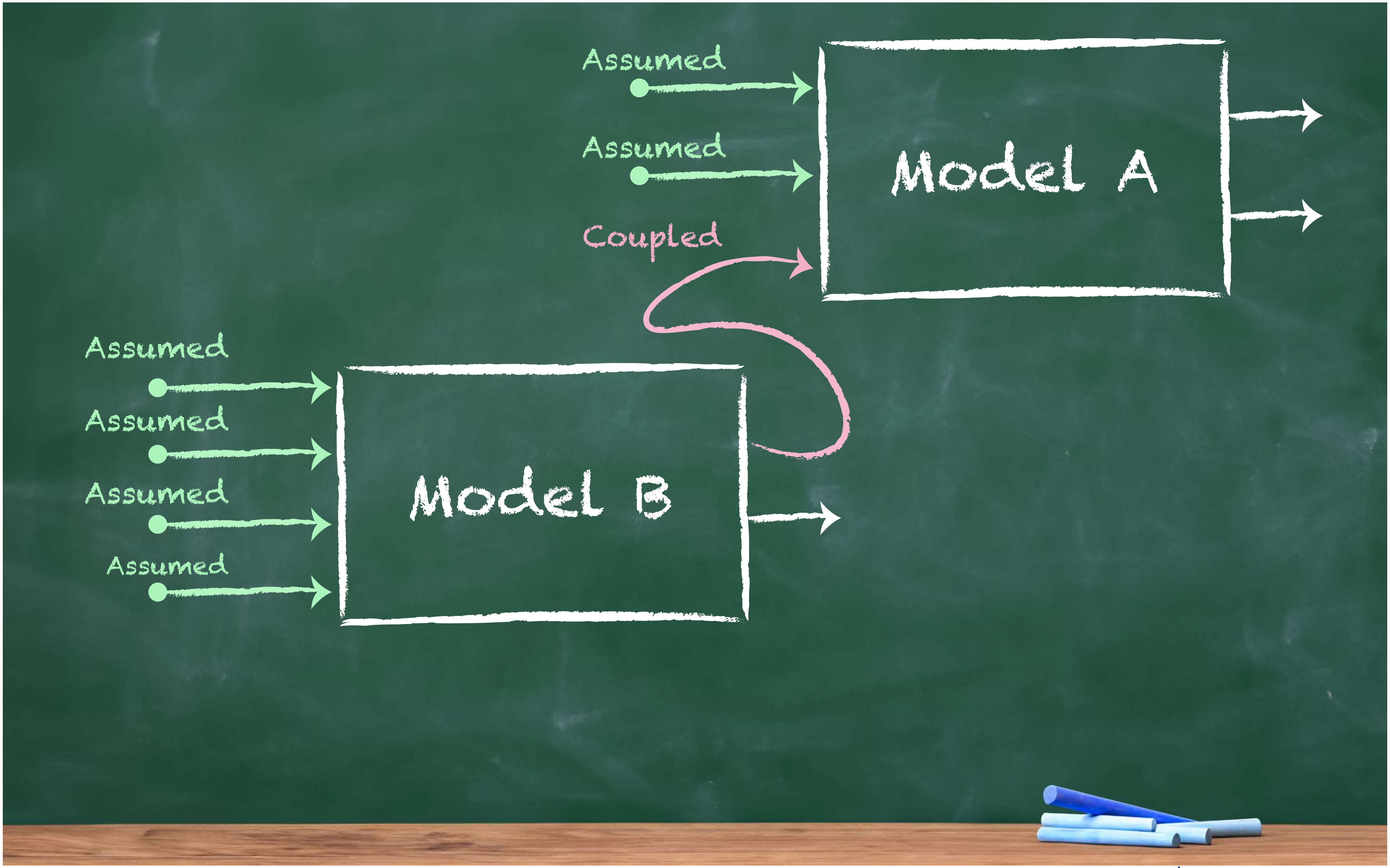


GENERAL ATOMICS

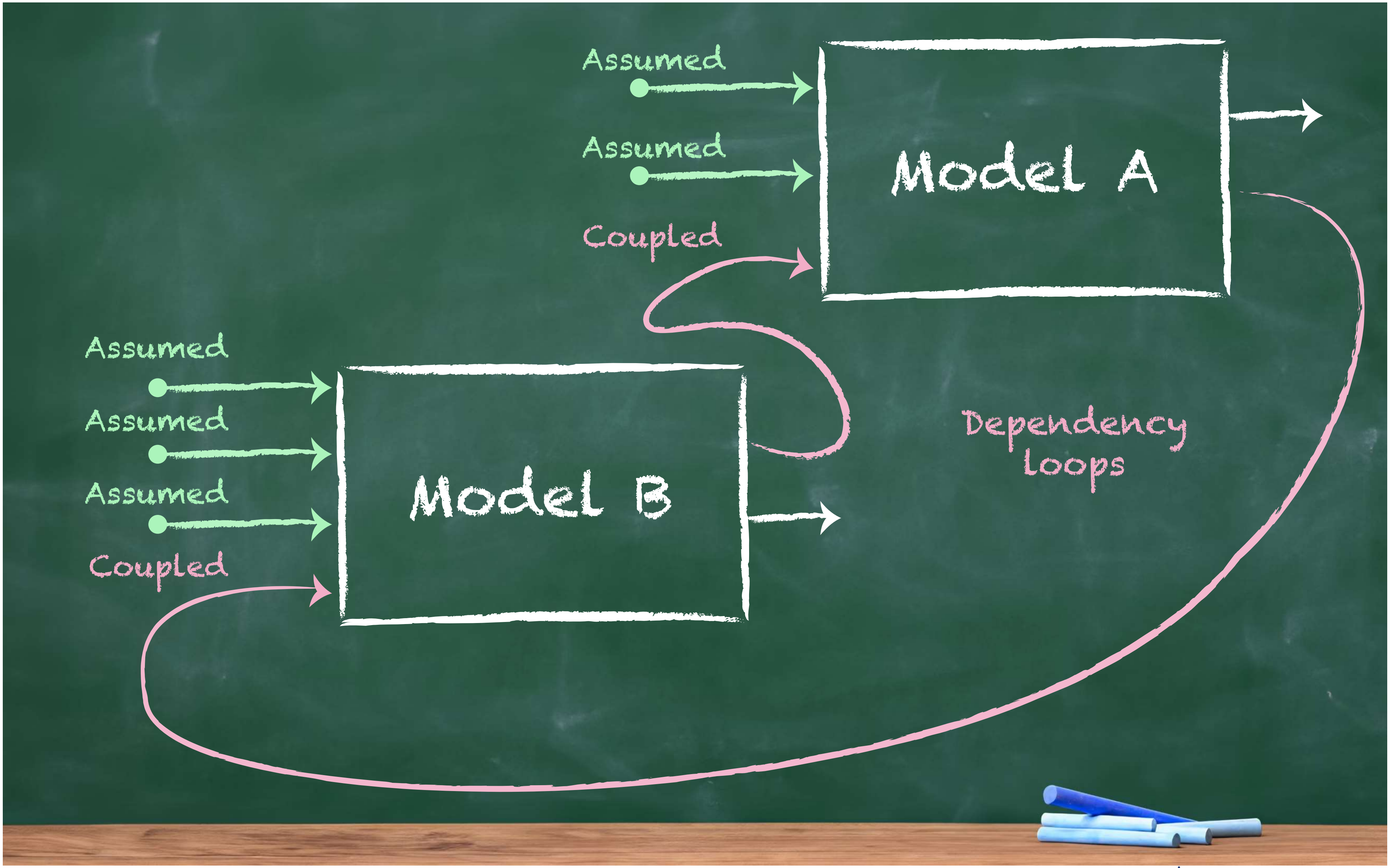
How do we get rid of input assumptions?



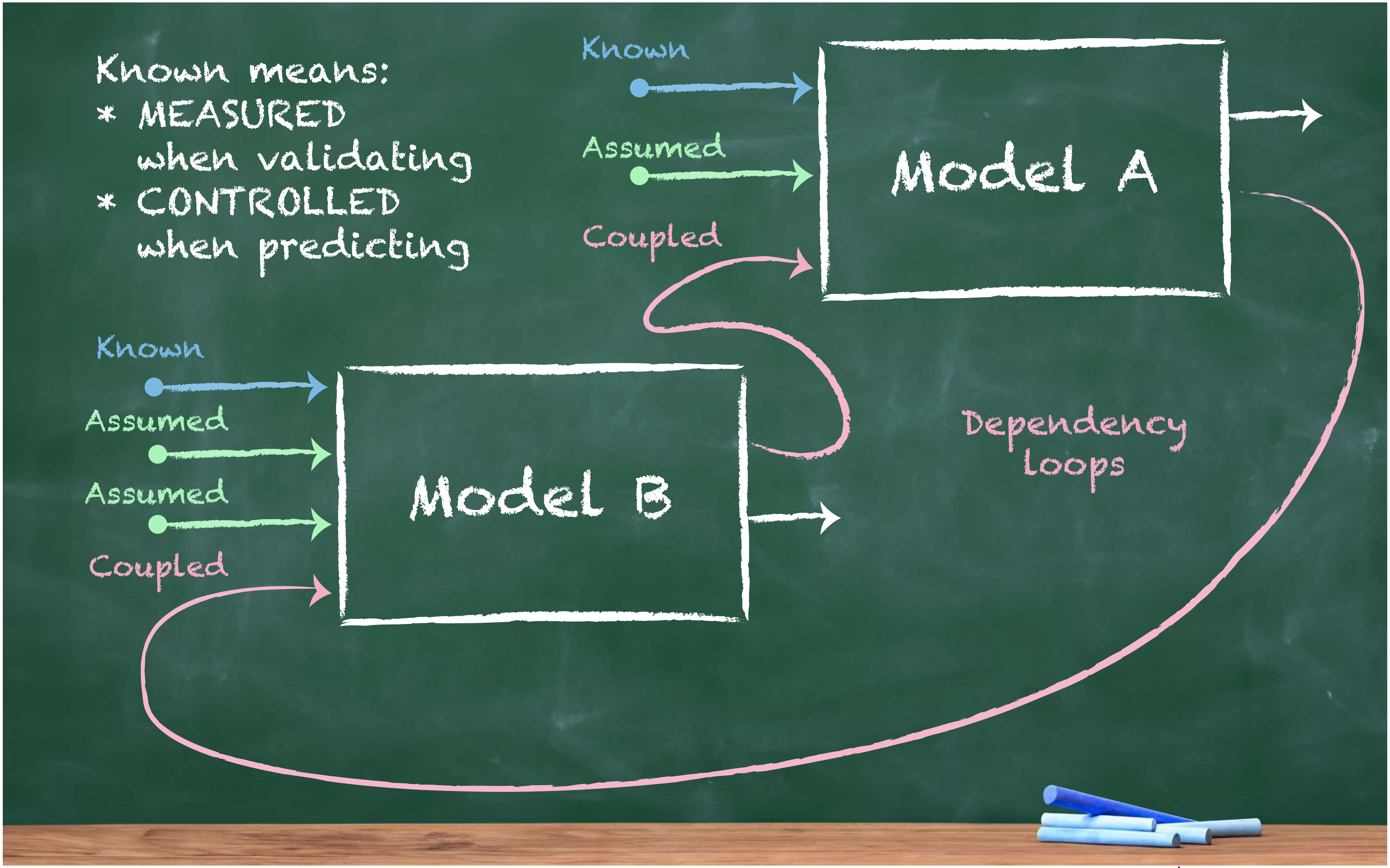
Couple another model ⇒ Why we do integrated modeling!!! Does coupling more codes always add value?



Couplings can lead to recursion, which need to be resolved with different algorithms



Some inputs are known...but what does it mean to know an input? When do we stop coupling codes?



Uncertainty quantification tells us where integration should stop

Quantities measured with sufficient accuracy

- **Stop when** you reach parameters that are well-characterized
- UQ quantifies whether current accuracy meets requirements

Controllable parameters

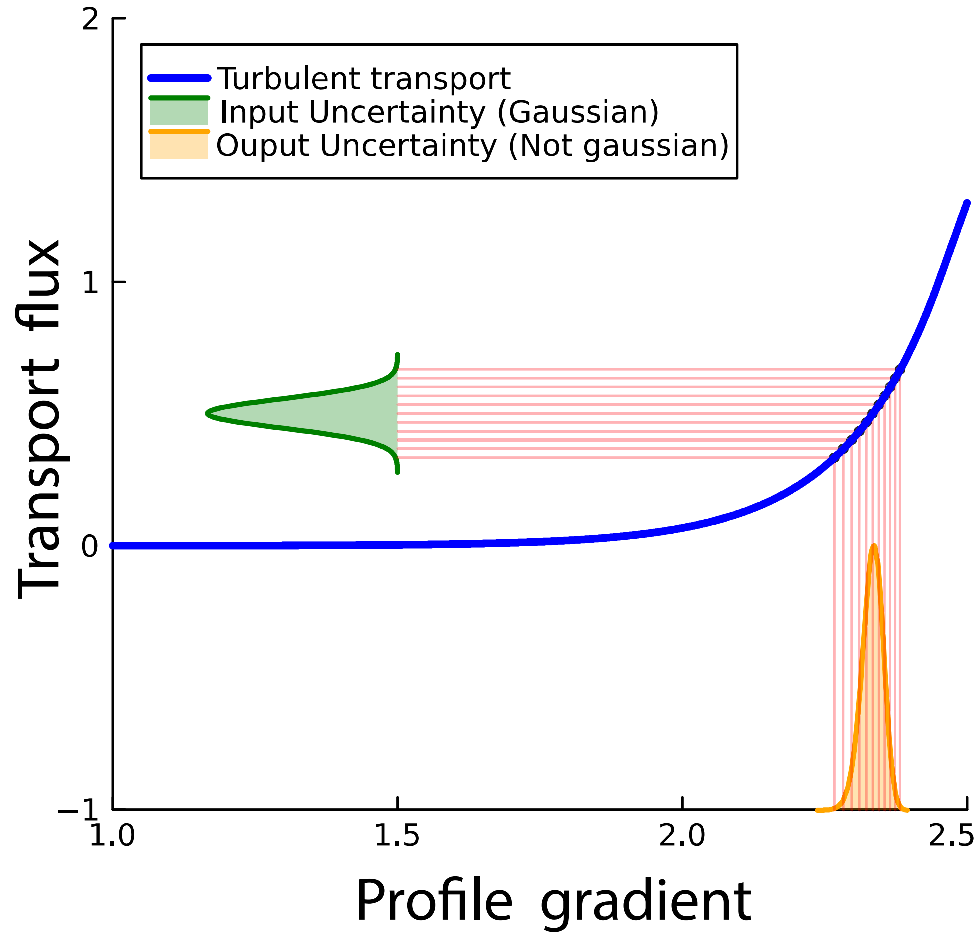
- **Stop when** model inputs that can be directly controlled
- Managing uncertainty becomes a design/control problem

Uncertainty-dominated regimes

- **Stop when** propagated uncertainties exceed model fidelity error
- Sensitivity analysis becomes more valuable than detailed modeling
- Focus should shift to establishing methods for robust control

Choice of uncertainty propagation method depends on model non-linearity over range of input uncertainty

- **Linear**
  - $f(x_0 + \epsilon) = f(x_0) + f'(x_0)\epsilon + \cancel{\frac{1}{2} f''(x_0)\epsilon^2} + \ldots$
    Analytic, finite differences, or automatic differentiation
  - Cheap: evaluate $f'(x_0)$
- **Non-linear**
  - Sampling-based methods
    Monte Carlo, unscented transform, chaos polynomials, distribution particles, ...
  - Expensive

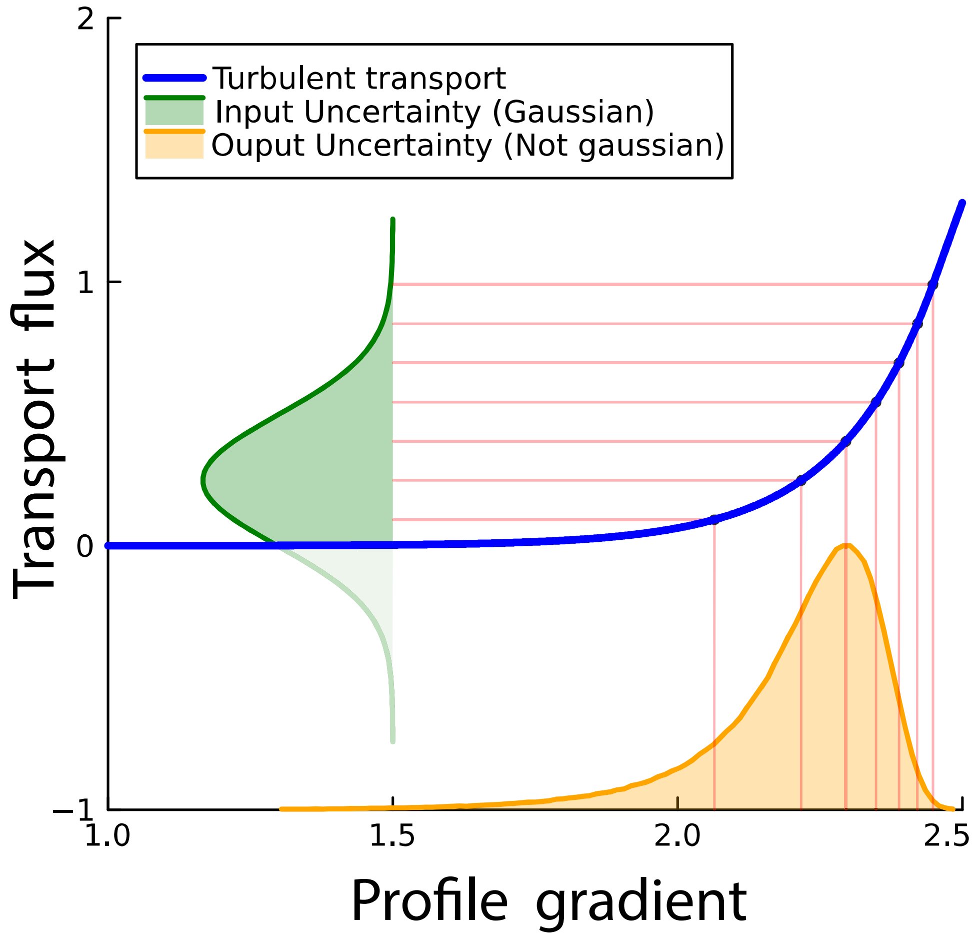


Practical solution: Lean on **ML surrogates**, which naturally support AD and are fast enough to support sampling methods

Choice of uncertainty propagation method depends on model non-linearity over range of input uncertainty

- **Linear**
  - $f(x_0+\epsilon) = f(x_0)+f'(x_0)\epsilon+\cancel{\frac{1}{2}f''(x_0)\epsilon^2}+\ldots$ Analytic, finite differences, or automatic differentiation
  - Cheap: evaluate $f'(x_0)$

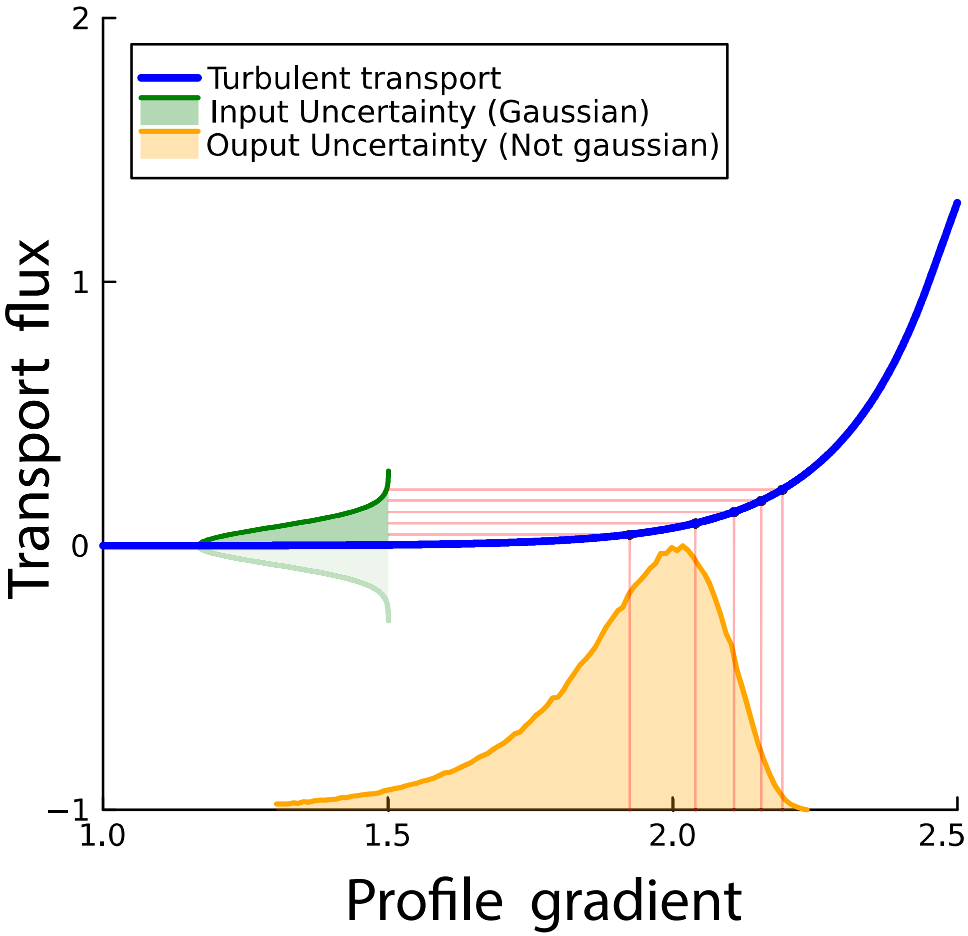
- **Non-linear**
  - Sampling-based methods Monte Carlo, unscented transform, chaos polynomials, distribution particles, ...
  - Expensive



Practical solution: Lean on **ML surrogates**, which naturally support AD and are fast enough to support sampling methods

Choice of uncertainty propagation method depends on model non-linearity over range of input uncertainty

- **Linear**
  - $f(x_0+\epsilon) = f(x_0) + f'(x_0)\epsilon + \cancel{\frac{1}{2}f''(x_0)\epsilon^2} + \ldots$
    Analytic, finite differences, or automatic differentiation
  - Cheap: evaluate $f'(x_0)$

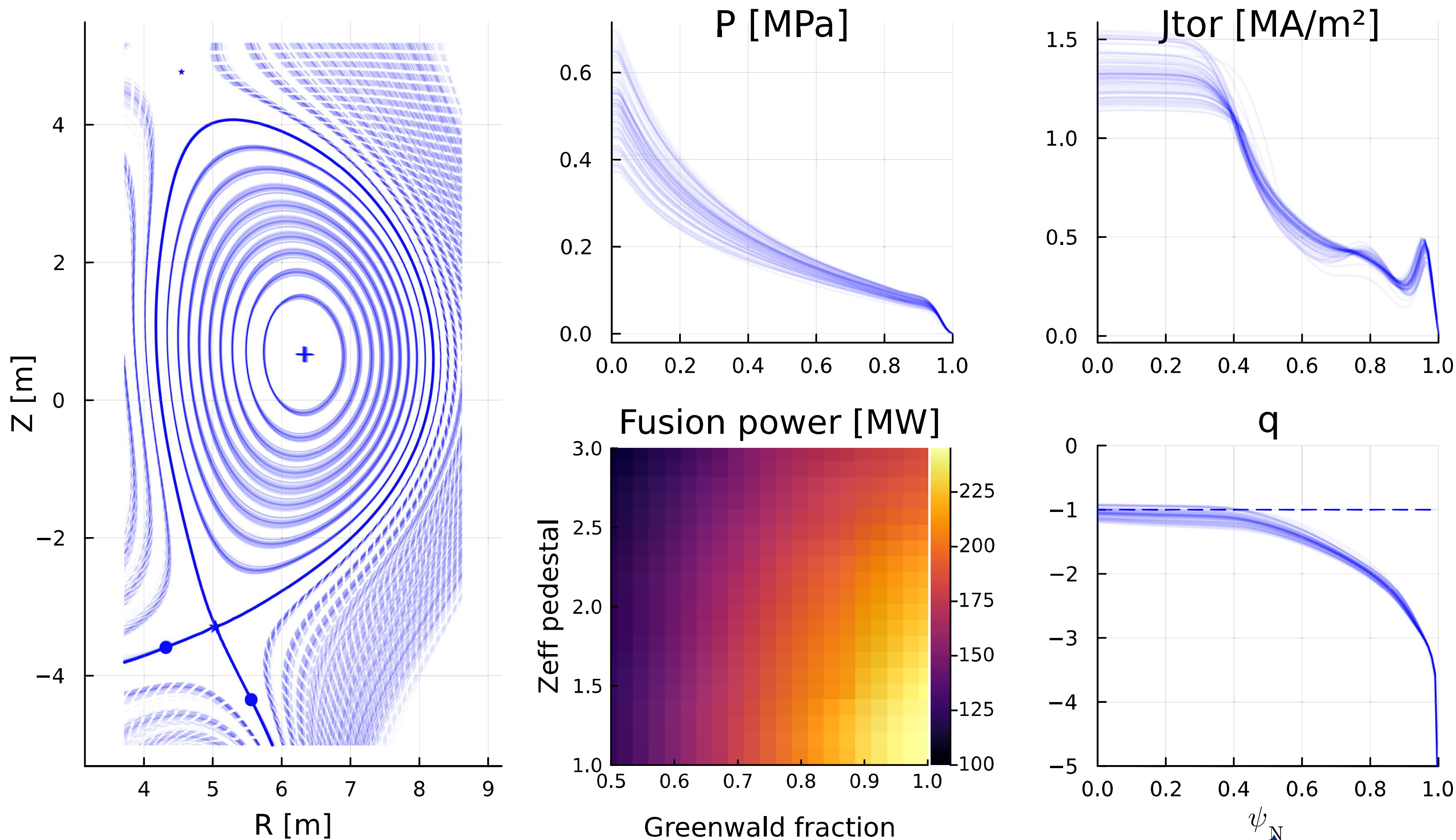


- **Non-linear**
  - Sampling-based methods
    Monte Carlo, unscented transform, chaos polynomials, distribution particles, ...
  - Expensive

Practical solution: Lean on **ML surrogates**, which naturally support AD and are fast enough to support sampling methods

When assumptions are made, we should use uncertainty propagation to evaluate their effects on the solution

Eg. ITER core-pedestal-equilibrium integration w/o SOL model,
but assuming $0.5 < f_{GW} < 1.0$ and $1.0 < Z_{\mathrm{eff,ped}} < 3.0$



GENERAL ATOMICS

Outline

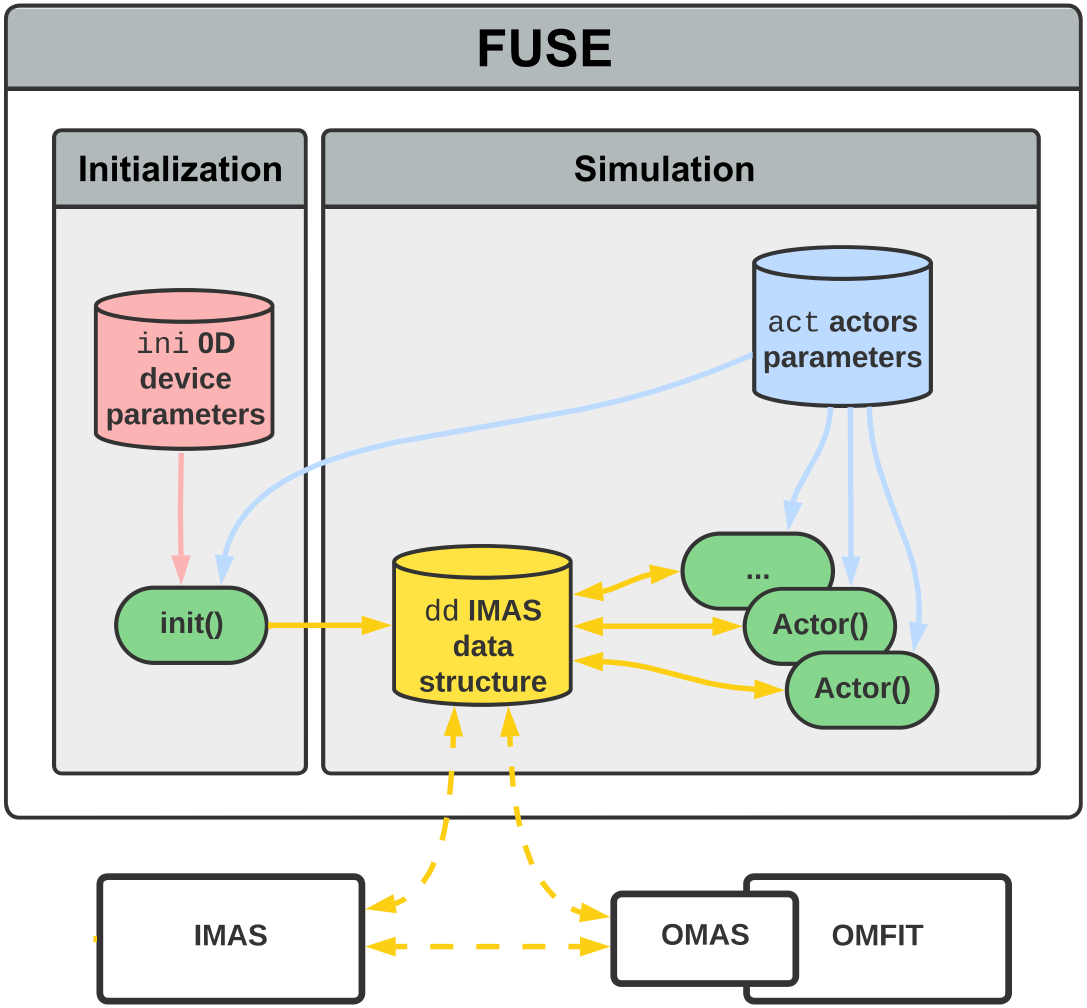
- Fundamentals of V&V and UQ
  - Validation and Verification
  - Fidelity hierarchy for V&V
  - Statistical validation
  - Uncertainty propagation
- FUSE: Implementing V&V and UQ in practice
  - Machine optimization
  - Integrated data analysis
  - Pulse design
- Key Takeaways

FUSE: Putting V&V and UQ principles into practice

- Born in 2021 to support nascent **FPP design** industry
- Applying **lessons learned** from GA modeling expertise
  OMFIT, OMAS, STEP, TGYRO, TGLF-NN, EPED-NN, EFIT-AI, TokSys, GASC, ...
- Built from scratch, all in one language: **Julia**
  - High-level like Python
  - As fast as C
  - Auto-differentiable
- Uses **ITER IMAS** ontology

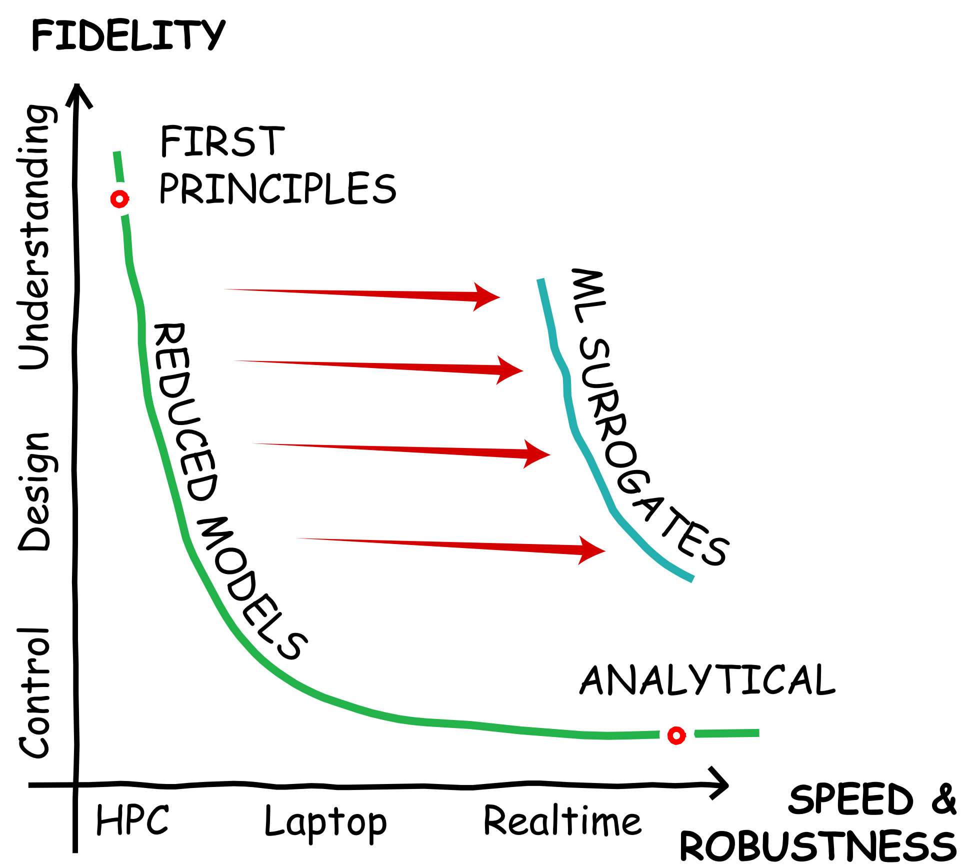
Few (strict) design rules enable high degree of modularity

1. All data is stored in a centralized `dd` data structure (IMAS based)
2. Actors only talk via `dd`
3. Actor functionality set by `act` parameters
4. `dd` can be initialized from 0D `ini` parameters
5. FUSE interfaces to outside world only via `dd`



Whole fidelity spectrum is supported, but strategically balance fidelity with speed and use ML when advantageous

- Want to **capture realistic system dynamics**
  1. Whenever possible, use of **theory-based** (reduced) models
  2. **Sufficient fidelity** to capture critical interface physics between subsystems, so in-depth high-fidelity studies do not upend couplings
- While enabling **rapid design iterations**
  1. **Julia** for high performance
  2. **Tightly coupling** of models
  3. Break efficiency-fidelity tradeoff with **ML surrogates**



How to integrate models:

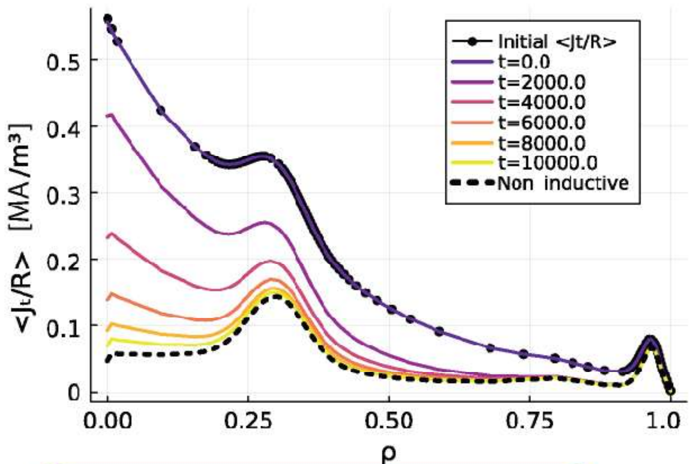
- (re)-write in Julia (preferred)
- In memory coupling
- File-based (last resort)

GENERAL ATOMICS

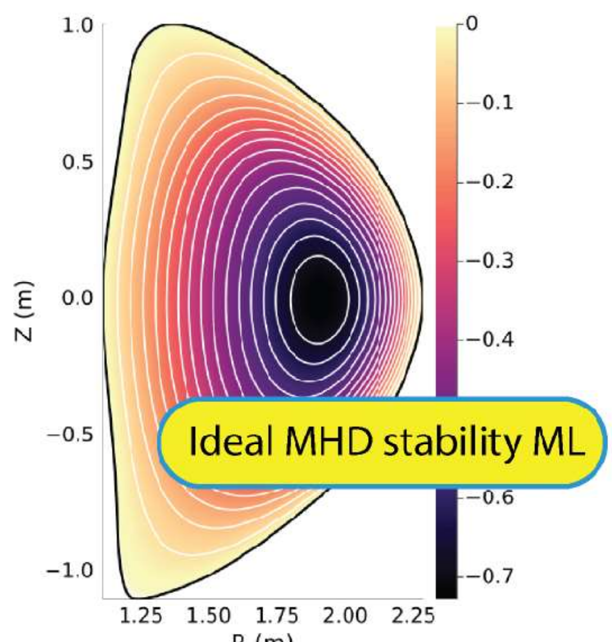
FUSE models span from the plasma core to the site boundary

FUSE models span from the plasma core to the site boundary

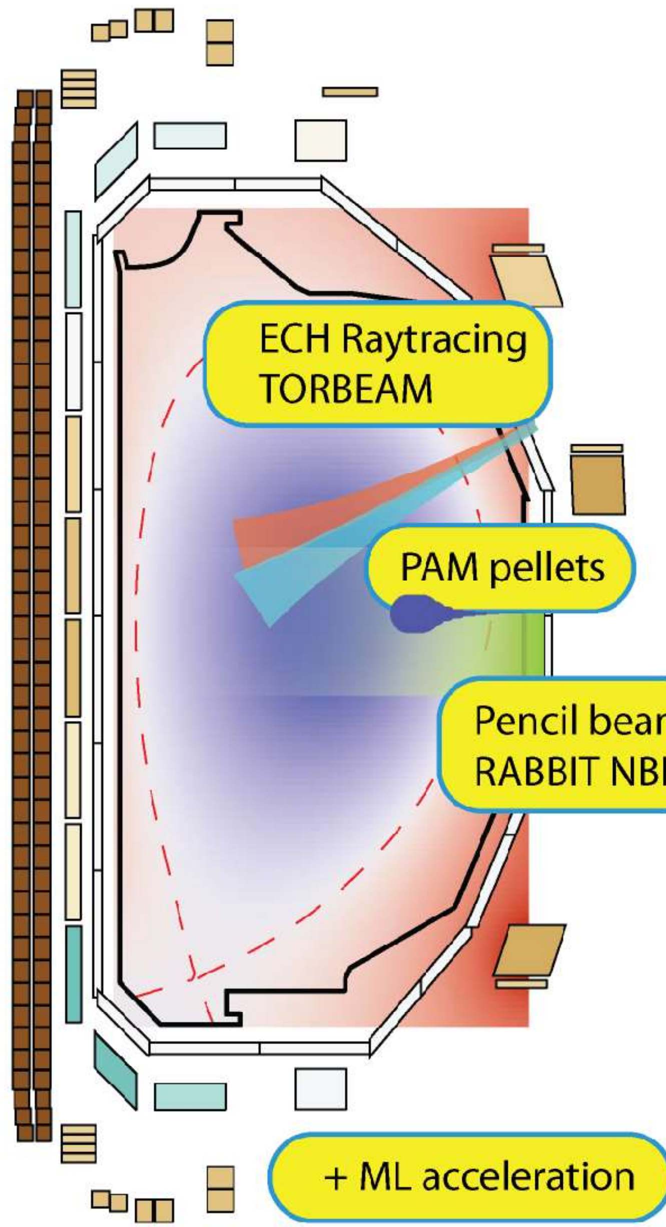
Current evolution w/ sawteeth



Fixed boundary equilibrium



Ideal MHD stability ML



ECH Raytracing TORBEAM

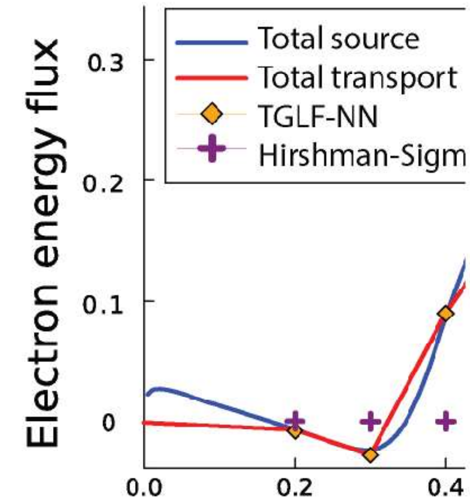
PAM pellets

Pencil beam RABBIT NBI

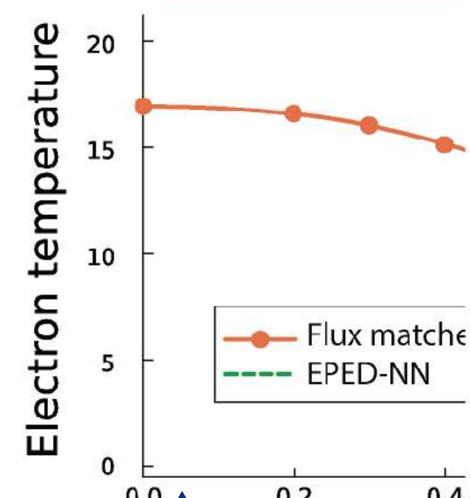
+ ML acceleration

Free boundary equilibrium w/ coil + passive couplings

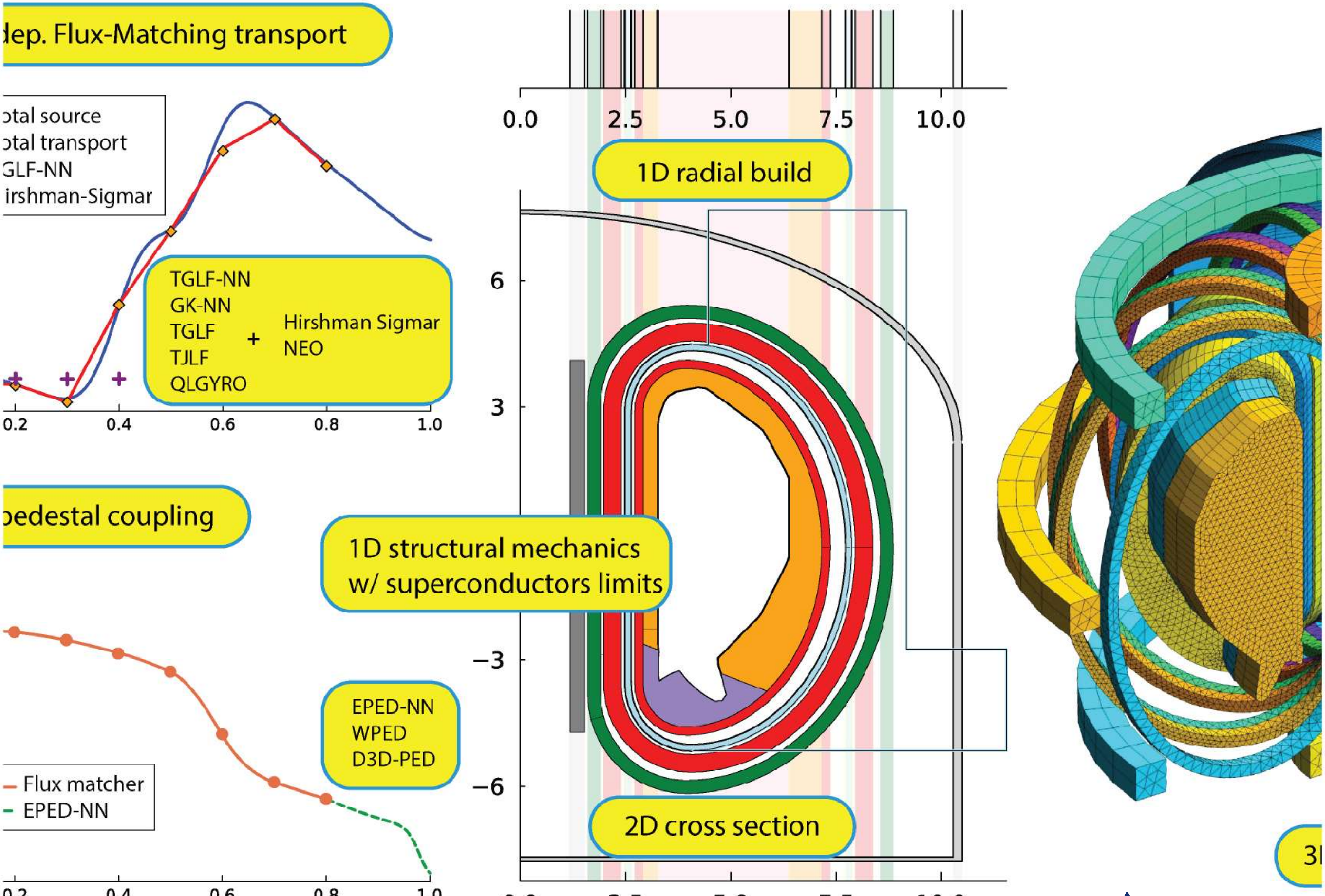
Time dep. Flux-M



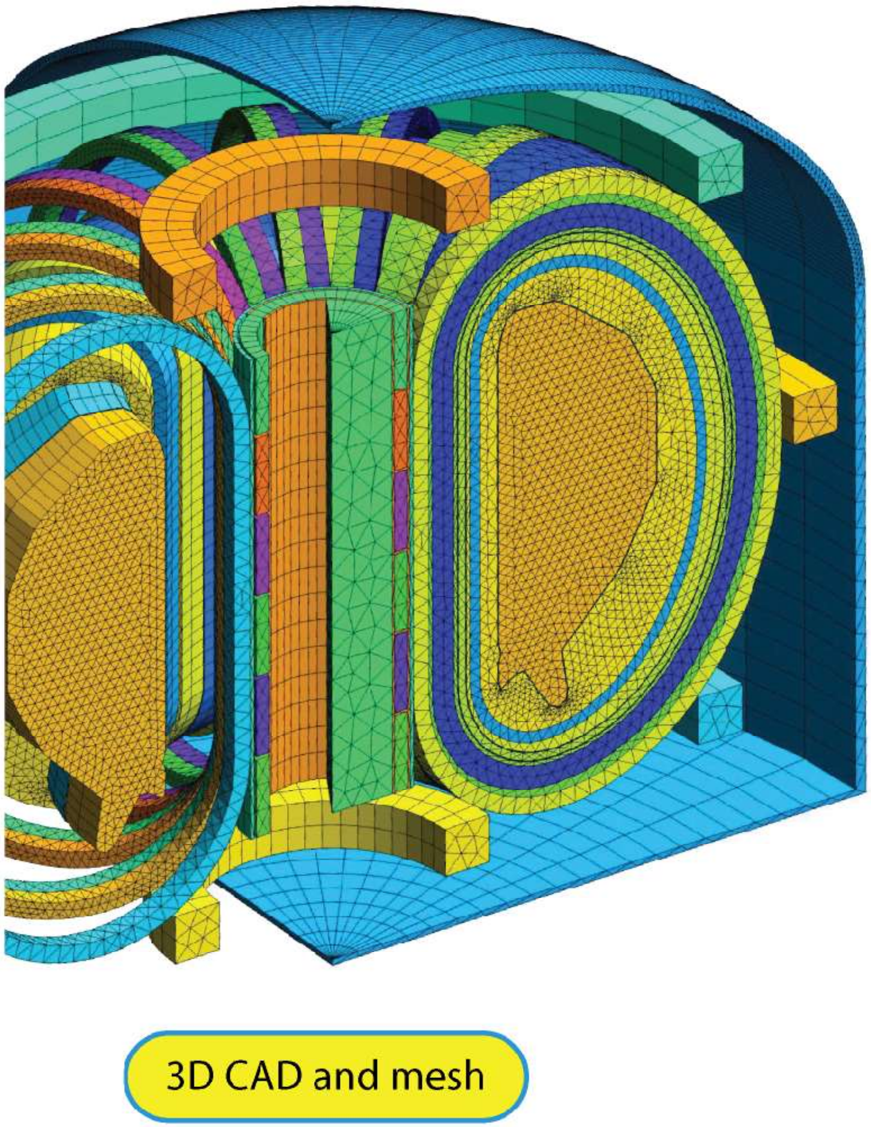
Core-pedestal co

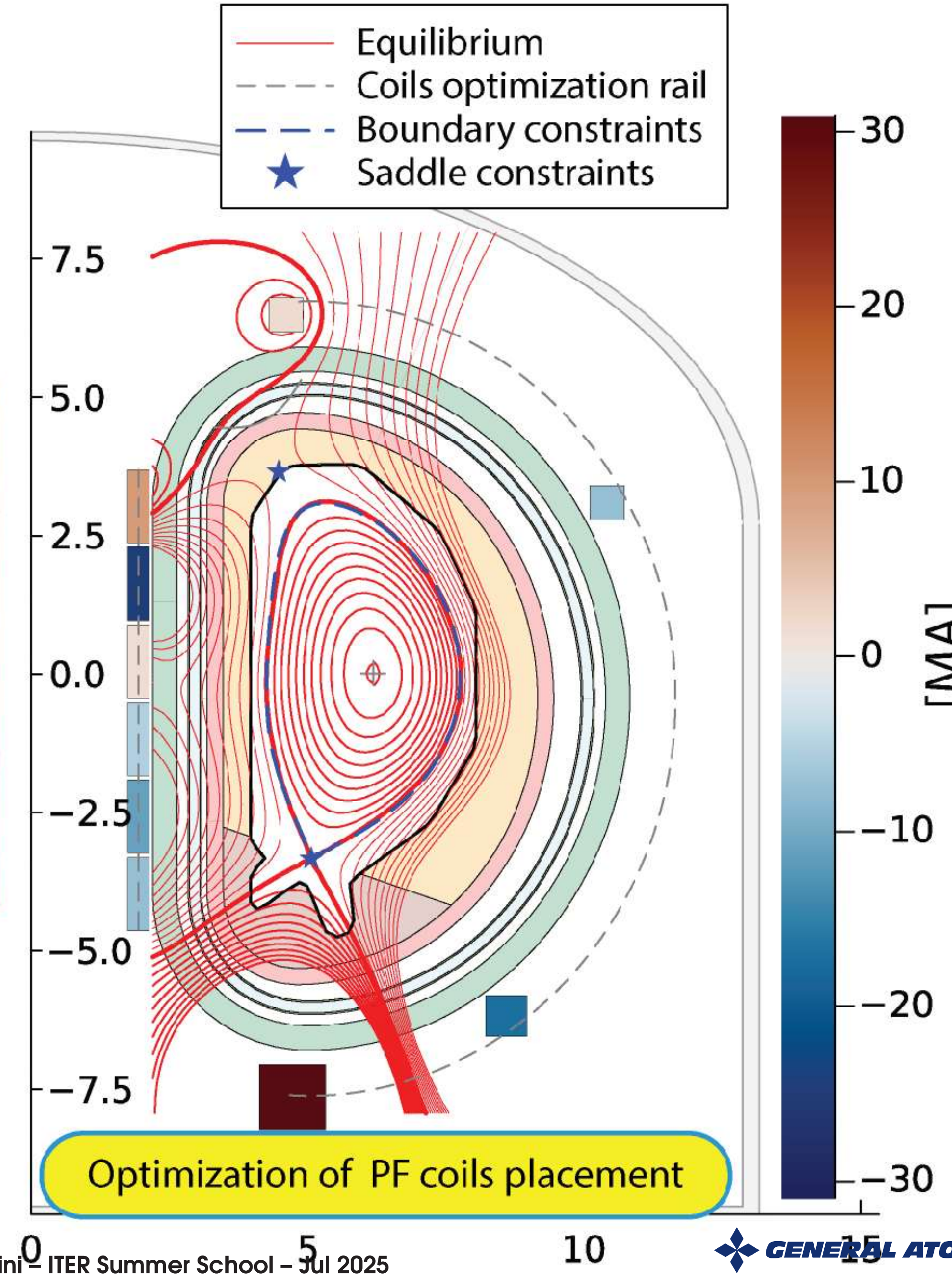


FUSE models span from the plasma core to the site boundary

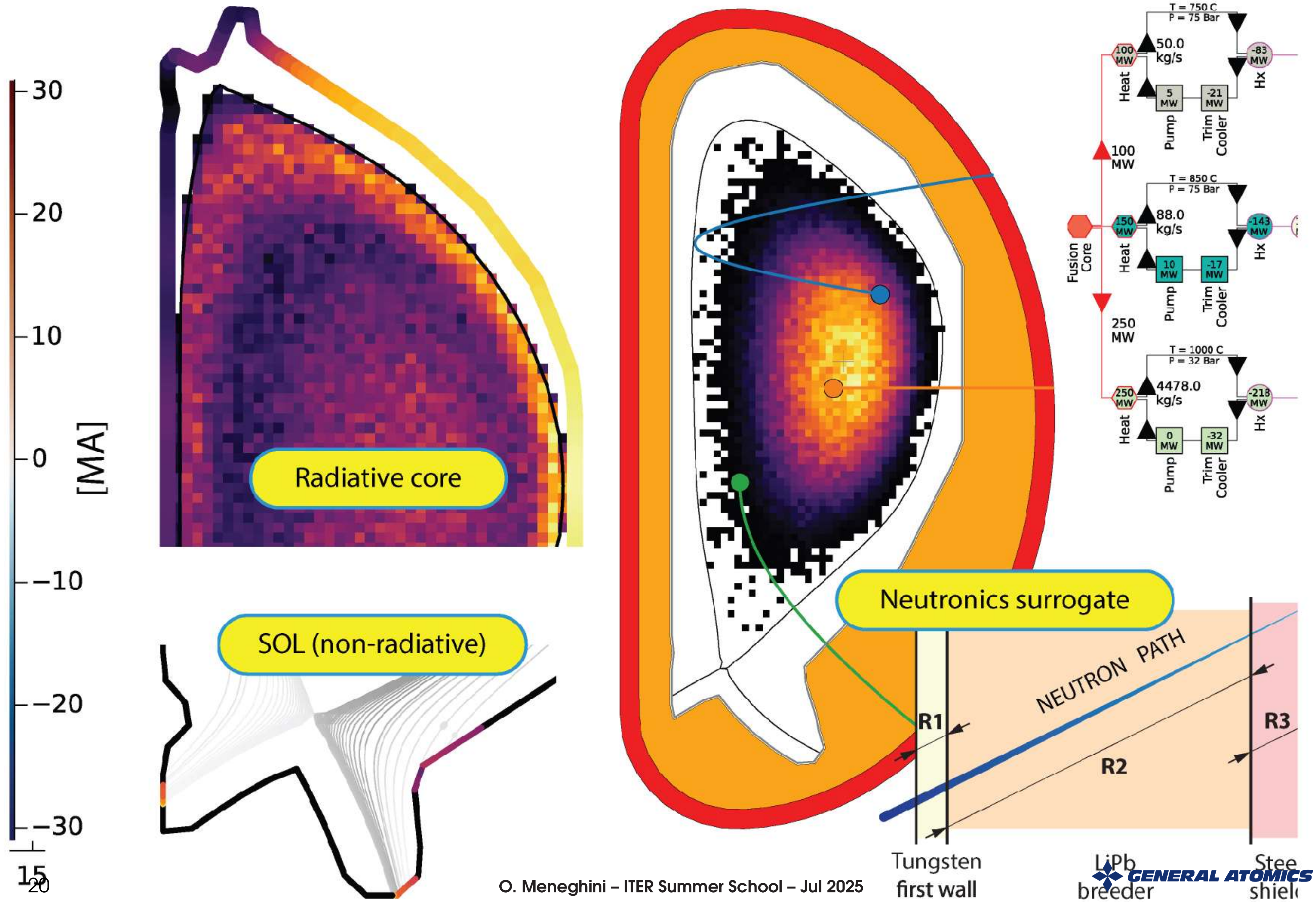


FUSE models span from the plasma core to the site boundary



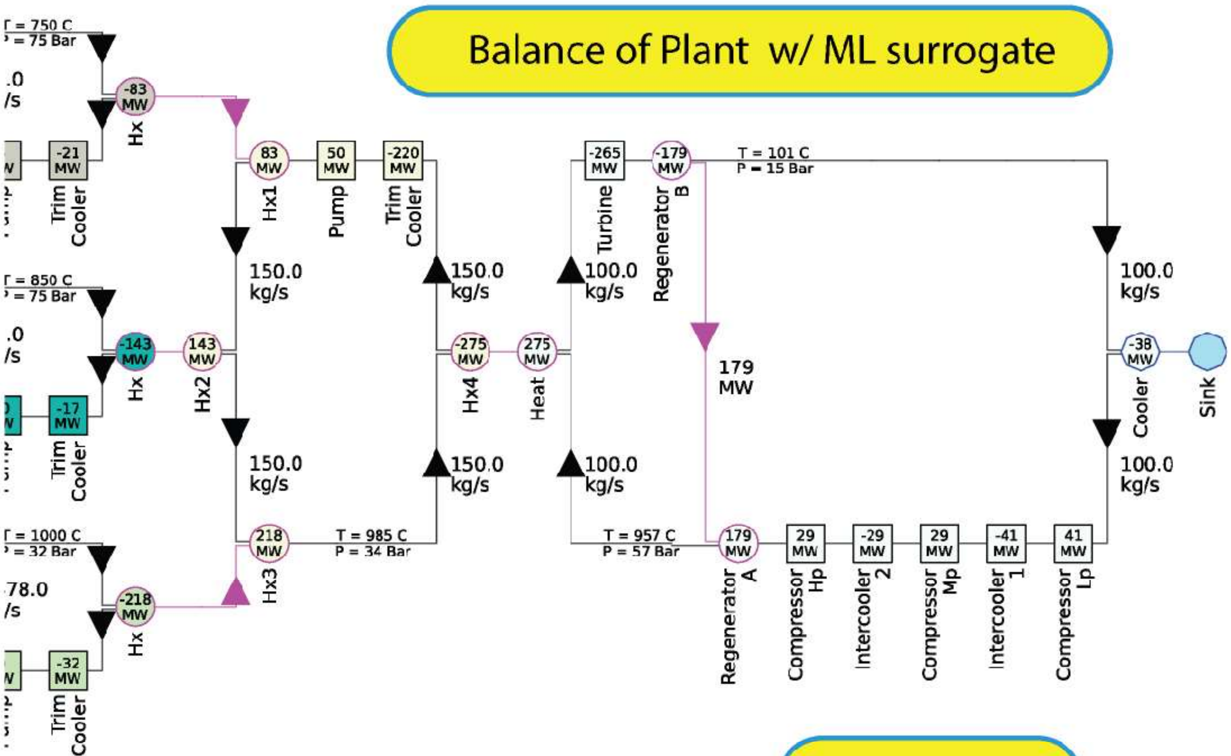


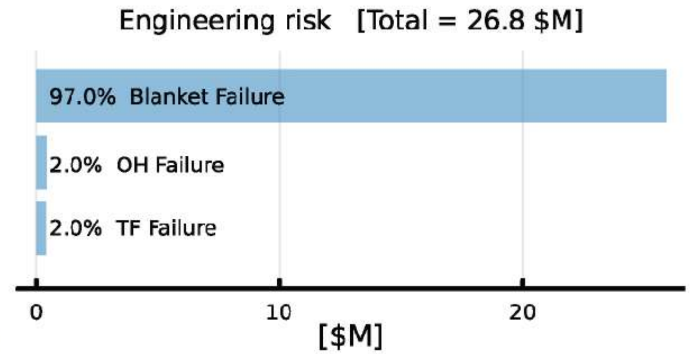
FUSE models span from the plasma core to the site boundary

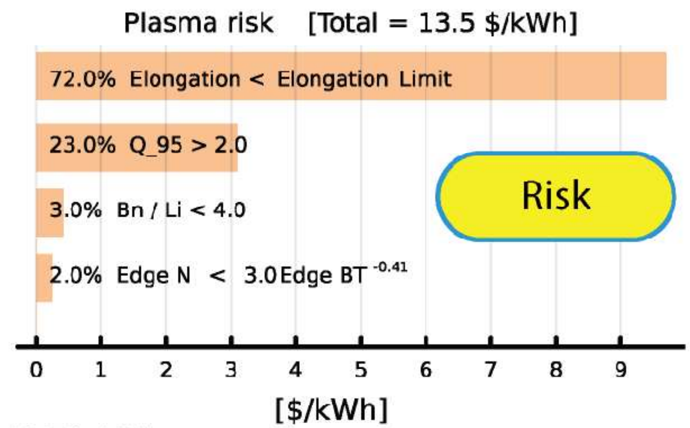


GENERAL ATOMICS

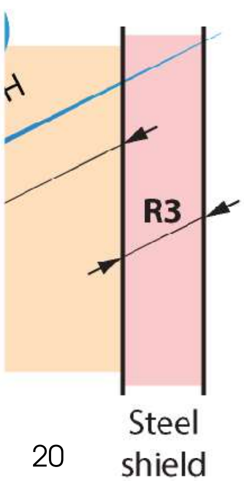
FUSE models span from the plasma core to the site boundary

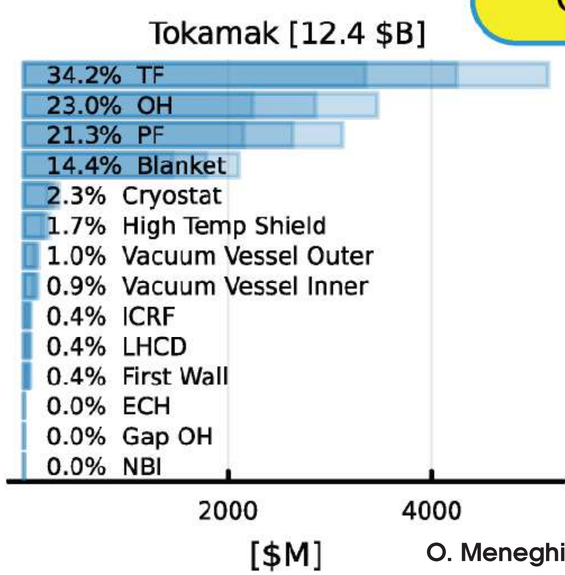


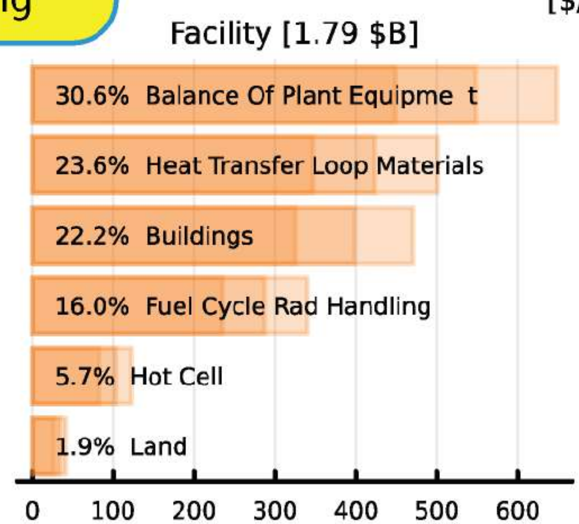




Costing

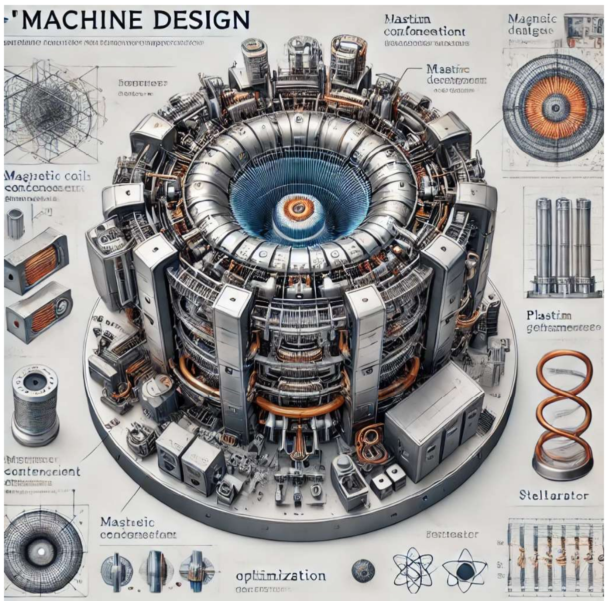


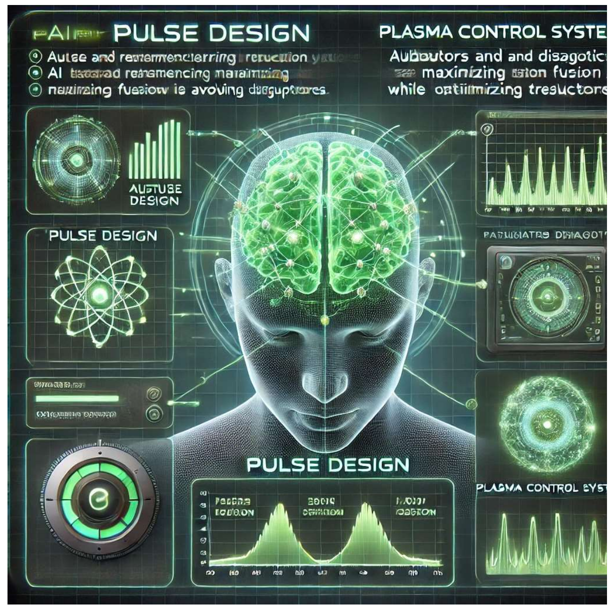


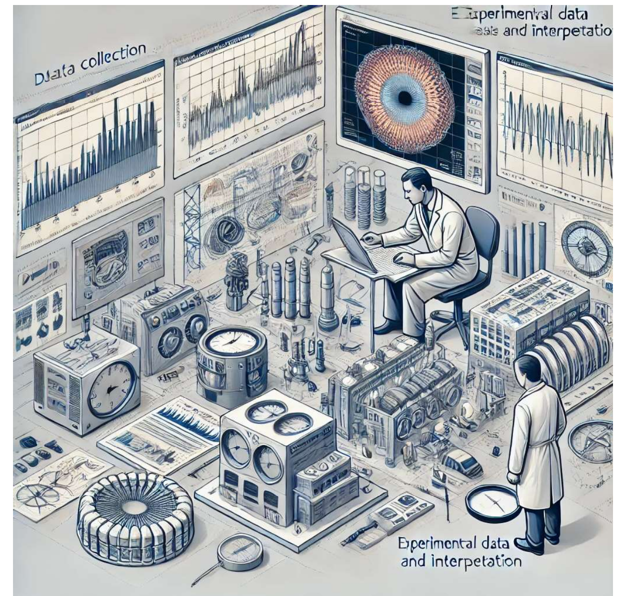


FUSE is designed to support all three main applications of integrated modeling

1 **MACHINE DESIGN**	2 **PULSE DESIGN**	3 **DATA ANALYSIS**





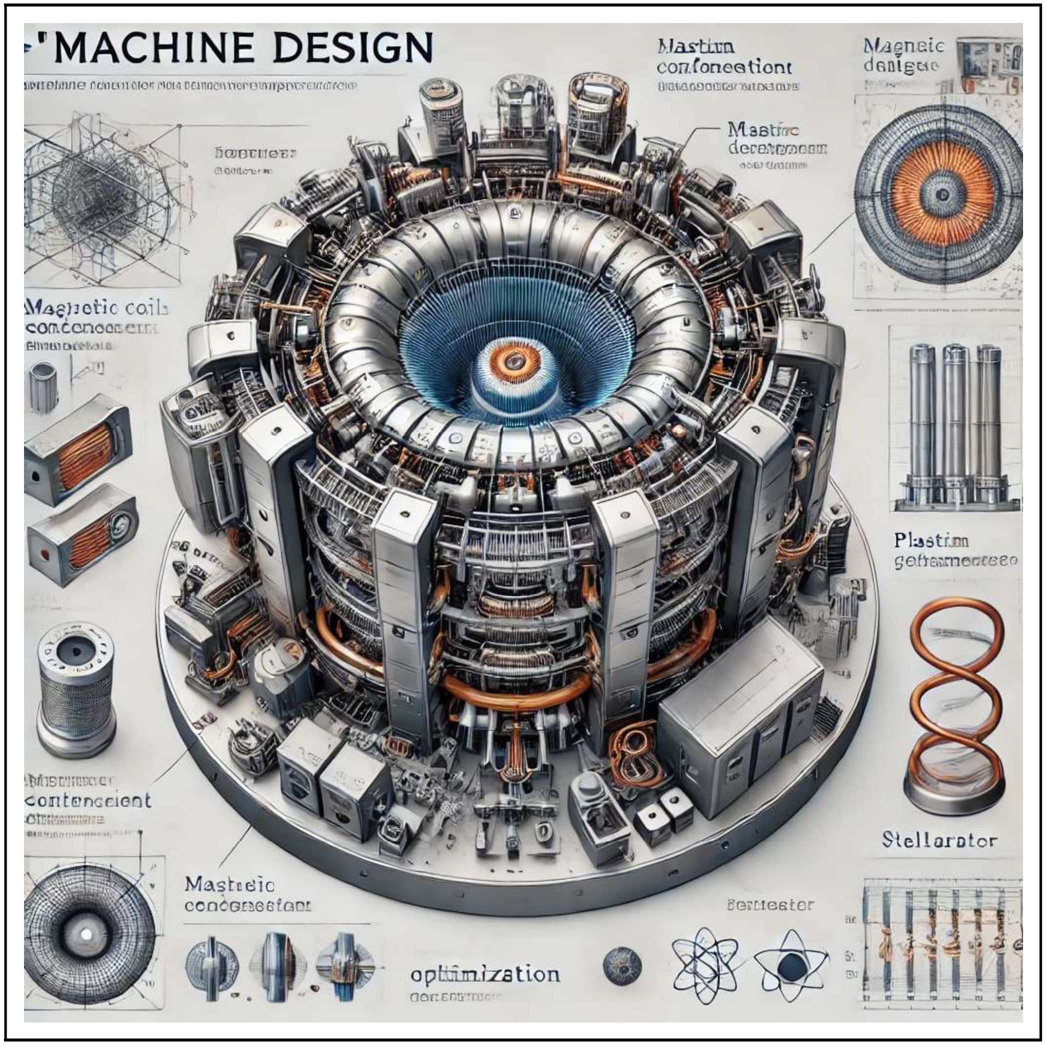


- Same theory-based models
- Same act./diag. models
- Same machine-agnosticity
- Same integrated workflows
- Same data structures
- Same need for speed, always!

There's a great potential to exploit these synergies!

GOAL: From idea to pre-conceptual designs in minutes and evaluate wildly different concepts on same footing

1) MACHINE DESIGN



FUSE uses a multi-objective constrained optimization workflow to enable design explorations and trade studies

OBJECTIVES

1. min capital cost
2. max q_{95}

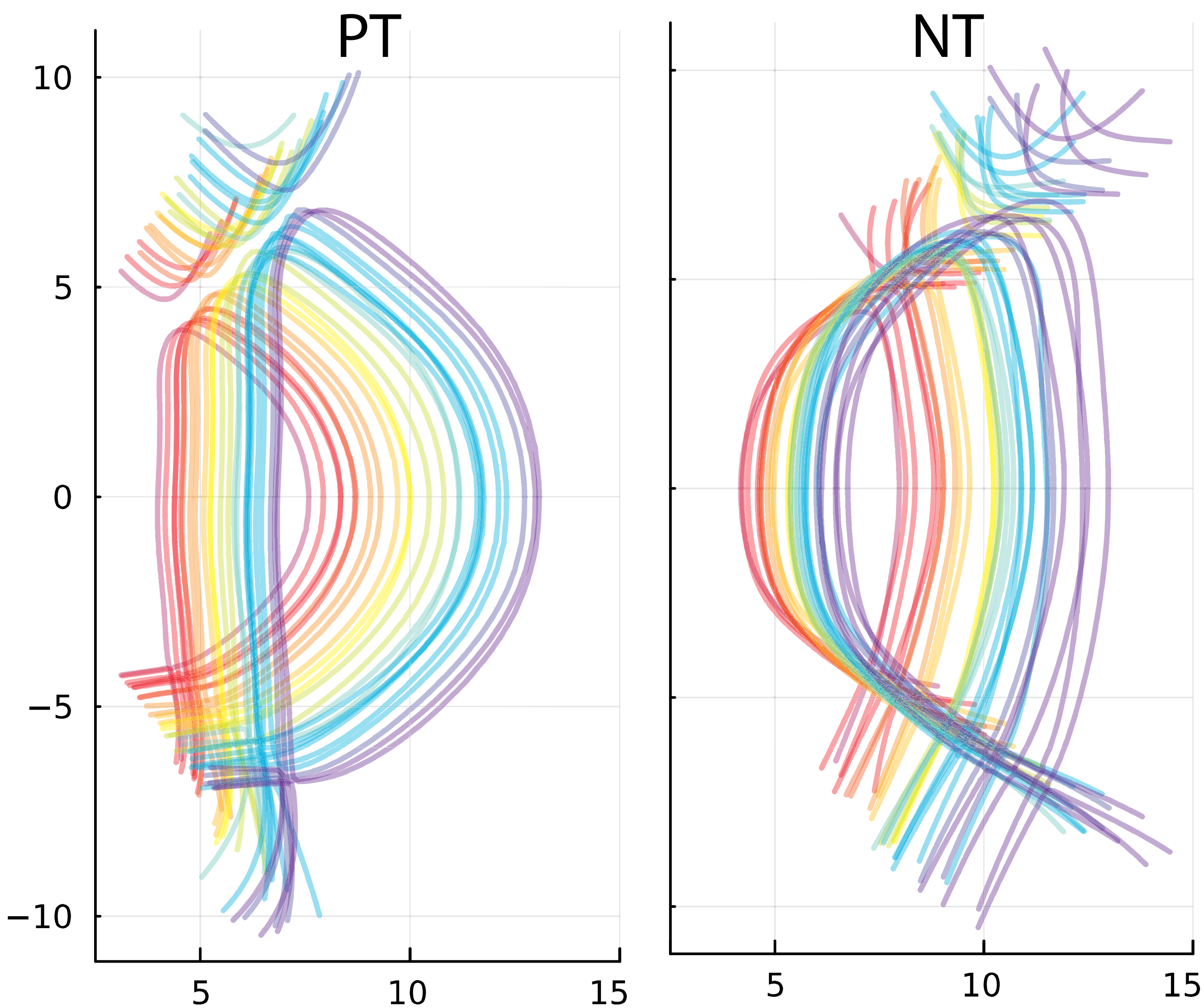
CONSTRAINTS

- $P_{\text{electric}} = 250 \pm 50$ MW
- flattop $= 1.0 \pm 0.1$ (h)
- TBR $= 1.1 \pm 0.1$
- $P_{\text{sol}}/P_{\text{LH}} > 1.1$ (for $+\delta$)
- $P_{\text{sol}}/R < 15$ (MW/m)

ACTUATORS

- $5.0 < R_0 < 10.0$ (m)
- $3.0 < B_0 < 15.0$ (T)
- $4.0 < I_p < 22$ (MA)
- $1.5 < \kappa < 2.2$
- $|\delta| < 0.7$
- $1.1 < z_{\text{eff,ped}} < 3.5$
- $0.4 < f_{\text{GW,ped}} < 0.85$
- Impurity: Ne, Ar, Kr
- $0 < P_{\text{EC}} < 100$ (MW)
- $0 < \rho_{\text{EC}} < 0.9$
- $0 < P_{\text{NB}} < 50$ (MW)

eg. Trade study for positive-δ VS negative-δ FPP

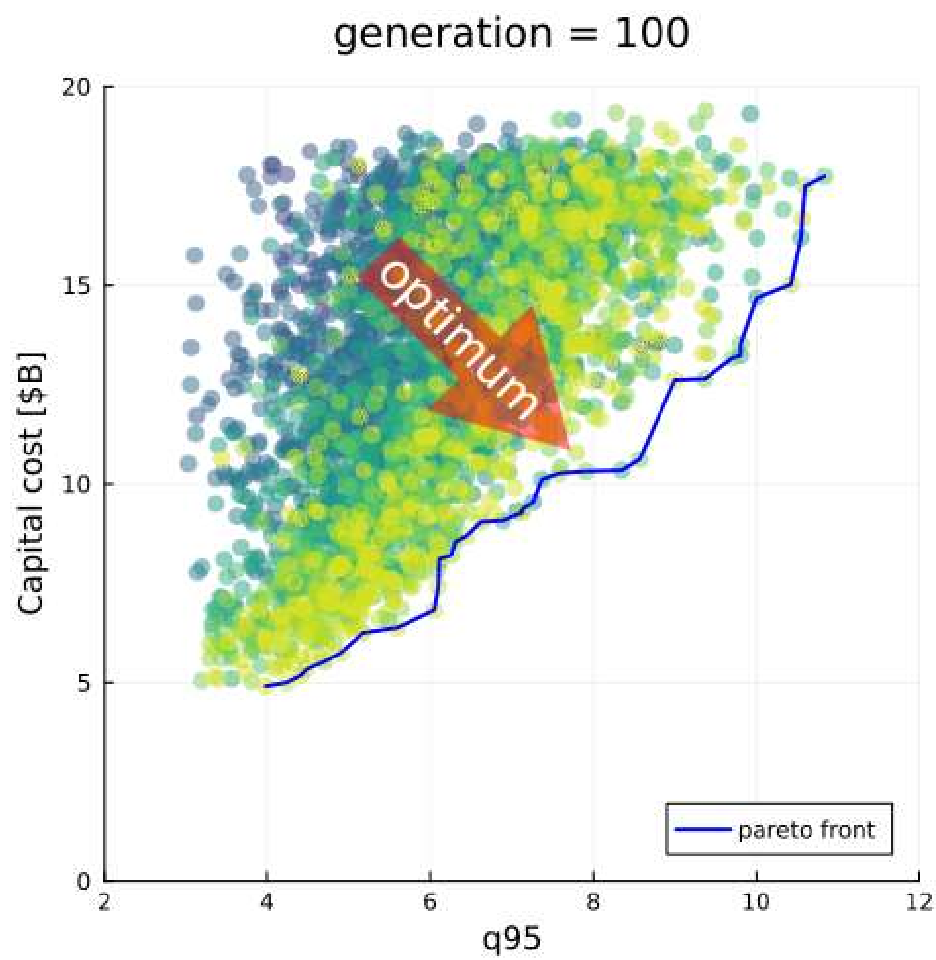


GENERAL ATOMICS

Multi-objective constrained optimization workflow enables designs exploration and trade studies

A **Genetic algorithm** steers solution towards the **Pareto front**

- **Each point is a full machine design** that takes $\sim$ 1 min to run
- Highlights **complex system dynamics** and exposes **objectives trade-offs**
- Helps different stakeholders **identify a target design**
  (scientists, investors, policymakers,...)
- **Scalable parallel execution** runs 10k+ cases in few hours on small cluster

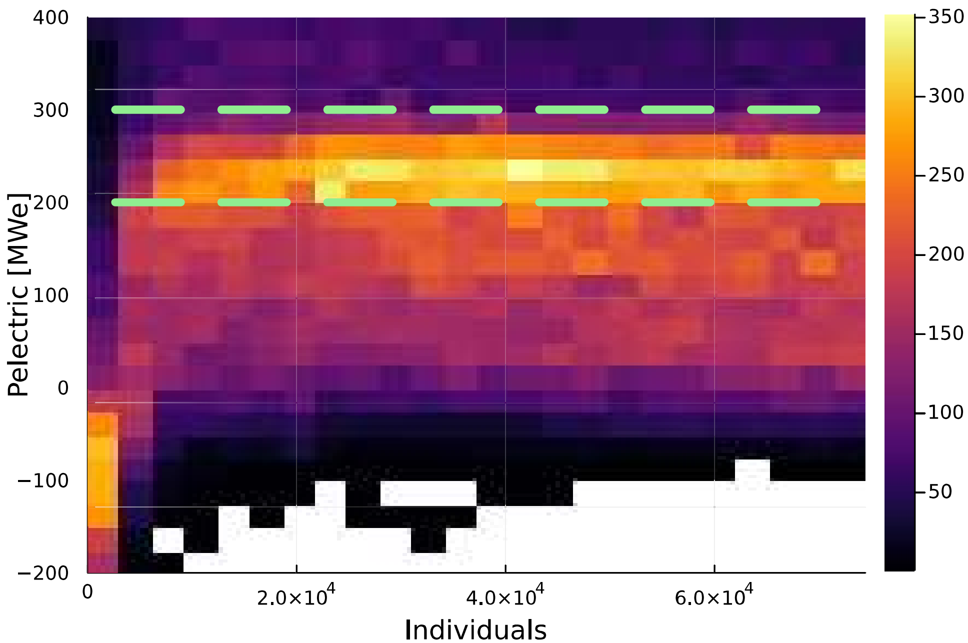


Multi-objective constrained optimization workflow enables designs exploration and trade studies

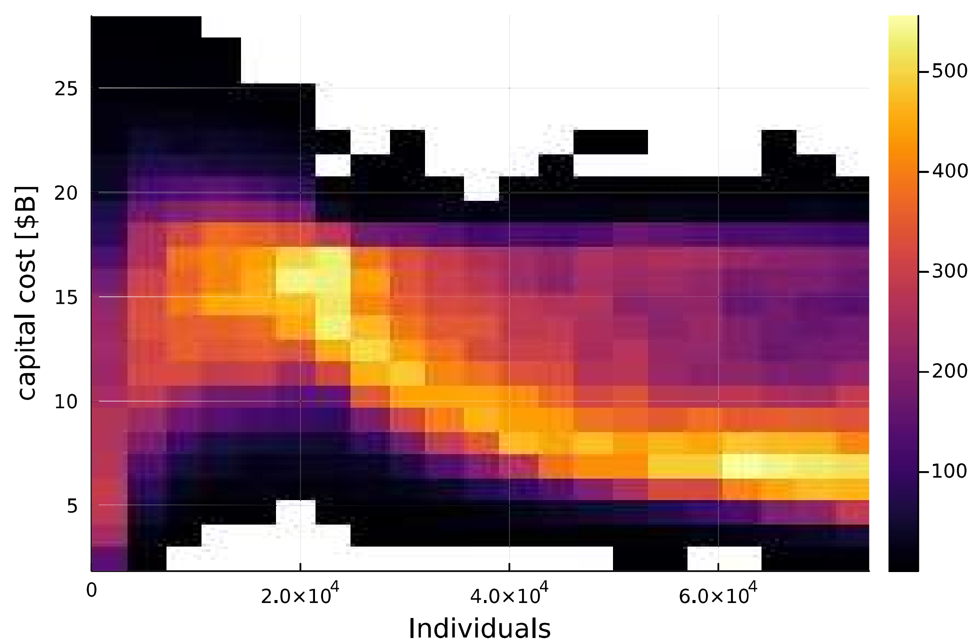
A **Genetic algorithm** steers solution towards the **Pareto front**

It takes 10's of thousands of full designs to find optimal solutions that satisfies the constraints. Eg:

Power generation constraint

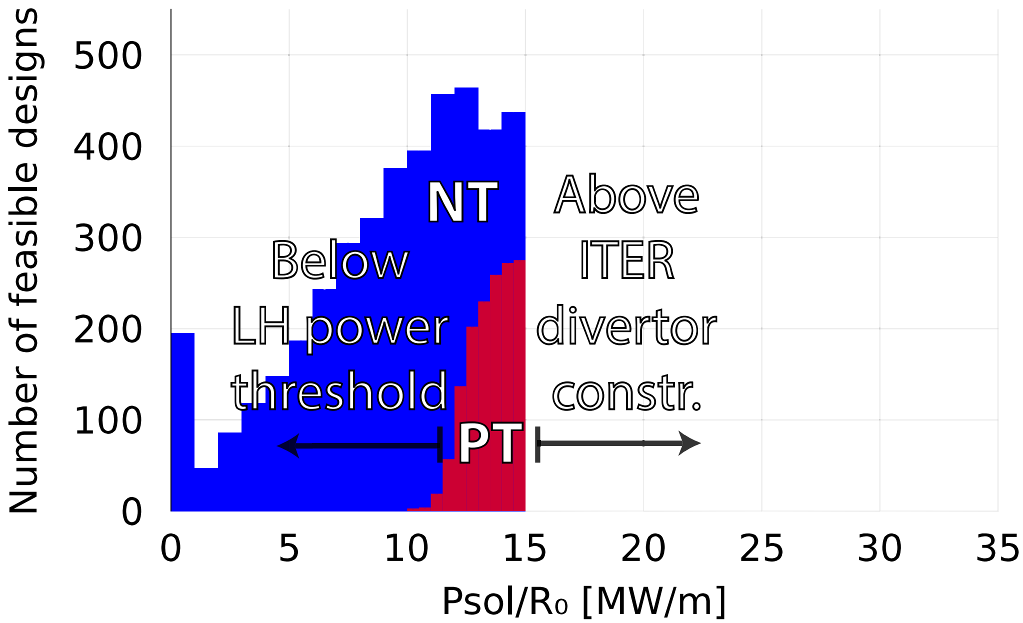


Minimum cost objective



Accuracy, speed, scalability, and robustness are all key

We can use optimization datasets to relate uncertainties in key parameters to cost. Relation to risk = cost × probability

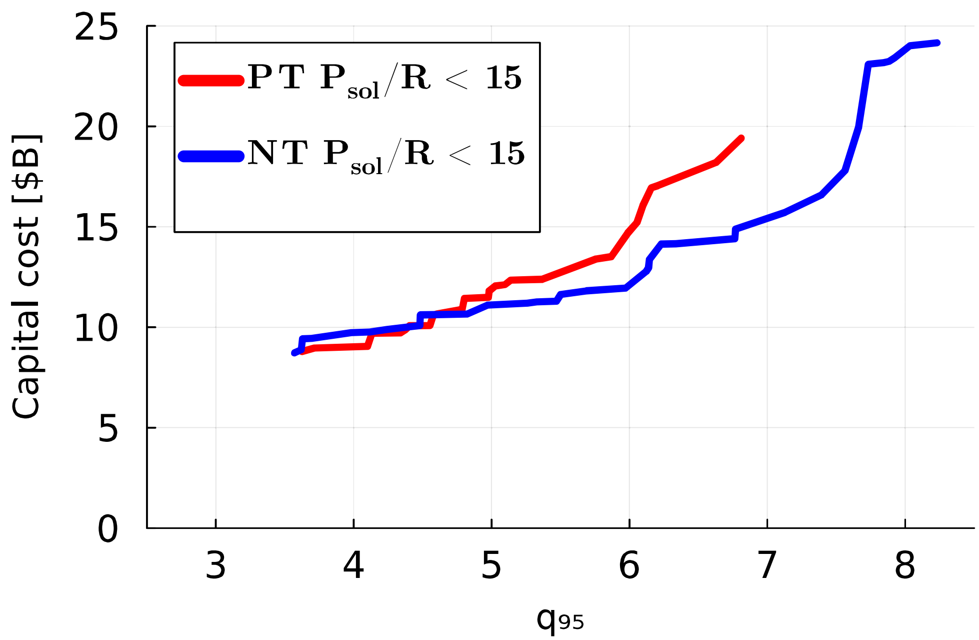


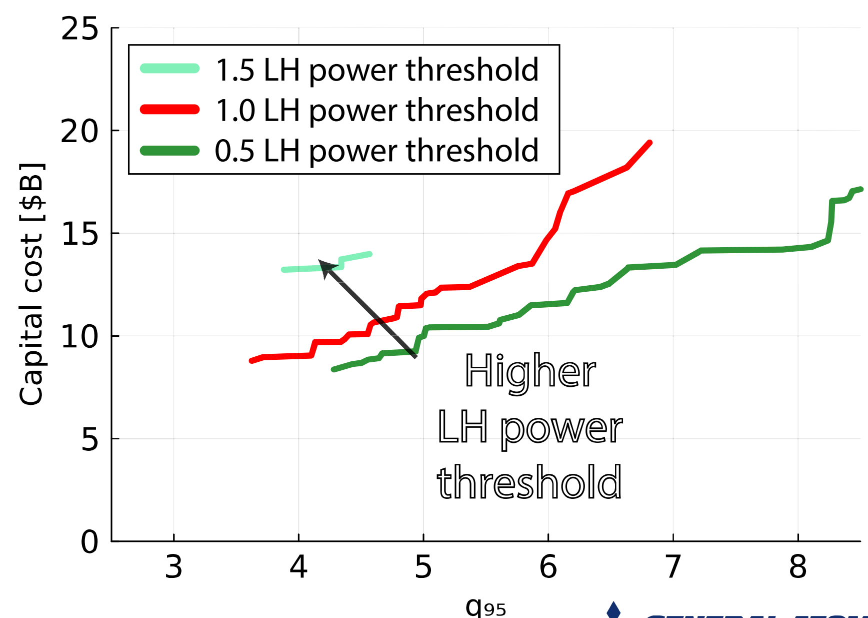
PT Mainly limited by:

- LH power threshold
- Divertor power exhaust

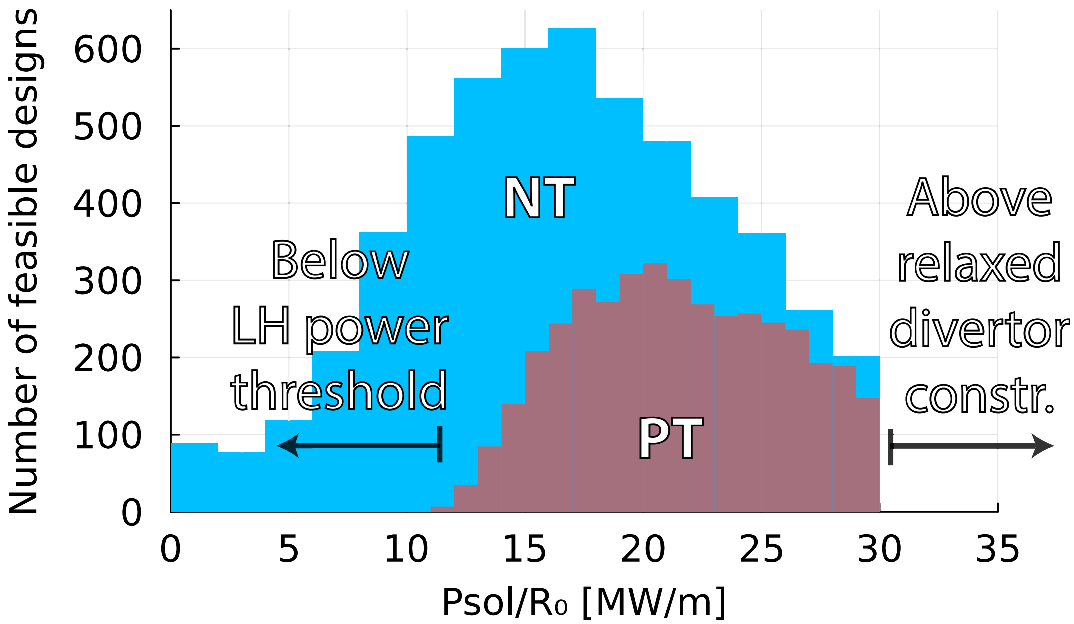
NT mainly limited by:

- Confinement (no pedestal and higher radiation)





We can use optimization datasets to relate uncertainties in key parameters to cost. Relation to risk = cost × probability

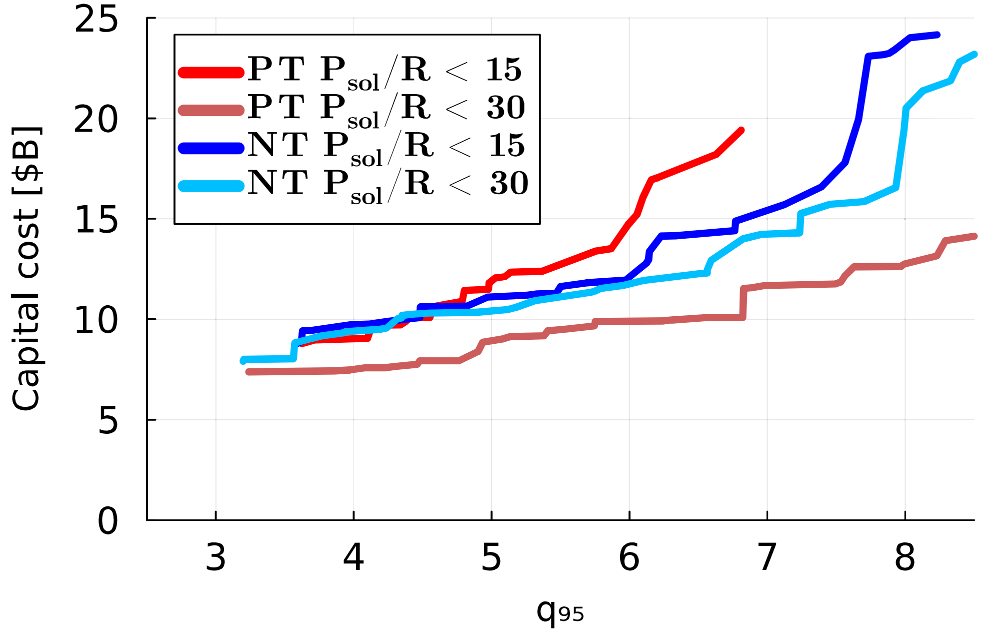


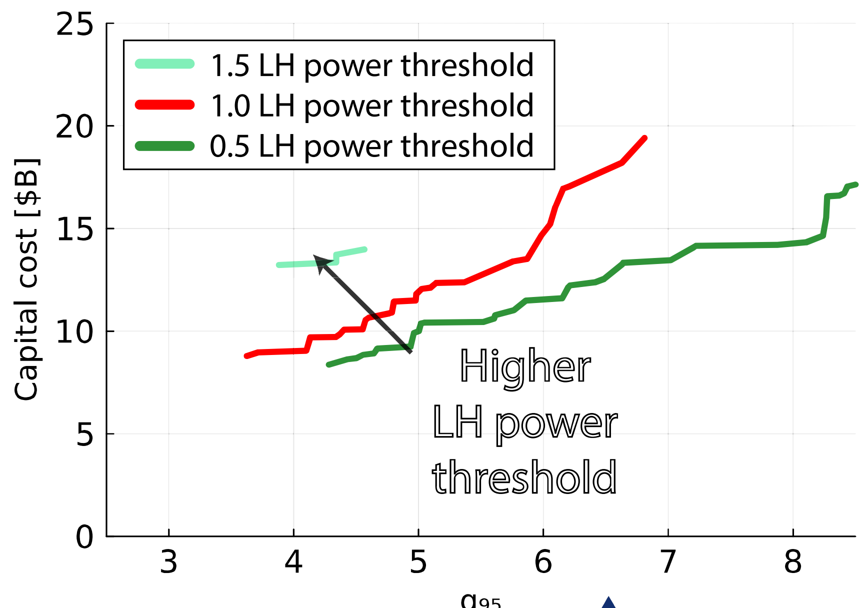
PT Mainly limited by:

- LH power threshold
- Divertor power exhaust

NT mainly limited by:

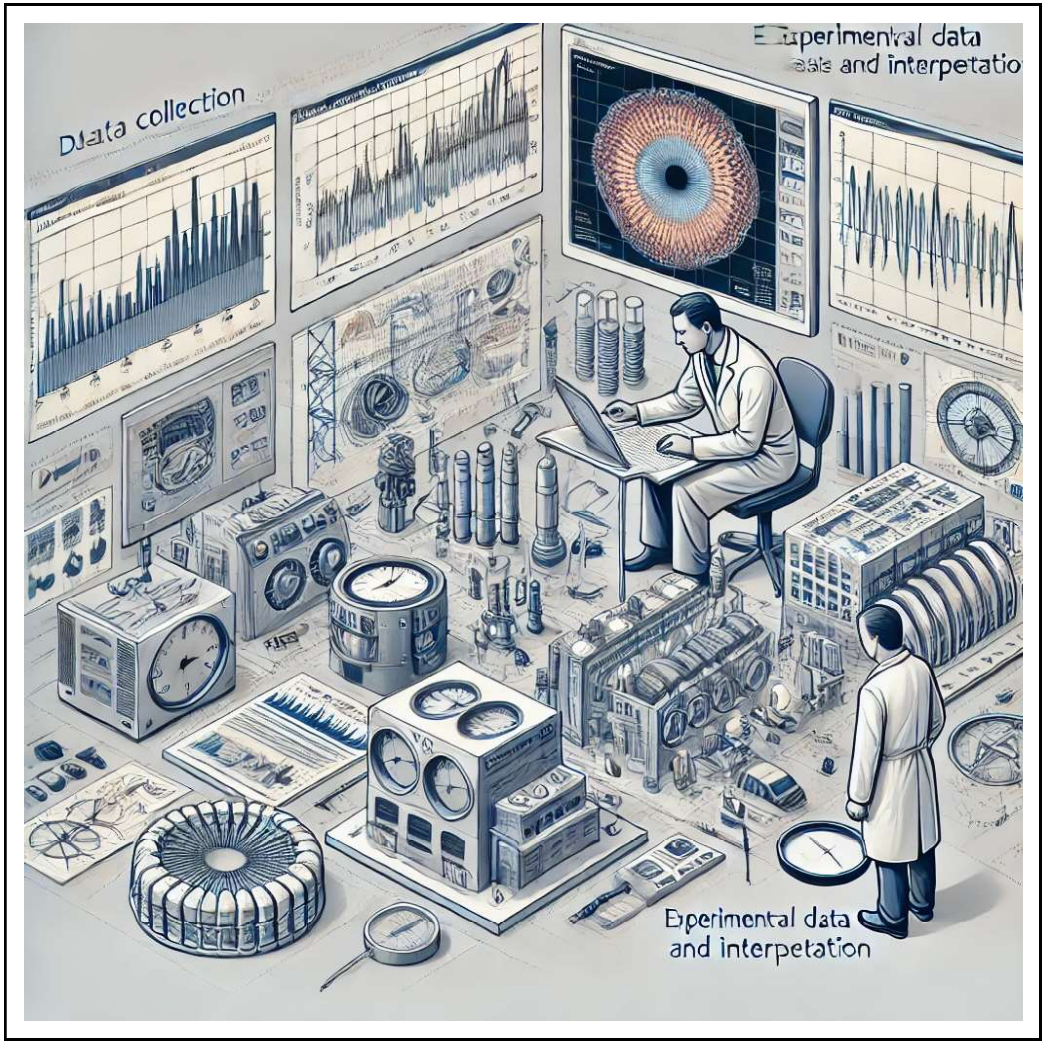
- Confinement
  (no pedestal and
  higher radiation)



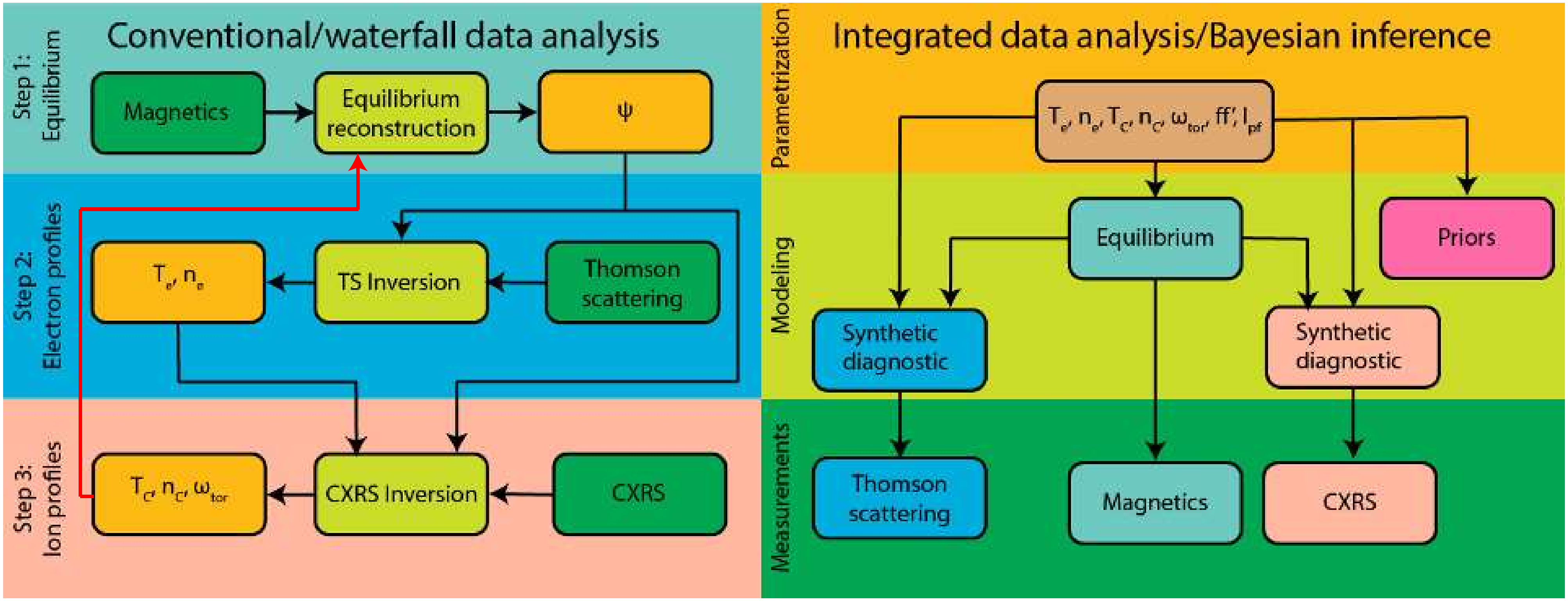


GOAL: Enable high-throughput accurate analyses that are essential for comprehensive model validation

DATA ANALYSIS



Conventional Vs integrated analysis is a strategic choice



- **Separate** inverse problems
- Iteration to reconcile couplings
- **Computationally efficient** (sec, mins)
  ⇒ applied to many shots/times

- A **single** Bayesian inference problem
- Rigorous uncertainty propagation
- **Computationally intensive** (hrs, days)
  ⇒ applied to selected shots/times

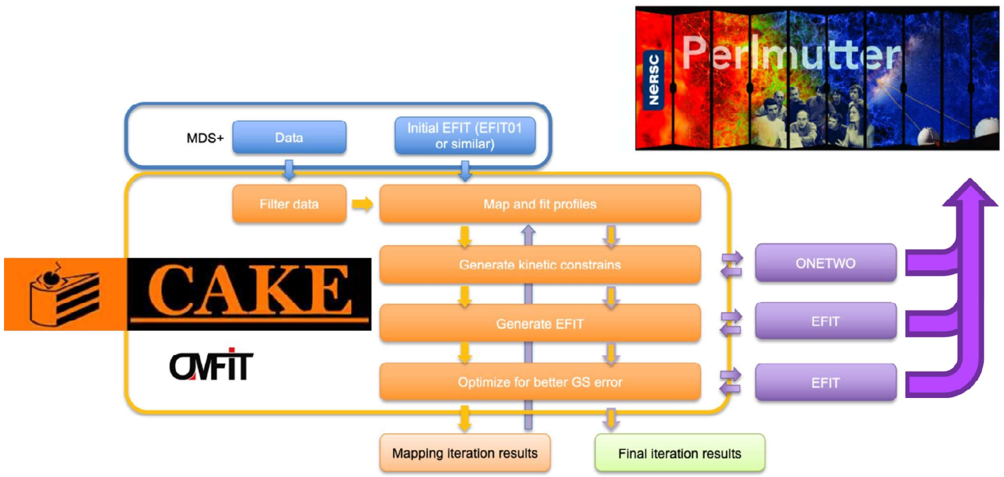
Eventually we will want to run integrated analysis for everything, but for now
Breadth + **Depth** = **Robust validation** within computational constraints
(conventional) (integrated)

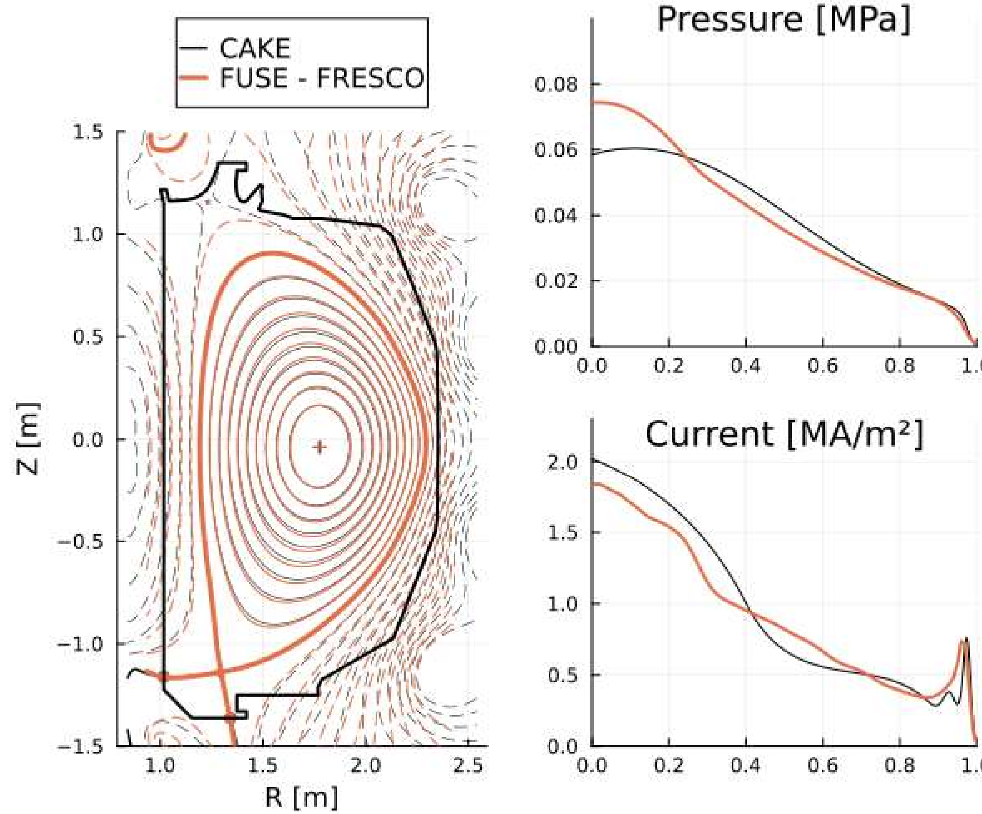
GENERAL ATOMICS

High-throughput conventional data analysis provides the large datasets needed for statistical validation (and ML)

eg. Kinetic equilibrium reconstructions and transport analyses on DIII-D:

	2013	2017	2019	2023	2025
Manual	**KineticEFIT**	**KineticEFITtime**	**CAKE**	**CAKE @ NERSC**	**FUSE**
Days	Hour	Hour	Hour	20 minutes	1 min
1 timeslice	1 timeslice	10 timeslices	50 timeslices	50 timeslices	100 timeslices





1.3M high-quality DIII-D kinetic EFIT reconstructions using 10k node-hours

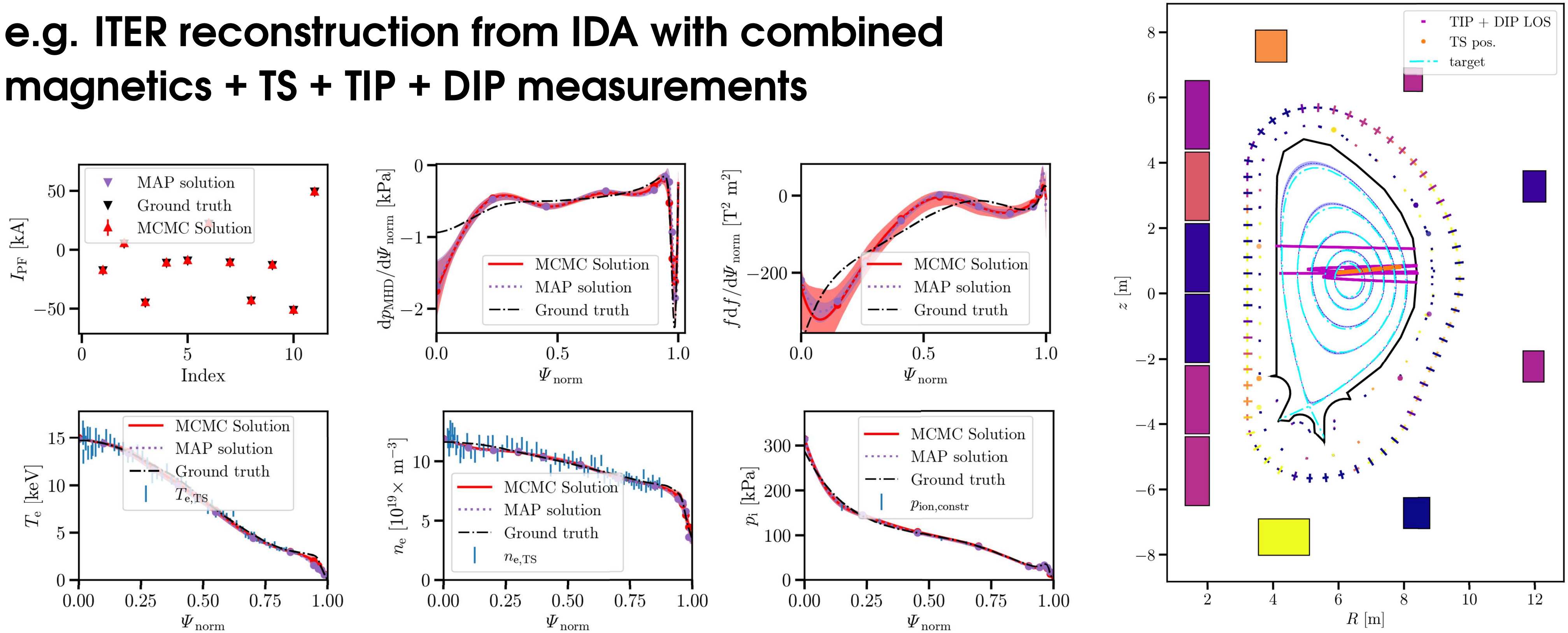
- Starting point for any exp. analysis
- Train DIII-D ML eq. surrogates

GENERAL ATOMICS

Integrated data analysis maximizes the information yield, especially critical for next-step devices

Fusion pilot plants will have few and more limited diagnostics than current experiments, because of nuclear and radiation constraints

e.g. ITER reconstruction from IDA with combined magnetics + TS + TIP + DIP measurements



NOTE: To be practical any **IDA must rely heavily on ML forward models**

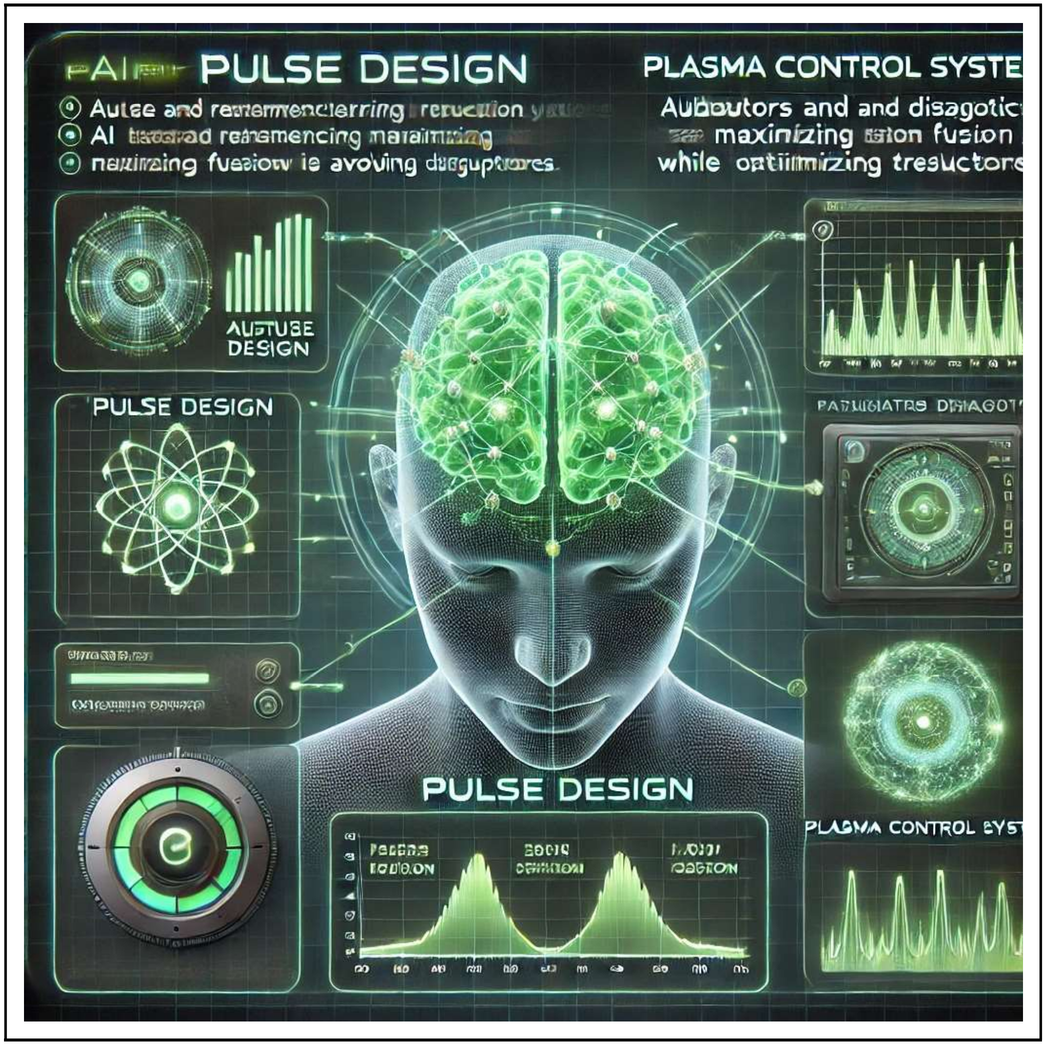
- FUSE used to create 50k self-consistent eq. and transport solutions used for training ITER free-boundary equilibrium ML surrogate

GOAL: Time-dependent capabilities for fast, high-fidelity, machine-agnostic pulse design with PCS integration

Feed-forward simulation

For scientists

PULSE DESIGN



Feed-back simulation

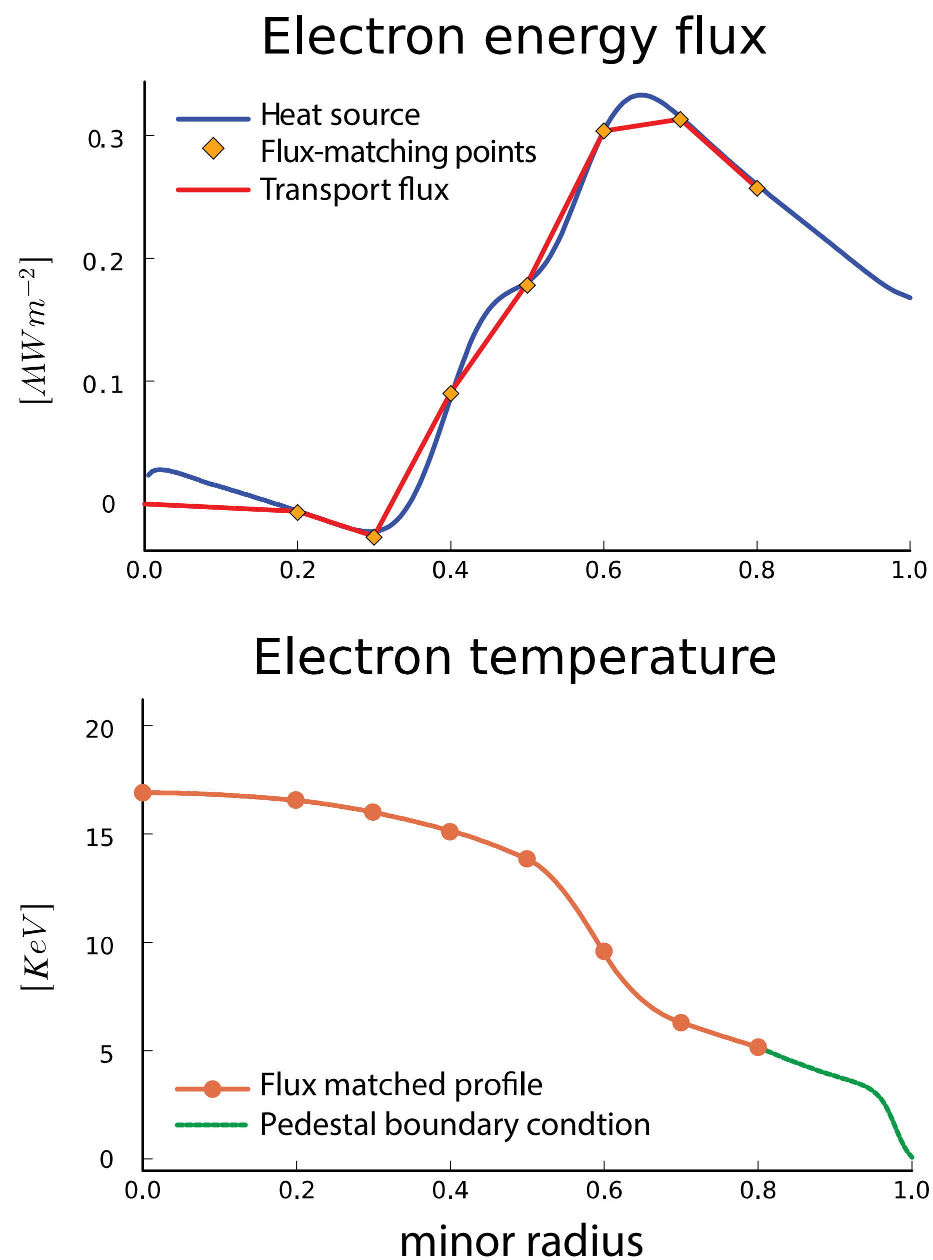
For physics-operators

To maximize computational efficiency FUSE uses an implicit time-dependent flux-matching transport solver

$$\frac{\partial}{\partial t} + \langle \nabla \cdot \Gamma \rangle = S$$

STATIONARY

- No time dependence
  $t \rightarrow \infty$ and $\dfrac{\partial X}{\partial t} = 0$
  $\cancel{\dfrac{\partial}{\partial t}} + \langle \nabla \cdot \Gamma \rangle = S$
- Flux-matching at few radial locations + linear profiles inverse-scale-length



To maximize computational efficiency FUSE uses an implicit time-dependent flux-matching transport solver

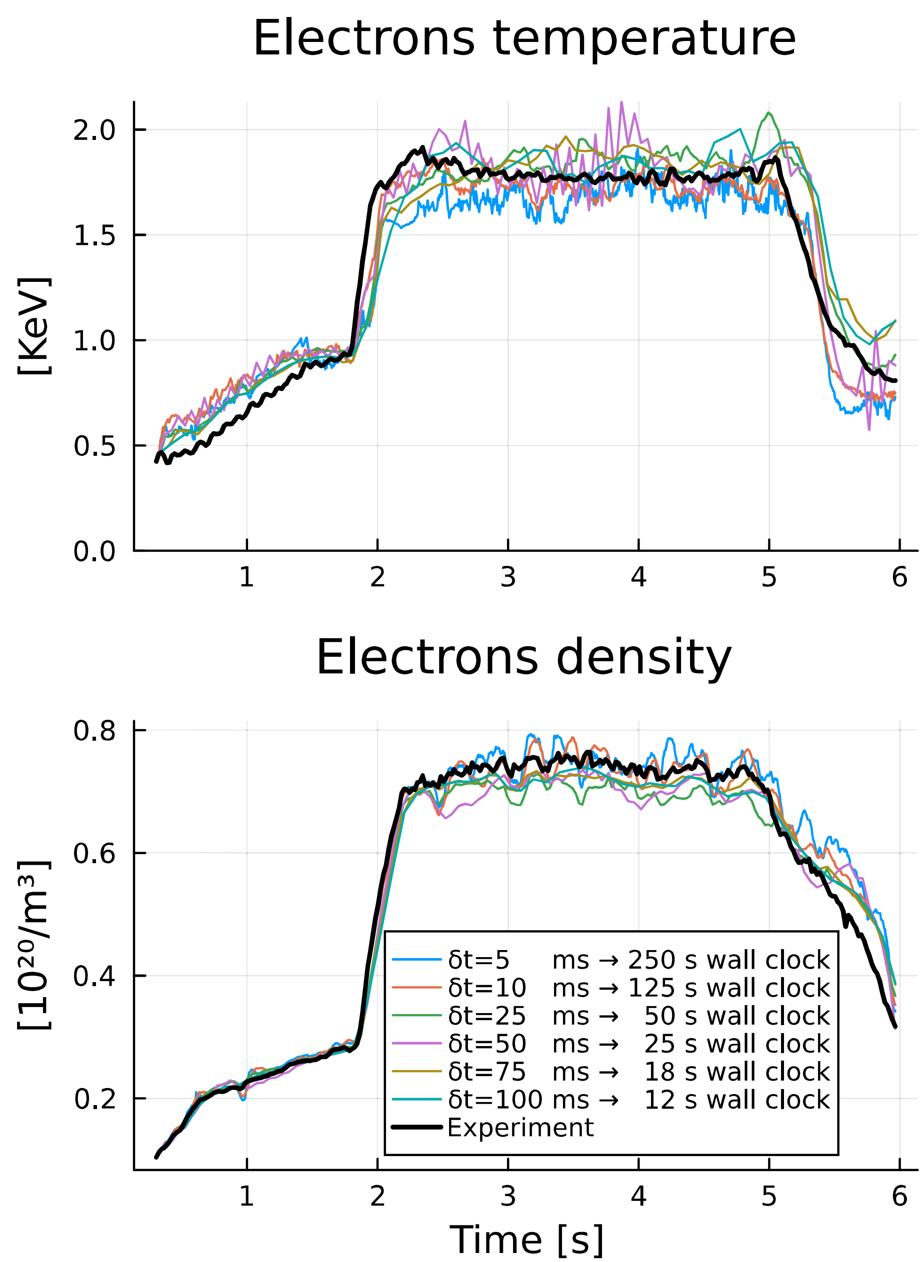
$$\frac{\partial}{\partial t} + \langle \nabla \cdot \Gamma \rangle = S$$

STATIONARY

- No time dependence
  $t \to \infty$ and $\frac{\partial X}{\partial t} = 0$
  $\cancel{\frac{\partial}{\partial t}} + \langle \nabla \cdot \Gamma \rangle = S$
- Flux-matching at few radial locations + linear profiles inverse-scale-length

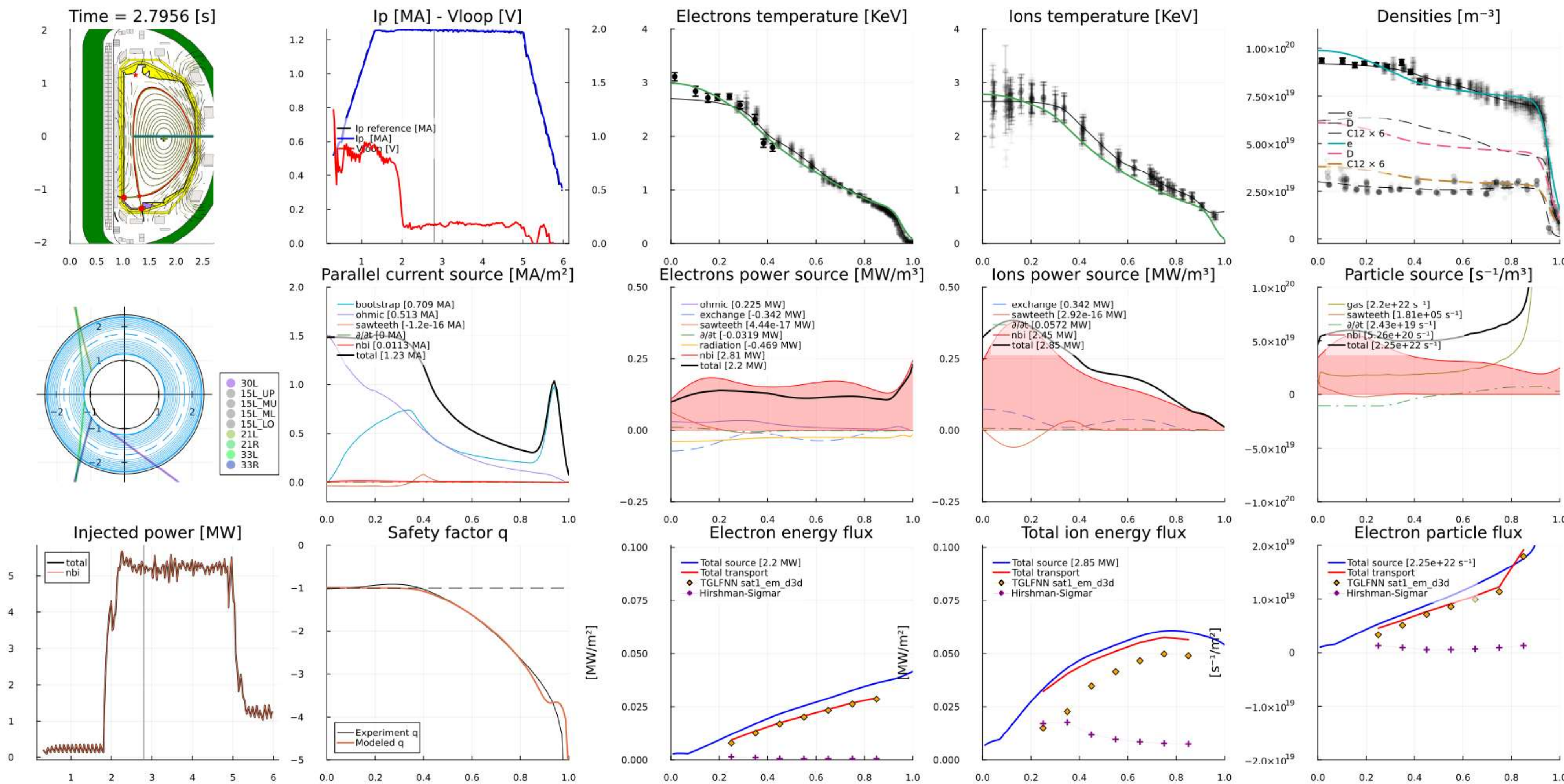
DYNAMIC

- Time-derivative as a source
  $\langle \nabla \cdot \Gamma \rangle = S - \frac{\partial}{\partial t}$
- Implicit time stepping, allows taking larger steps



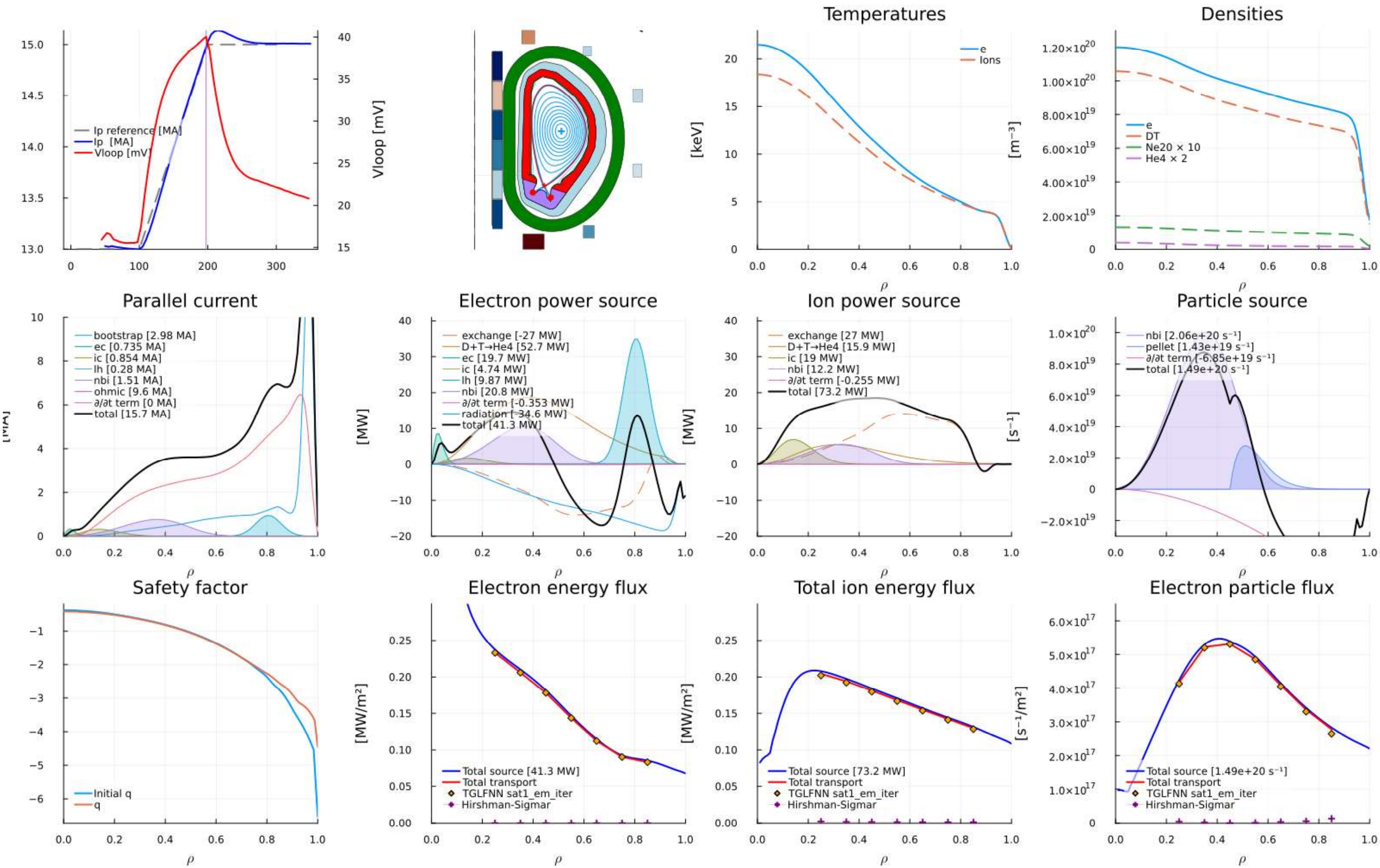
GENERAL ATOMICS

Time continuity provides a stringent validation of models and couplings. From rampup, through flattop, to rampdown



→ see FUSE in action ←

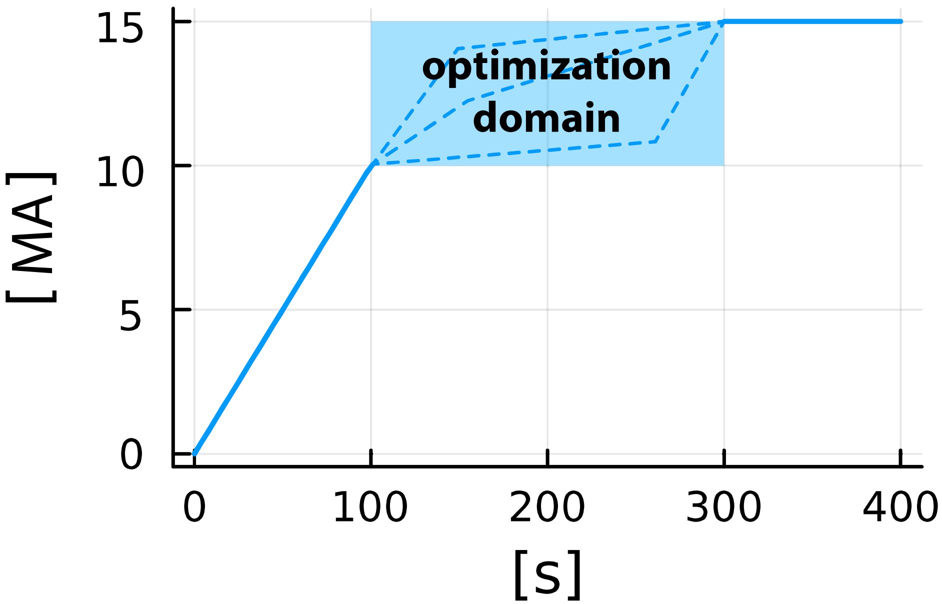
The validated models are then used to make predictions of new experiments and devices



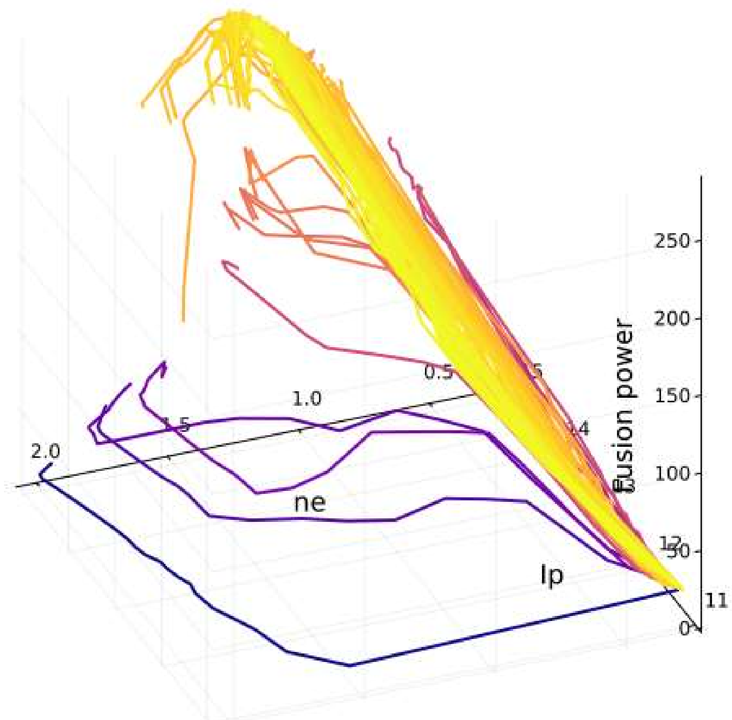
GENERAL ATOMICS

Time dependence + Optimization = Trajectory optimization with built-in sensitivity analysis

Leverage **same optimization infrastructure used for machine design**



eg. Find optimal Ip and n_e ramp rates to max ITER fusion energy

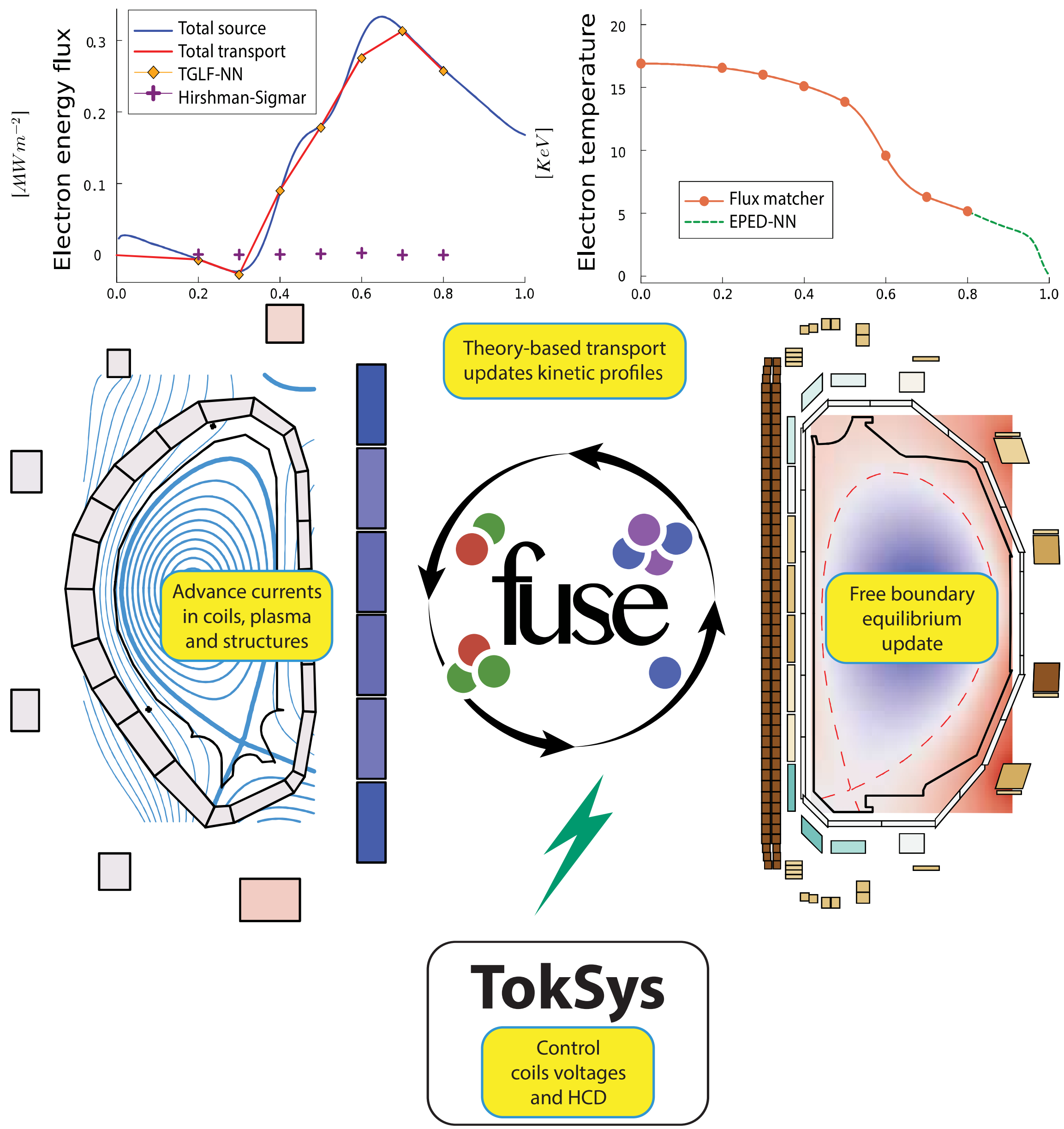


- Define optimization domains for actuators time traces
- Define time-depedent objectives and constraints
- Take full advantage of HPC

GENERAL ATOMICS

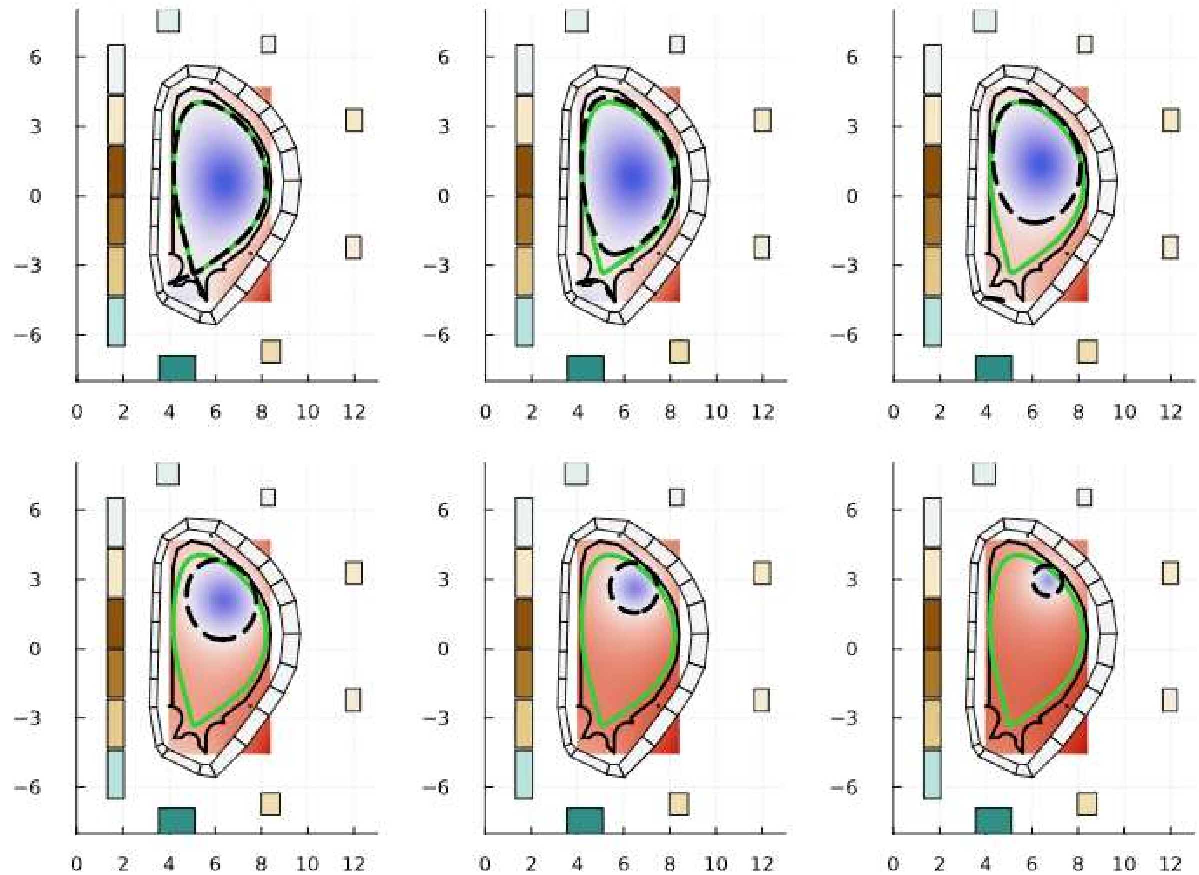
Grad-Hogan solver under development to model plasma dynamics in combination with control system

1. Free-boundary solver
2. Theory-based transport
3. Inductive coupling of plasma, PF coils, and conducting structures
4. Co-simulation with control system
   - Initially with **TokSys**
   - Developing coupling with **DIII-D PCS**



Grad-Hogan solver under development to model plasma dynamics in combination with control system

1. Free-boundary solver
2. Theory-based transport
3. Inductive coupling of plasma, PF coils, and conducting structures
4. Co-simulation with control system
   - Initially with **TokSys**
   - Developing coupling with **DIII-D PCS**



VDE modeling in FUSE

Outline

- Fundamentals of V&V and UQ
    - Validation and Verification
    - Fidelity hierarchy for V&V
    - Statistical validation
    - Uncertainty propagation
- FUSE: Implementing V&V and UQ in practice
    - Machine optimization
    - Integrated data analysis
    - Pulse design
- Key Takeaways

Integrated modeling with proper validation, verification, and uncertainty quantification matters now more than ever

The Stakes

- ITER: $20B investment
- FPPs: $B decisions ahead
- Limited shots for learning
- No room for surprises

The Opportunity

- Design with confidence
- Optimize before building
- Learn from every shot
- Accelerate deployment

V&V/UQ Fundamentals

1. Validation transforms models from theoretical tools into trusted predictive capabilities
2. Statistical validation across parameter ranges is essential. Not just point comparisons!
3. ML surrogates bridge the gap between computational efficiency and physics fidelity
4. UQ guides integration boundaries: stop when parameters are well-controlled or uncertainties dominate

FUSE puts V&V and UQ principles into practice:
Machine optimization – Data Analysis – Pulse Design

Open source ecosystem

- Apache 2.0 (OK commercial)
- 25+ packages
- 200K+ lines of Julia
- Documented
- Regression tested
- Preprint on Arxiv

http://fuse.help

Consider Julia for your next software project!

- High-level, fast, auto-diff
- Enthusiastic community
- Most Julia devs were former Python devs ;)