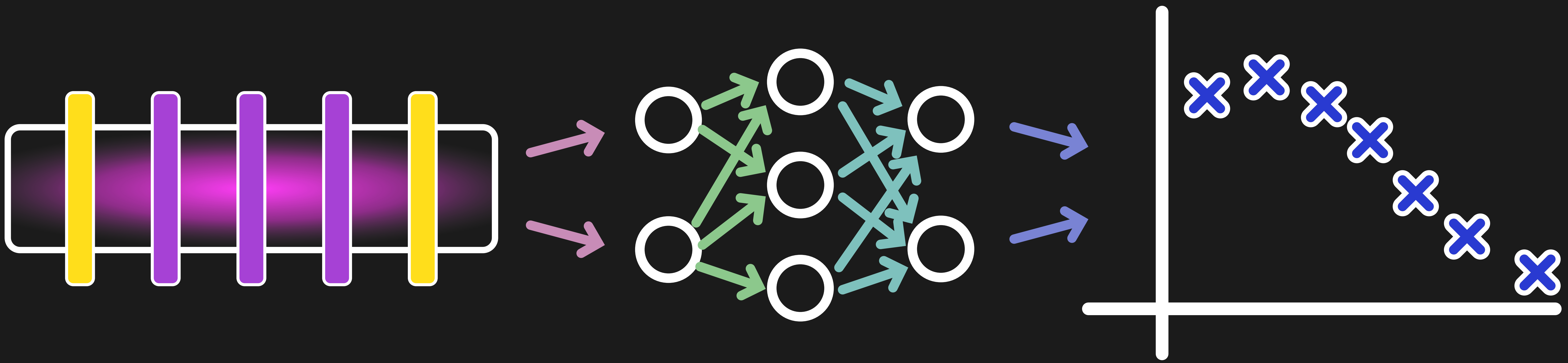
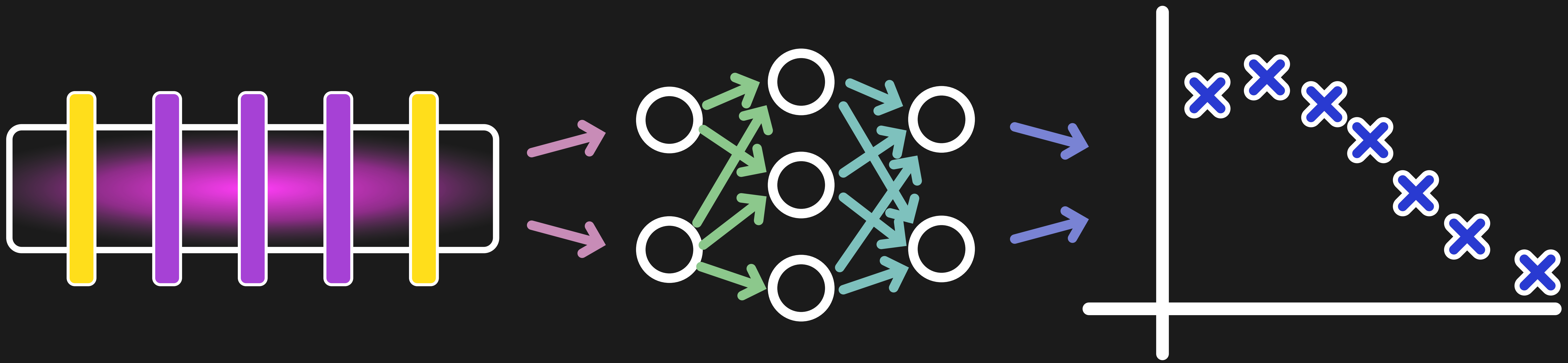
Building a generative LAPD transport model



ATC
Phil Travis
5/11/2021

Building a generative LAPD transport model



LAPD | **Generative** | **Transport**

LAPD

- 20 m long, 1m wide
- T_e ~ 5+ eV
- Shots ~ 16 ms long
- Rep rate: 1 Hz

Generative

- Learns distribution $p(S)$
- Can sample $s \sim p(S)$
- Can *generate* data

Transport

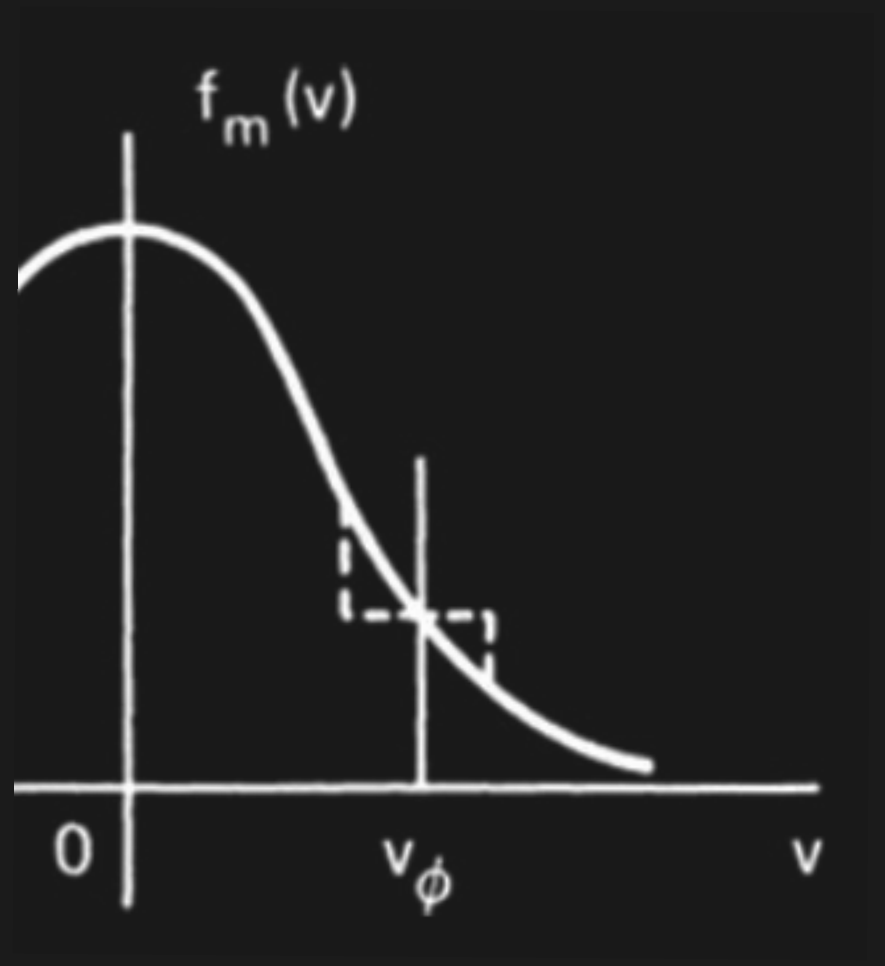
- Heat and particle loss
- Cross-field
- Primary concern in MCF

ATC Outline: building a generative LAPD transport model

- Why ML matters
  - Case studies: machine learning is important for the future of fusion science
  - Accelerating science with ML and energy based models
- What I'm doing
  - LAPD project: comprehensive transport characterization
  - Side projects: optimizing plasma quality, autonomous search
- Looking forward: long-term potential is huge
- Quick demo

ML provides a new prediction paradigm

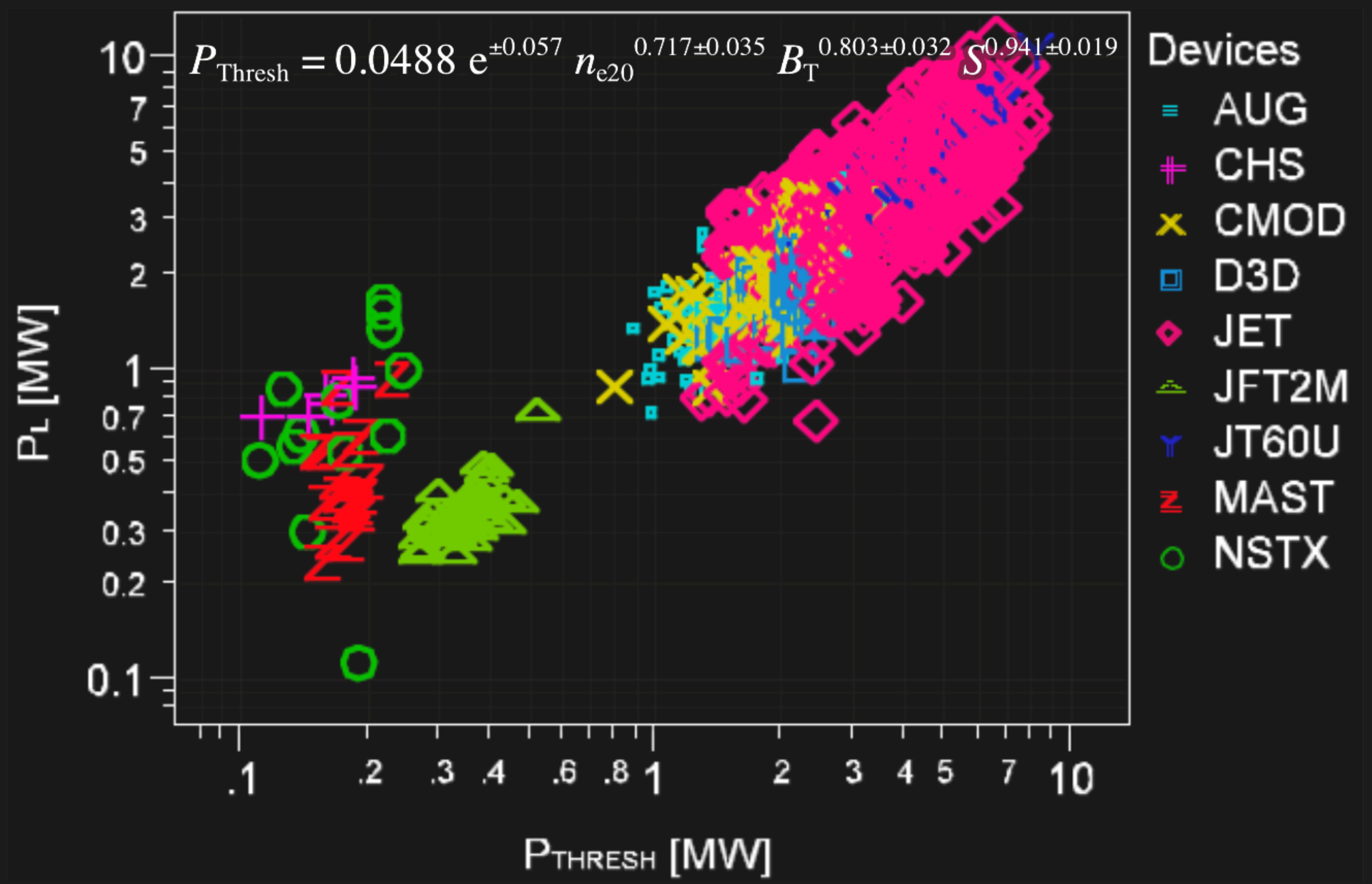
Theory



Landau damping (Chen)

- Provides insight
- Wide scale range
- Very hard

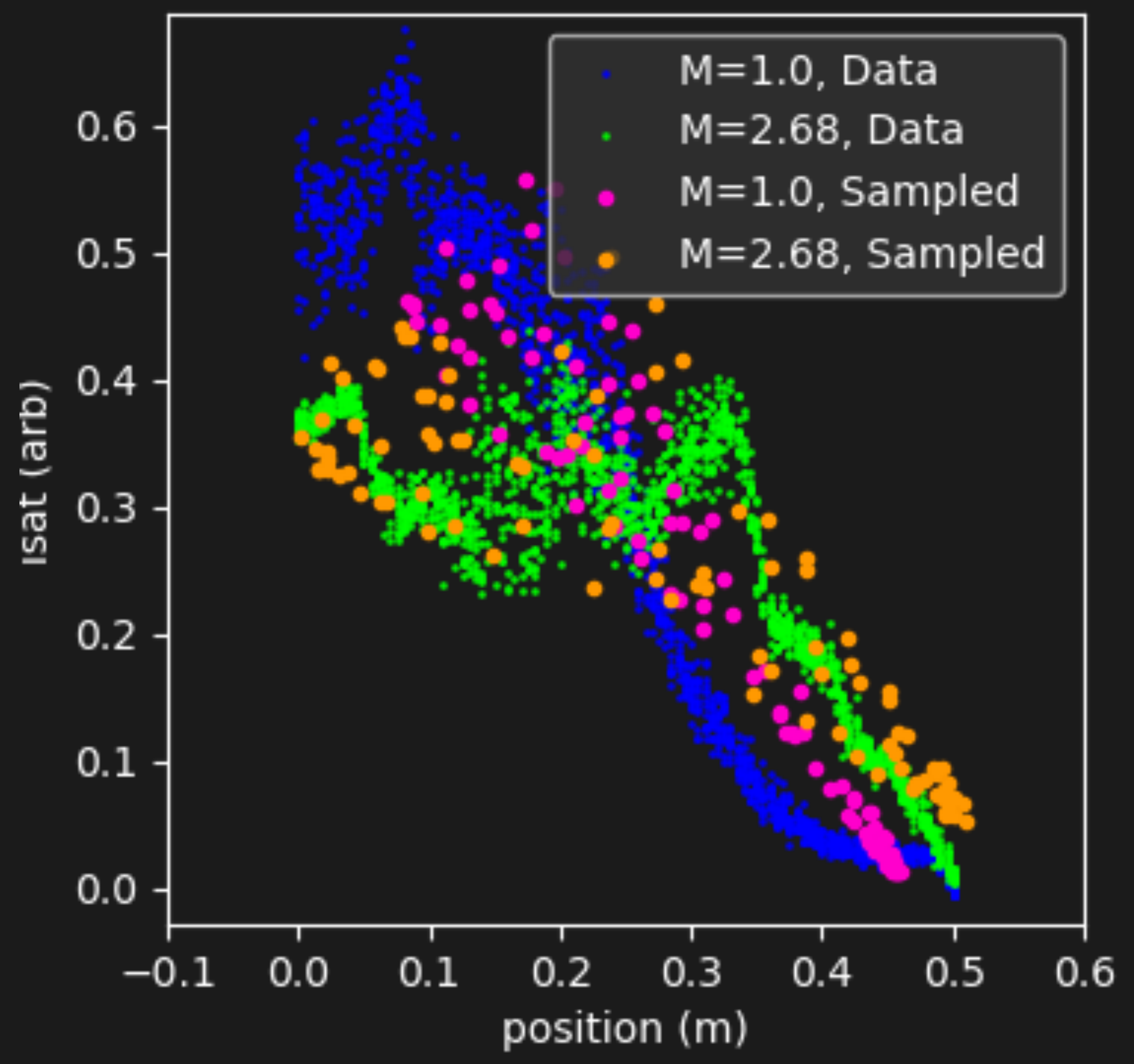
Empirical approach



H-mode threshold scaling (Martin et al. 2008)

- Simple
- High bias

Machine learning

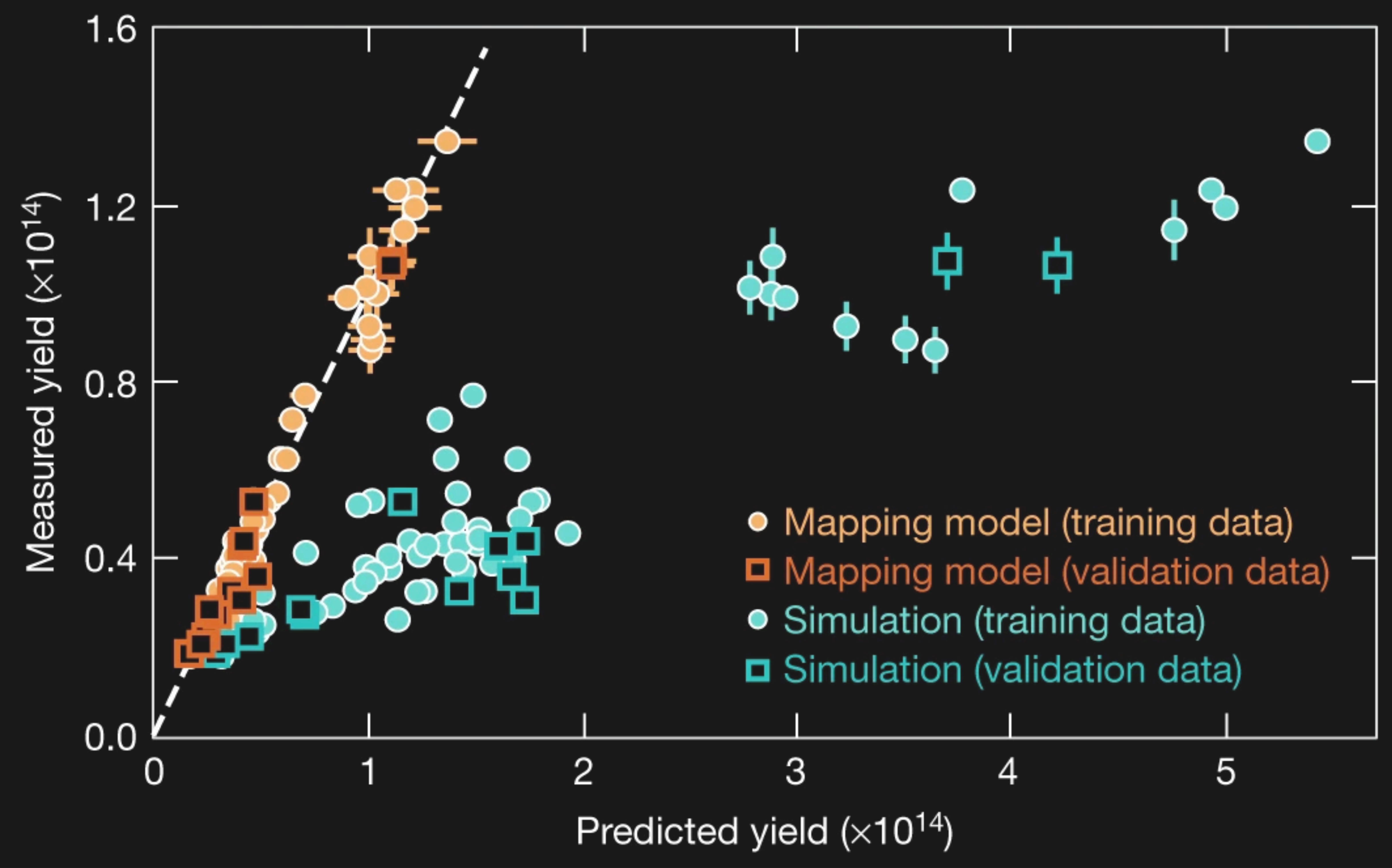


Sampling LAPD profiles via an EBM

- Low bias*
- Learns structure
- Opaque

Learned models are useful for correcting our models

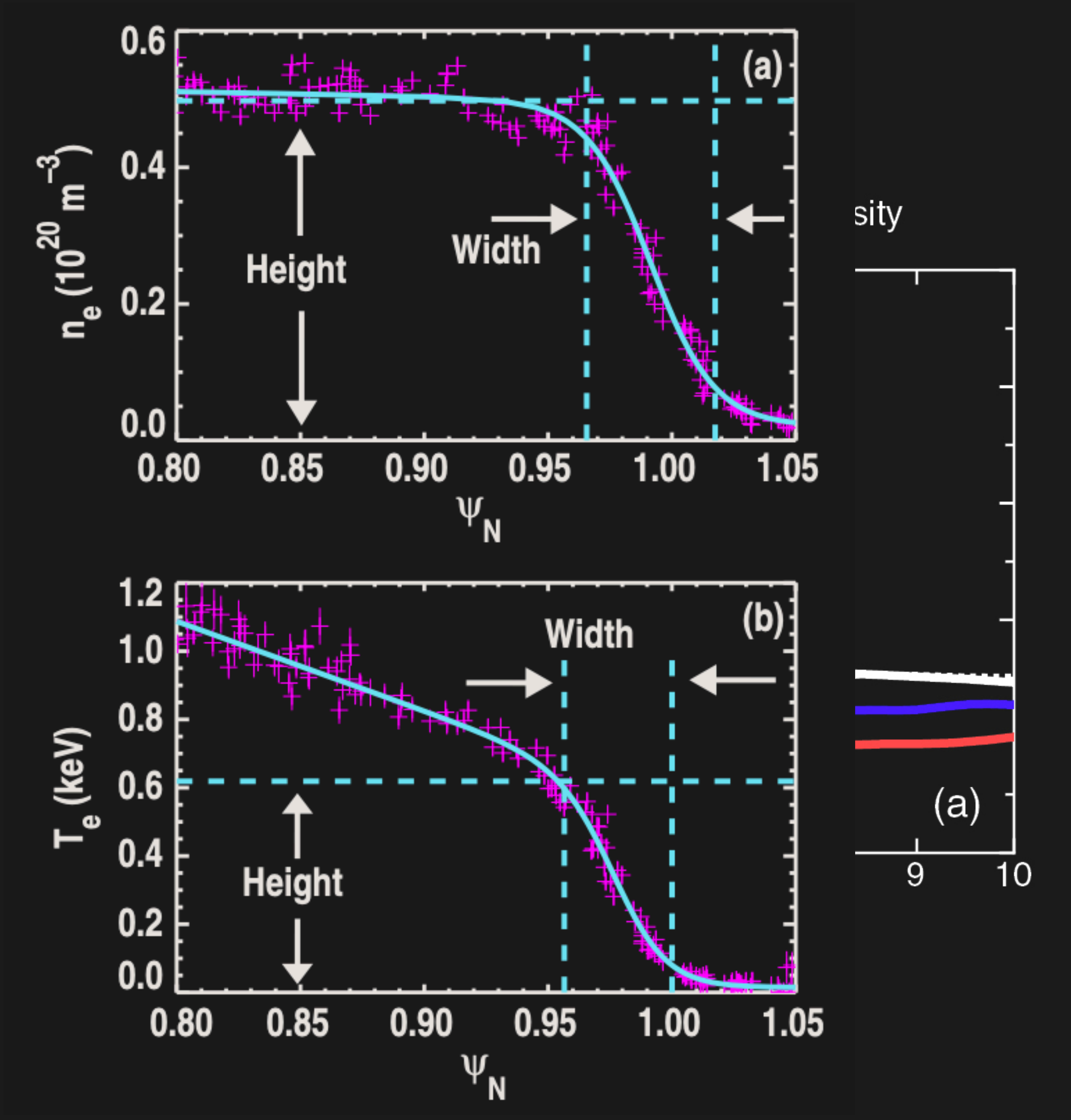
- OMEGA laser pulse and target optimization
  V. Gopalaswamy et al. 2019
  - Statistically mapped simulations to experiments for given inputs
  - Tripled fusion yield by varying target parameters and pulse shape
- Model-based methods enable *planning*
  Reinforcement Learning, Sutton



V. Gopalaswamy et al. 2019

Learned models can compensate for lack of understanding

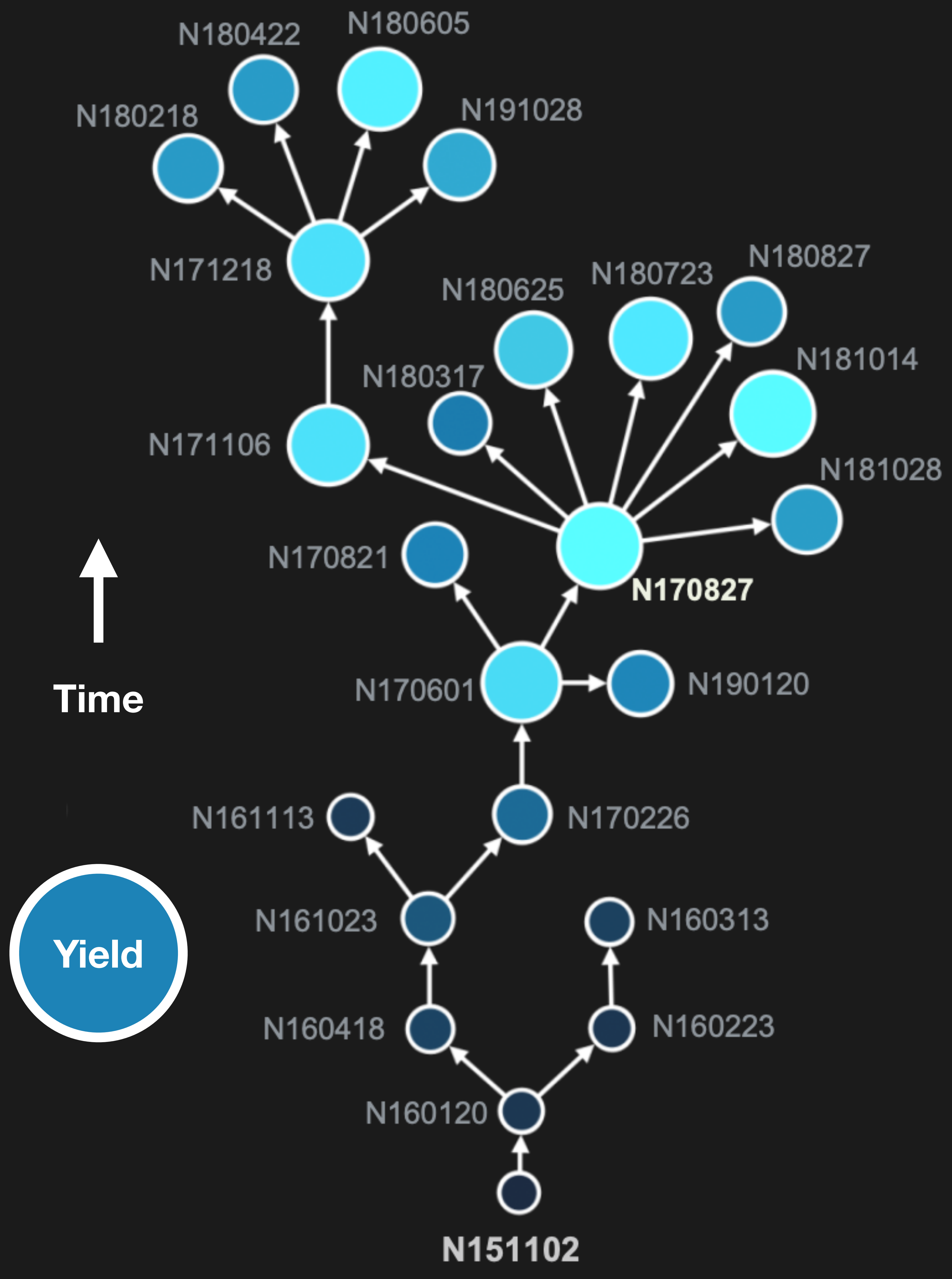
- Example: super H-mode prediction
  (Snyder et al. 2015)
  - Strongly-shaped confinement improvement
  - May have been observed years before prediction
    (Matthias Knolker, private communication)
- Counterfactual: what else has been observed that we cannot use?
  - ML models enable use of information without theoretical prediction



Groebner et al. 2009

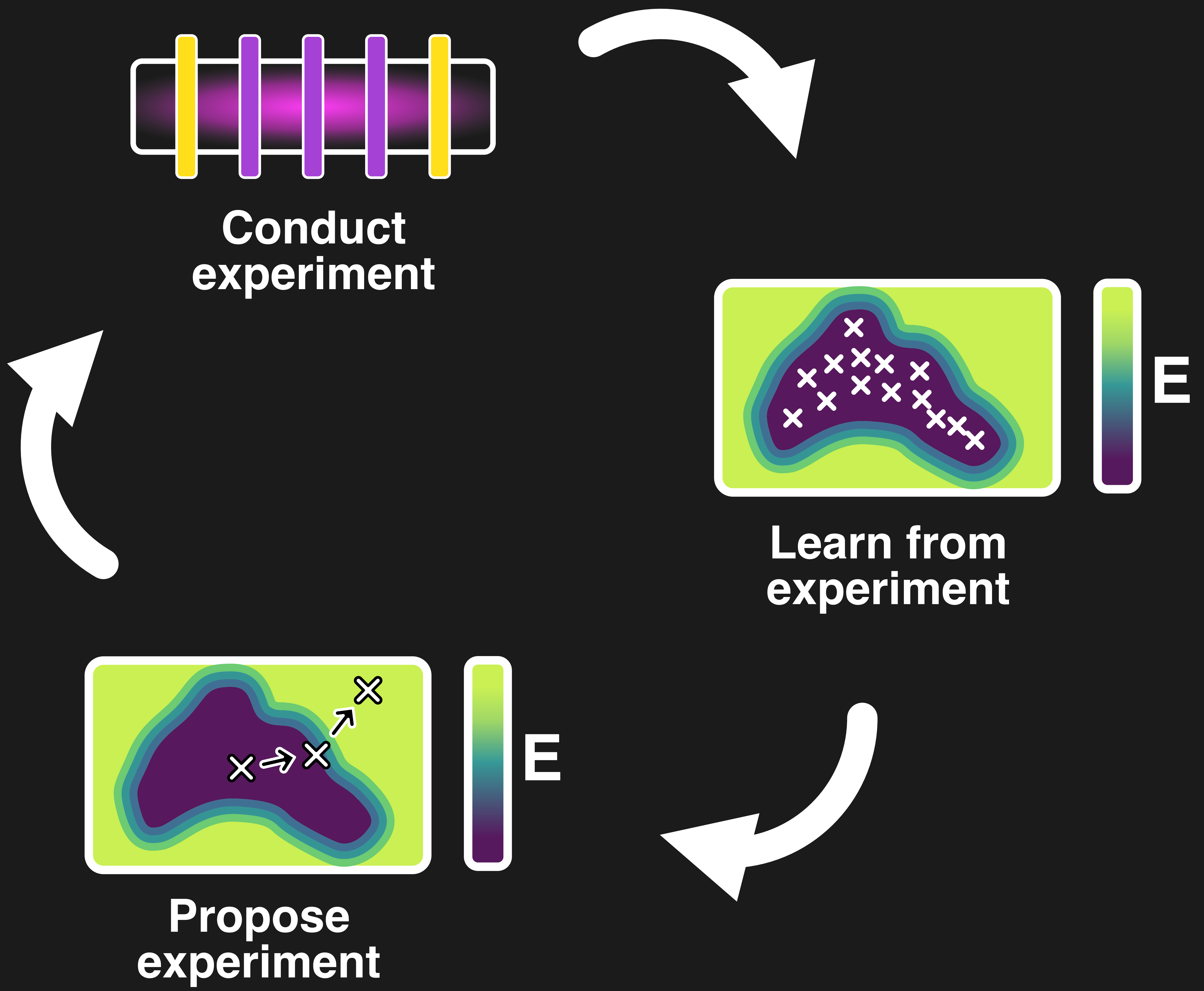
Learned models may alleviate human shortcomings

- Evidence of bias in NIF experiment planning
  - Assumption relaxation —> largest yield increases
  - Lack of disconfirming experiments
  - Theory introduces bias
- Is there bias in our MCF experiment designs?
- A learned model has no preconceived notions*



Finnegan @ SIAM MDS 2020

Long-term goal: automate fusion science



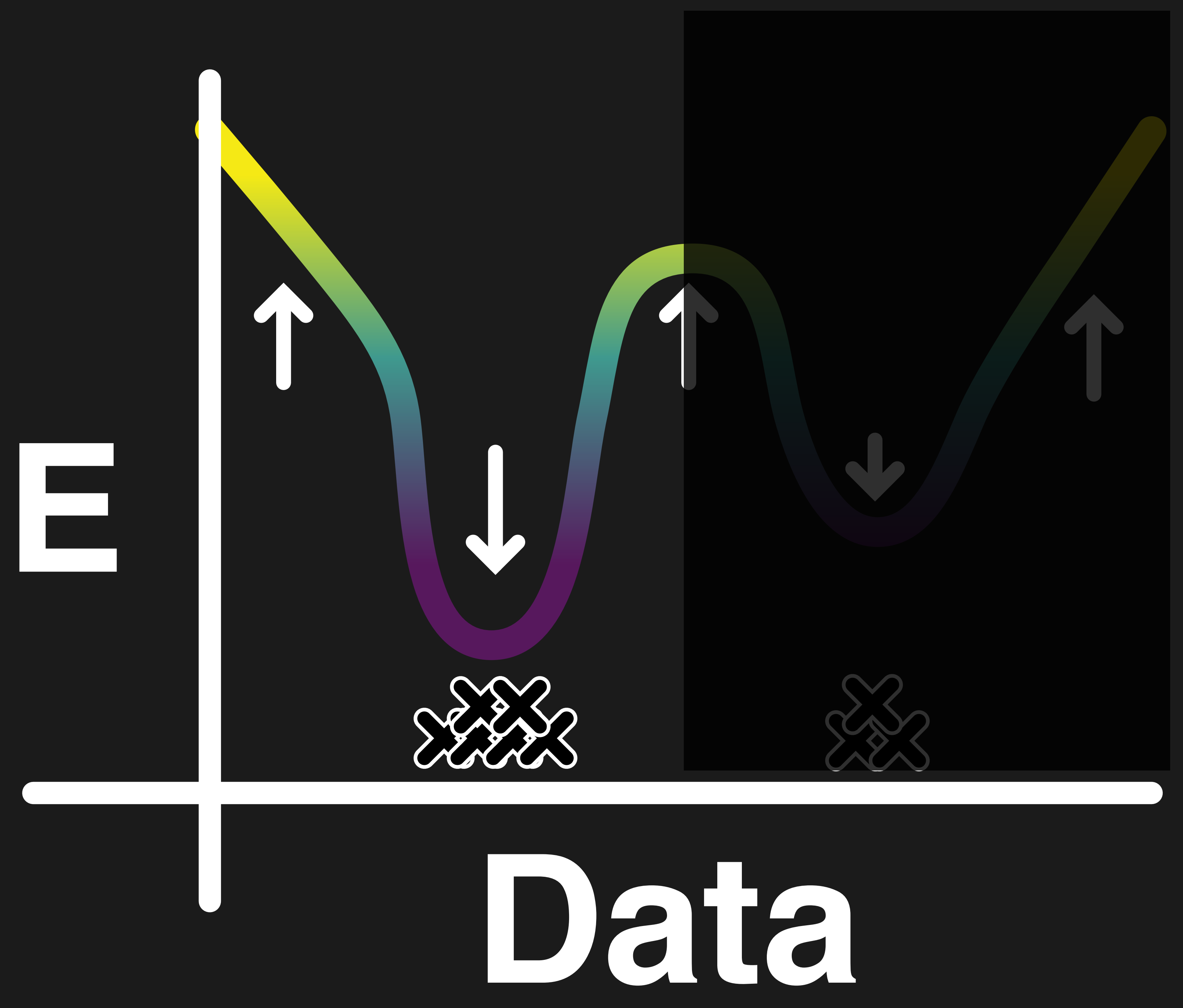
- Learning structures permits automated exploration
  - encourages rapid device iteration
  - memorization is not sufficient
- Something needs to understand fusion reactors
  - it need not be humans
- My goal: set us on this path

Energy based models (EBMs) learn the distribution of the data

LeCun 2006

$$p(x) \sim e^{-\beta E(x)}$$ <- generative ML model

Data → NN → E

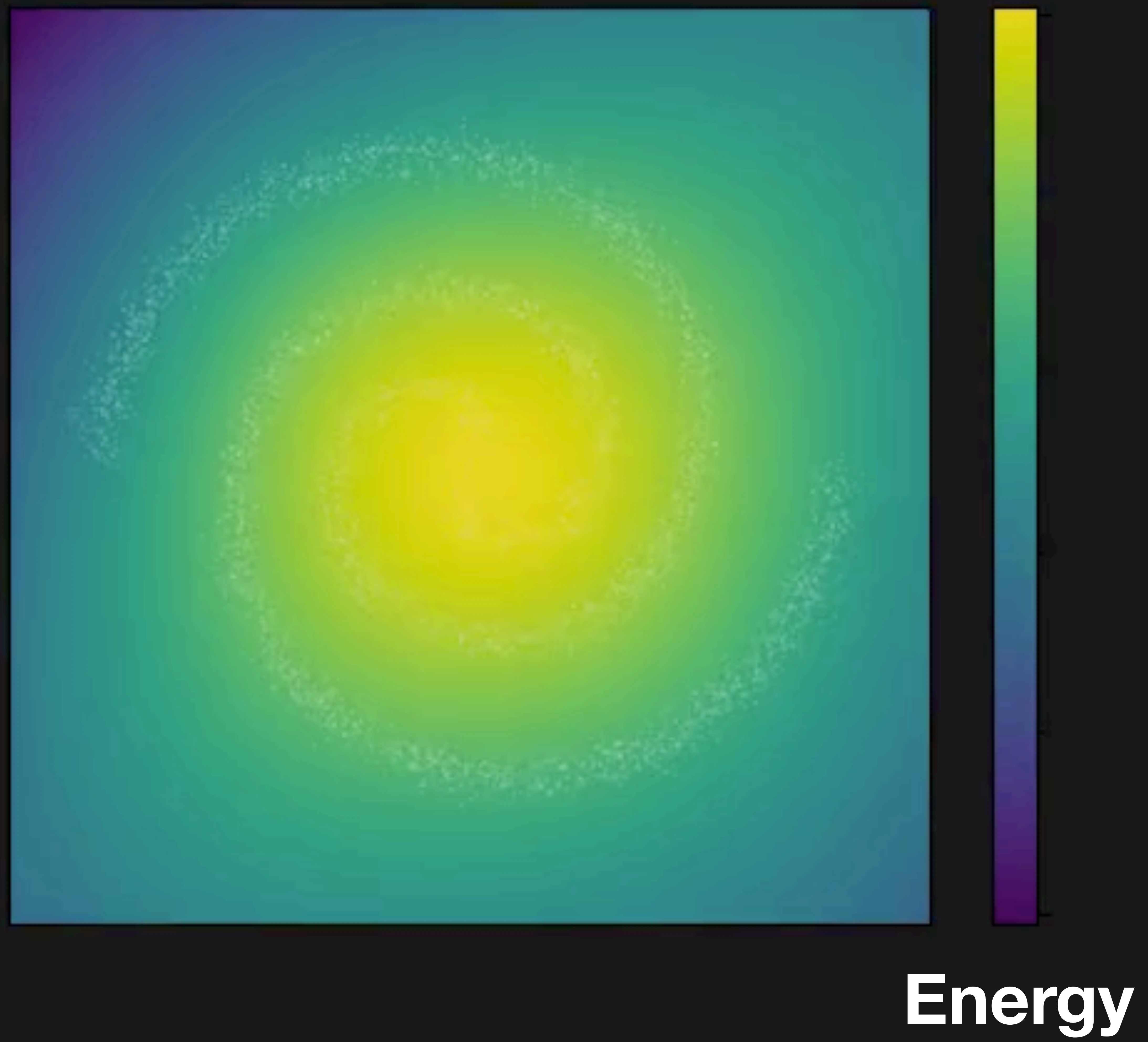


- Learns an energy surface
- Sample from this energy surface

$$x_L \sim p(X \mid Side = L)$$

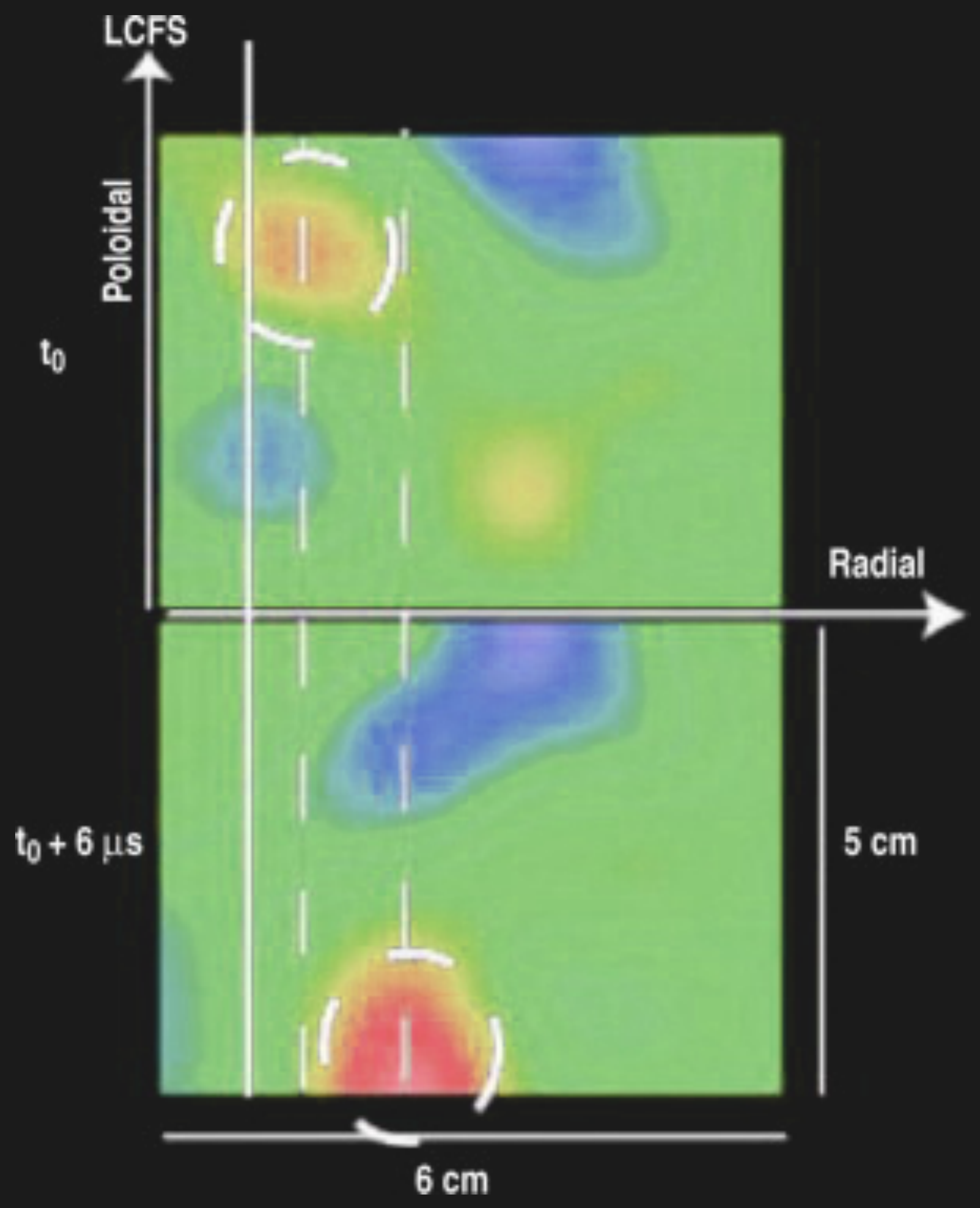
- Learn how everything is related

EBM: training and sampling

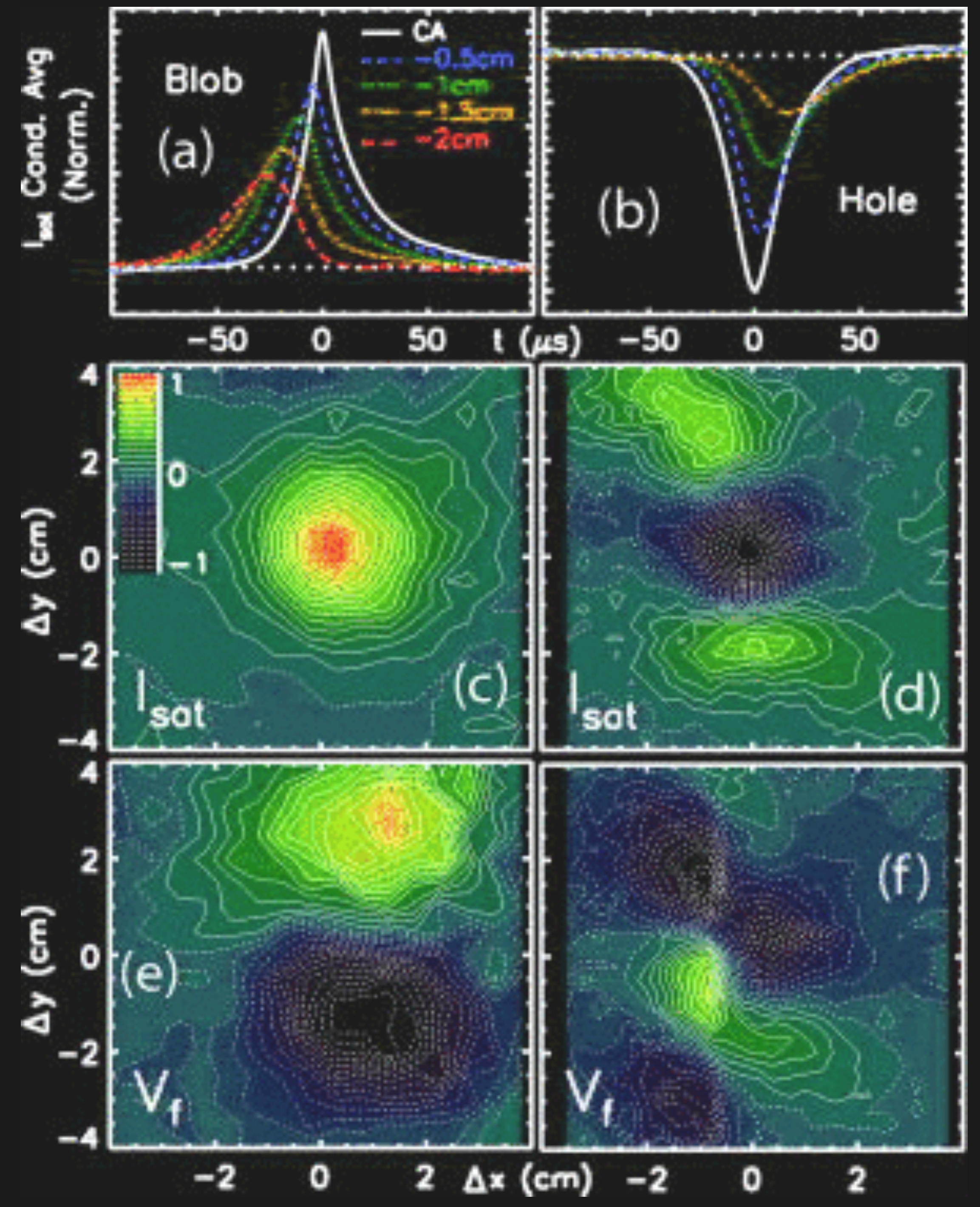


The LAPD is an ideal testbed for EBMs + fusion

- High rep rate, flexible, accessible
  (Gekelman et al. 2016)
  - 31 million experiments per year
- Is capable of performing fusion relevant studies (mostly edge)
  - Sheared flow turbulence suppression
    (Schaffner 2013)
  - Intermittent filamentary structures
    (Boedo 2009, Carter 2006)
  - Drift wave turbulence
    (Tynan et al. 2009)



"Blobs" in DIII-D (Boedo 2009)



"Blobs" in LAPD (Carter 2006)

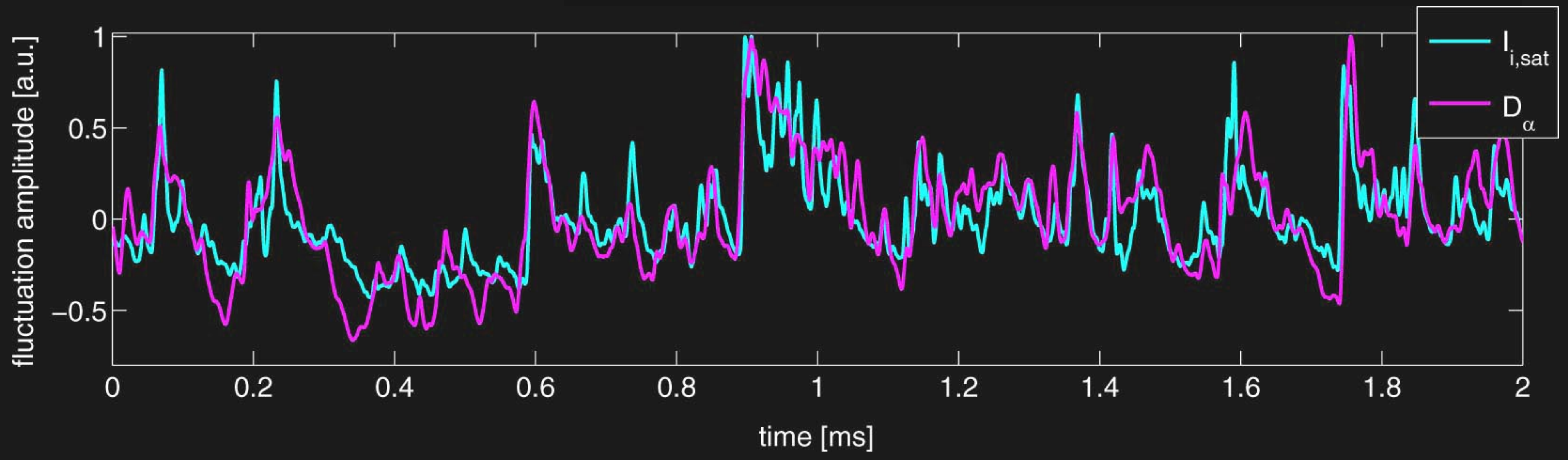
Projects: a comprehensive LAPD transport study

- Project 0: auxiliary diagnostics system augmentation
- Project 1: profile reconstruction
- Project 2: theory incorporation
- Project 3: wide-range characterization
- Side projects

- Why: simplest problem that can demonstrate generative model efficacy

Project 0: augmenting an auxiliary diagnostics pipeline

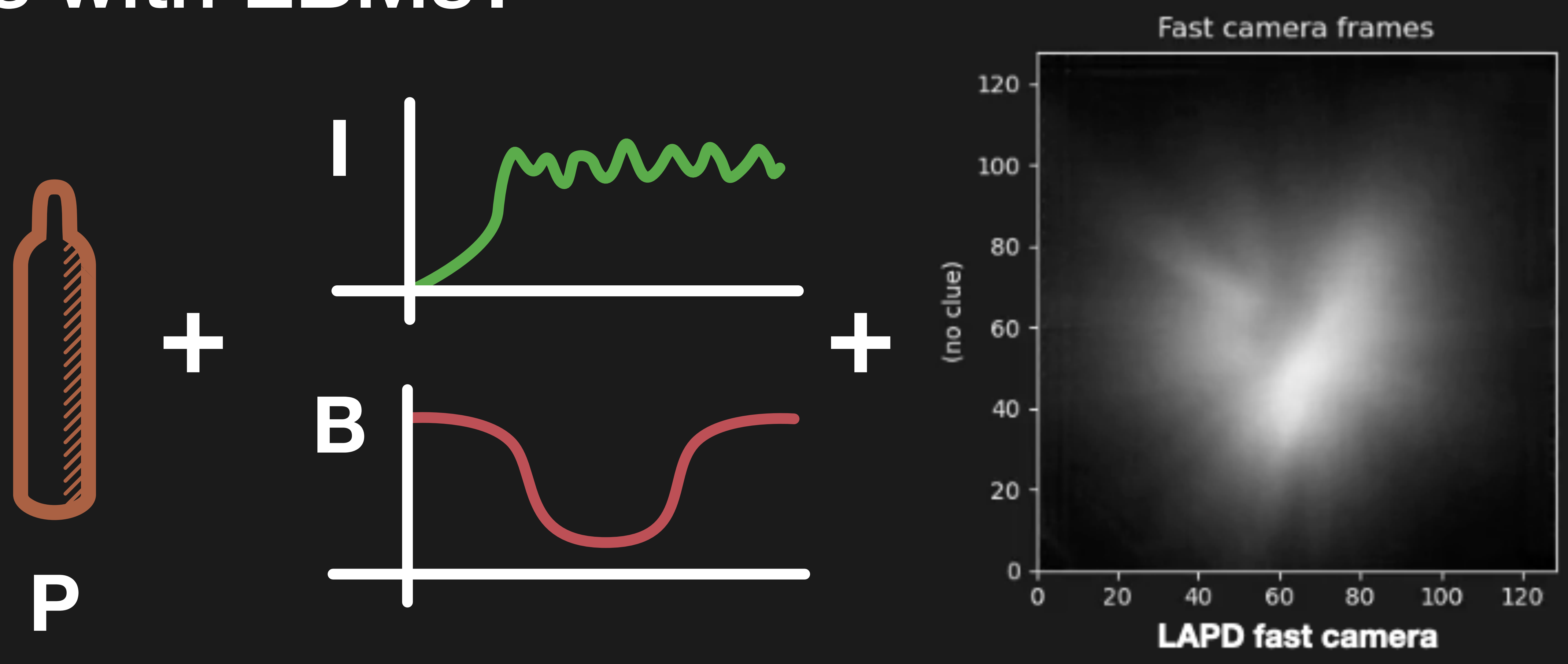
- Limitations restraining ML methods:
  - Lack of machine state information (MSI)
  - Diagnostics are highly localized (probes)
- Goal: build auxiliary data acquisition system
  - Fast camera
  - Diamagnetic loop
  - Spectrometers
  - May try algorithmic control



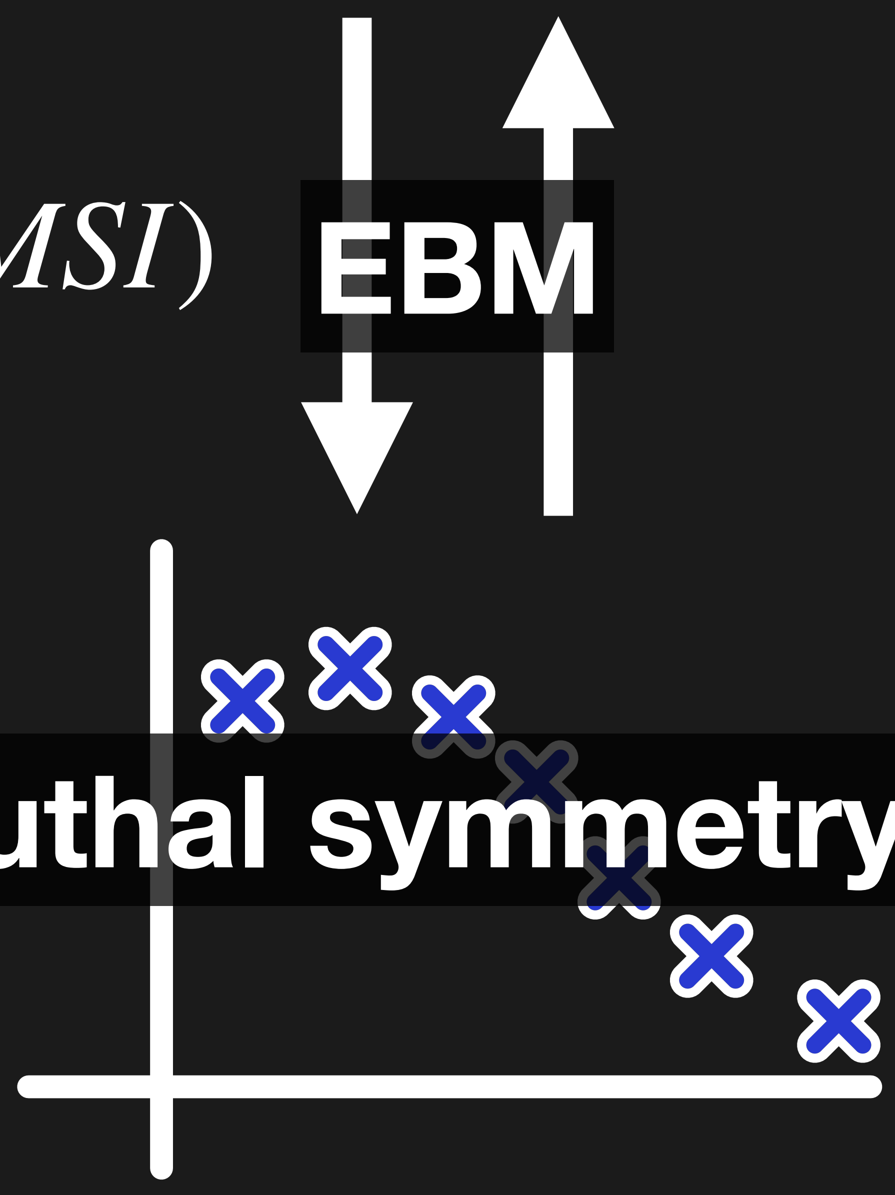
isat-GPI time series (Grulke et al. 2014)

Project 1: can we reconstruct profiles with EBMs?

- Combine MSI + global diagnostics (project 0) + probes
- Gather data with various: fill pressures, discharge currents, field strengths
- Goals: using an EBM,
  - from few probe measurements, infer profile
  - from a profile, infer machine state



$n(x) \sim p(N\,|\,MSI)$



Relax azimuthal symmetry assumption?

Project 2: can we deduce transport by incorporating theory?

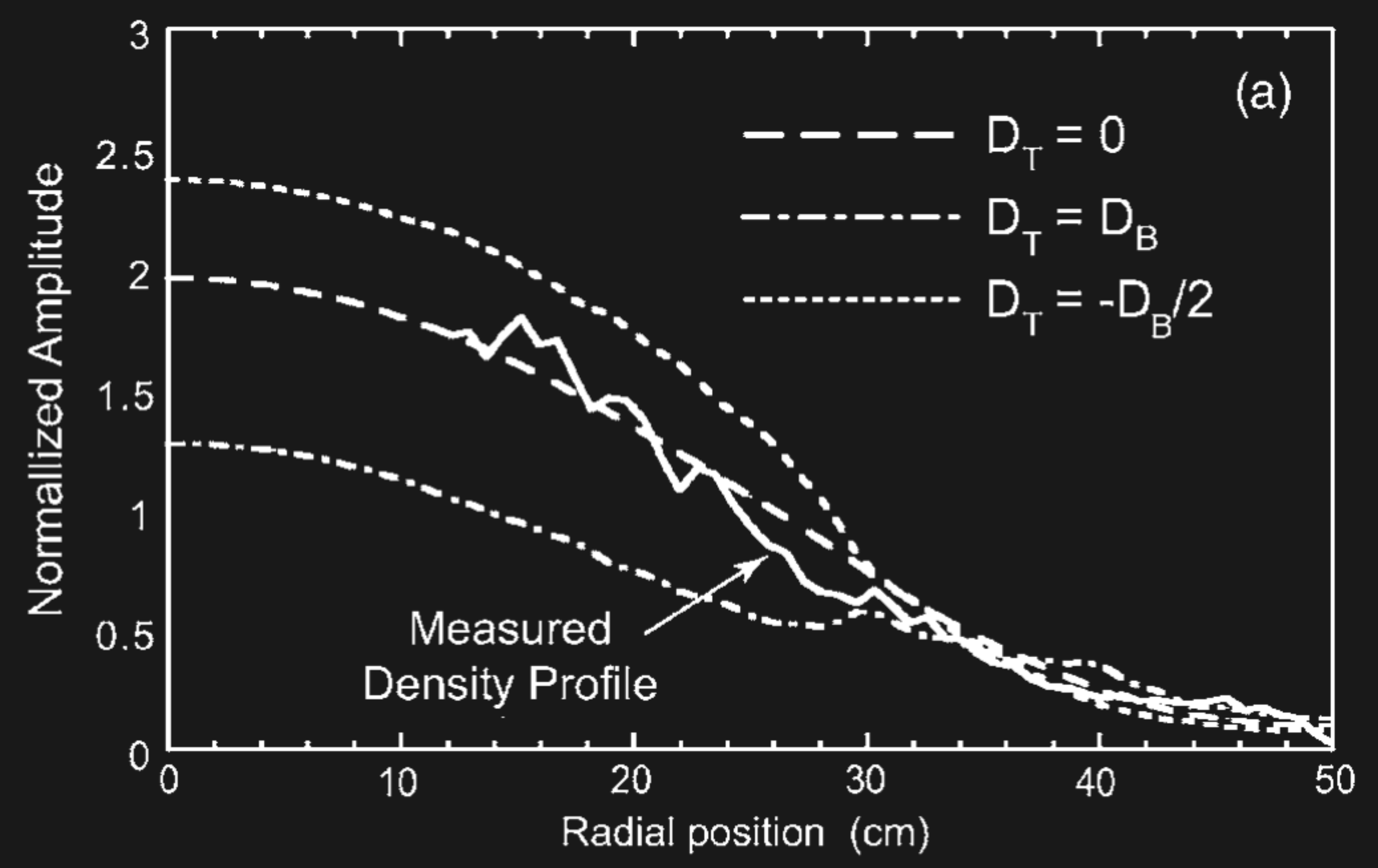
- Incorporate LAPD transport model
  (Maggs et al. 2007)
  - Includes radial dynamics
- Probably use a surrogate model
- Goals:
  - predict diffusion coefficients
  - could predict temperature profiles
  - evaluate model via experiments

$$\Gamma_{e,r} = -D_F \frac{\partial n}{\partial r} - D_T n \frac{\partial \ln(T_e)}{\partial r}$$

Radial particle flux

Bulk flow diffusion

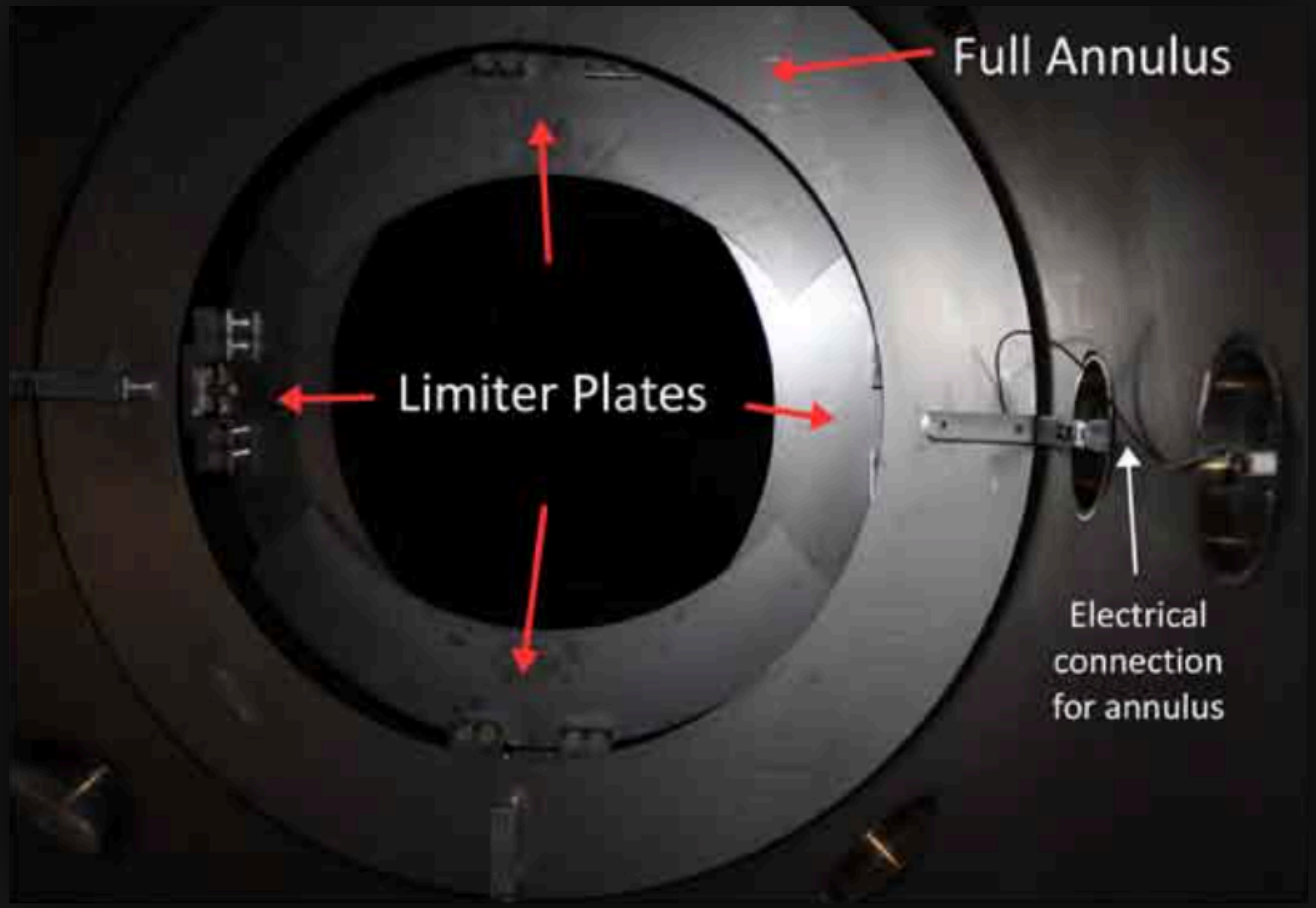
Thermal diffusion



Maggs et al. 2007

Project 3: can we characterize LAPD transport over a wide range?

- Expand applicability of model over the wide range of parameters
  - Also vary: edge bias, field profiles
  - Possibly vary: antenna waveforms, powers
- Goals:
  - learn and provide confinement trends
  - attempt learning of non-stationary variables (relax reproducibility assumption)



Schaffner 2013

Timeline: ~6 months per project

Auxiliary diagnostics system — MSI + camera + one other

Profile reconstruction — within 20%*

Theory incorporation — diffusivity within 20%*

long-term behavior within 50%* — Wide range characterization

Now | Fall 2021 | Spring 2022 | Fall 2022 | Spring 2022

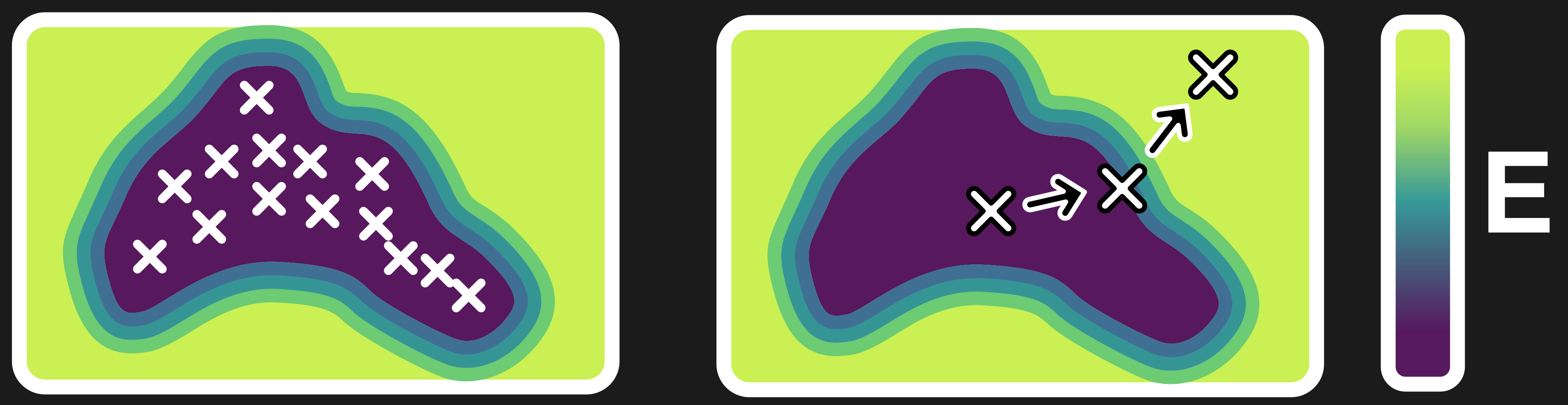
- Contingency: can do simpler analysis manually (not high-dimensional)
  - do human learning instead

Project review: ML can help us out with characterizing LAPD transport

- Augment auxiliary diagnostics system, build LAPD ML model, incorporate theory
- Why use ML anyways for a transport study?
  - Current paradigm: isolate variables
  - 10 actuators, 5 values for each = 10^5 experiments
  - Some actuators / confounders cannot be isolated
  - ML / generative models learn correlations & structure
  - Can sample from this structure

Side project: making good quality plasmas in strong mirrors

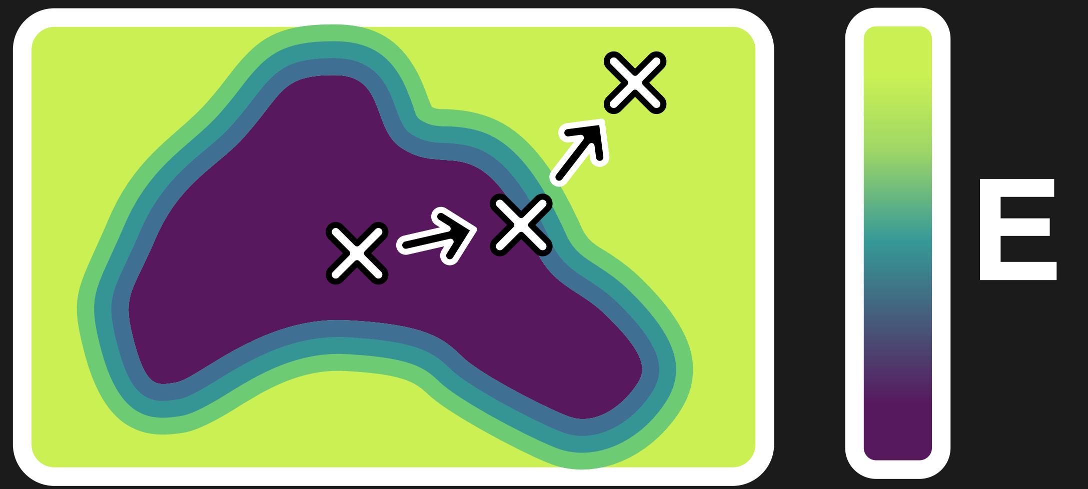
- At high fields, plasma quality suffers
- Find regions “good” regions of parameter space



Sampling **Exploring**

Side project: novelty search

- Sample unknown space
- Vary some external actuator (antenna, biasing)
- Actively seek out bad predictions / novelty
- Holy grail: prediction and detection of something unexpected / unknown



Machine parameter space

ML will become increasingly important

- Some fields consider ML a pillar of science alongside theory, simulation, and experiment
- Autonomous characterization of fusion devices* could radically speed up performance
- Integrate learned plasma model with natural language processing

OpenAI's DALL-E (Ramesh 2021)

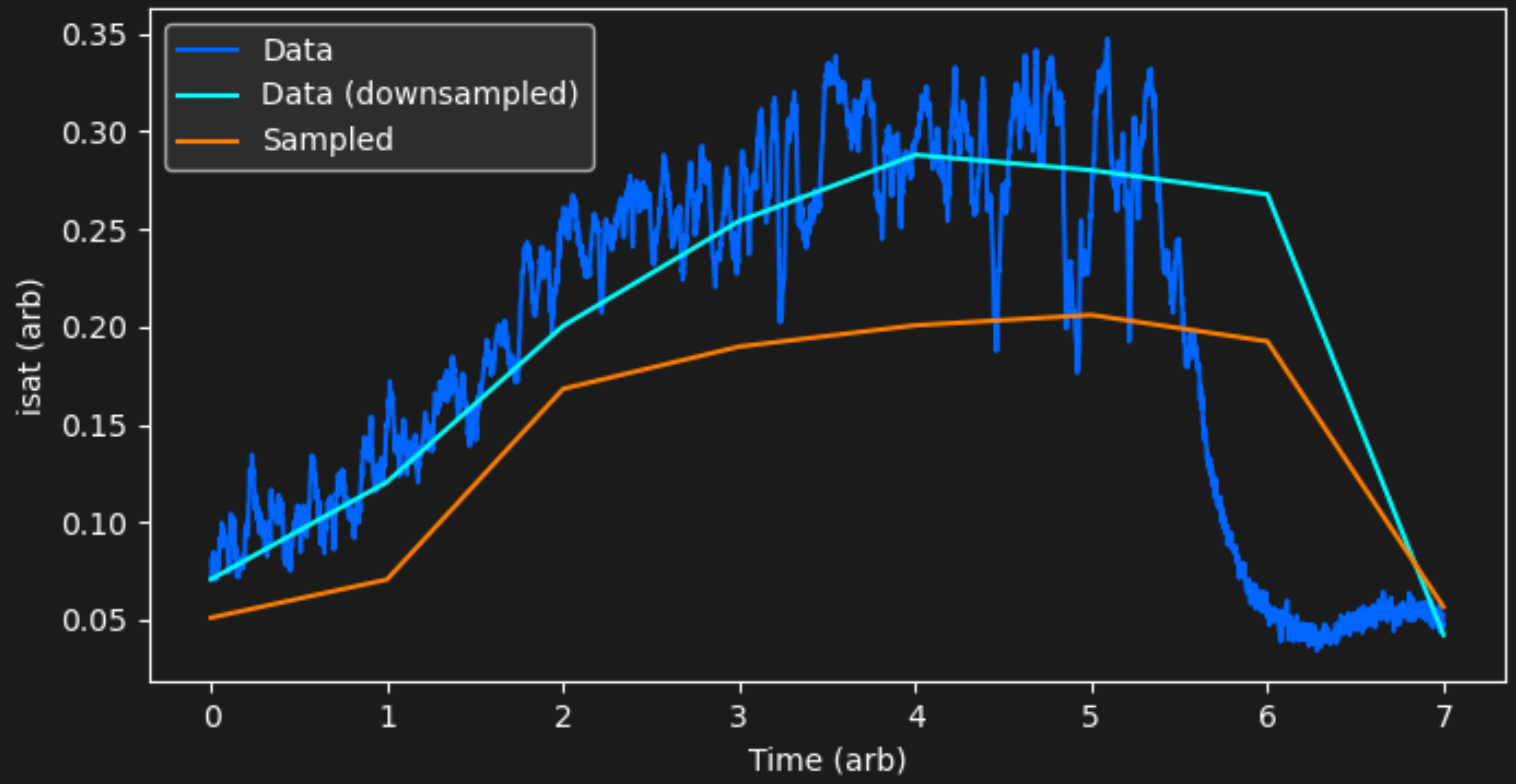
Autonomous learning

Emergent tool use from multi-agent autocurricula
Baker et al. 2020

Demonstration of learned EBM on LAPD data — free sampling

- 10 dimensional: position, mirror ratio, 8 time series points
- Sample from EBM: $s \sim p(S)$

Dataset:
7575 shots
M = 1, 1.47, 1.9, 2.3, 2.68
x = 0 - 50 cm, dx = 0.5 cm

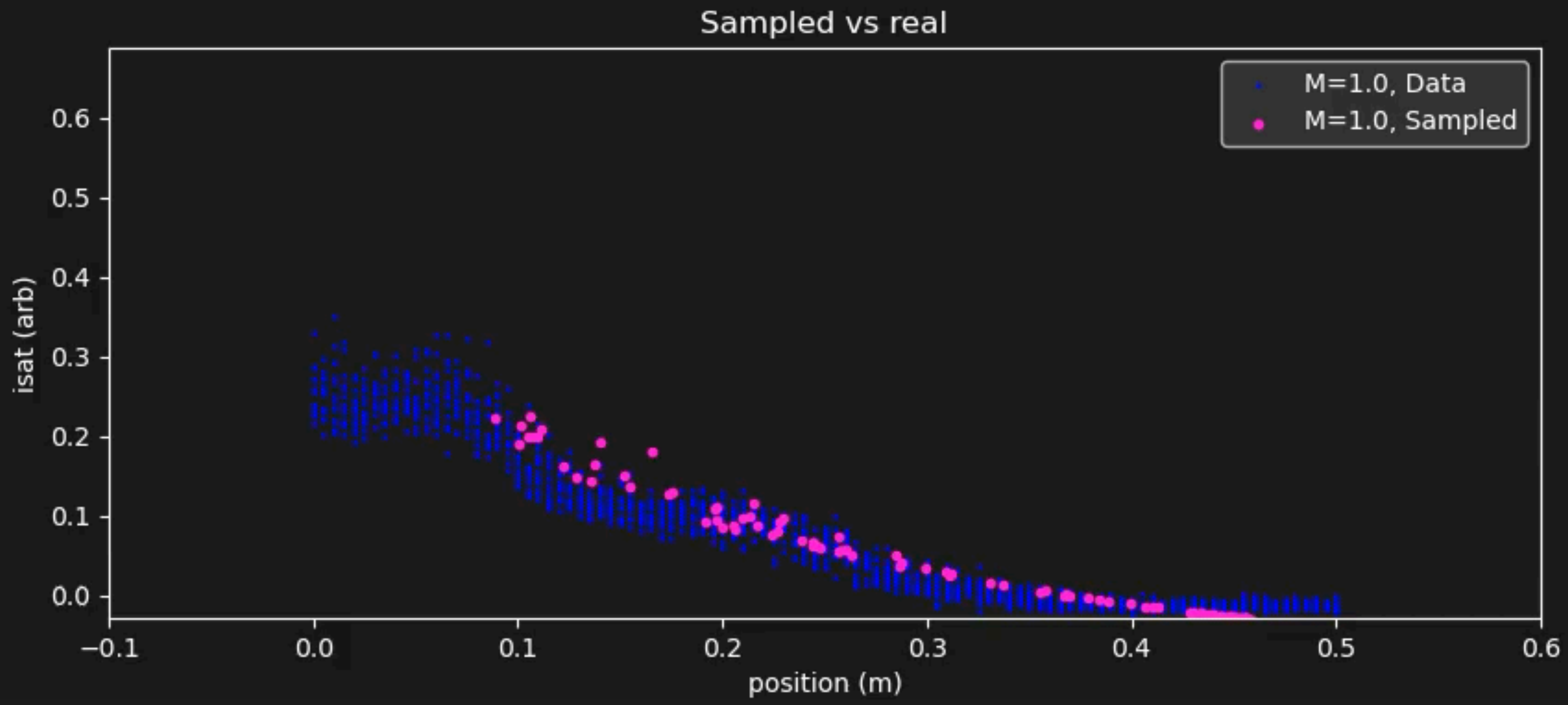


Data: M=1.9 at 33.5 cm
Sampled: M=1.7 at 33.4 cm

Used IGEBM-style architecture
(Du et al. 2020)

Conditional sampling on LAPD data $s \sim p(s \mid M = 1)$

- Model figures out correlations between M, position, and time series



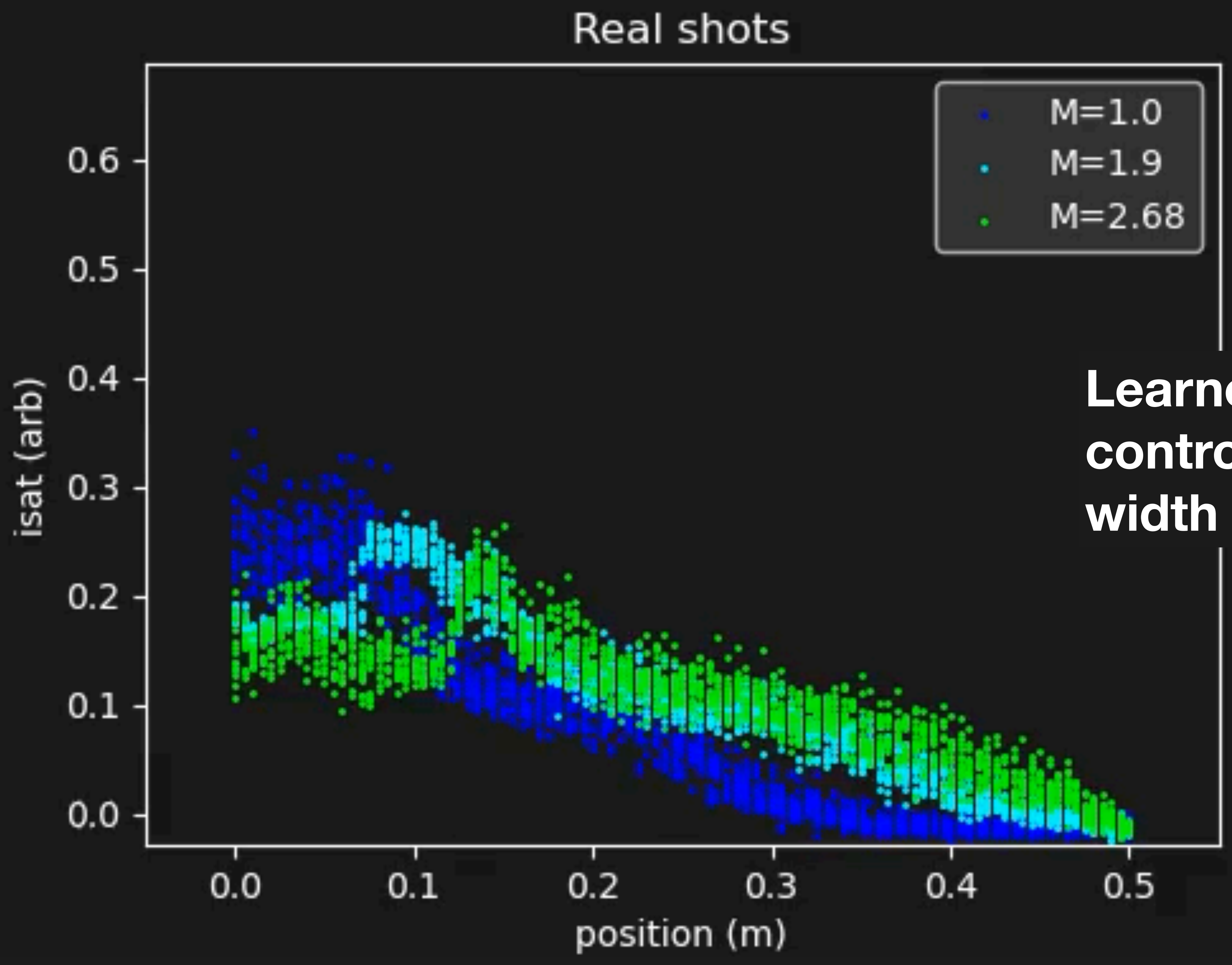
Learned the variance

Tracks profile in time

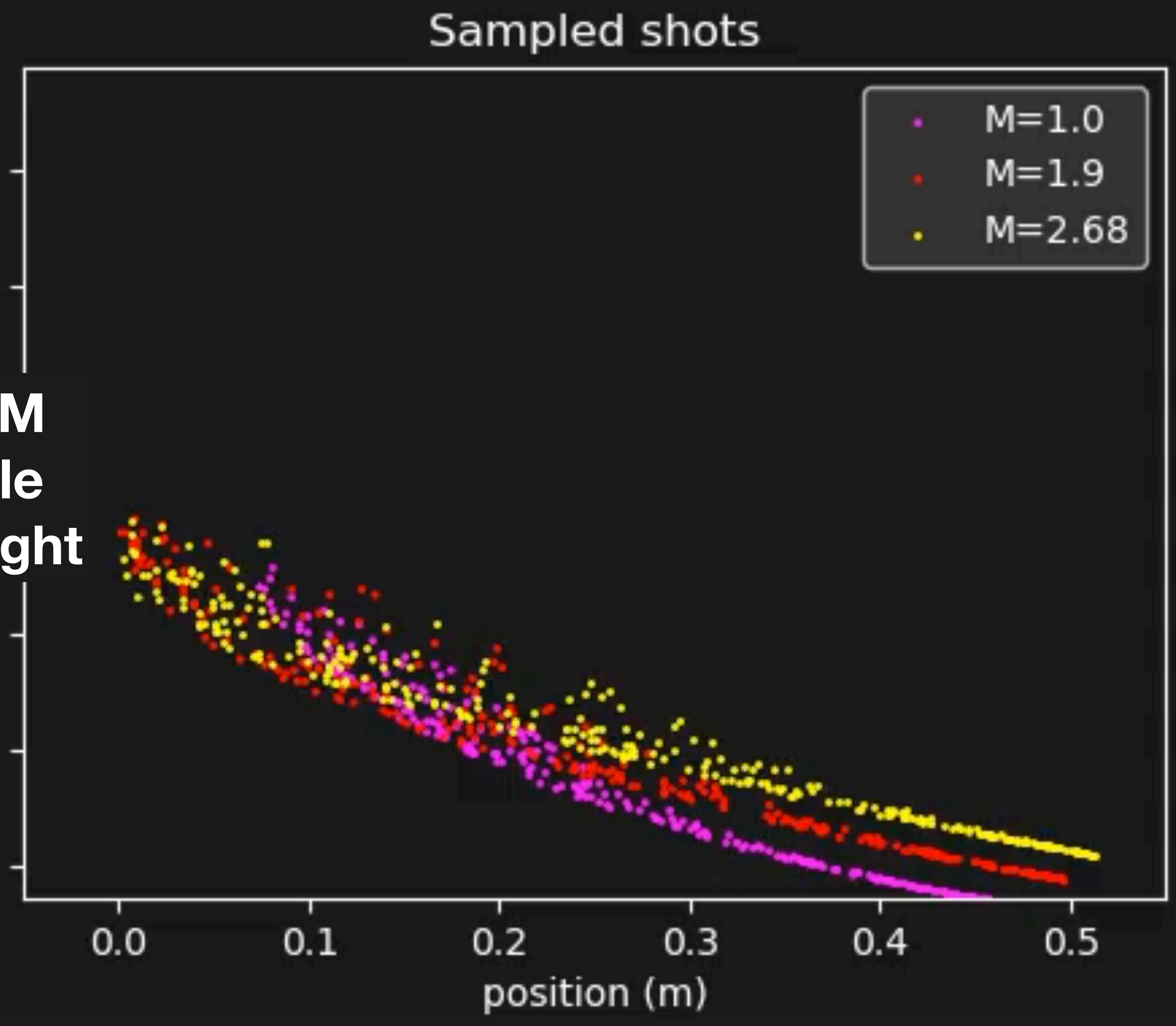
Conditional sampling on LAPD data

$$s_m \sim p(s \mid M = m)$$

- Model figures out correlations between M, position, and time series



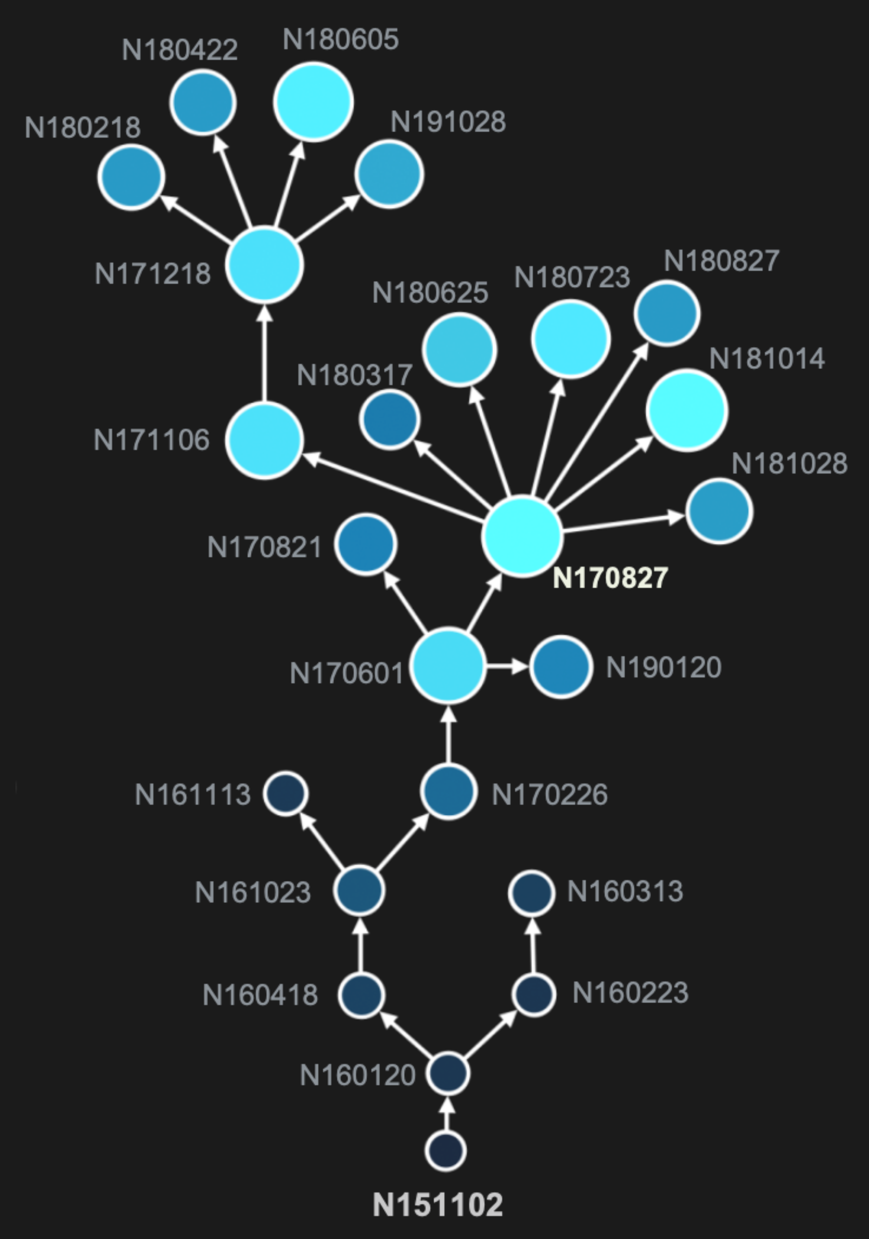
Learned that M controls profile width and height



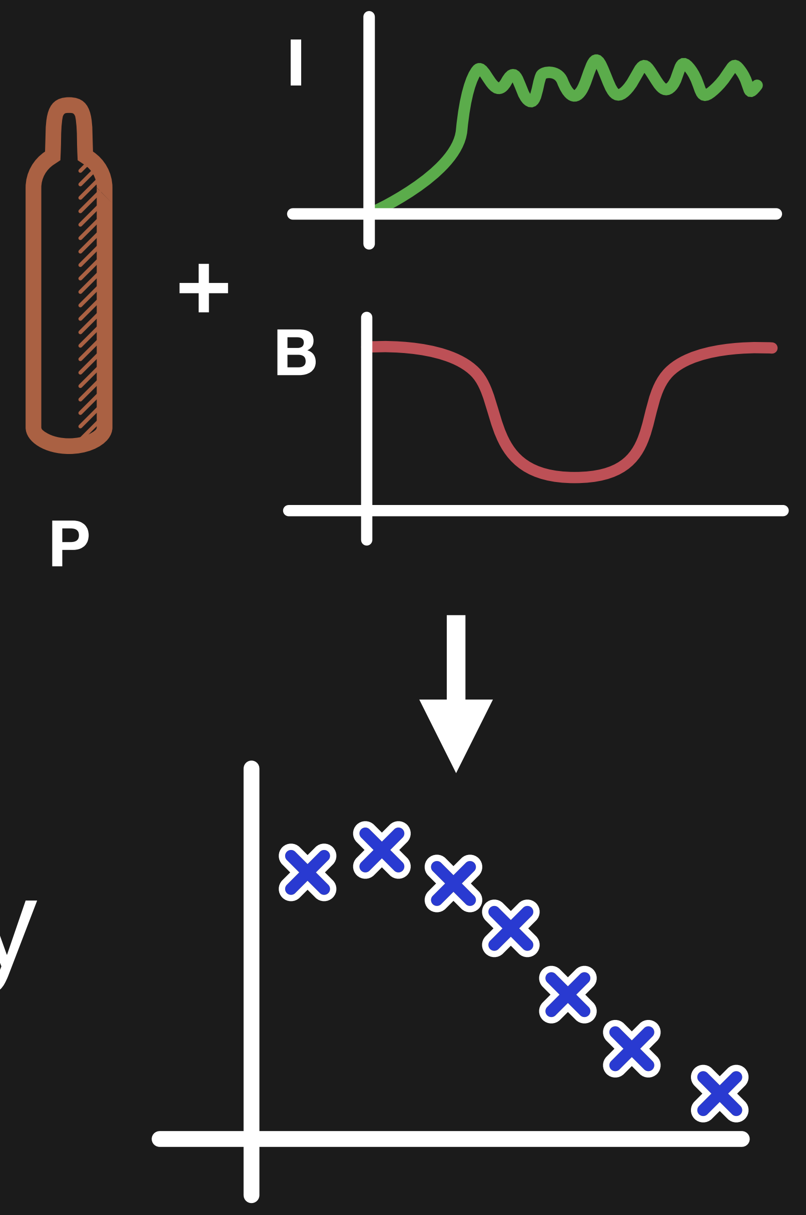
Towards automating fusion science

- ML is a new prediction paradigm

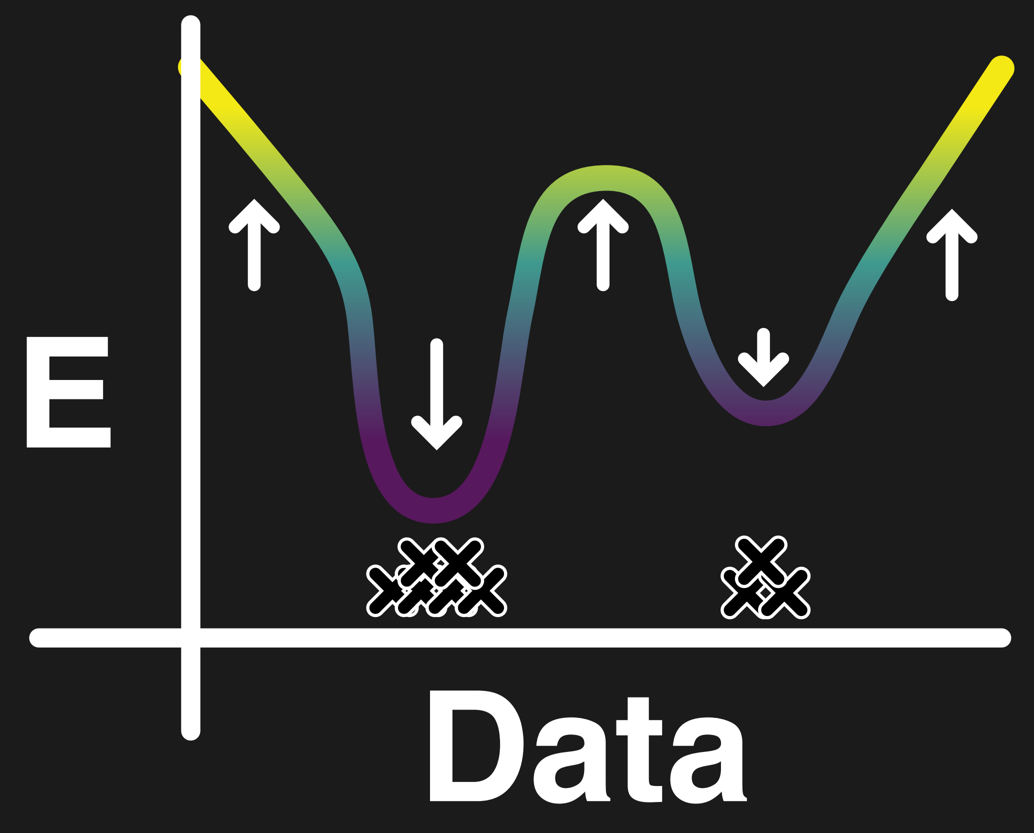
- ML-enabled methods may help alleviate our shortcomings



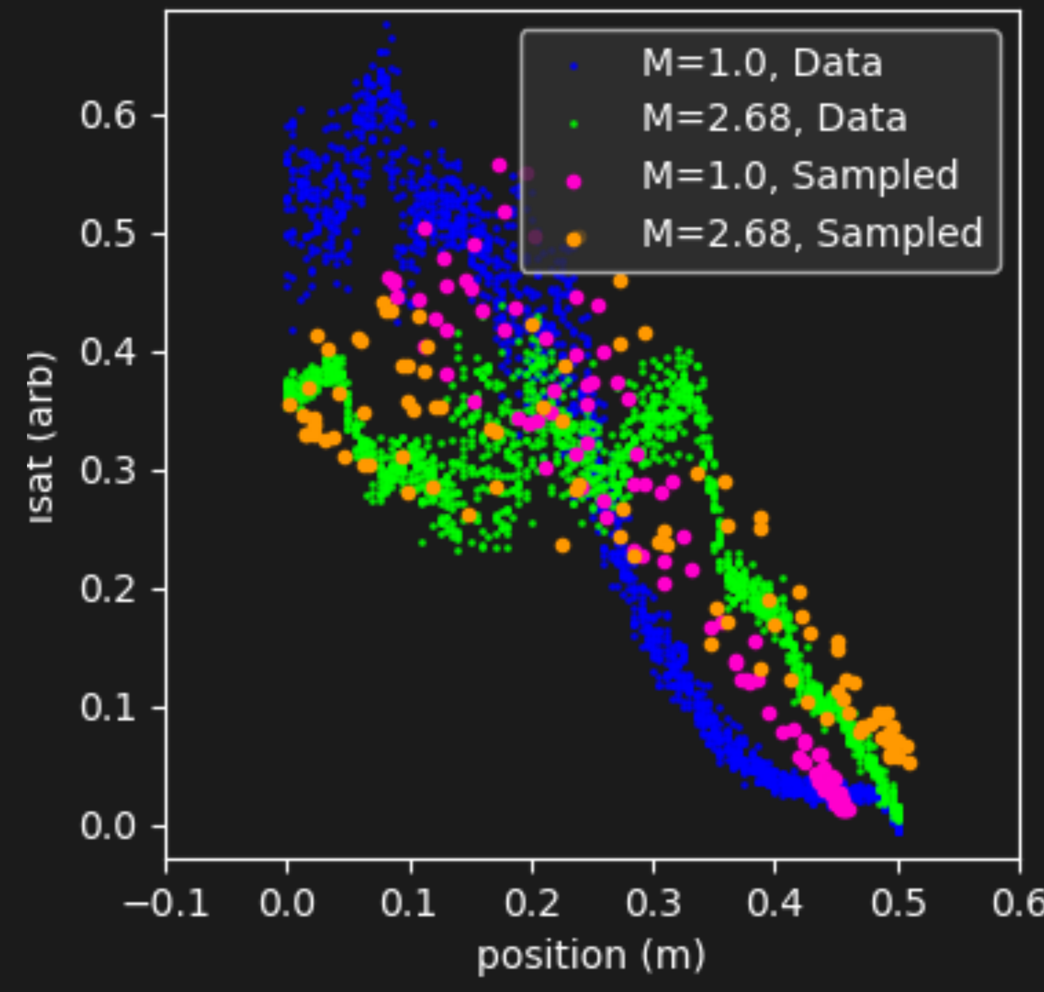
- Plan:
  augment diagnostics,
  predict profiles,
  integrate theory,
  perform expansive study



- LAPD is a great testbed for EBMs + fusion



- Recent work is encouraging



Backup slides

Question: how is using an EBM better than Gaussians on data points?

- In high dimensional spaces, gaussians will have a massive amount of probability mass in not useful directions
- How about we have a covariance matrix for these gaussians?
  - giant covariance matrix, painful / intractable to train
  - —> learn reduced representation, then place Gaussians? VAE with extra steps

Question: how do you sample evenly with uneven data coverage?

- EBM learns distribution of configurations
  - want value at any point: use conditional sampling
- Could penalize energy gradient wrt to an input during training
- Could also add a very large bias via adding an energy function just for, say, magnetic field configuration