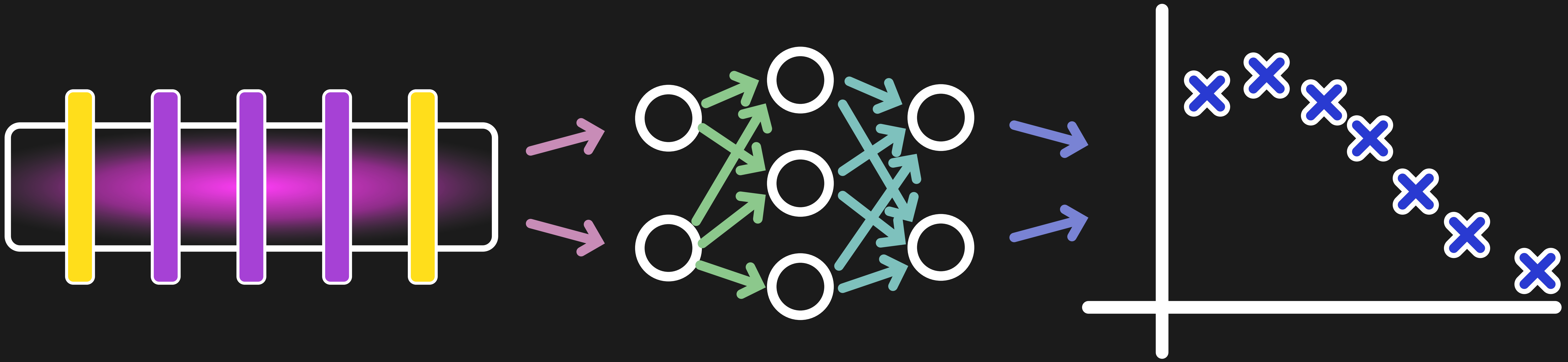
Generating synthetic LAPD discharges using energy-based models (EBMs)



Phil Travis*, Steve Vincena

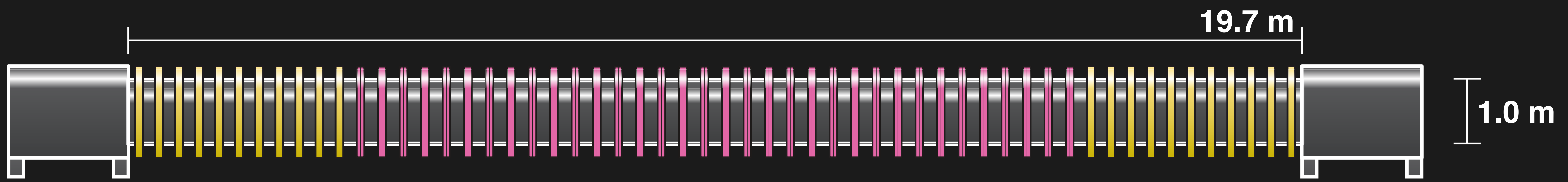
*phil@physics.ucla.edu

UCLA Basic Plasma Science Facility

(Thanks: Gurleen Bal, Tom Look, Yhoshua Wug, **Troy Carter**)

Work supported by the US DOE and NSF, and performed at the Basic Plasma Science Facility (BaPSF), UCLA.

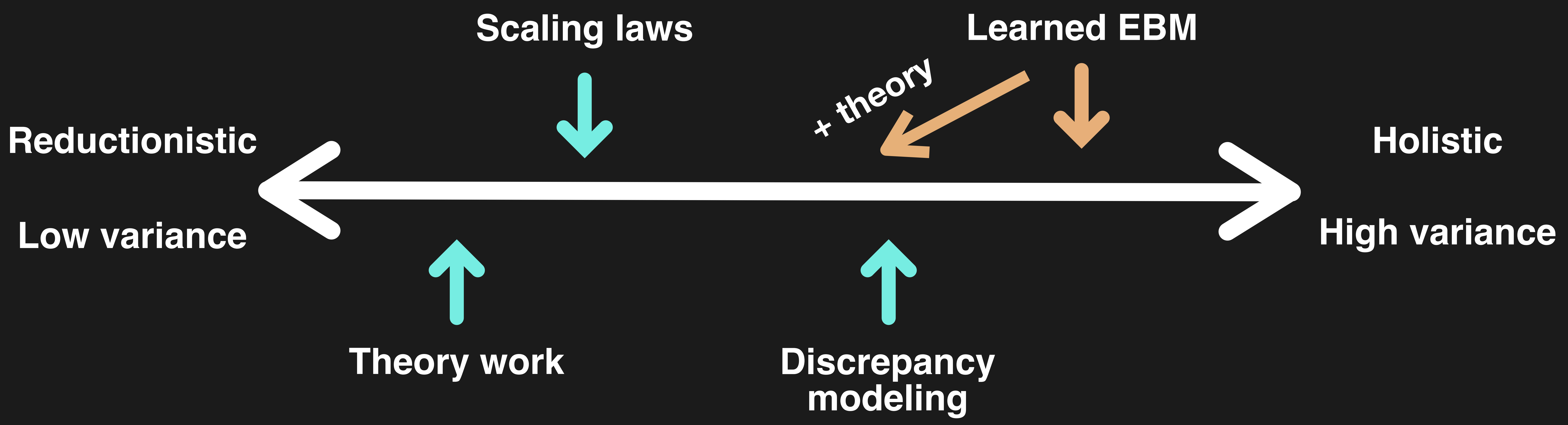
The Large Plasma Device (LAPD) is flexible and accessible



- $T_e \sim$ 5-10 eV, $n_e \sim 10^{13}$
- Wide variety of basic & space plasma experiments
- Up to ~31m shots per year: data rich
- **This work: collecting MSI & diagnostics**

- Machine state information (MSI)
  - Discharge I & V, B profile, gas pressure, RGA
- Diagnostics
  - Interferometer, visible light diodes
  - Fast framing camera
  - Probes (high spatial resolution)

A high-variance approach to learning from plasmas can be useful



- All effects accounted for in prediction
  - Discharges may contain currently unexploitable information
- Model has few preconceived notions (low bias)
- Can introduce bias (theoretical models) later

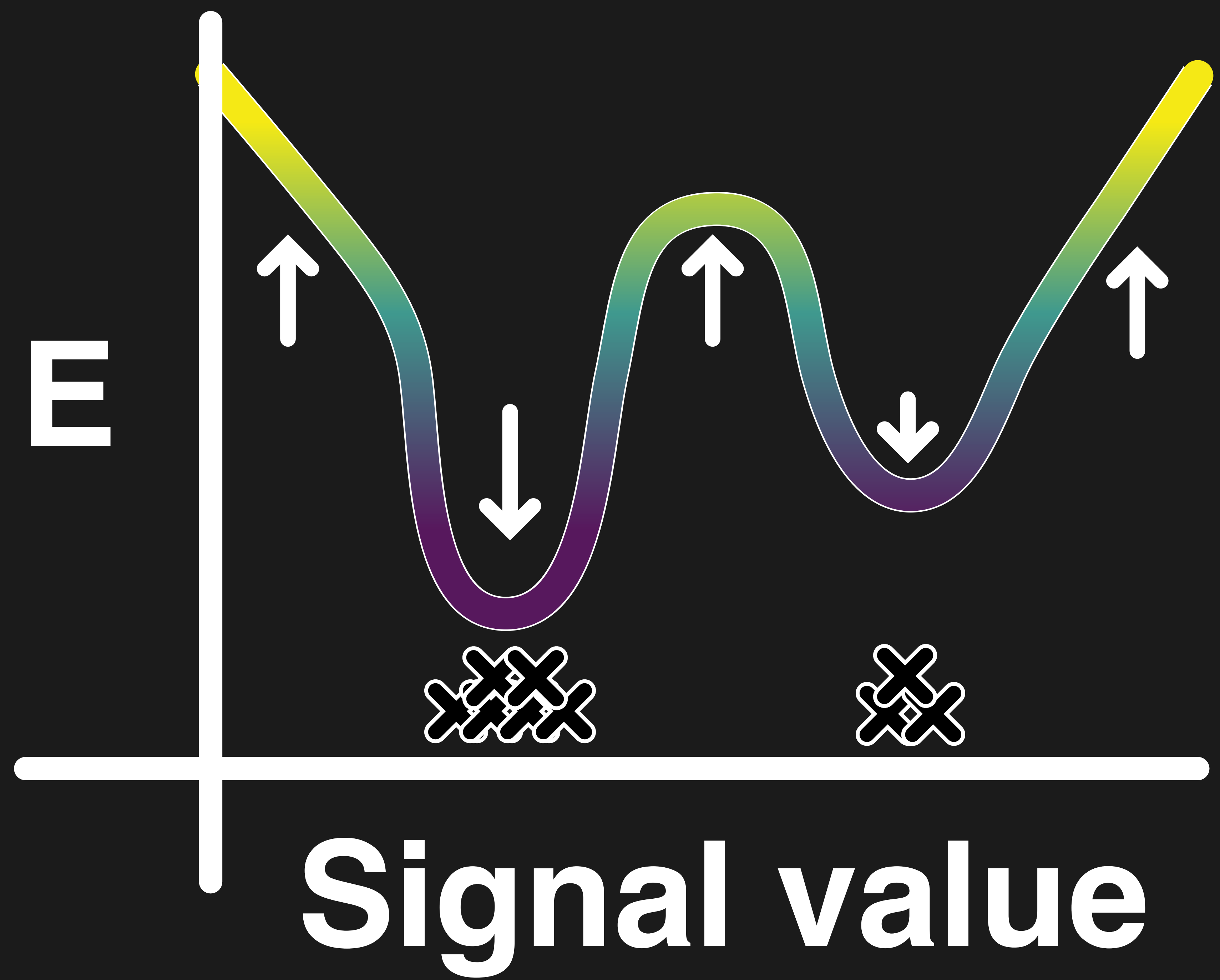
Energy-based models (EBMs) have useful qualities for science tasks

$$p(x) \sim e^{-\beta E(x)}$$

Data → NN → E

$dim(x) = 1715$

$=$ total # of samples



- Generative: learns a probability distribution
  - Predicts everything from everything else
  - Conditional sampling is built-in
- Energies are composable — can add models together
- Doesn't suffer from mode collapse / spurious modes

Review: Deep Generative Modelling — Bond-Taylor et al. (2021) arXiv:2103.04922

Synthetic discharges can be sampled from a learned probability distribution

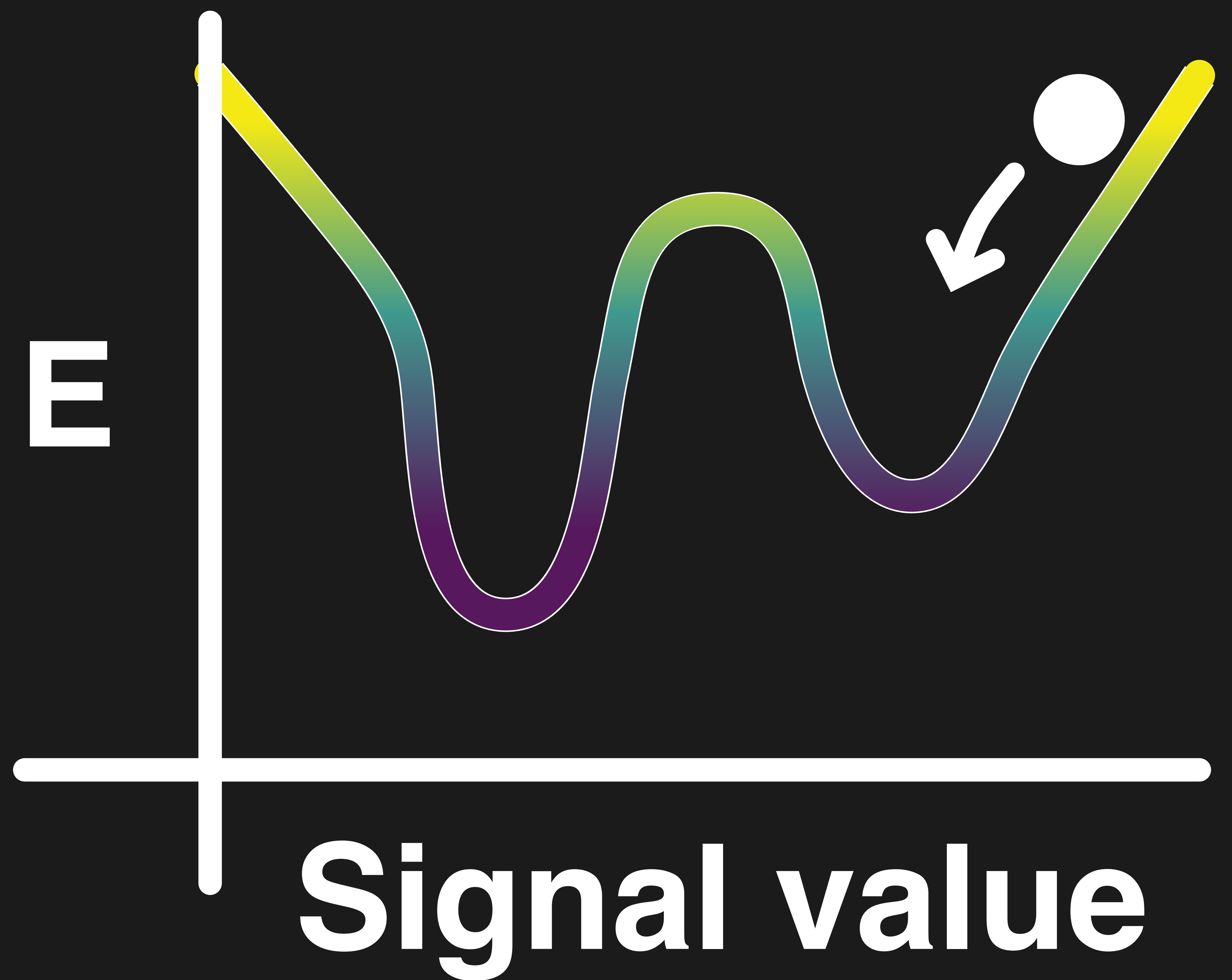
$$x \sim p(x) \qquad dim(x) = 1715$$

Langevin dynamics:

$$\ddot{x} = -\nabla E(x) + \sqrt{T\mathcal{N}(0,1)}$$

potential gradient

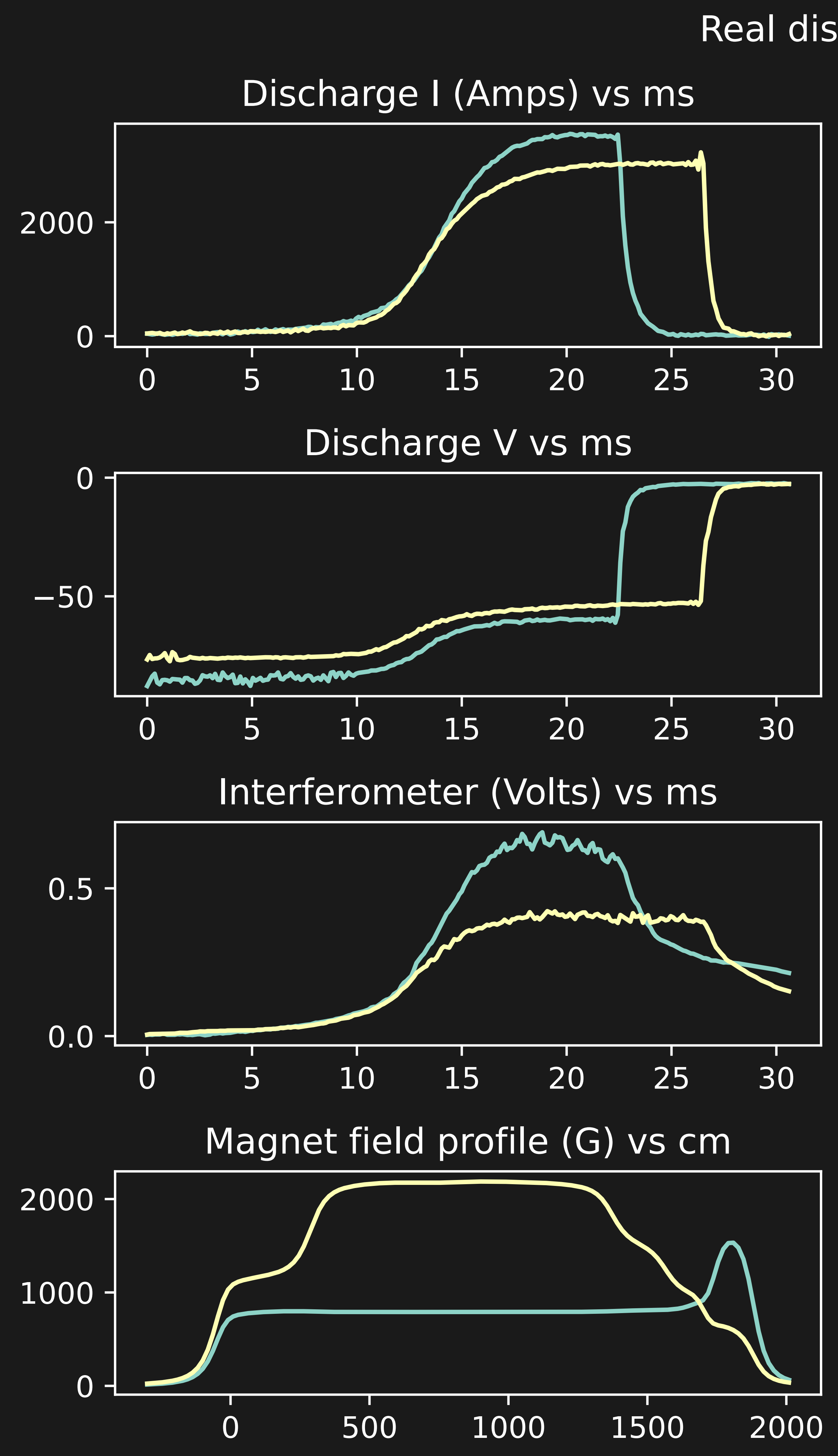
gaussian process

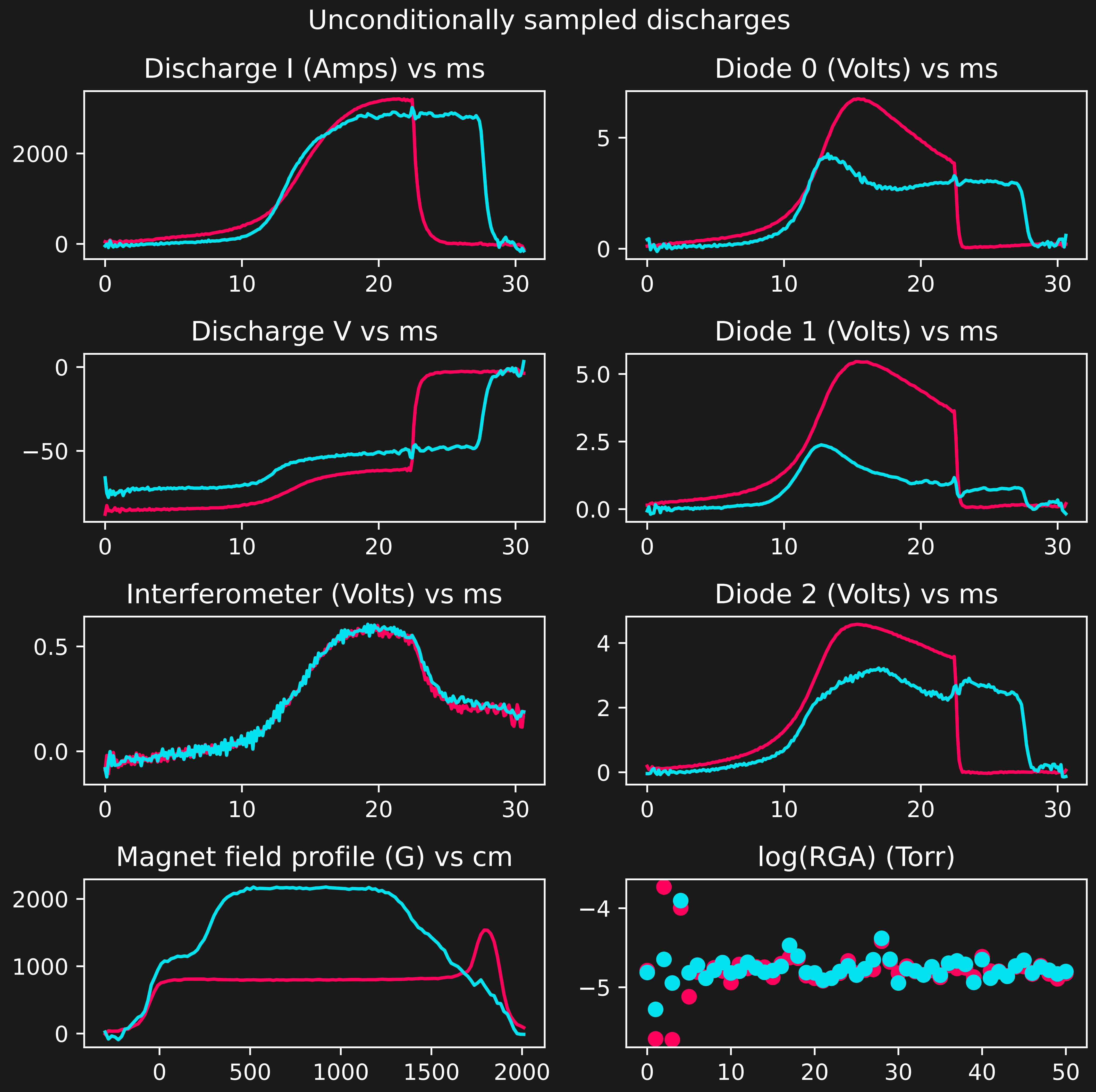


Du & Mordatch (2019) arXiv:1903.08689v6

Du et al. (2020) arXiv:2012.01316v4

Synthetic discharges can be sampled from a learned probability distribution



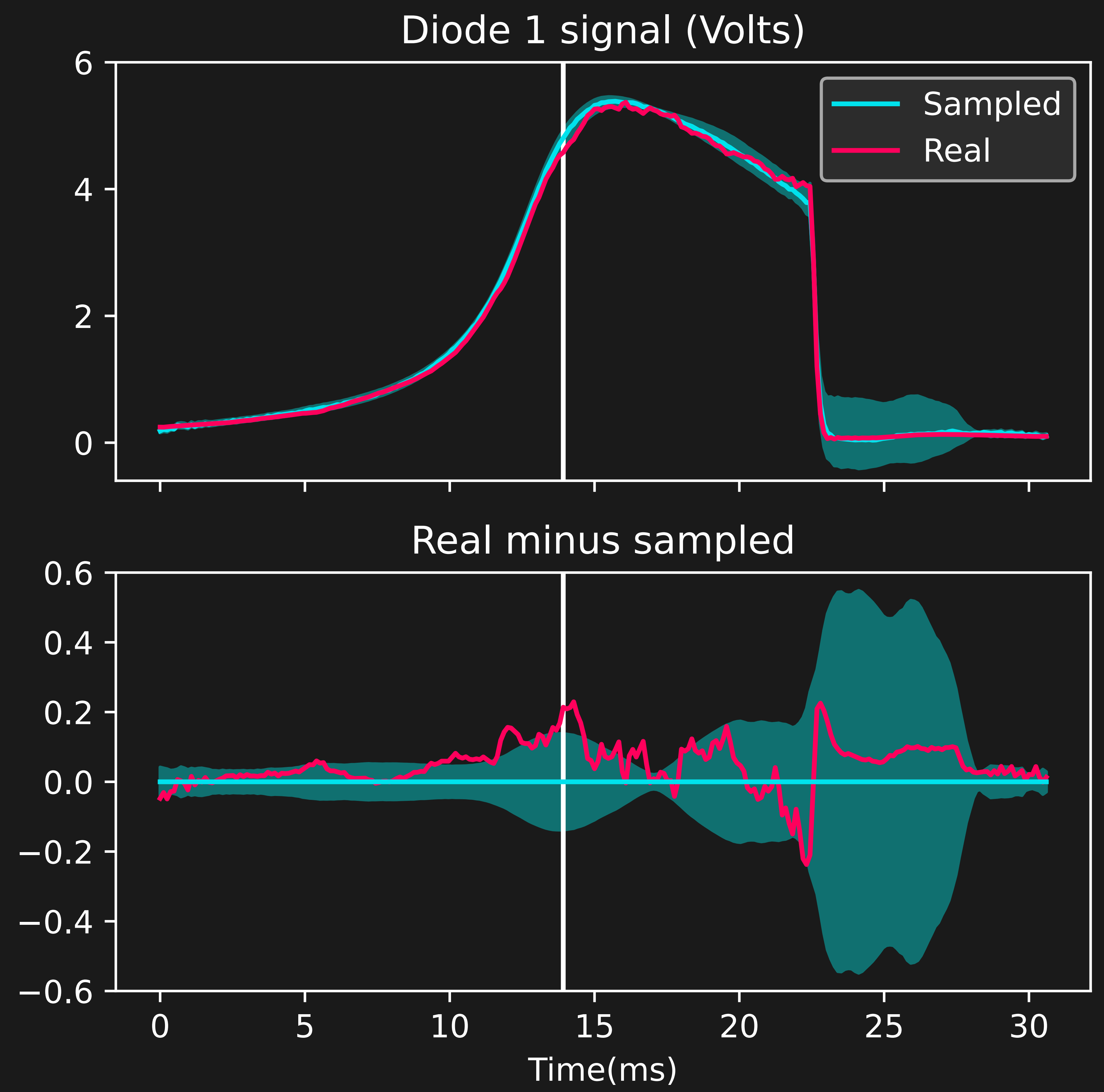


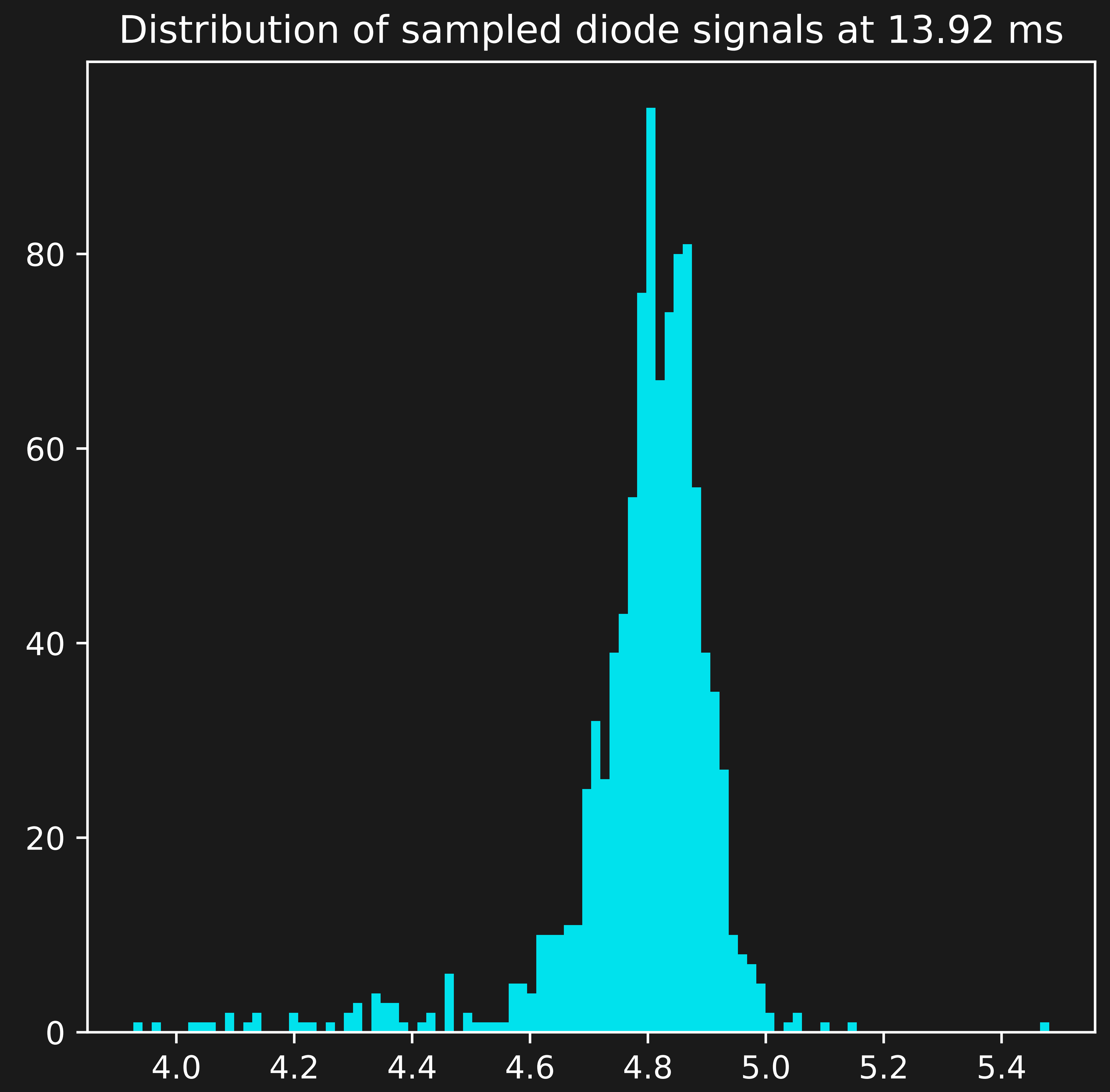
Conditional sampling can be performed to fill in missing signals

$$y \sim p(y \mid x)$$

x: discharge parameters

y: missing diode signal

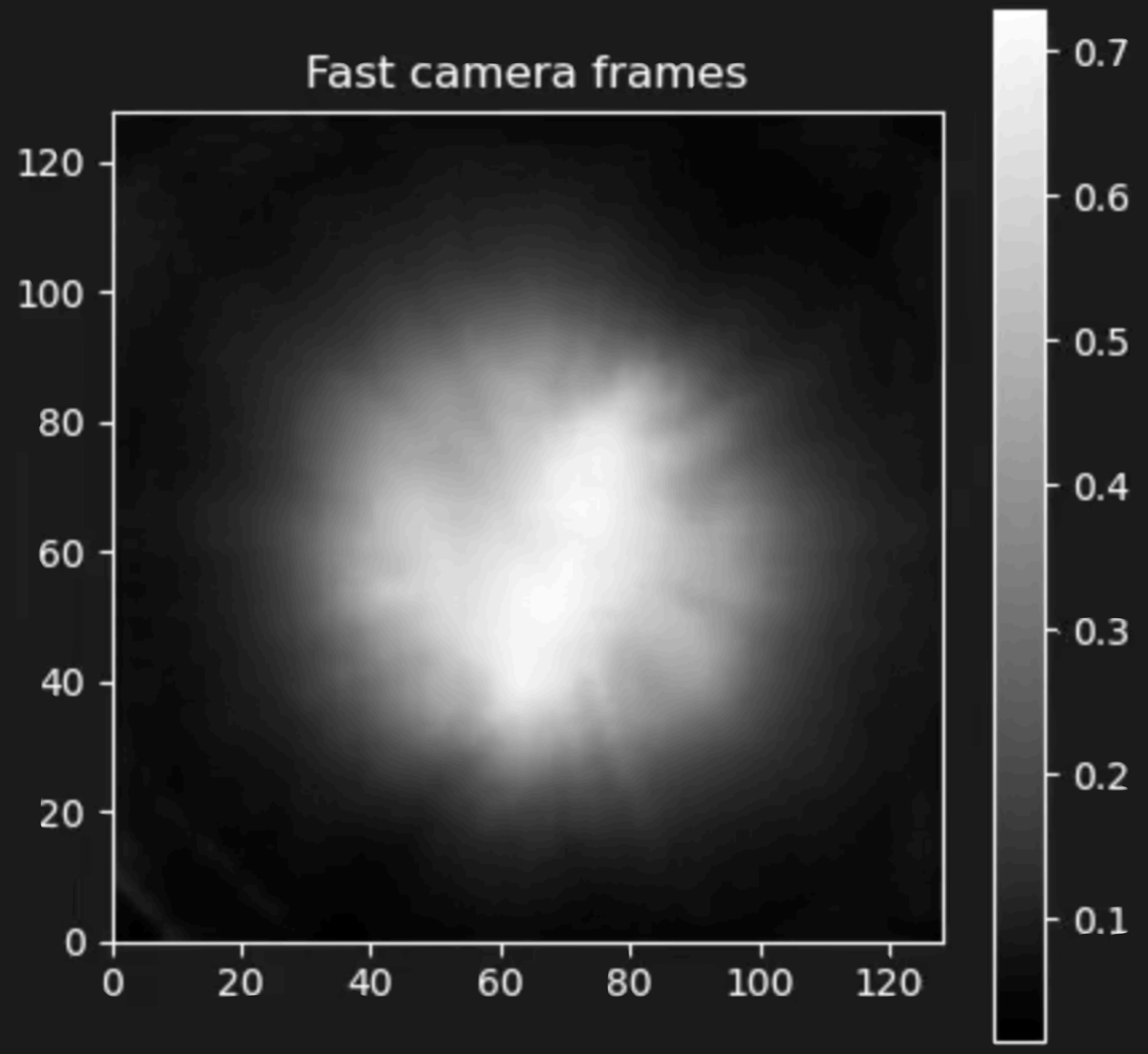




We get conditional sampling and a measure of uncertainty for free

Future work: accelerate science campaigns + find good operating regimes

- Add fast framing camera data, probe signals
  - Infer probe signals from other diagnostics
  - Infer 2d profiles in a few discharges
- New LAPD source characterization
  - Develop model architectures
  - Interrogate model
- Long term: use model to uncover correlations, form basis for theoretical work



Summary: EBMs can be used to sample discharges

- Data pipeline was constructed for machine state information and diagnostics
- Energy-based models (EBMs) learn a probability distribution
  - assigns an energy value to each input configuration
- EBMs can be unconditionally sampled to generate synthetic discharges
- EBMs can be conditionally sampled to fill in missing signals

phil@physics.ucla.edu

(thanks for not running off to lunch)