Advances in AI/ML for theoretical nuclear physics

Felix Ringer

Stony Brook University

EINN conference, Cyprus, 2025

Stony Brook University

Recent advances in AI/ML

- Two directions: General and single-purpose models

Language Models

- E.g.

ChatGPT

Recent advances in AI/ML

- Two directions: General and single-purpose models

Language Models

↓

Foundation models

- E.g. ChatGPT

- Multi modal

Recent advances in AI/ML

- Two directions: General and single-purpose models

Language Models

↓

Foundation models

↓

Agentic models

- E.g. ChatGPT

- Multi modal

Recent advances in AI/ML

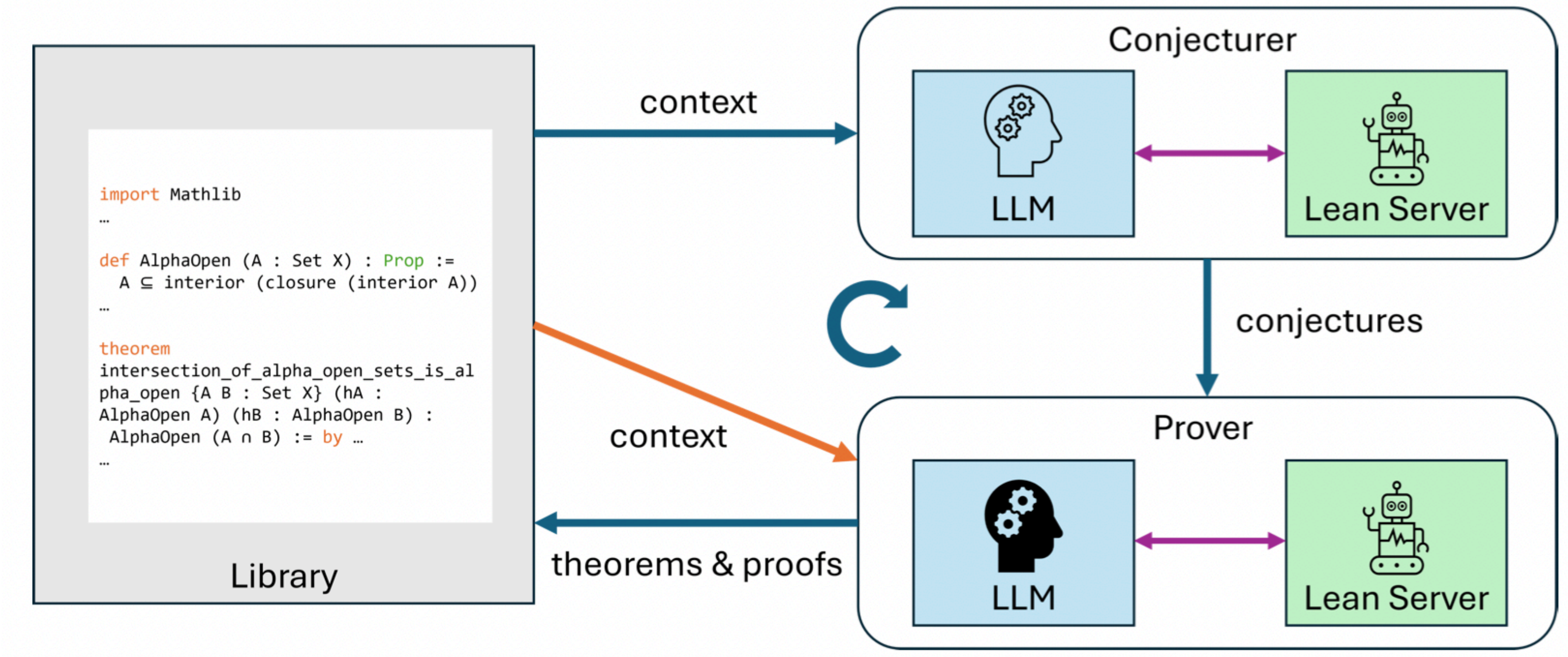
- Two directions: General and single-purpose models

Language Models

- E.g. ChatGPT

Foundation models

- Multi modal



Agentic models

- Scientific rigor? E.g. Math proofs with "Lean"

Recent advances in AI/ML

- Two directions: General and single-purpose models

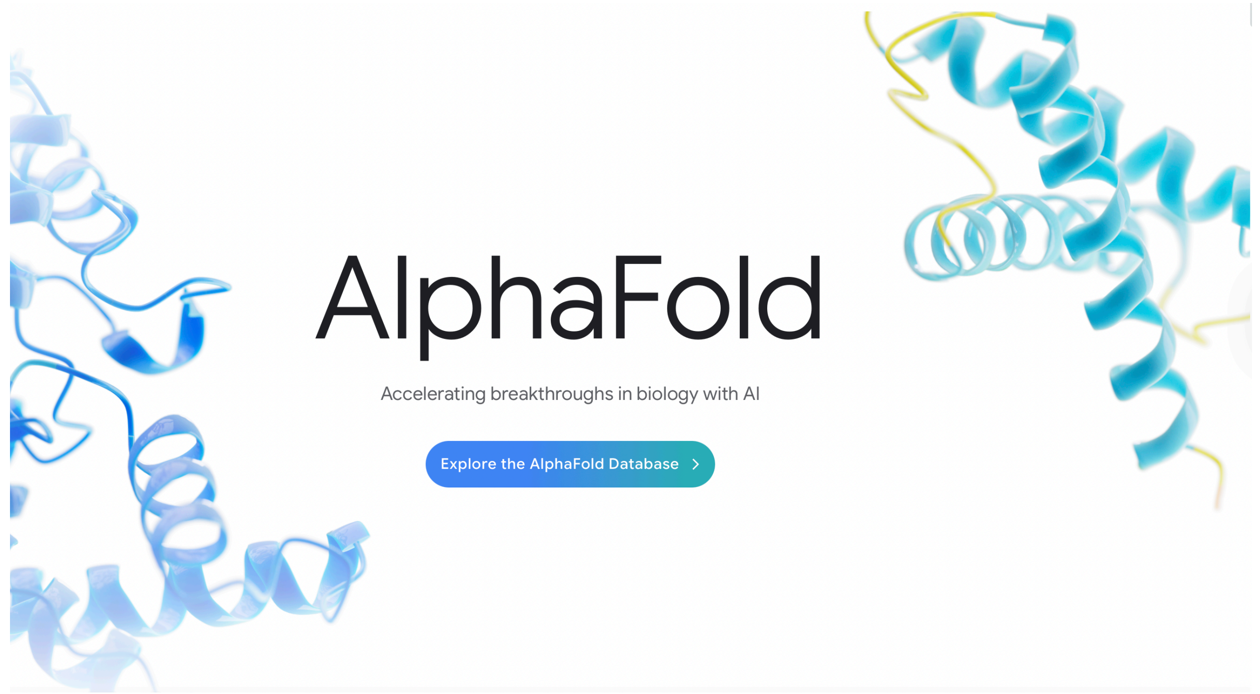
Language Models

Foundation models

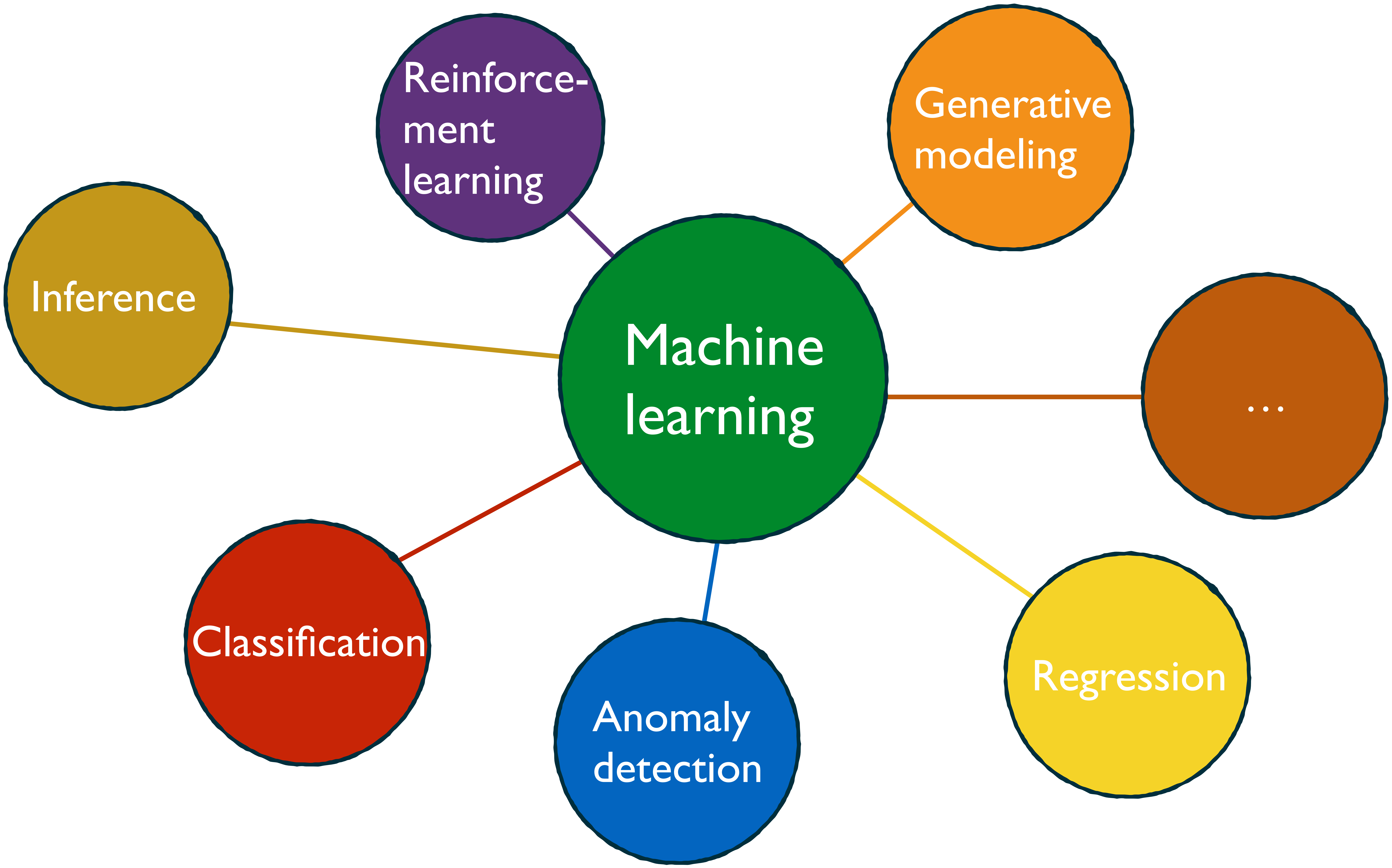
Agentic models

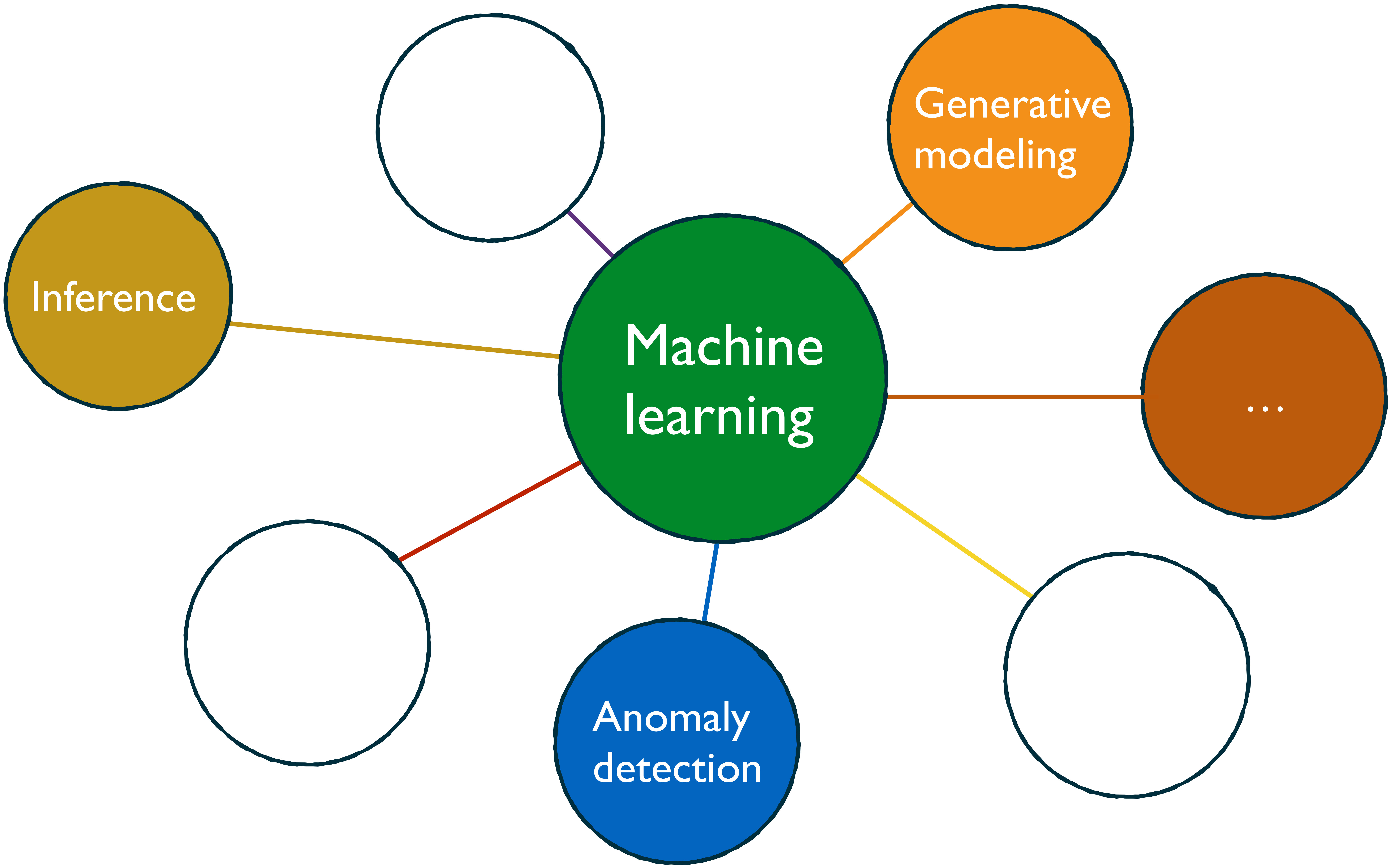
Single (few) purpose models, include expert knowledge

- Playing Go, Chess
- Alphafold



2024

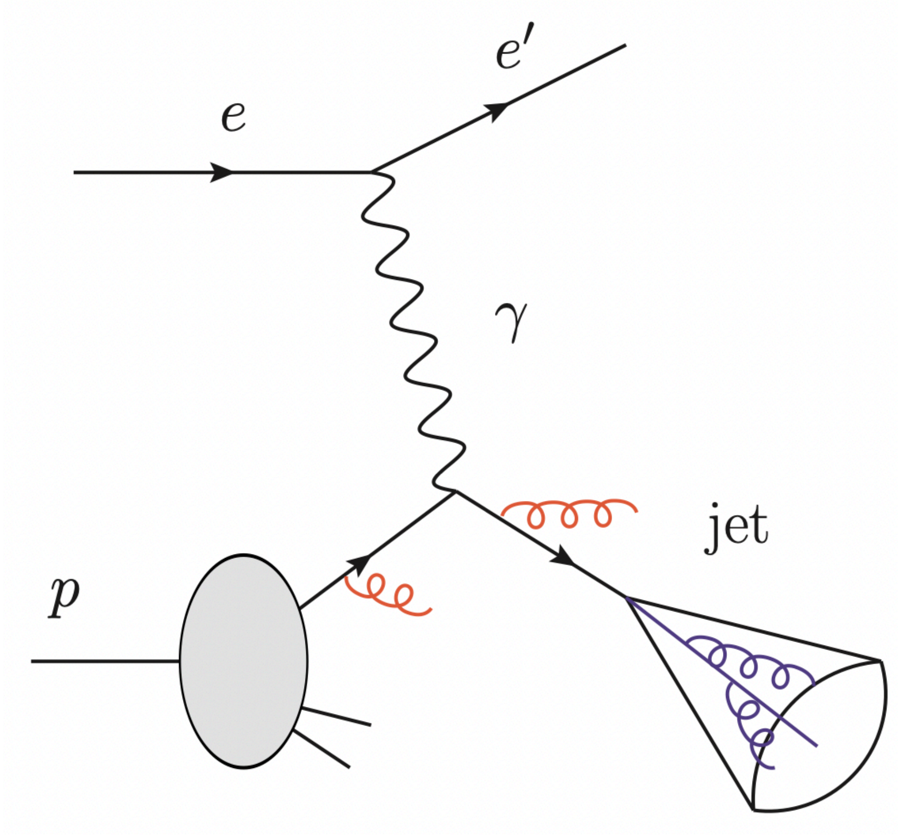

Reinforce-
ment
learning
Generative
modeling
Inference
Machine
learning
…
Classification
Anomaly
detection
Regression

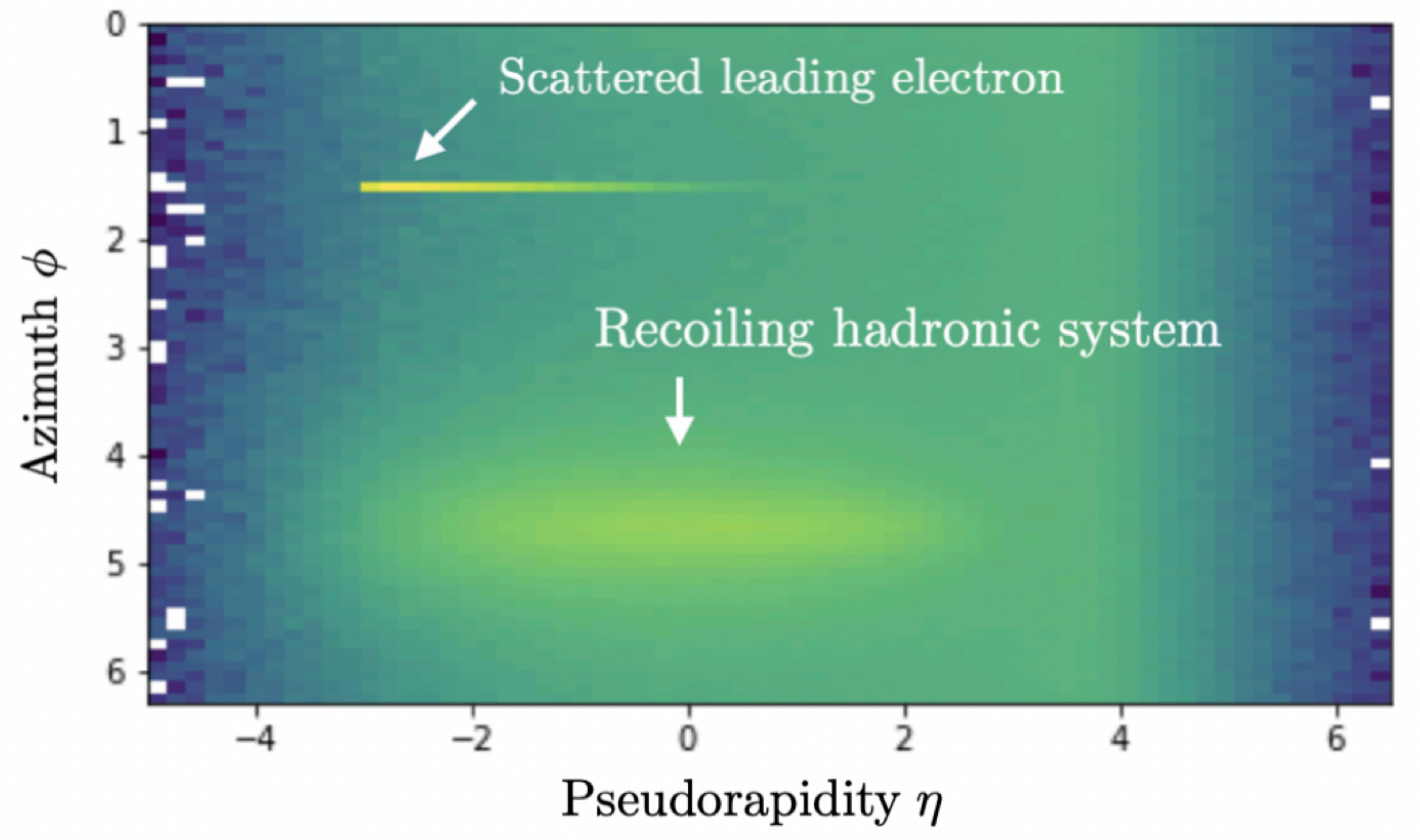

Generative modeling
Inference
Machine learning
...
Anomaly detection

Generative models for EIC events

- Full ep events at $\sqrt{s} = 105\ \mathrm{GeV}$

Develop a generative model





Applications:

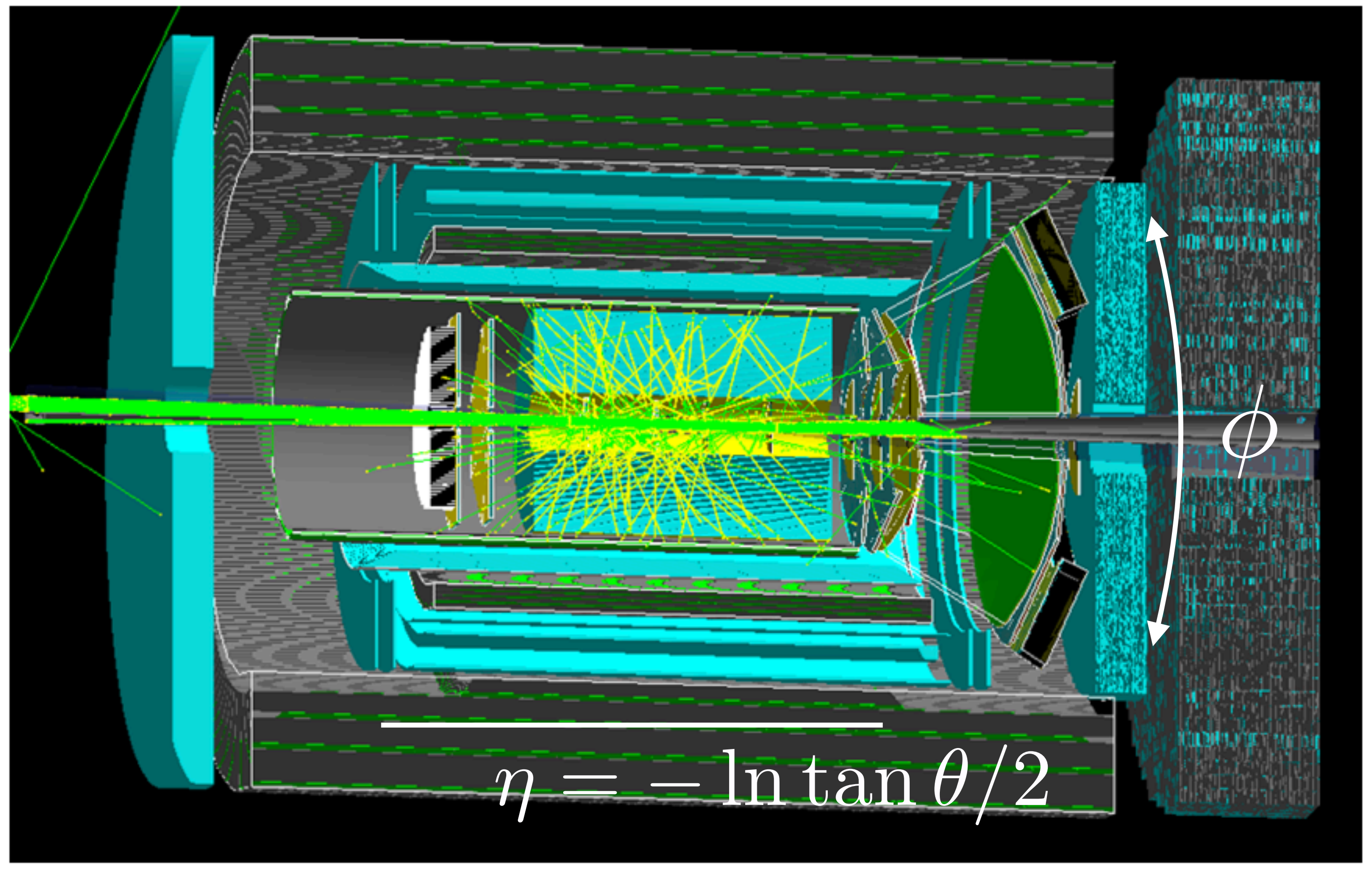
- Searches of physics beyond the Standard Model
- Neural Simulation-Based Inference (SBI) at the event level
- Unfolding

$\textsc{Pythia}8,\ Q > 5\ \mathrm{GeV}$

w/o photoproduction & CC

Devlin, Qiu, FR, Sato `23
see *also Arratia, Mikuni, Nachman et al.*

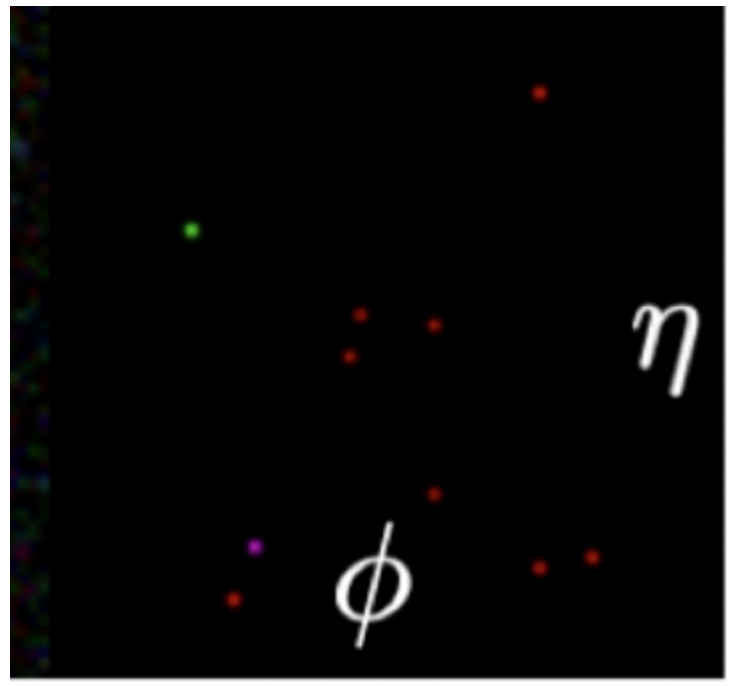
High-energy nuclear collisions



Data representations examples

- Images

EIC



Sparse

Devlin, Qiu, FR, Sato `23

- Point clouds

$$\text{Event}_1: \ (p_{T1}, \eta_1, \phi_1, \text{PID}_1), \ (p_{T2}, \eta_2, \phi_2, \text{PID}_2), \ldots, \ (p_{Tn}, \eta_n, \phi_n, \text{PID}_n)$$

Variable length

$$\text{Event}_2: \ (p_{T1}, \eta_1, \phi_1, \text{PID}_1), \ldots, \ (p_{Tm}, \eta_m, \phi_m, \text{PID}_m)$$

$$\vdots$$

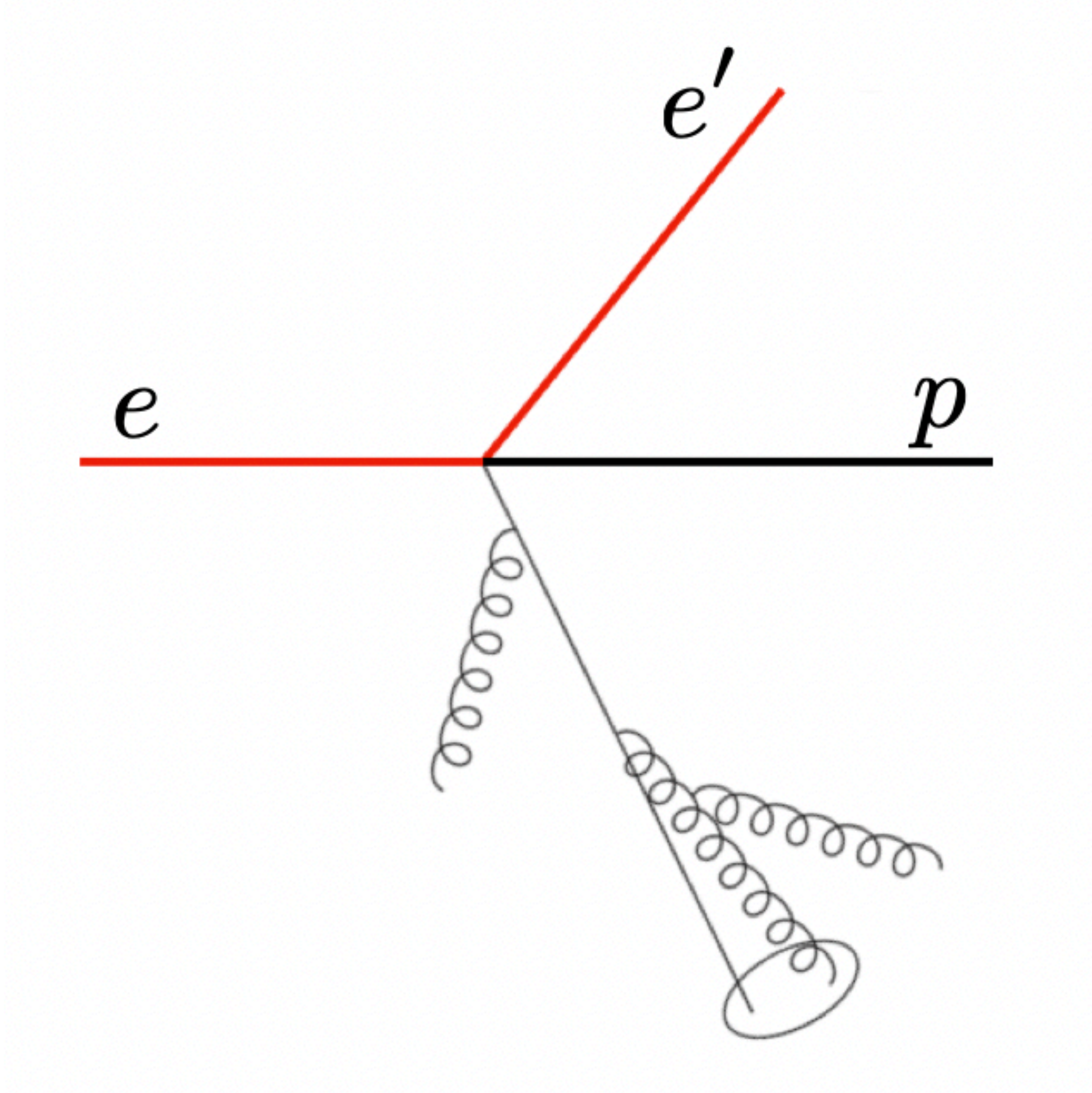
See also Arratia, Torales Acosta, Nachman, Mikuni et al, Kasieczka, Thaler et al.

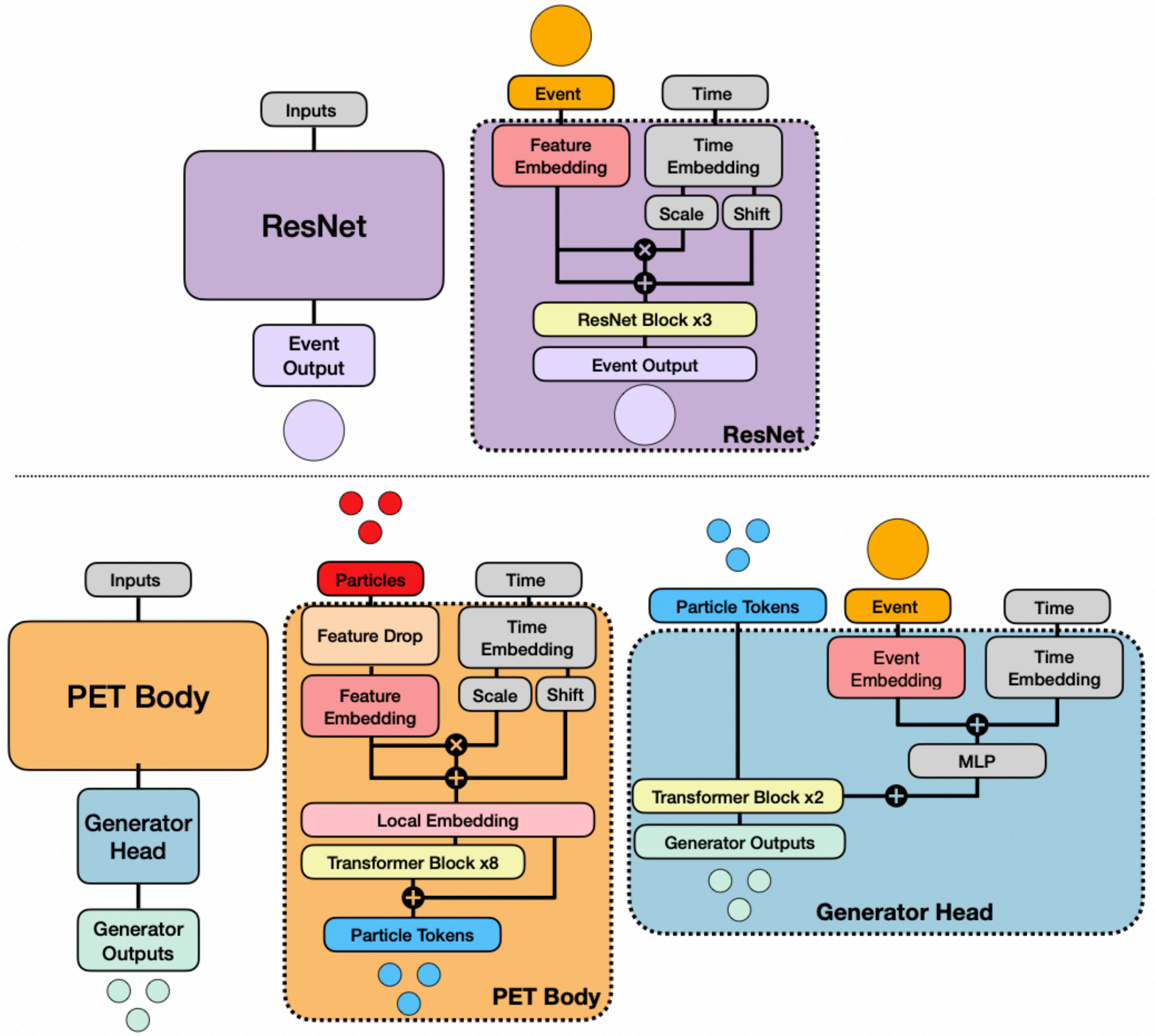
Diffusion models

- Point clouds (w/o pixelation)

Araz, Mikuni, FR, Sato, Torales-Acosta, Whitehill `24

Two-step diffusion process

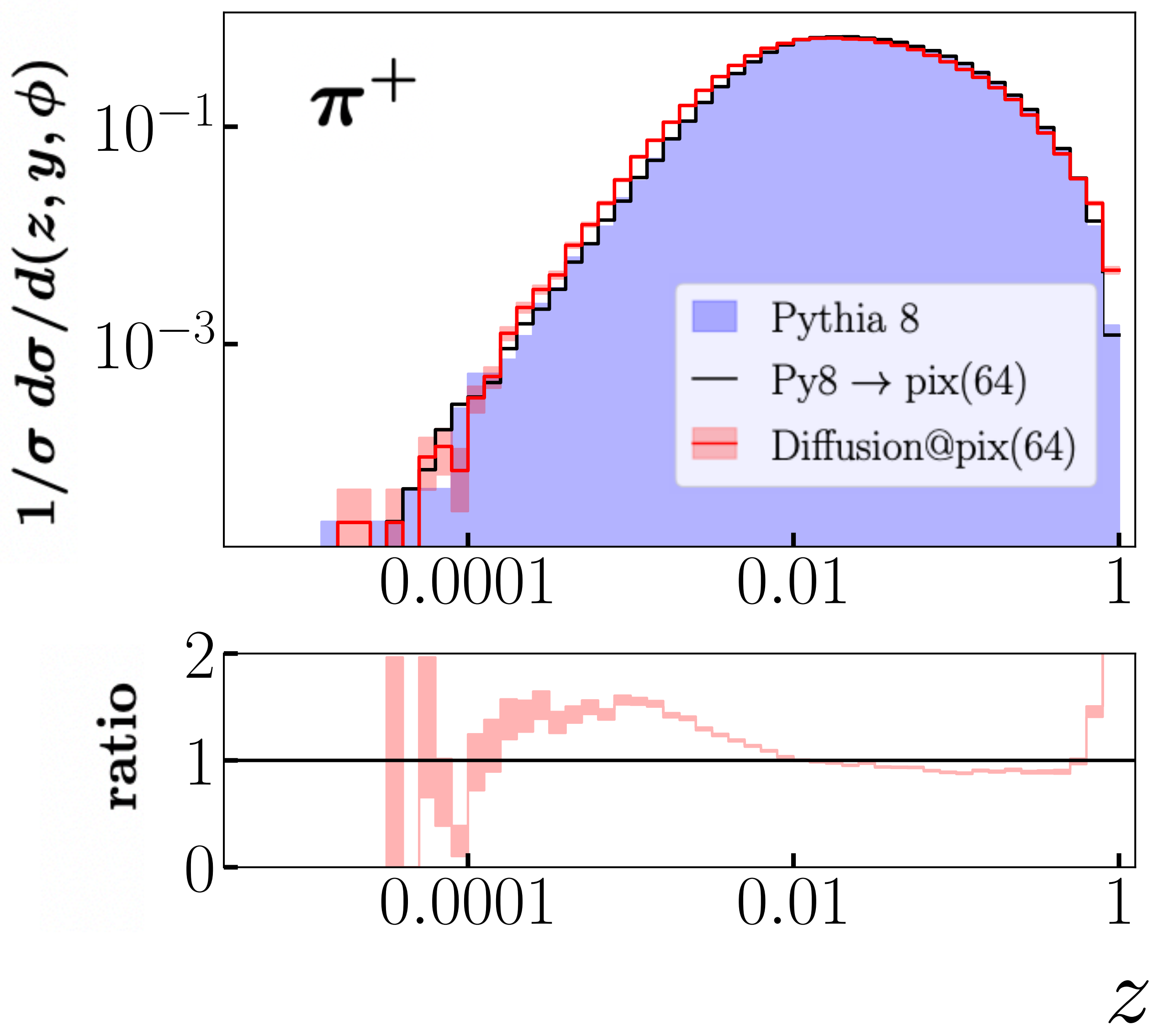




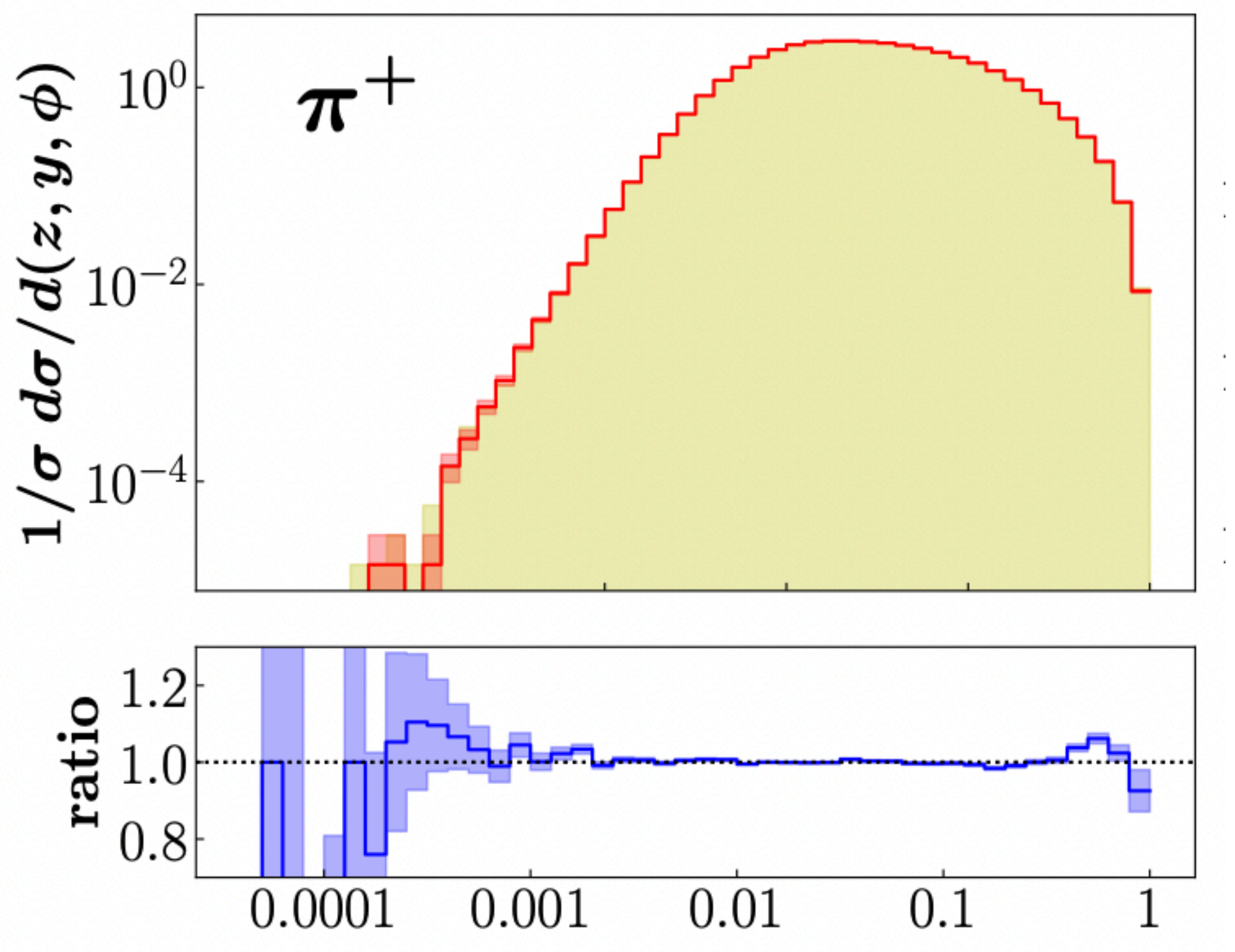
Diffusion models for the EIC

- E.g. pion momentum distributions

Araz, Mikuni, FR, Sato, Torales-Acosta, Whitehill `24



Pixelated images



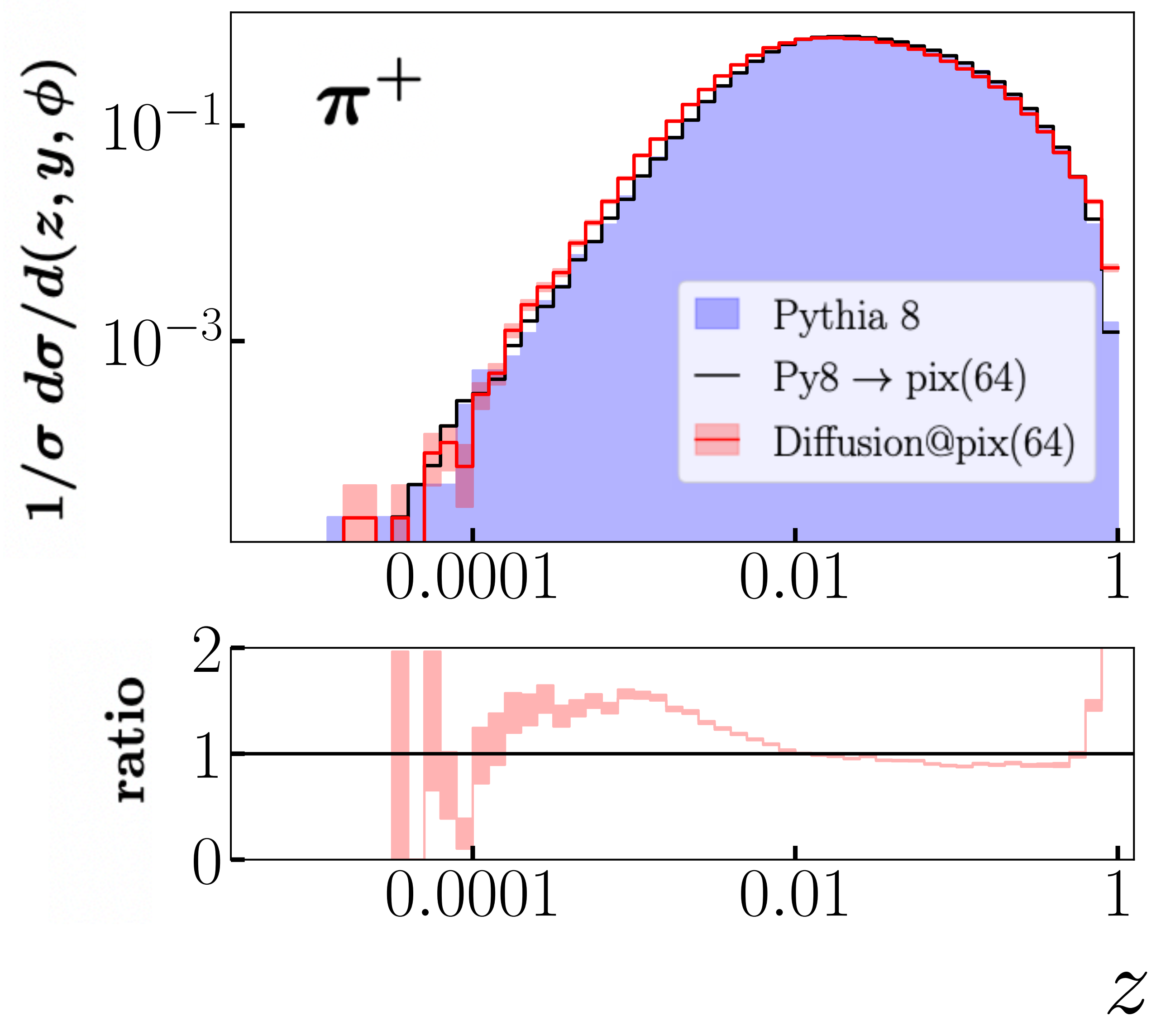
Point cloud

$$z = \frac{2p_T}{\sqrt{s}} \cosh \eta$$

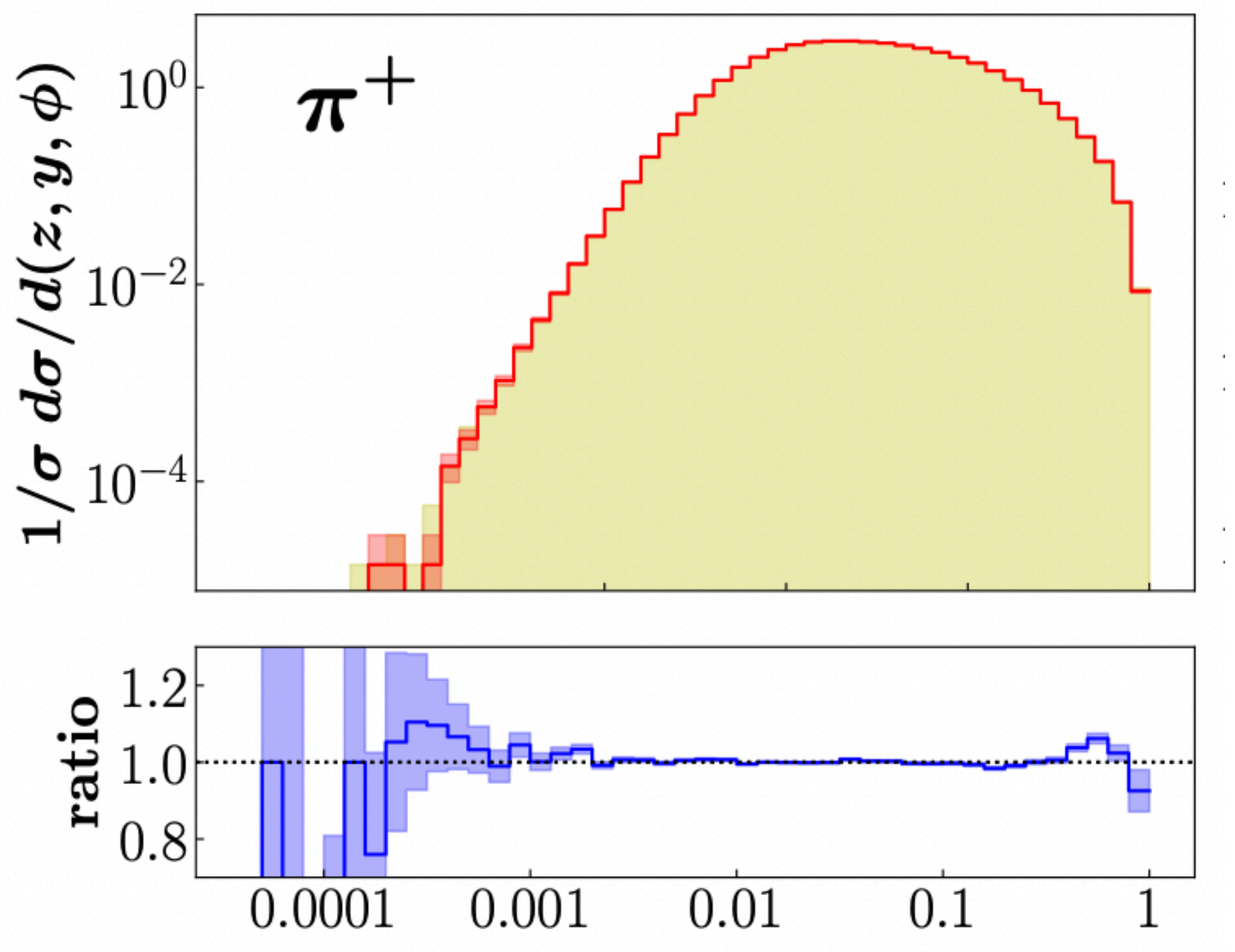
Diffusion models for the EIC

Araz, Mikuni, FR, Sato, Torales-Acosta, Whitehill `24

- E.g. pion momentum distributions



Pixelated images



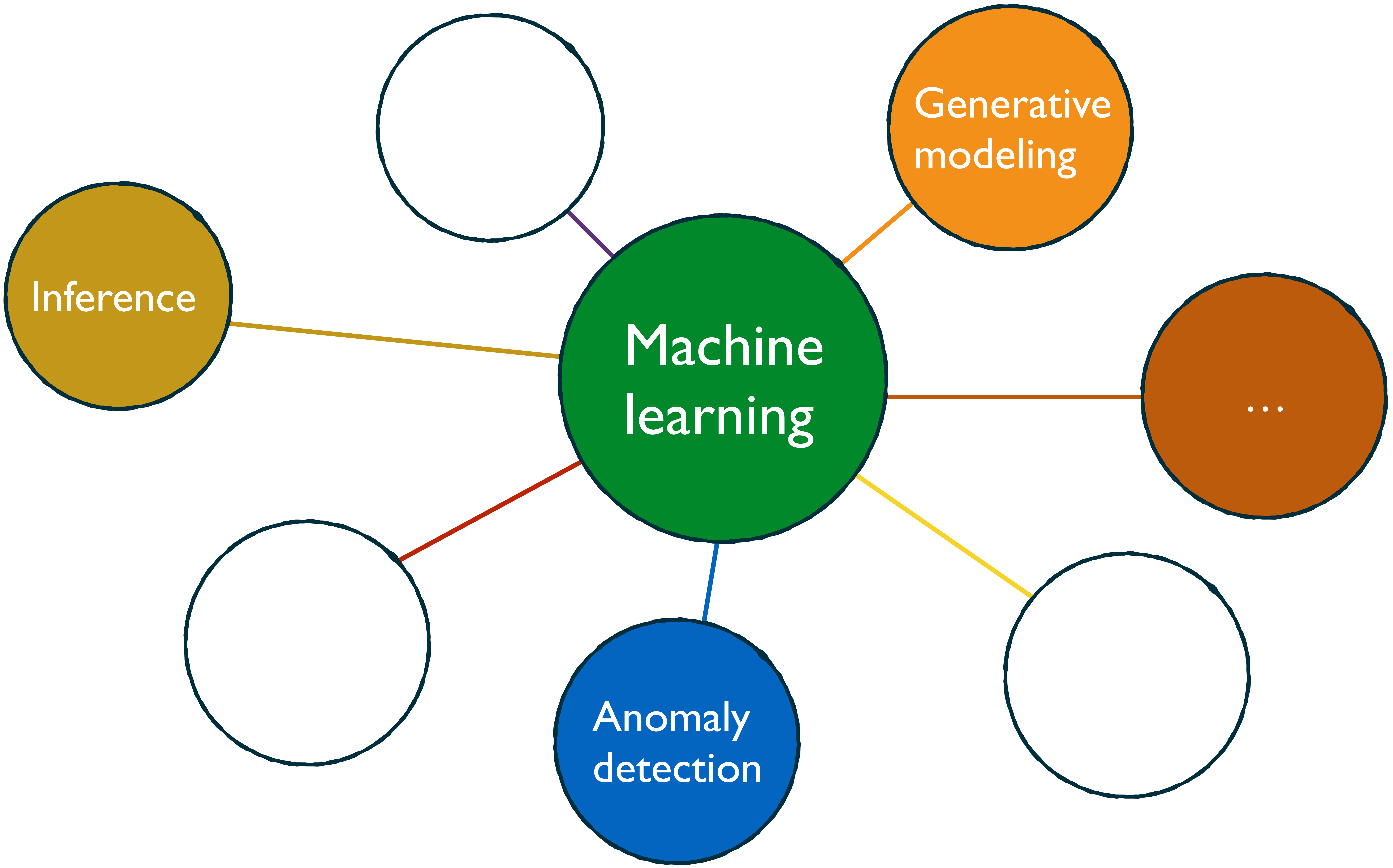
Point cloud

$$z = \frac{2p_T}{\sqrt{s}} \cosh \eta$$

- Event-wide conservation laws

Point clouds

Momentum: 0.999(3)
Baryon number: 0.995(2)
Lepton number: 1.001(2)


Generative modeling
Inference
Machine learning
...
Anomaly detection

Neural simulation based inference

Lersch, Qiu, FR, Sato et al. `25

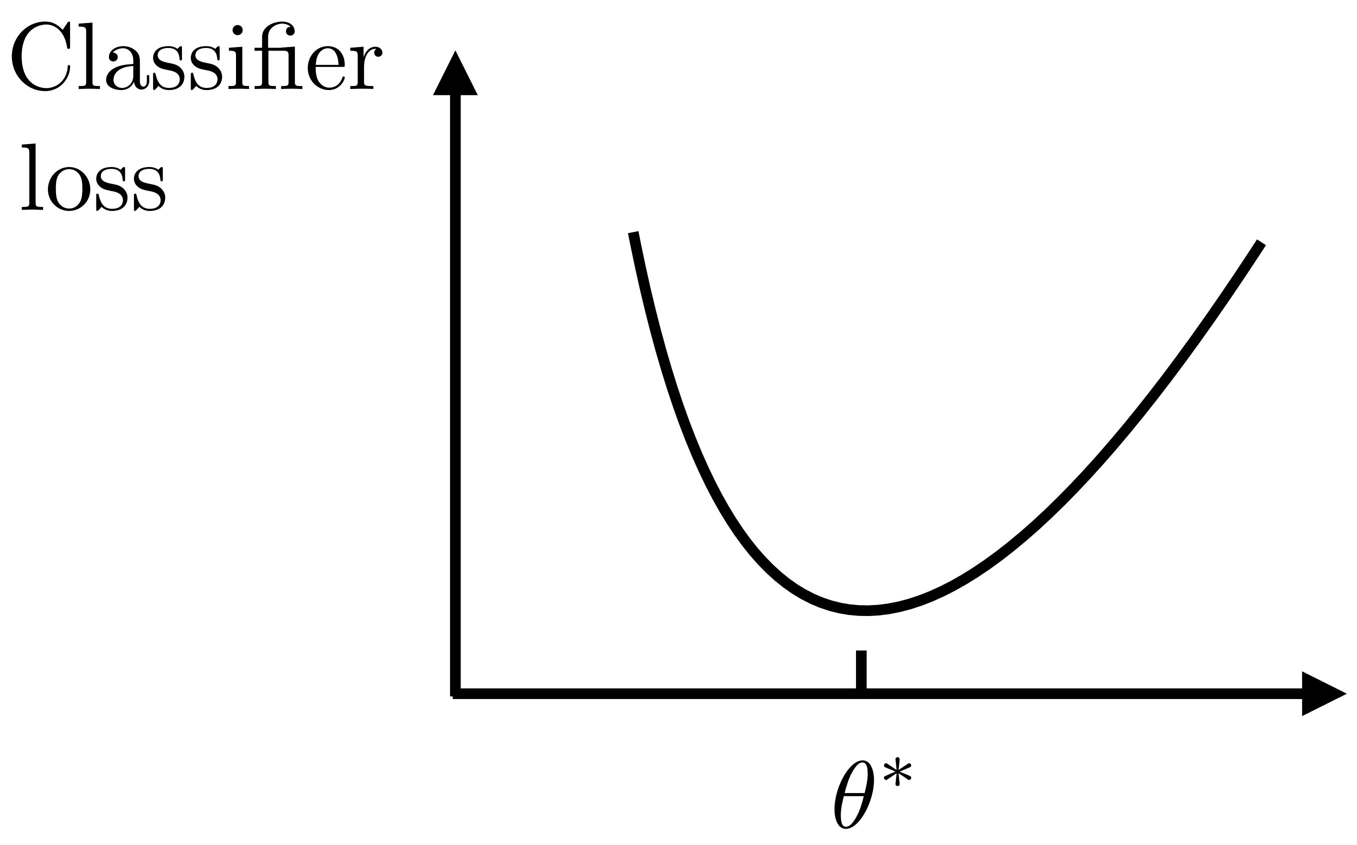
- Use full event level information

Theory θ → Events x_i → Detector level $\tilde{x}_i$

Classifier

Experiment $\tilde{y}_i$

Parametrizing e.g. nonperturbative physics

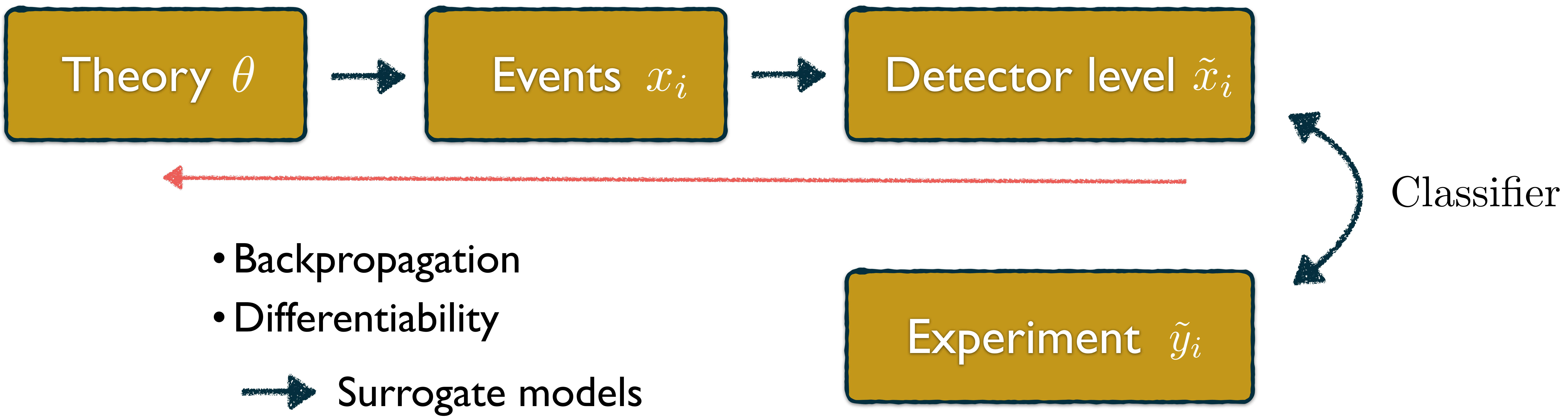


See also Nachman et al., Cranmer et al.

Neural simulation based inference

Lersch, Qiu, FR, Sato et al. `25

- Use full event level information



- Backpropagation
- Differentiability

→ Surrogate models

- See e.g. recent ATLAS measurement of SMEFT operators

See also Nachman et al., Cranmer et al.

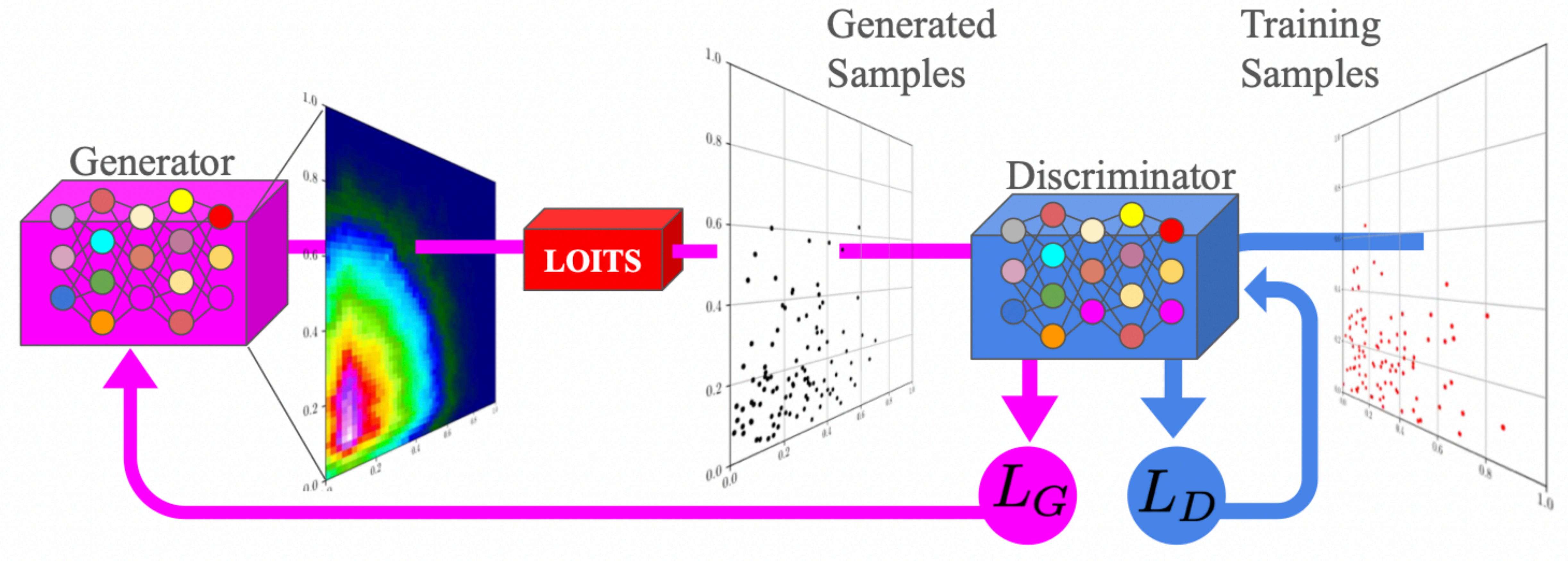
Neural simulation based inference

Lersch, Qiu, FR, Sato et al. `25

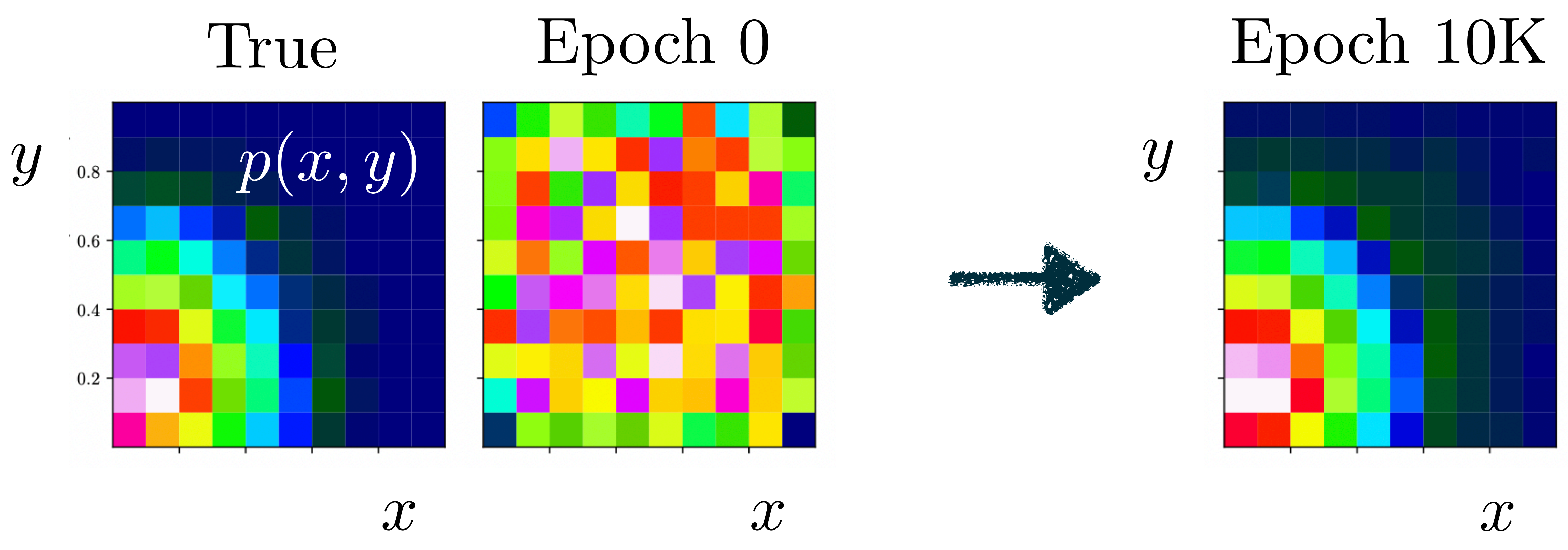
- Use full event level information

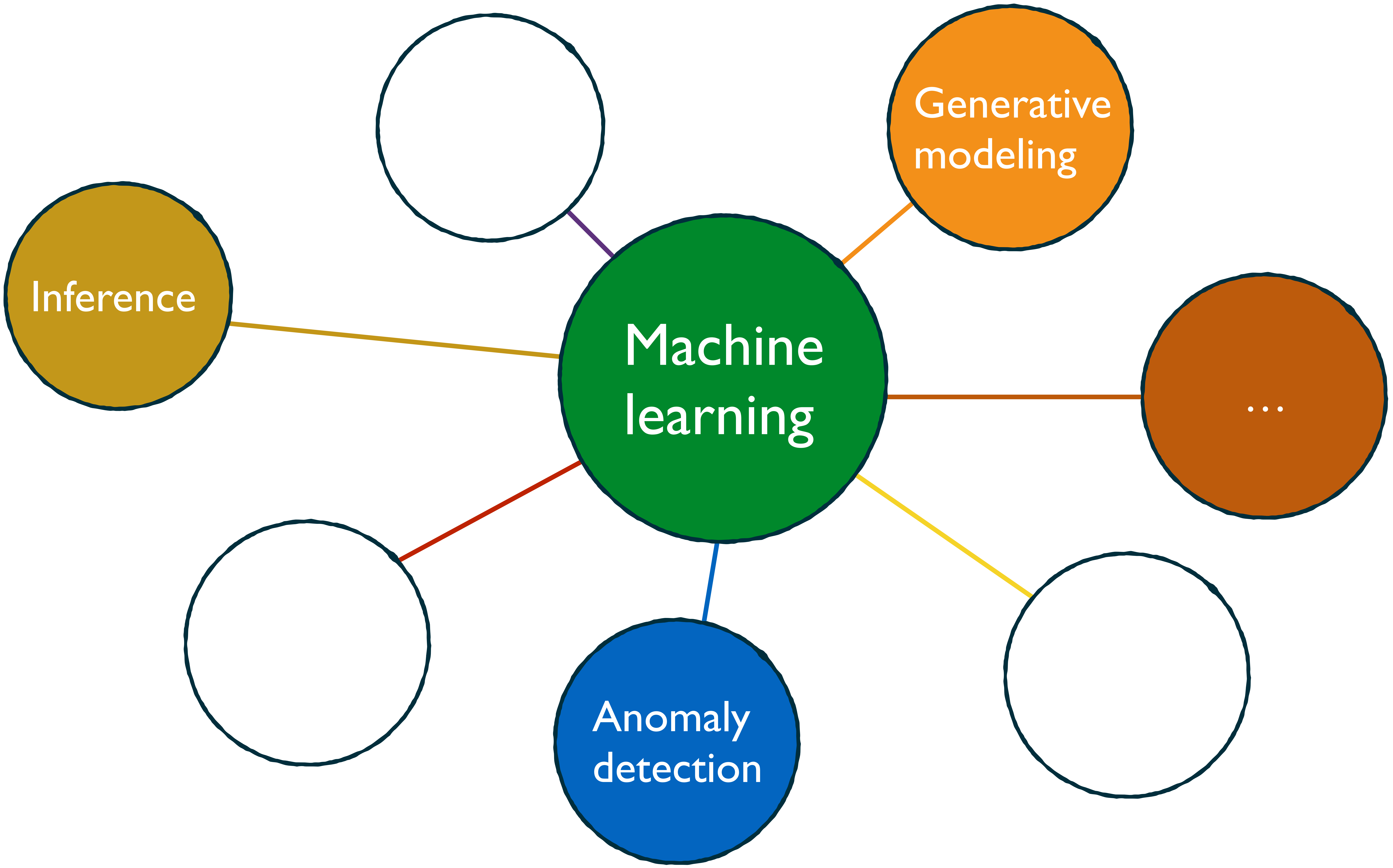
- Differential sampling algorithm

  E.g. analytical calculations



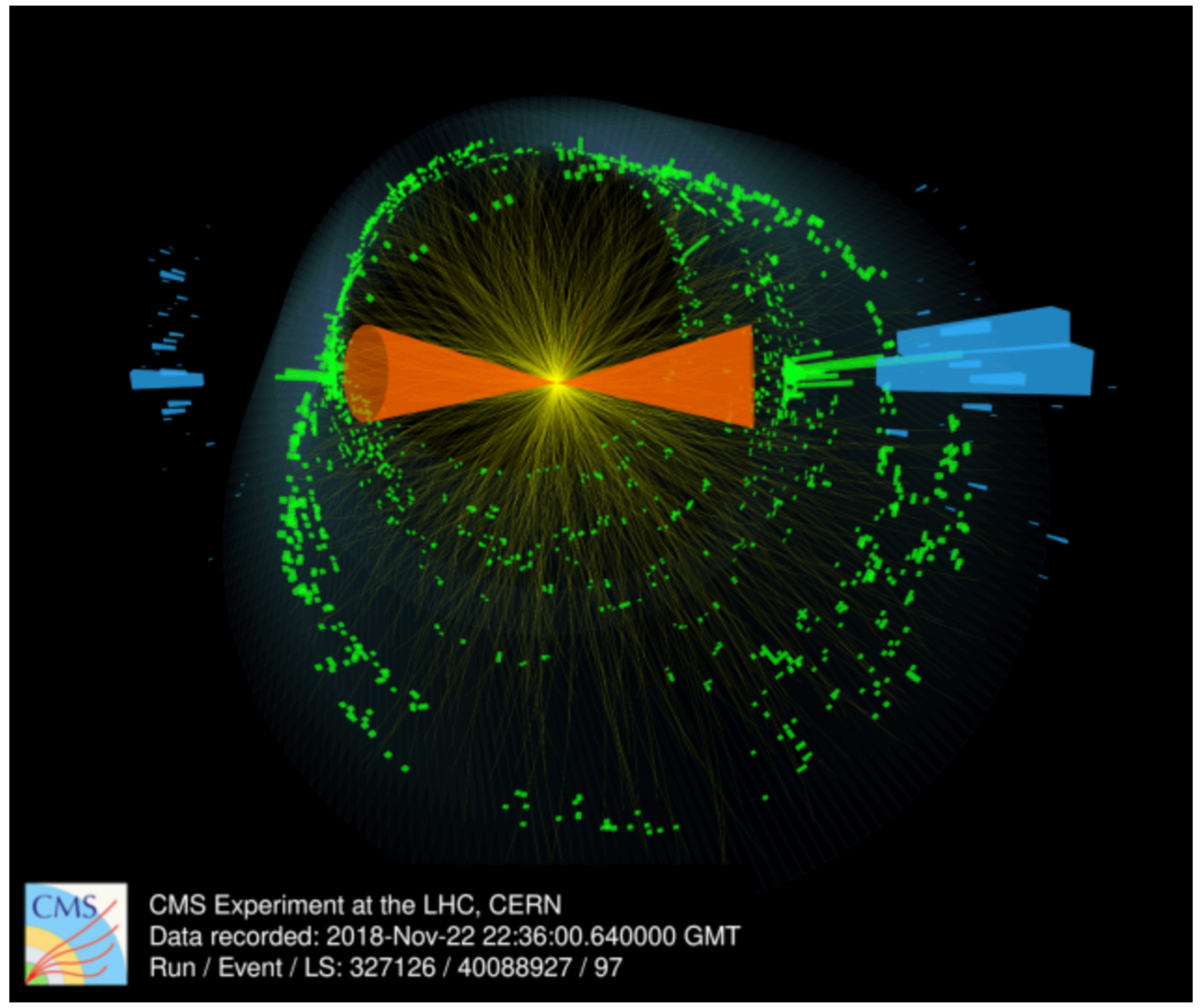
- Closure test: Train GAN to learn parameters from simulated events

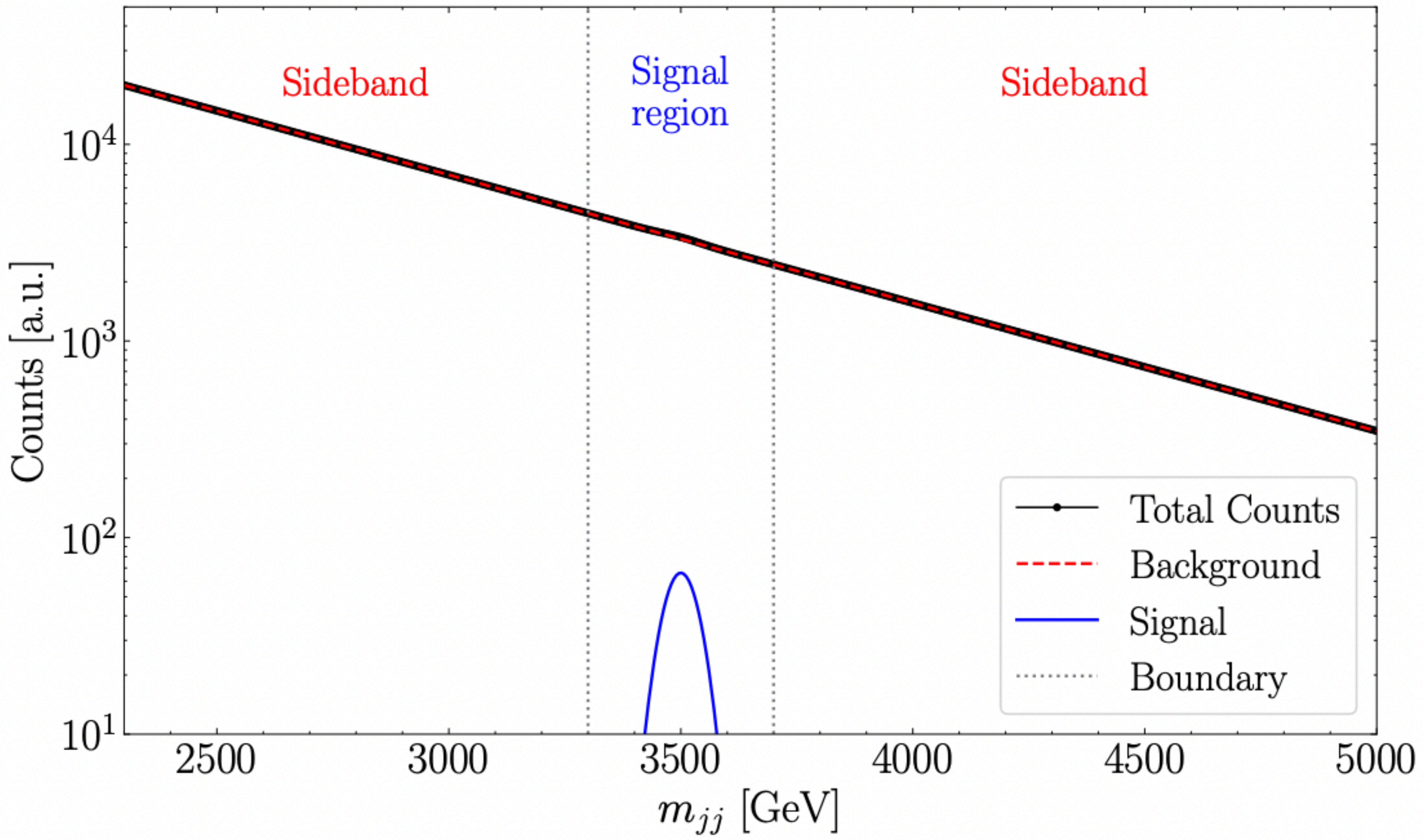



Generative modeling
Inference
Machine learning
...
Anomaly detection

Anomaly detection and BSM physics

- Theory agnostic BSM searches
  - Weakly supervised
  - Unsupervised



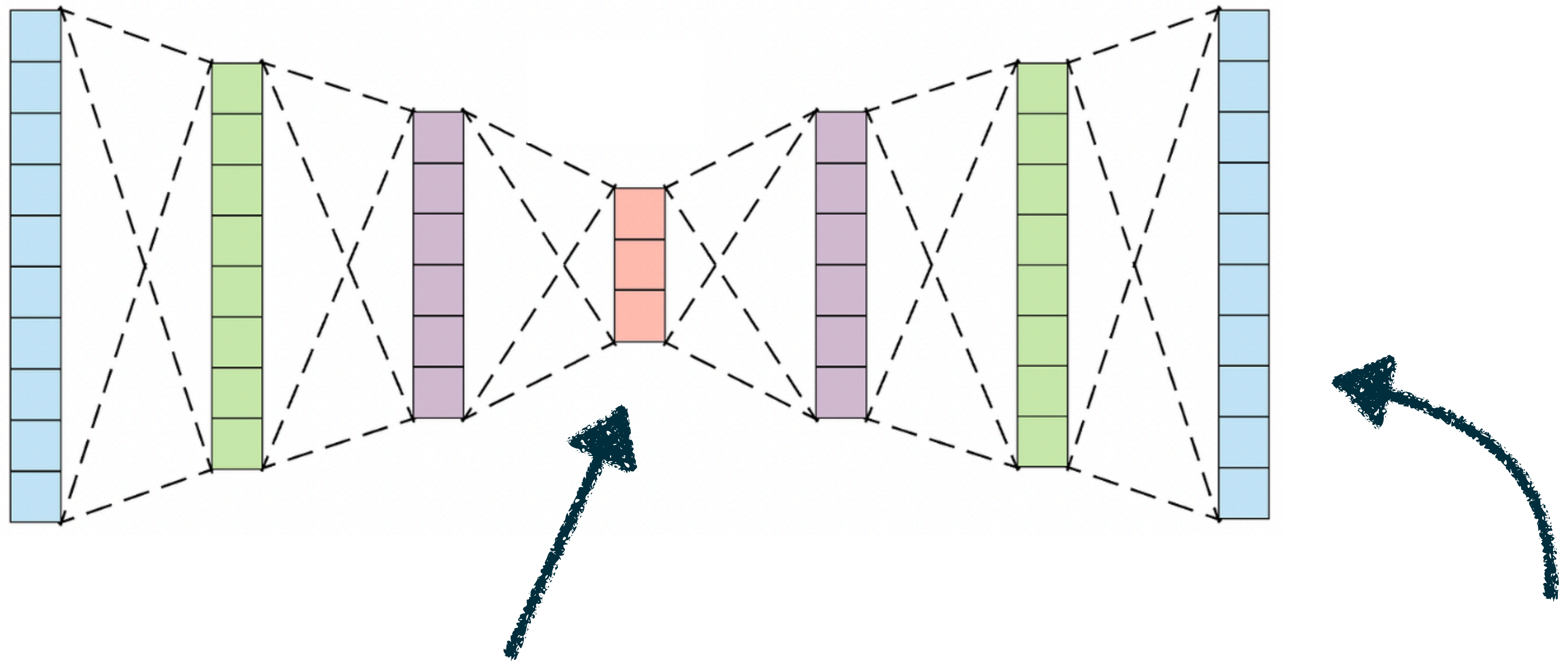


See e.g. Nachman, Thaler, Mikuni, Plehn, Spannowsky, Kasieczka, et al.

Anomaly detection and BSM physics

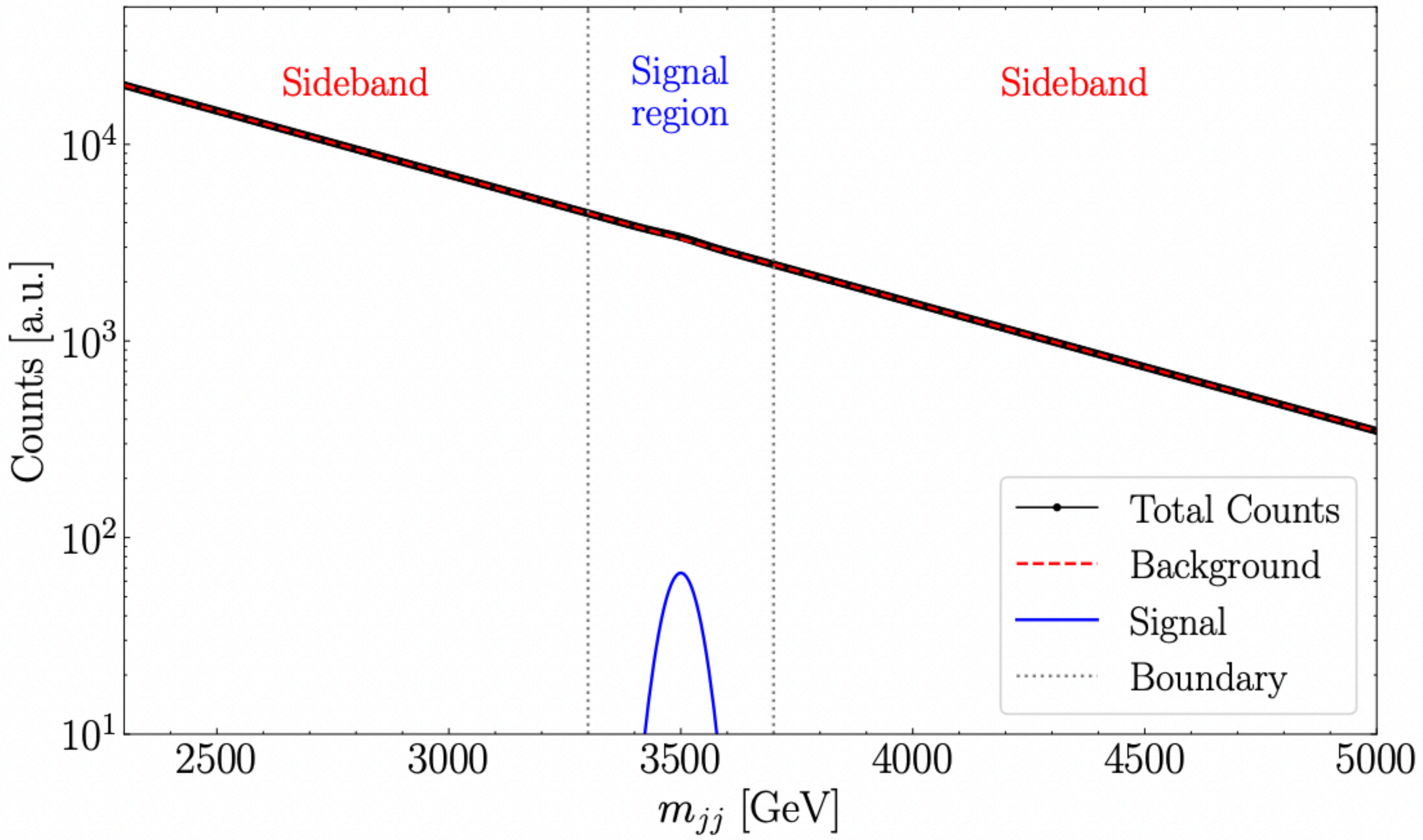
- Theory agnostic BSM searches
  - Weakly supervised
  - Unsupervised

  Autoencoder



Bottleneck

Anomaly score via reconstruction loss



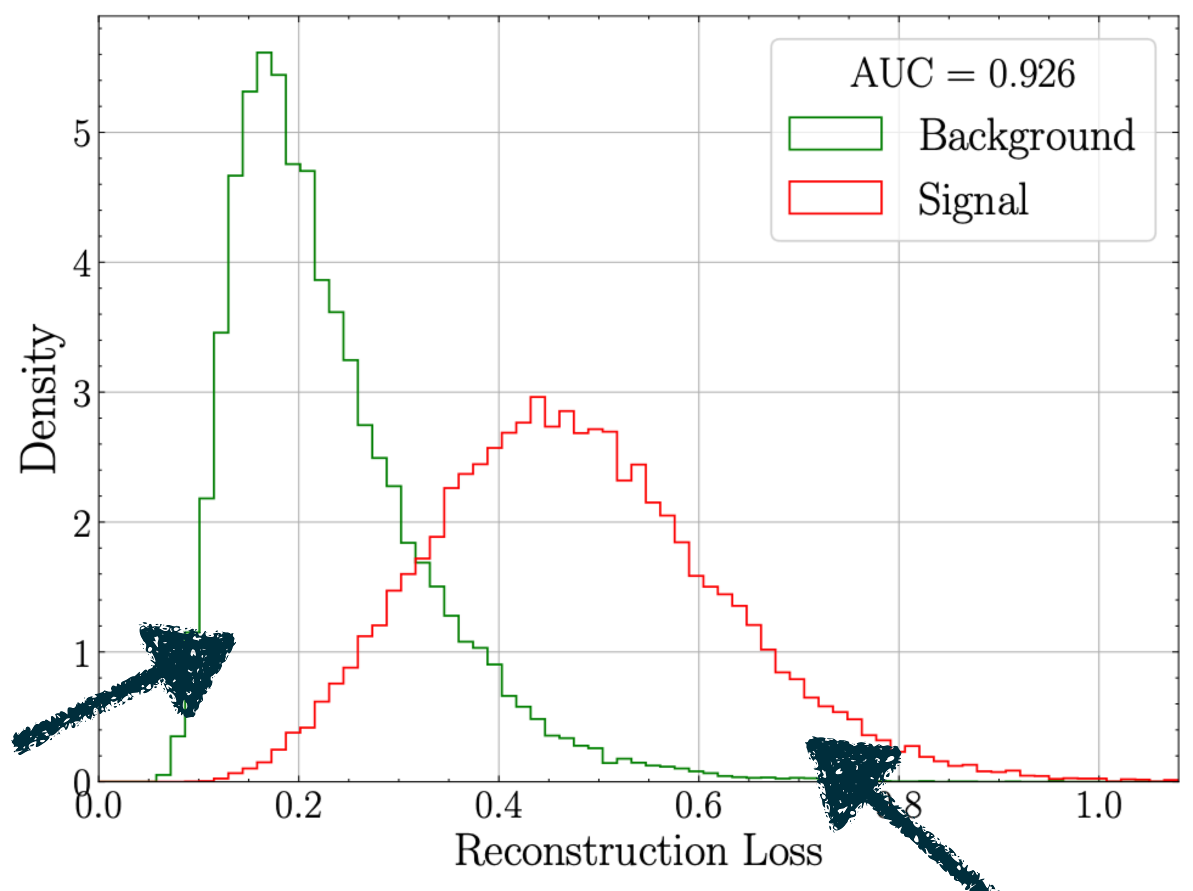
See e.g. Nachman, Thaler, Mikuni, Plehn, Spannowsky, Kasieczka, et al.

Anomaly detection and BSM physics

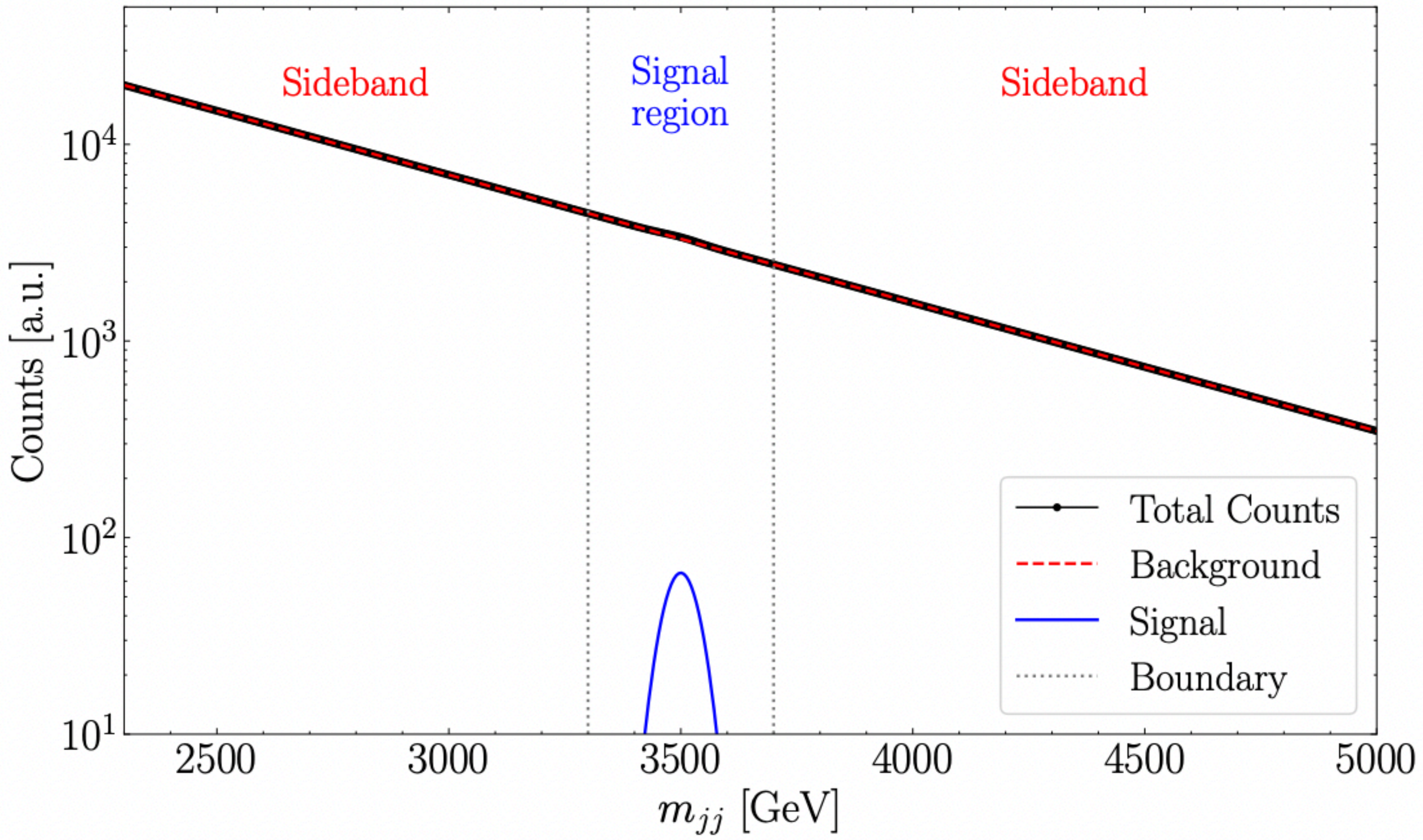
- Theory agnostic BSM searches
  - Weakly supervised
  - Unsupervised

    Autoencoder

Signal



Background



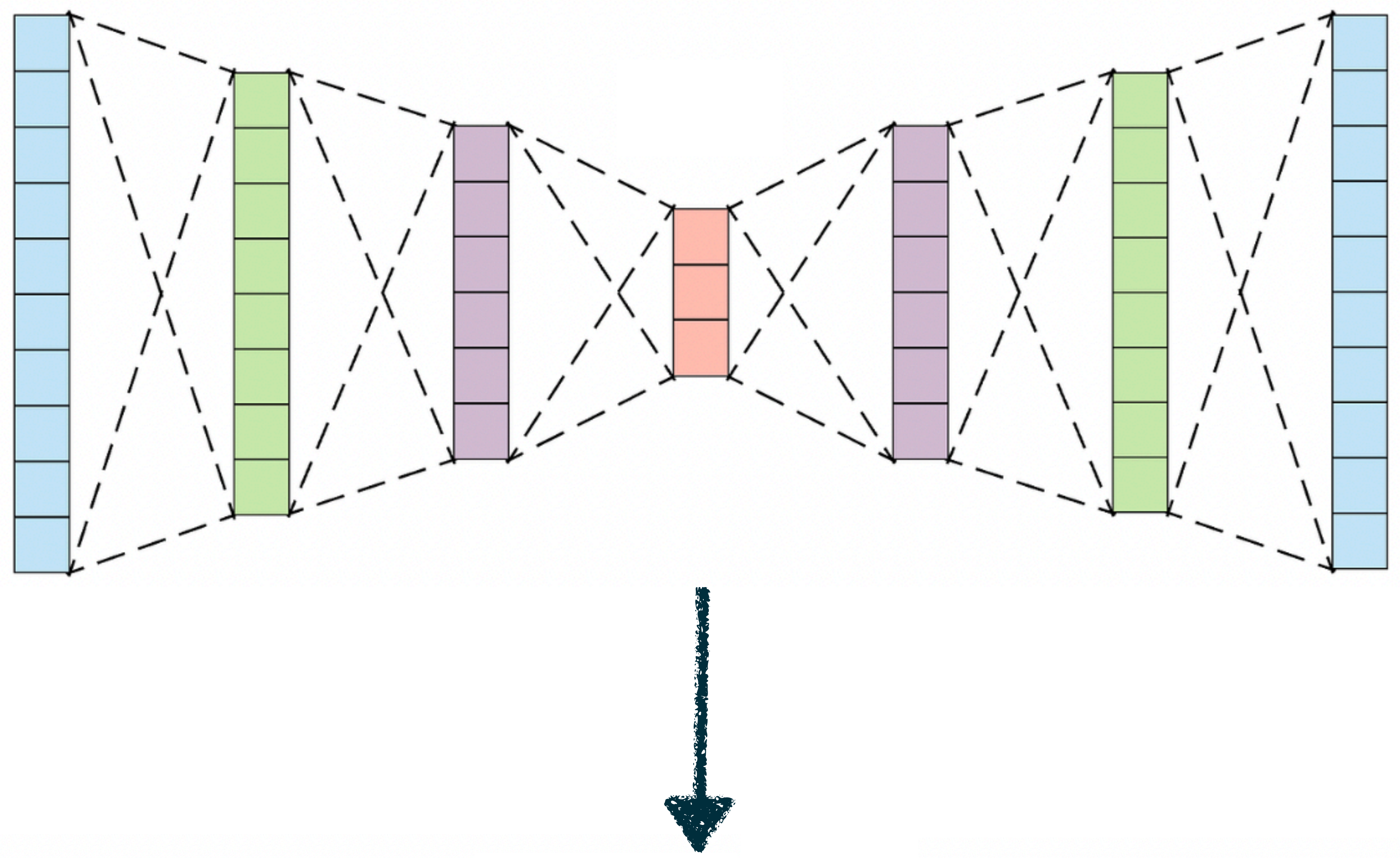
See e.g. Nachman, Thaler, Mikuni, Plehn, Spannowsky, Kasieczka, et al.

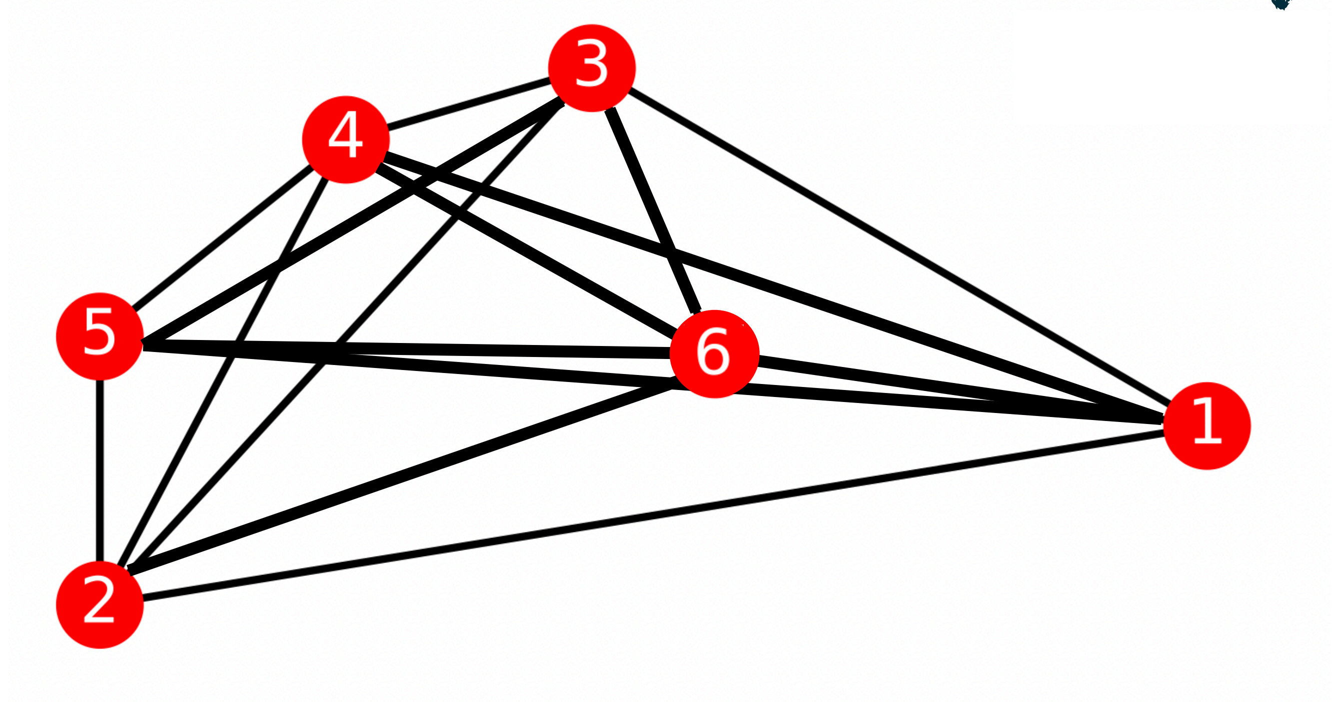
Anomaly detection and BSM physics

Araz, Athanasakos, Ploskon, FR `25

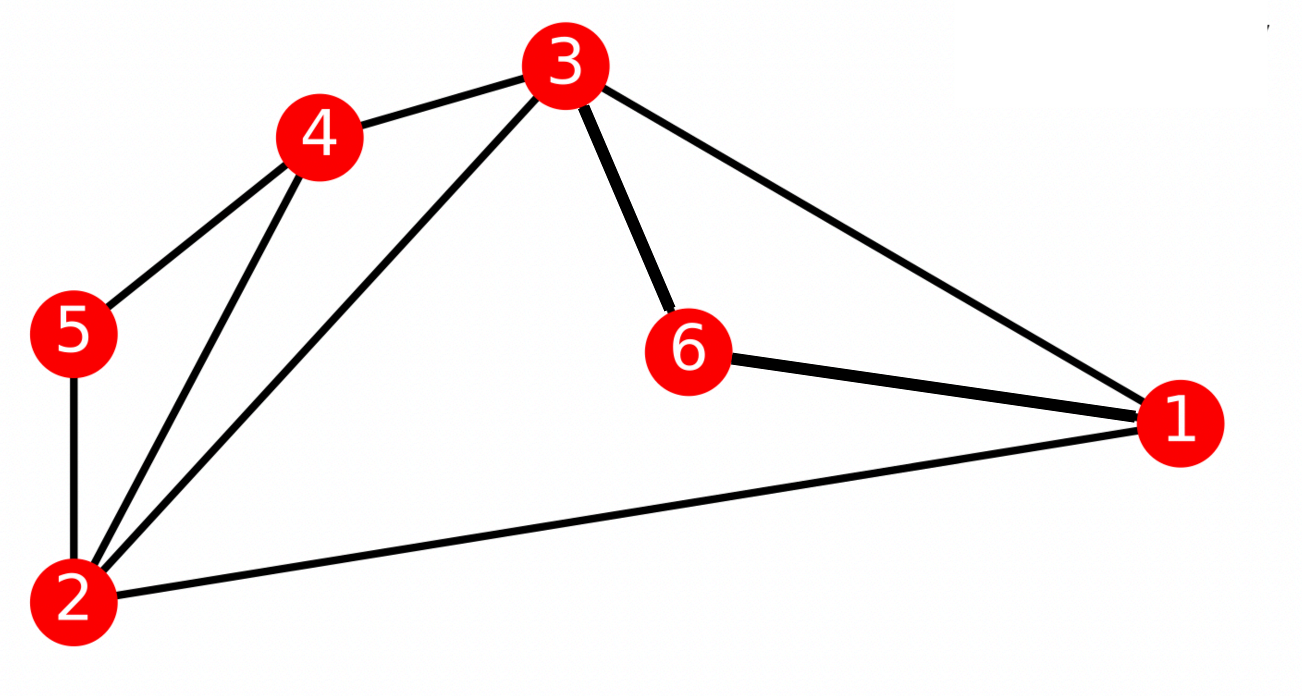
- Graph-based autoencoder

  Based on graph structure





Fully connected graph



Sparse "Laman" graph

Anomaly detection and BSM physics

Araz, Athanasakos, Ploskon, FR `25

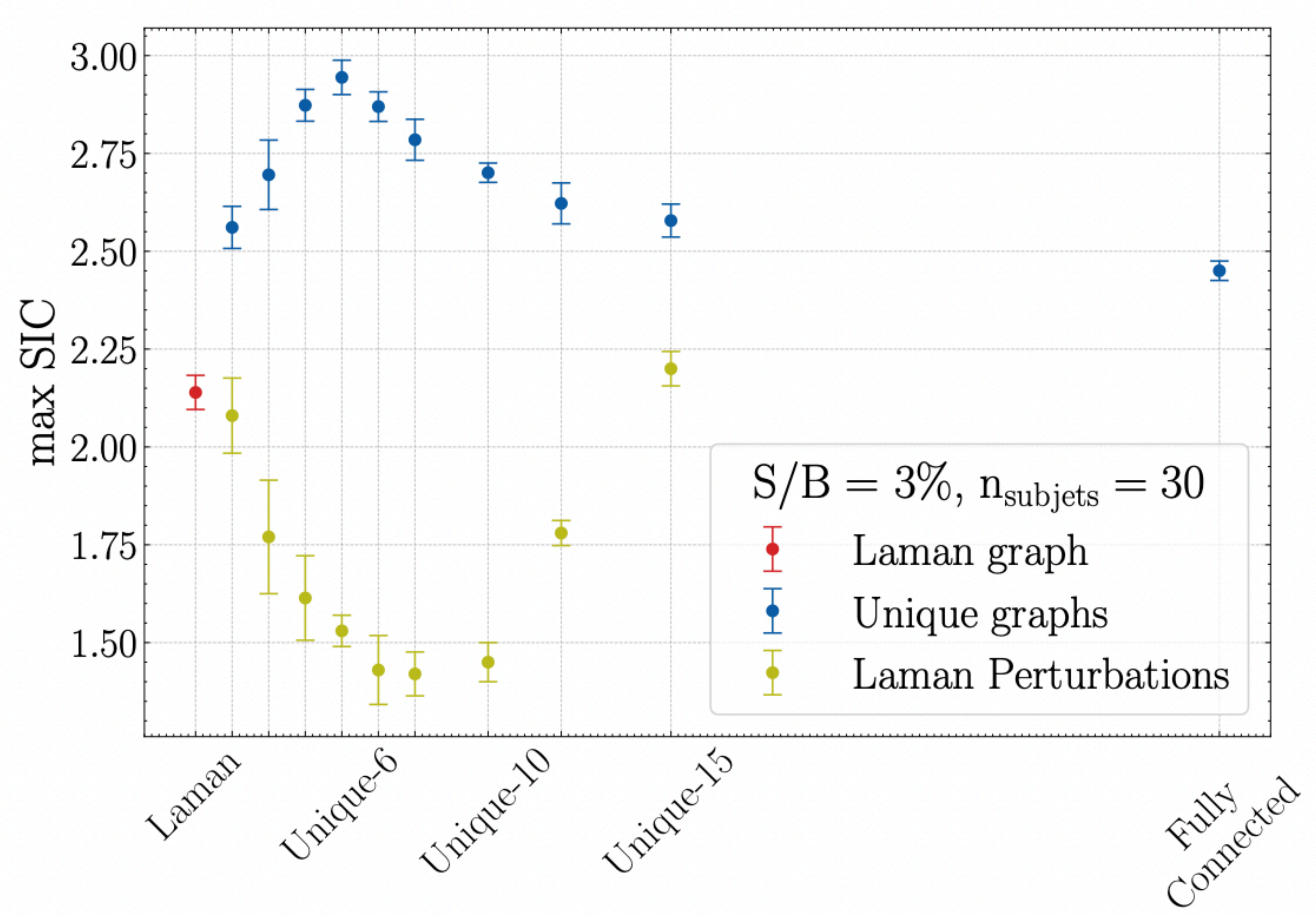
- Graph-based autoencoder

  State of the art for unsupervised searches using LHC benchmark

  Signal Improvement Characteristic (SIC)

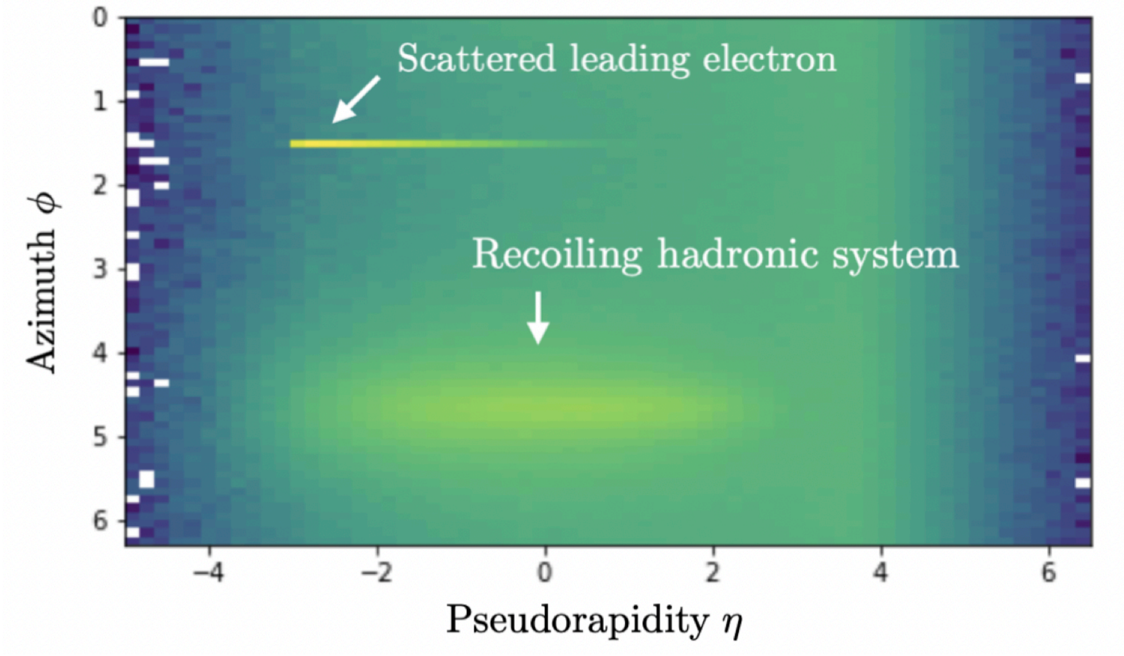
  Full event information e.g. $2\sigma \to 6\sigma$

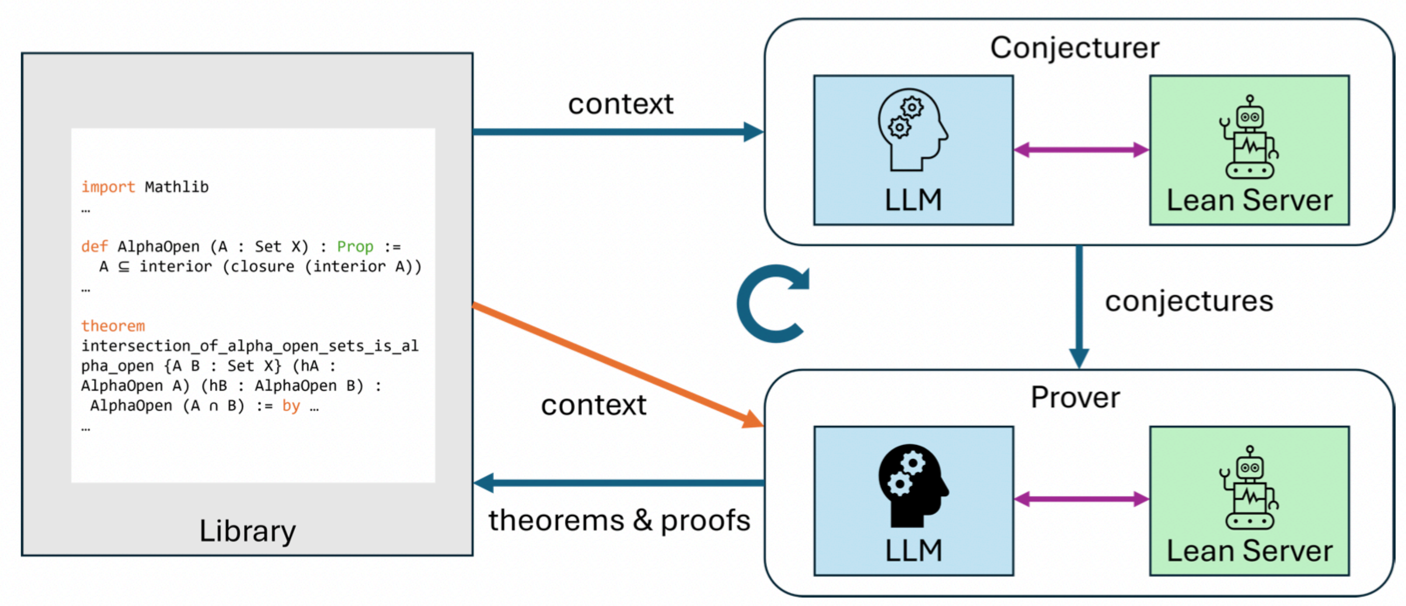
⟶ Applications at the future EIC?

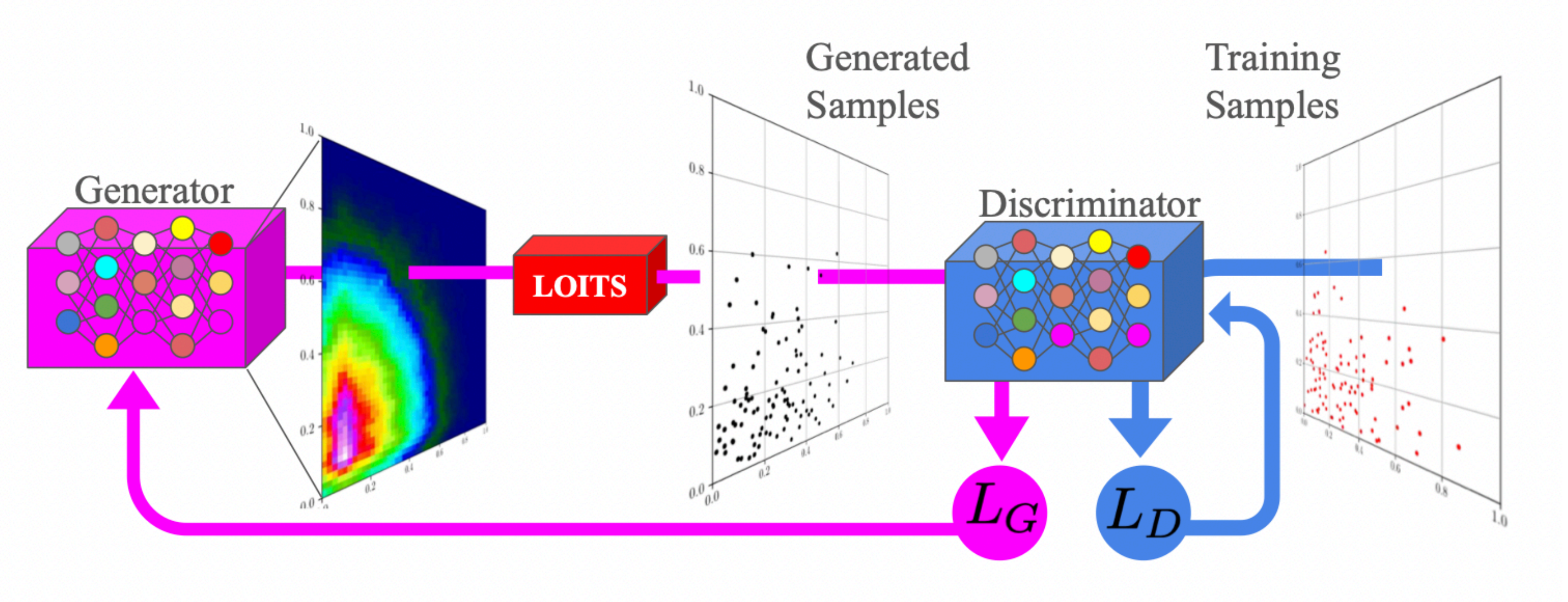


Conclusions and outlook

- Rapid progress in AI/ML
- Various applications using event-level analyses
- New collaborations between theory and experiment
- Use of general purpose AI in nuclear physics?



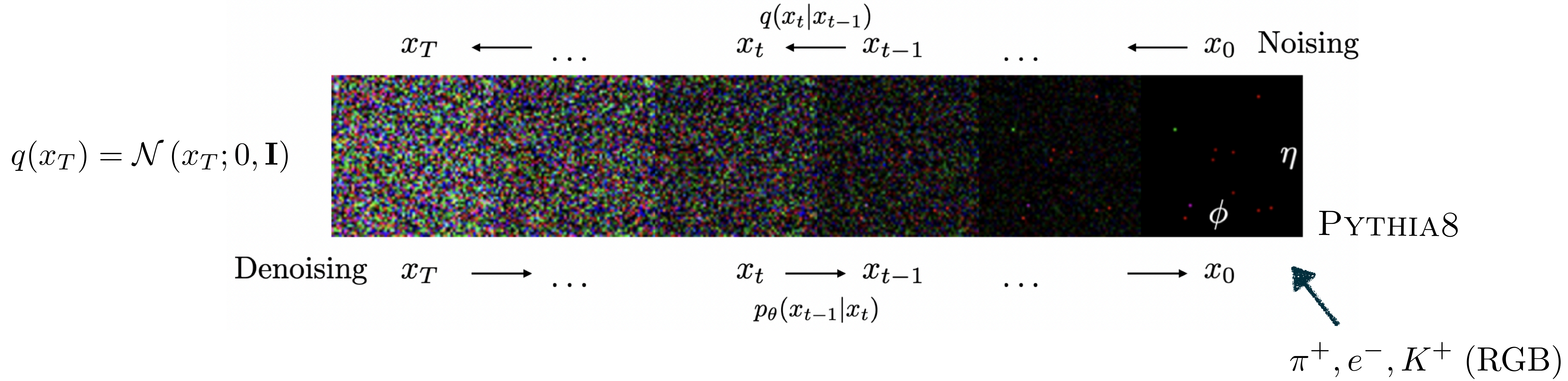




Diffusion models

Devlin, Qiu, FR, Sato `23

- Represent events as images (pixelated)



- Markovian noising process $$q(x_1, \dots, x_T|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1})$$

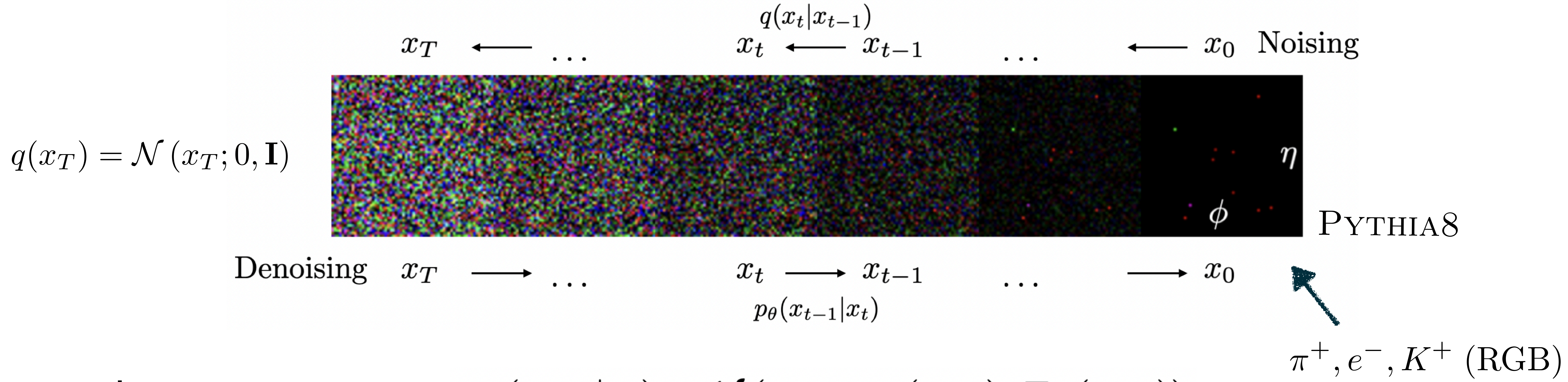
  adding Gaussian noise $$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t \mathbf{I})$$

Sohl-Dickstein et al.`15
Ho, Jain, Abbeel `20
Nichol, Dhariwal `21

Diffusion models

Devlin, Qiu, FR, Sato `23

- Represent events as images (pixelated)



- Learn denoising process $p_\theta\left(x_{t-1}|x_t\right) = \mathcal{N}\left(x_{t-1}; \mu_\theta\left(x_t, t\right), \Sigma_\theta\left(x_t, t\right)\right)$

Stochastic differential equation in continuum limit

Train convolutional U-Net

Sohl-Dickstein et al.`15
Ho, Jain, Abbeel `20
Nichol, Dhariwal `21