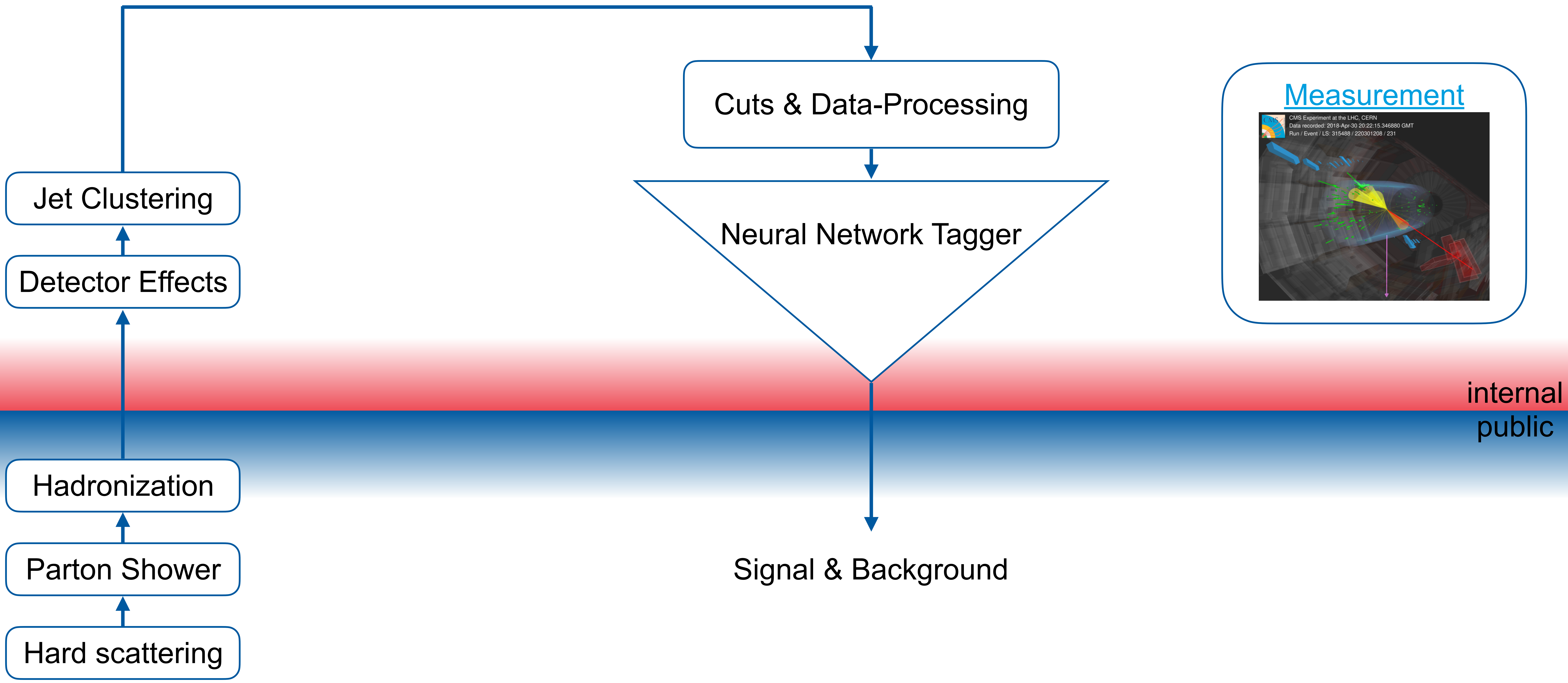
Generalization Properties of Jet Classification

Sebastian Bieringer, Gregor Kasieczka, Jan Kieseler

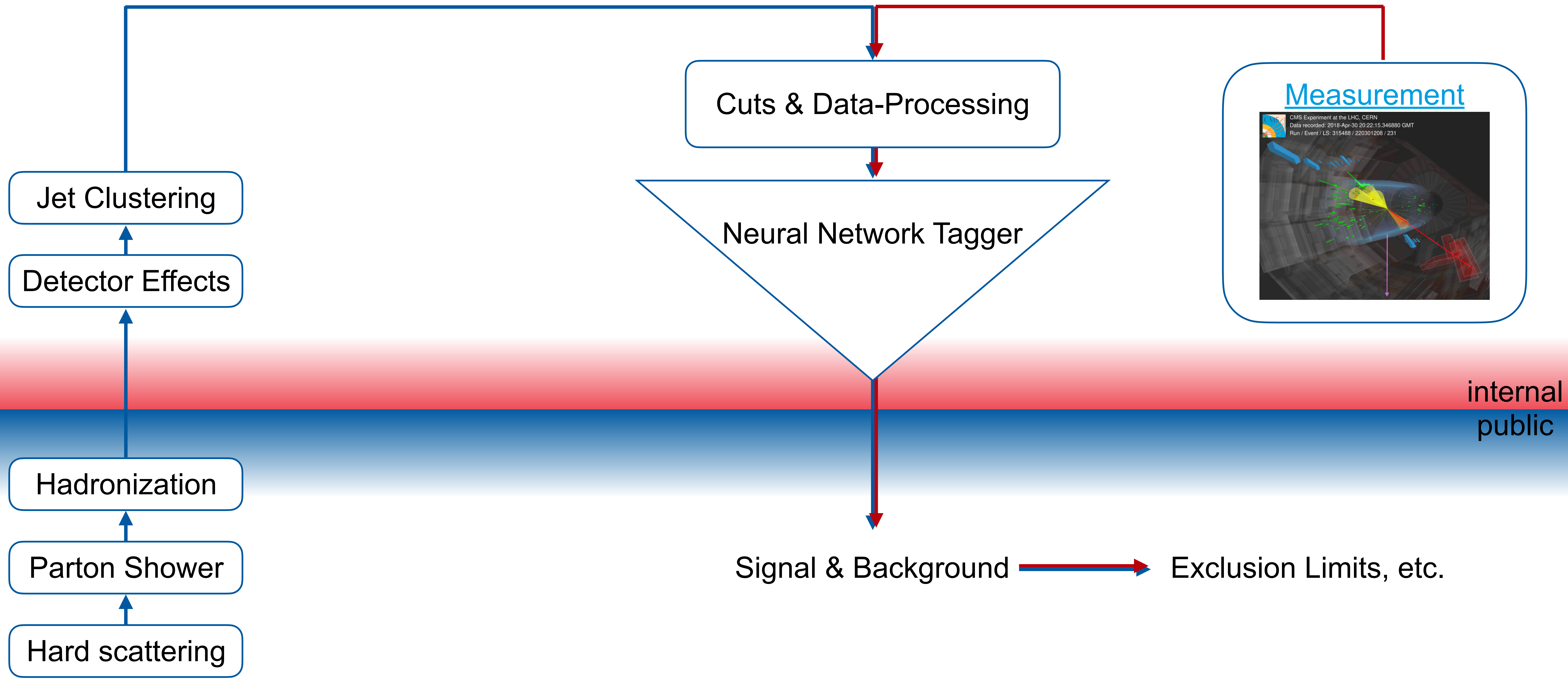
Institut für Experimentalphysik, Universität Hamburg, Germany
sebastian.guido.bieringer@uni-hamburg.de

09.11.2023 - ML4jets 2023

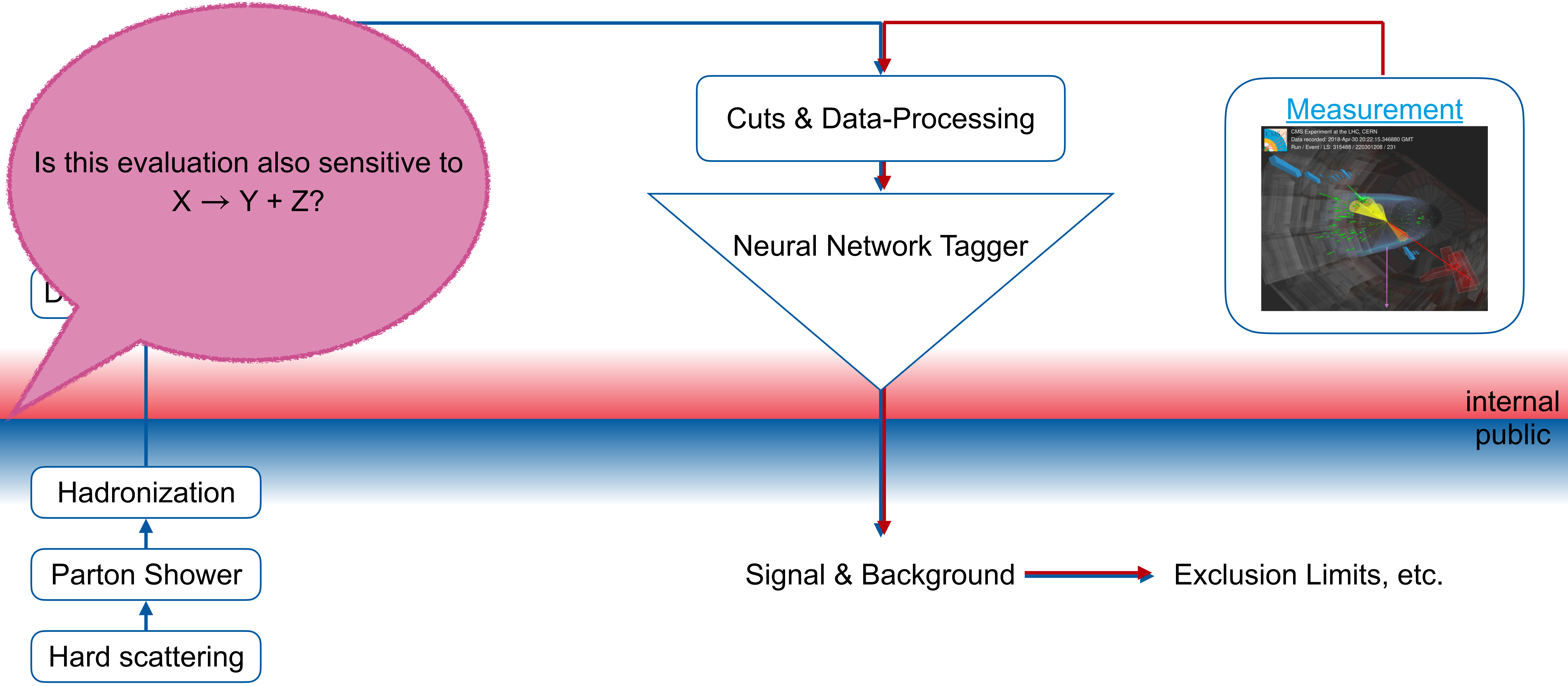
Classification Surrogates



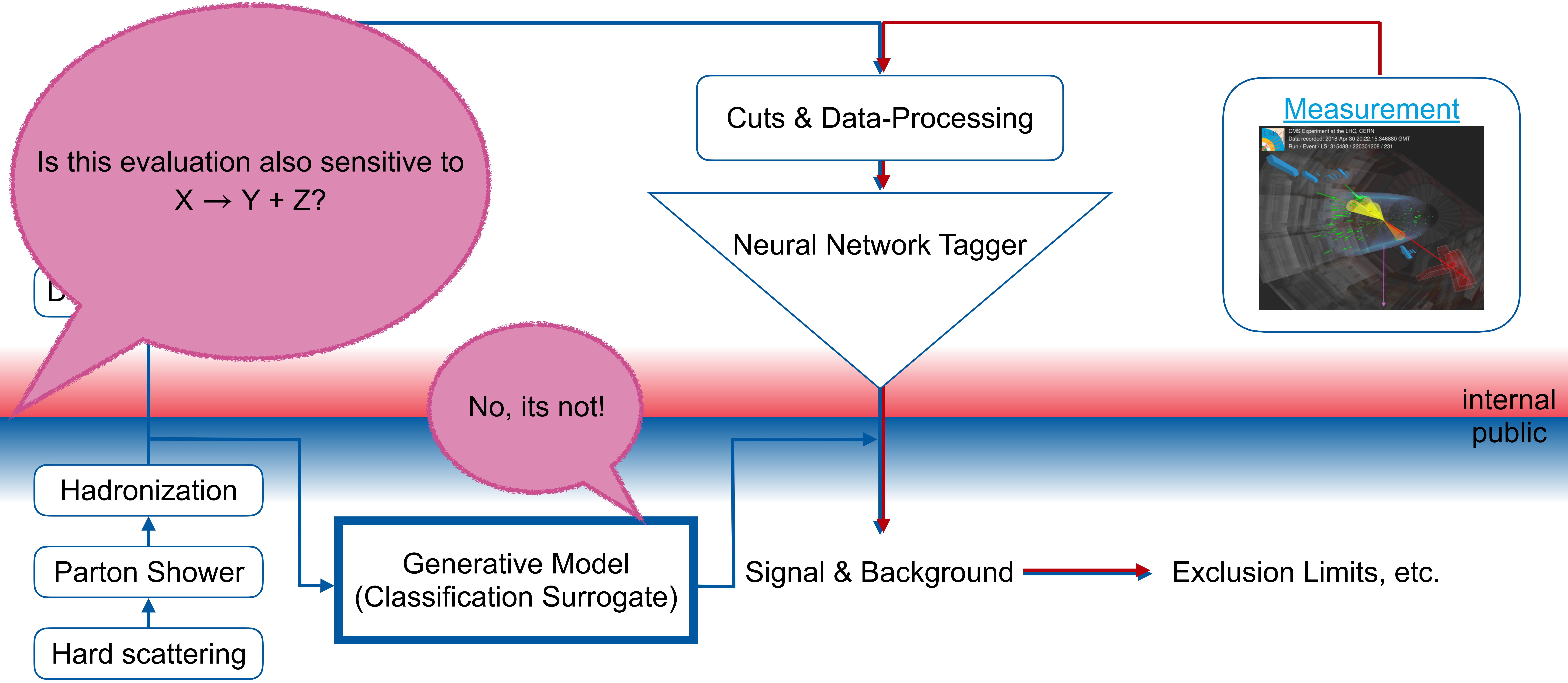
Classification Surrogates



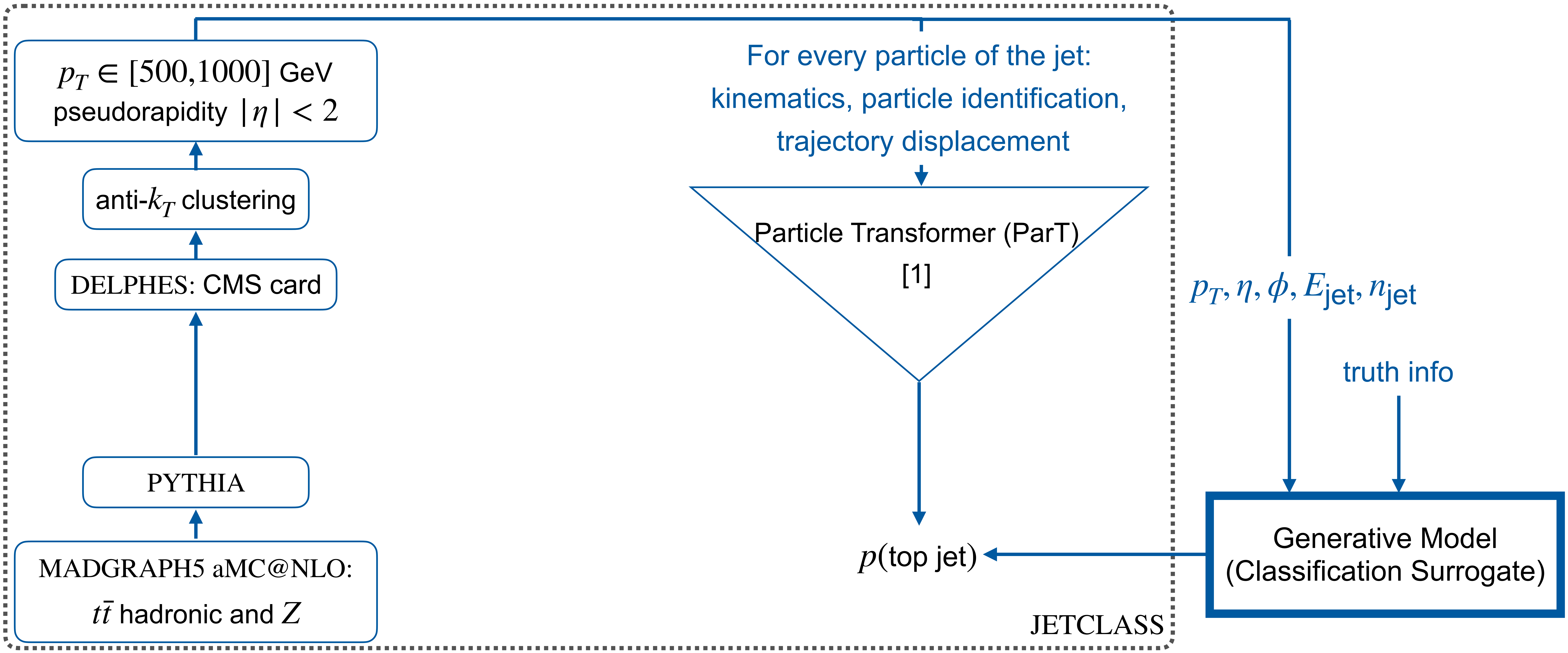
Classification Surrogates



Classification Surrogates

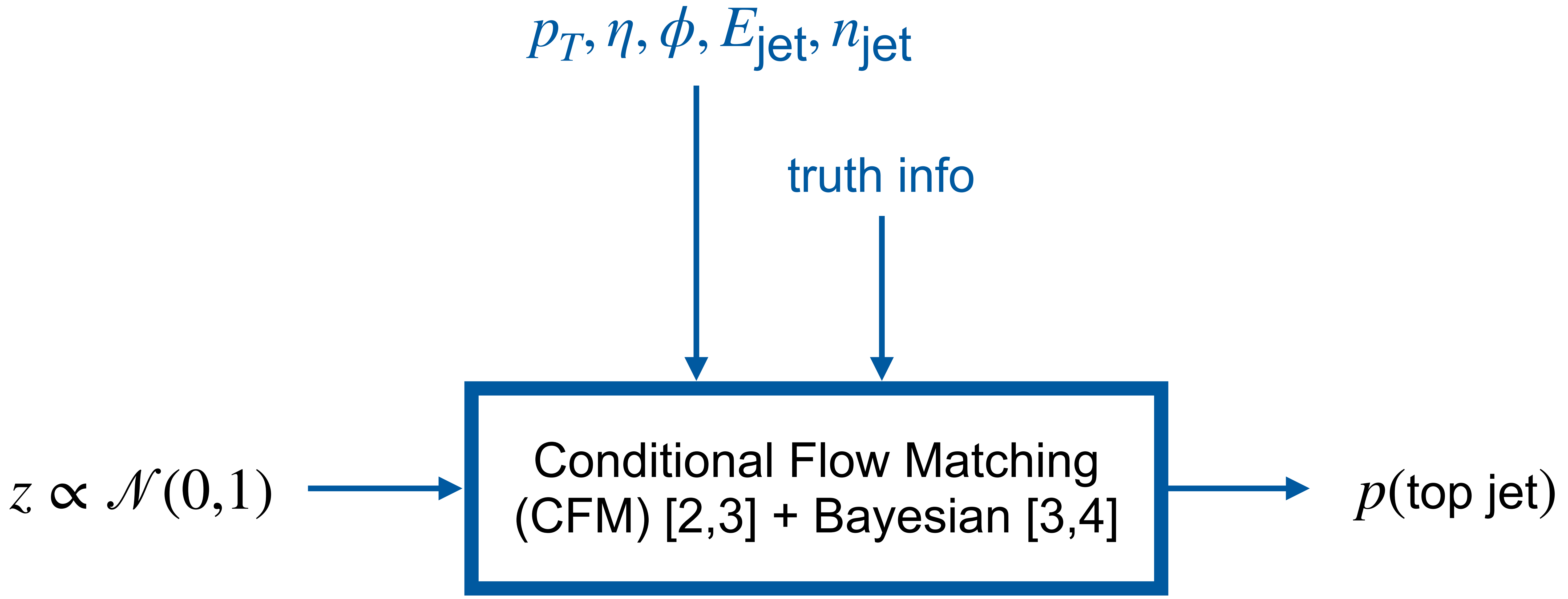


The Toy Setup



[1] https://arxiv.org/pdf/2202.03772

The Generative Model



[2] https://gist.github.com/francois-rozet/fd6a820e052157f8ac6e2aa39e16c1aa
[3] https://arxiv.org/pdf/2210.02747
[4] https://github.com/IntelLabs/bayesian-torch
[5] https://arxiv.org/pdf/2305.10475

Bayesian CFM

Continuous Normalizing Flow:

- Flow $\phi : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ defined via

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = v_t(\phi_t(x)) = \tilde{v}_t(x, \theta)$$

- solve the ODE to train and sample
- linear trajectory
- transforms probability distributions

$$p_t(x) = p_0\left(\phi_t^{-1}(x)\right) \det\left[\frac{\partial \phi_t^{-1}}{\partial x}(x)\right]$$

Conditional Flow Matching:

- loss that does not ODE solving

$$\mathscr{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{t,p_t(x)} \left\| v_t(x) - \tilde{v}_t(x,\theta)) \right\|^2$$

- by choice of p_t and v_t

$$\mathscr{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{t,p_t(x),\epsilon}\left[\tilde{v}_t\left((1-t)x_0 + t\epsilon, \theta\right) - \left(\epsilon - x_0\right)\right]^2$$

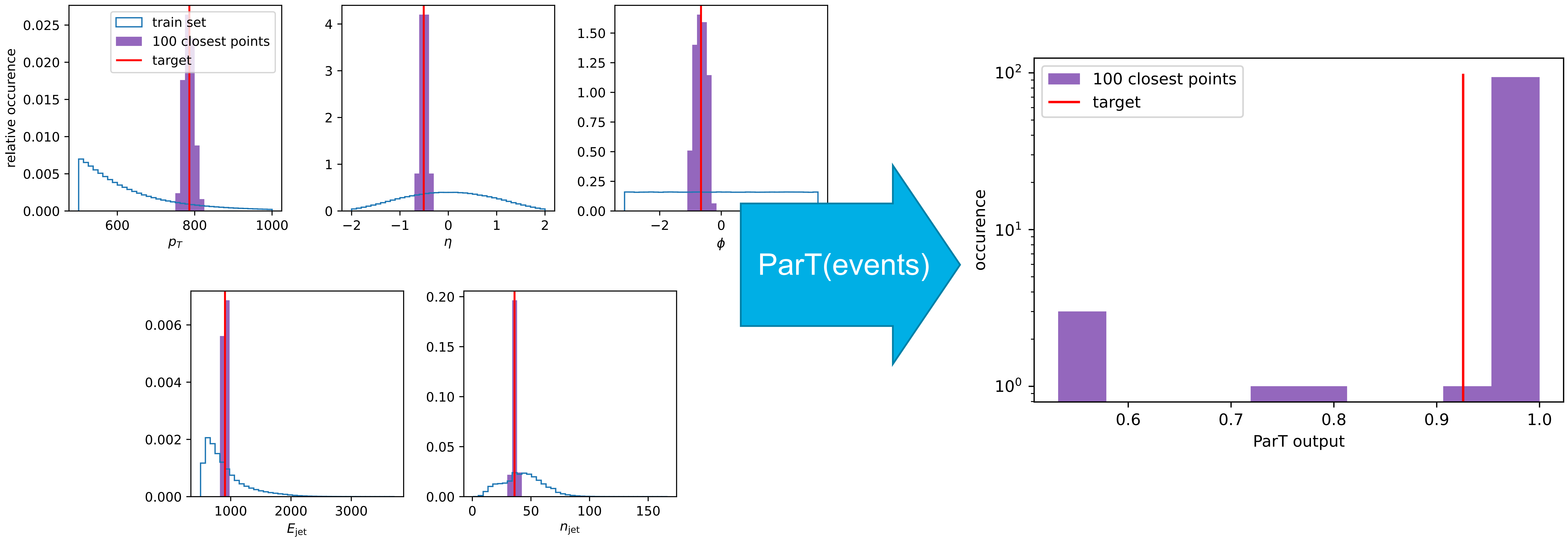
- not a $\log$-Likelihood loss

Bayesian Conditional Flow Matching:

- Bayesian loss $\mathscr{L}_{\mathrm{BNN}} = \mathrm{KL}\left[q(\theta), p\left(\theta \mid x\right)\right] = -\int \mathrm{d}\theta\, q(\theta) \log p\left(x \mid \theta\right) + \mathrm{KL}[q(\theta), p(\theta)] +$ const.
- connect both $\mathscr{L}_{\mathrm{B-CFM}} = \left\langle \mathscr{L}_{\mathrm{CFM}} \right\rangle_{\theta \sim q(\theta)} + c\mathrm{KL}[q(\theta), p(\theta)]$, with $q(\theta)$ uncorrelated Gaussian shape

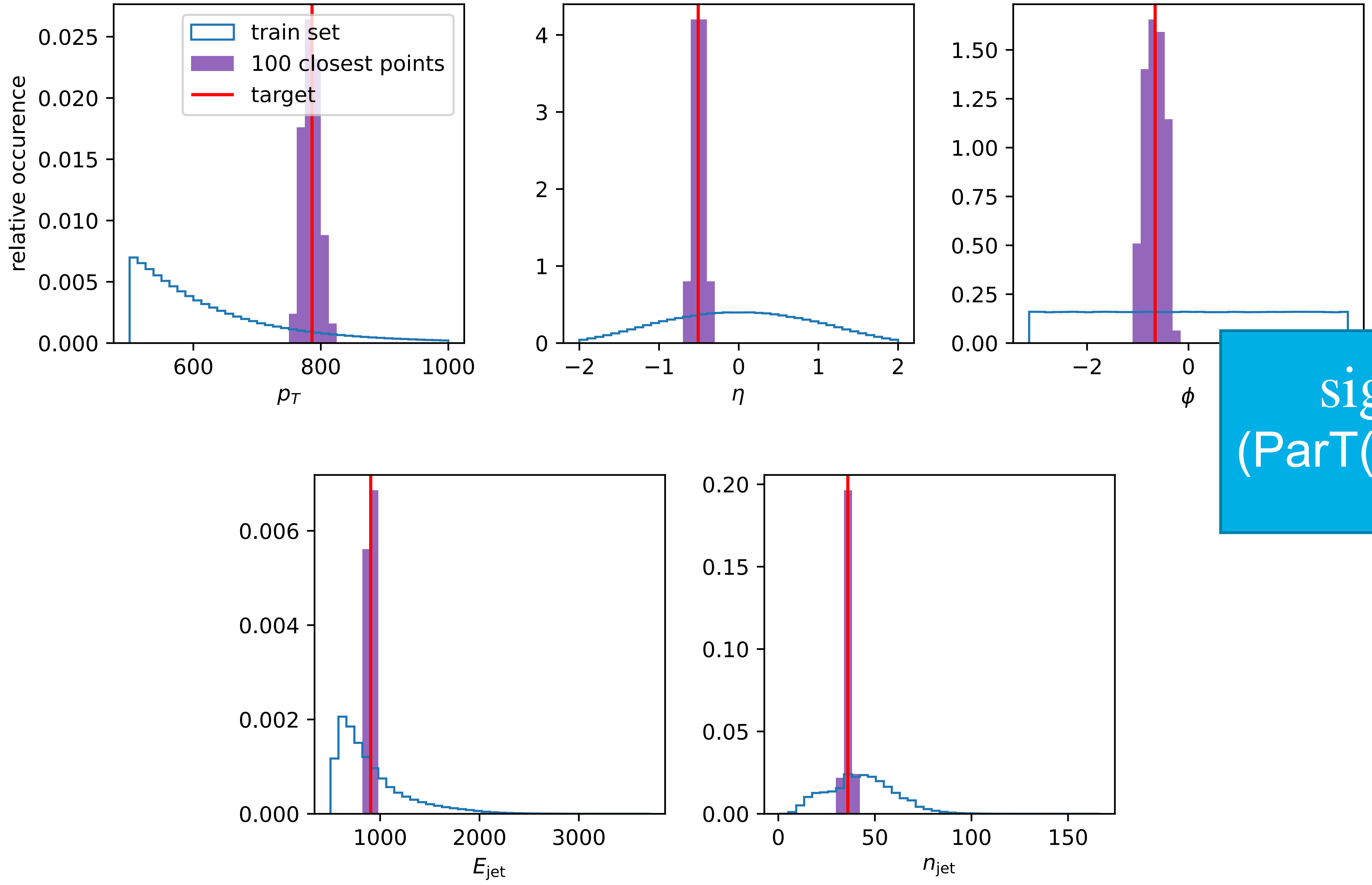
Detector Smearing Distribution

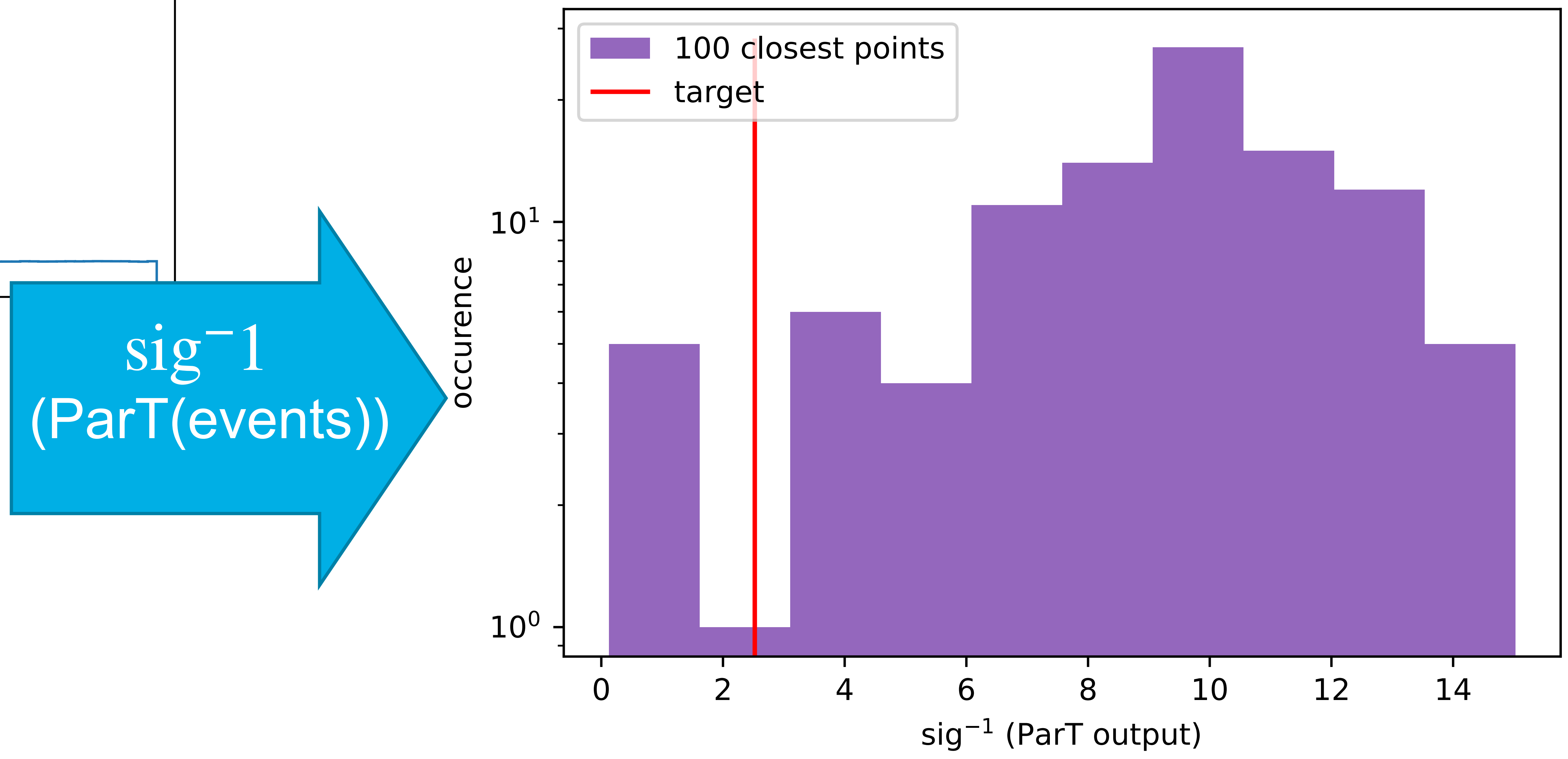
- pick a jet event
- select the 100 events with $p_T, \eta, \phi, E_{\text{jet}}, n_{\text{jet}}$ closest



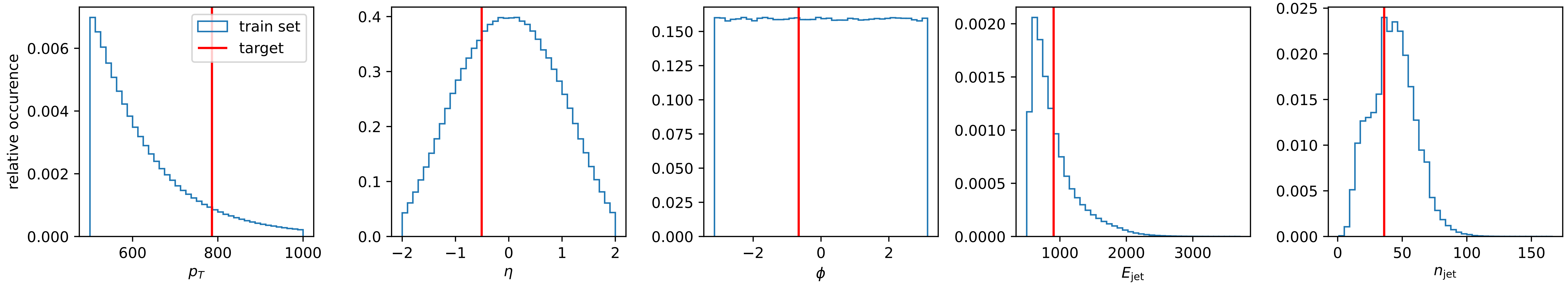
Detector Smearing Distribution

- pick a jet event
- select the 100 events with $p_T, \eta, \phi, E_{\text{jet}}, n_{\text{jet}}$ closest





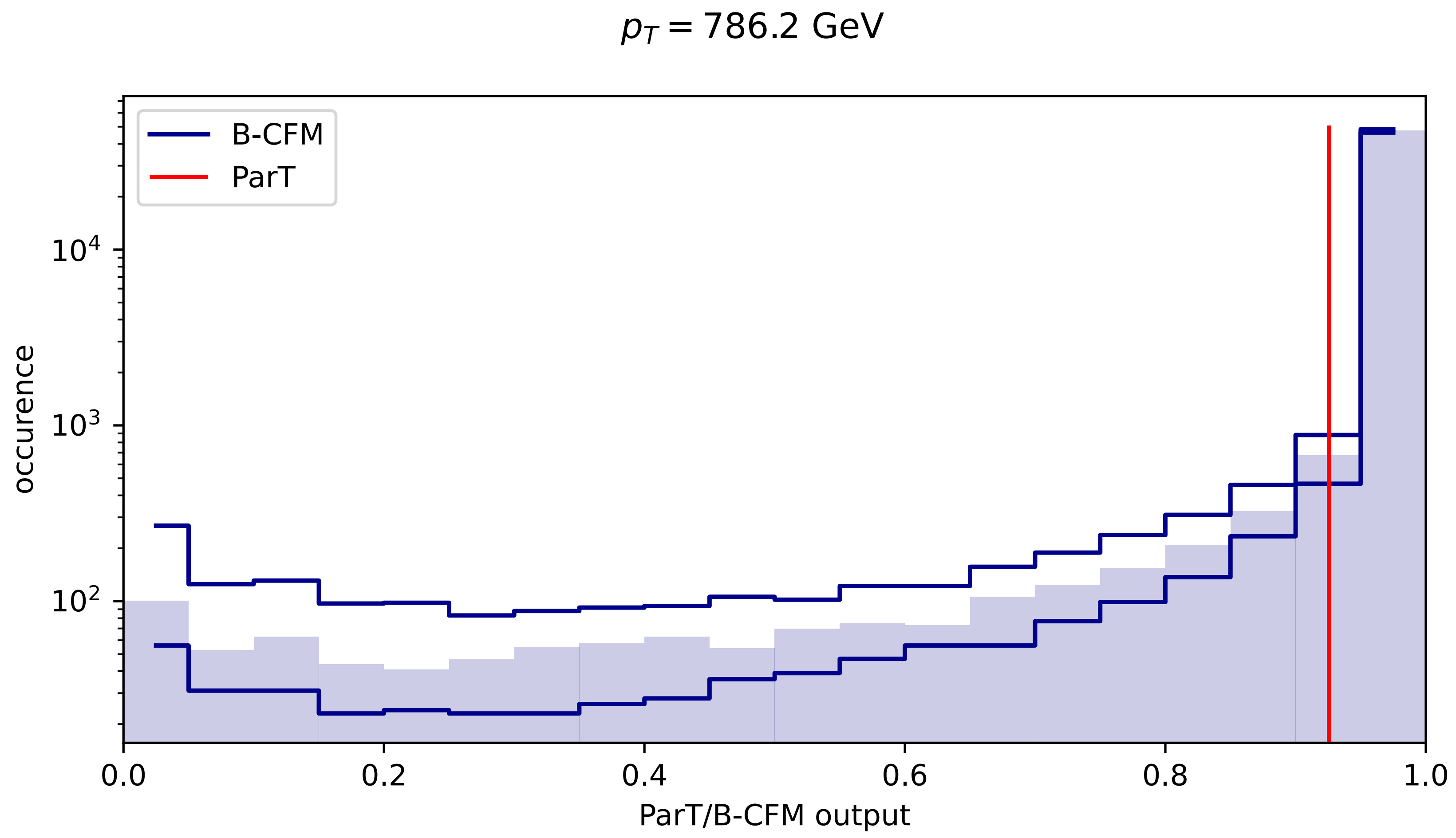
Learned Detector Smearing Distribution



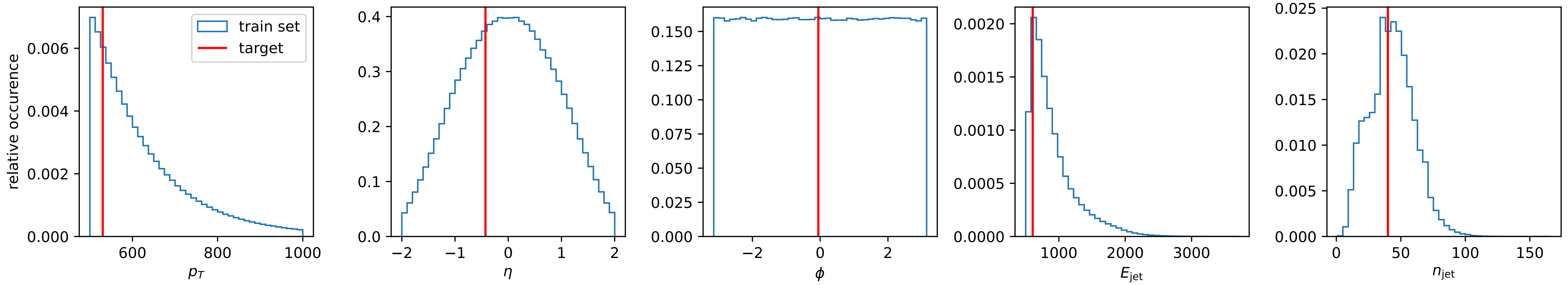
1

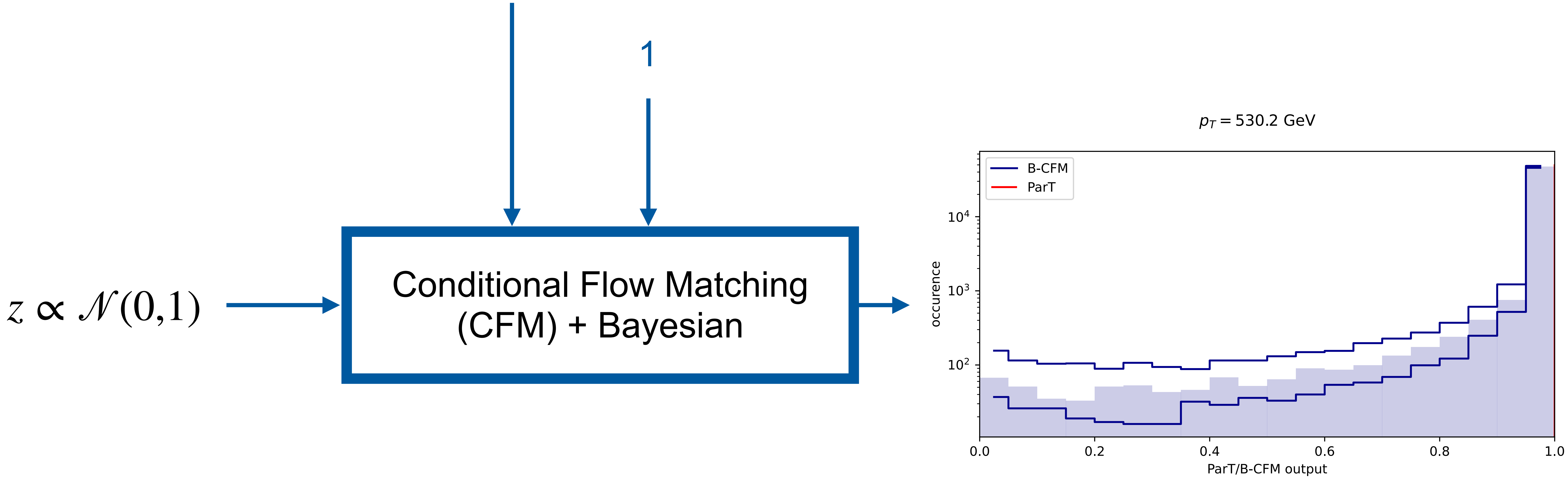
$z \propto \mathcal{N}(0,1)$

Conditional Flow Matching (CFM) + Bayesian

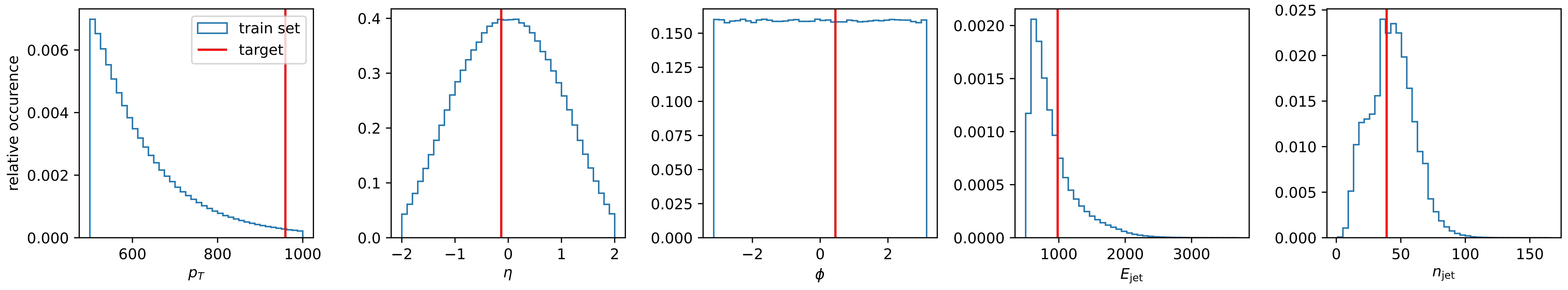


Learned Detector Smearing Distribution





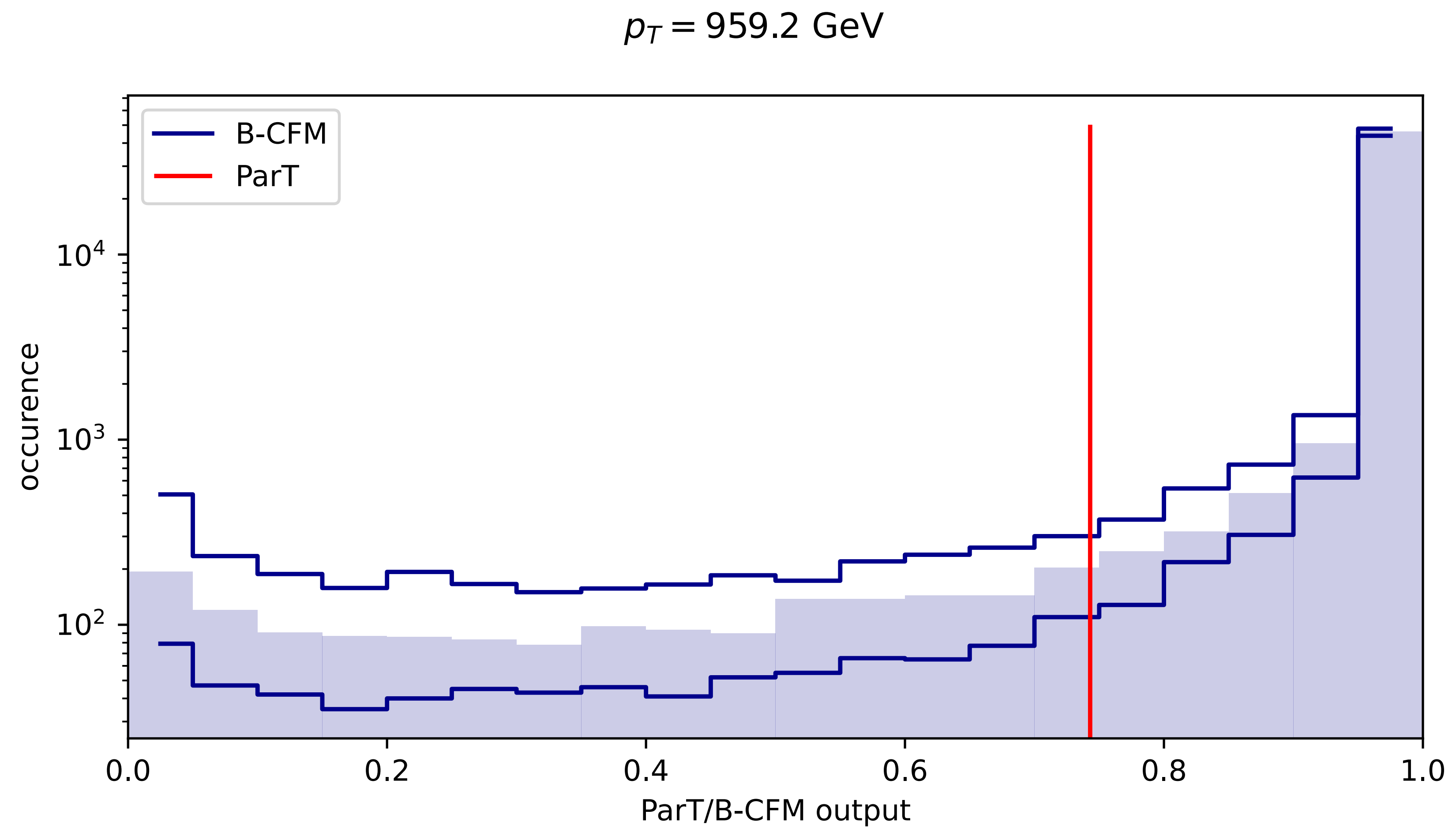
Learned Detector Smearing Distribution



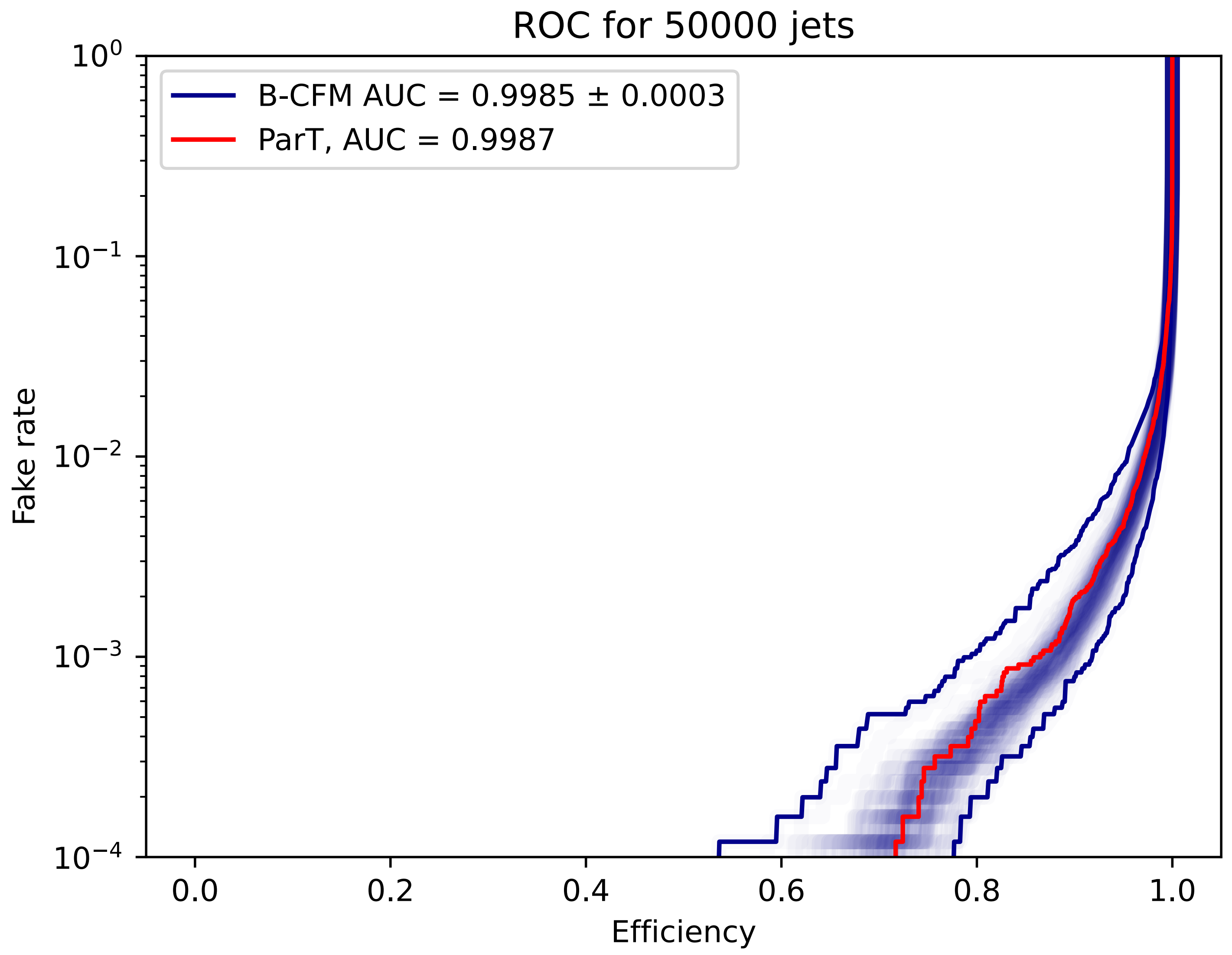
1

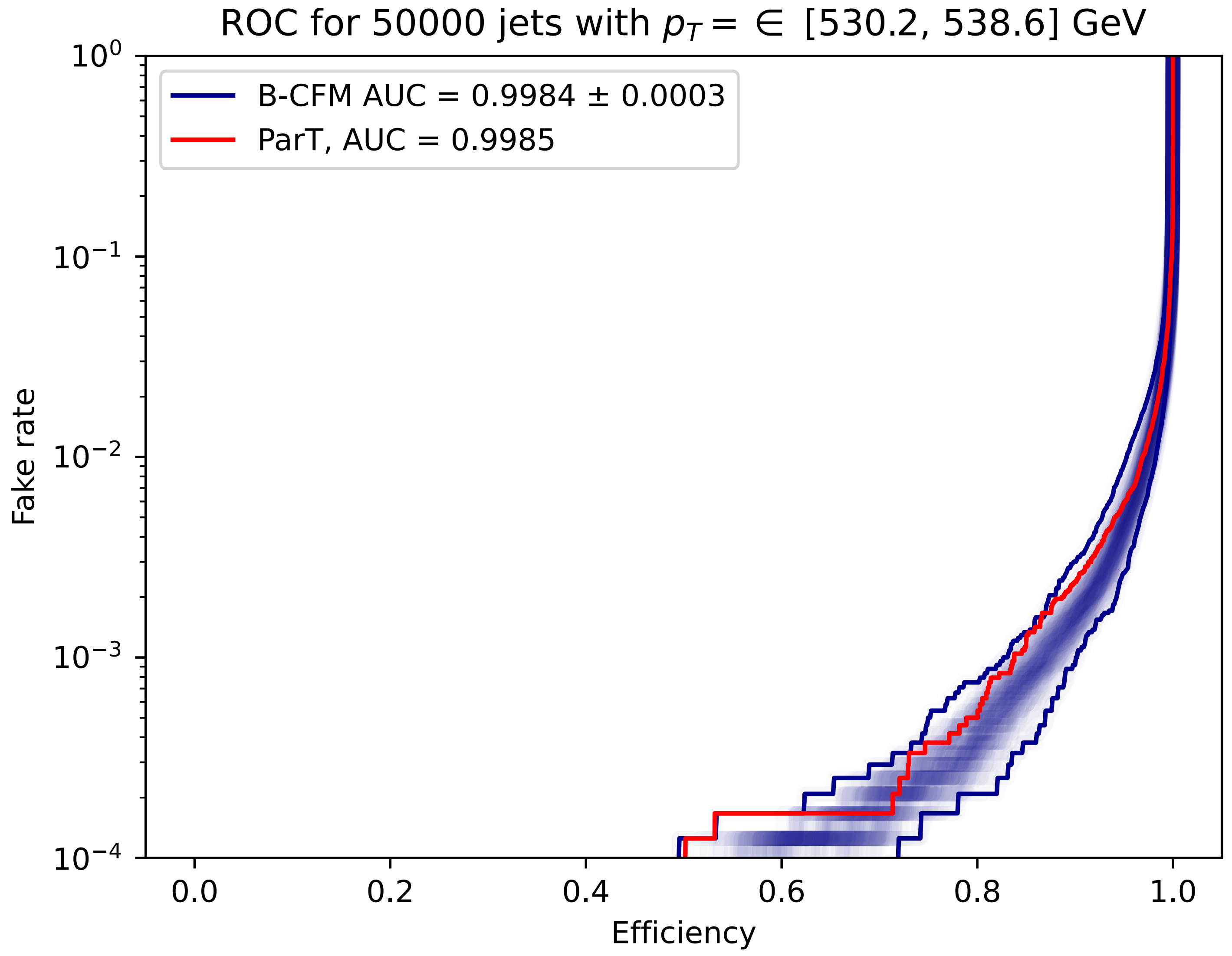
$z \propto \mathcal{N}(0,1)$

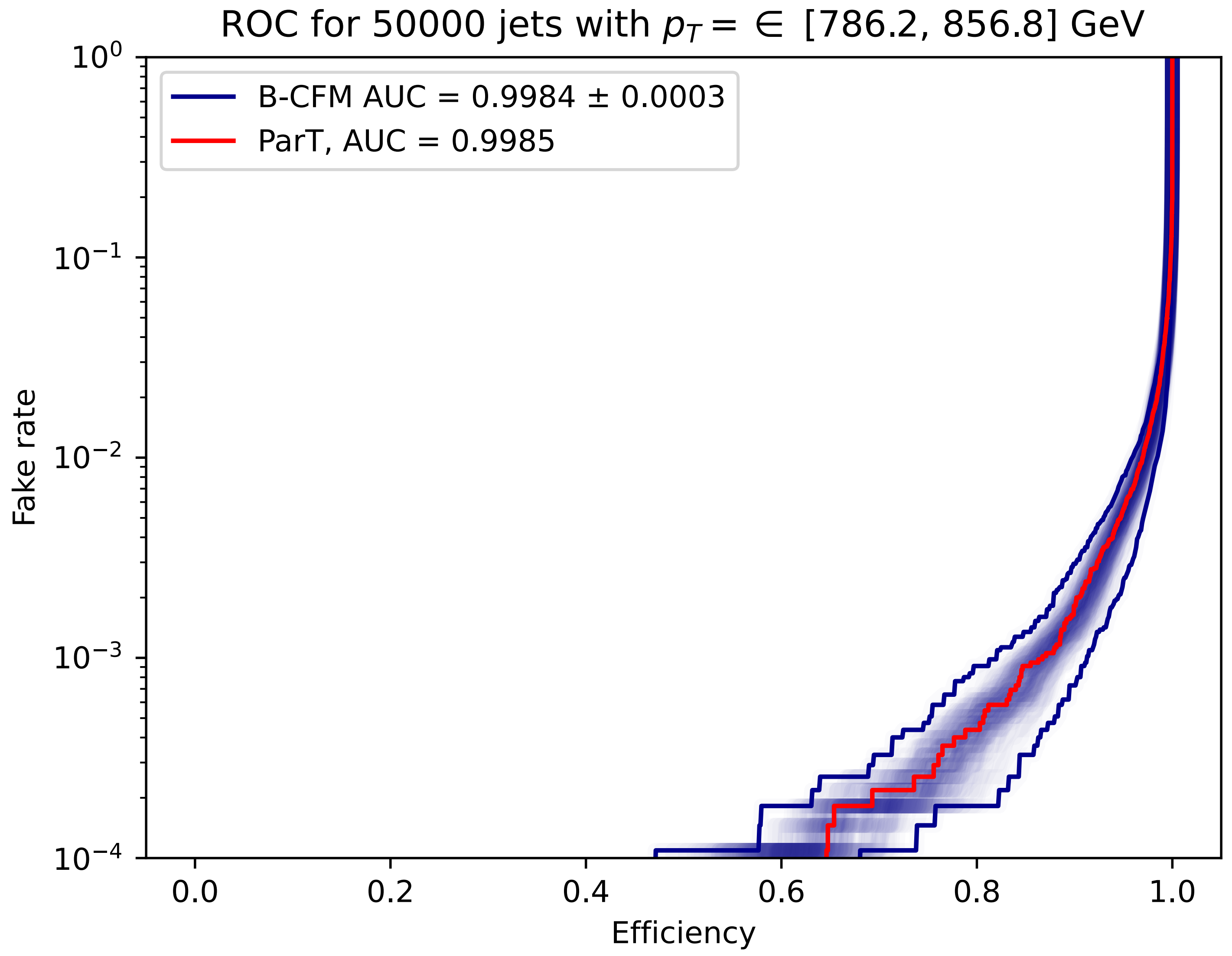
Conditional Flow Matching (CFM) + Bayesian

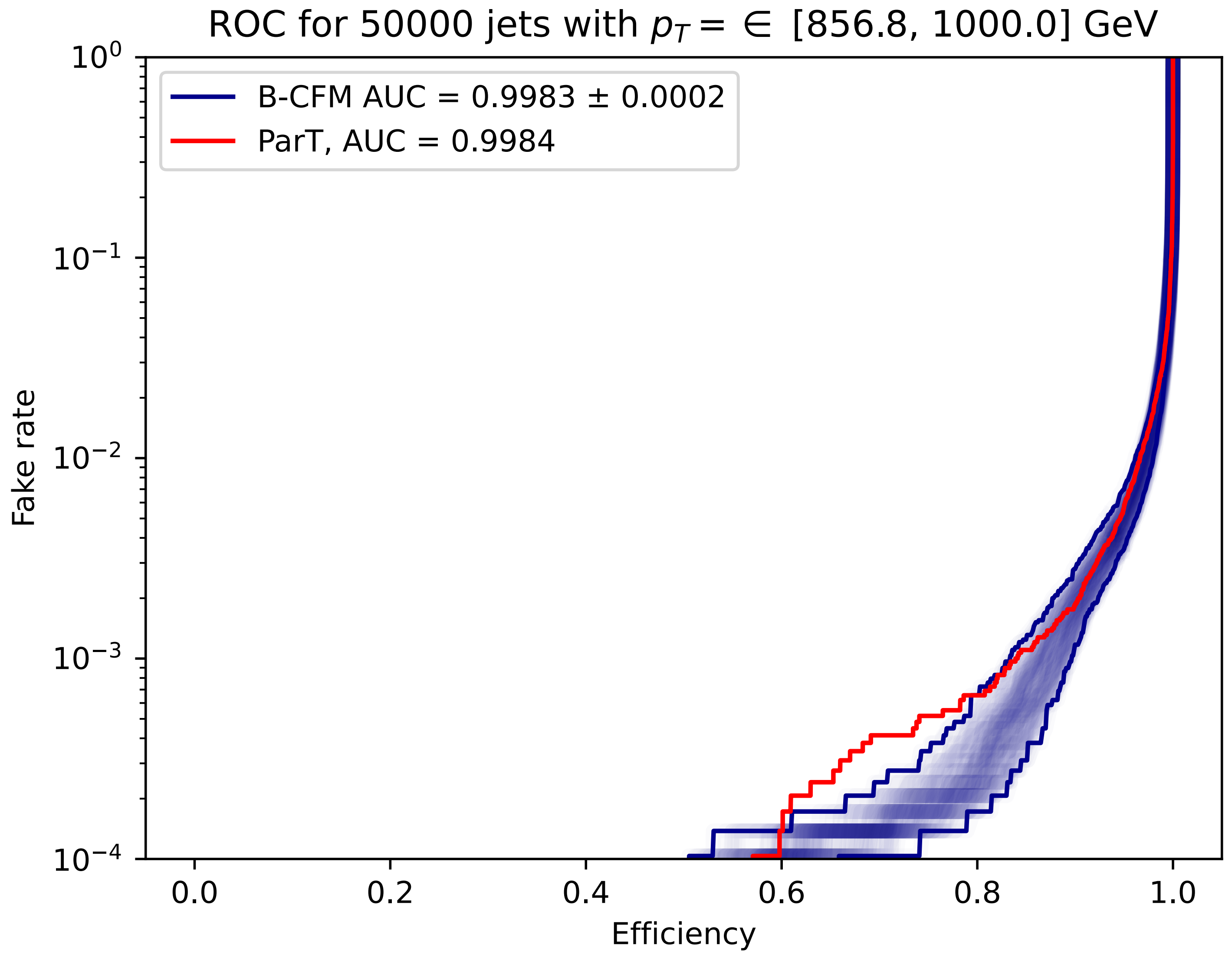


Predicted ROC

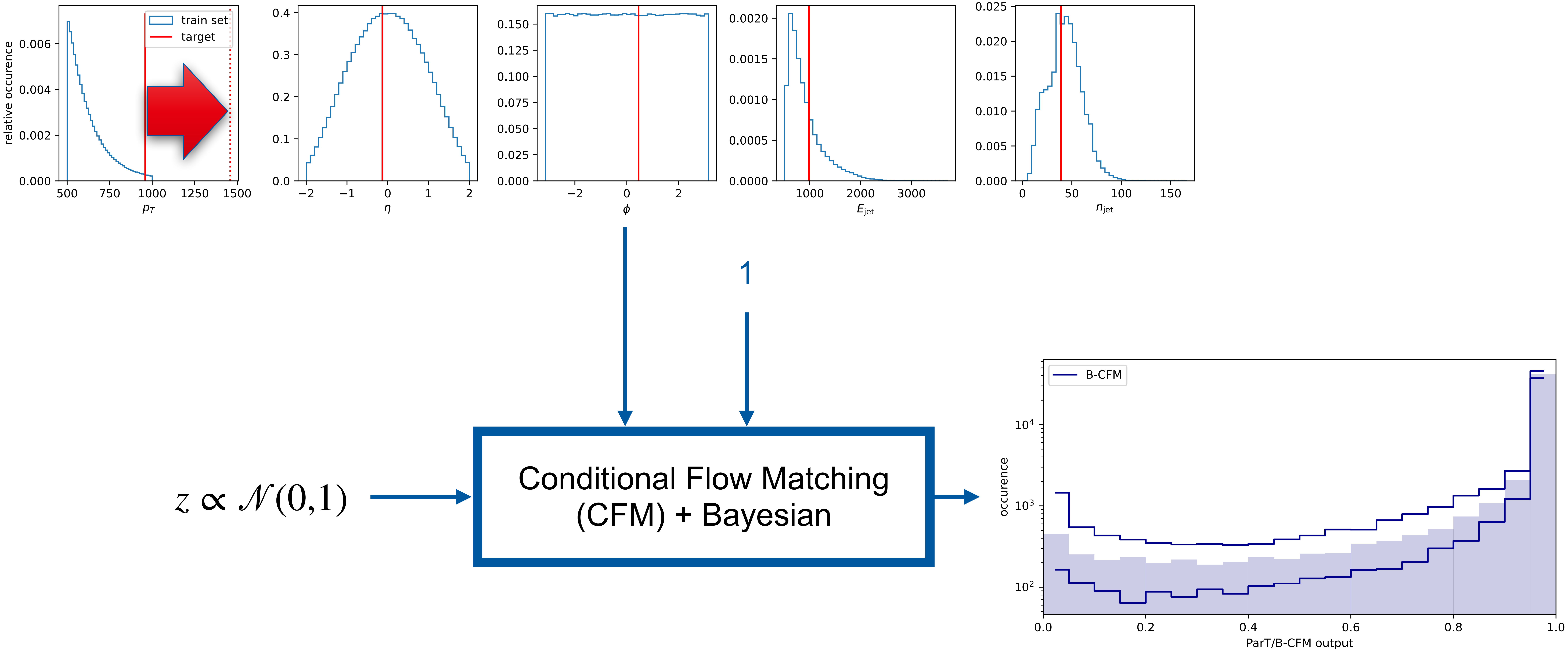




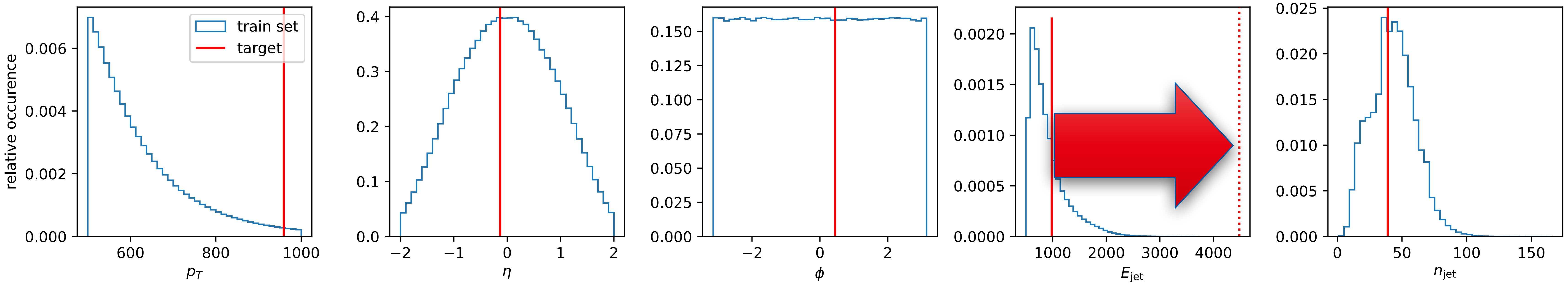




Unphysical Inputs



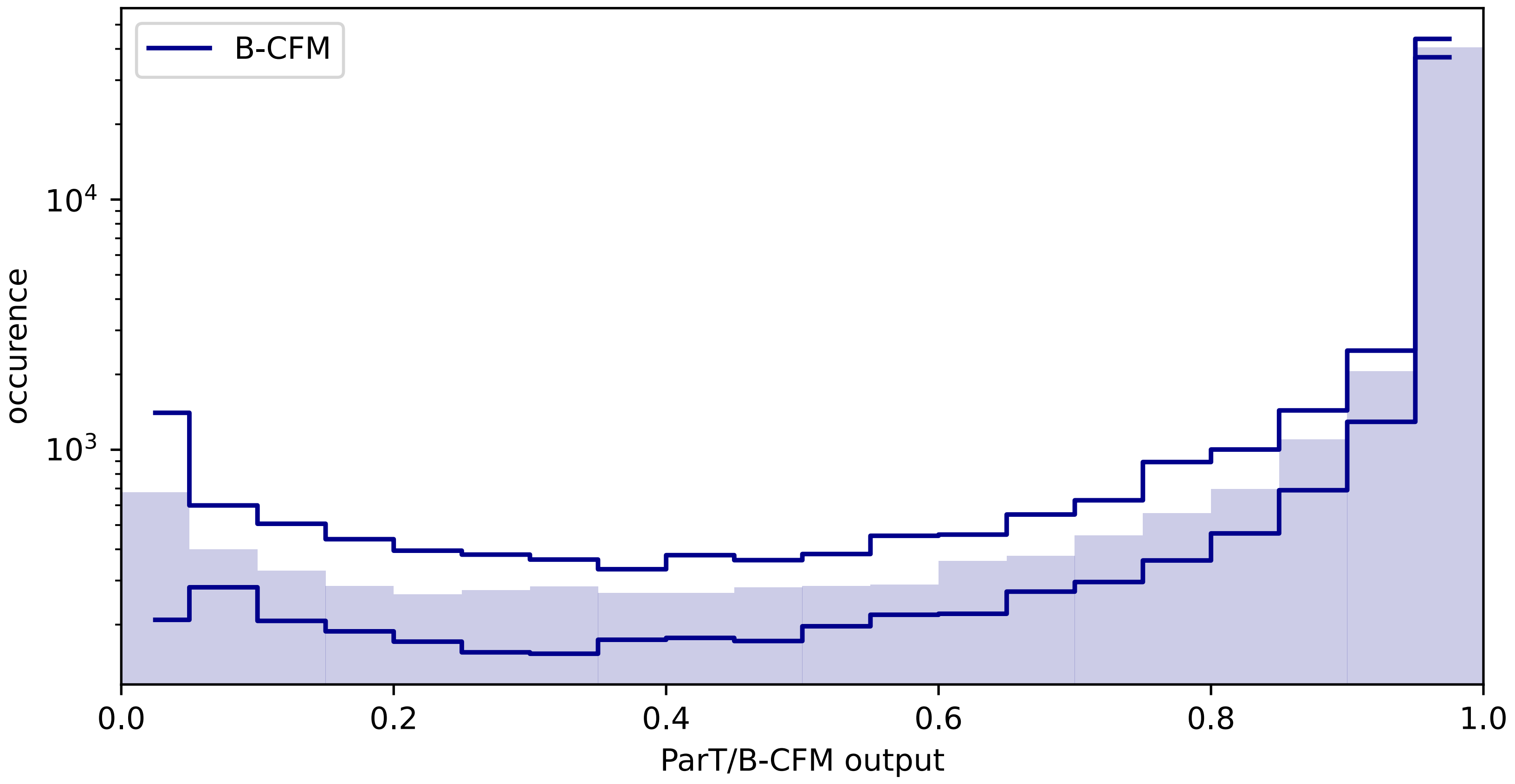
Unphysical Inputs



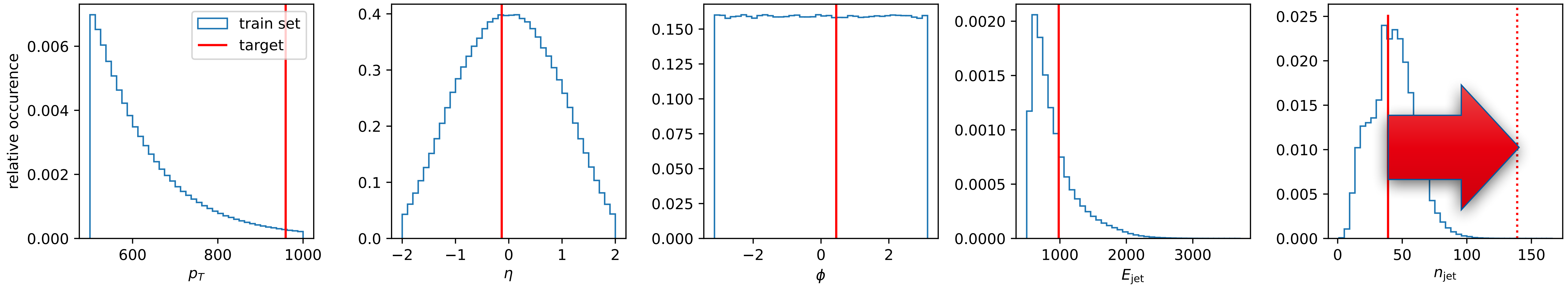
1

$z \propto \mathcal{N}(0,1)$

Conditional Flow Matching (CFM) + Bayesian



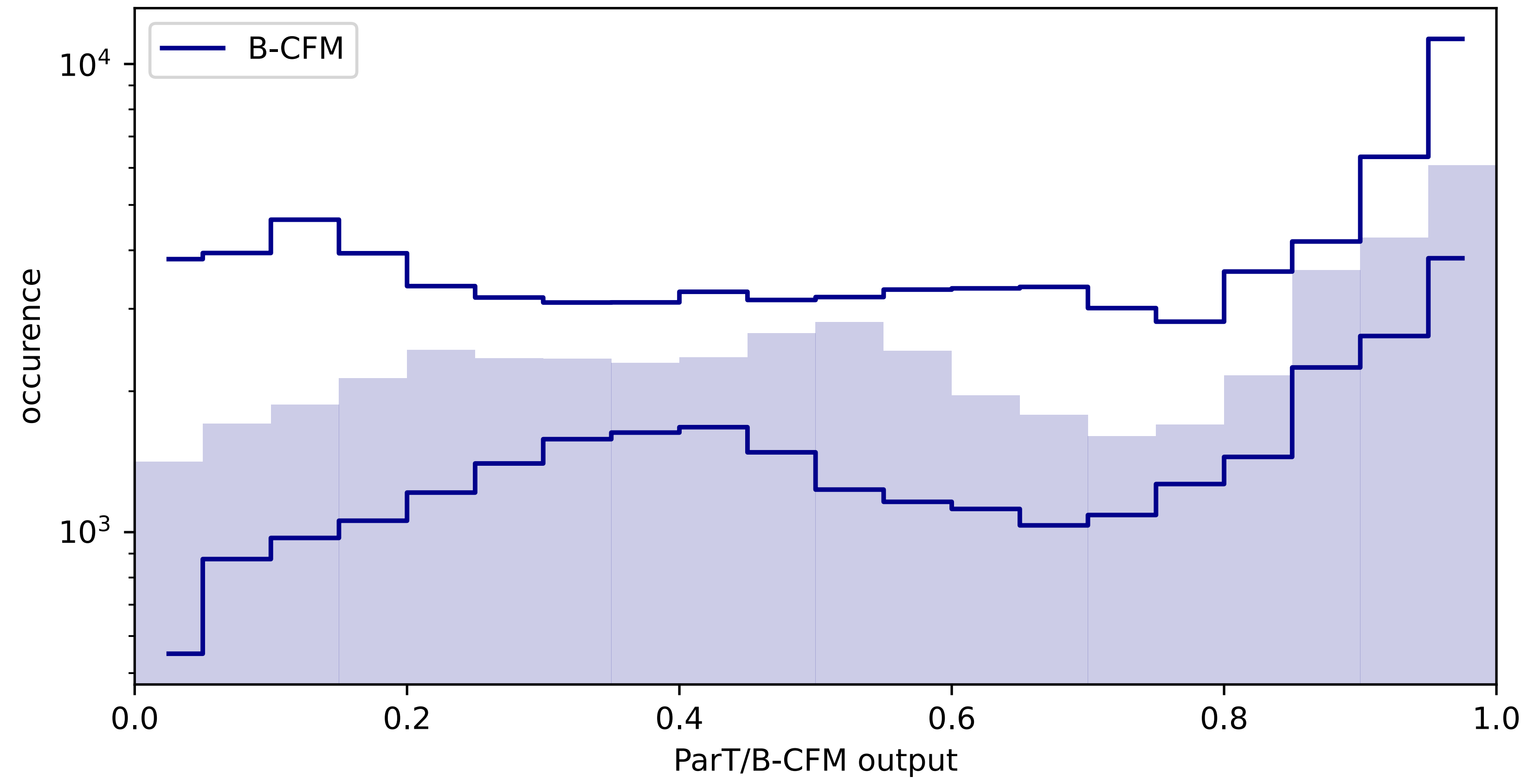
Unphysical Inputs



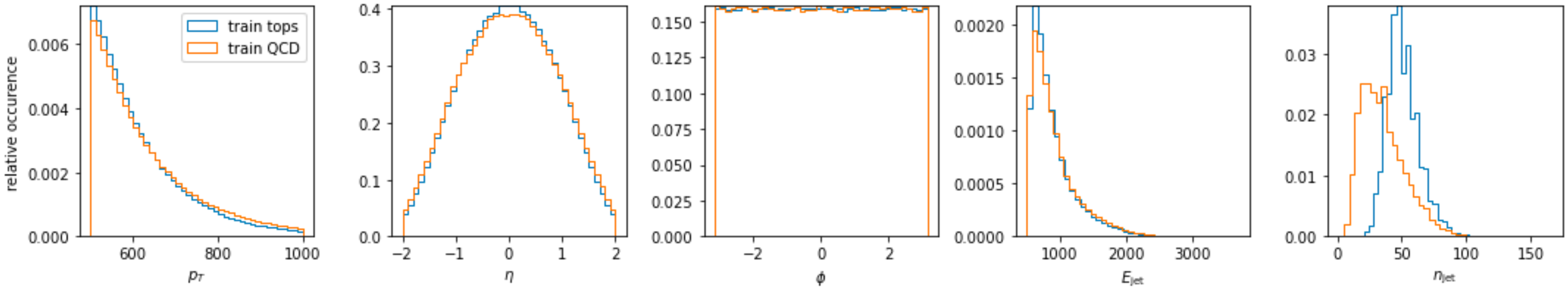
1

$z \propto \mathcal{N}(0,1)$

Conditional Flow Matching (CFM) + Bayesian



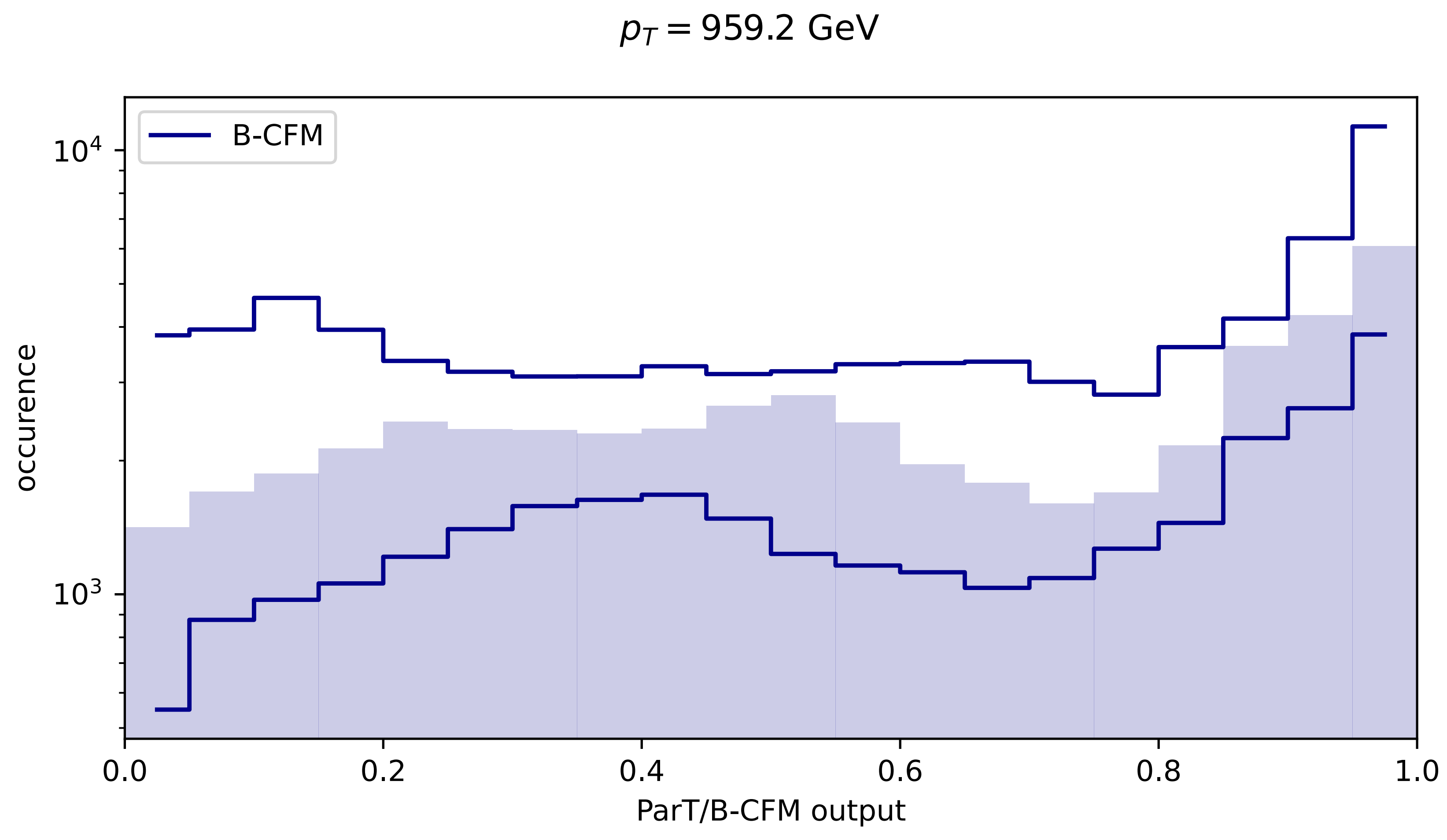
Top vs. QCD jets

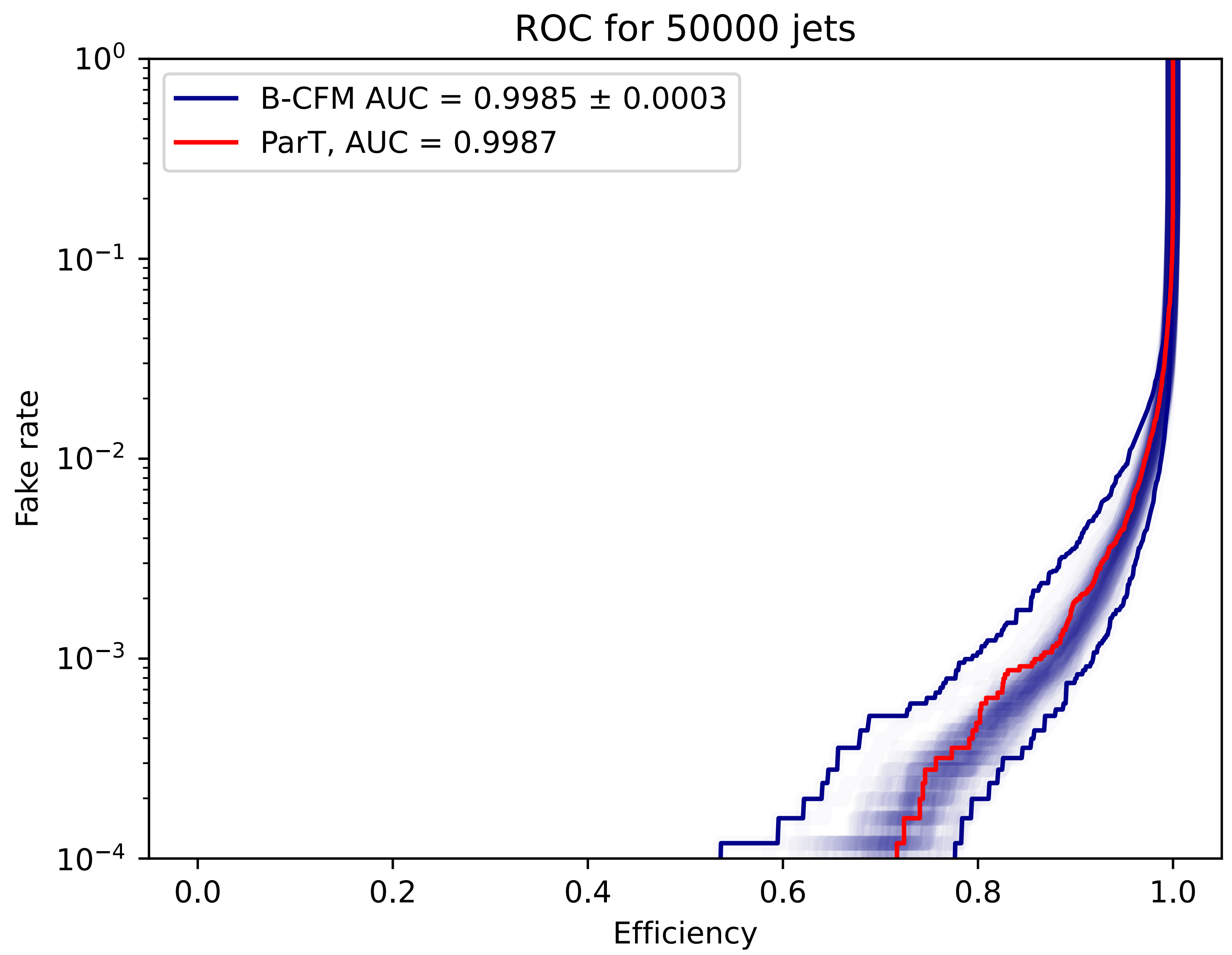


Conclusion

- CFM model can can predict the in-distribution behavior of a large classifier well
  - Independent of detector-level data
  - Can be shared with in analysis

- Further investigation of Bayesian methods to fix out-of-distribution predictions for all dimensions



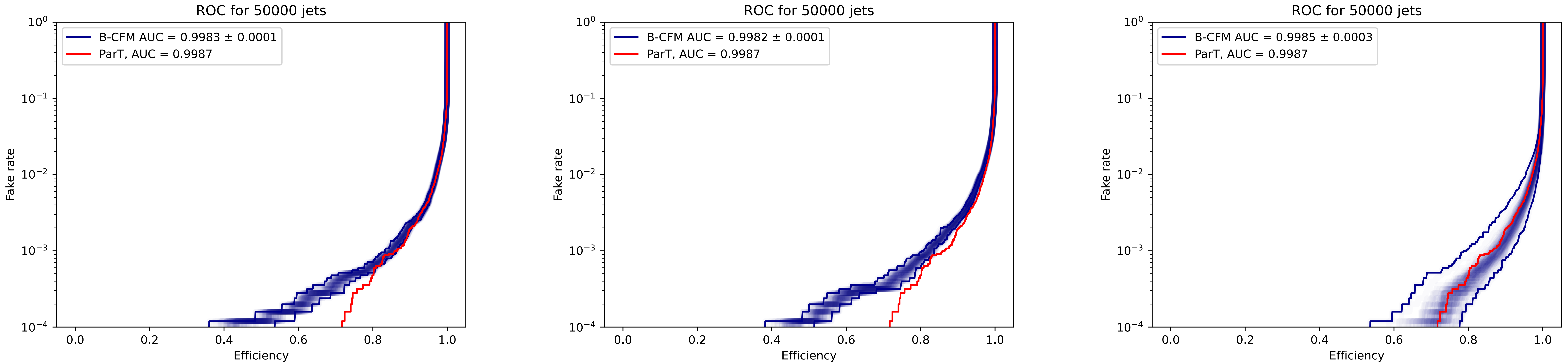


Effects of the Prior Parameter c

Bayesian Conditional Flow Matching:

- Bayesian loss $\mathscr{L}_{\text{BNN}} = \text{KL}\left[q(\theta), p\left(\theta \mid x\right)\right] = -\int \mathrm{d}\theta\, q(\theta)\, \log p\left(x \mid \theta\right) + \text{KL}[q(\theta), p(\theta)] +$ const.
- connect both $\mathscr{L}_{\text{B-CFM}} = \left\langle \mathscr{L}_{\text{CFM}} \right\rangle_{\theta \sim q(\theta)} + c\text{KL}[q(\theta), p(\theta)]$, with $q(\theta)$ uncorrelated Gaussian shape

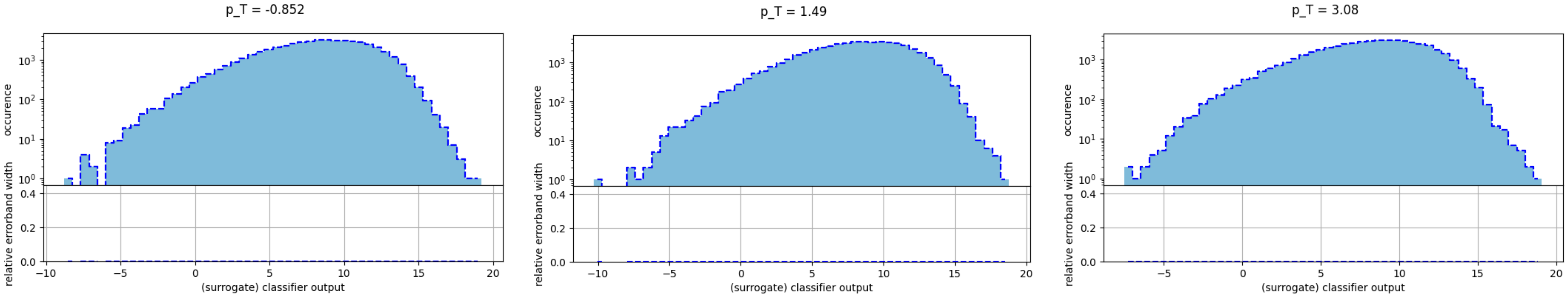
$c = 0.01$ $c = 1$ $c = 100$



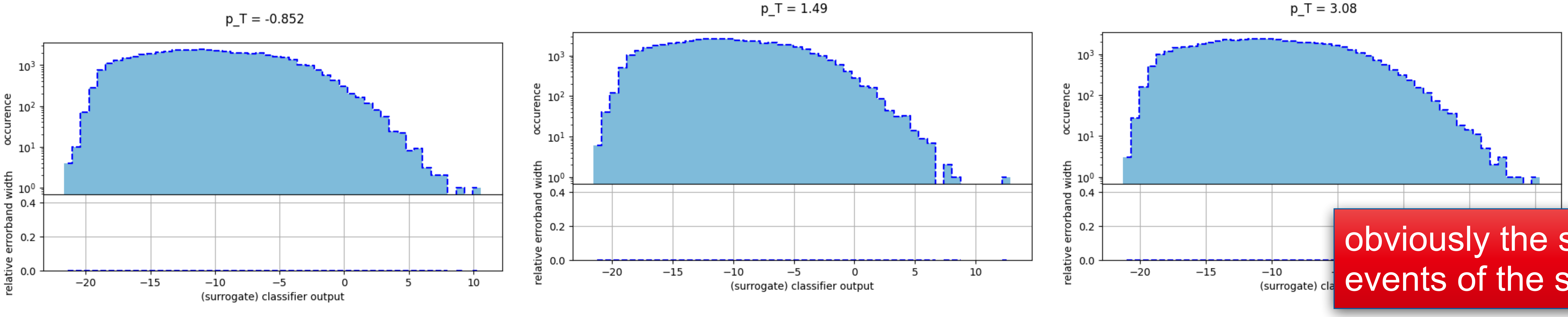
What if only trained on truth?

DASHH

top jets



not top jets



obviously the same for events of the same class