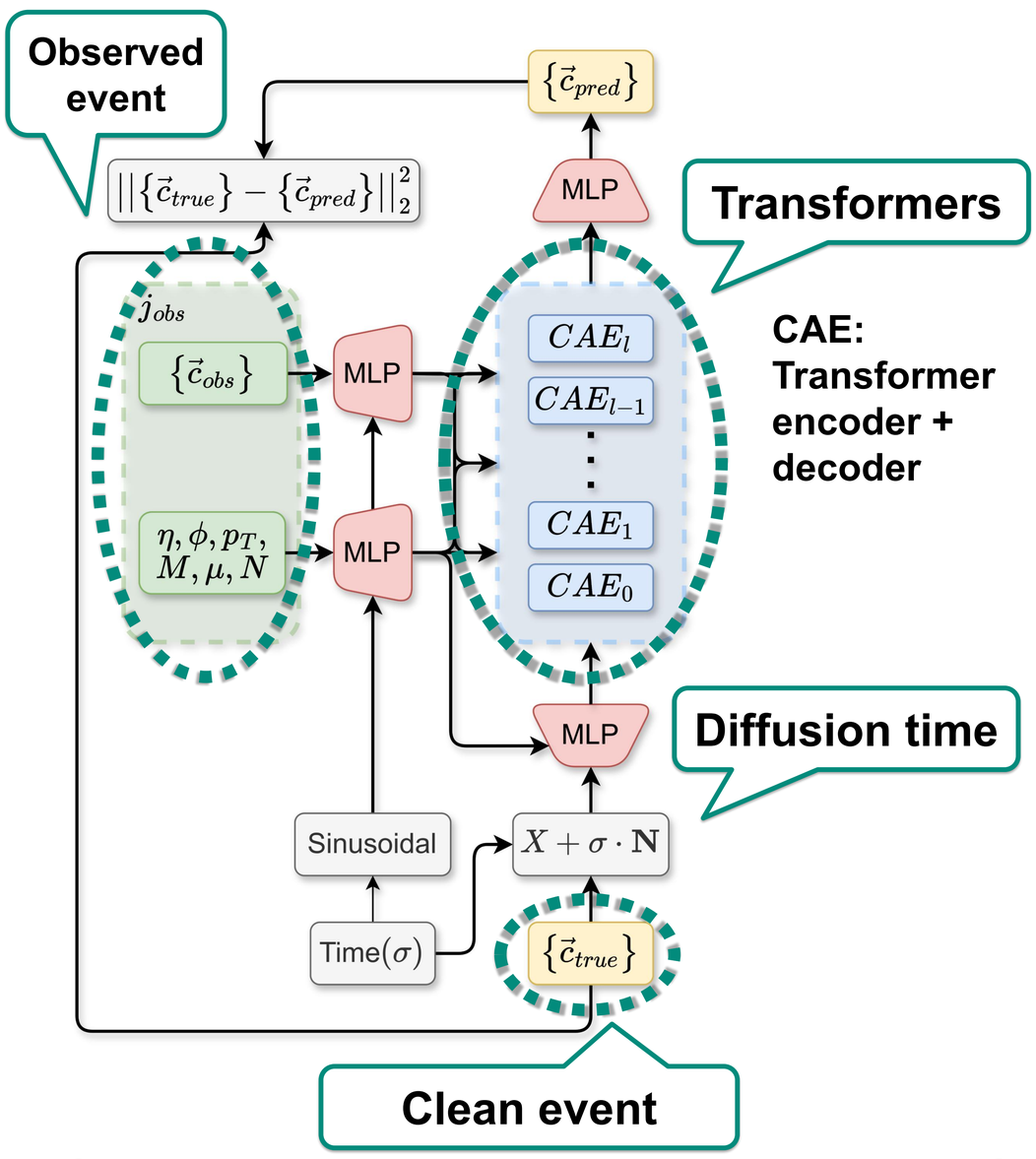
Noise removal using diffusion models

Variational inference for pile-up removal (VIPR)

WARWICK
THE UNIVERSITY OF WARWICK

Set-to-set diffusion

Training setup for EDM for set-to-set diffusion

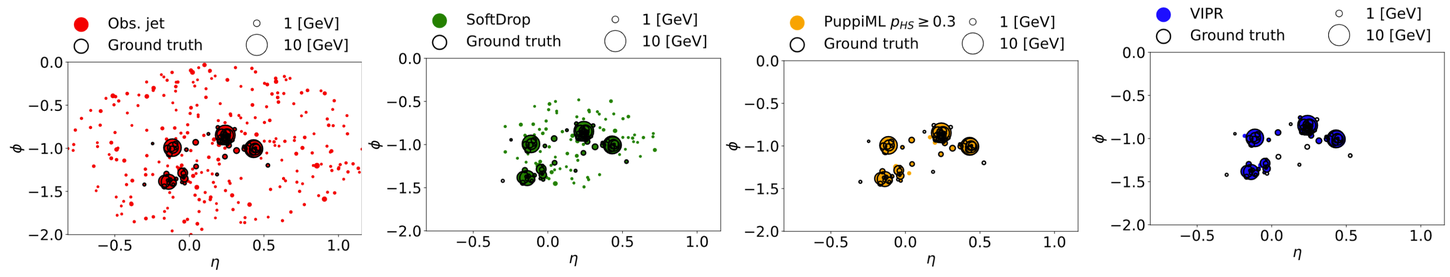


CAE: Transformer encoder + decoder

Results of set-to-set generation

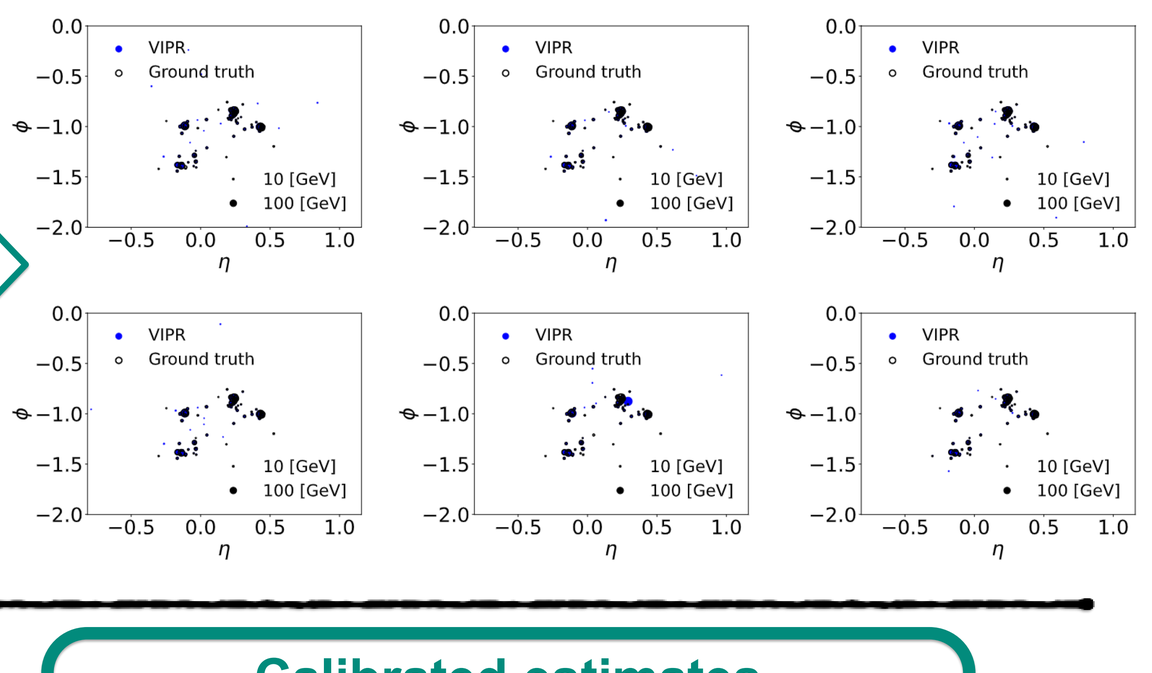
The model is trained to generate a pile-up free jet from an observed jet contaminated with pile-up. Standard methods use classification schemes (SoftDrop, PuppiML)

Generated Jets:



(a) Observed jet (b) SoftDrop jet (c) PuppiML jet (d) Vipr jet

Multiple solutions can be generated from the posterior of an event



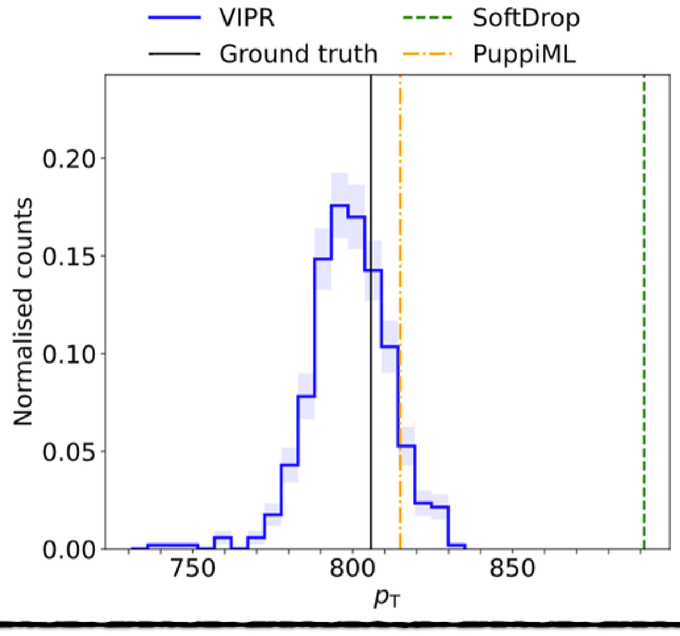
Posterior estimation using VIPR

Using a generative model for probabilistic regression to generate the posterior distribution of an estimate

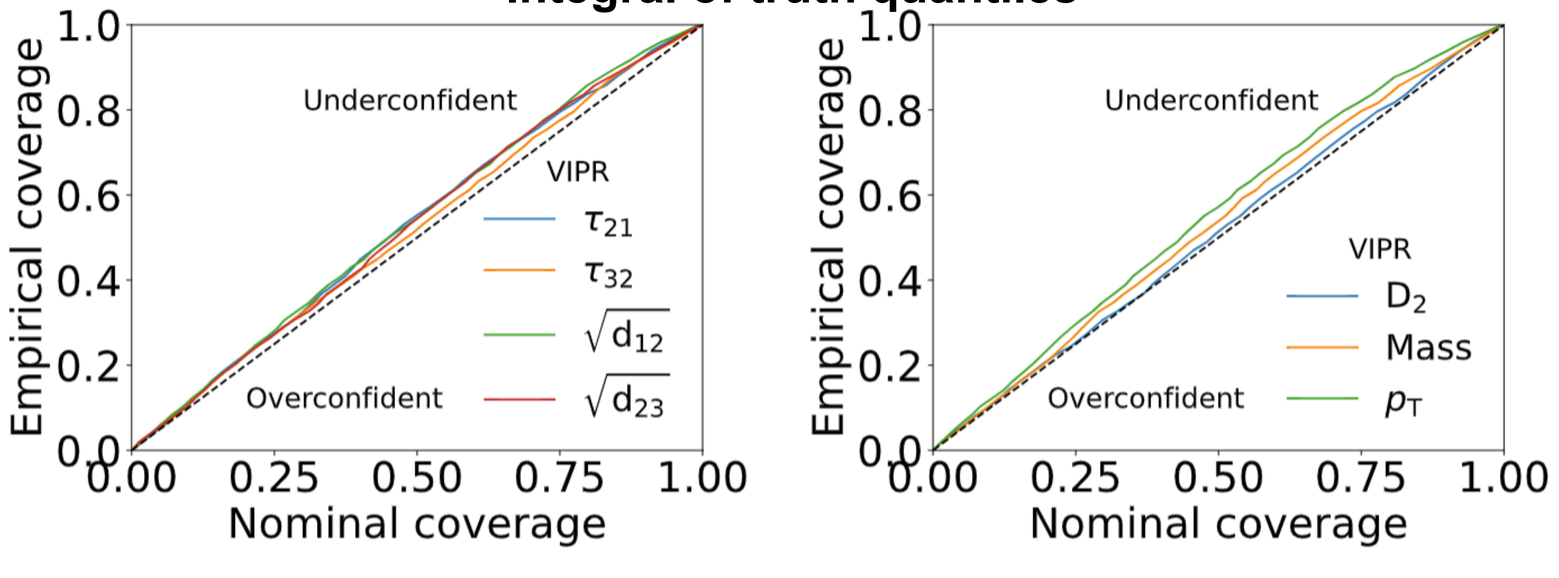
Calibrated estimates

If your posteriors are calibrated, the nominal and empirical coverage should be identical

Posterior of a single event in p_T

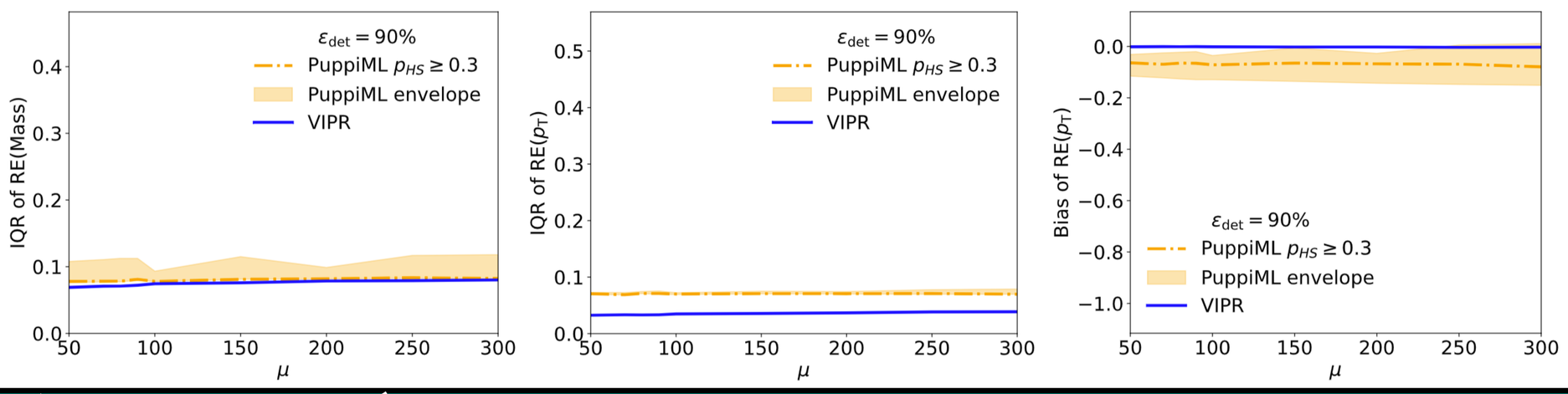


Integral of truth quantiles



Accounting for inefficiencies in the detector:

Let's say your detector merges some particles together or does not reconstruct all particles. Generative models can account for missing observed particles or other inefficiencies, whereas classifier-based models cannot.
We introduce a flat inefficiency of 10% (i.e., 10% of particles are missing):



Link to VIPR

UNIVERSITÉ DE GENÈVE

Malte Algren, Tobias Golling, Christopher Pollard and John Andrew Raine

EuCAIFCon 2025