Generative Networks for Precision Enthusiasts

Theo Heimel

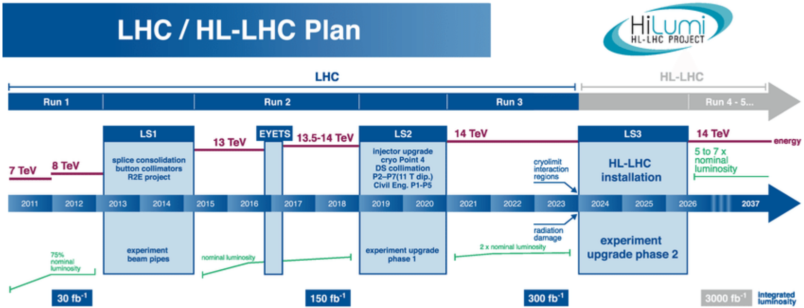
May 2022

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

arXiv: 2110.13632

A. Butter, TH, S. Hummerich, T. Krebs, T. Plehn, A. Rousselot, S. Vent

Introduction



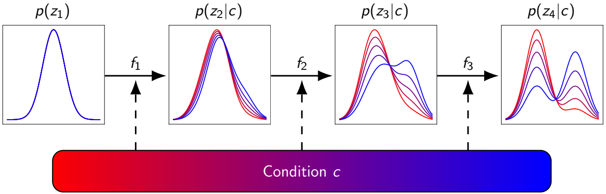
- Amount of LHC data will increase drastically
  $\rightarrow$ need similar amount of simulated events
- Not enough computational resources for event generation
- **Machine learning to accelerate event generation**

Deep learning in event generation

- Network architectures used for event generation
  → Generative adversarial networks [Butter et al, 1907.03764]
  → Variational autoencoders [Howard et al, 2101.08944]
  → **Normalizing flows / Invertible Neural Networks (INNs)** [Verheyen, Stienen, 2011.13445]
- Other physics applications of INNs
  → Phase space generation [Bothmann et al, 2001.05478] [Gao et al, 2001.05486] [Gao et a;, 2001.10028] [Chen et al, 2009.07819]
  → Detector simulation [Krause et al, 2110.11377]
  → Anomaly detection [Nachman, Shih, 2001.04990]
  → Density estimation [Brehmer, Cranmer, 2003.13913]
  → Parton shower unfolding [Bellagente et al, 2006.06685]

Invertible Neural Networks (INNs)

- INNs (normalizing flows): **chain of learnable, invertible transformations**
- Transform latent distribution (e.g. Gaussian) into distribution of interest

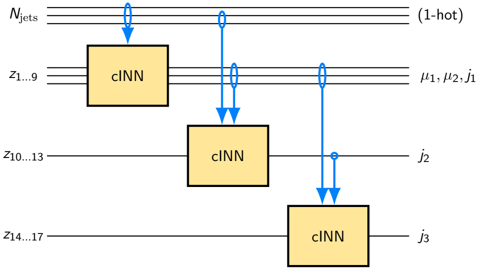


- Training: Evaluate in backward direction to get z_1 (latent space)
  $\rightarrow$ maximize log-likelihood (from change of variables formula)

$$\mathcal{L} = \log p(z_n) = \log p(z_1) + \log \left| \det \frac{\partial f^{-1}}{\partial z_n} \right|$$

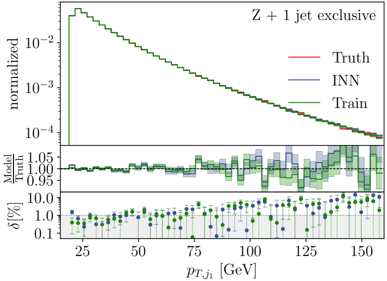
- Sampling: Sample from $p(z_1)$, evaluate forward to get z_n

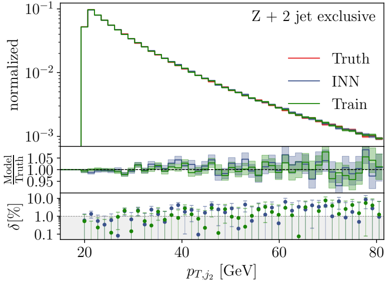
Physics process and network architecture



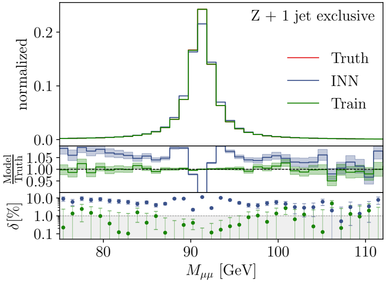
- Training data:
  - → leptonic Z decay with 1-3 jets
    $pp \to \mu_+ \mu_- + jets$
  - → include shower and hadronization
  - → no detector effects
- INNs work on fixed dimension
  → **chain of conditional INNs for variable jet multiplicity**
- INN built from cubic spline coupling blocks [Durkan et al, 1906.02145]

Event generation results

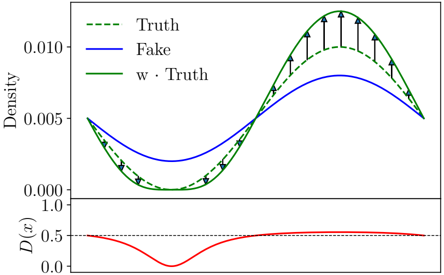




- Percent-level accuracy in bulk
- Challenges:
  - → Features hidden in correlations e.g. Z mass peak
  - → Topological problems e.g. ΔR cuts



Joint training with a discriminator



- Improve generative model with discriminator reweighting

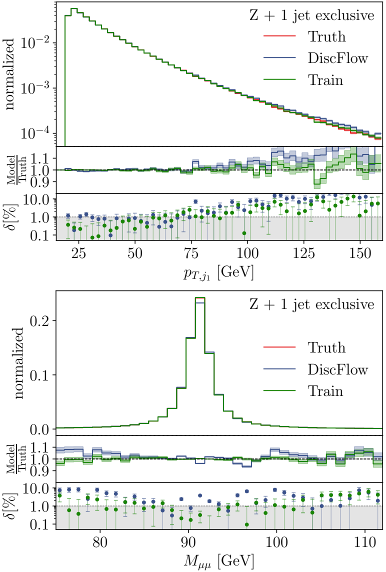
  [Diefenbacher et al, 2009.03796]

  → Discriminate Truth vs. INN events

  → Problem: weighted events
- **Solution: Joint training of INN and discriminator**
- Reweight INN training data to over-exaggerate features

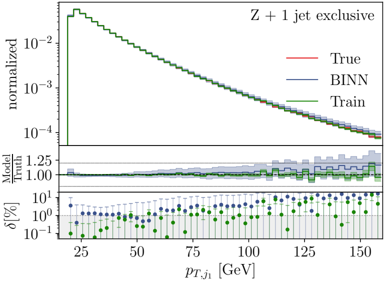
  → No Nash equilibrium between two networks needed
- Give challenging observables to discriminator as additional inputs

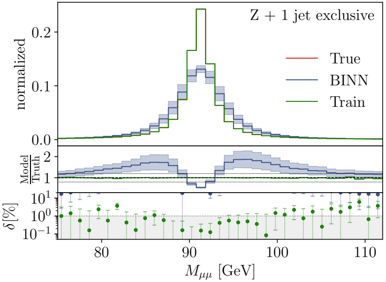
DiscFlow results



- Discriminator input: INN input, $M_{\mu\mu}$, $\Delta R_{i,j}$
- Similar precision as without DiscFlow for most observables
- **Improved precision for mass peak**
- However: only effective in regions populated by training data

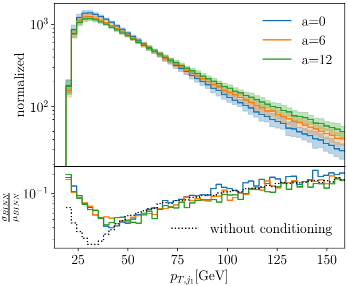
Training uncertainty

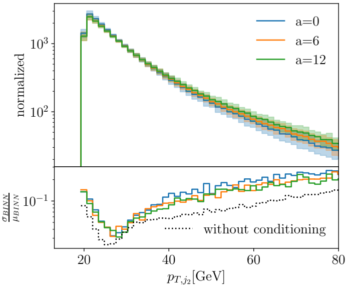




- Estimate training uncertainty from ensemble of networks
  $\rightarrow$ uses a lot of resources
- Solution: Bayesian neural networks
  [MacCay, 1995] [Neal, 2012]
  [Bellagente et al., 2104.04543]
  $\rightarrow$ parameters drawn from Gaussian
  $\theta_i \sim \mathcal{N}(\mu_i, \sigma_i)$
  $\rightarrow$ new loss term to learn μ_i, σ_i
- Bin-wise means and standard deviations over multiple samples from parameter distribution
  $\rightarrow$ errorbars for histograms
- Caveat: unlearnable features not part of error

Systematic uncertainties



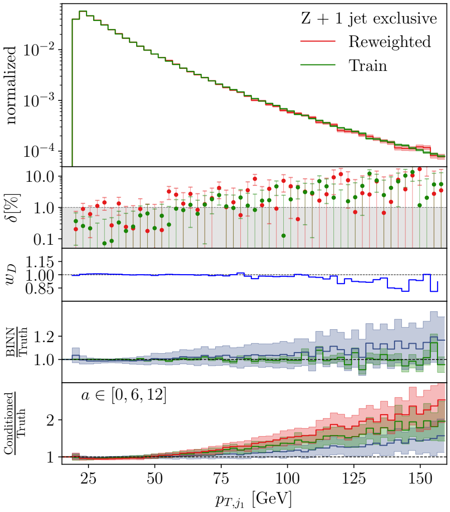


- Toy model for a theory uncertainty
  $\rightarrow$ nuisance parameter a
  $\rightarrow$ shift in event weights

$$w = 1 + a \left(\frac{p_{T,j_1} - 15\text{ GeV}}{100\text{ GeV}} \right)^2$$

- Condition network on a
  $\rightarrow$ sample a during training
- Vary a during sampling
- Larger BINN uncertainty due to additional sources of correlation

Summary



- Percent-level accuracy in bulk of distribution
- Joint training to use discriminator feedback
- Control training uncertainties with Bayesian INNs
- Understand systematic uncertainties by adding conditions
- **Event generation with INNs: High precision while having control over uncertainties**