Unfolding with conditional invertible networks

[2006.06685]

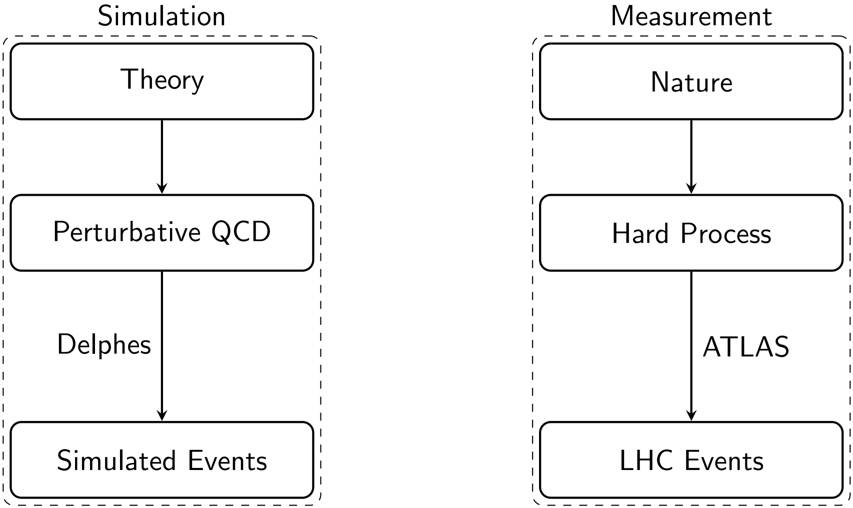
Armand Rousselot

with Marco Bellagente, Anja Butter, Gregor Kasieczka, Tilman Plehn, Ramon Winterhalder, Lynton Ardizzone, and Ullrich Köthe

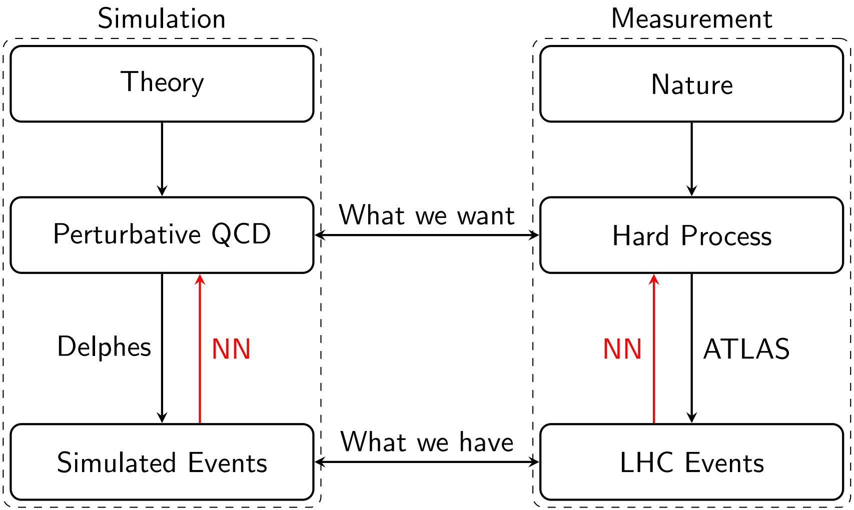
ITP, Universität Heidelberg

05 November 2020

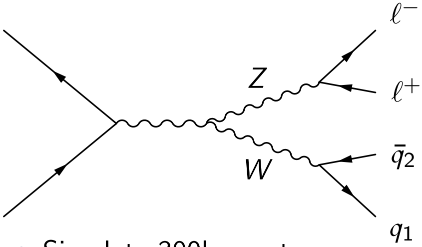
Inverting the simulation



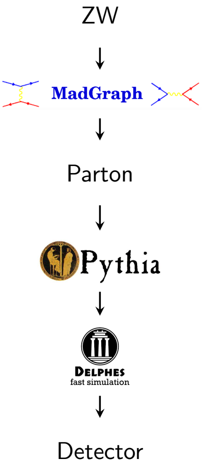
Inverting the simulation



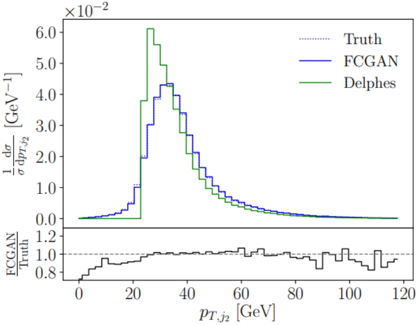
Simulating and Inverting Detector Effects

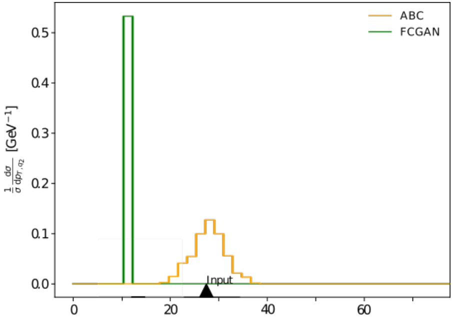


- Simulate 300k events
- Learn detector $\rightarrow$ parton
- Want **probabilistic** mapping



"Just GAN it"



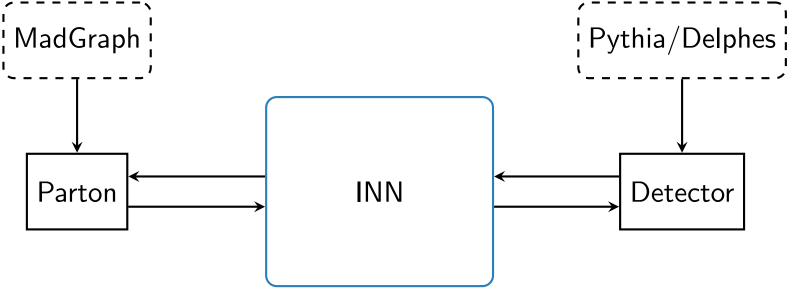


- Good results on full phase space

- Unfold single event 3200 times
- Compare to closest dataset events

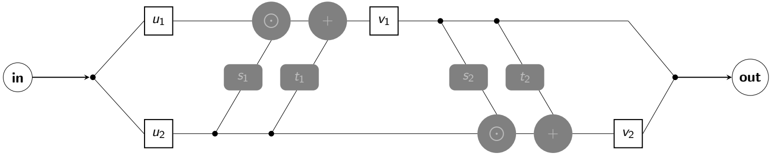
Details: [1912.00477]

Invertible Neural Networks



- Learn bijective mapping → inverse direction for free
- Same dimensionality for both sides
- Extra noise for probabilistic application → extend with noise (eINN)

Coupling Blocks

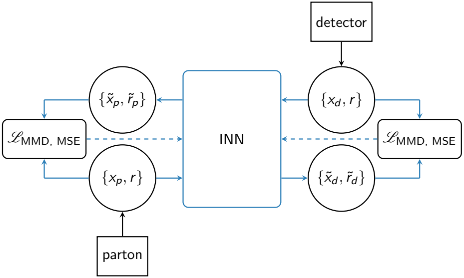


Forward: $v_1 = u_1 \times e^{s_1(u_2)} + t_1(u_2)$ $\quad v_2 = u_2 \times e^{s_2(v_1)} + t_2(v_1)$

Backward: $u_2 = (v_2 - t_2(v_1)) \times e^{-s_2(v_1)}$ $\quad u_1 = (v_1 - t_1(u_2)) \times e^{-s_1(u_2)}$

- Split input in u_1, u_2
- Invertible for arbitrary s, t
- Jacobian triangular $\rightarrow$ easily tractable

INN Training



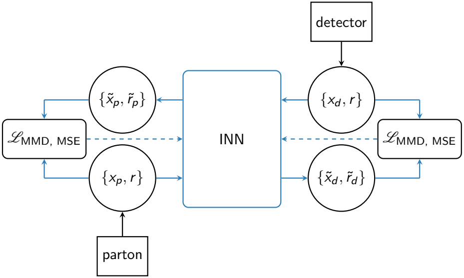
INN Training

Detector → Parton

$$\mathscr{L}_y = \mathscr{L}_{MMD}(x_p, \tilde{x}_p) + \mathscr{L}_{MSE}(x_p, \tilde{x}_p)$$

Parton → Detector

$$\mathscr{L}_x = \mathscr{L}_{MMD}(x_d, \tilde{x}_d) + \mathscr{L}_{MSE}(x_d, \tilde{x}_d)$$



Invariant mass distribution

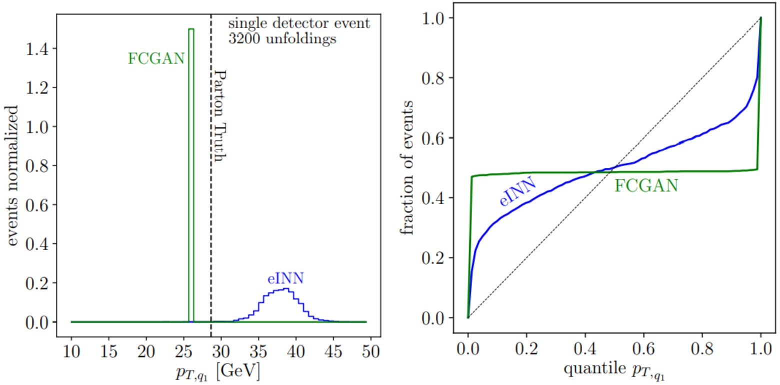
$$\mathscr{L}_M = \mathscr{L}_{MMD}(M(x_d), M(\tilde{x}_d)) + \mathscr{L}_{MMD}(M(x_p), M(\tilde{x}_p))$$

Gaussian latent distribution

$$\mathscr{L}_z = \mathscr{L}_{MMD}(r, \tilde{r}_p) + \mathscr{L}_{MMD}(r, \tilde{r}_d)$$

Maximum Mean Discrepancy: $\mathscr{L}_{MMD} = k(X, X) + k(Y, Y) - 2k(X, Y)$

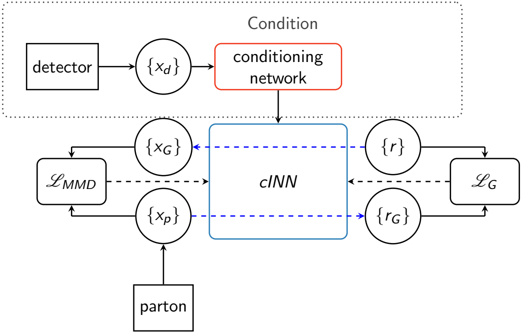
INN Results



Calibration Curves

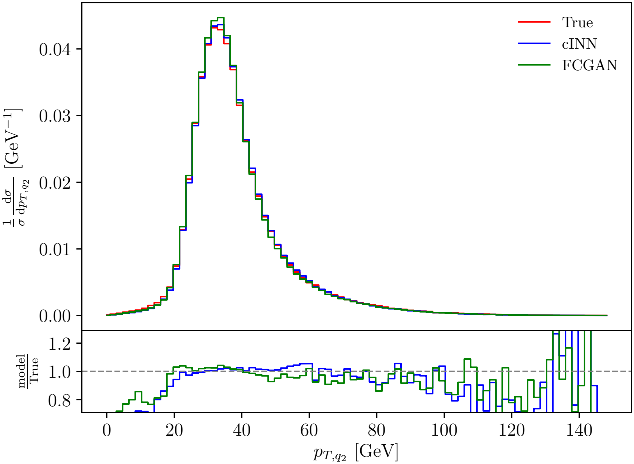
- Get parton distribution for fixed detector-level (noise sampling)
- Plot frequency of ground truth appearing in quantile x
- Expect diagonal → Noise parton correlation incorrect

Conditional Invertible Neural Networks



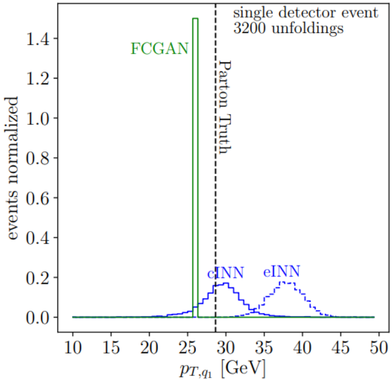
- Network conditioned on detector information
- Map parton to normal distribution
- Minimize $\mathscr{L}_G = \mathbb{E}_{i\in 1,\ldots,N}\left[\frac{||f(x_{p,i}|x_{d,i},\theta)||_2^2}{2} - \log|J_i|\right] + \tau||\theta||_2^2$

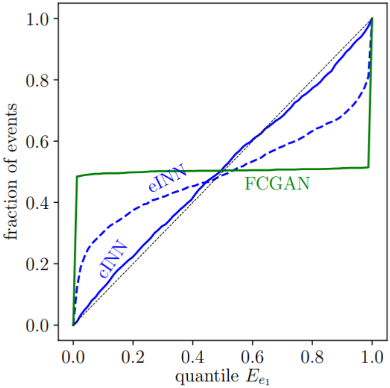
cINN Results



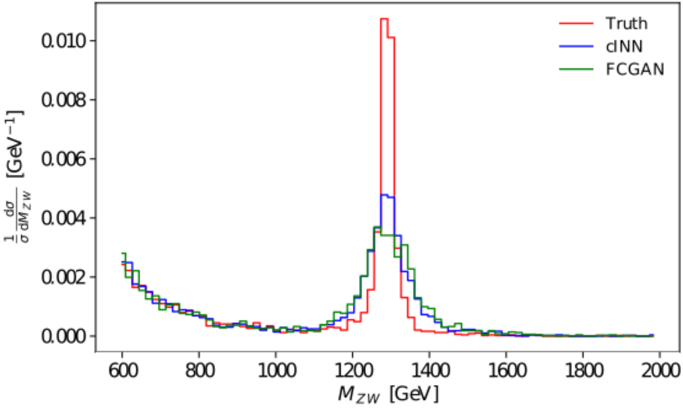
cINN and cGAN both perform well on full phase space

Calibration



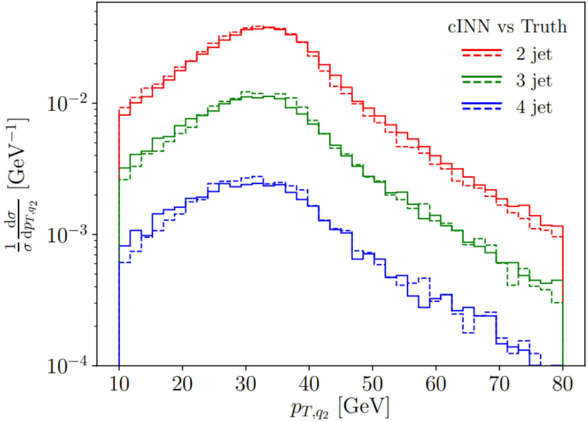


Hidden Mother Particle



- Introduce new W' particle into test data → decays into W and Z
- Adding their four-vectors results in distinct mass peak

Include Initial State Radiation



Train on inclusive channel
Evaluate on exclusive 2/3/4 jets

Conclusion

- Unfolding of ZW Process works with cGAN/eINN/cINN
- Statistically coherent results with cINN
- Can reconstruct unseen structures on parton level (e.g. W' particle)
- Simultaneous unfolding of different numbers of jets