**The Landscape of Unfolding with Machine Learning Paper Review**

Paper Review by Tyler Kim

**Motivation**

* Predictions from quantum field theory are at partons level which require precise and detailed simulations
* Original approach was to use *forward inferencing* which fold predictions using QCD effects, hadronization, and detector response
  + There are couple of problems with *forward inferencing*
    - Requires access to data
    - Requires access to detector simulation
* *Unfolding* is an alternative approach has advantages and problems
  + *Unfolding*: data are adjusted to provide and estimate of their pre-detector distributions
  + Advantages
    - Data analysis more possible for broader community
    - Enabling efficient combination of data
  + Problems
    - Widely used methods work on small dimensionality

**Proposal**

* Propose to use unfolding as an alternative for predicting parton level data

**ML-Unfolding Methods**

* Simulated samples come in pairs
* Two approaches
  + Reweight simulated samples
  + Generate unfolded samples from conditional probabilities

A diagram of a person's body

Description automatically generated with medium confidence

* Reweighting
* **Omnifold (Reweighting)**
  + Re-weighting based on Neyman-Pearson lemma
    - *Neyman-Pearson lemma*: optimally trained, calibrated classifier C, will learn the likelihood ratio of two underlying phase space distribution
  + Computes classifier weights at reco-level and uses paired simulated data to pull weights from reco-level to particle-level

A diagram of weights and weights

Description automatically generated

* + Bayesian network allows for learning distributions
* Mapping Distributions
  + Assumes that and describe same features at reco level
  + Train network to map event distributions form  to based on paired or unpaired simulated events and apply mapping to to generate

A diagram of a diagram

Description automatically generated

* **Schrodinger Bridge**
  + Uses forward-time stochastic differential equations (SDE) as a time-dependent process to define the transformation between to
    - * : deterministic part
      * : noise schedule
      * : noise infinitesimal
    - Need to find and
    - as the score
    - There is a lot of physics that makes solving this SDE easier
* **Direction Diffusion**
  + Describes time evolution between particle and reconstruction levels
  + Paired
    - Learn the velocity field () that transforms the density such that using the equation: for paired
    - Loss function:
  + Unpaired
    - Needed equation is:
    - Loss function:
  + Can add Bayesian layers, gaussian distributions, and KL-term for more distribution approach
* Generative Unfolding
  + Uses conditional generative networks to learn conditional probability describing the inverse simulation

A black text on a white background

Description automatically generated

* + Building forward surrogate network uses same data and has close to same setup going backwards due to Baye’s Theorem
* **Conditional INN**
  + Conditional invertible neural network
  + Creates a bijective mapping between the latent and phase space as an invertible function conditioned at the reco-level event
    - Learned density:
    - Loss function:
  + Transformer extension translate sequences of reco-level momenta into a sequence of particle-level momenta
* **Conditional Flow Matching**
  + Same as direction diffusion except CFM samples from a Gaussian latent distribution, conditional on a reco-level event
  + Advantage to Direct Diffusion is that this approach allows them to unfold the same reco-level event repeatedly with different noise from as a starting point
* **Transformer Conditional Flow Matching**
  + Steps
    - Reco-level and particle-level dimensions are individually mapped into a higher-dimensional embedding space
    - Reco-level embeddings are then fed to the transformer encoder
    - Updated embeddings are fed into a cross-attention block resolves the combinatorics between reco-level and particle-level objects
    - Outputs a final condition
  + For unfolding, sample from latent distribution and solve:
* **Latent Variational Diffusion** 
  + The goal is o reduce disparity between parameterizations of the set of observables to enable a more robust network
  + Map observables from particle/parton phase space to a latent space
  + Fixed length reco-level objects encoding mapping is learned by a deep feed-forward neural network
  + Variable-length inputs are used by a transformer encoder

**Key Findings**

**Notation**

: measured parton/particle level data

: parton level input

: reconstructed level data