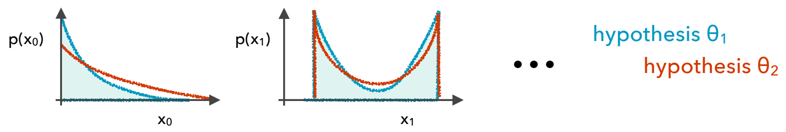
Expressive density models for high energy physics using a custom latent space

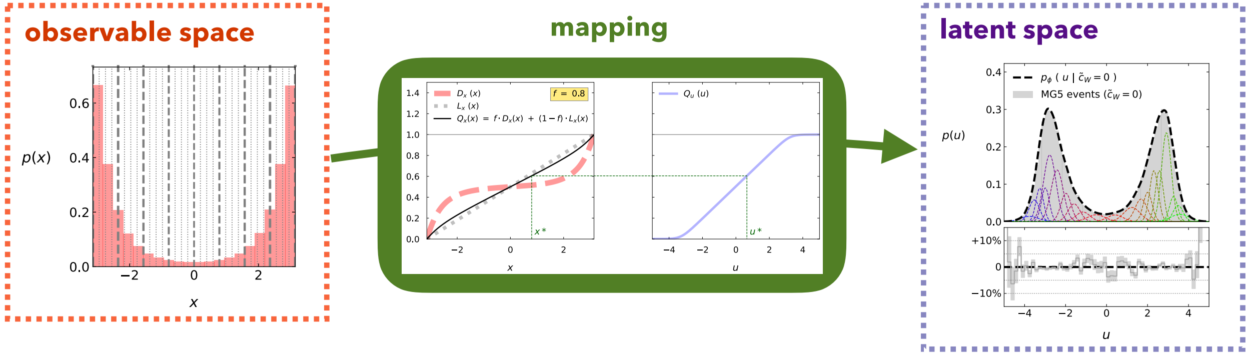
1. Motivation

- Particle physics datasets have **many correlated observables**
- Modified laws of physics expected to **deform these distributions**



- **Discovery potential** is maximised when we can model all observables $x = \{x_0, x_1, \dots\}$ simultaneously, and capture the dependence on parameters of nature θ
- Achieved with **neural density models** $p(x|\theta)$
- **This work:** present novel modelling method which captures two key features:
  - existence of abrupt physical boundaries (x_1 in above)
  - parameterised deformations of spectra
- Demonstrate on particle physics example
- Can be used to model data with these features in any domain

2. Explanation of method



- Data c.d.f. added with fraction **f** to linear function, creating a response curve $Q_x(x)$
- Response curve $Q_u(u)$ over a latent space smoothly approaches 0 and 1
- Mapping transforms data into a latent distribution (i) with **smooth edges** and (ii) well described by a **Gaussian mixture model with multiple narrow modes**
- Parameter variations deform spectra by modifying the amplitudes, means and widths of the Gaussian modes local to the deformation, leading to a **highly expressive model**

3. Experimental setup

- Auto-regressive: latent observables modelled as conditional chain

$$p(u|\theta) = p(u_0|\theta) \times p(u_1|u_0,\theta) \times p(u_2|u_1,u_0,\theta) \times \dots$$

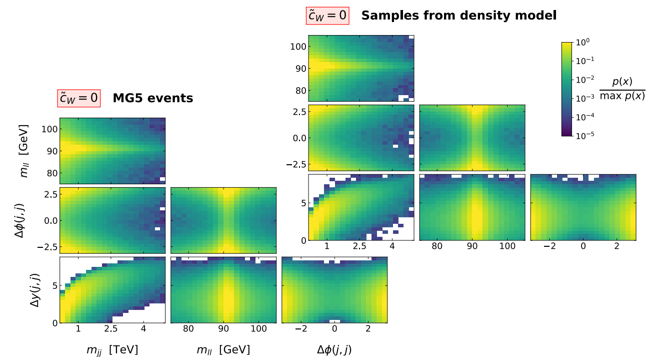
- Gaussian amplitudes/means/widths modelled using a neural net which learns conditional dependencies

$$p(u_n|u_{<n},\theta) = \sum_g f_g(u_{<n},\theta) \cdot \mathcal{N}(u_n;\mu_g(u_{<n},\theta), \log\sigma_g(u_{<n},\theta))$$

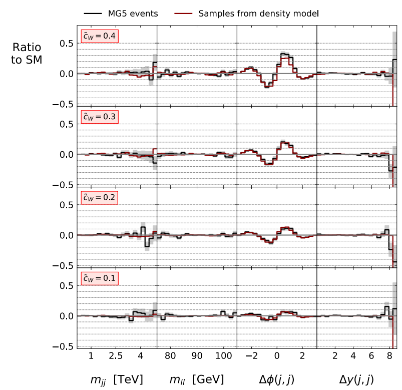
- Particle physics example: electroweak production of Z + 2 jets at LHC; four observables $\{m_{jj}, m_{ll}, \Delta\phi(j,j), \Delta y(j,j)\}$ and one parameter of interest $\{\tilde{c}_W\}$ from SMEFT describing modified physics

4. Results

- Test by sampling from the model and comparing with the ground truth



2D marginals show correlations are well modelled



Ratios show parameter dependence is well modelled

S. Menary (stephen.menary@manchester.ac.uk) and D. Price (darren.price@manchester.ac.uk), University of Manchester, ML and the Physical Sciences Workshop, NeurIPS, 11th Dec 2020