Turbo-Sim: a generalised generative model with a physical latent space

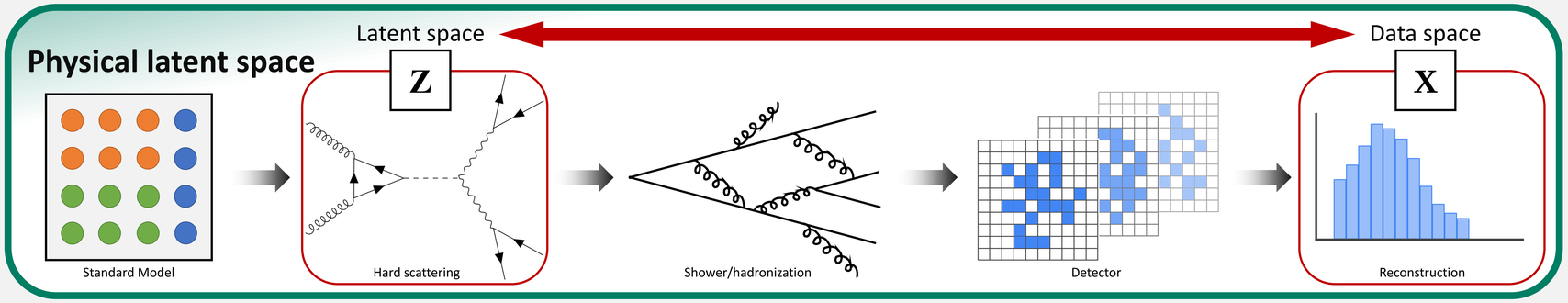
Fast simulation and clustering

Introduction

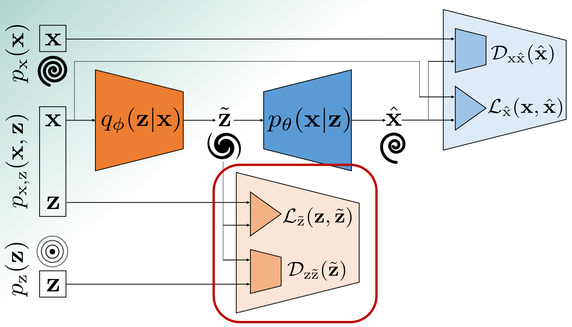
Turbo-Sim is a generalised framework which gives an interpretation derived from principles of information theory to several deep learning models.

- It maximises mutual information between data and latent spaces.
- It can be used as a generative model for particle physics.

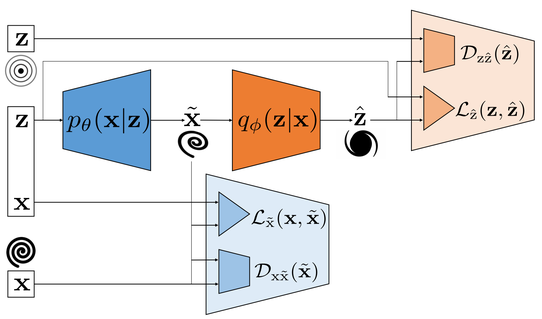
Physical latent space



Direct **direction**



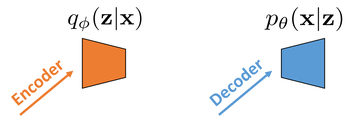
Reverse **direction**



Maximises the mutual information between $\mathbf{x}$ and $\tilde{\mathbf{z}}$

Theory

- The joint probability density is approximated thanks to two neural networks (mappers).



$$p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z}) = p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})p_{\mathrm{x}}(\mathbf{x}) = p_{\mathrm{x|z}}(\mathbf{x}|\mathbf{z})p_{\mathrm{z}}(\mathbf{z})$$

- Different decompositions give several lower bounds on the mutual information.

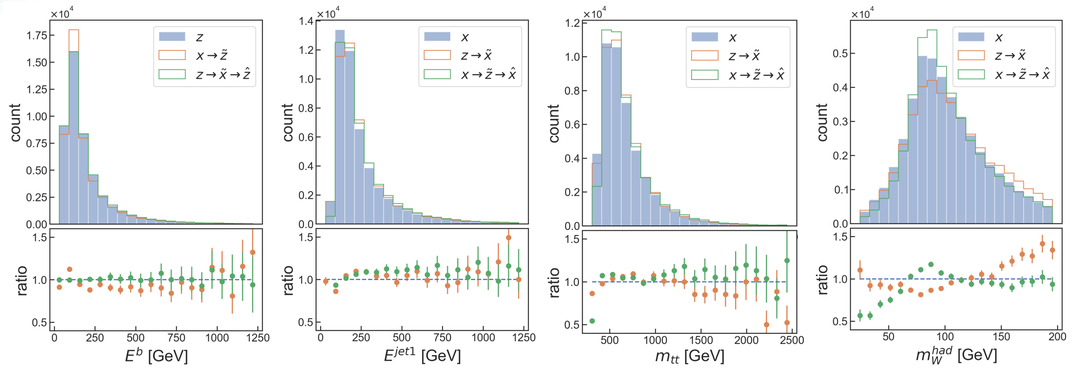
$$\mathcal{I}^{\mathrm{z}}_{\phi}(\mathbf{X};\mathbf{Z}) = \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p_{\mathrm{z}}(\mathbf{z})}\frac{\tilde{q}_\phi(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] \geq \underbrace{\mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log q_\phi(\mathbf{z}|\mathbf{x})\right]}_{-\mathcal{L}_{\tilde{z}}(\mathbf{z},\tilde{\mathbf{z}})} - \underbrace{D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z}))}_{\mathcal{D}_{z\tilde{z}}(\tilde{\mathbf{z}})}$$

$$\mathcal{I}^{\mathrm{x}}_{\phi,\theta}(\mathbf{Z};\mathbf{X}) = \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{p_\theta(\mathbf{x}|\mathbf{z})}{p_{\mathrm{x}}(\mathbf{x})}\frac{\hat{p}_\theta(\mathbf{x})}{\hat{p}_\theta(\mathbf{x})}\right] \geq \underbrace{\mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right]}_{-\mathcal{L}_{\hat{x}}(\mathbf{x},\hat{\mathbf{x}})} - \underbrace{D_{\mathrm{KL}}(p_{\mathrm{x}}(\mathbf{x})\|\hat{p}_\theta(\mathbf{x}))}_{\mathcal{D}_{x\hat{x}}(\hat{\mathbf{x}})}$$

- Many common adversarial auto-encoders (AAEs) and generative adversarial networks (GANs) deep learning models can be expressed using subparts of the Turbo-Sim framework.

Results

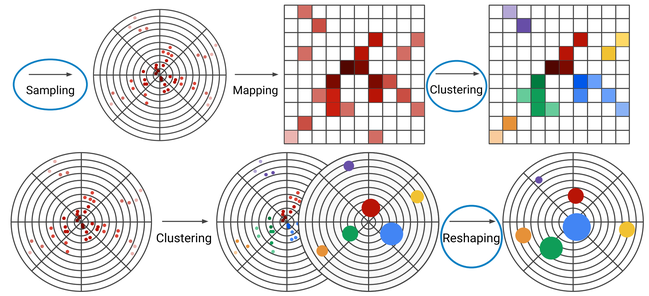
We focus on double top quarks production in proton–proton collision subsequently decaying into b-quarks and W-bosons which in turn decay into a pair of light leptons and a pair of light quarks: $pp \to t\bar{t} \to e^- \bar{\nu}_e\, b\bar{b}\, u\bar{d}$



Fast simulation and clustering

ATLAS data analysis needs a huge amount of simulated data, which is produced via very complex and slow Monte Carlo tools. Alternatives exist, but always with a trade-off between speed and accuracy.

- We can use machine learning to sample from precomputed average distributions of energy deposits.
- We can use machine learning to reshape the clusters of energy deposits.
- We develop fast clustering algorithms to apply and to validate these techniques.



Physics Applications of AI Day
03.11.2021

Contact: guillaume.quetant@unige.ch

Guillaume Quétant, Mariia Drozdova, Vitaliy Kinakh, Johnny Raine, Tobias Golling, Slava Voloshynovskiy

UNIVERSITÉ DE GENÈVE