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

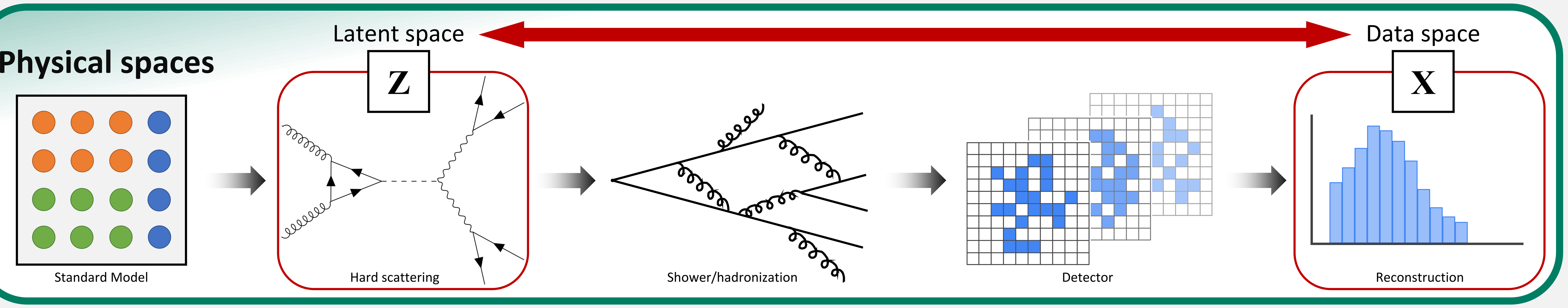
UNIVERSITÉ DE GENÈVE

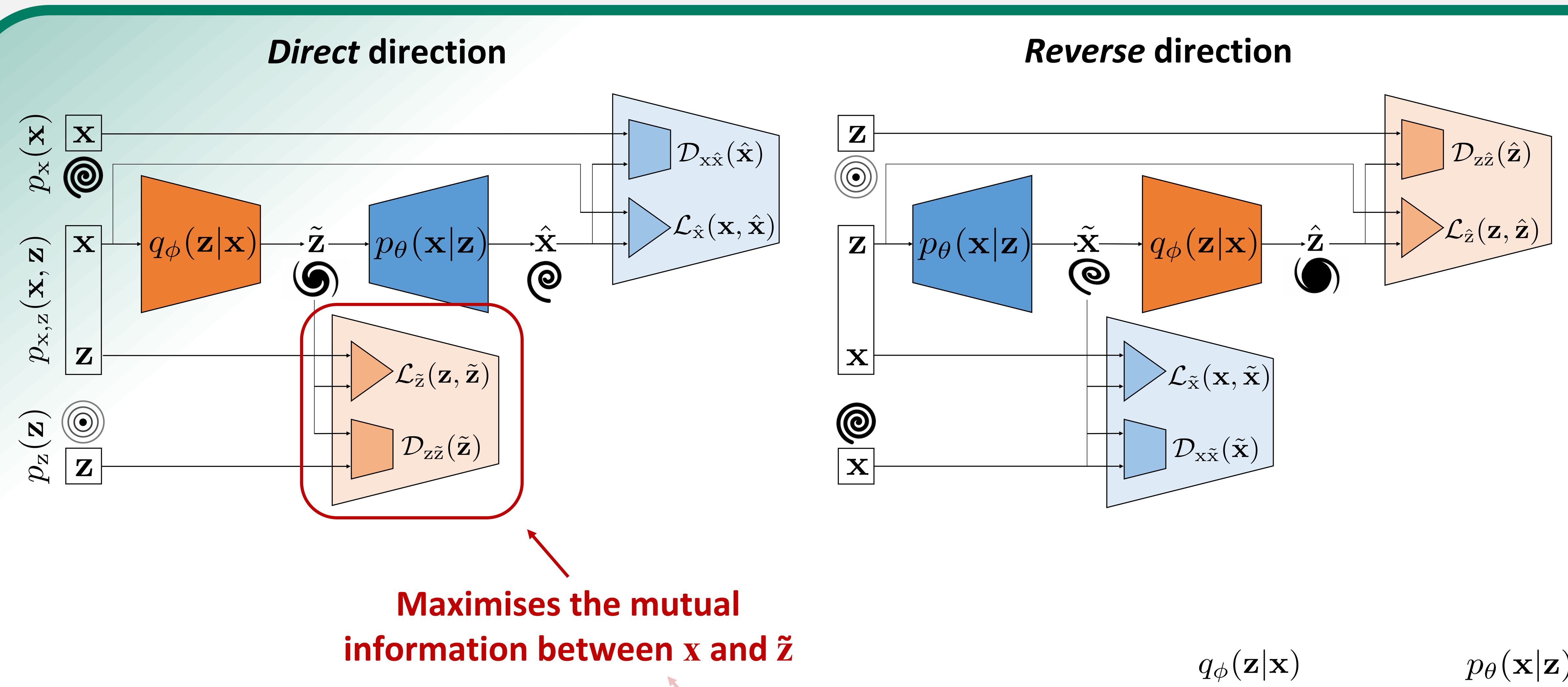
Introduction

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

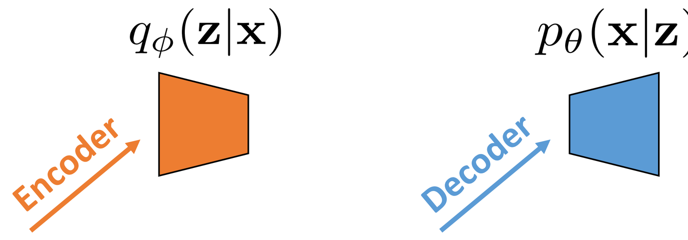
- It maximises mutual information between data and latent spaces.
- It can be used as a generative model for particle physics.

Physical spaces





Theory



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

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

- Several lower bounds to the true mutual information are in turn lower bounded.

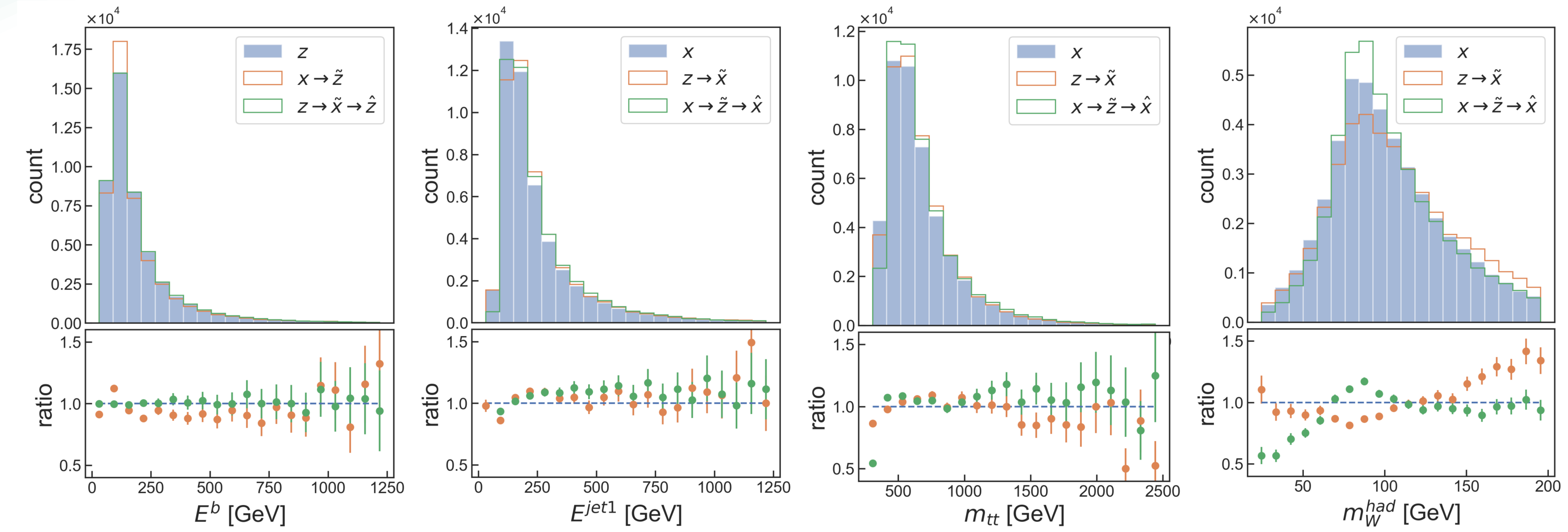
$$I(\mathbf{X};\mathbf{Z}) \geq \mathcal{I}^{\mathrm{z}}_{\phi}(\mathbf{X};\mathbf{Z}) = \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p_{\mathrm{z}}(\mathbf{z})}\frac{\tilde{q}_\phi(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] \geq \boxed{\underbrace{\mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{q_\p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

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

- 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: $pp \to t\bar{t} \to e^- \bar{\nu}_e b\bar{b} u\bar{d}$



Comparison

We compare our results with another method applied to the same data thanks to the Kolmogorov-Smirnov distance metric.

	Z space			**X** space			Reconstructed physics			
Model	p_y^b	p_z^b	E^b	p_y^{jet1}	p_z^{jet1}	E^{jet1}	m_{tt}	m_W^{had}	m_t^{lep}	m_t^{had}
Turbo-Sim	5.28	7.28	3.96	**2.89**	10.3	**4.43**	**2.97**	**7.72**	**5.20**	**8.52**
OTUS	**1.59**	**1.23**	**2.76**	3.78	**2.39**	5.75	15.8	11.7	14.1	24.9

Turbo-Sim is better at:

- Sampling the data space i.e. being a generative model
- Reconstructing underlying objects i.e. learning the physics

References

- Slava Voloshynovskiy et al. *Information bottleneck through variational glasses*. 2019. arXiv: 1912.00830
- Jessica N. Howard et al. *Foundations of a Fast, Data-Driven, Machine-Learned Simulator*. 2021. arXiv: 2101.08944

Fourth Workshop on Machine Learning and the Physical Sciences

NeurIPS 2021

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

Contact: guillaume.quetant@unige.ch

NEURAL INFORMATION PROCESSING SYSTEMS