Turbo-Sim

a generalised generative model with a physical latent space

5th Inter-experiment Machine Learning workshop - 12.05.2022

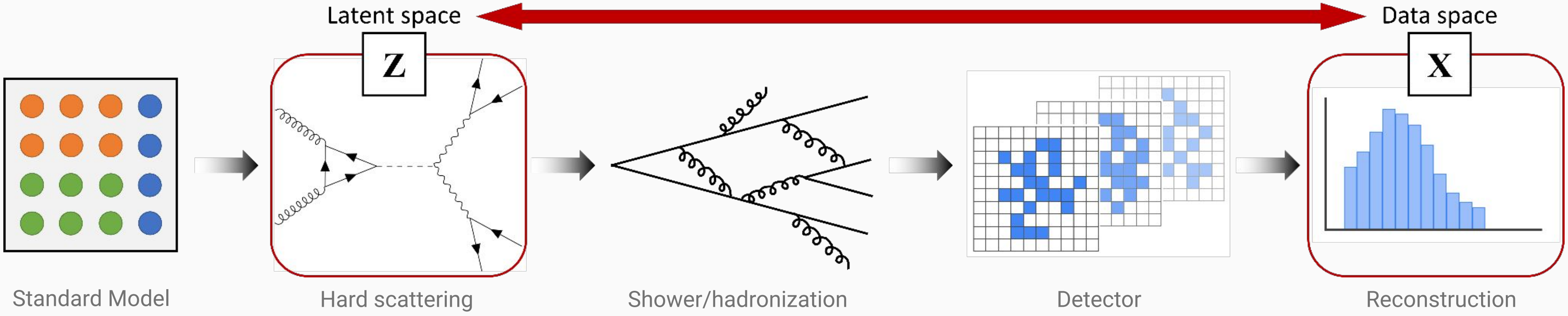
Guillaume Quétant, Mariia Drozdova, Vitaliy Kinakh, Tobias Golling, Slava Voloshynovskiy
arXiv

UNIVERSITÉ DE GENÈVE

Introduction

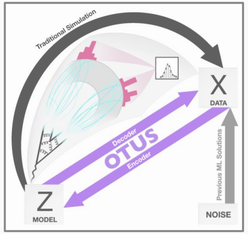
Physical motivation

Transform from theory to observation



- ML network to learn the transformation
- Learn both directions
- Everything is stochastic

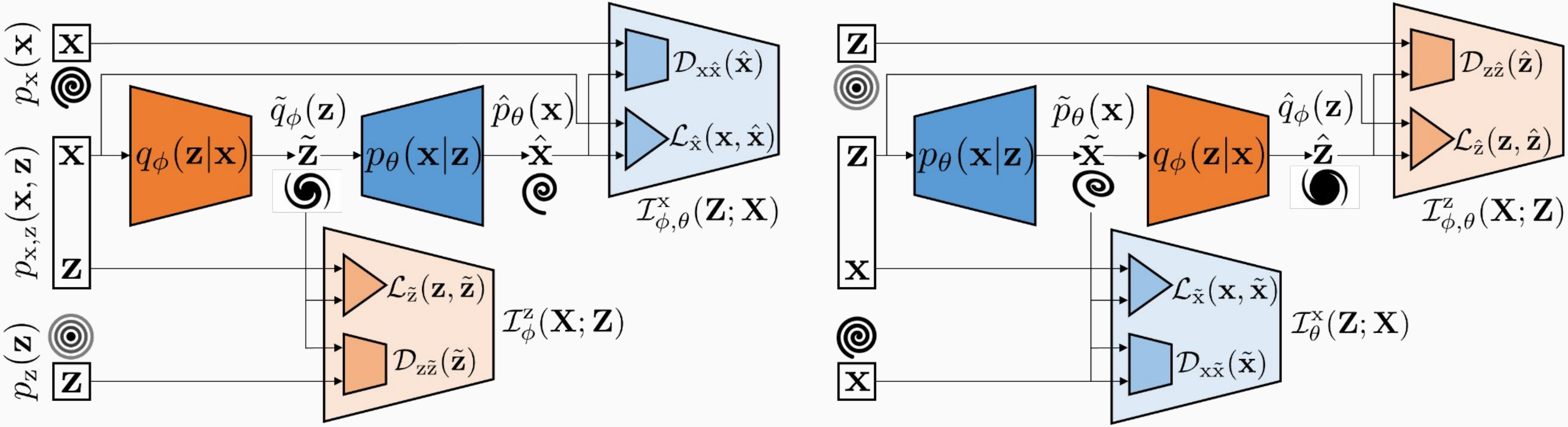
Jessica N. Howard et al.
Foundations of a Fast, Data-Driven, Machine-Learned Simulator. 2021.
arXiv: 2101.08944



Turbo-Sim

Walkthrough the model

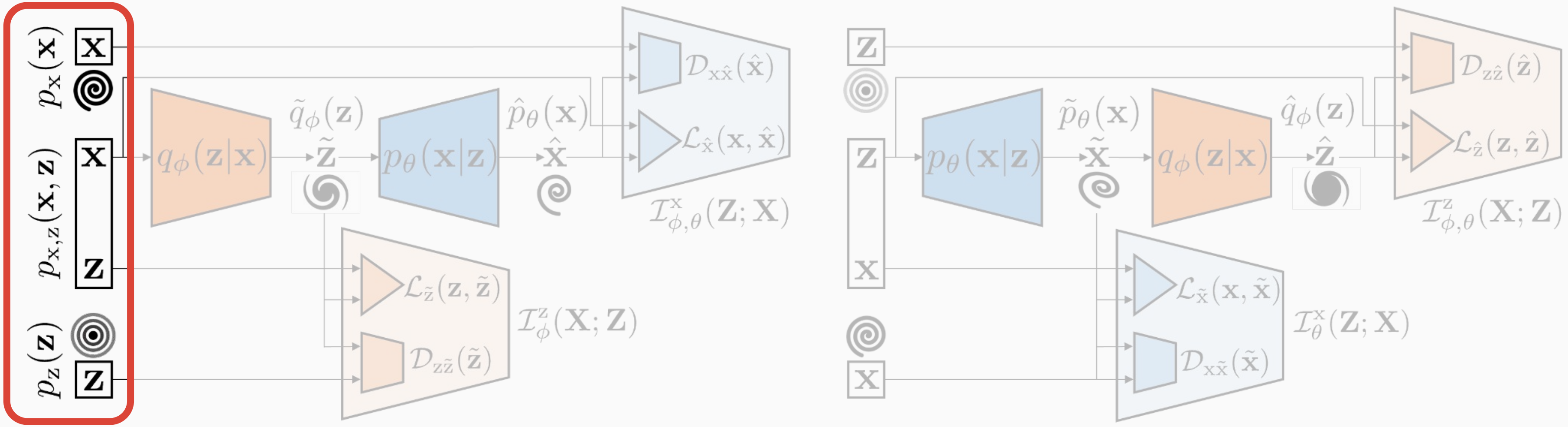
Turbo-Sim: complete framework



Turbo-Sim is a generalised autoencoder: encoder, decoder, eight loss terms and a joint data distribution

→ Let us dig into it

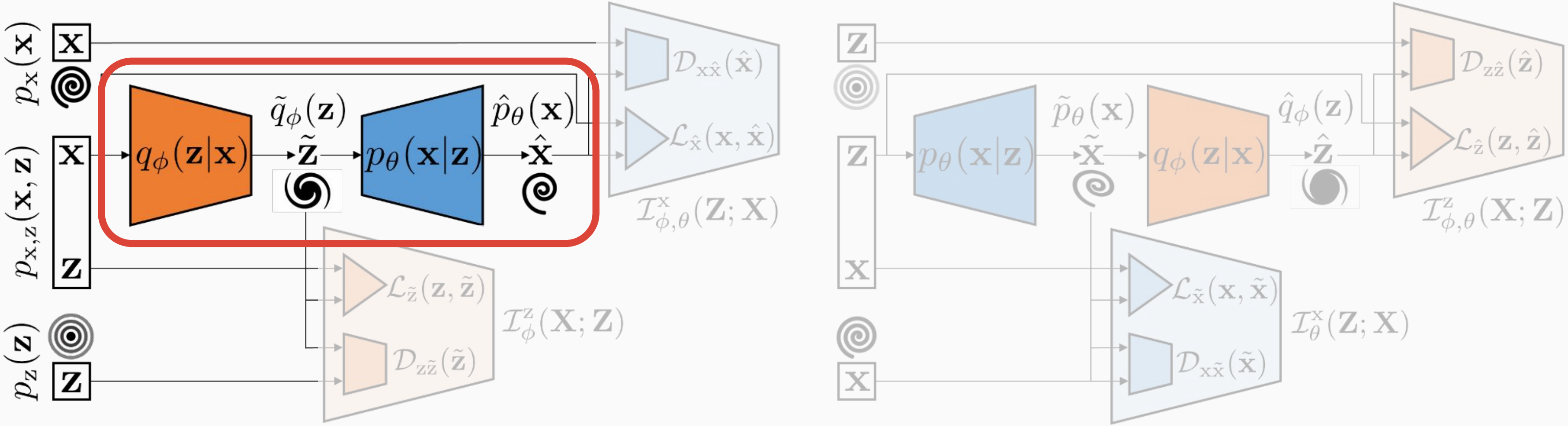
Turbo-Sim: complete framework



- Data and latent points are sampled jointly *(they may be independent)*
- Marginal distributions are available as well

→ In our HEP experiment: Z = hard scattering space / X = reconstruction space

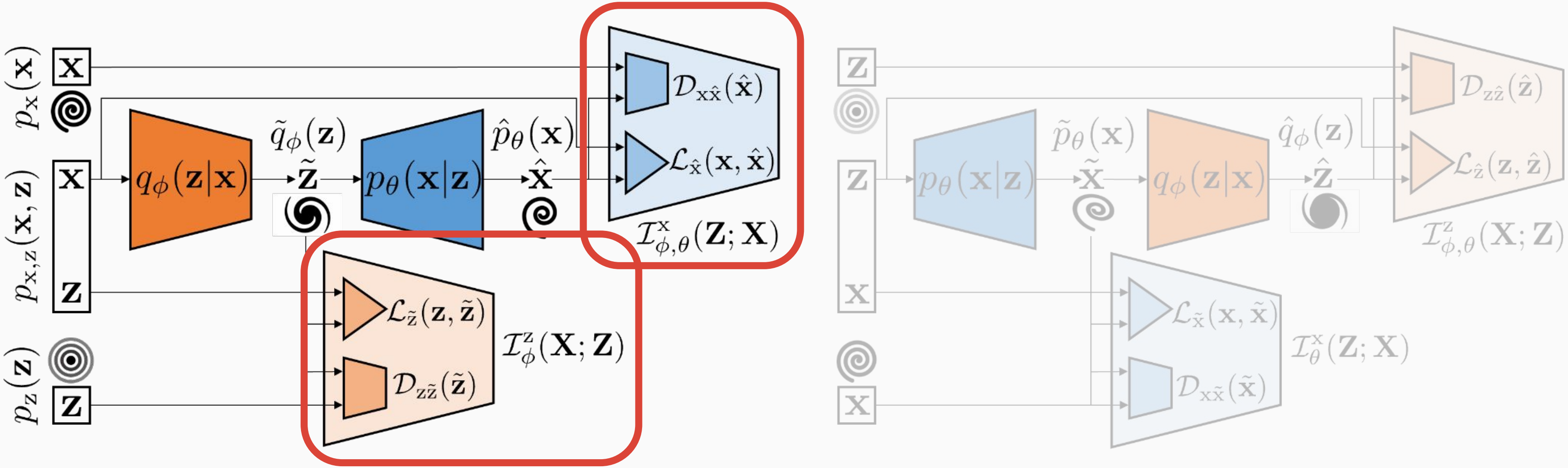
Turbo-Sim: complete framework



- Autoencoder outputs a latent space and a reconstructed space distributions
- Encoder and decoder should be (stochastic) inverse of each other

→ In our HEP experiment: encoder = unfolding / decoder = simulation

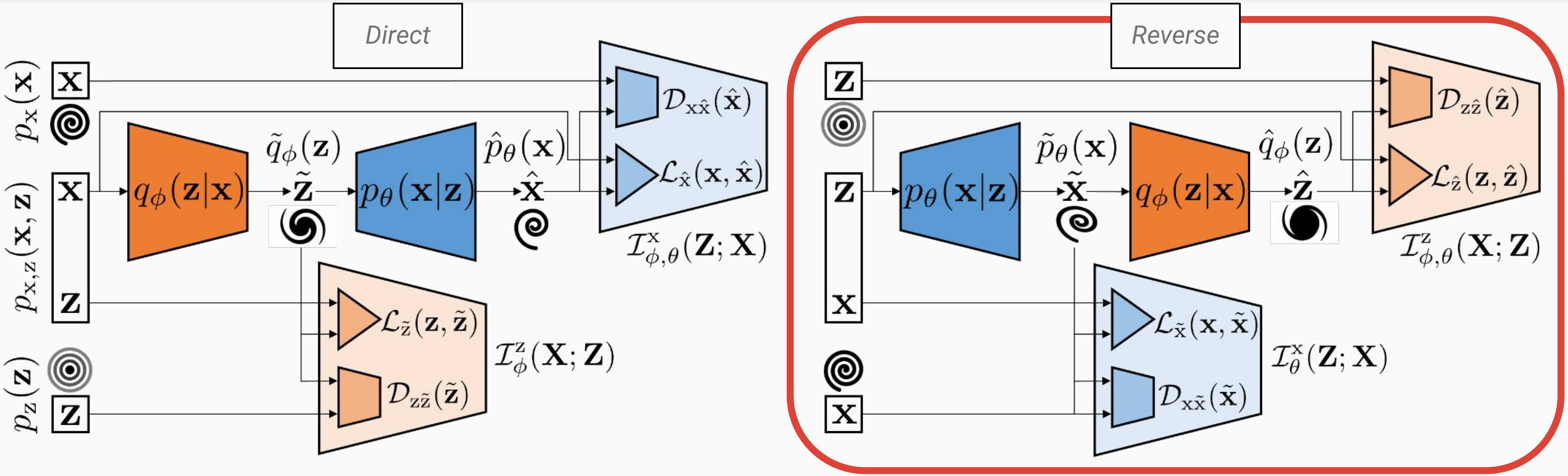
Turbo-Sim: complete framework



- Supervised and unsupervised losses in both latent and reconstructed spaces
- Information theory interpretation of several usual models (AAE, GAN)

→ In our HEP experiment: {z, x} pairs for supervised / {z}, {x} marginals for unsupervised

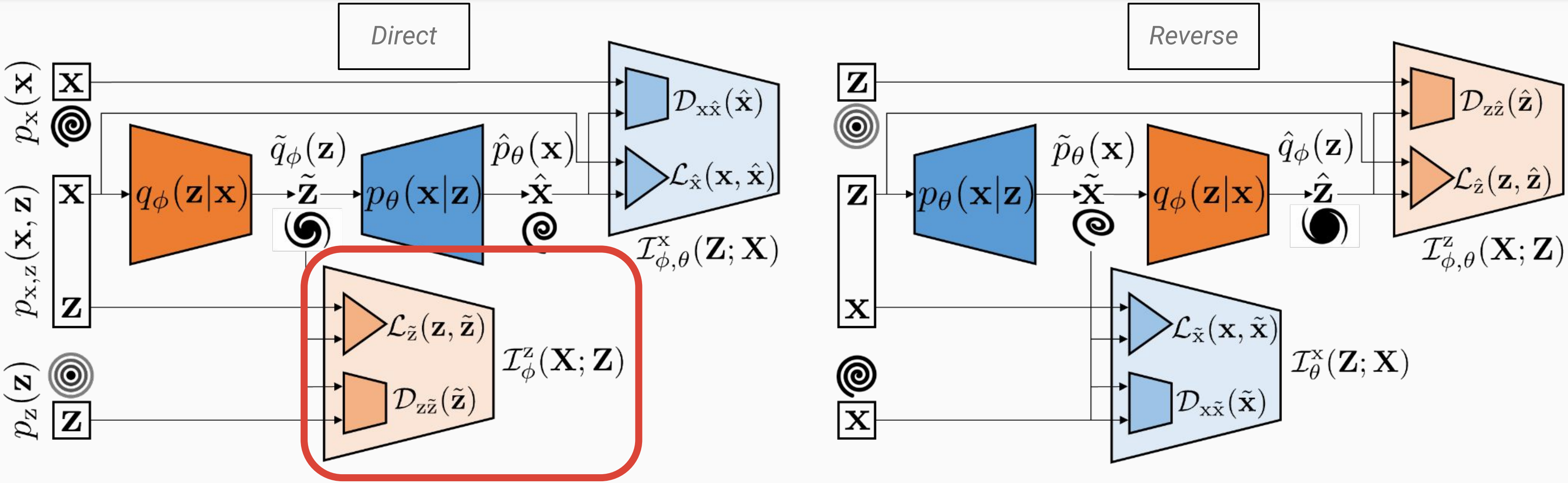
Turbo-Sim: complete framework



- Everything is symmetrical with Z ↔ X
- Four more loss terms and two ways of training *(direct and reverse)*

→ In our HEP experiment: should helps the network to be good in unfolding and simulation

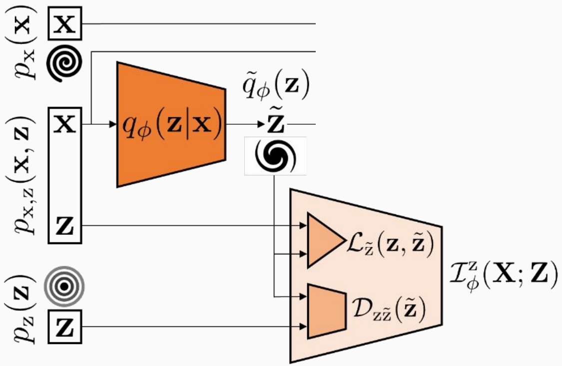
Turbo-Sim: complete framework



→ Let us focus on the derivation of these two terms

Lower bound to the true mutual information

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

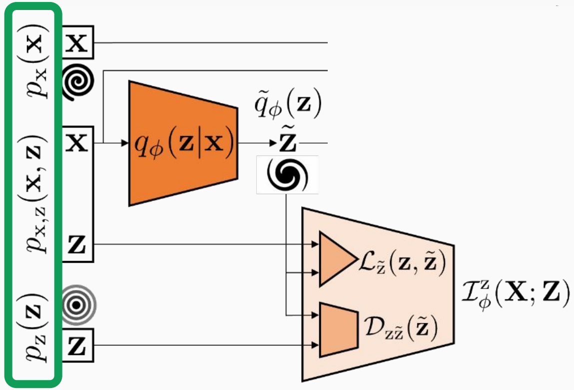


Equality reached for

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_\phi(\mathbf{z}|\mathbf{x})$$

Lower bound to the true mutual information

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

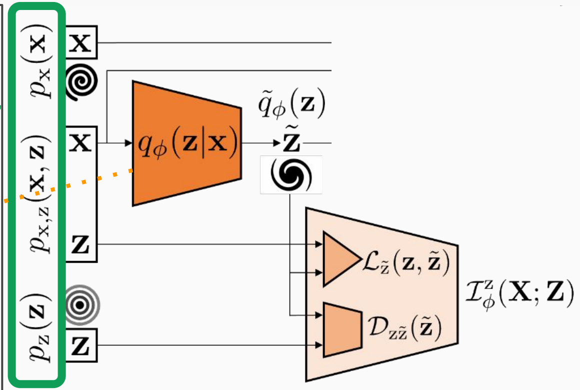


Equality reached for

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_\phi(\mathbf{z}|\mathbf{x})$$

Lower bound to the true mutual information

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

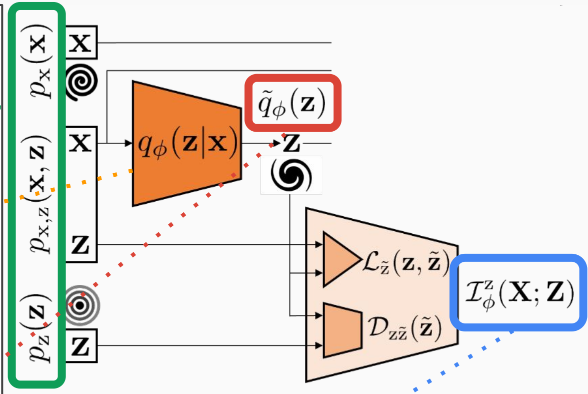


Equality reached for

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_\phi(\mathbf{z}|\mathbf{x})$$

Lower bound to the true mutual information

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

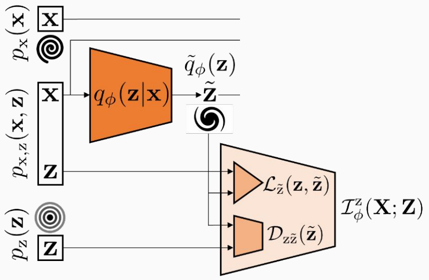


Equality reached for

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_{\phi}(\mathbf{z}|\mathbf{x})$$

Lower bound to the approximate mutual information

$$
\begin{aligned}
\mathcal{I}_\phi^{\mathrm{z}}(\mathbf{X};\mathbf{Z}) &= \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p_{\mathrm{z}}(\mathbf{z})}\frac{\tilde{q}_\phi(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] \\
&= \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{p_{\mathrm{z}}(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] - \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}[\log \tilde{q}_\phi(\mathbf{z})] \\
&= \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - \mathbb{E}_{p_{\mathrm{z}}(\mathbf{z})}\left[\log \frac{p_{\mathrm{z}}(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] - \mathbb{E}_{p_{\mathrm{z}}(\mathbf{z})}[\log \tilde{q}_\phi(\mathbf{z})] \\
&= \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z})) + H(p_{\mathrm{z}}(\mathbf{z});\tilde{q}_\phi(\mathbf{z})) \\
&\geq \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z})) \\
&\simeq \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z})),
\end{aligned}
$$

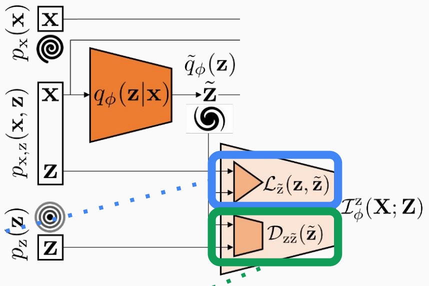


Equality *not* reached for (but the maximum value is)

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_\phi(\mathbf{z}|\mathbf{x})$$

Lower bound to the approximate mutual information

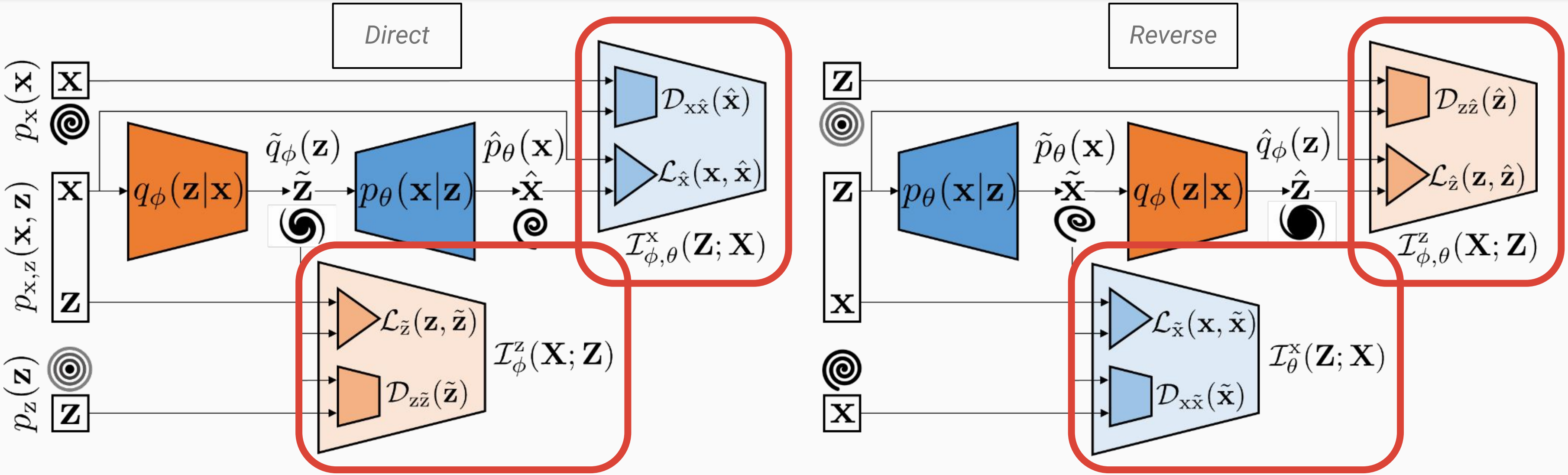
$$
\begin{aligned}
\mathcal{I}^{\mathrm{z}}_{\phi}(\mathbf{X};\mathbf{Z}) &= \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p_{\mathrm{z}}(\mathbf{z})}\frac{\tilde{q}_\phi(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] \\
&= \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}\left[\log \frac{p_{\mathrm{z}}(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] - \mathbb{E}_{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}[\log \tilde{q}_\phi(\mathbf{z})] \\
&= \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - \mathbb{E}_{p_{\mathrm{z}}(\mathbf{z})}\left[\log \frac{p_{\mathrm{z}}(\mathbf{z})}{\tilde{q}_\phi(\mathbf{z})}\right] - \mathbb{E}_{p_{\mathrm{z}}(\mathbf{z})}[\log \tilde{q}_\phi(\mathbf{z})] \\
&= \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z})) + H(p_{\mathrm{z}}(\mathbf{z});\tilde{q}_\phi(\mathbf{z})) \\
&\geq \mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{p_{\mathrm{z|x}}(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})] - D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z})) \\
&\simeq \boxed{\mathbb{E}_{p_{\mathrm{x}}(\mathbf{x})}\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log q_\phi(\mathbf{z}|\mathbf{x})]}\ \boxed{- D_{\mathrm{KL}}(p_{\mathrm{z}}(\mathbf{z})\|\tilde{q}_\phi(\mathbf{z}))},
\end{aligned}
$$



Equality *not* reached for (but the maximum value is)

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = q_\phi(\mathbf{z}|\mathbf{x})$$

Turbo-Sim: complete framework



→ Similar decompositions lead to the six other terms

Turbo-Sim: losses

Turbo-Sim vs Bounded Information Bottleneck AE:

$$\bar{\mathcal{L}}^{\mathrm{Direct}}(\phi, \theta) = -\mathcal{I}^{\mathrm{z}}_{\phi}(\mathbf{X};\mathbf{Z}) - \alpha\mathcal{I}^{\mathrm{x}}_{\phi,\theta}(\mathbf{Z};\mathbf{X})$$

$$\mathcal{L}_{\mathrm{BIB-AE}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = I_{\boldsymbol{\phi}}(\mathbf{X};\mathbf{Z}) - \beta I^{\mathrm{U}_L}_{\boldsymbol{\theta},\boldsymbol{\phi}}(\mathbf{Z};\mathbf{X})$$

(1912.00830)

Mutual information between input and latent space

- Turbo-Sim: maximised *(loss is minimised)*
- BIB-AE: minimised

Turbo-Sim: losses

Turbo-Sim vs Bounded Information Bottleneck AE:

$$\bar{\mathcal{L}}^{\text{Direct}}(\phi, \theta) = -\mathcal{I}^{\text{z}}_{\phi}(\mathbf{X}; \mathbf{Z}) - \alpha \mathcal{I}^{\text{x}}_{\phi,\theta}(\mathbf{Z}; \mathbf{X})$$

$$\mathcal{L}_{\text{BIB-AE}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = I_{\boldsymbol{\phi}}(\mathbf{X}; \mathbf{Z}) - \beta I^{\text{U}_L}_{\boldsymbol{\theta},\boldsymbol{\phi}}(\mathbf{Z}; \mathbf{X})$$

(1912.00830)

Mutual information between input and latent space

- Turbo-Sim: maximised *(loss is minimised)*
- BIB-AE: minimised

Turbo-Sim: losses

Direct and reverse losses:

$$\bar{\mathcal{L}}^{\text{Direct}}(\phi, \theta) = \mathcal{L}_{\tilde{\text{z}}}(\mathbf{z}, \tilde{\mathbf{z}}) + \mathcal{D}_{\text{z}\tilde{\text{z}}}(\tilde{\mathbf{z}}) + \alpha\, \mathcal{L}_{\hat{\text{x}}}(\mathbf{x}, \hat{\mathbf{x}}) + \alpha\, \mathcal{D}_{\text{x}\hat{\text{x}}}(\hat{\mathbf{x}})$$

$$\bar{\mathcal{L}}^{\text{Reverse}}(\phi, \theta) = \mathcal{L}_{\tilde{\text{x}}}(\mathbf{x}, \tilde{\mathbf{x}}) + \mathcal{D}_{\text{x}\tilde{\text{x}}}(\tilde{\mathbf{x}}) + \beta\, \mathcal{L}_{\hat{\text{z}}}(\mathbf{z}, \hat{\mathbf{z}}) + \beta\, \mathcal{D}_{\text{z}\hat{\text{z}}}(\hat{\mathbf{z}})$$

Total loss:

$$\bar{\mathcal{L}}^{\text{Turbo}}(\phi, \theta) = \bar{\mathcal{L}}^{\text{Direct}}(\phi, \theta) + \gamma\, \bar{\mathcal{L}}^{\text{Reverse}}(\phi, \theta)$$

Turbo-Sim: losses

Direct and reverse losses:

Adversarial AutoEncoder (AAE)

$$\bar{\mathcal{L}}^{\text{Direct}}(\phi, \theta) = \mathcal{L}_{\tilde{\mathrm{z}}}(\mathbf{z}, \tilde{\mathbf{z}}) + \boxed{\mathcal{D}_{\mathrm{z}\tilde{\mathrm{z}}}(\tilde{\mathbf{z}})} + \boxed{\alpha\, \mathcal{L}_{\hat{\mathrm{x}}}(\mathbf{x}, \hat{\mathbf{x}})} + \alpha\, \mathcal{D}_{\mathrm{x}\hat{\mathrm{x}}}(\hat{\mathbf{x}})$$

$$\bar{\mathcal{L}}^{\text{Reverse}}(\phi, \theta) = \mathcal{L}_{\tilde{\mathrm{x}}}(\mathbf{x}, \tilde{\mathbf{x}}) + \mathcal{D}_{\mathrm{x}\tilde{\mathrm{x}}}(\tilde{\mathbf{x}}) + \beta\, \mathcal{L}_{\hat{\mathrm{z}}}(\mathbf{z}, \hat{\mathbf{z}}) + \beta\, \mathcal{D}_{\mathrm{z}\hat{\mathrm{z}}}(\hat{\mathbf{z}})$$

Total loss:

$$\bar{\mathcal{L}}^{\text{Turbo}}(\phi, \theta) = \bar{\mathcal{L}}^{\text{Direct}}(\phi, \theta) + \gamma\, \bar{\mathcal{L}}^{\text{Reverse}}(\phi, \theta)$$

Turbo-Sim: losses

Direct and reverse losses:

Adversarial AutoEncoder (AAE)

$$\bar{\mathcal{L}}^{\text{Direct}}(\phi,\theta) = \mathcal{L}_{\tilde{\mathrm{z}}}(\mathbf{z},\tilde{\mathbf{z}}) + \boxed{\mathcal{D}_{\mathrm{z}\tilde{\mathrm{z}}}(\tilde{\mathbf{z}})} + \boxed{\alpha\,\mathcal{L}_{\hat{\mathrm{x}}}(\mathbf{x},\hat{\mathbf{x}})} + \alpha\,\mathcal{D}_{\mathrm{x}\hat{\mathrm{x}}}(\hat{\mathbf{x}})$$

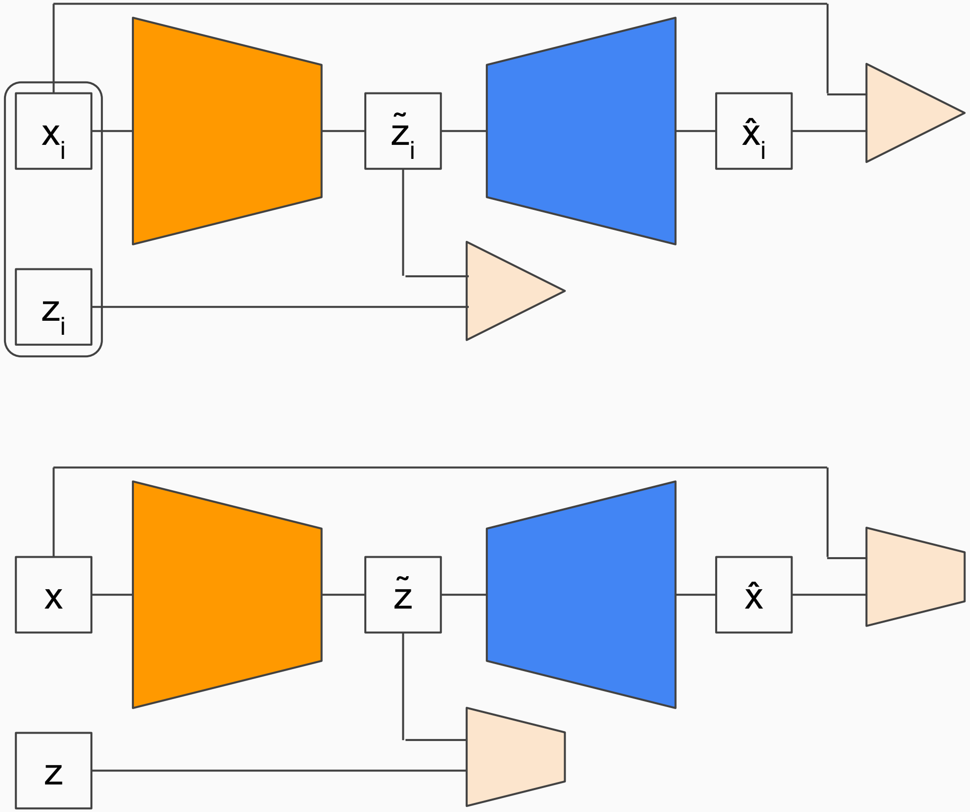
$$\bar{\mathcal{L}}^{\text{Reverse}}(\phi,\theta) = \mathcal{L}_{\tilde{\mathrm{x}}}(\mathbf{x},\tilde{\mathbf{x}}) + \boxed{\mathcal{D}_{\mathrm{x}\tilde{\mathrm{x}}}(\tilde{\mathbf{x}})} + \beta\,\mathcal{L}_{\hat{\mathrm{z}}}(\mathbf{z},\hat{\mathbf{z}}) + \beta\,\mathcal{D}_{\mathrm{z}\hat{\mathrm{z}}}(\hat{\mathbf{z}})$$

Generative Adversarial Network (GAN)

Total loss:

$$\bar{\mathcal{L}}^{\text{Turbo}}(\phi,\theta) = \bar{\mathcal{L}}^{\text{Direct}}(\phi,\theta) + \gamma\,\bar{\mathcal{L}}^{\text{Reverse}}(\phi,\theta)$$

Training scheme



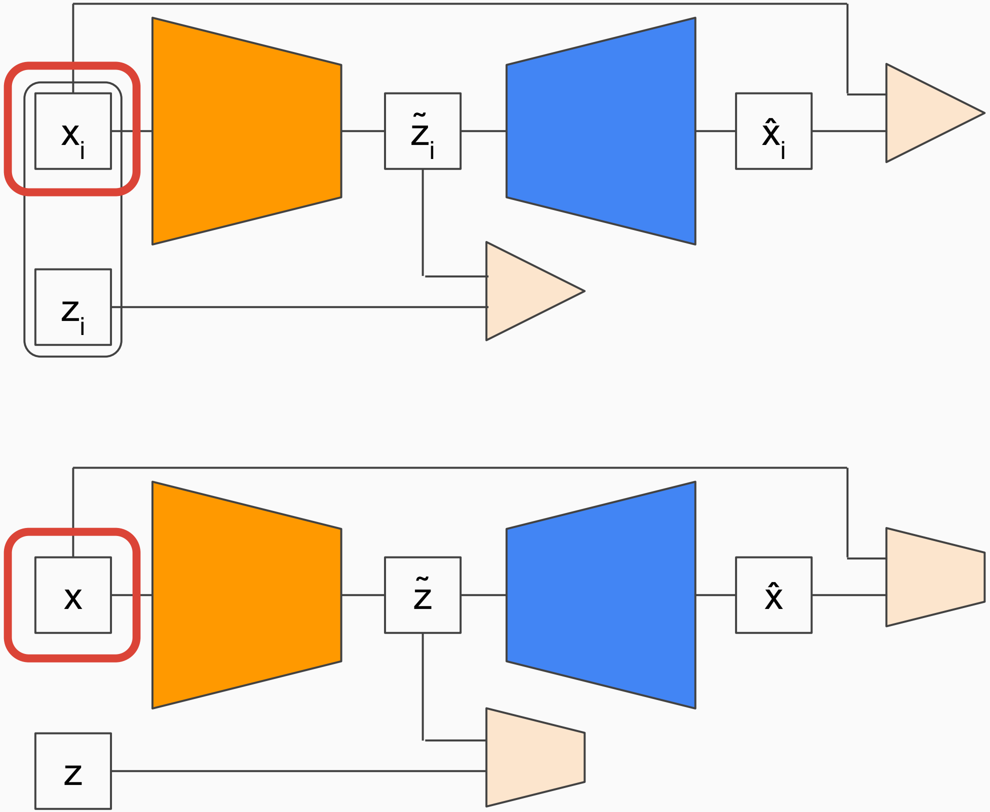
Direct

- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

Reverse:

- The same is done swapping $x \leftrightarrow z$

Training scheme



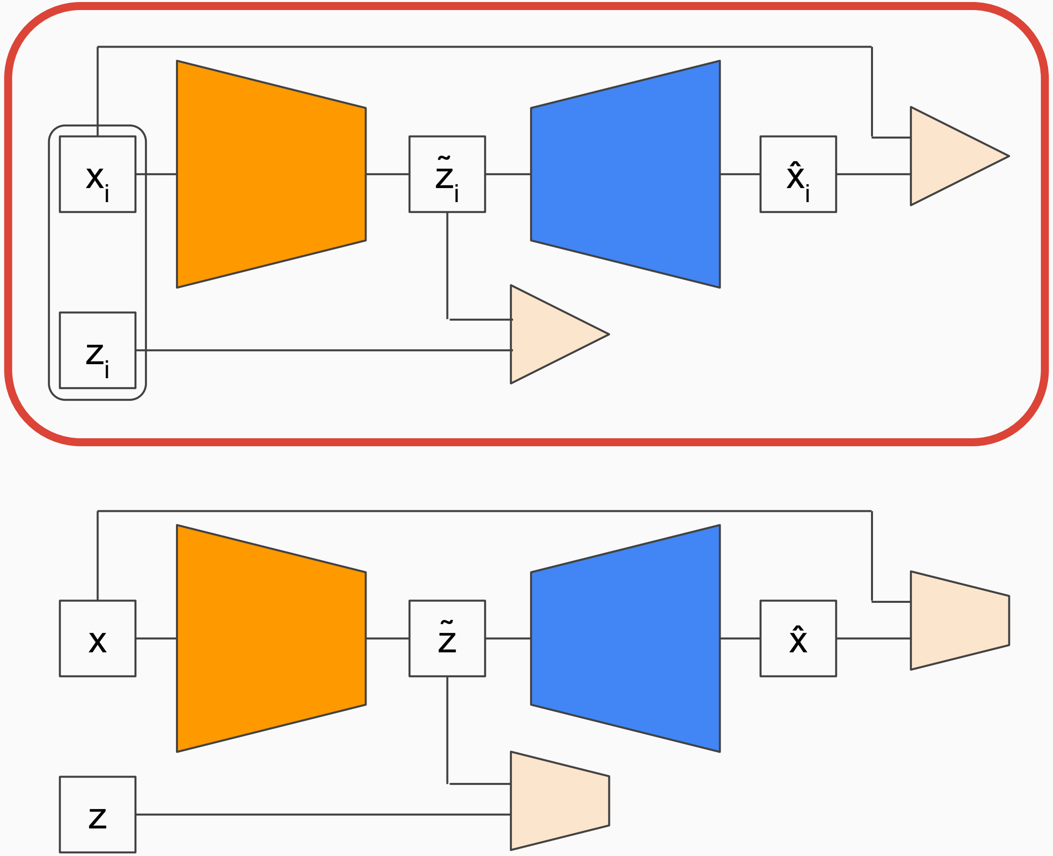
Direct

- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

Reverse:

- The same is done swapping $x \leftrightarrow z$

Training scheme



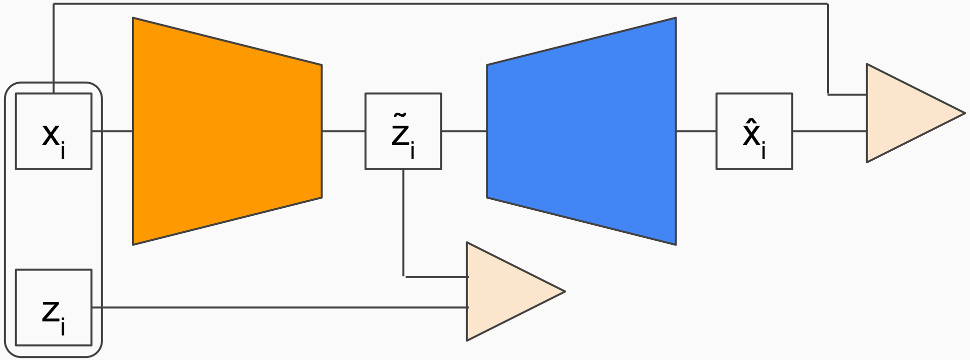
Direct

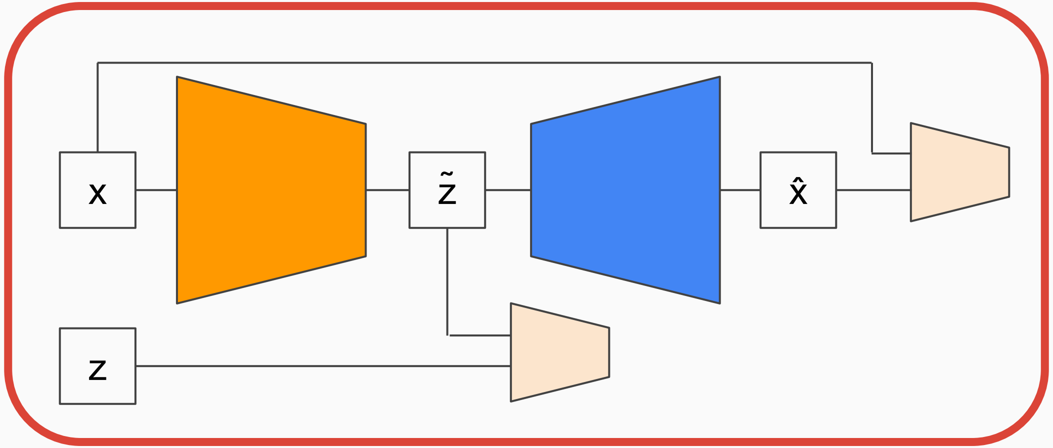
- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

Reverse:

- The same is done swapping x ↔ z

Training scheme





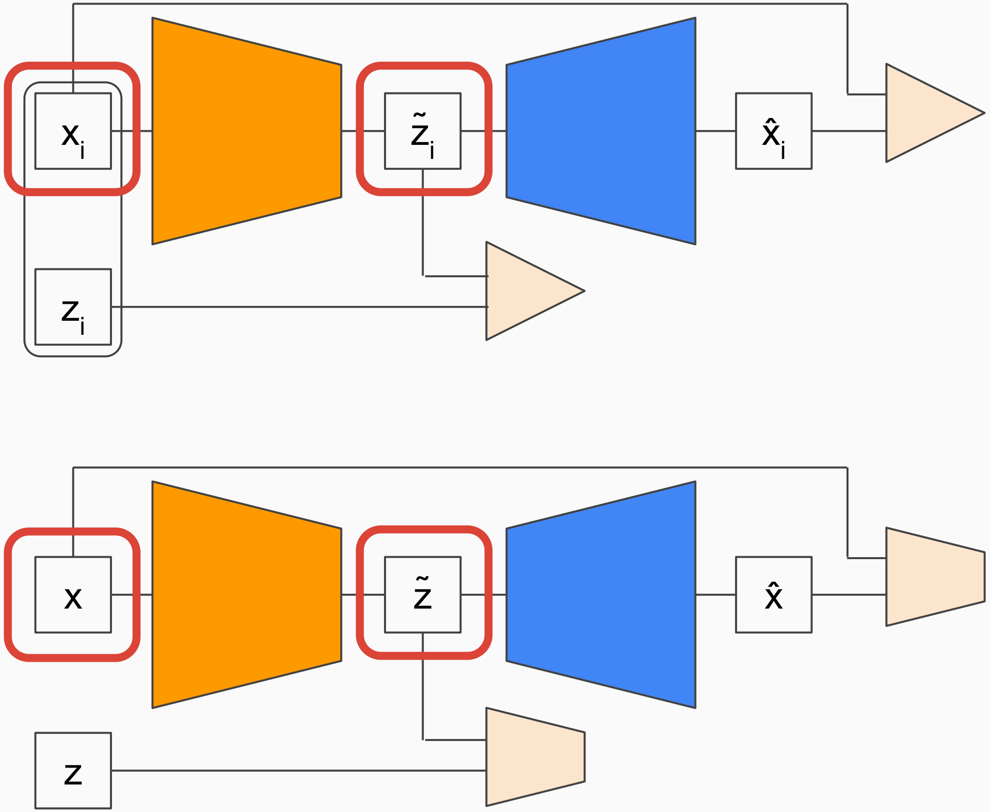
Direct

- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

Reverse:

- The same is done swapping $x \leftrightarrow z$

Training scheme



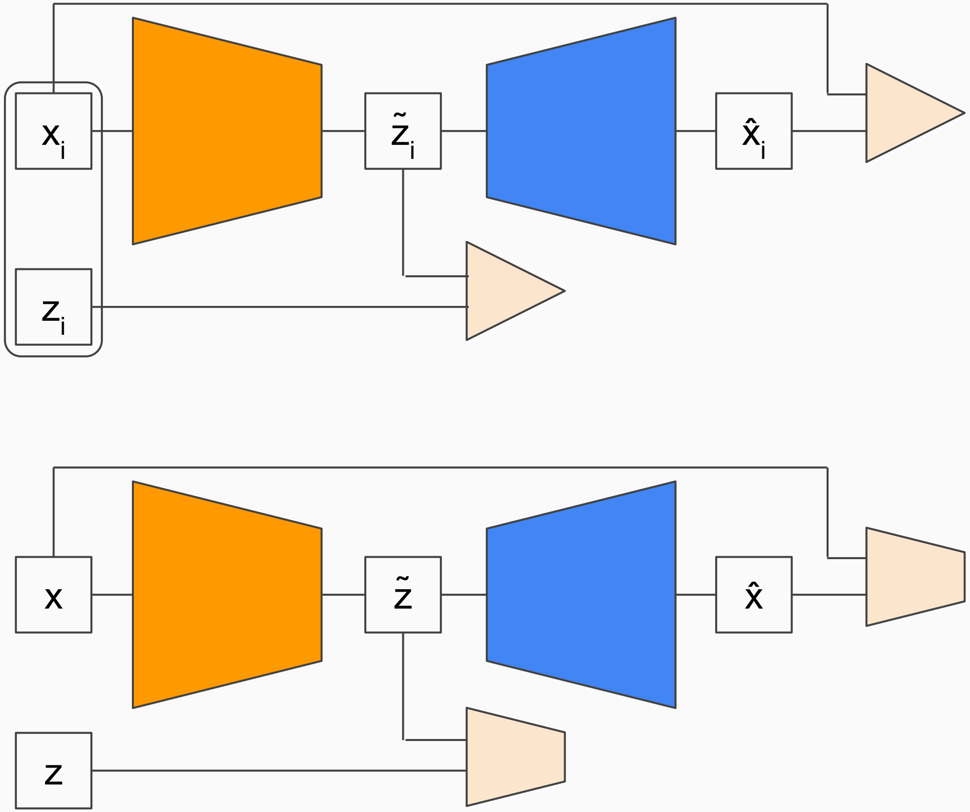
Direct

- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

Reverse:

- The same is done swapping $x \leftrightarrow z$

Training scheme



Direct

- A minibatch of $\{x_i\}$ and $\{x\}$ is passed through the network
- Supervised losses are computed between $\{z_i, x_i\}$ and $\{\tilde{z}_i, \hat{x}_i\}$
- Unsupervised losses are computed between $\{z, x\}$ and $\{\tilde{z}, \hat{x}\}$
- Random noise is added to the input of both the encoder and the decoder

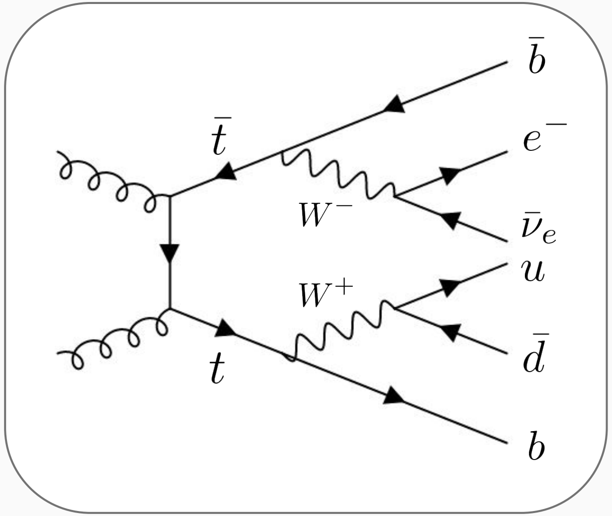
Reverse:

- The same is done swapping $x \leftrightarrow z$

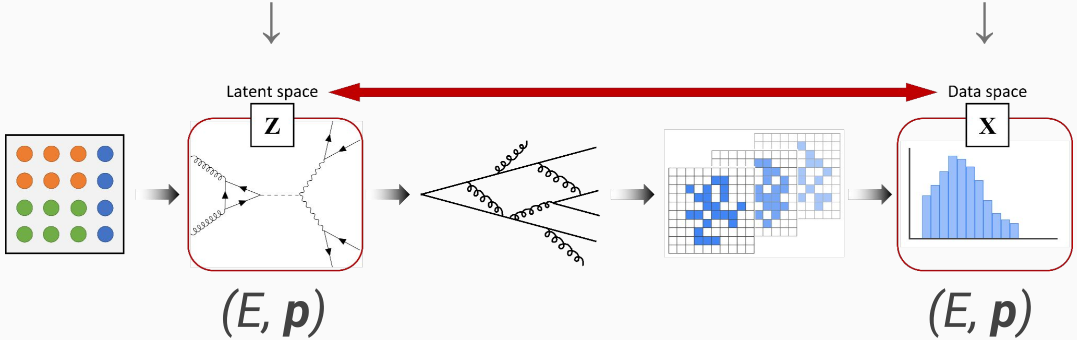
HEP experiment

Dataset and results

Dataset



SM $\rightarrow$ MadGraph $\rightarrow$ Pythia $\rightarrow$ Delphes



$$z = \{E_{e^-}, \mathbf{p}_{e^-}, E_{\bar{\nu}_e}, \mathbf{p}_{\bar{\nu}_e}, E_b, \mathbf{p}_b, E_{\bar{b}}, \mathbf{p}_{\bar{b}}, E_u, \mathbf{p}_u, E_{\bar{d}}, \mathbf{p}_{\bar{d}}\}$$

$$x = \{E_{e^-}, \mathbf{p}_{e^-}, E_{\text{miss}}, \mathbf{p}_{\text{miss}}, E_1, \mathbf{p}_1, E_2, \mathbf{p}_2, E_3, \mathbf{p}_3, E_4, \mathbf{p}_4\}$$

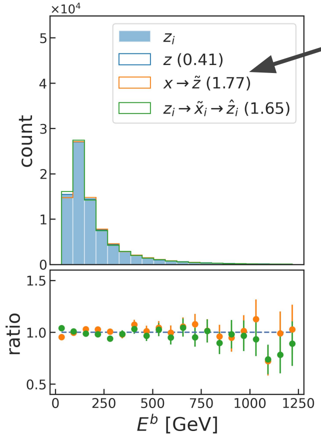
Results

Kolmogorov-Smirnov

x10^{-2}

Legend

- Input data
  {z, x} pairs and {z}, {x} marginals
  (KS for the dataset itself)
- Generated data
  i.e. simulation and unfolding
- Reconstructed data
  i.e. autoencoder output



In parenthesis:
Kolmogorov-Smirnov distance to truth (the lower the better)

Distributions of several observables

Trained on:

- *(E,* ***p****)* of final state particles
- *(E,* ***p****)* of reconstructed particles

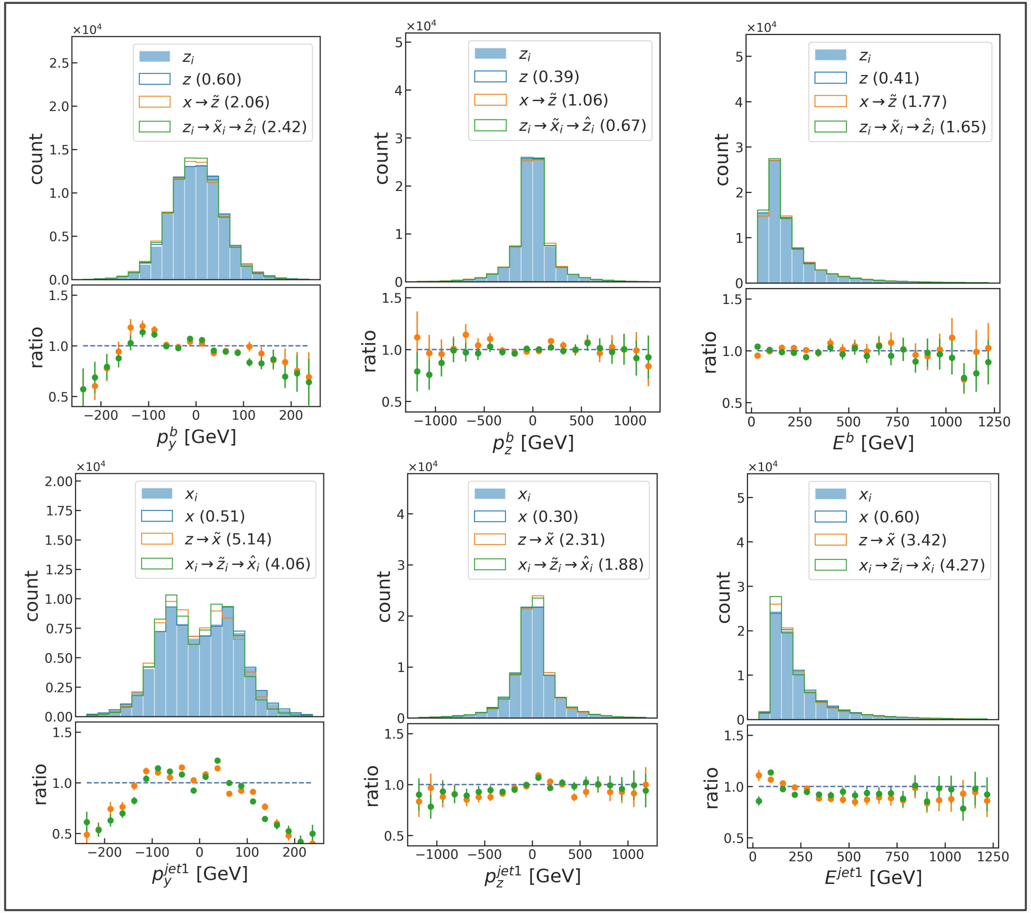
Not trained on:

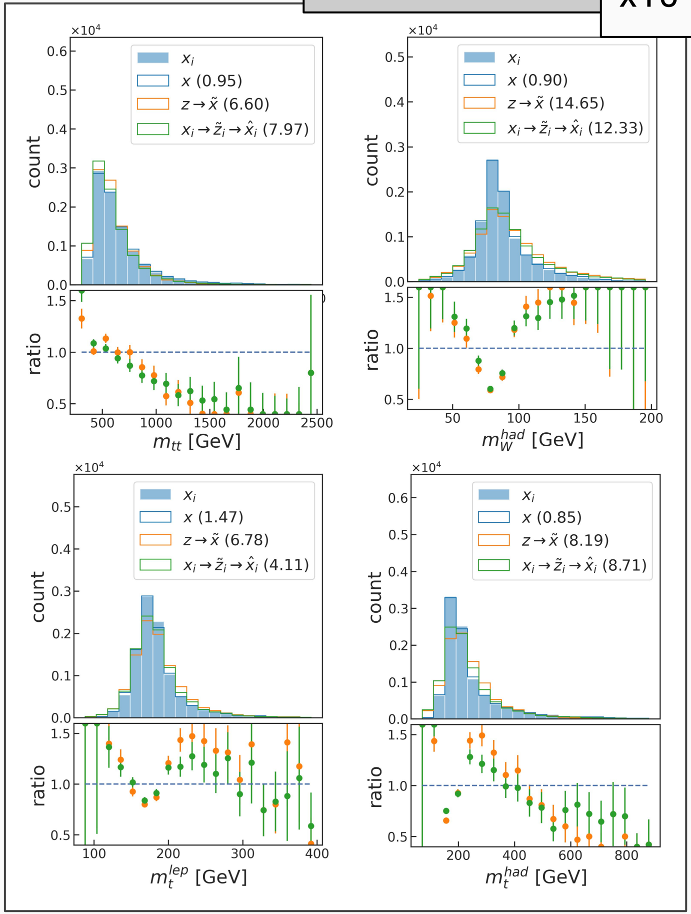
- *mass* of underlying particles (*W*, *t*)

Results

Kolmogorov-Smirnov

x10^{-2}

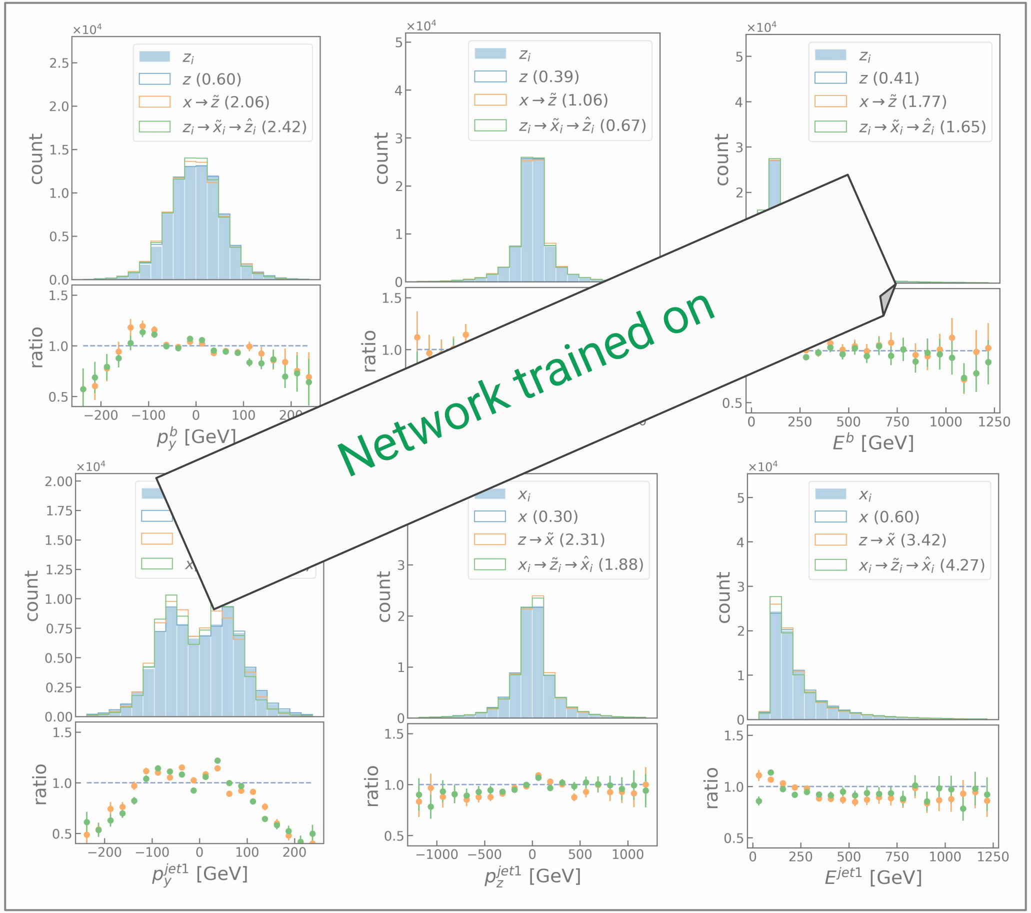


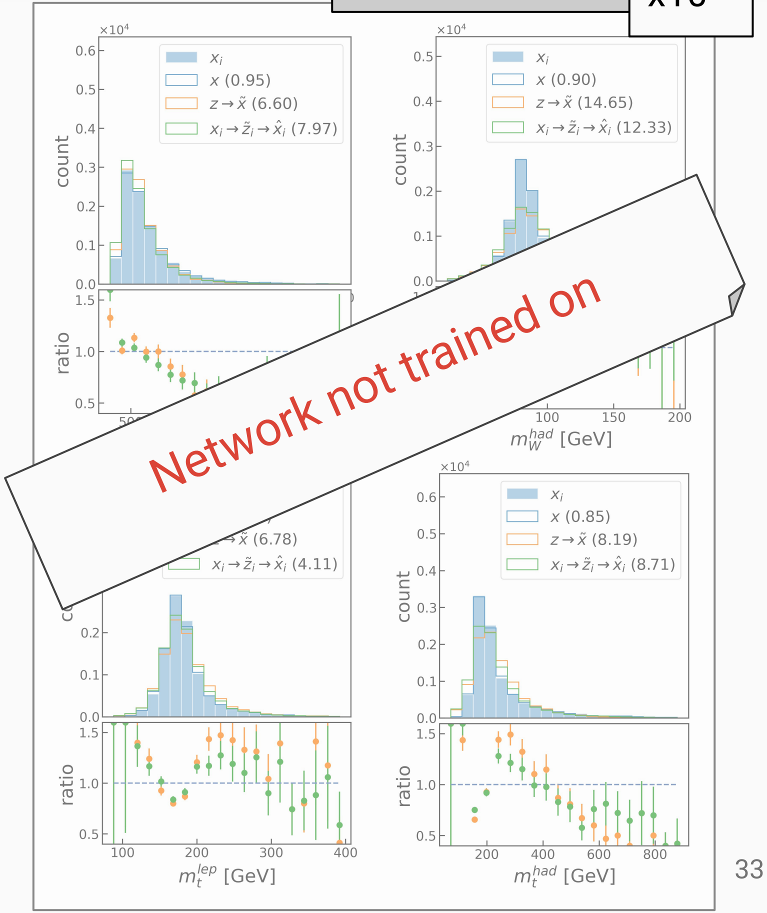


Results

Kolmogorov-Smirnov

x10^{-2}





Results

Kolmogorov-Smirnov

x10^{-2}

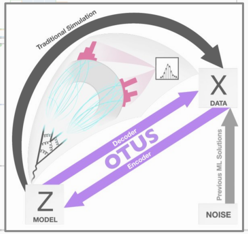
Comparison for unfolding (Z space) and simulation (X space + reco. physics)

	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	2.06	**1.06**	**1.77**	5.14	**2.31**	**3.42**	**6.60**	14.7	**6.78**	**8.19**
OTUS	**1.59**	1.23	2.76	**3.78**	2.39	5.75	15.8	**11.7**	14.1	24.9

Trained on

Not trained on

2101.08944



Conclusion

Summary and outlook

Summary

- Generalised autoencoder based on information theory principles
  - Maximise mutual information between input and latent space
  - Give an interpretation to several architecture (AAE, GAN, etc.)

- Applied to HEP simulation and unfolding
  - Transform four-momenta from hard scattering to detector reconstruction
  - Achieve very good results with basic building blocks

arXiv: 2112.10629

Outlook

- Other scattering processes
- Improve stochasticity *(more expressive network)*
- Permutation invariance *(input ordering ≠ output ordering)*
    - Graph/attention based architecture
- Transform distributions *(specialised architecture)*
    - Conditional Normalising Flow *(see yesterday talk)*
- but variable dimensionality *(of input/output)*
    - Funnels *(see next talk)*

Thanks!