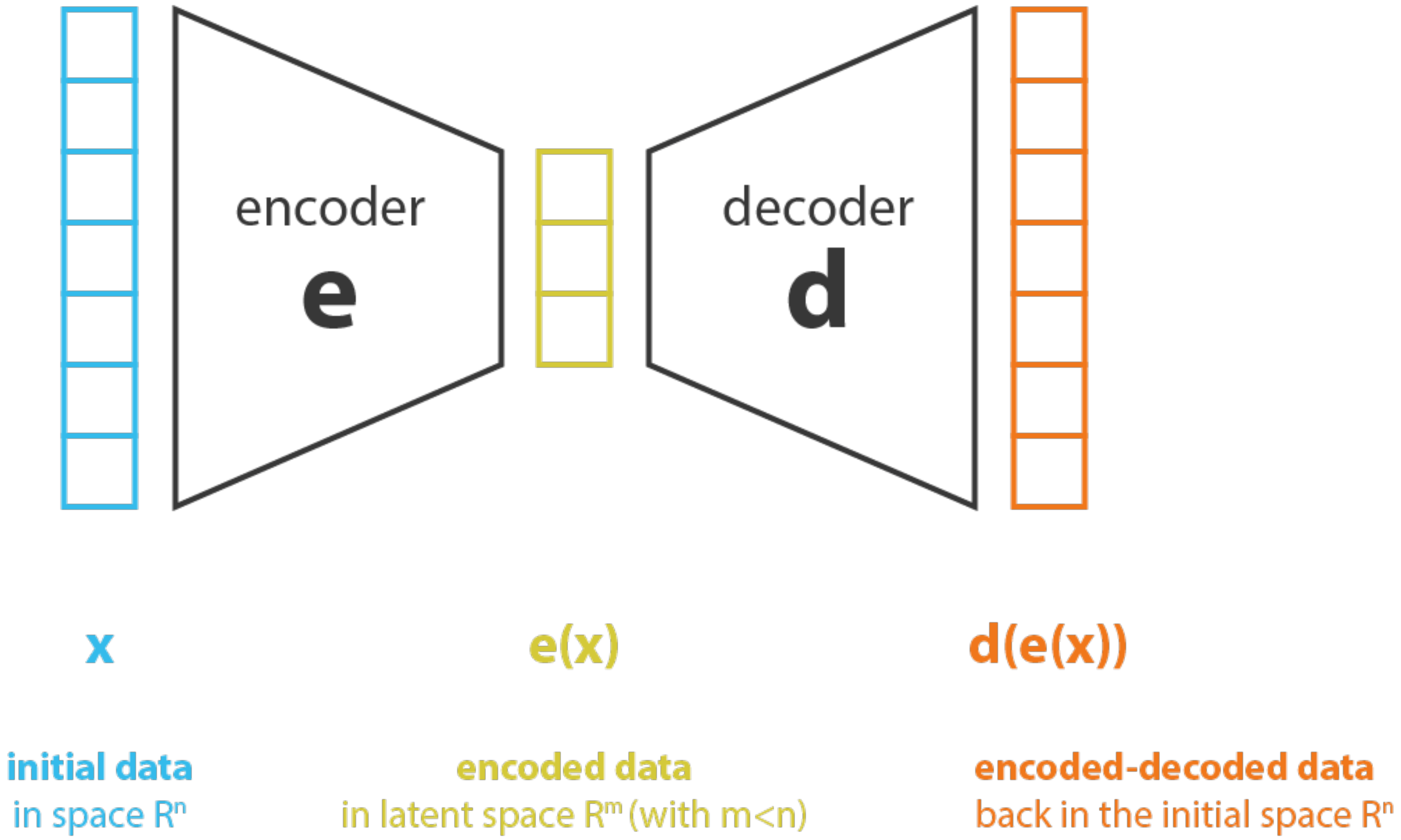
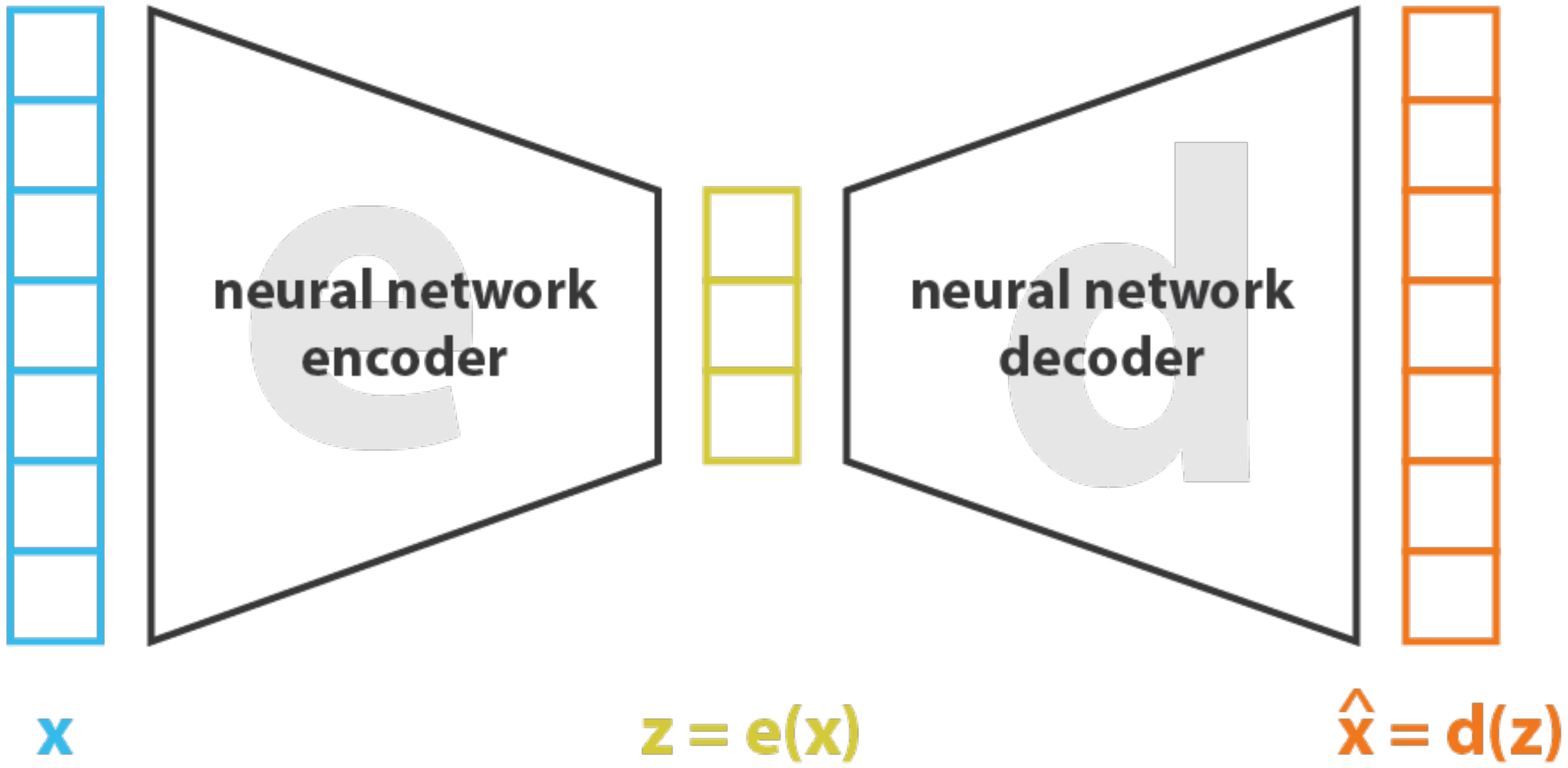
Variational Autoencoder

Dalin Guo, Kuei-Da Liao, Corey Zhou

Encoder and Decoder

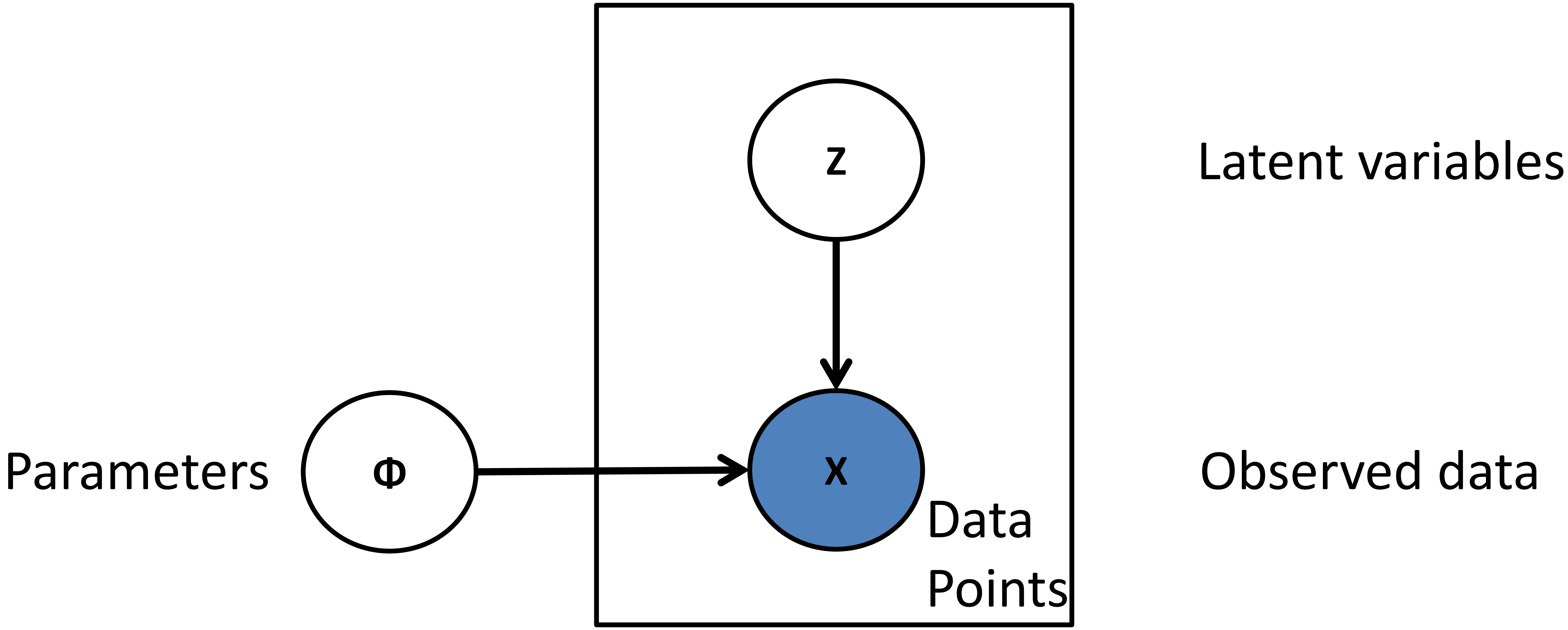


Autoencoder



$$\text{loss} = \| x - \hat{x} \|^2 = \| x - d(z) \|^2 = \| x - d(e(x)) \|^2$$

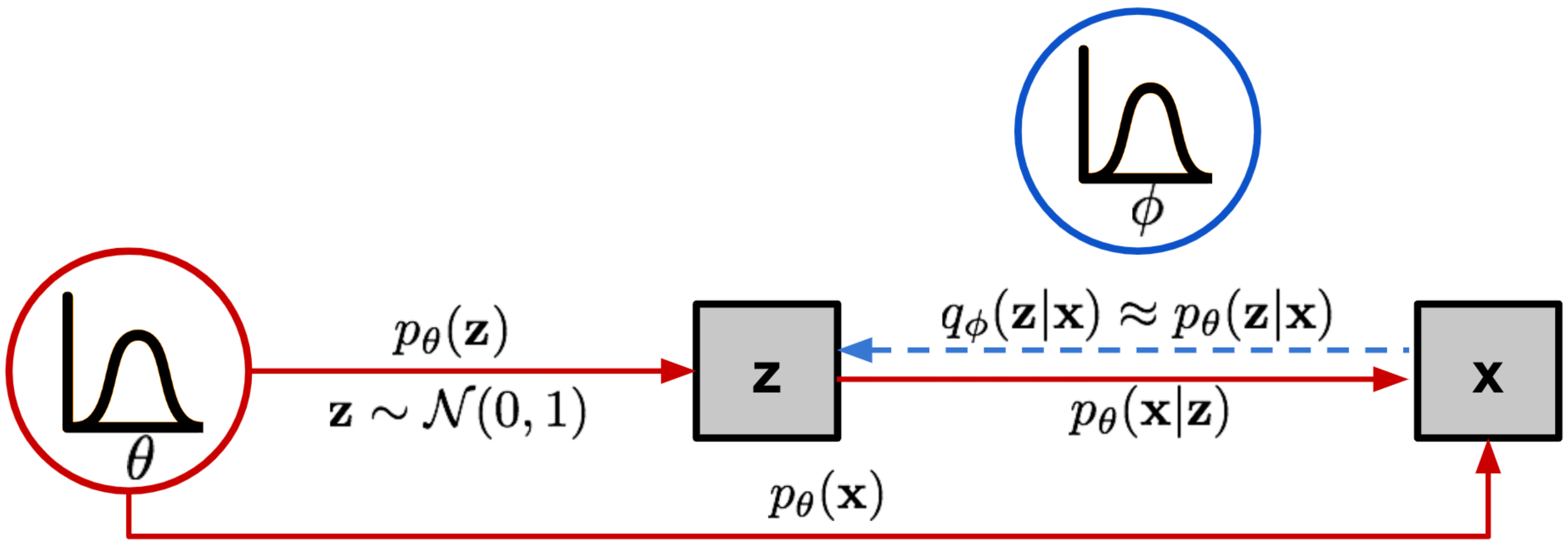
Latent Variable Models

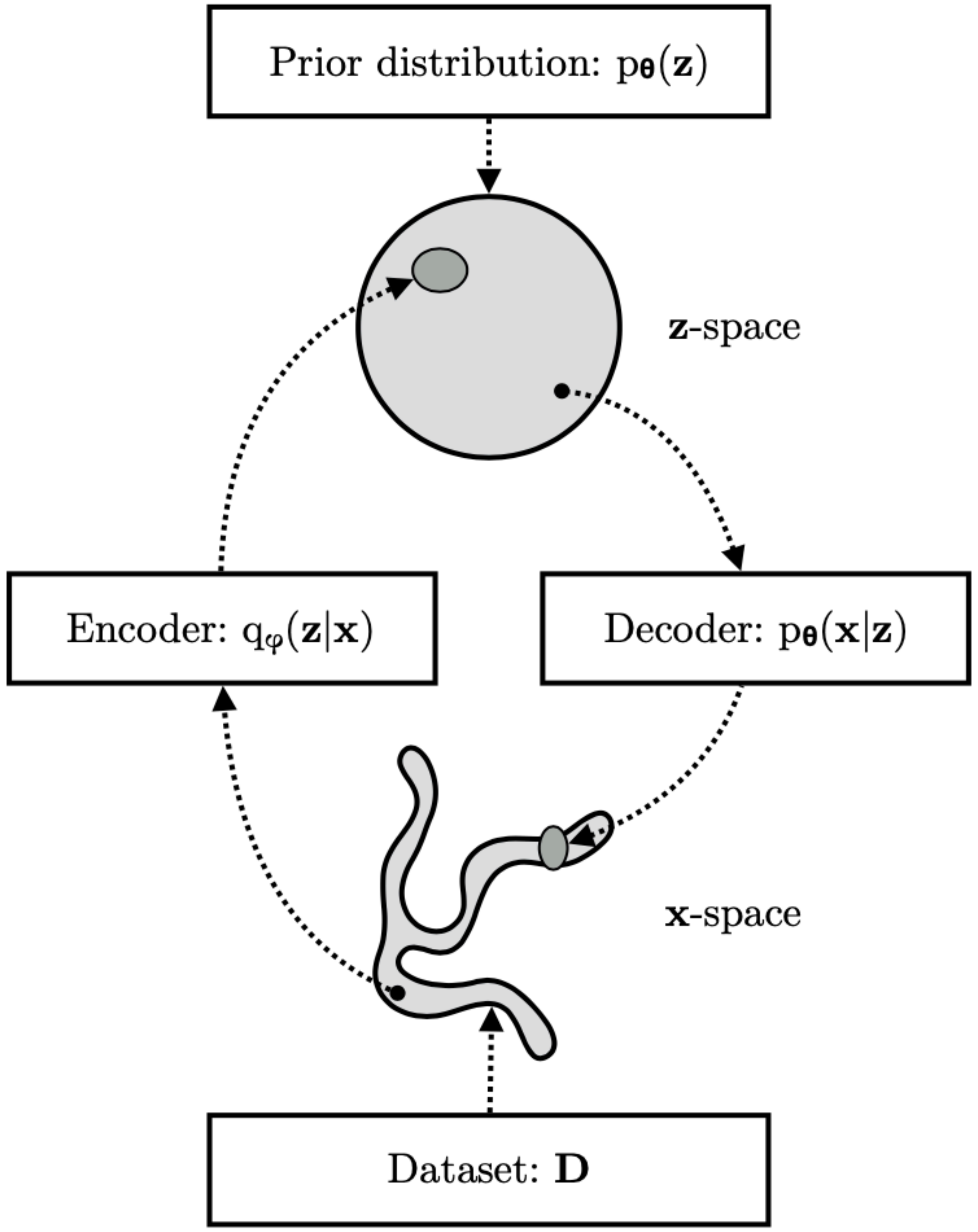


Dimensionality(X) >> dimensionality(Z)

Z is a **bottleneck**, which finds a **compressed**, **low-dimensional representation** of X

Variational Autoencoder (VAE)





AE vs. VAE

	input		latent representation		input reconstruction		
simple autoencoders	**x**	encoding $\rightarrow$	**z = e(x)**	decoding $\rightarrow$	**d(z)**		

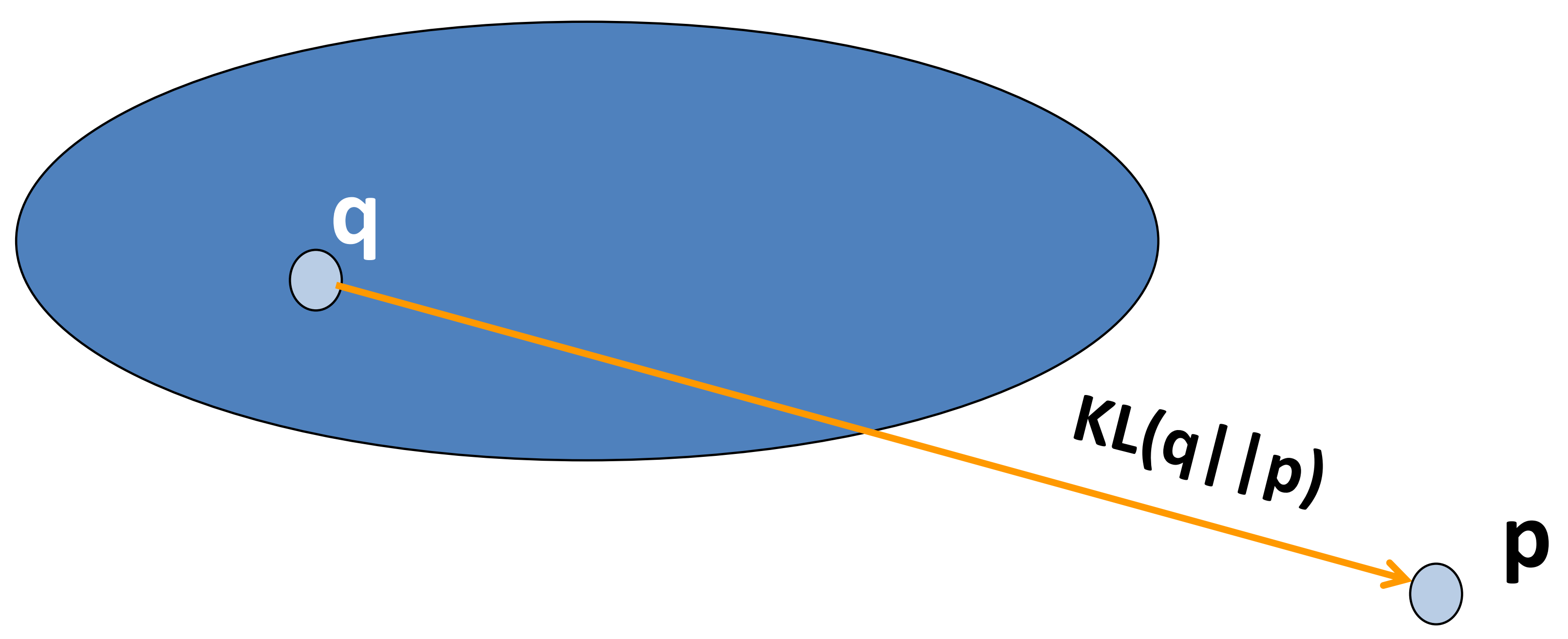
	input		latent distribution		sampled latent representation		input reconstruction
variational autoencoders	**x**	encoding $\rightarrow$	**p(z|x)**	sampling $\rightarrow$	**z ~ p(z|x)**	decoding $\rightarrow$	**d(z)**

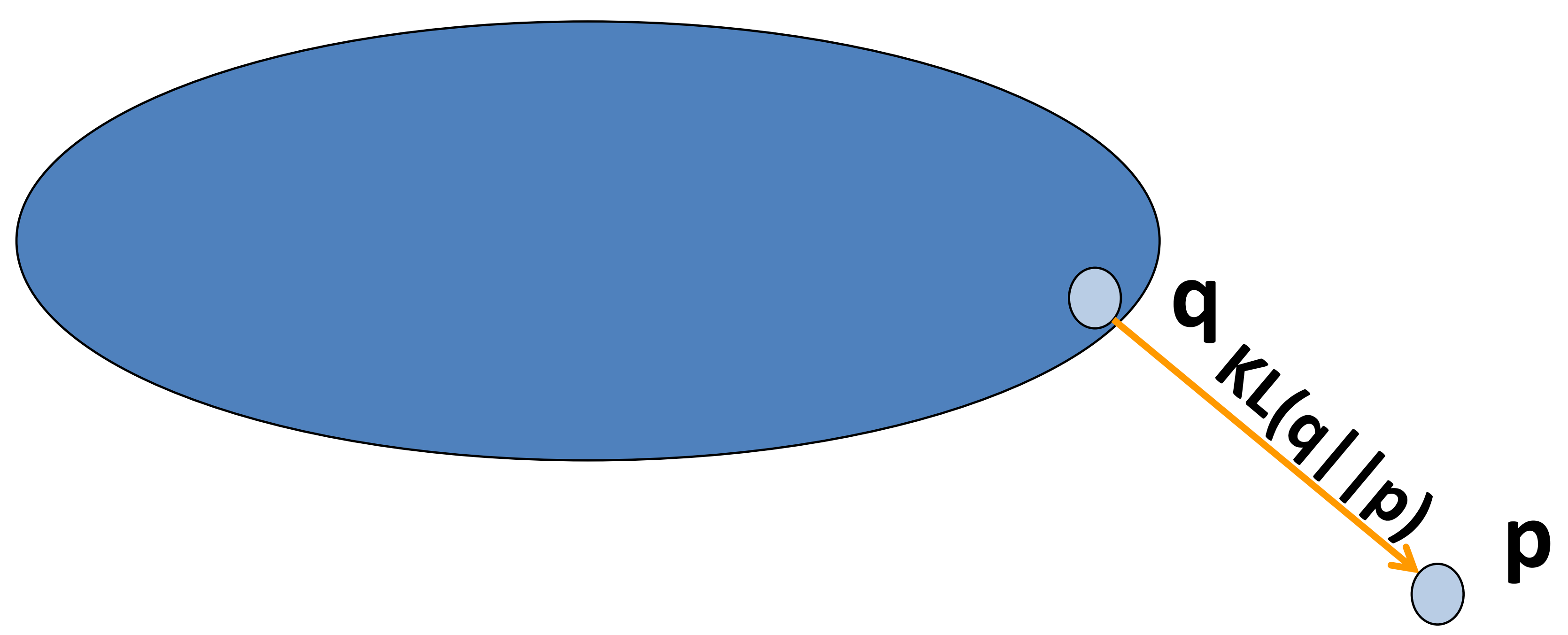
Approximate Inference

- **Optimization approaches**
  - EM
  - Variational inference
    - Variational Bayes
    - Message passing
  - Laplace approximation
- **Simulation approches** (Monte Carlo methods)
  - MCMC: Gibbs sampling, etc
  - …

Variational Inference

- Approximate p(z) with q(z)
- Make q(z) tractable to work with
- Solve an optimization problem, e.g. KL(q(z)||p(z))





Variational Inference

$$KL(P||Q) = \sum_z p(z) \log \frac{p(z)}{q(z)}$$
$$= \mathbb{E}_{p(z)} \left[\log \frac{p(z)}{q(z)} \right]$$

Monte Carlo Methods

$$E_{P(\mathbf{x})}[f(\mathbf{x})] \approx \frac{1}{S}\sum_{i=1}^{S} f(\mathbf{x}^{(i)}), \mathbf{x}^{(i)} \sim P(\mathbf{x})$$

- This suggests the procedure:
  - Draw S samples from P(x)
  - Compute f(x) for each of the samples
  - Approximate E[f(x)] by the sample average

Evidence Lower BOund (ELBO)

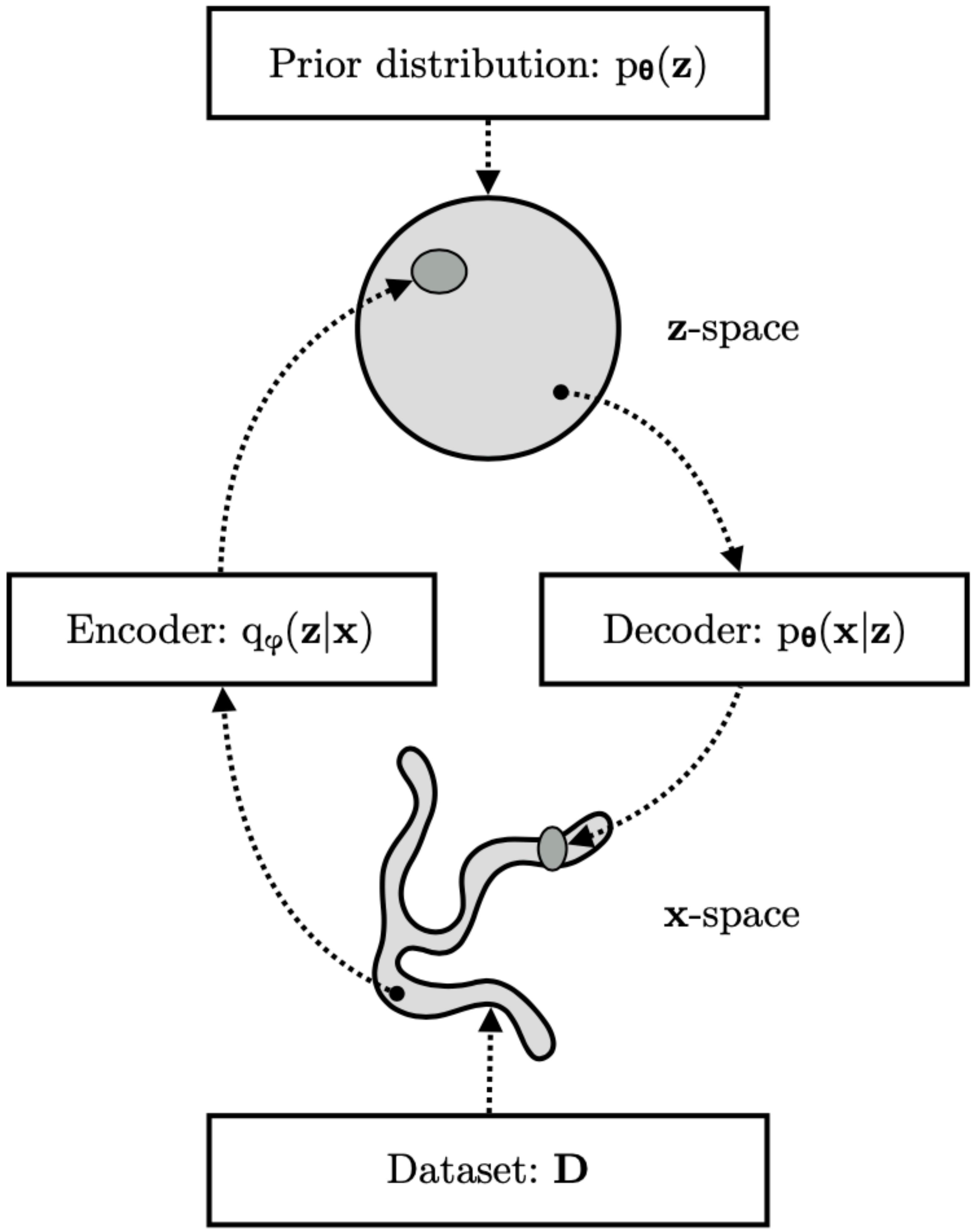
$$\log p_{\boldsymbol{\theta}}(\mathbf{x}) = \mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\left[\log p_{\boldsymbol{\theta}}(\mathbf{x})\right] \tag{2.5}$$

$$= \mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\left[\log\left[\frac{p_{\boldsymbol{\theta}}(\mathbf{x},\mathbf{z})}{p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})}\right]\right] \tag{2.6}$$

$$= \mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\left[\log\left[\frac{p_{\boldsymbol{\theta}}(\mathbf{x},\mathbf{z})}{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\frac{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}{p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})}\right]\right] \tag{2.7}$$

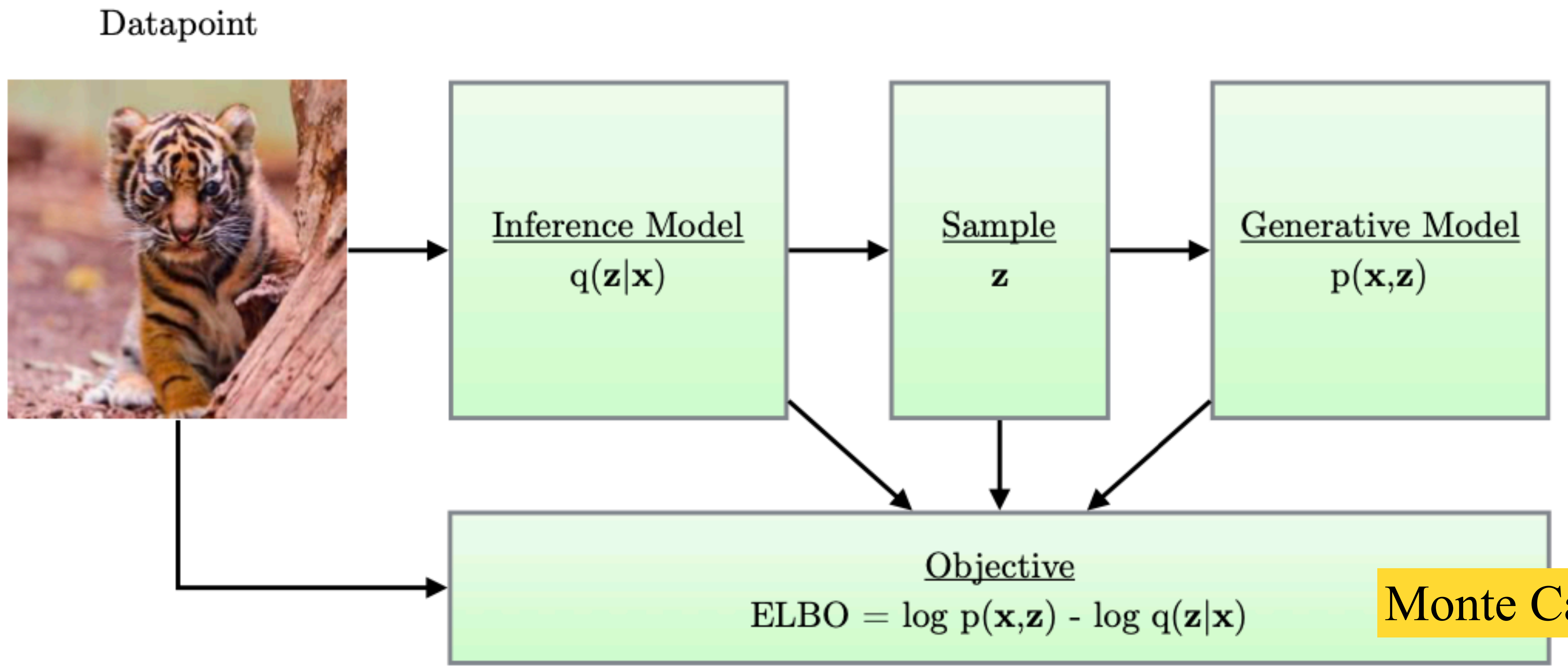
$$= \underbrace{\mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\left[\log\left[\frac{p_{\boldsymbol{\theta}}(\mathbf{x},\mathbf{z})}{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\right]\right]}_{\substack{=\mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\phi}}(\mathbf{x}) \\ \text{(ELBO)}}} + \underbrace{\mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}\left[\log\left[\frac{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}{p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})}\right]\right]}_{=D_{KL}(q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})||p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x}))} \tag{2.8}$$

$$\begin{aligned}\mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\phi}}(\mathbf{x}) &= \log p_{\boldsymbol{\theta}}(\mathbf{x}) - D_{KL}(q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})||p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})) \\ &\leq \log p_{\boldsymbol{\theta}}(\mathbf{x})\end{aligned}$$



ELBO

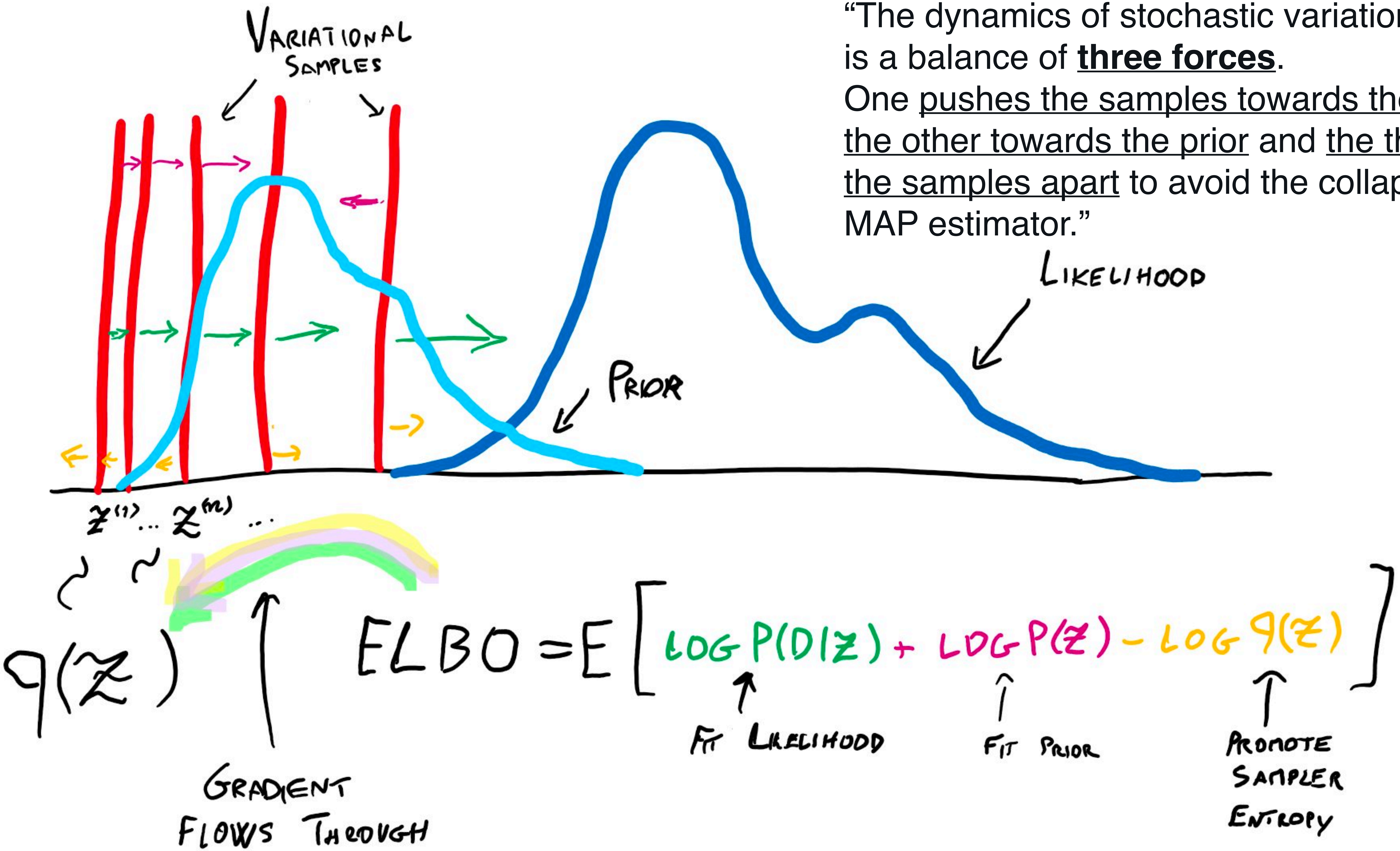
$$
\begin{aligned}
\mathcal{L}_{\theta,\phi}(\mathbf{x}) &= \log p_\theta(\mathbf{x}) - D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) \\
&= \int q_\phi(\mathbf{z}|\mathbf{x}) \log p_\theta(\mathbf{x}) d\mathbf{z} + \int q_\phi(\mathbf{z}|\mathbf{x}) \log p_\theta(\mathbf{z}|\mathbf{x}) d\mathbf{z} - \int q_\phi(\mathbf{z}|\mathbf{x}) \log q_\phi(\mathbf{z}|\mathbf{x}) d\mathbf{z} \\
&= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x},\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})] \\
&= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})]
\end{aligned}
$$



Monte Carlo estimator

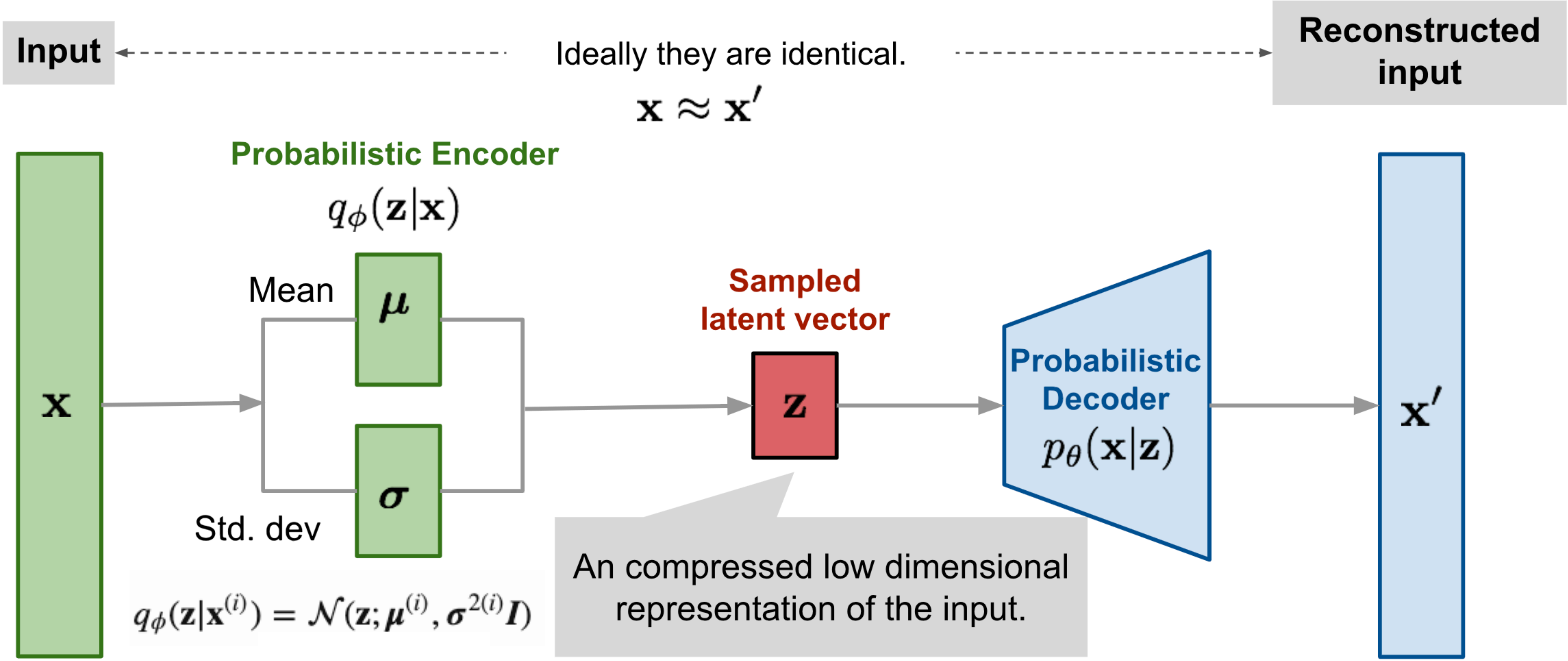
“The dynamics of stochastic variational inference is a balance of **three forces**.
One pushes the samples towards the likelihood, the other towards the prior and the third pushes the samples apart to avoid the collapse into the MAP estimator.”

@LucaAmb



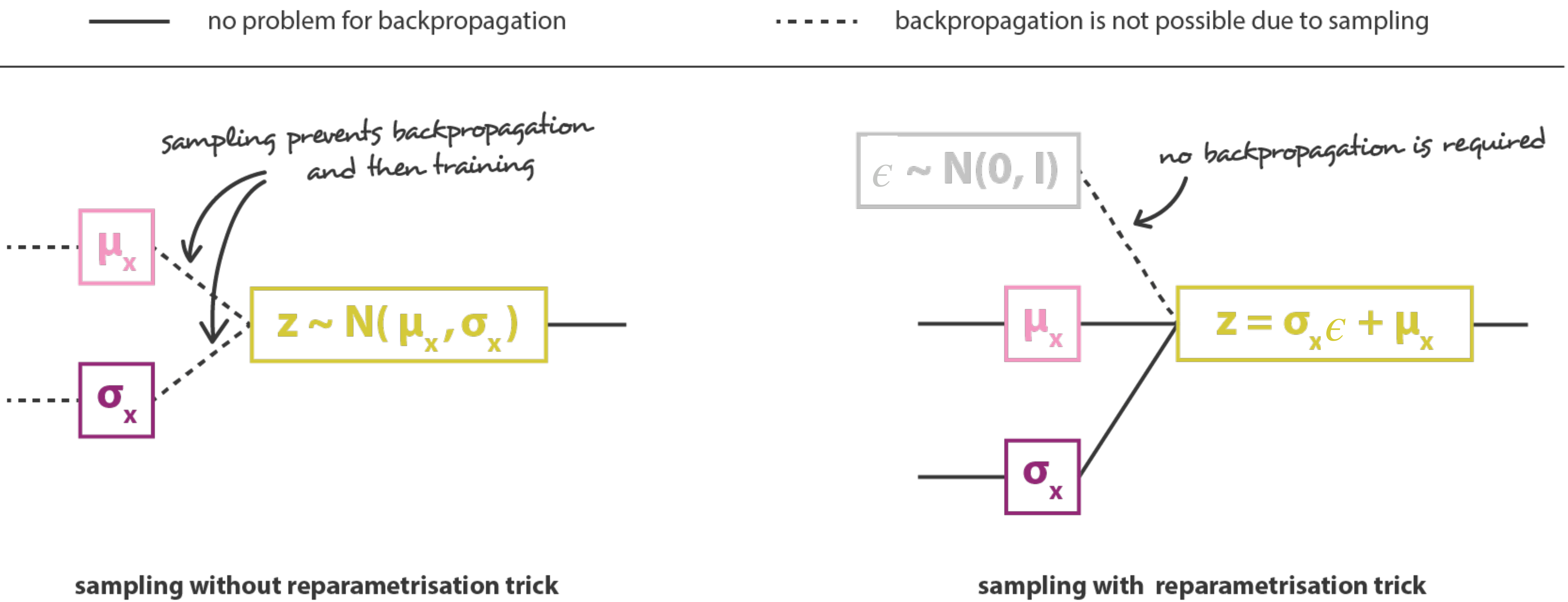
$$\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})]$$

(vanilla) VAE

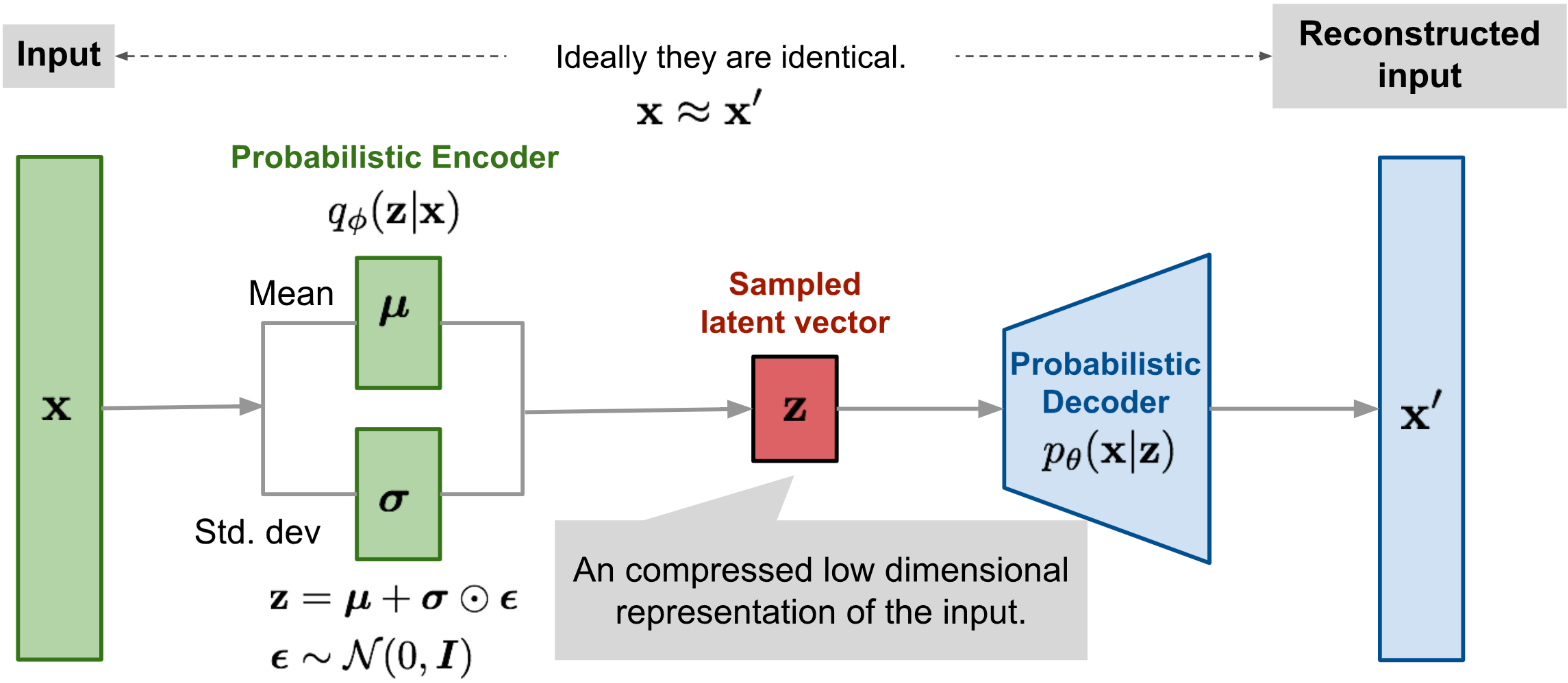


Input
Ideally they are identical.
$\mathbf{x} \approx \mathbf{x}'$
Reconstructed input
Probabilistic Encoder
$q_\phi(\mathbf{z}|\mathbf{x})$
Mean
$\boldsymbol{\mu}$
Std. dev
$\boldsymbol{\sigma}$
$\mathbf{x}$
Sampled latent vector
$\mathbf{z}$
Probabilistic Decoder
$p_\theta(\mathbf{x}|\mathbf{z})$
$\mathbf{x}'$
$q_\phi(\mathbf{z}|\mathbf{x}^{(i)}) = \mathcal{N}(\mathbf{z}; \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{2(i)}\boldsymbol{I})$
An compressed low dimensional representation of the input.

Reparameterization trick



(vanilla) VAE



Input
Ideally they are identical.
$\mathbf{x} \approx \mathbf{x}'$
Reconstructed input
Probabilistic Encoder
$q_\phi(\mathbf{z}|\mathbf{x})$
$\mathbf{x}$
Mean
$\boldsymbol{\mu}$
$\boldsymbol{\sigma}$
Std. dev
$\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \boldsymbol{\epsilon}$
$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{I})$
Sampled latent vector
$\mathbf{z}$
An compressed low dimensional representation of the input.
Probabilistic Decoder
$p_\theta(\mathbf{x}|\mathbf{z})$
$\mathbf{x}'$

ELBO

$$
\begin{aligned}
\mathcal{L}_{\theta,\phi}(\mathbf{x}) &= \log p_\theta(\mathbf{x}) - D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) \\
&= \int q_\phi(\mathbf{z}|\mathbf{x}) \log p_\theta(\mathbf{x}) d\mathbf{z} + \int q_\phi(\mathbf{z}|\mathbf{x}) \log p_\theta(\mathbf{z}|\mathbf{x}) d\mathbf{z} - \int q_\phi(\mathbf{z}|\mathbf{x}) \log q_\phi(\mathbf{z}|\mathbf{x}) d\mathbf{z} \\
&= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}, \mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})] \\
&= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})] \\
&= \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}
\end{aligned}
$$

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

Assuming a Gaussian likelihood for the decoder, i.e. $\mathbf{x}|\mathbf{z} = z \sim \mathcal{N}(\boldsymbol{\mu}_{\theta,z}, \boldsymbol{\Sigma}_{\theta,z})$, we have

$$\begin{aligned}
\log p_\theta(\mathbf{x}|\mathbf{z}) &= \log\left[(2\pi)^{-\frac{J}{2}}|\boldsymbol{\Sigma}_{\theta,z}|^{-\frac{1}{2}}\exp(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\theta,z})^T\boldsymbol{\Sigma}_{\theta,z}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\theta,z}))\right] \\
&= -\frac{1}{2}\left[(\mathbf{x}-\boldsymbol{\mu}_{\theta,z})^T\boldsymbol{\Sigma}_{\theta,z}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\theta,z}) + J\log(2\pi) + \log|\boldsymbol{\Sigma}_{\theta,z}|\right]
\end{aligned}$$

In the isotropic Gaussian case, this reduces to

$$\begin{aligned}
\log p_\theta(\mathbf{x}|\mathbf{z}) &= \sum_{i=1}^{J}\log\left[(2\pi)^{-\frac{1}{2}}\sigma_i^{-1}e^{-\frac{1}{2\sigma_i^2}(x_i-\mu_{\theta,z,i})^2}\right] \\
&= -\frac{1}{2}\sum_{i=1}^{J}\left[\frac{(x_i-\mu_{\theta,z,i})^2}{\sigma_i^2} + \log(2\pi) + \log\sigma_i^2\right],
\end{aligned}$$

where $\mu_{\theta,z,i}$ is the i^{th} element of $\boldsymbol{\mu}_{\theta,z}$, and σ_i^2 is the i^{th} diagonal entry of $\boldsymbol{\Sigma}_{\theta,z}$.

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

$$\begin{aligned}
\log p_\theta(\mathbf{z}) &= \sum_{i=1}^{J} \log \mathcal{N}(z_i; 0, 1) \\
&= \sum_{i=1}^{J} \log \left[(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2} z_i^2} \right] \\
&= -\sum_{i=1}^{J} \frac{1}{2} (z_i^2 + \log(2\pi))
\end{aligned}$$

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

Reparameterization

ϵ-> z:

Change of variables formula (vector case):

$$p(\boldsymbol{\epsilon}) = q_{\boldsymbol{\phi}}(\mathbf{g}(\boldsymbol{\epsilon}))\left|\det\left[\frac{\partial \mathbf{g}(\boldsymbol{\epsilon})}{\partial \boldsymbol{\epsilon}}\right]\right|$$

$$\log p(\boldsymbol{\epsilon}) = \log q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}) + \log\left|\det\left[\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}}\right]\right|$$

$$\log q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}) = \log p(\boldsymbol{\epsilon}) - \log\left|\det\left[\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}}\right]\right|$$

$$\log\left|\det\left[\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}}\right]\right| = \sum_{i=1}^{J} \log \sigma_i = \frac{1}{2}\sum_{i=1}^{J} \log \sigma_i^2$$

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

$$\begin{aligned}
\mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{z})] &= \mathbb{E}_{p(\epsilon)}\left[-\frac{1}{2}\sum_{i=1}^{J}(z_i^2 + \log(2\pi))\right] \\
&= -\frac{1}{2}\sum_{i=1}^{J}\mathbb{E}_{p(\epsilon)}[(\mu_i + \sigma_i \cdot \epsilon_i)^2] - \frac{1}{2}\sum_{i=1}^{J}\log(2\pi) \\
&= -\frac{1}{2}\sum_{i=1}^{J}(\mathbb{E}_{p(\epsilon)}[\mu_i^2] + \mathbb{E}_{p(\epsilon)}[\sigma_i^2\epsilon_i^2] + \mathbb{E}_{p(\epsilon)}[2\mu_i\sigma_i\epsilon_i]) - \frac{J}{2}\log(2\pi) \\
&= -\frac{1}{2}\sum_{i=1}^{J}(\mu_i^2 + \sigma_i^2\mathbb{E}_{p(\epsilon)}[\epsilon_i^2] + 2\mu_i\sigma_i\mathbb{E}_{p(\epsilon)}[\epsilon_i]) - \frac{J}{2}\log(2\pi) \\
&= -\frac{1}{2}\sum_{i=1}^{J}(\mu_i^2 + \sigma_i^2 + 2\mu_i\sigma_i \cdot 0) - \frac{J}{2}\log(2\pi) \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{J}(\mu_i^2 + \sigma_i^2)
\end{aligned}$$

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

$$\begin{aligned}
\mathbb{E}_{p(\epsilon)}[\log q_\phi(\mathbf{z}|\mathbf{x})] &= \mathbb{E}_{p(\epsilon)}\left[-\frac{1}{2}\sum_{i=1}^{J}(\epsilon_i^2 + \log(2\pi) + \log\sigma_i^2)\right] \\
&= -\frac{1}{2}\sum_{i=1}^{J}\mathbb{E}_{p(\epsilon)}[\epsilon_i^2] - \frac{1}{2}\sum_{i=1}^{J}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{J}\log\sigma_i^2 \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{J}(1 + \log\sigma_i^2).
\end{aligned}$$

ELBO

$$\mathcal{L}_{\theta,\phi}(\mathbf{x}) = \mathbb{E}_{p(\epsilon)}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})], \text{ where } \mathbf{z} = \boldsymbol{\sigma} \odot \boldsymbol{\epsilon} + \boldsymbol{\mu}$$

$$\mathcal{L}_{\theta,\phi}(x_j) \simeq \frac{1}{2}\sum_{i=1}^{J}(1 + \log((\sigma_i^{(j)})^2) - (\mu_i^{(j)})^2 - (\sigma_i^{(j)})^2) + \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(x_j|z_j)\right]$$

ELBO: full covariance

- Reparameterization

$$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$$

$$\mathbf{z} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\epsilon}$$

$$\begin{aligned}\boldsymbol{\Sigma} &= \mathbb{E}\left[(\mathbf{z} - \mathbb{E}[\mathbf{z}|])(\mathbf{z} - \mathbb{E}[\mathbf{z}|])^T\right] \\ &= \mathbb{E}\left[\mathbf{L}\boldsymbol{\epsilon}(\mathbf{L}\boldsymbol{\epsilon})^T\right] = \mathbf{L}\mathbb{E}\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\right]\mathbf{L}^T \\ &= \mathbf{L}\mathbf{L}^T\end{aligned}$$

$$\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}} = \mathbf{L}$$

$$\log|\det(\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}})| = \sum_i \log|L_{ii}|$$

$$(\boldsymbol{\mu}, \log \boldsymbol{\sigma}, \mathbf{L}') \leftarrow \text{EncoderNeuralNet}_{\phi}(\mathbf{x})$$

$$\mathbf{L} \leftarrow \mathbf{L}_{mask} \odot \mathbf{L}' + \text{diag}(\boldsymbol{\sigma})$$

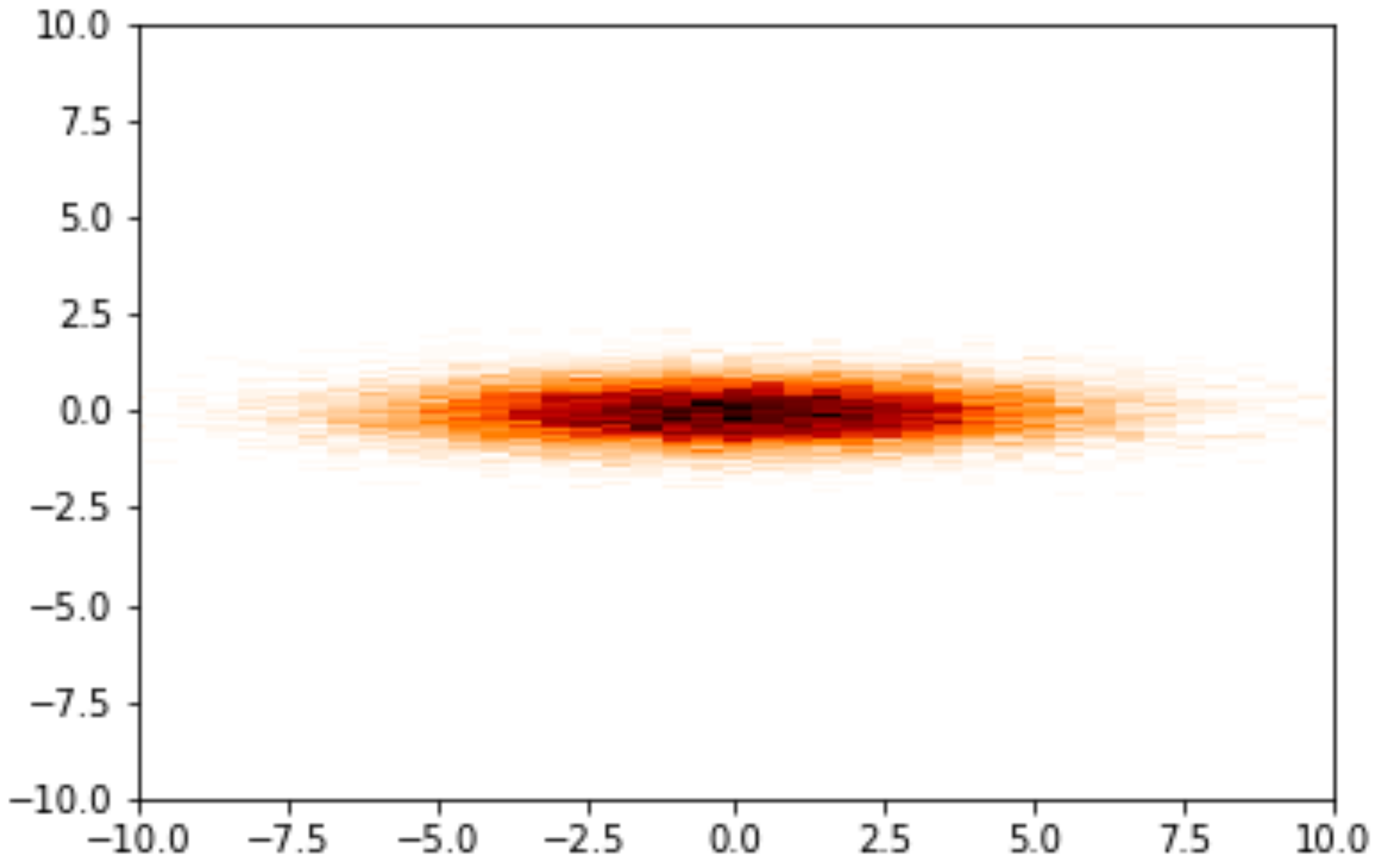
$$\log\left|\det\left(\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}}\right)\right| = \sum_i \log \sigma_i$$

Note that σ_i here is not standard deviation

Experiment

- VAE:
  - Encoder: 2 layers dense neural network (first layer shared), size [128, 128]
  - Decoder: same as encoder
- Dataset: synthetic data
  - 2d-Gaussian
  - Figure 8
  - 3 blobs
  - Mixture of Gaussian

Experiment: isotropic Gaussian



$$Z \sim N(0, I_2)$$

$$\epsilon \sim N(0, {\sigma_\epsilon}^2 \cdot I_2)$$

$$X|Z = \begin{bmatrix} 3 & 0 \\ 0 & 1/3 \end{bmatrix} \cdot Z + \epsilon$$

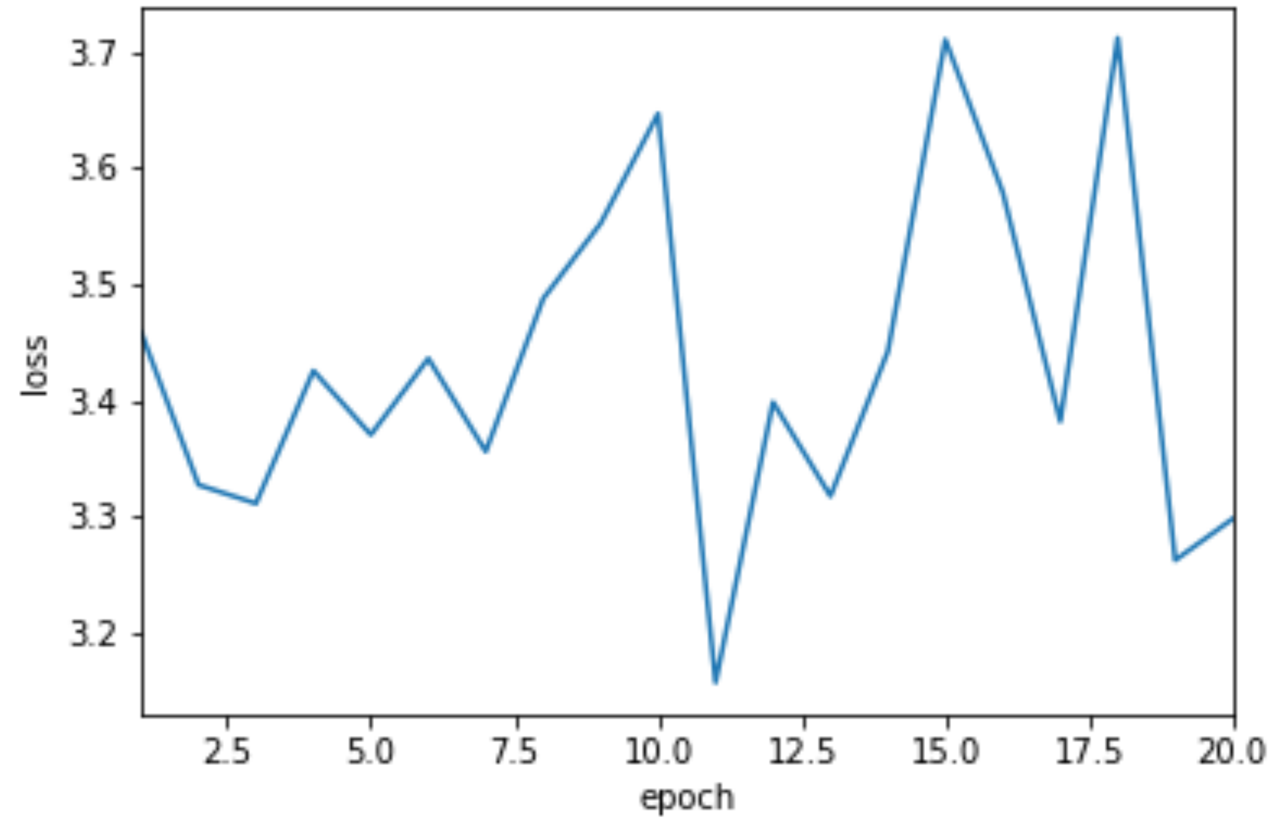
where ${\sigma_\epsilon}^2 = 0.2$

- Diagonal covariance model:
  - # of training samples: 20,000
  - latent dimension: 2
  - batch_size: 32
  - learning rate: 1e-3
  - # epoch: 20

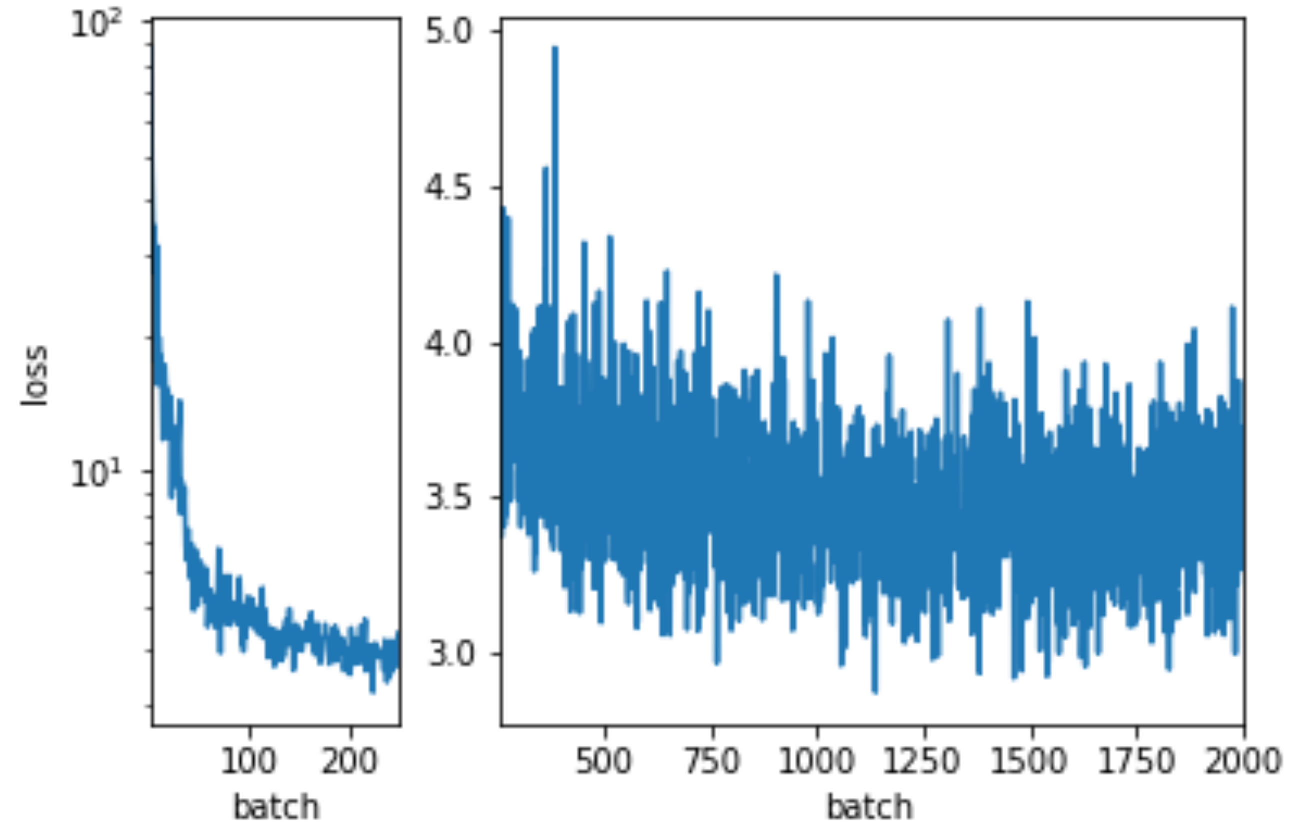
- Testing:
  - # generative samples: 10,000
  - # reconstruction of test set: 10,000

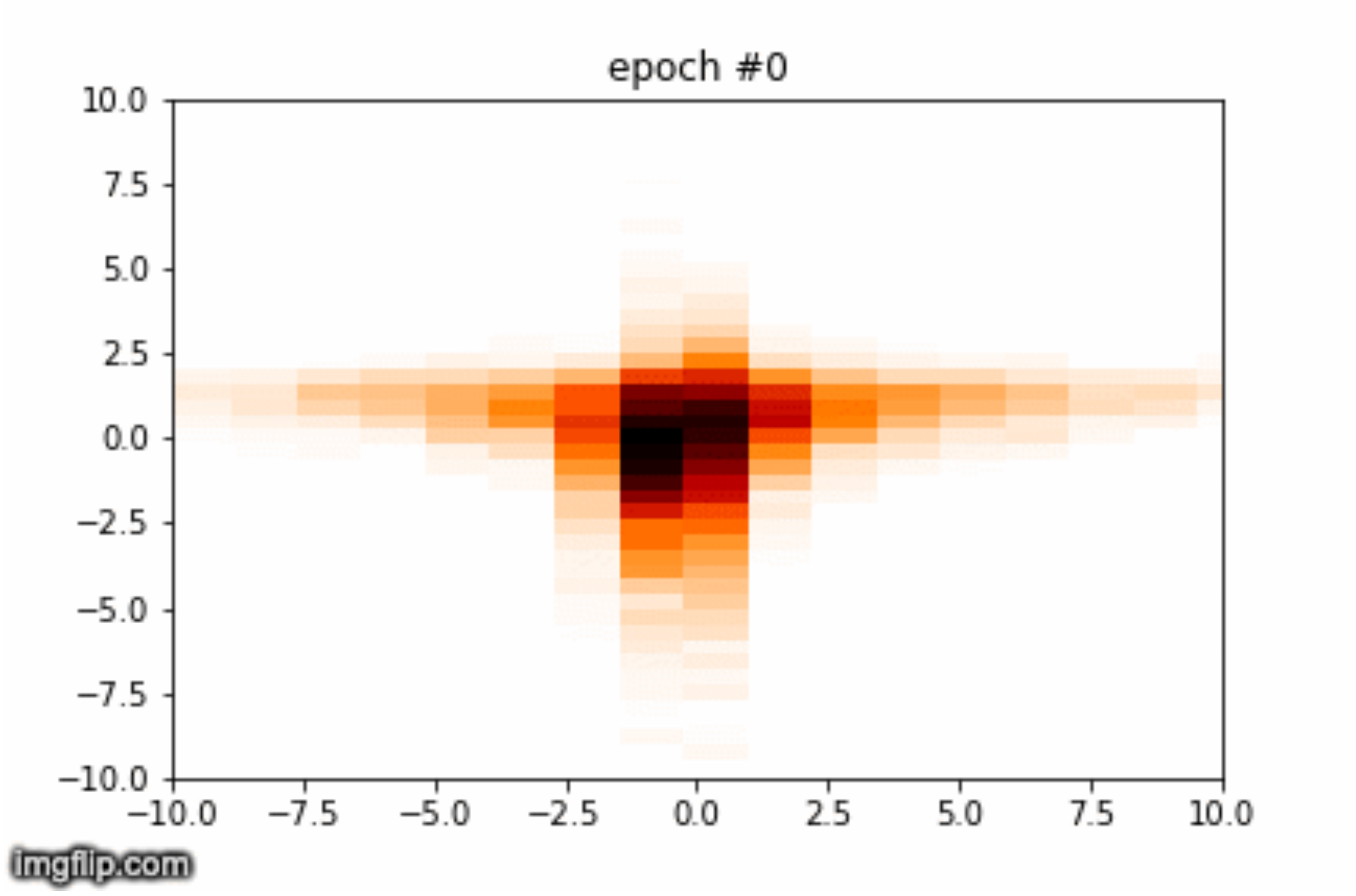
Experiment: isotropic Gaussian

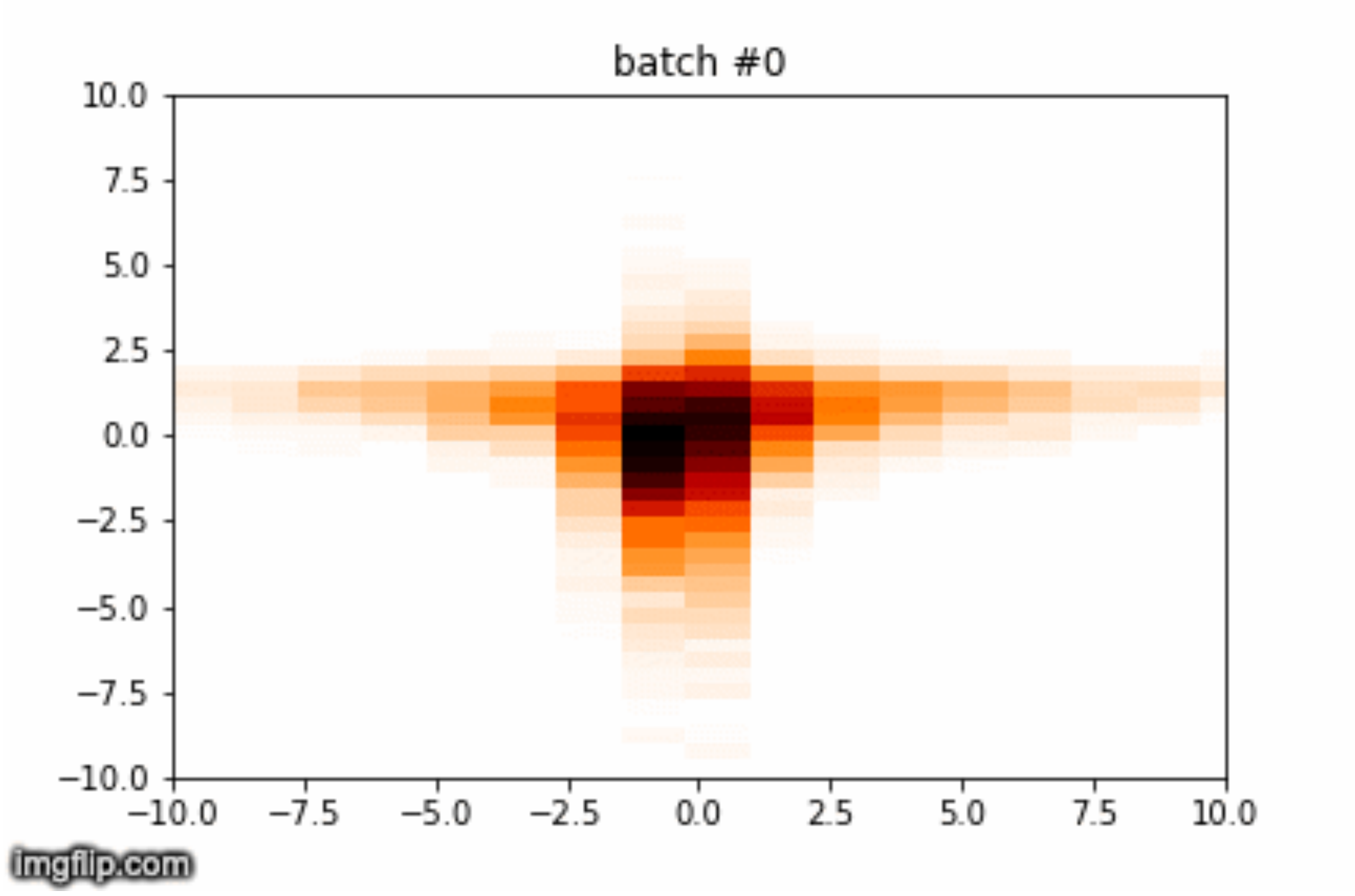
Epoch:



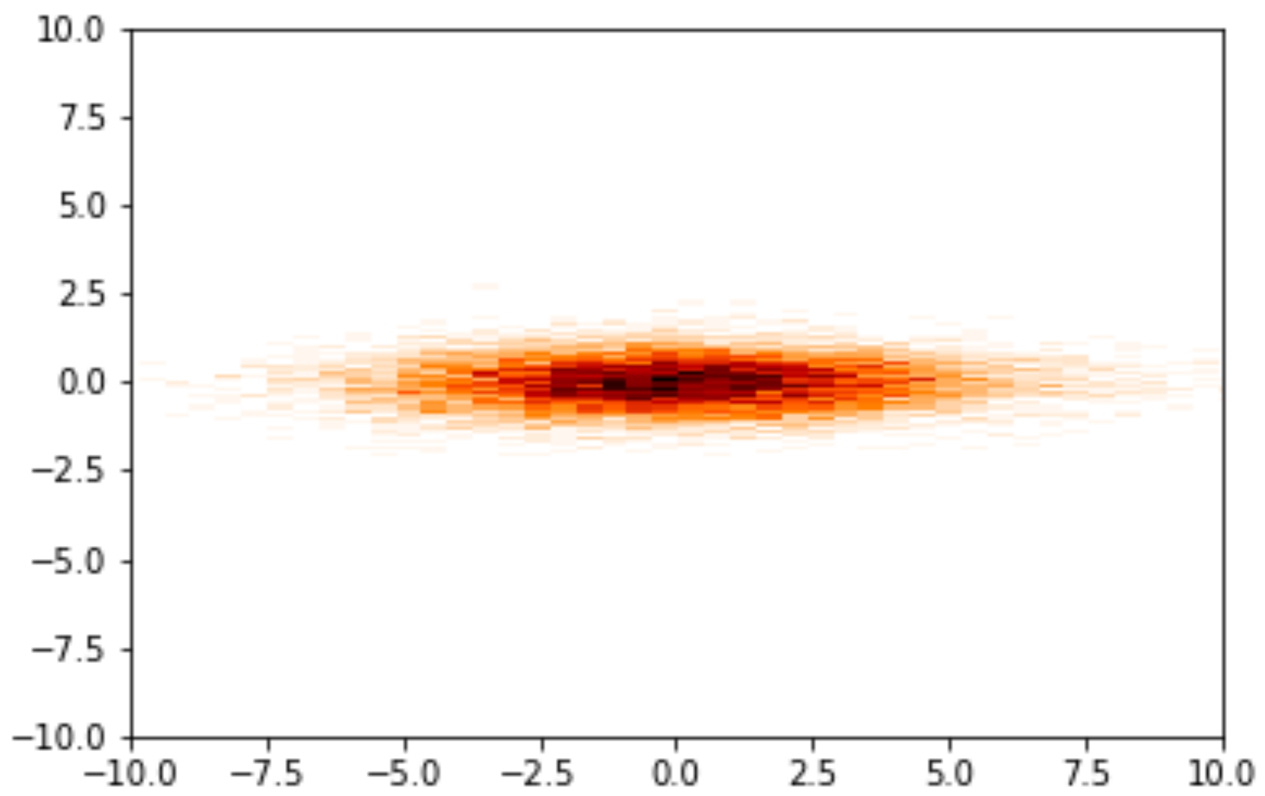
Batch:



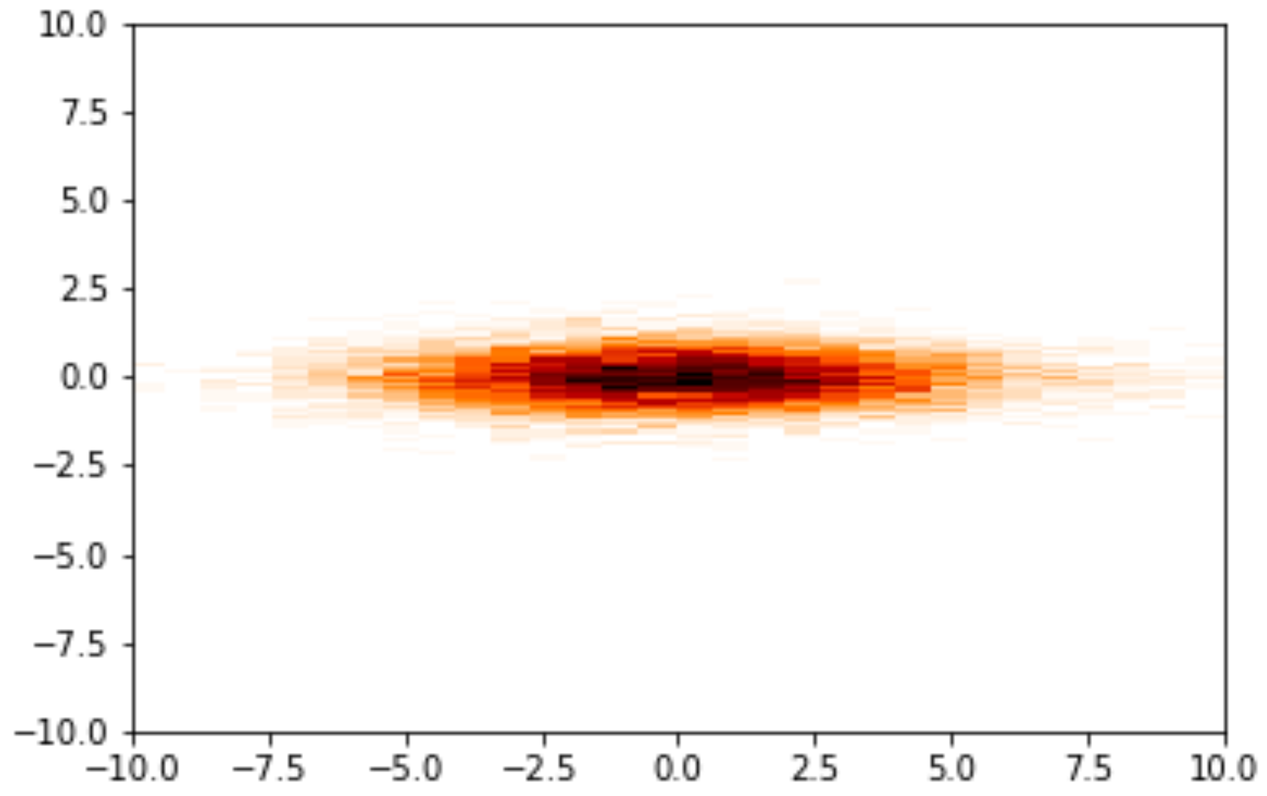




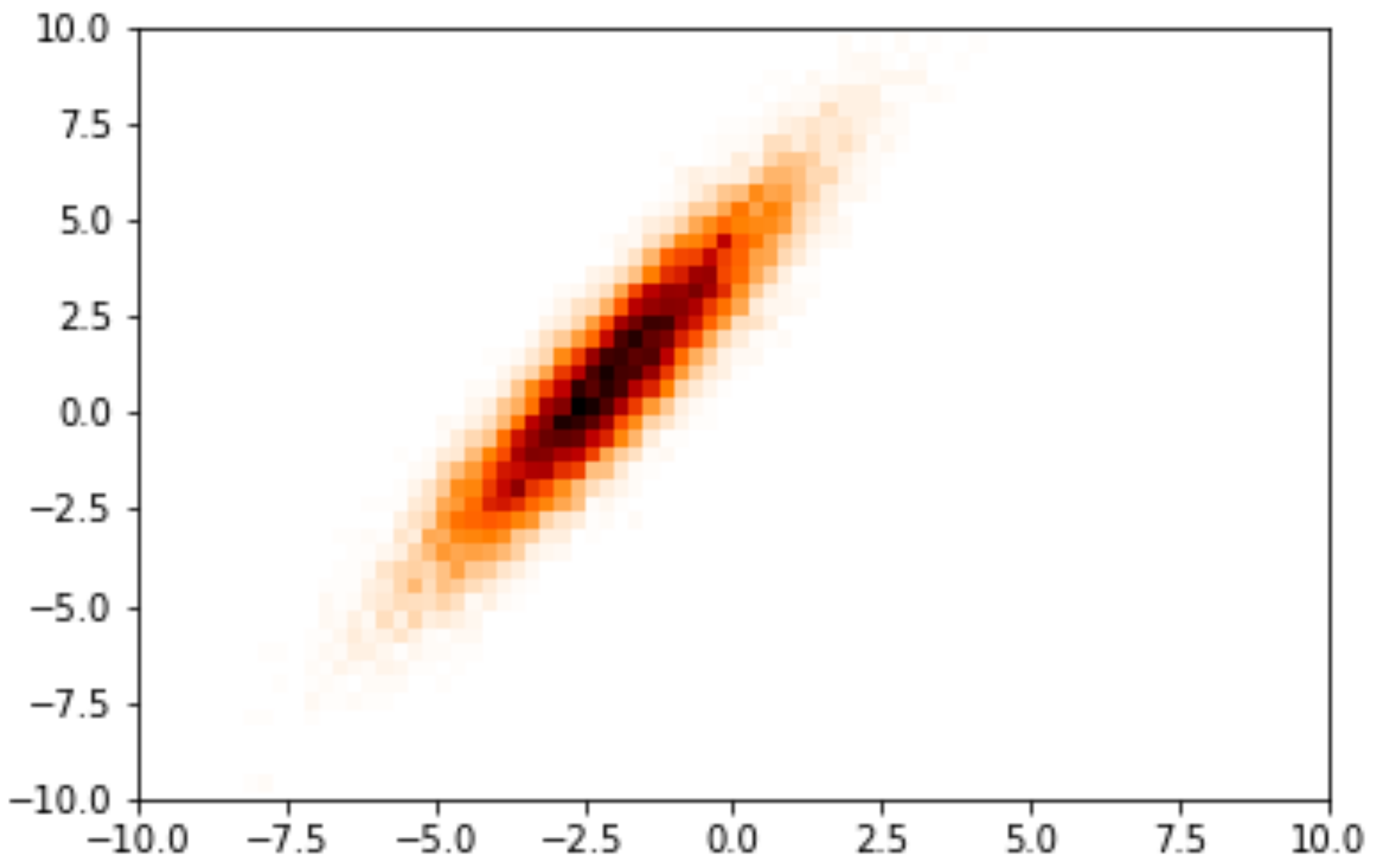
Generative samples



Reconstruction



Experiment: non-isotropic Gaussian



$$Z \sim N(0, I_2)$$

$$\epsilon \sim N(0, {\sigma_\epsilon}^2 \cdot I_2)$$

$$X|Z = \begin{bmatrix} cos(\pi/3) & \text{-}sin(\pi/3) \\ sin(\pi/3) & cos(\pi/3) \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 \\ 0 & 1/3 \end{bmatrix} \cdot Z + \epsilon$$

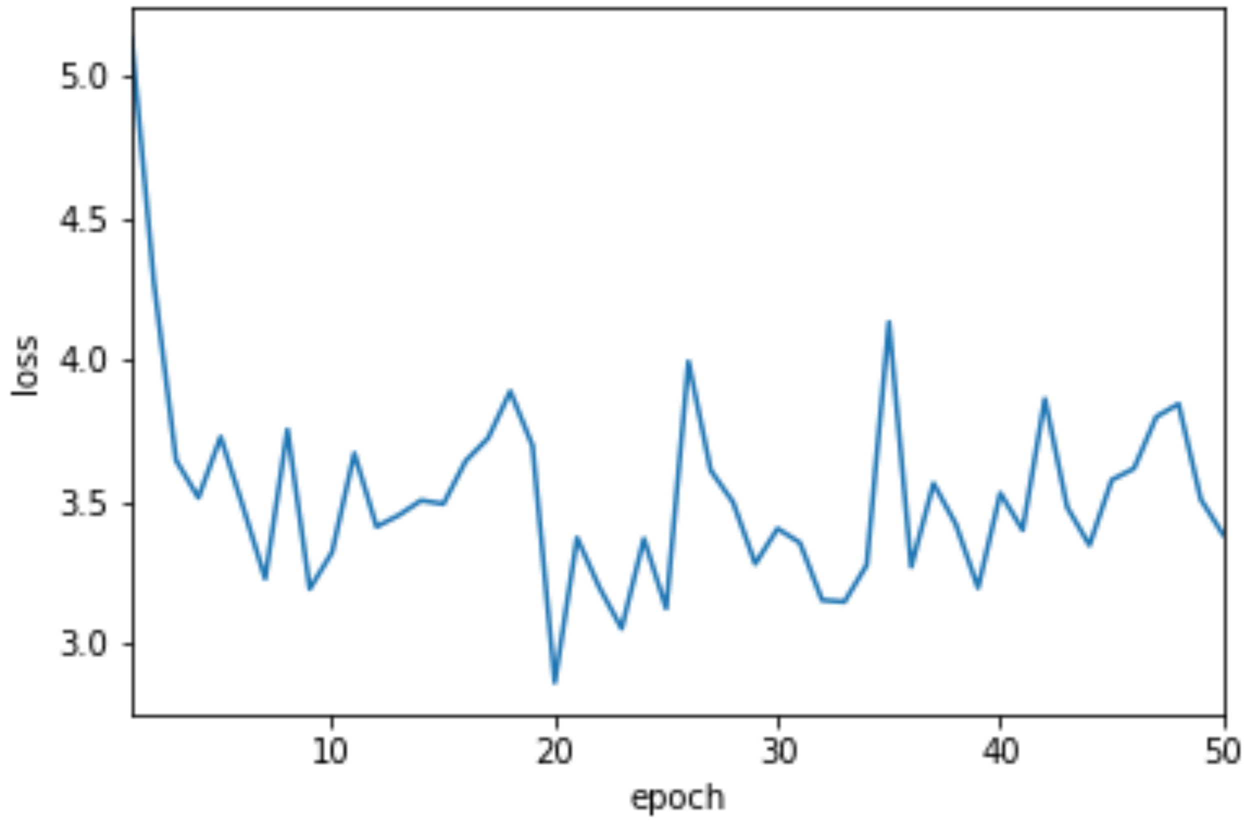
where ${\sigma_\epsilon}^2 = 0.2$

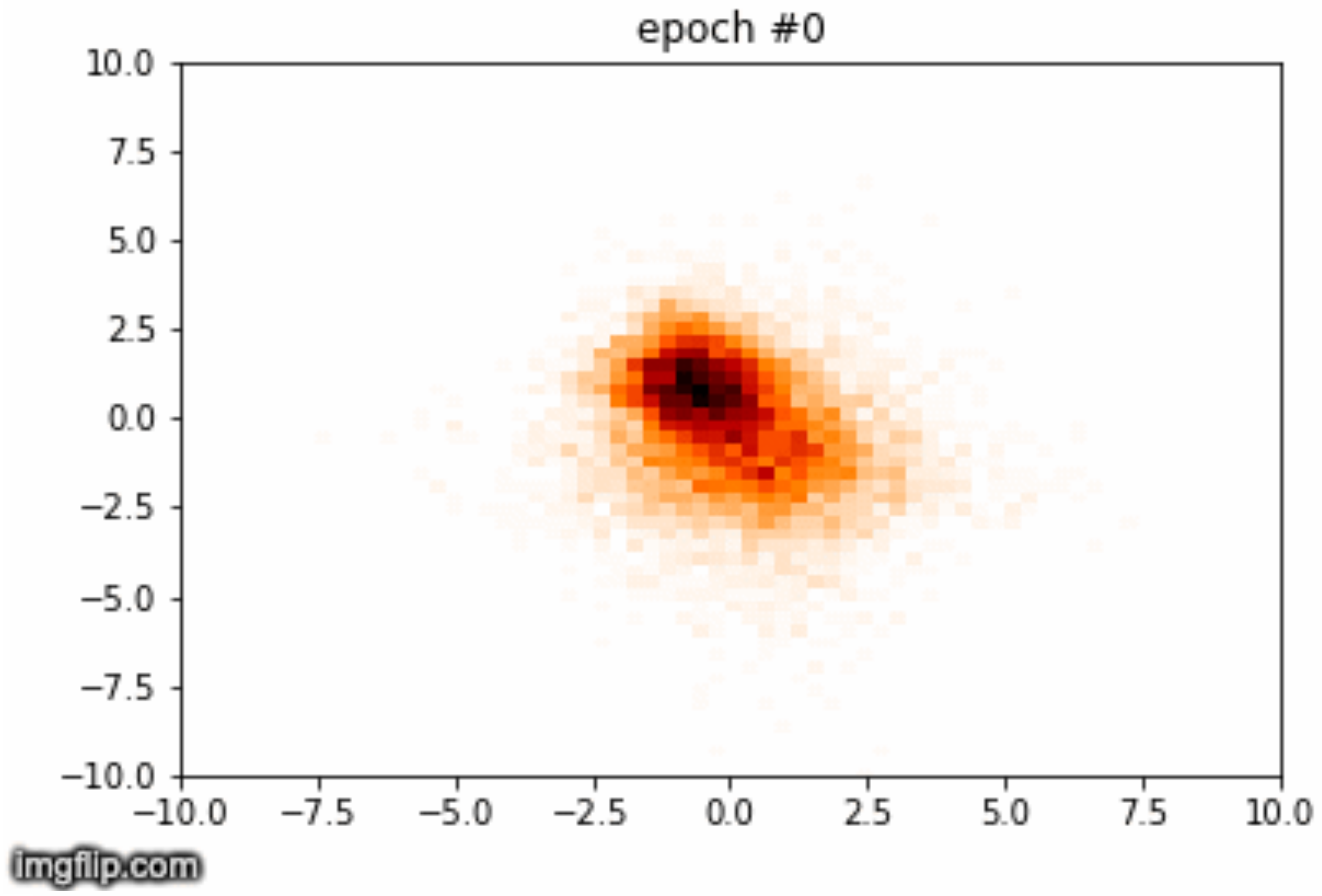
- Diagonal/Full covariance model:
  - # of training samples: 20,000
  - latent dimension: 2
  - batch_size: 32
  - learning rate: 1e-3
  - # epoch: 50

- Testing:
  - # generative samples: 10,000
  - # reconstruction of test set: 10,000

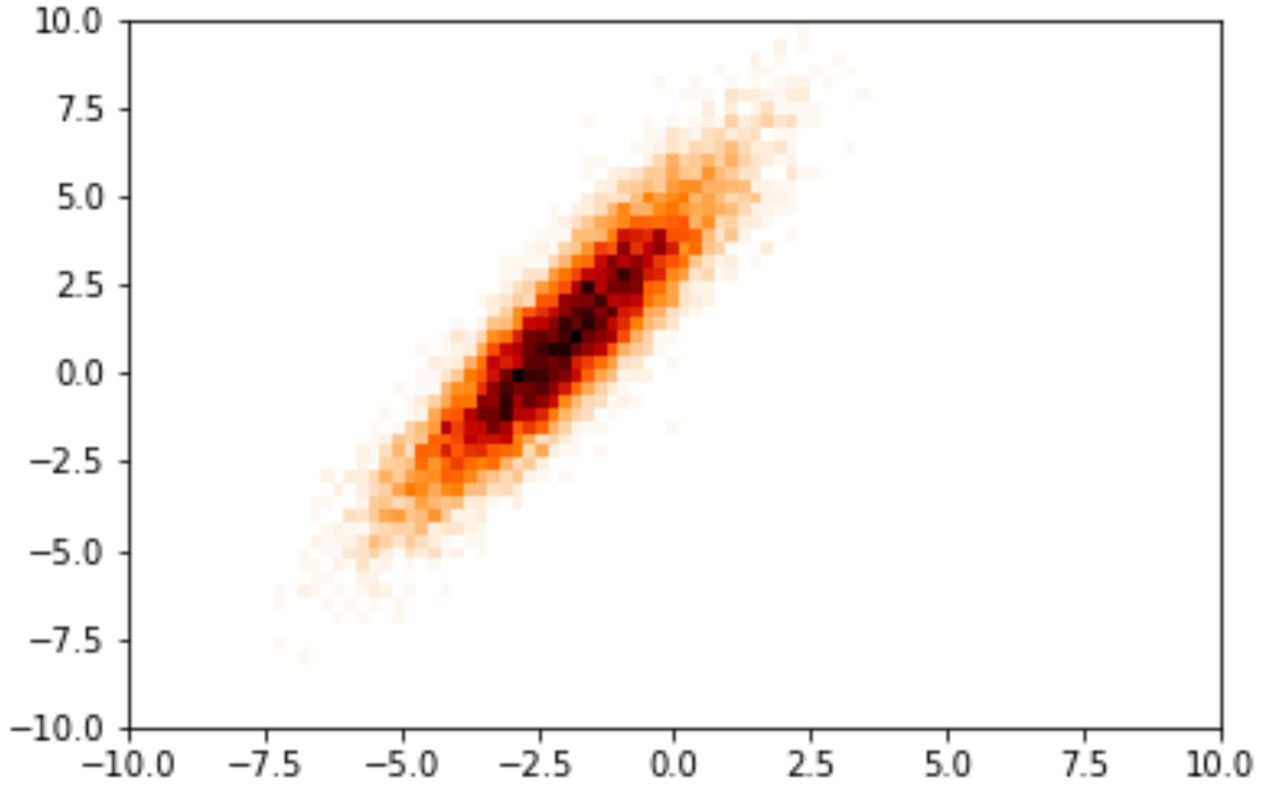
Experiment: non-isotropic Gaussian

Diagonal:

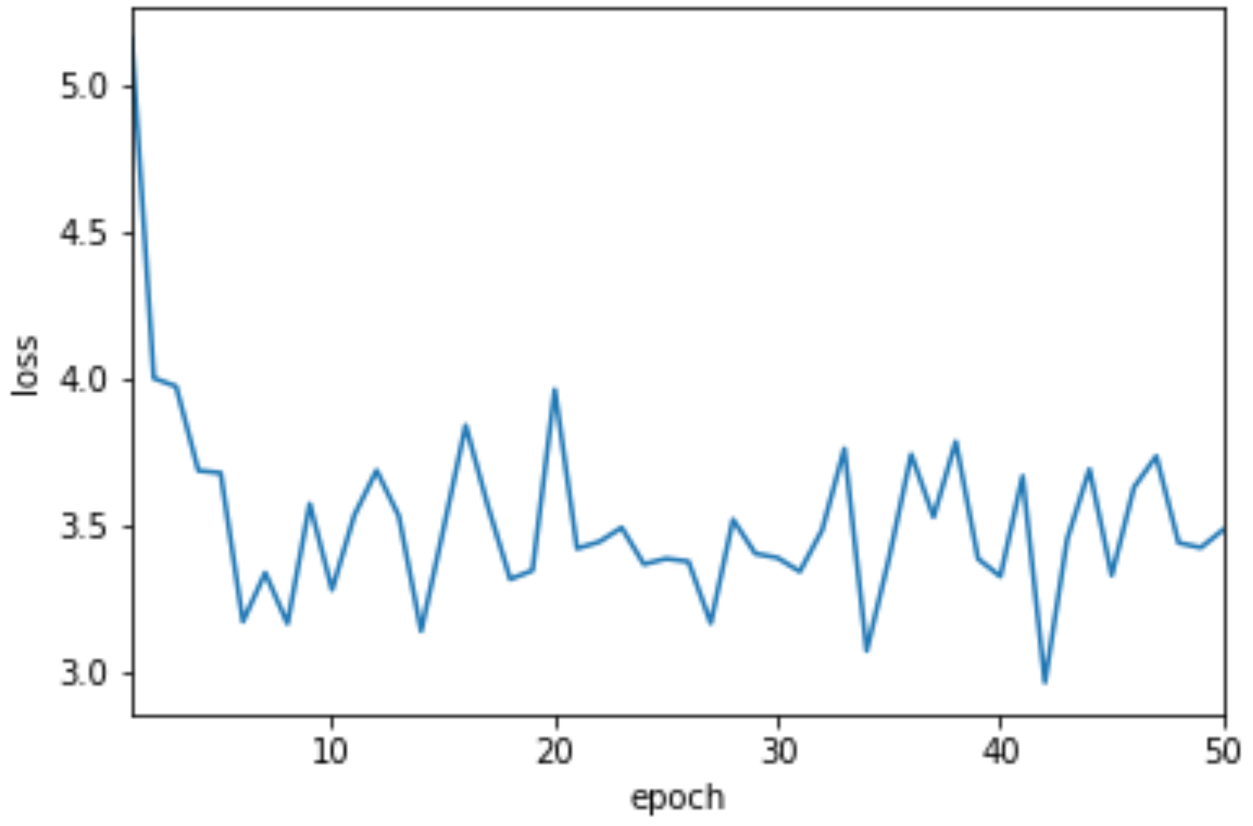


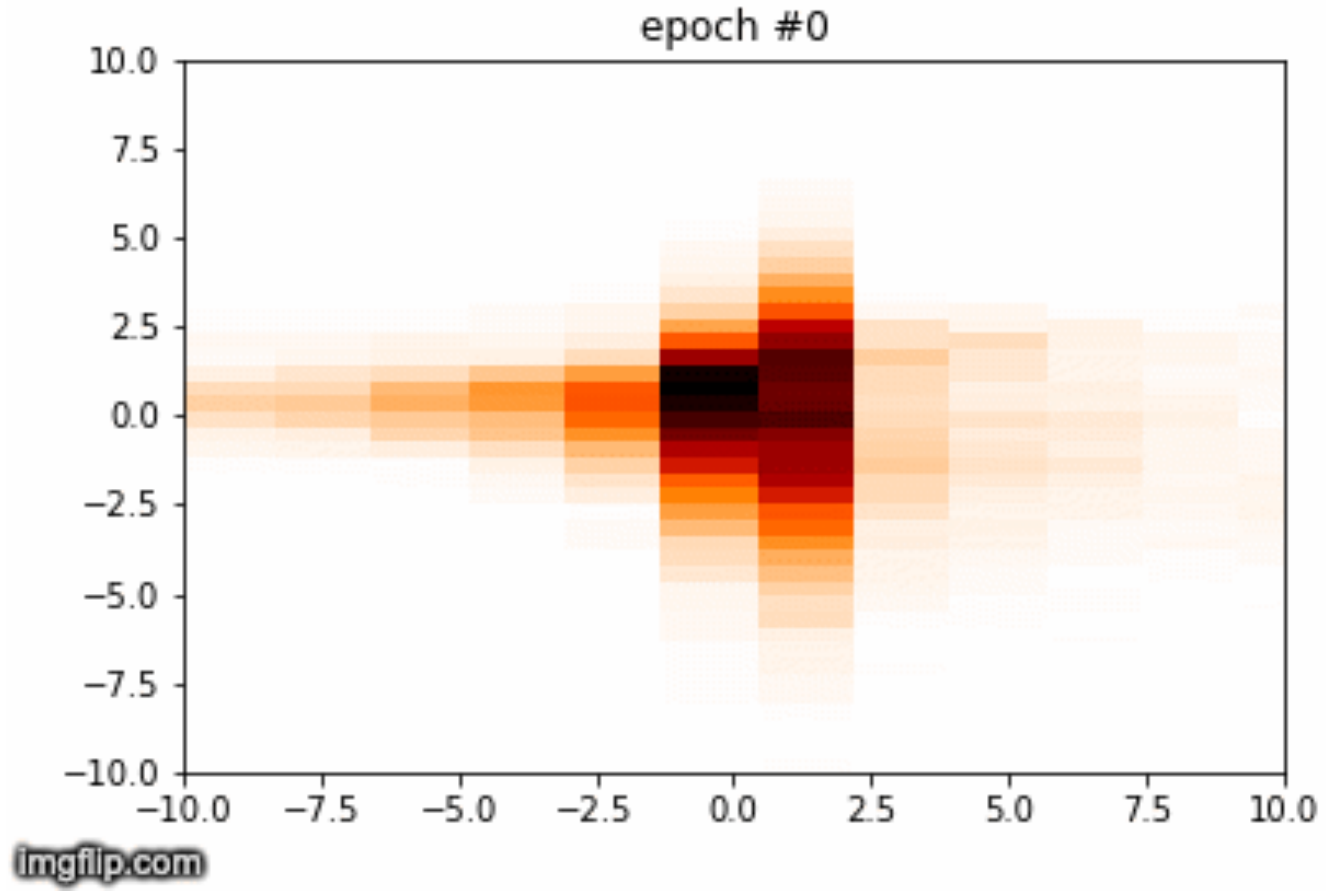


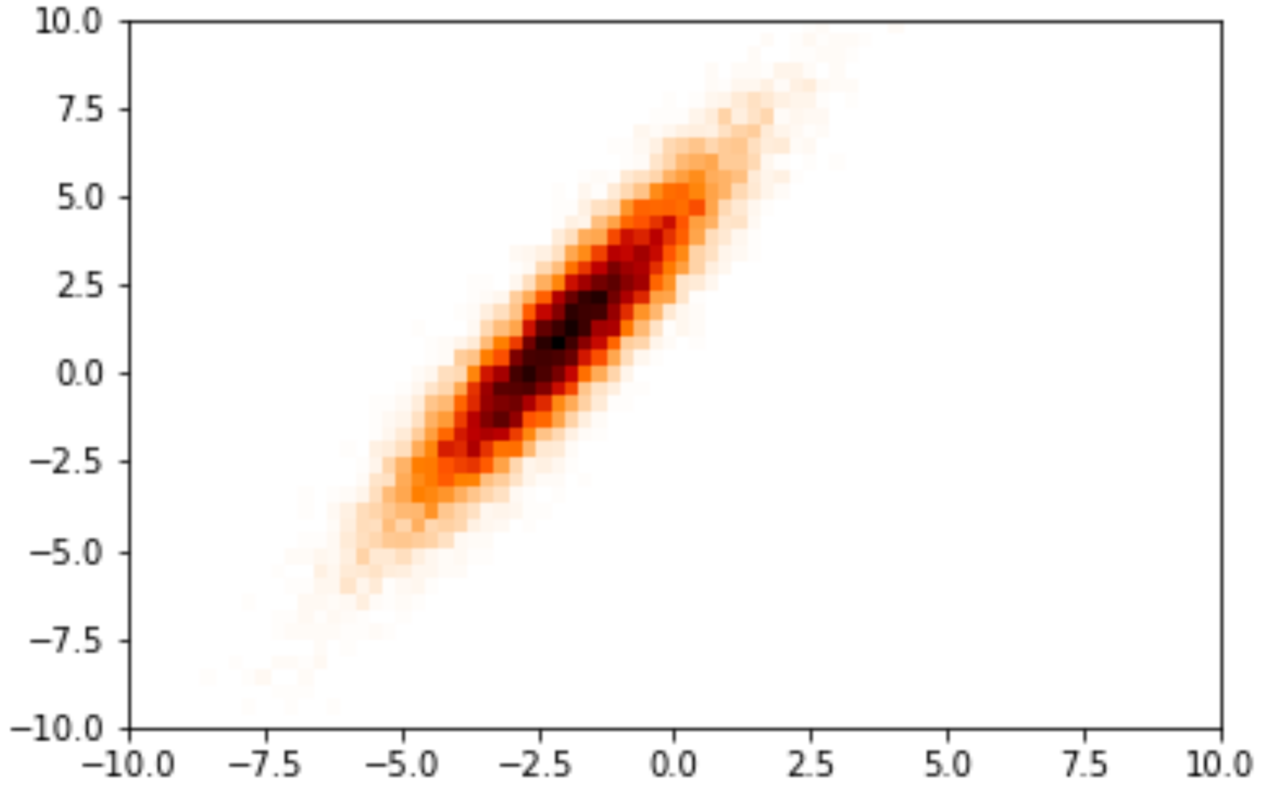
After training



Full:

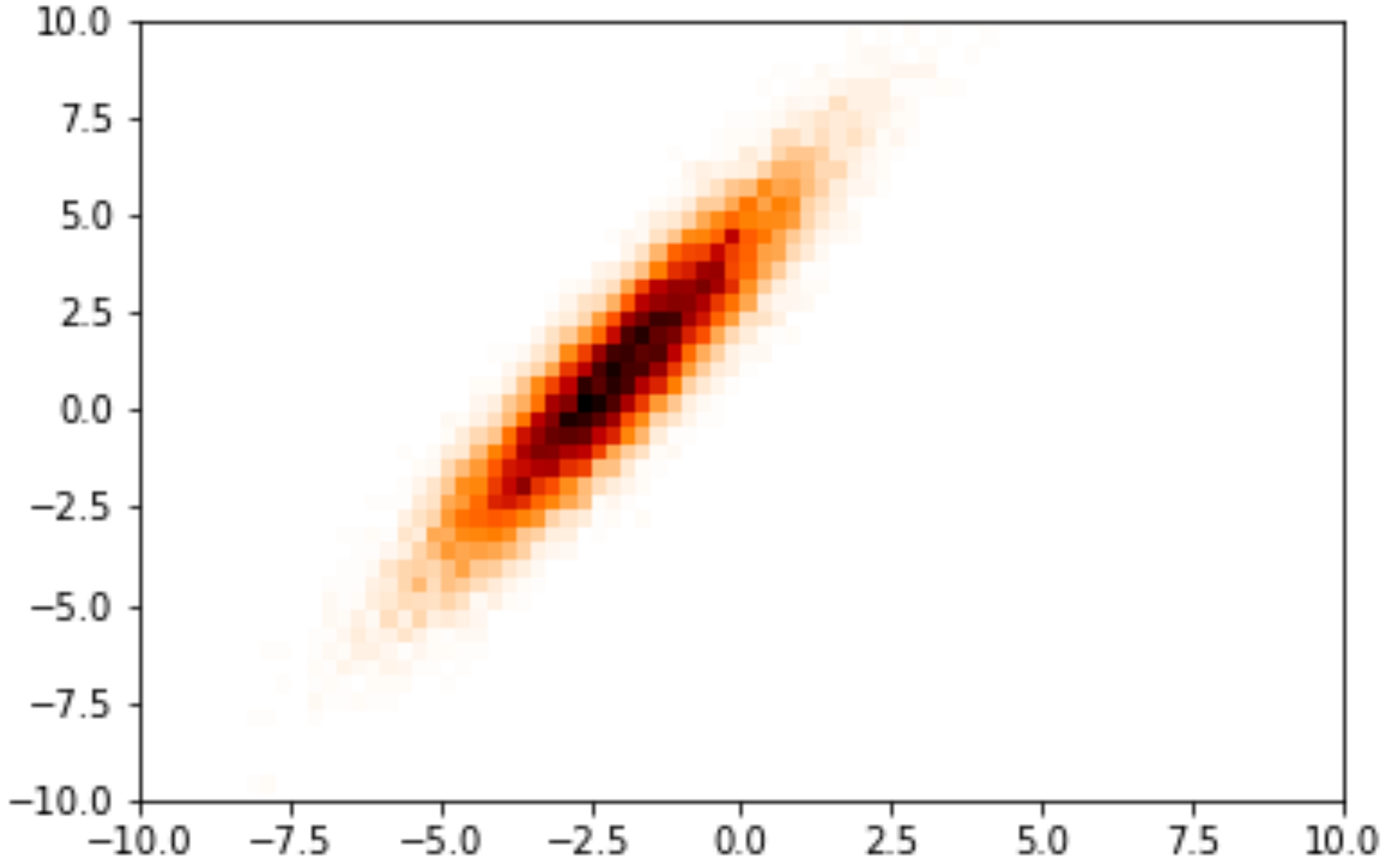






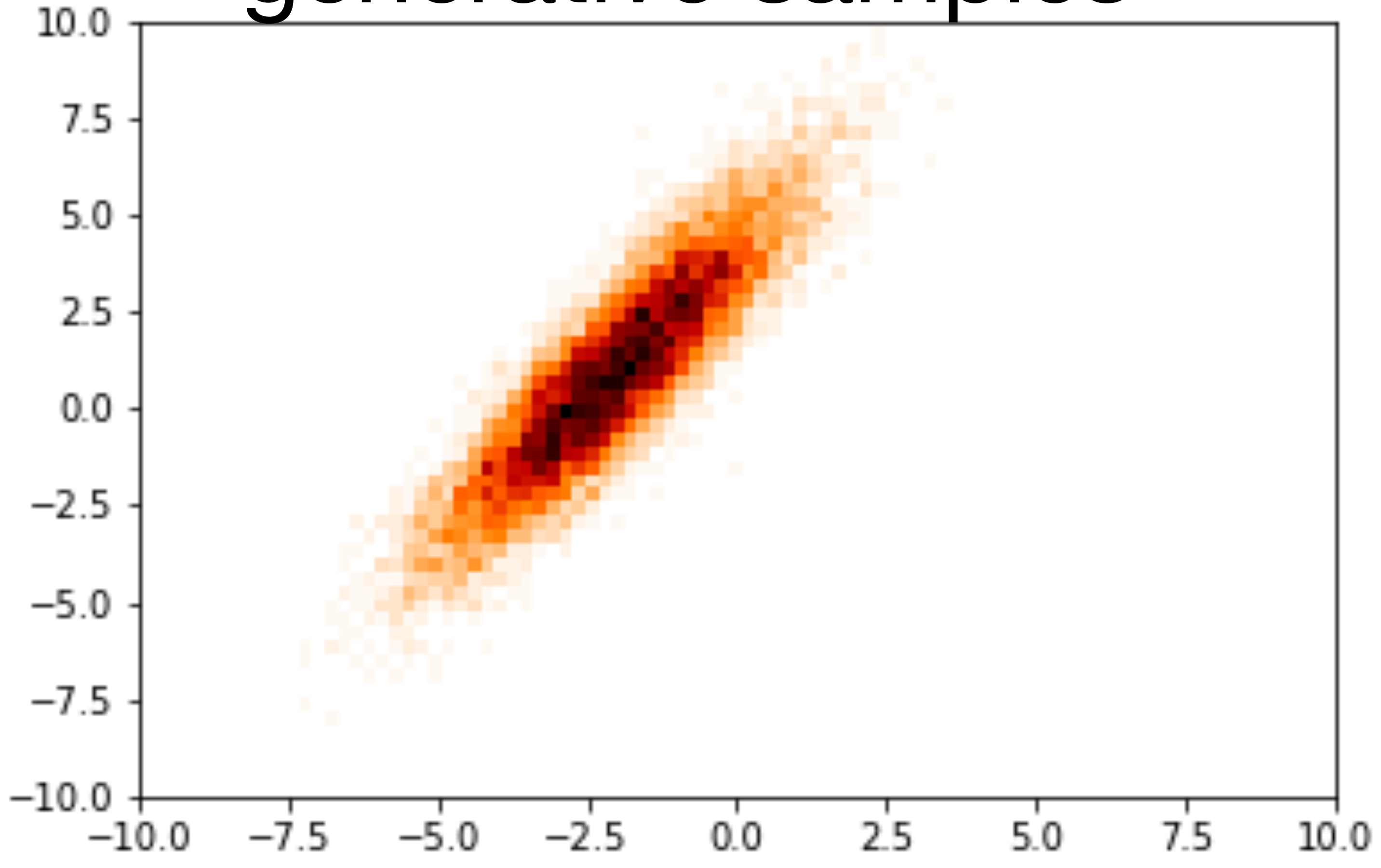
Experiment: non-isotropic Gaussian

$X|Z$

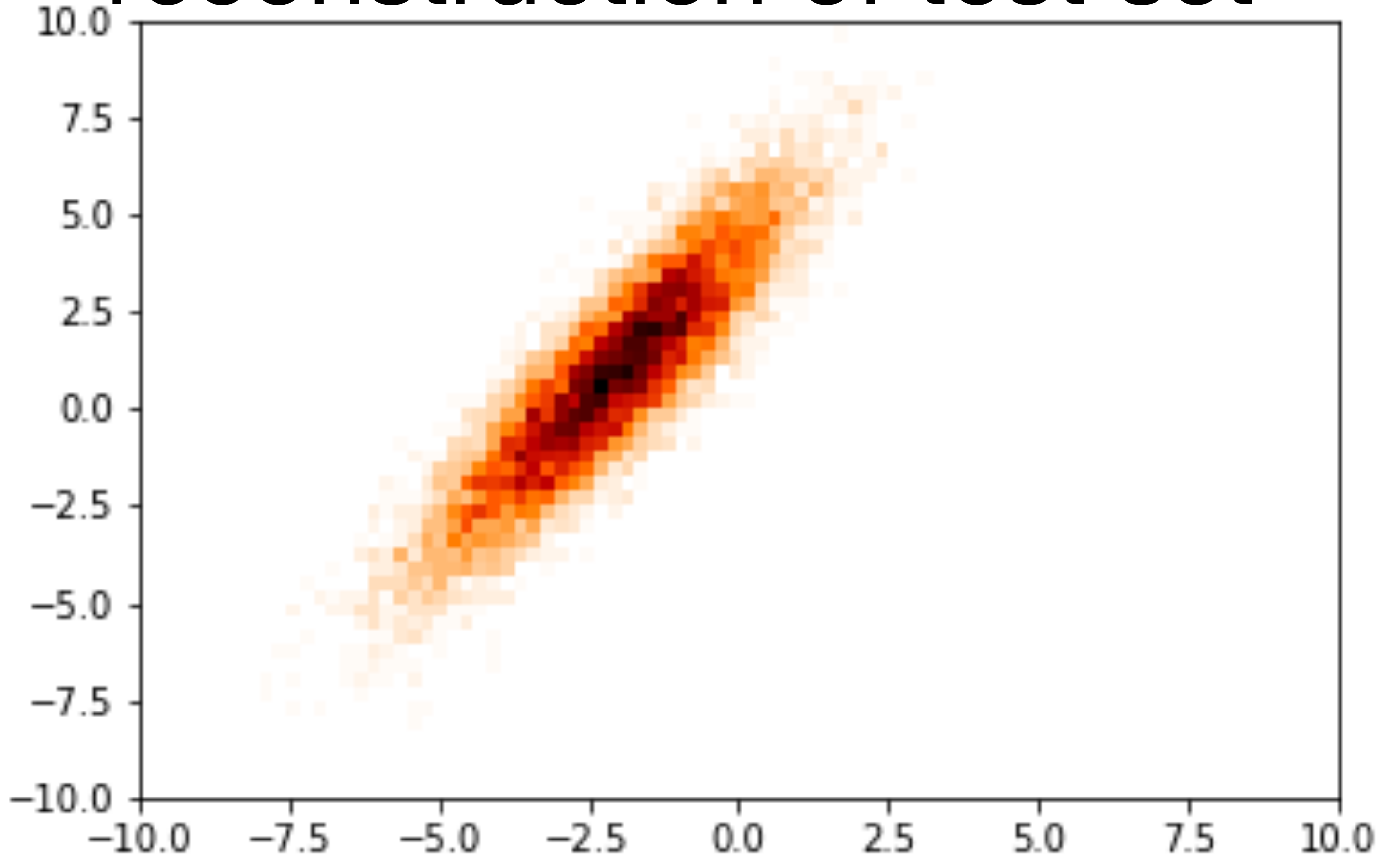


Diagonal:

generative samples

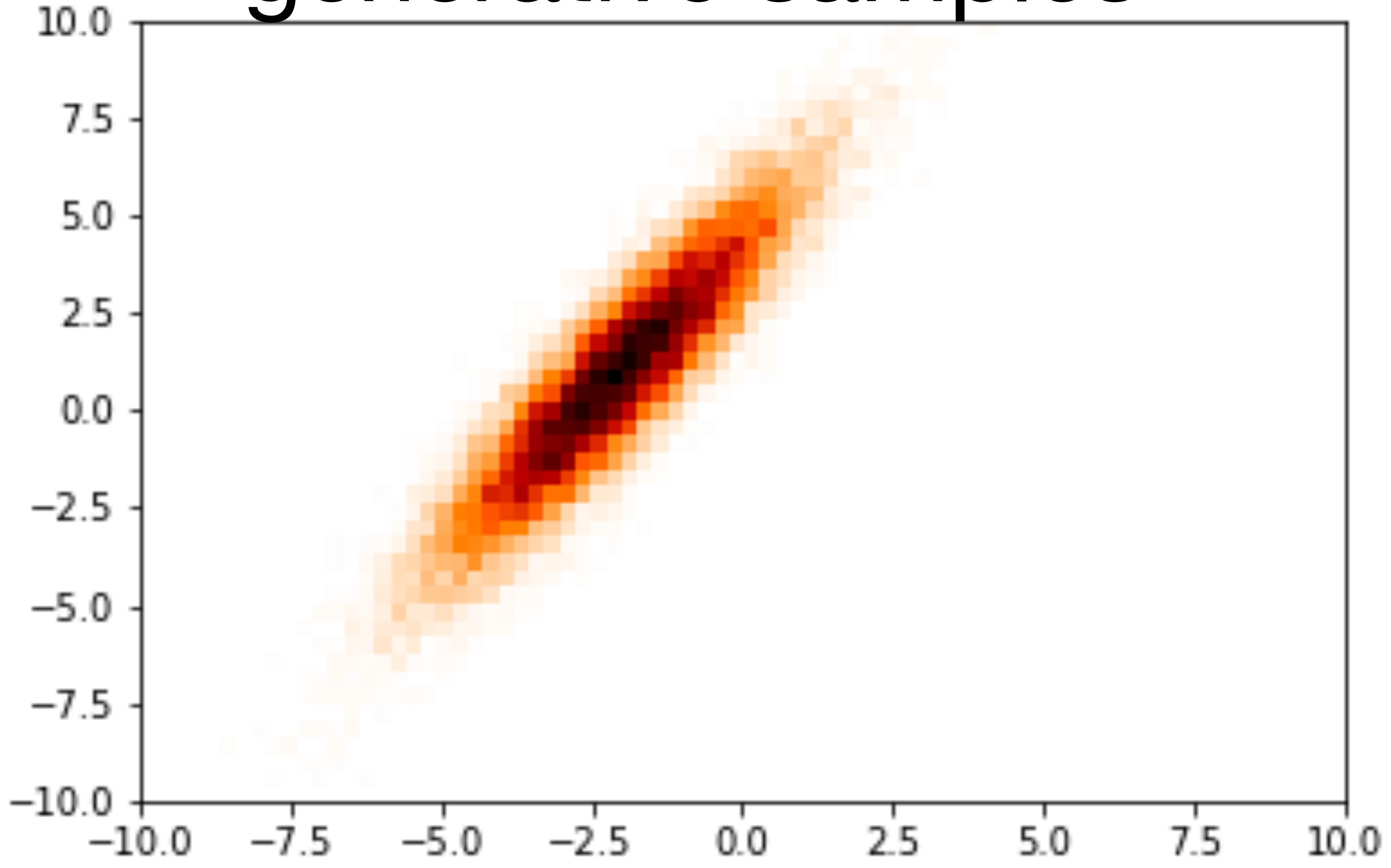


reconstruction of test set

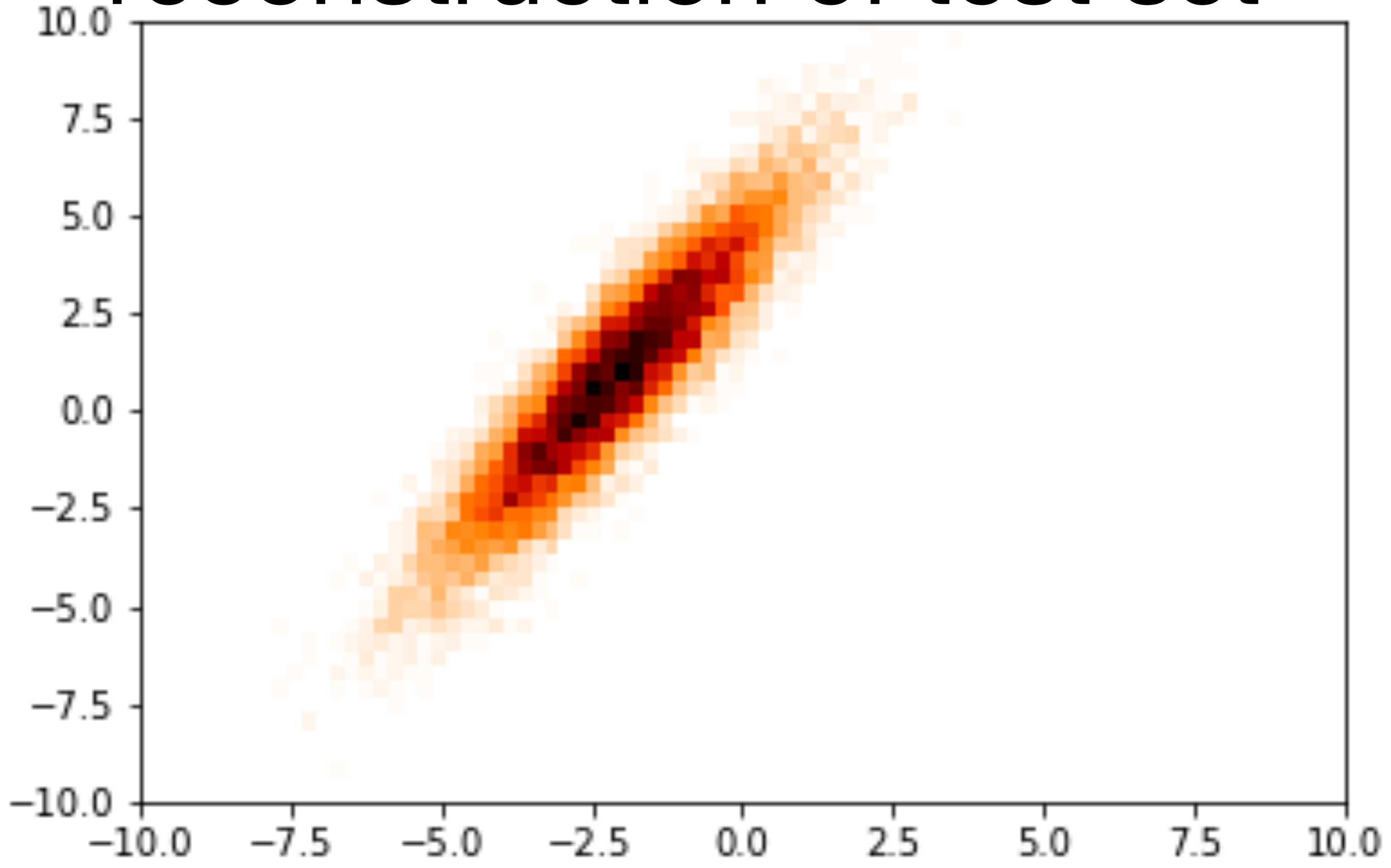


Full:

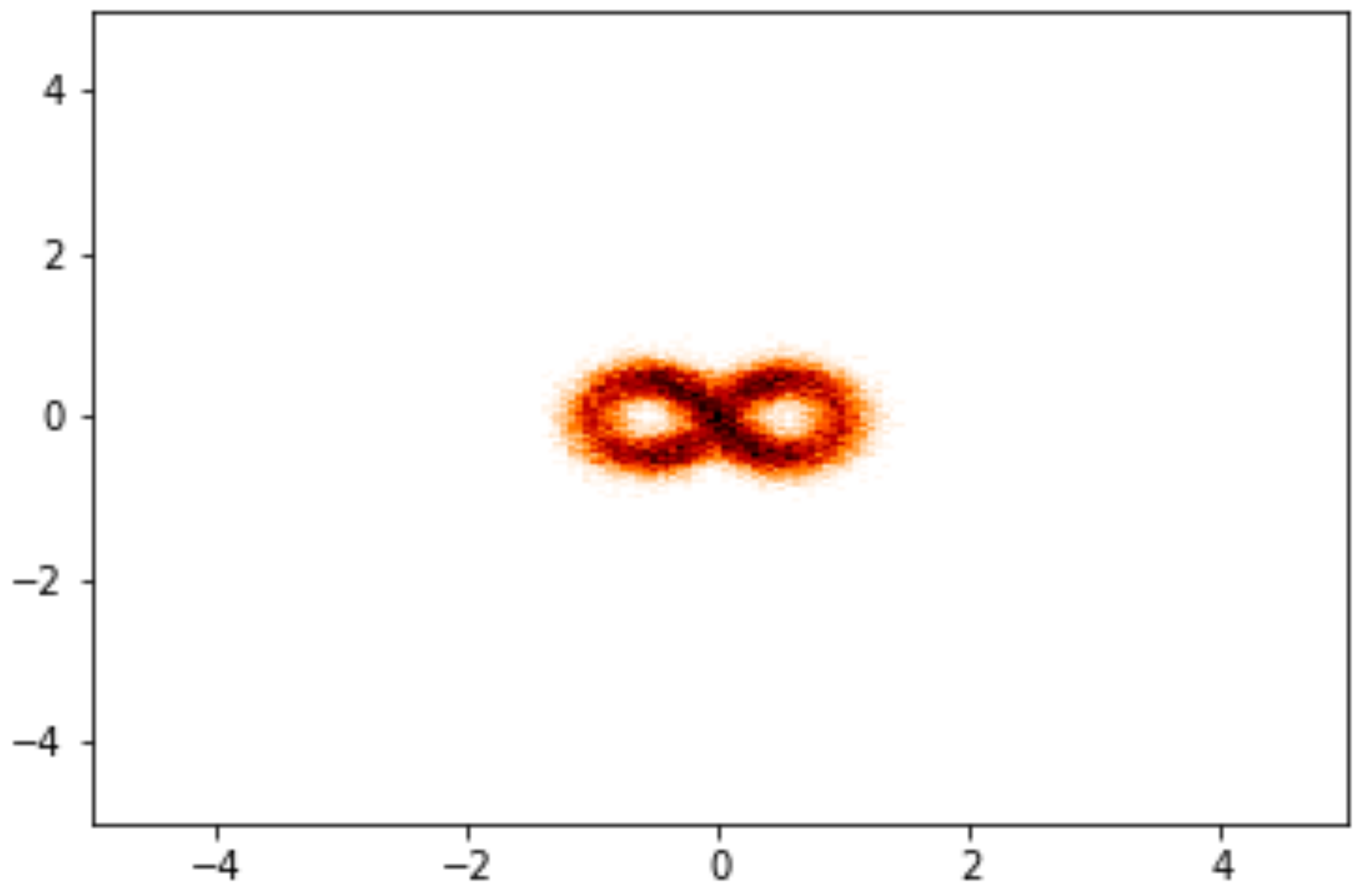
generative samples



reconstruction of test set



Experiment: figure 8



$$Z \sim N(0,1)$$

$$\epsilon \sim N(0, {\sigma_\epsilon}^2 \cdot I_2)$$

$$u(z) = \big(0.6 + 1.8 \cdot \Phi(z)\big)$$

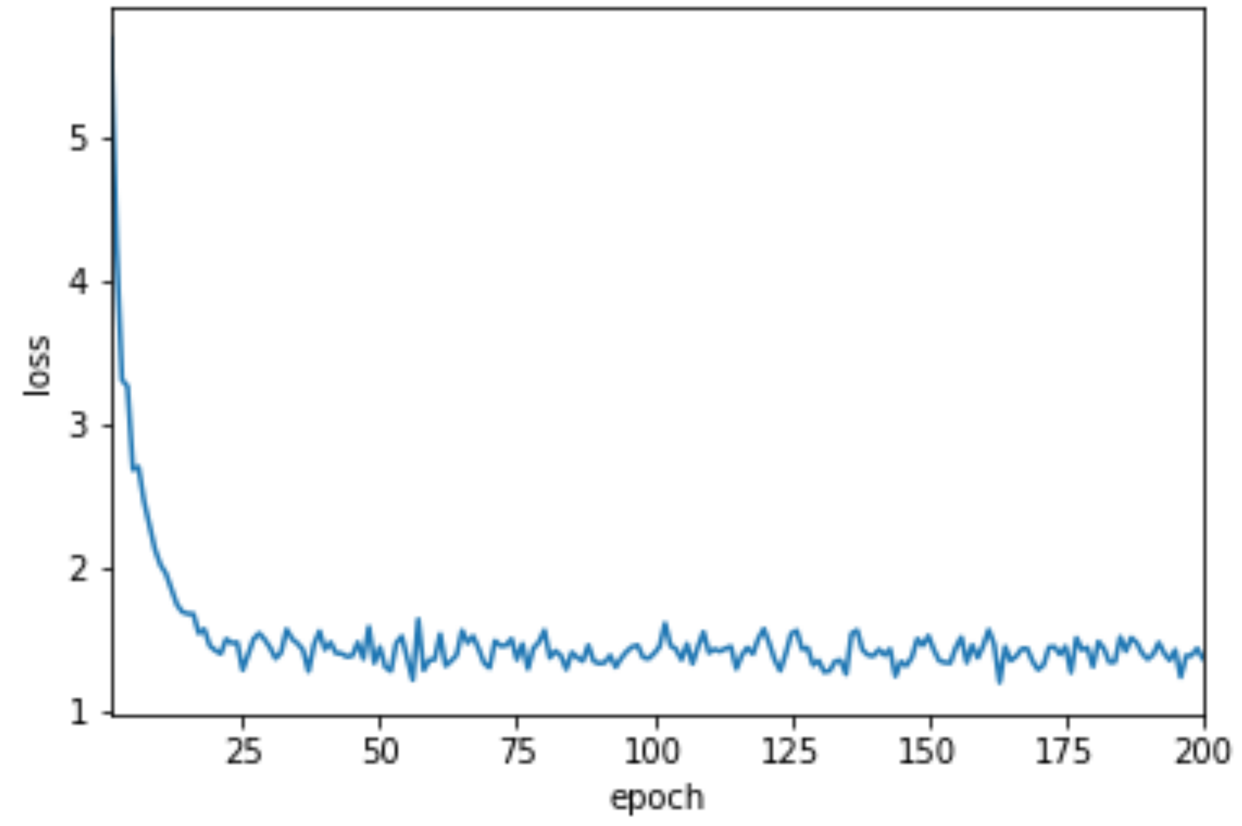
$$x|z = \begin{bmatrix} \frac{\sqrt{2}}{2} \cdot \frac{cos\big(u(z)\big)}{sin\big(u(z)\big)^2 + 1} \\ \sqrt{2} \cdot \frac{cos\big(u(z)\big) sin\big(u(z)\big)}{sin\big(u(z)\big)^2 + 1} \end{bmatrix} + \epsilon$$

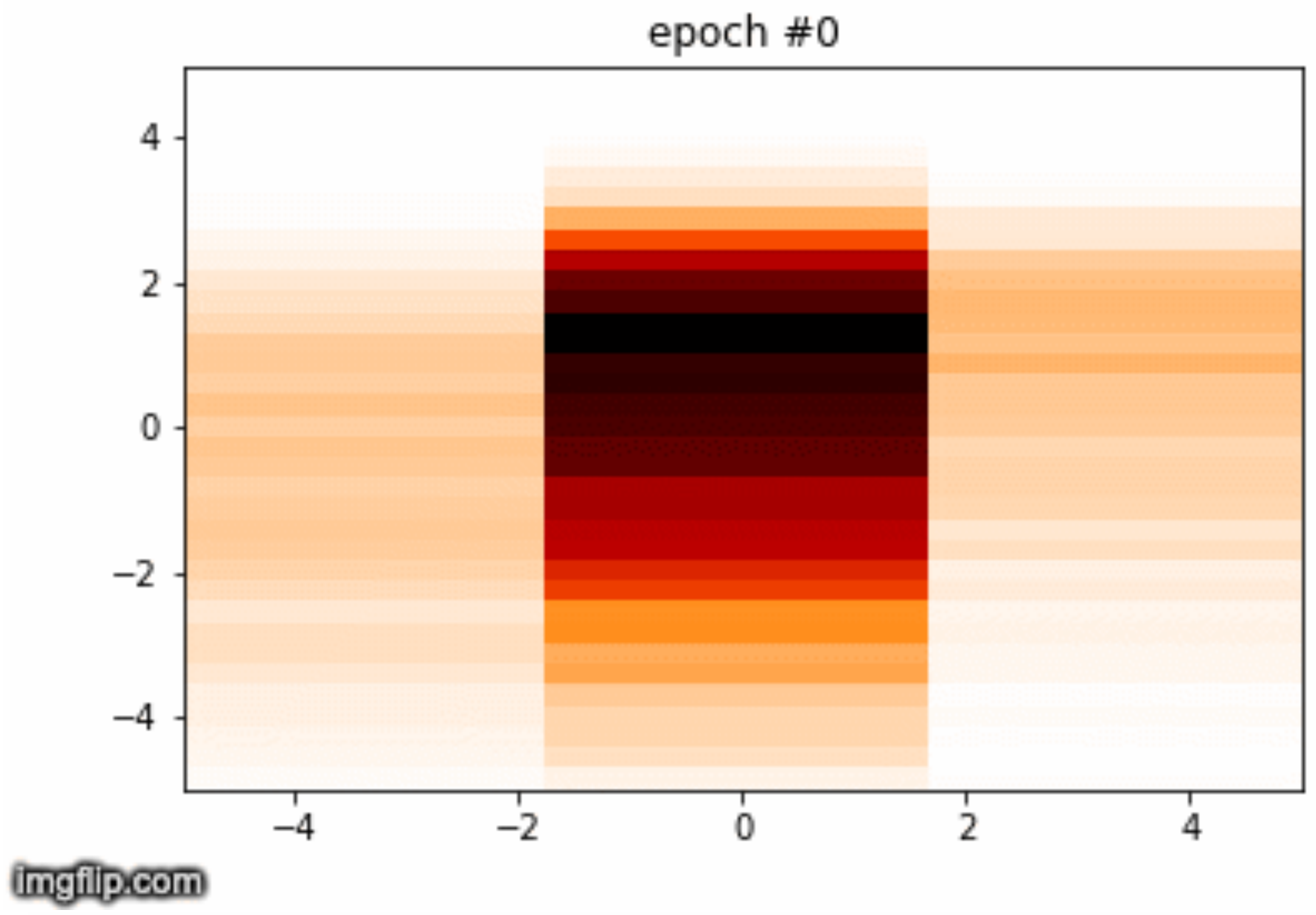
where $\Phi(z)$ is Gaussian CDF, and ${\sigma_\epsilon}^2 = 0.02$

- Diagonal/Full covariance model:
  - # of training samples: 40,000
  - latent dimension: 2
  - batch_size: 32
  - learning rate: 1e-3
  - # epoch: 200
- Testing:
  - # generative samples: 20,000
  - # reconstruction of test set: 20,000

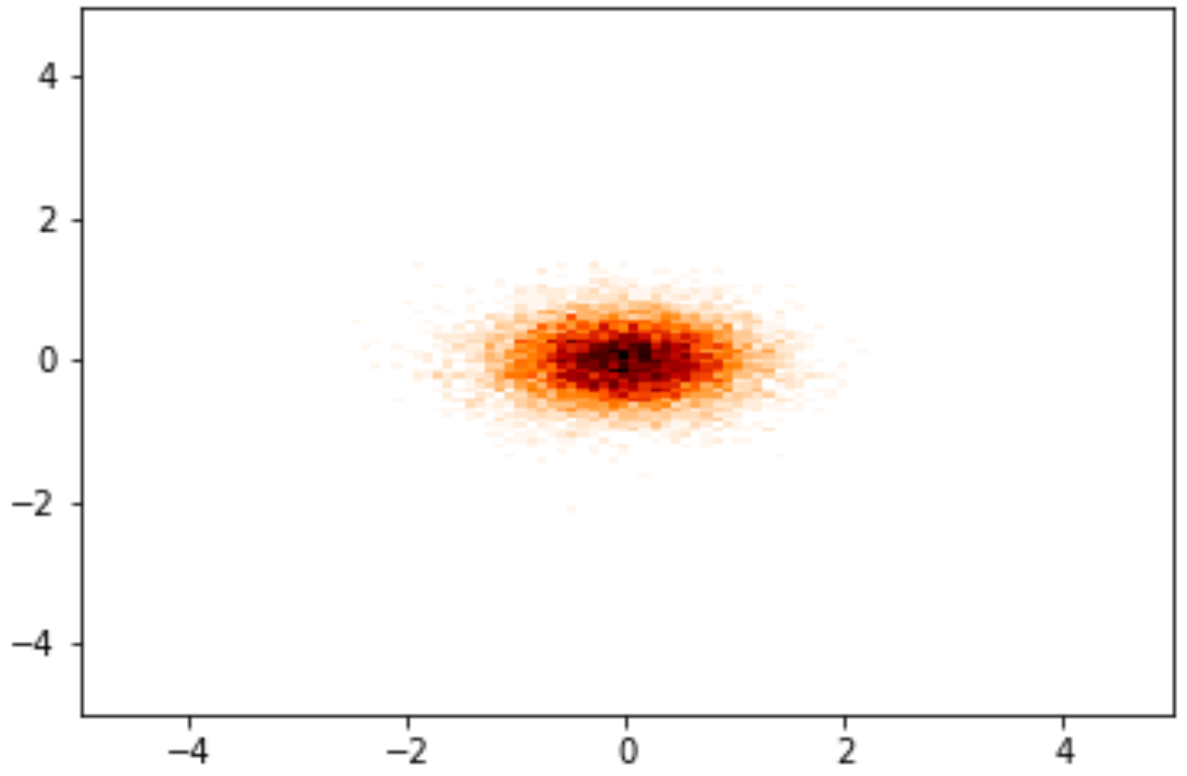
Experiment: figure 8

Diagonal:

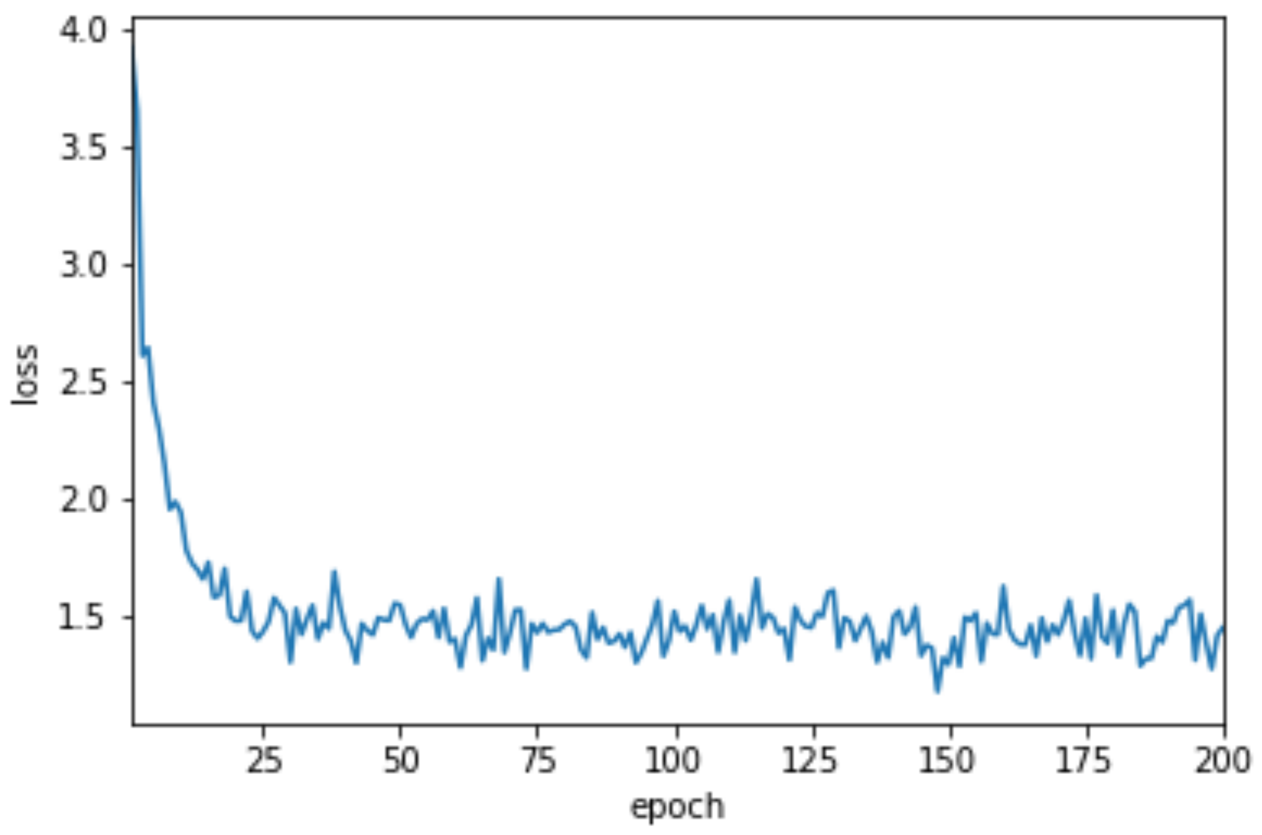


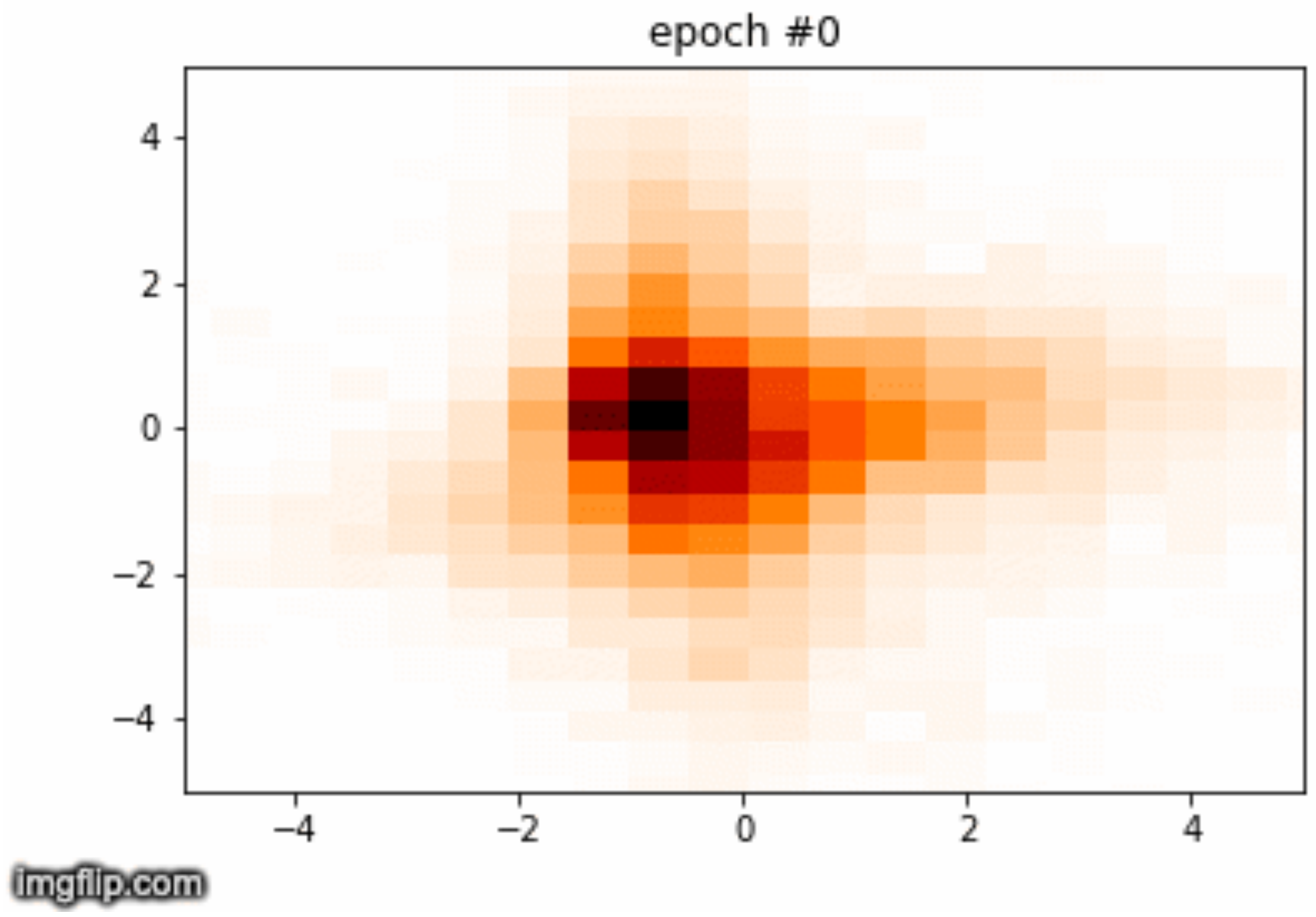


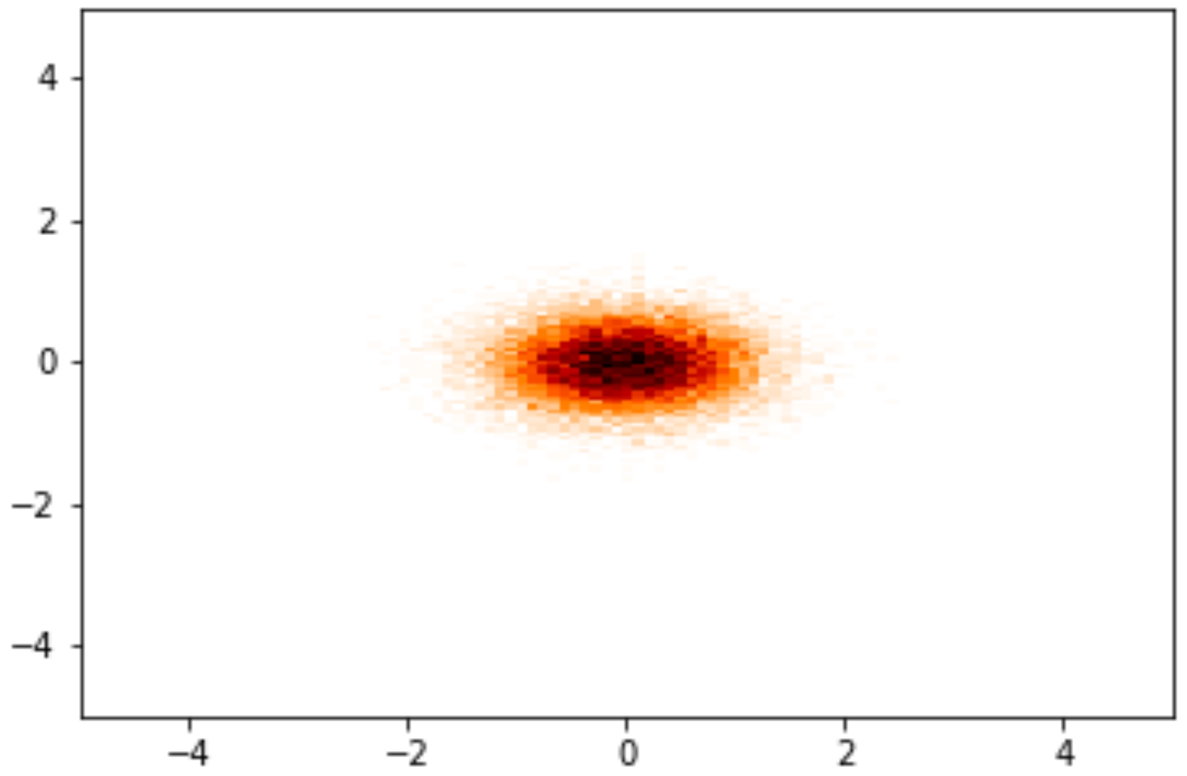
After training



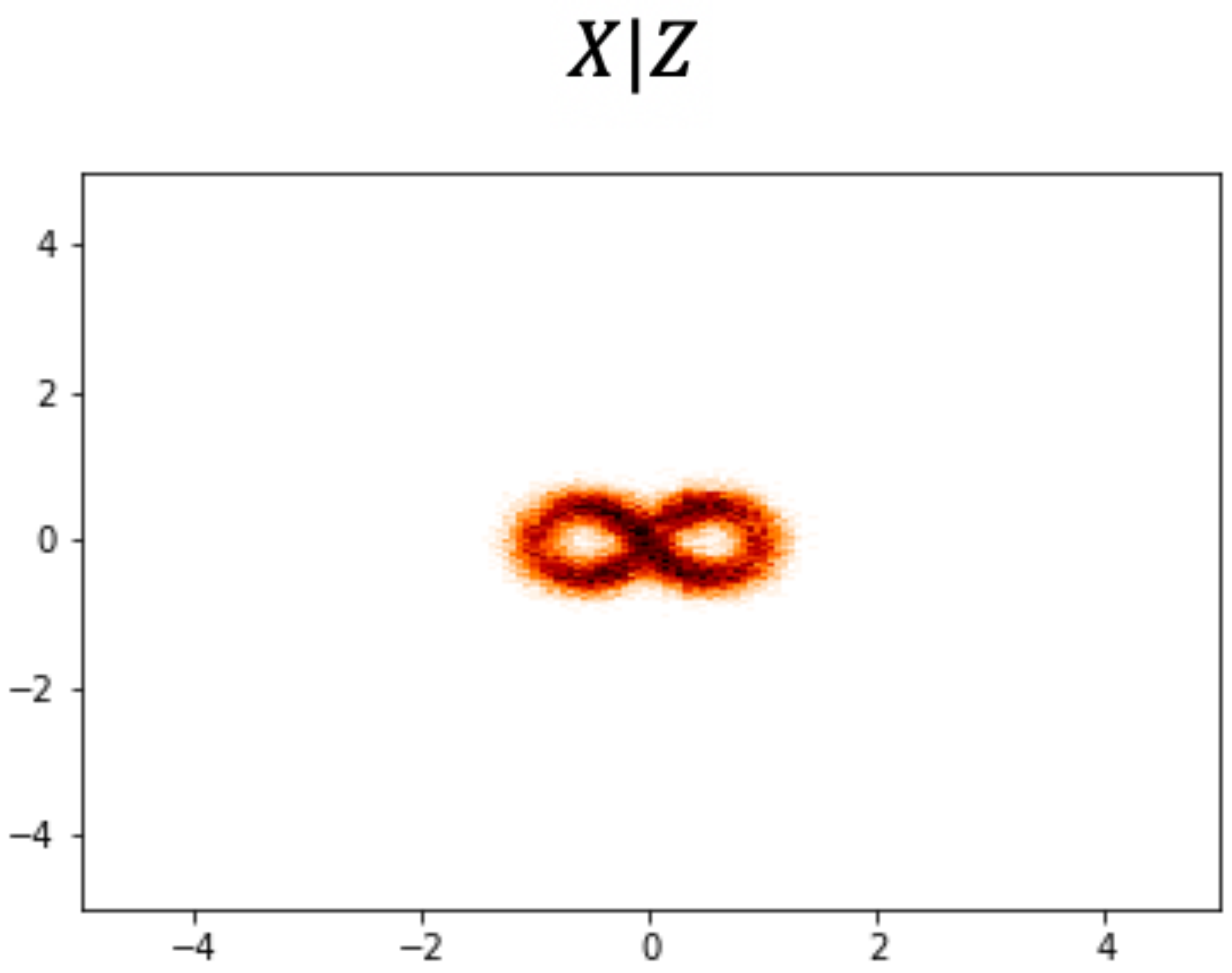
Full:





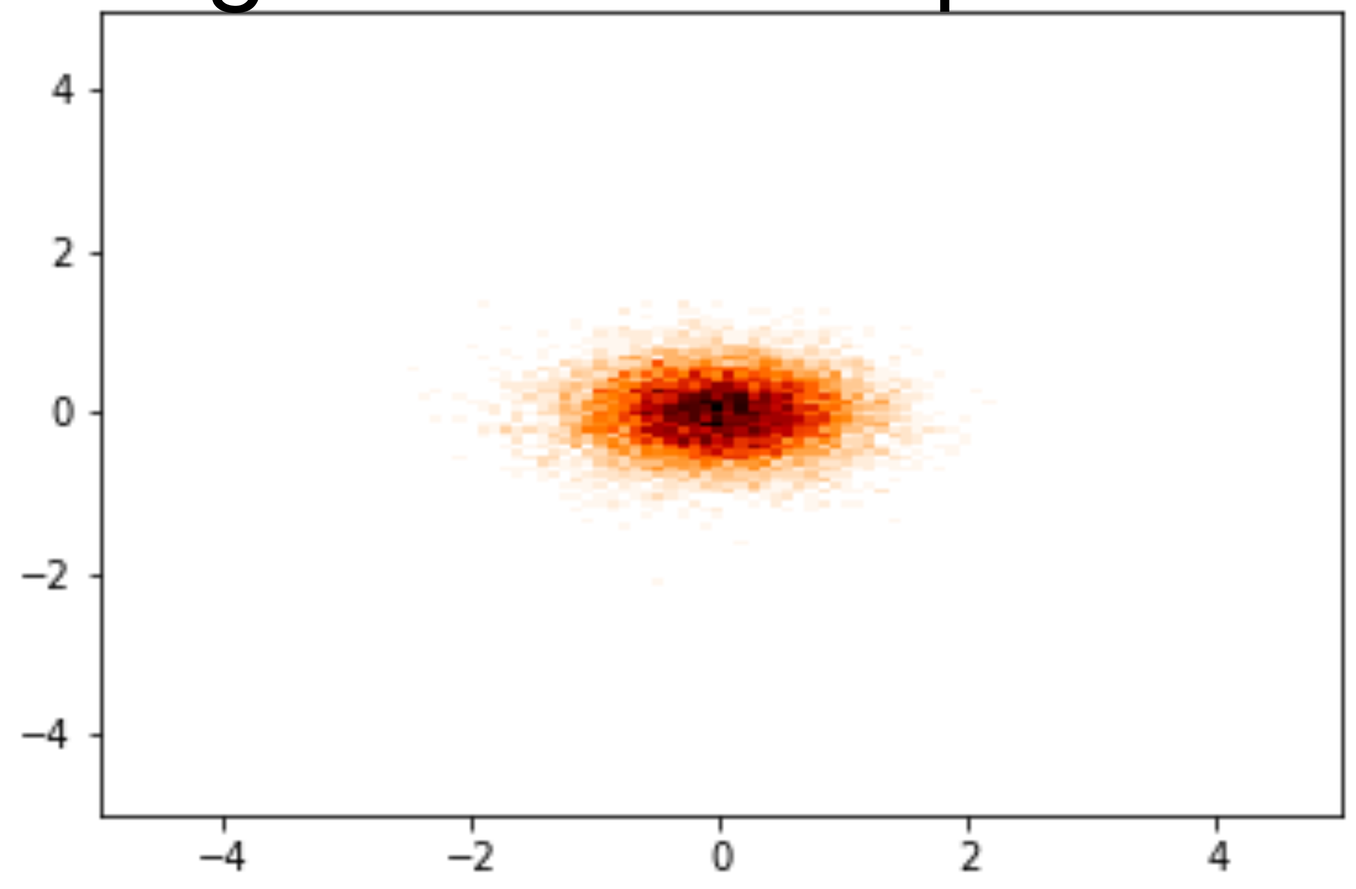


Experiment: figure 8

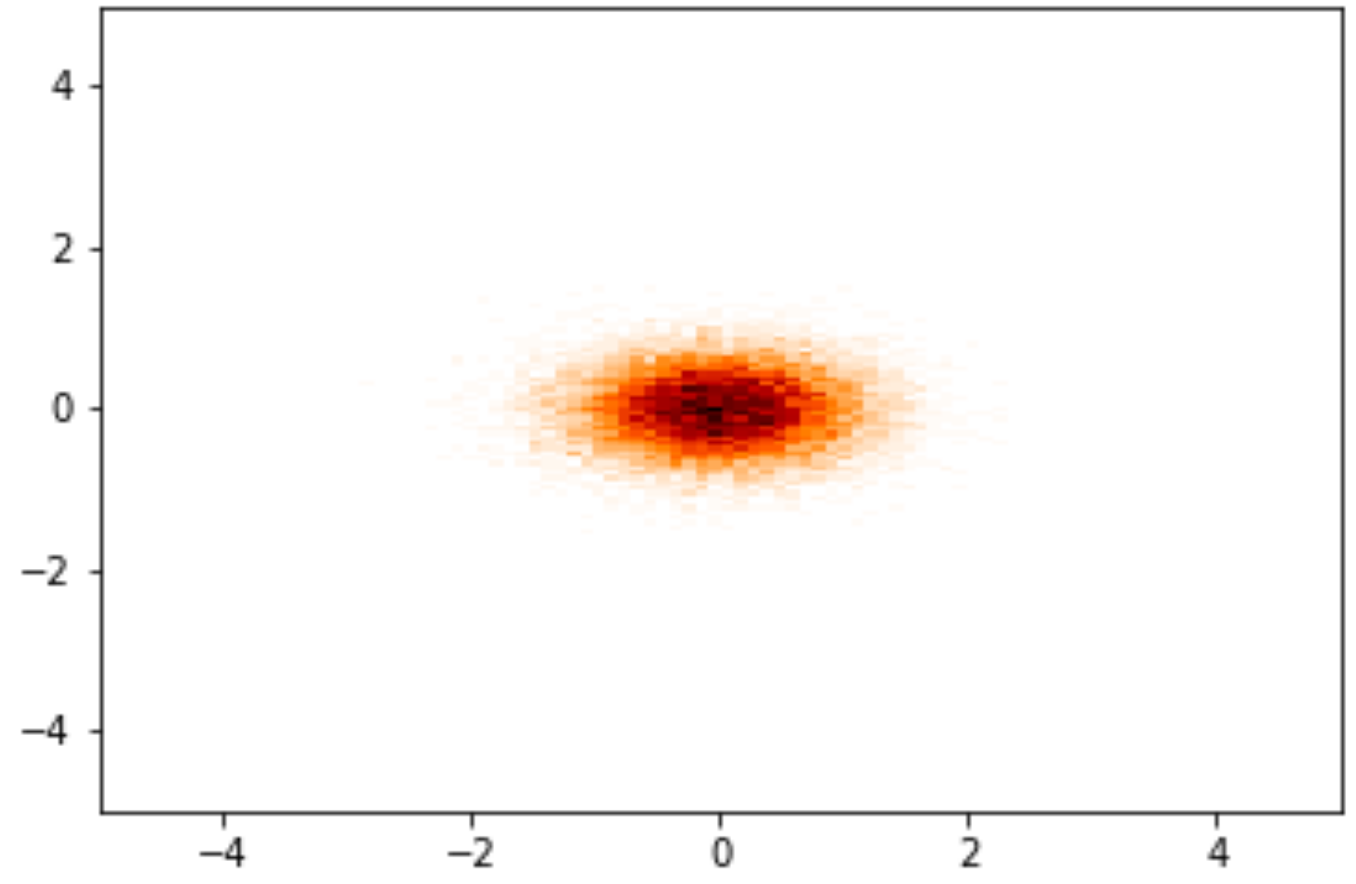


Diagonal:

generative samples

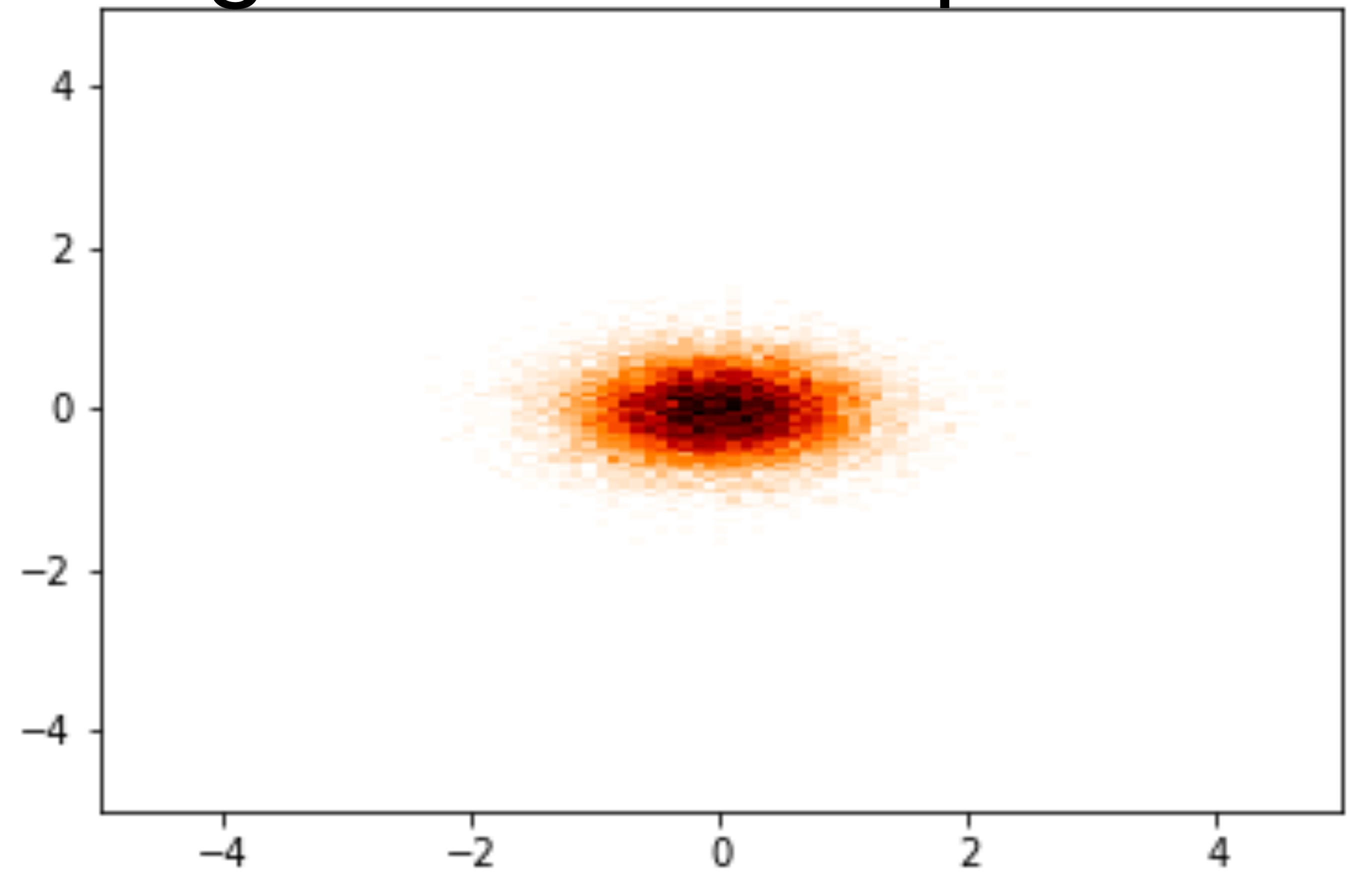


reconstruction of test set

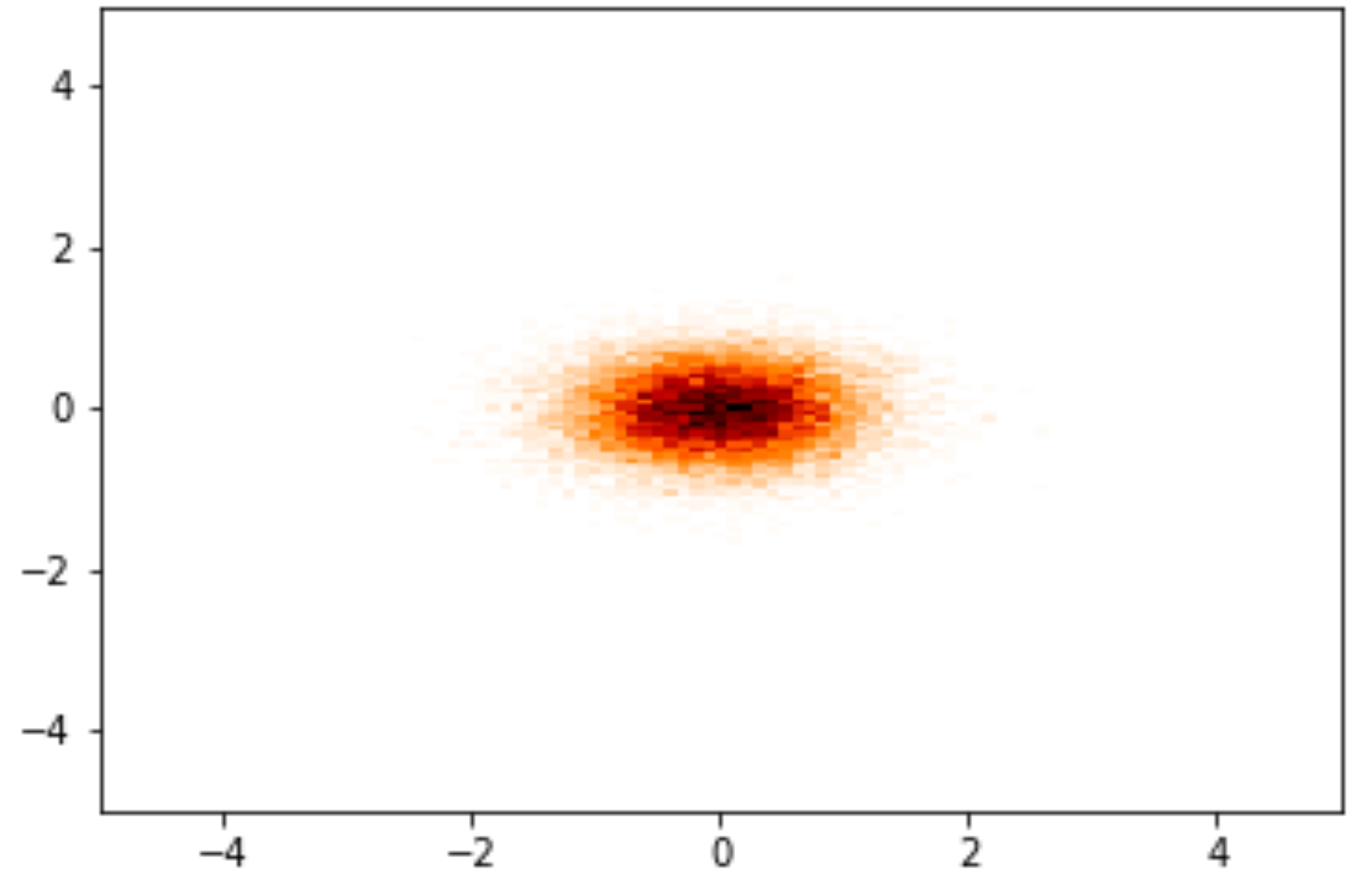


Full:

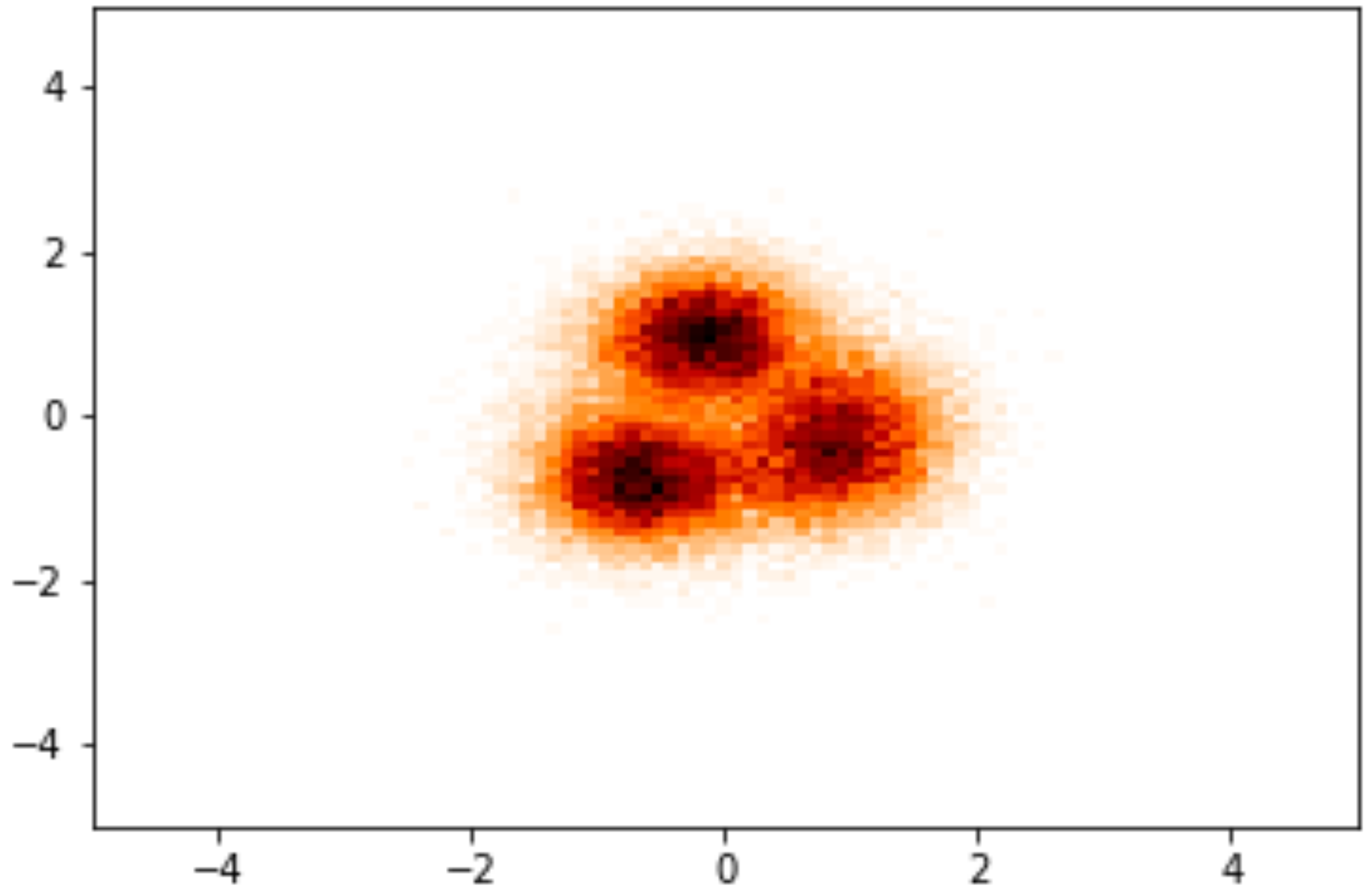
generative samples



reconstruction of test set



Experiment: three blobs



$$Z \sim N(0,1)$$

$$\epsilon \sim N(0, {\sigma_\epsilon}^2 \cdot I_2)$$

$$u(z) = \frac{2\pi}{1 + e^{-\frac{1}{2}\pi z}}$$

$$t(u) = 2 \cdot tanh(10 \cdot u - 20 \cdot \lfloor u/2 \rfloor - 10) + 4 \cdot \lfloor u/2 \rfloor + 2$$

$$x|z = \begin{bmatrix} cos\left(t\left(u(z)\right)\right) \\ \sin\left(t\left(u(z)\right)\right) \end{bmatrix} + \epsilon$$

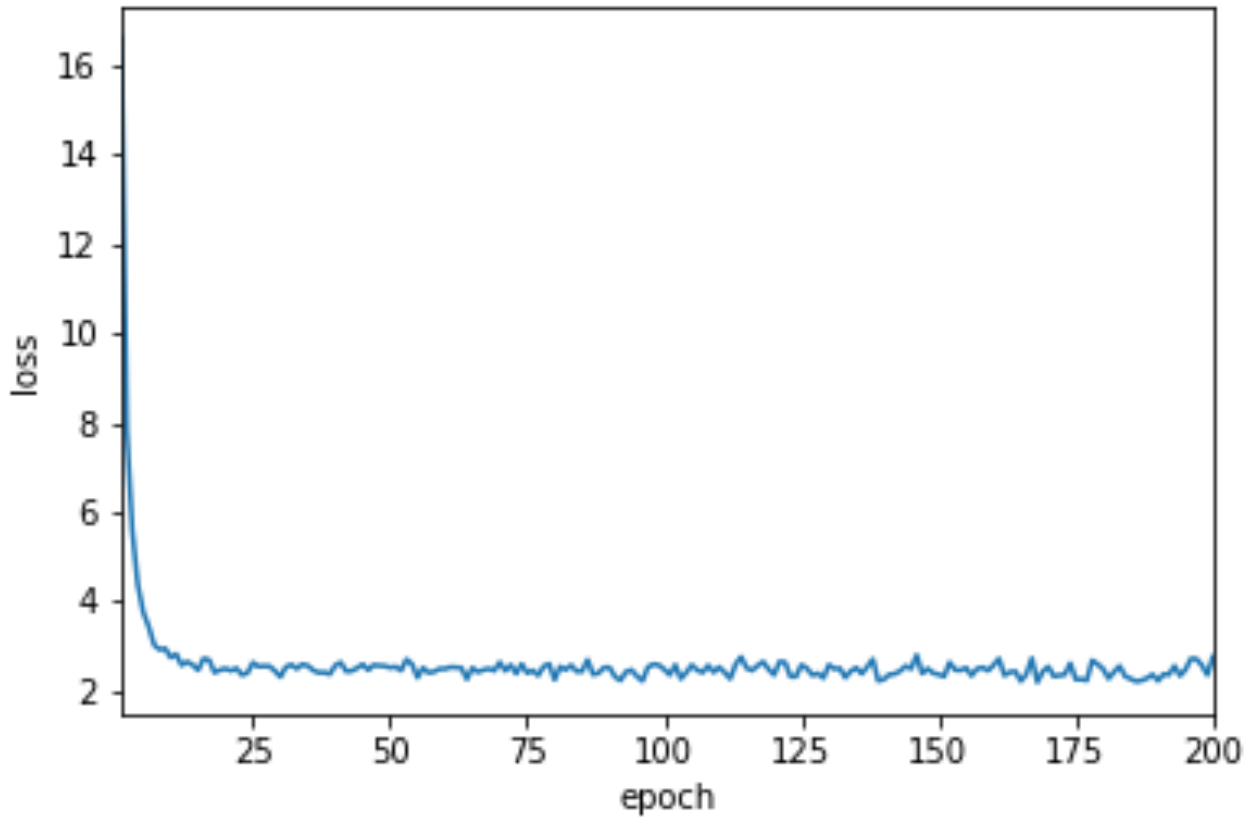
where ${\sigma_\epsilon}^2 = 0.2$

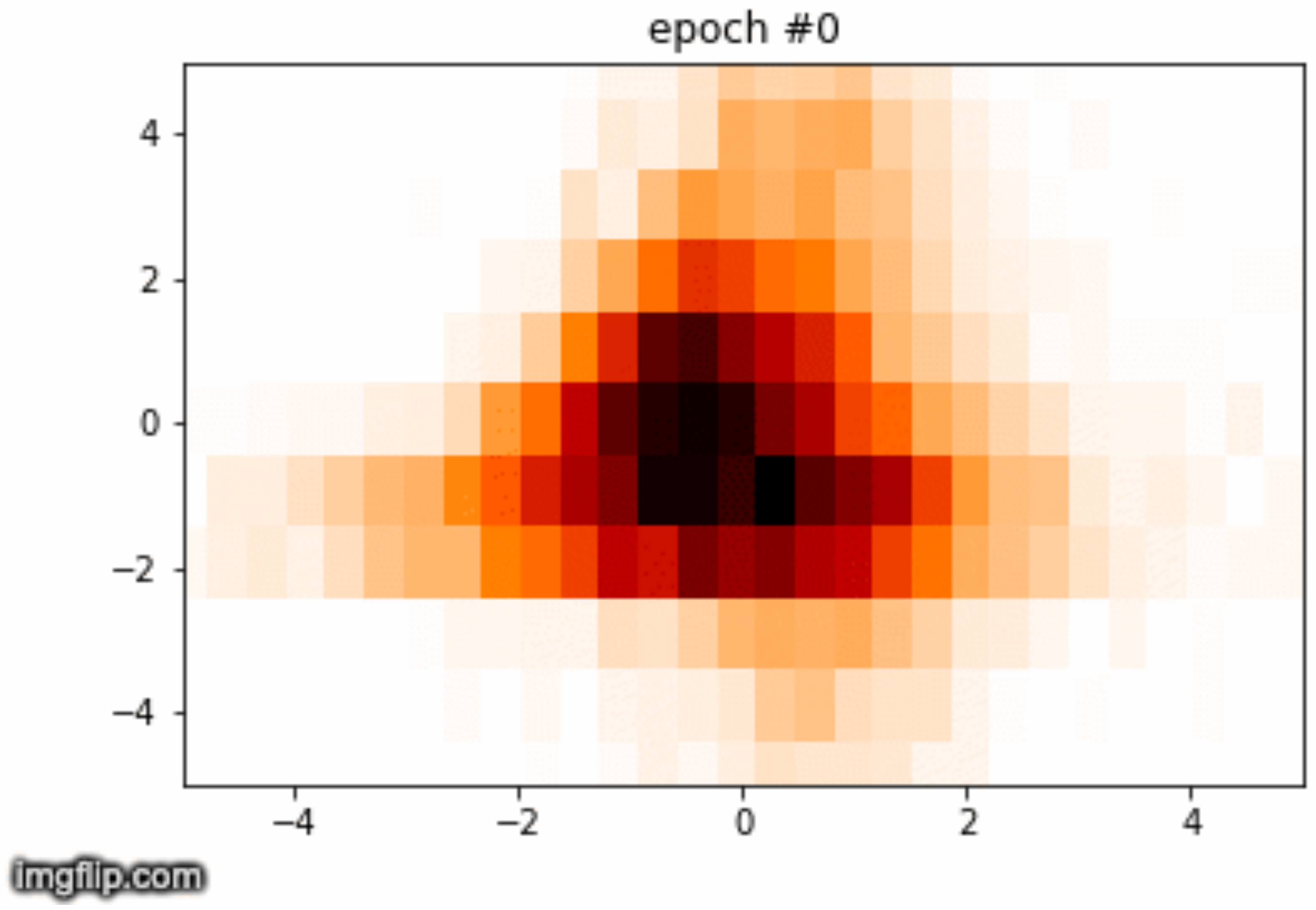
- Diagonal/Full covariance model:
  - # of training samples: 40,000
  - latent dimension: 2
  - batch_size: 32
  - learning rate: 1e-3
  - # epoch: 200

- Testing:
  - # generative samples: 20,000
  - # reconstruction of test set: 20,000

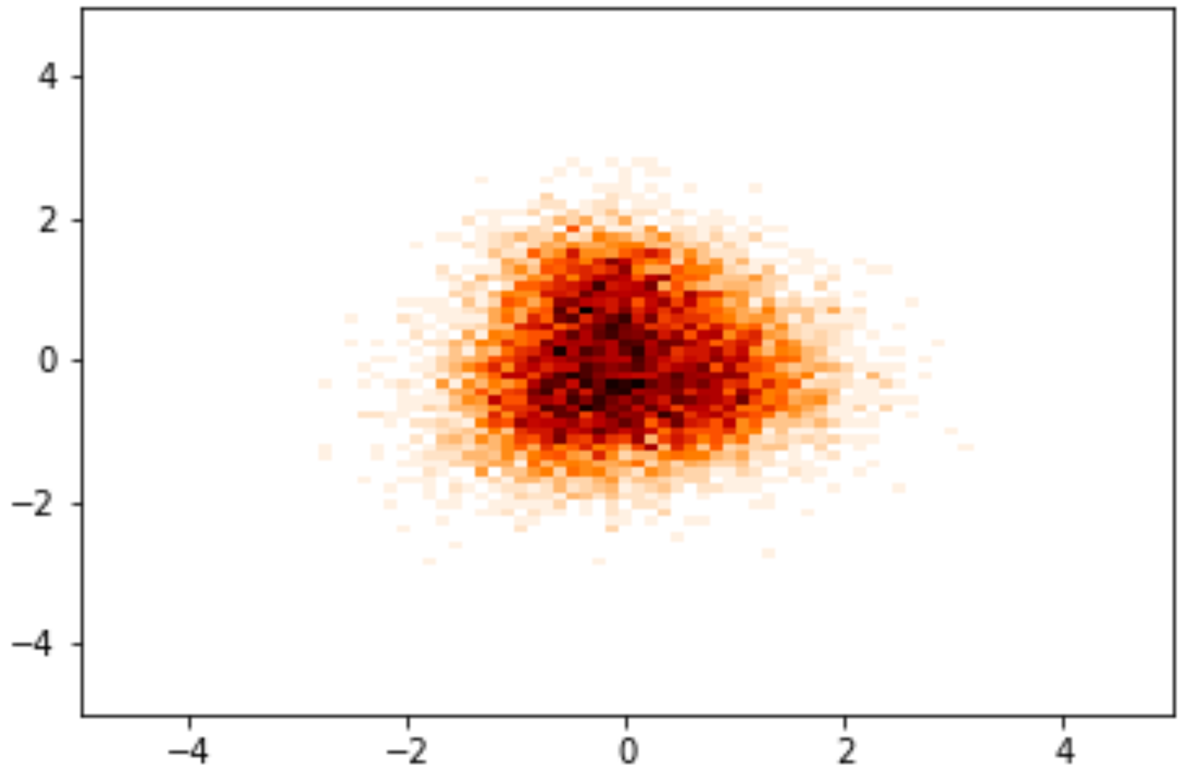
Experiment: three blobs

Diagonal:

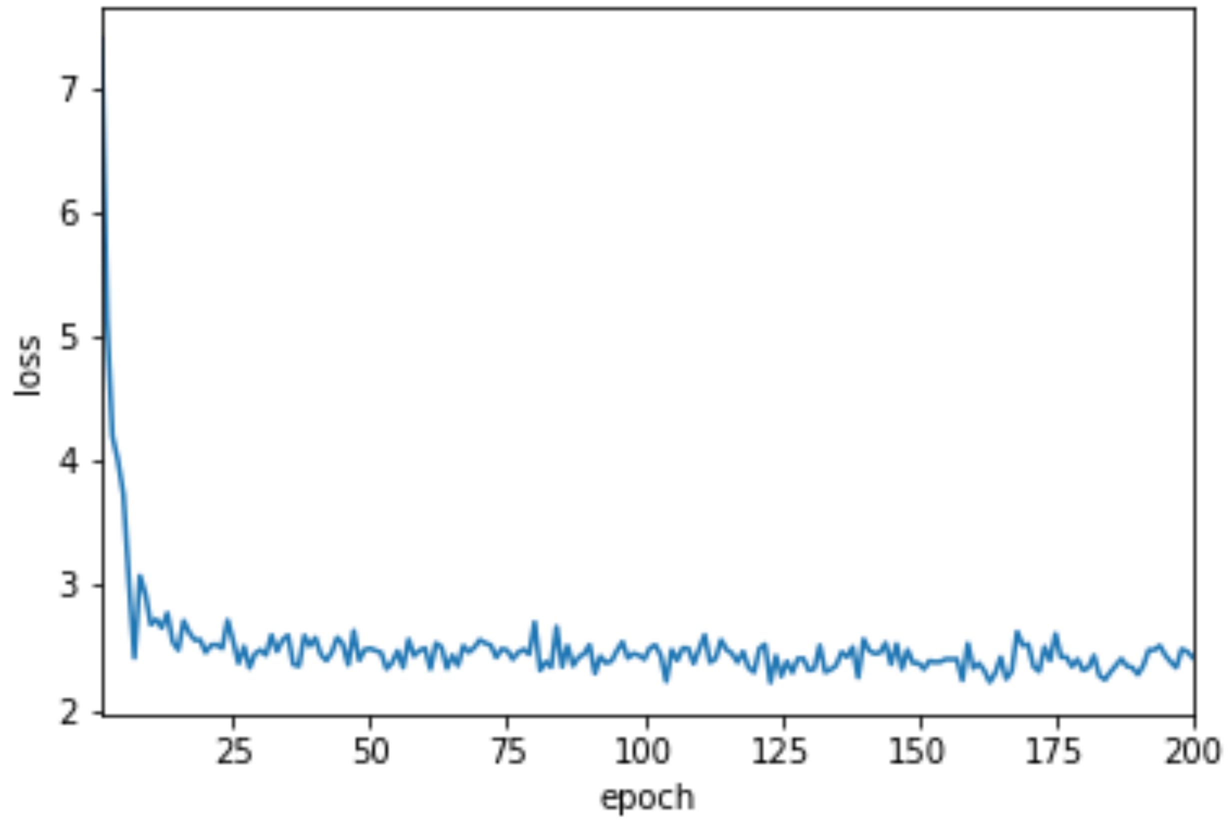


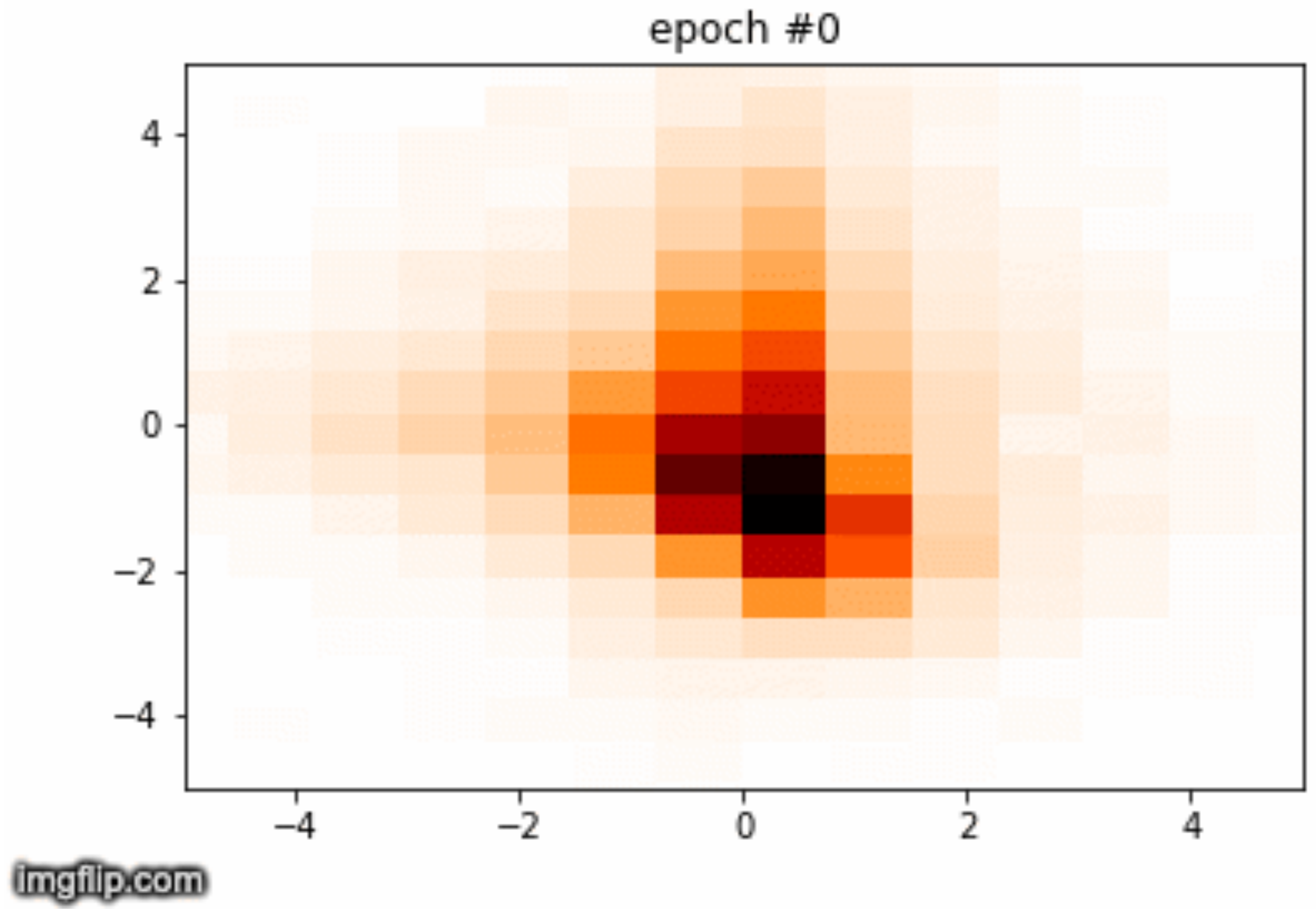


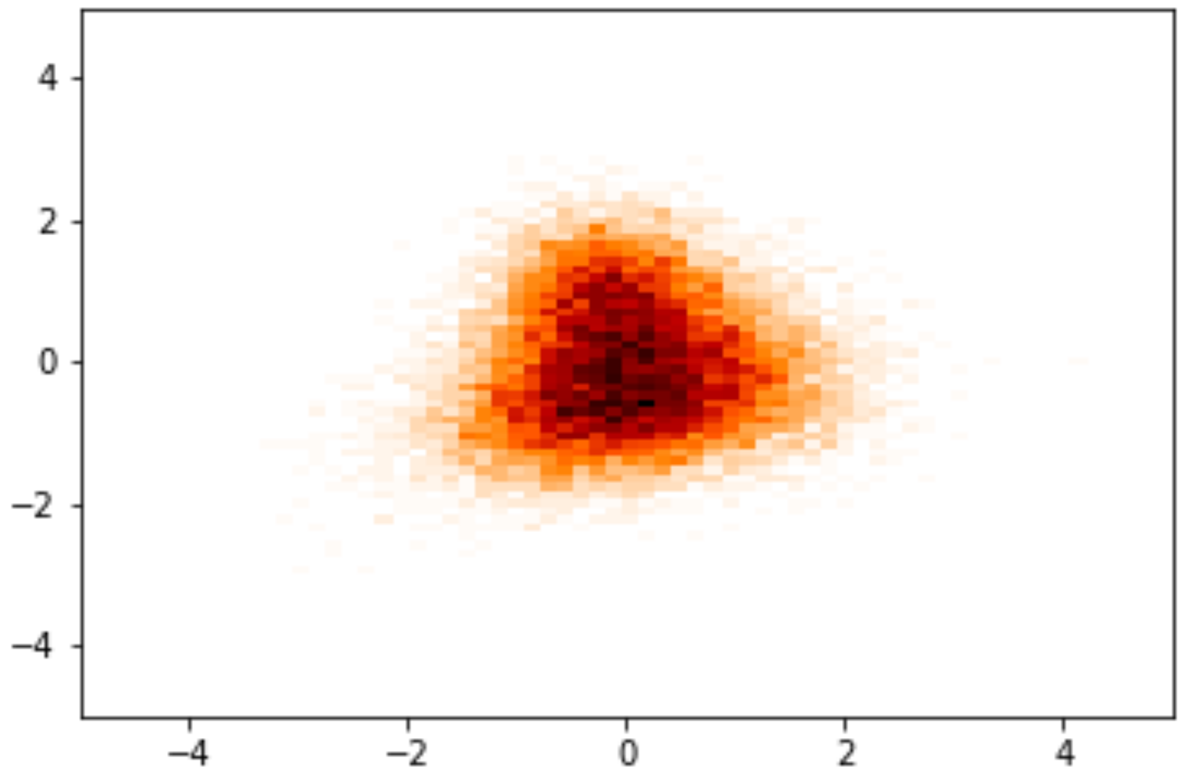
After training



Full:

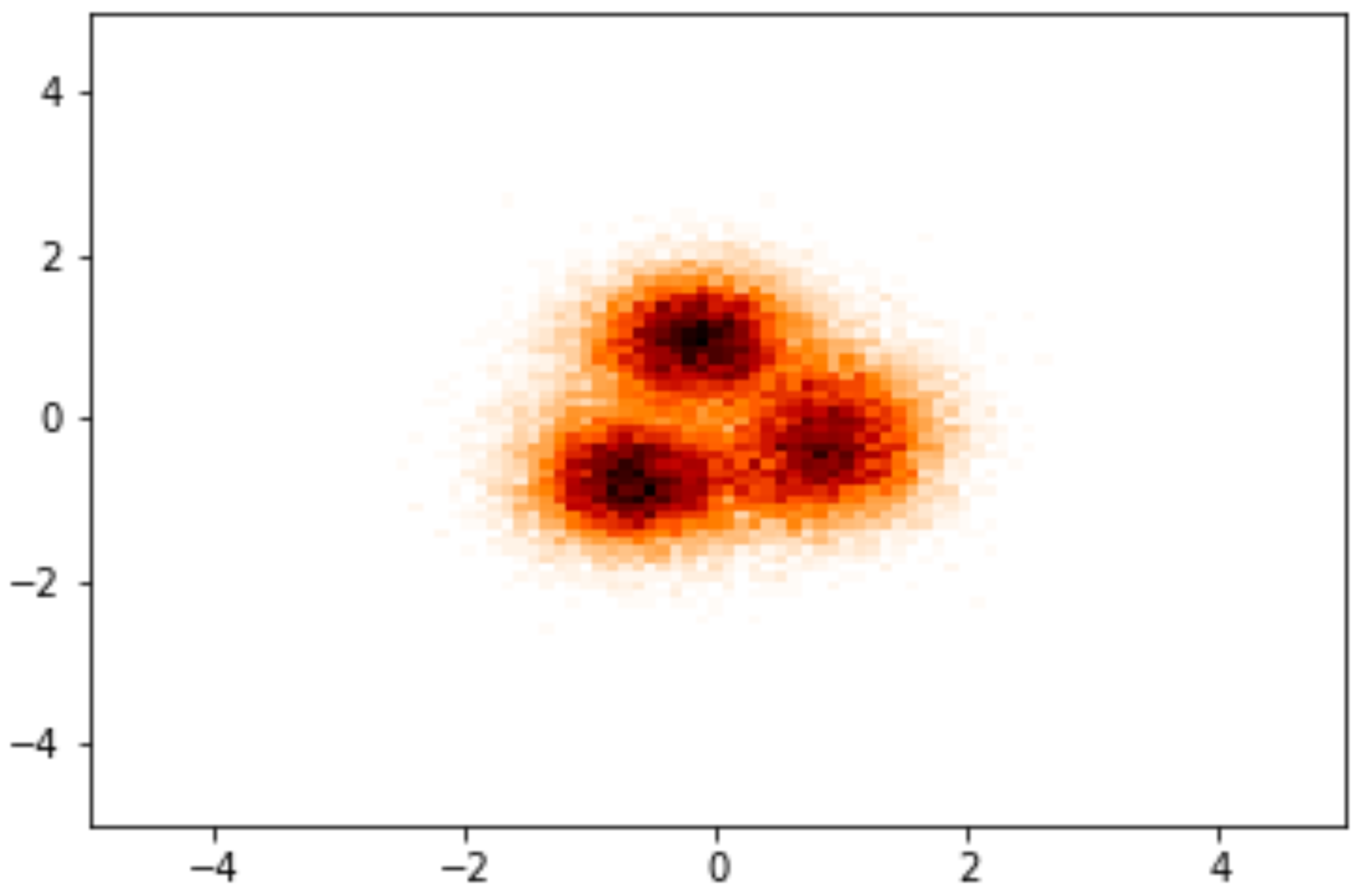






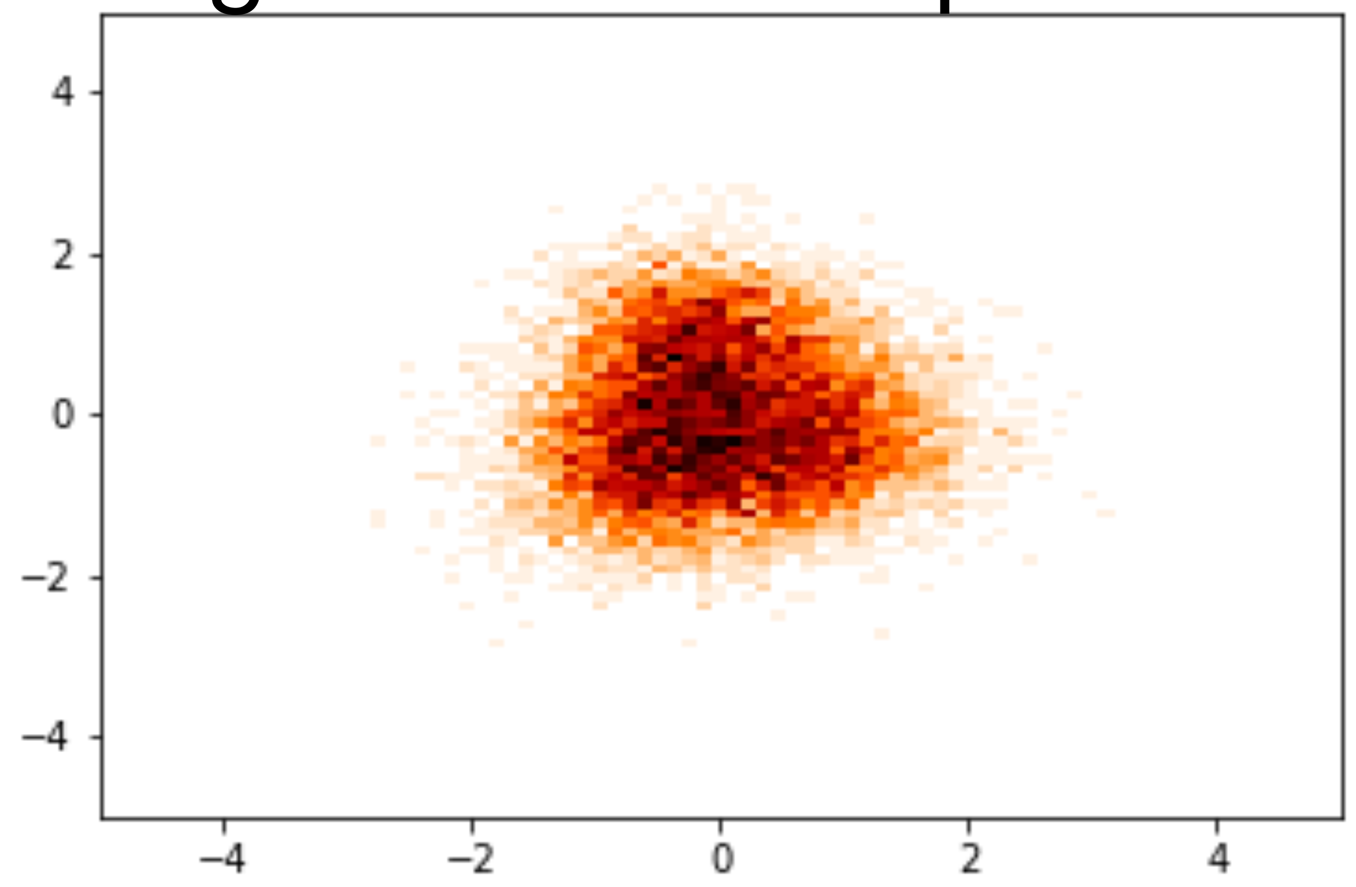
Experiment: three blobs

$X|Z$

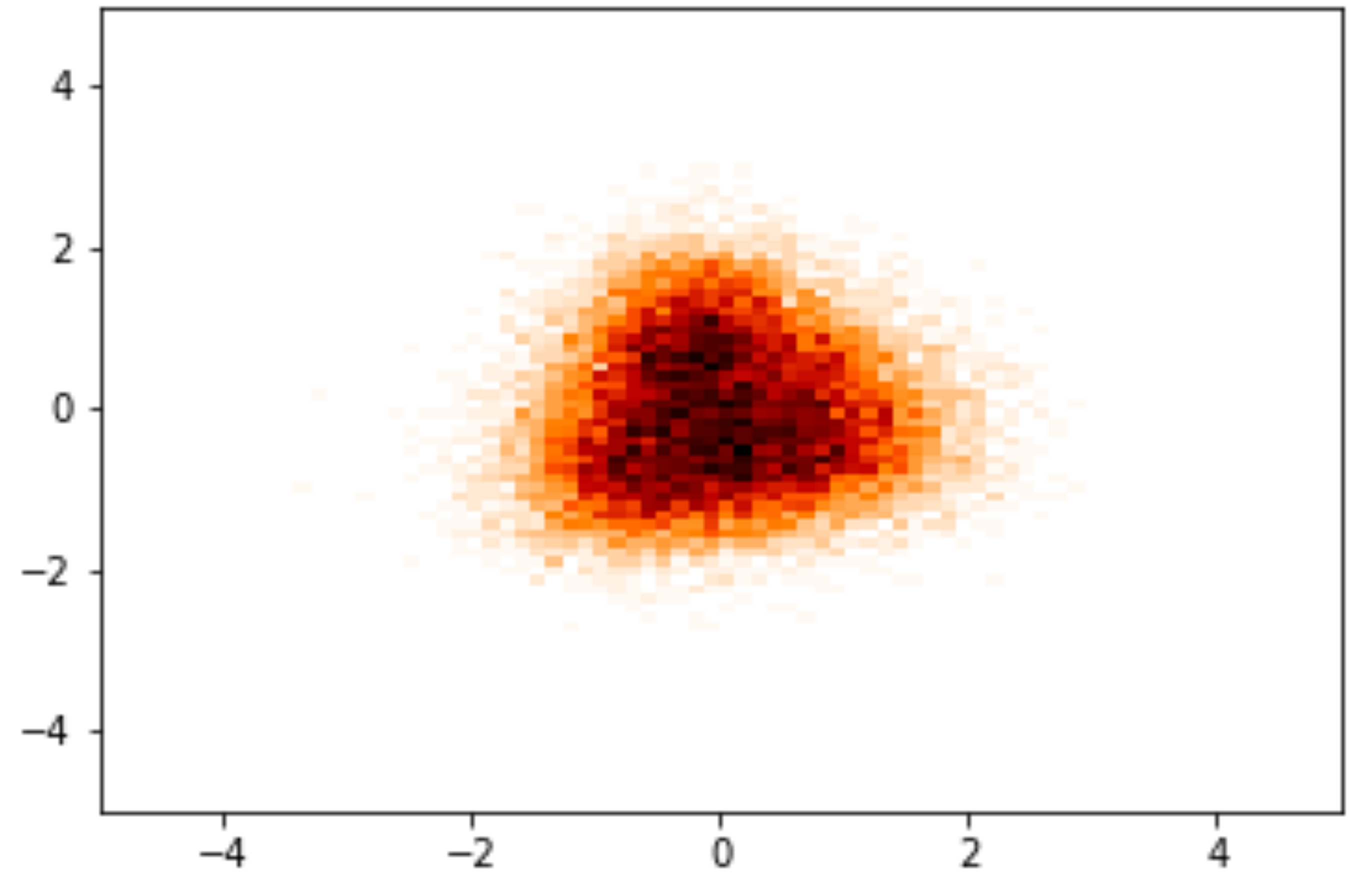


Diagonal:

generative samples

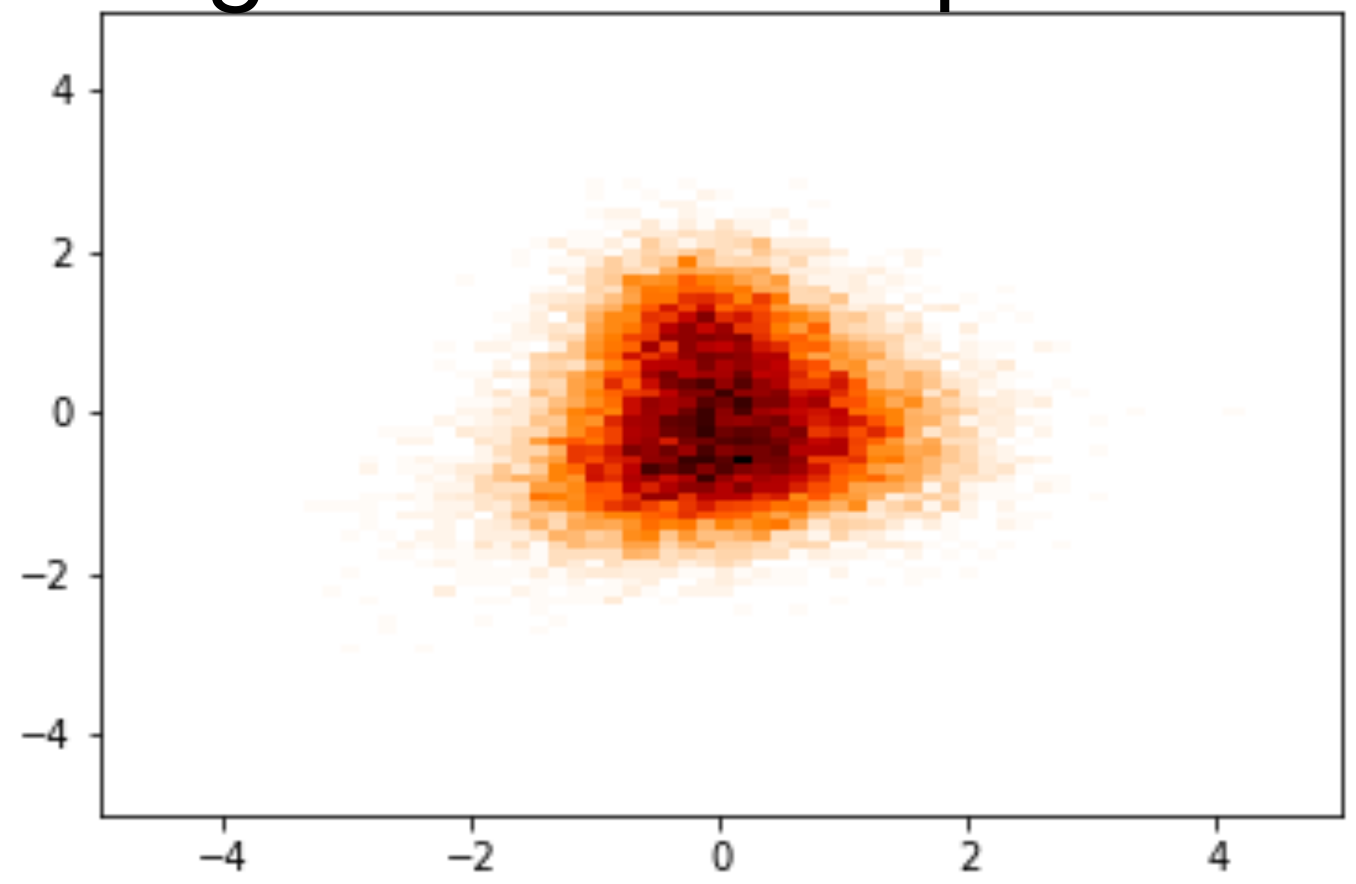


reconstruction of test set

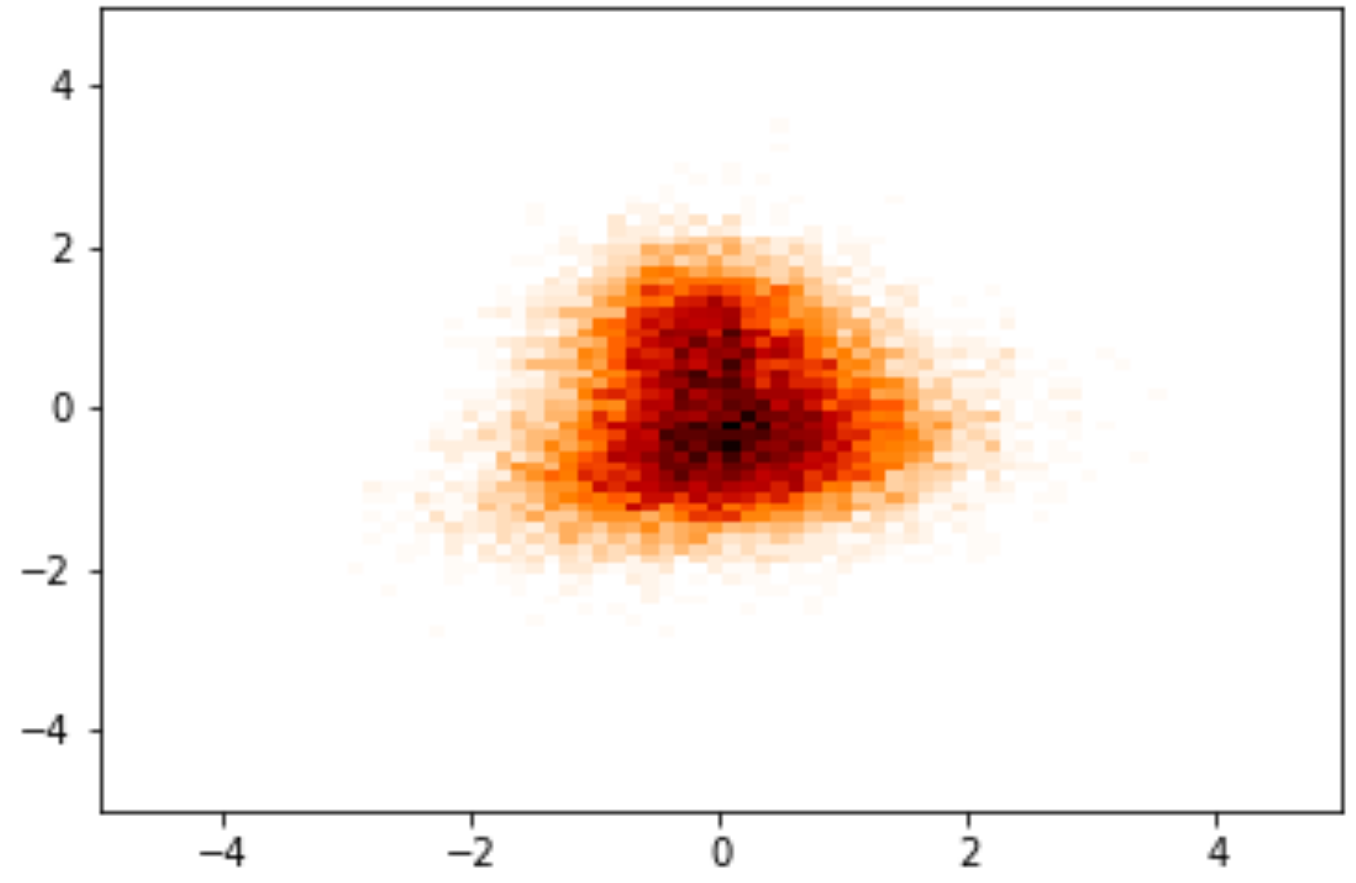


Full:

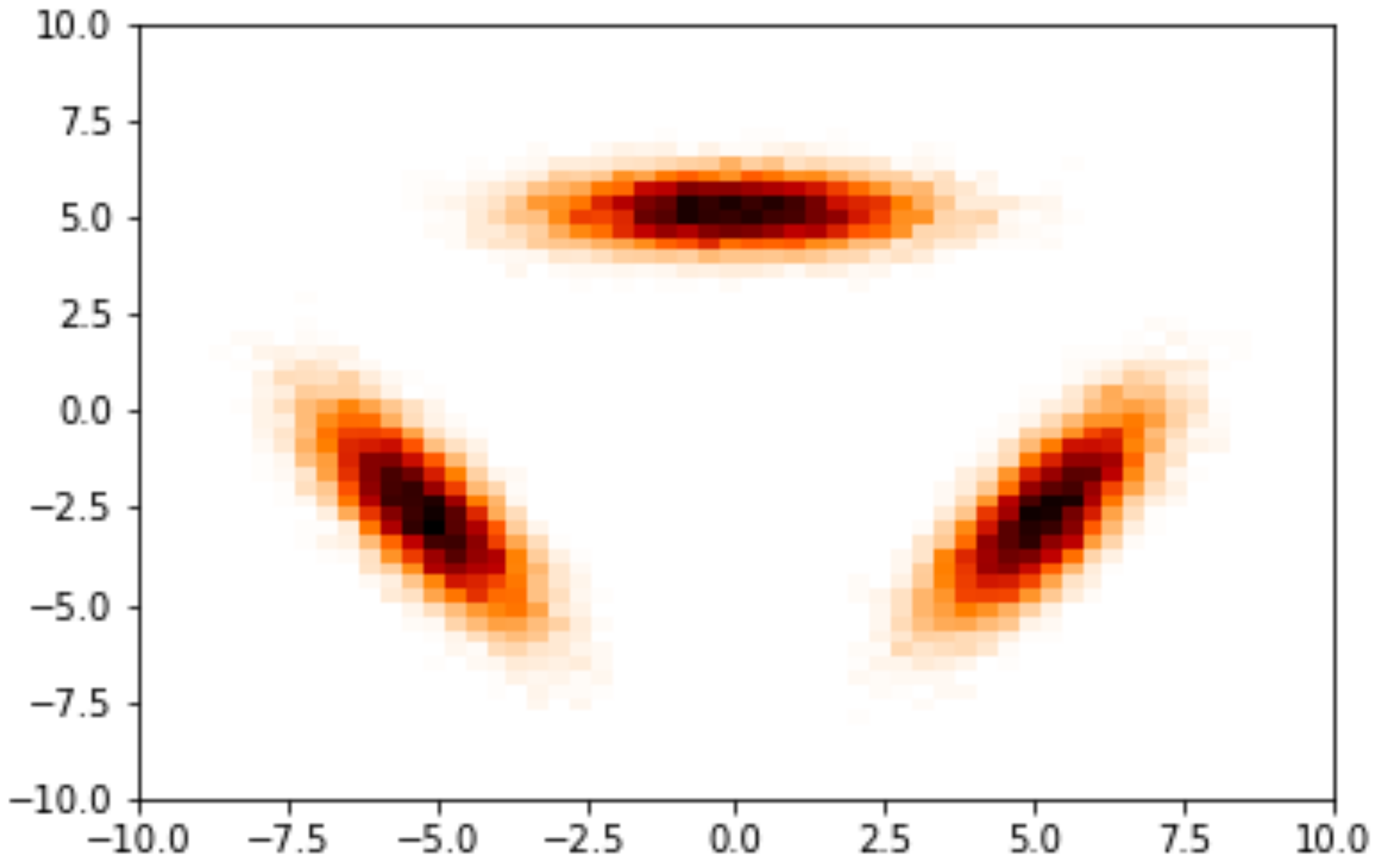
generative samples



reconstruction of test set



Experiment: mixture of Gaussians



$$Y_i \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}3 & 0\\0 & 1/3\end{bmatrix}\right), i \in 1,2,3$$

$$X_1 = Y_1 + \begin{bmatrix}0\\3\sqrt{3}\end{bmatrix}$$

$$X_2 = \begin{bmatrix}cos\left(\frac{2\pi}{3}\right) & \text{-}sin\left(\frac{2\pi}{3}\right)\\ sin\left(\frac{2\pi}{3}\right) & cos\left(\frac{2\pi}{3}\right)\end{bmatrix} \cdot Y_2 + \begin{bmatrix}\text{-}3\sqrt{3}\\ \text{-}\frac{3\sqrt{3}}{2}\end{bmatrix}$$

$$X_3 = \begin{bmatrix}cos\left(\frac{\pi}{3}\right) & \text{-}sin\left(\frac{2}{3}\right)\\ sin\left(\frac{\pi}{3}\right) & cos\left(\frac{\pi}{3}\right)\end{bmatrix} \cdot Y_3 + \begin{bmatrix}3\sqrt{3}\\ \text{-}\frac{3\sqrt{3}}{2}\end{bmatrix}$$

$$Z \sim U[0,1]$$

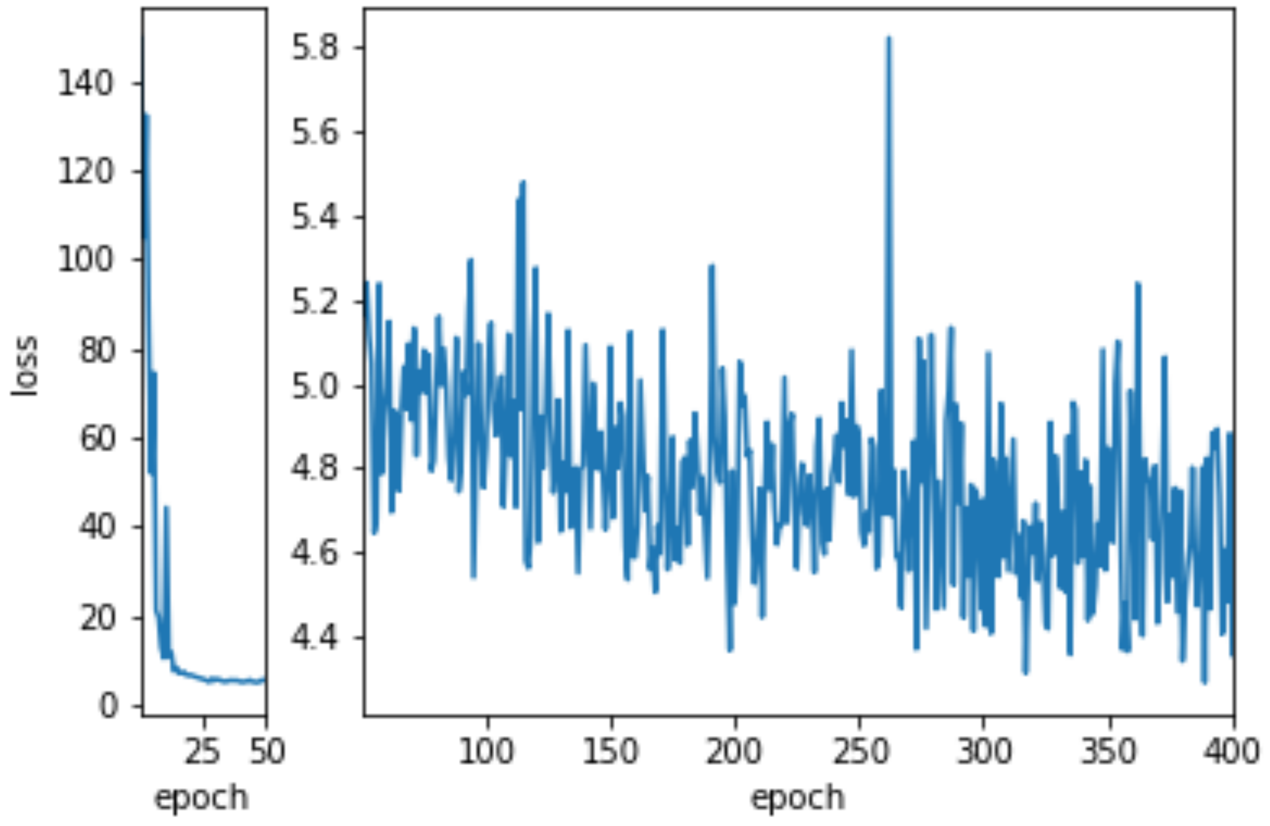
$$x|z = \begin{cases} x_1 & \text{, for } z < \frac{1}{3} \\ x_2 & \text{, for } \frac{1}{3} \leq z < \frac{2}{3} \\ x_3 & \text{, for } z \geq \frac{2}{3} \end{cases}$$

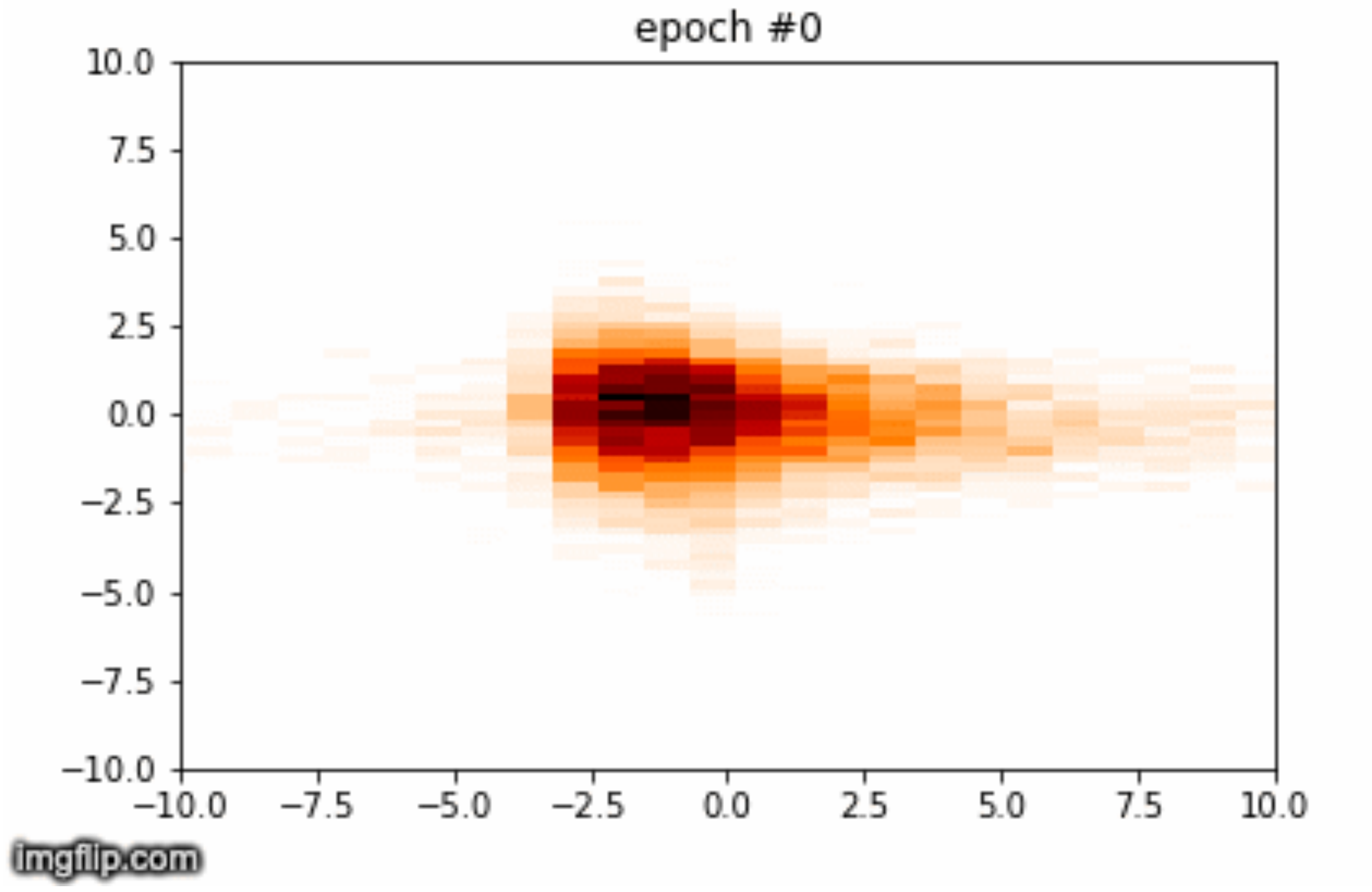
- Diagonal/Full covariance model:
  - # of training samples: 40,000
  - latent dimension: 2
  - batch_size: 32
  - learning rate: 1e-3
  - # epoch: 400

- Testing:
  - # generative samples: 20,000
  - # reconstruction of test set: 20,000

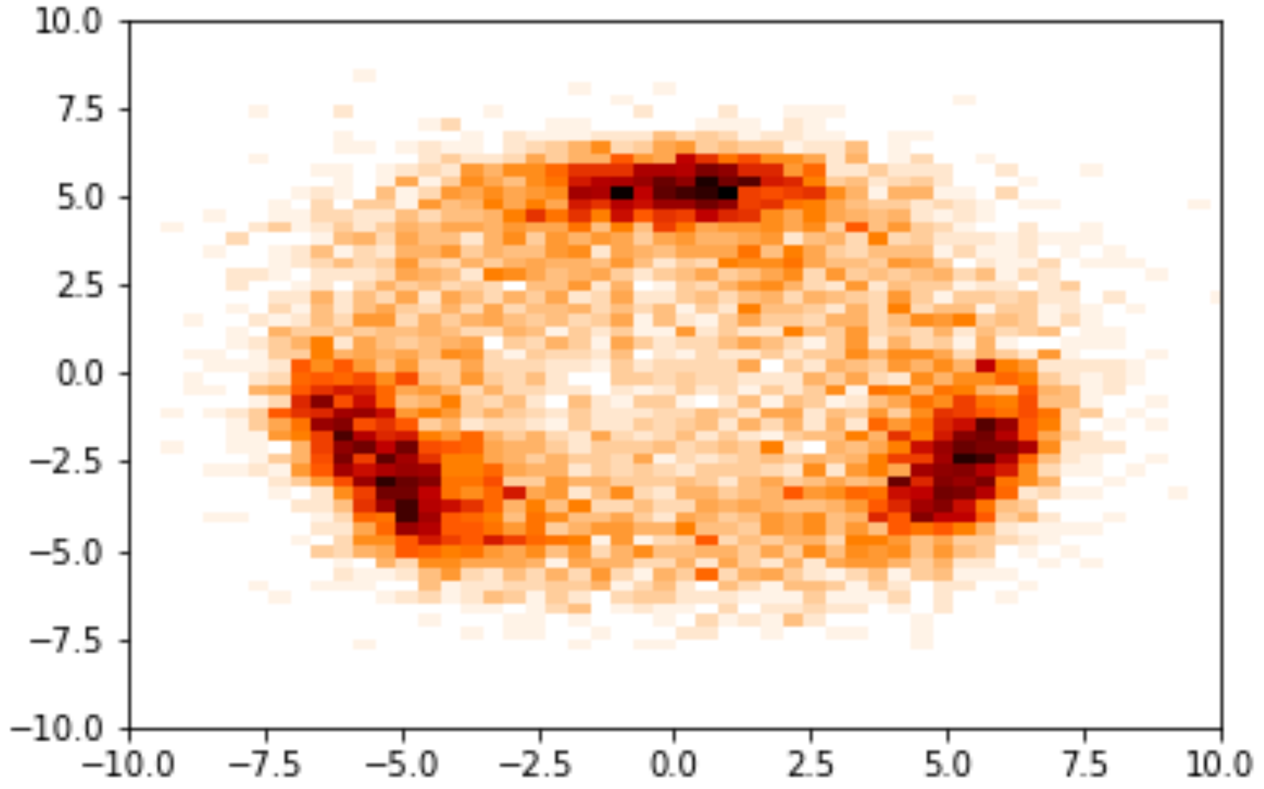
Experiment: mixture of Gaussians

Diagonal:

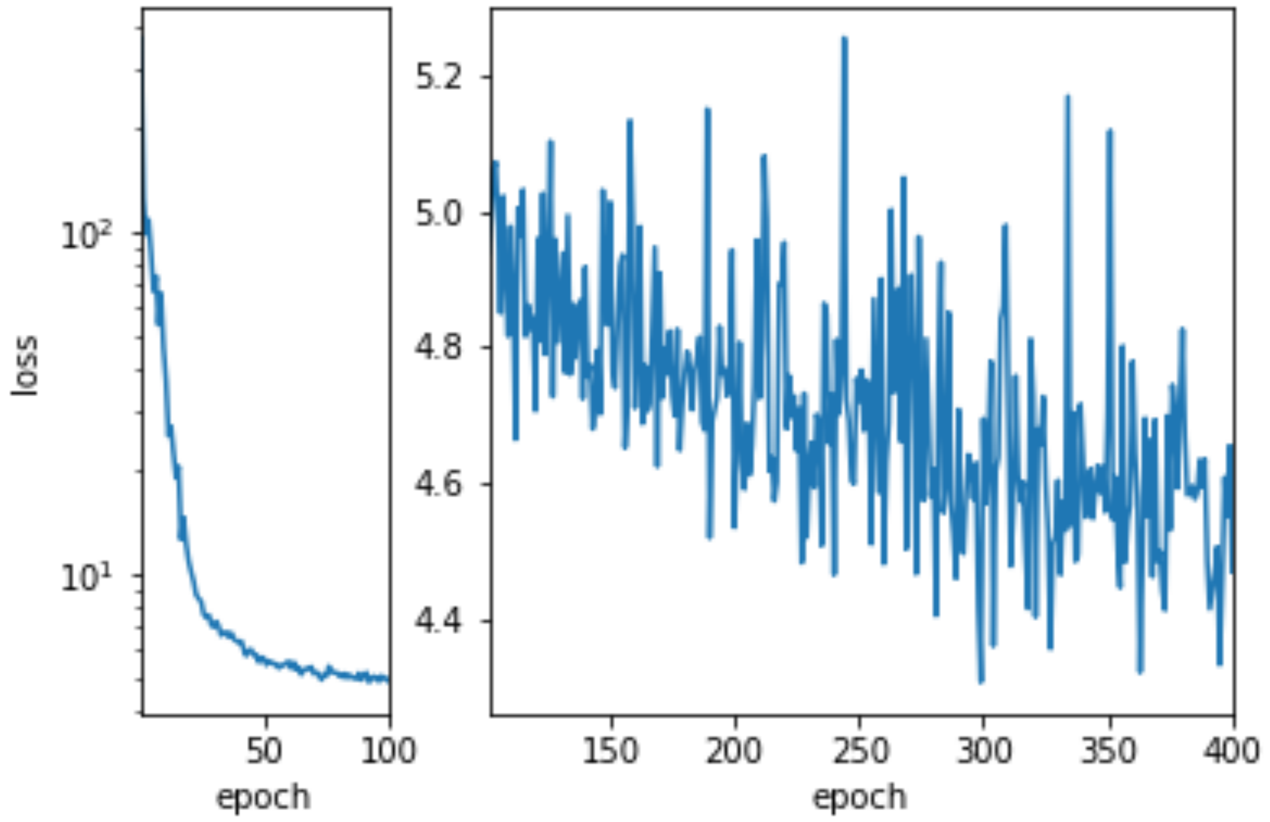


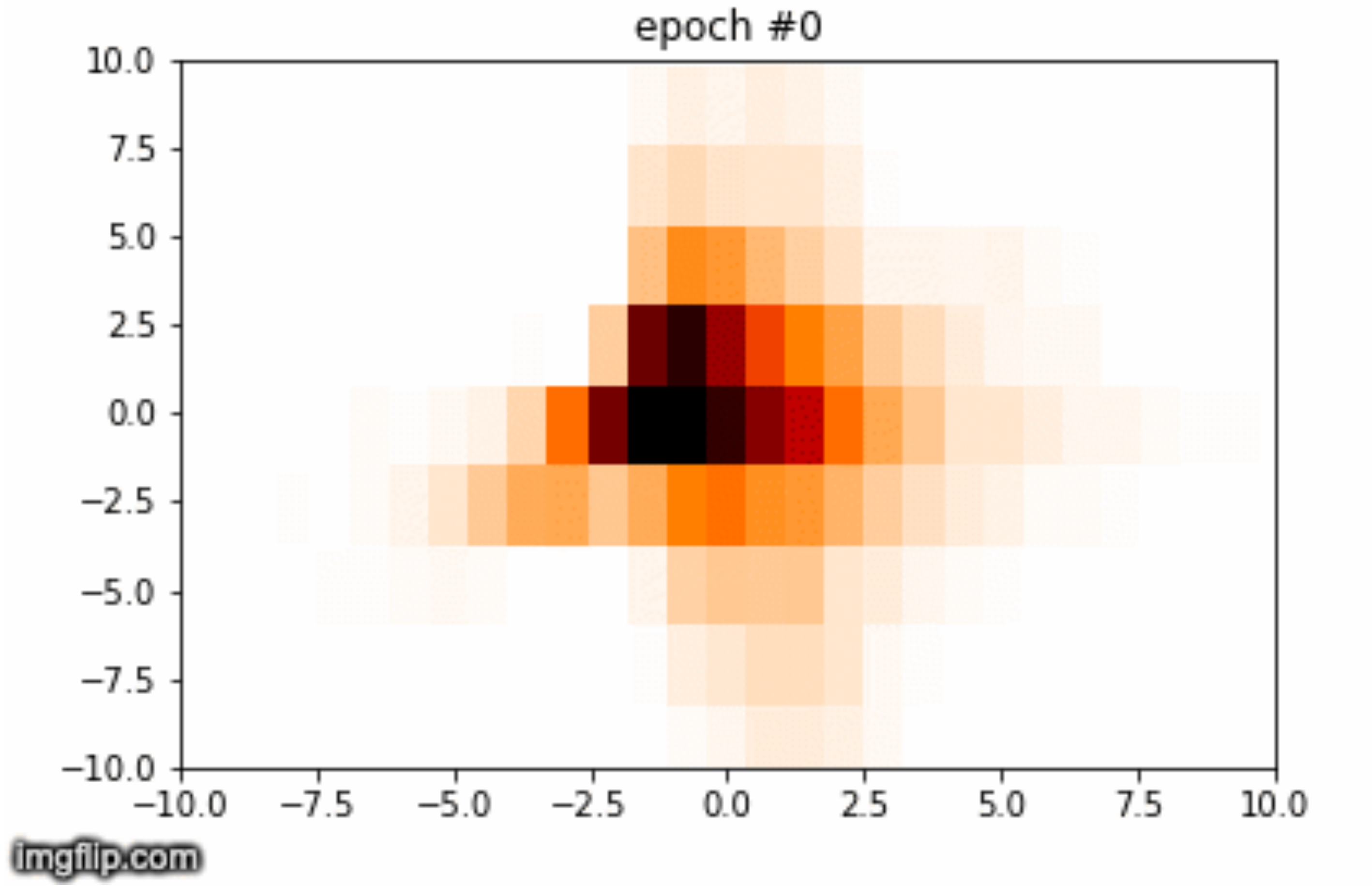


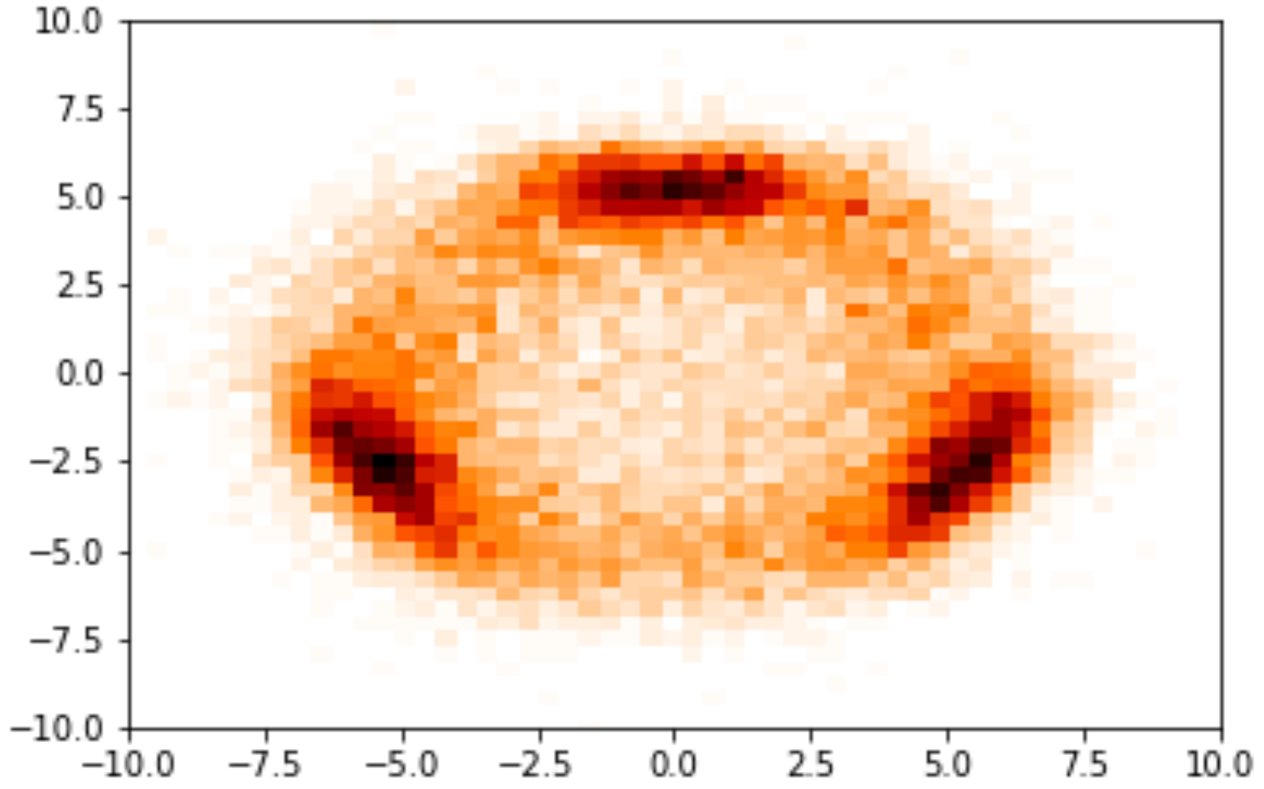
After training



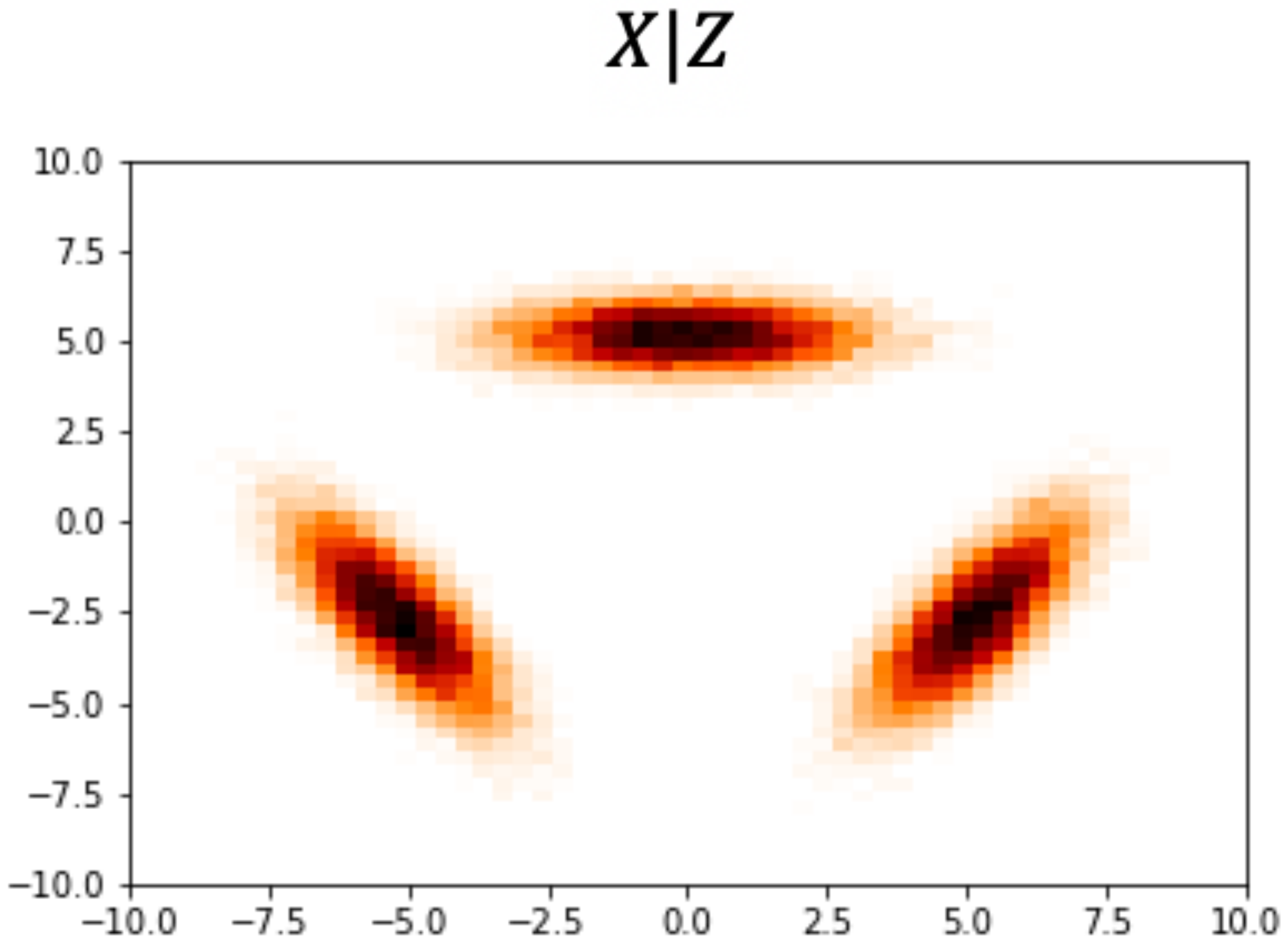
Full:





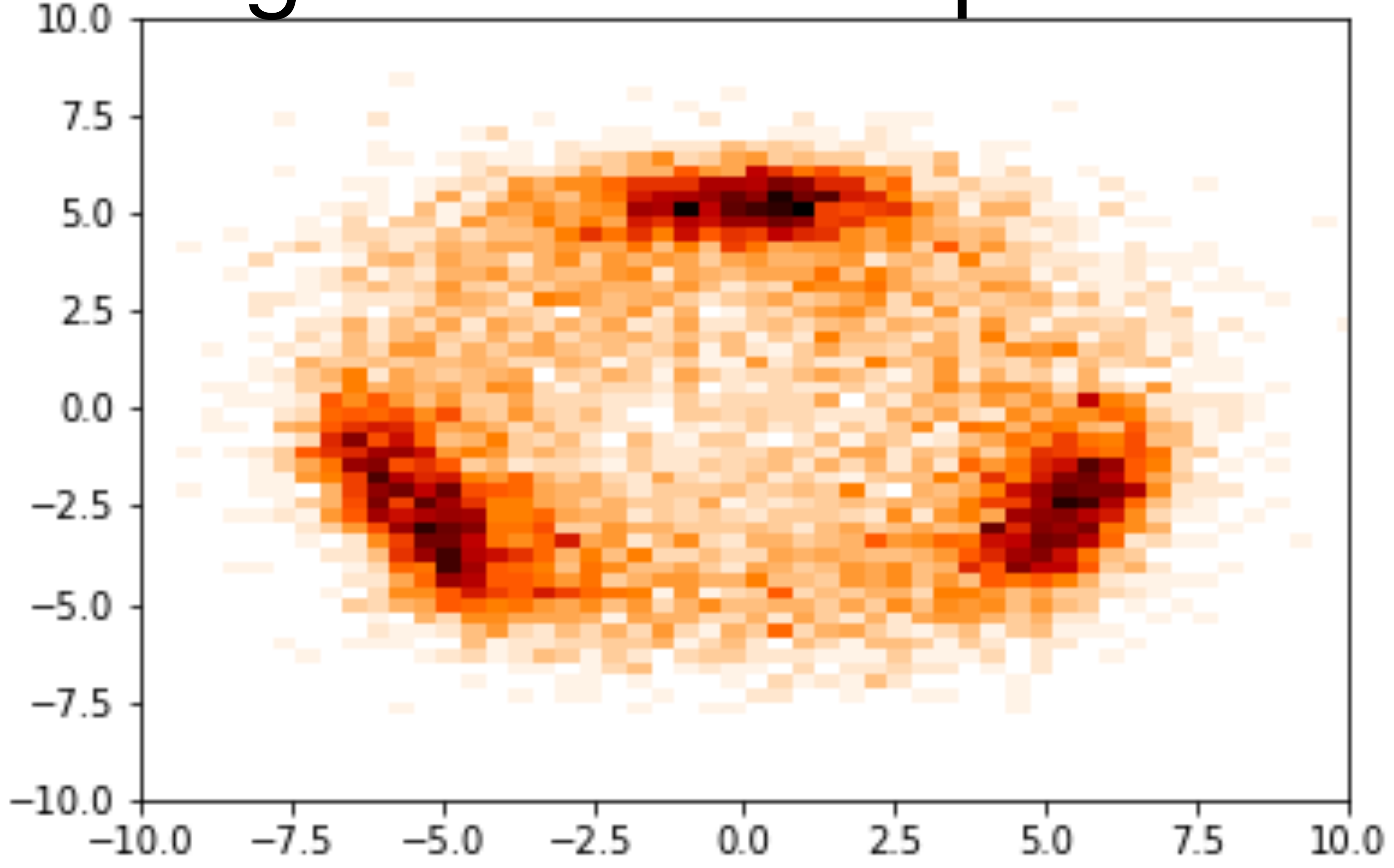


Experiment: mixture of Gaussians

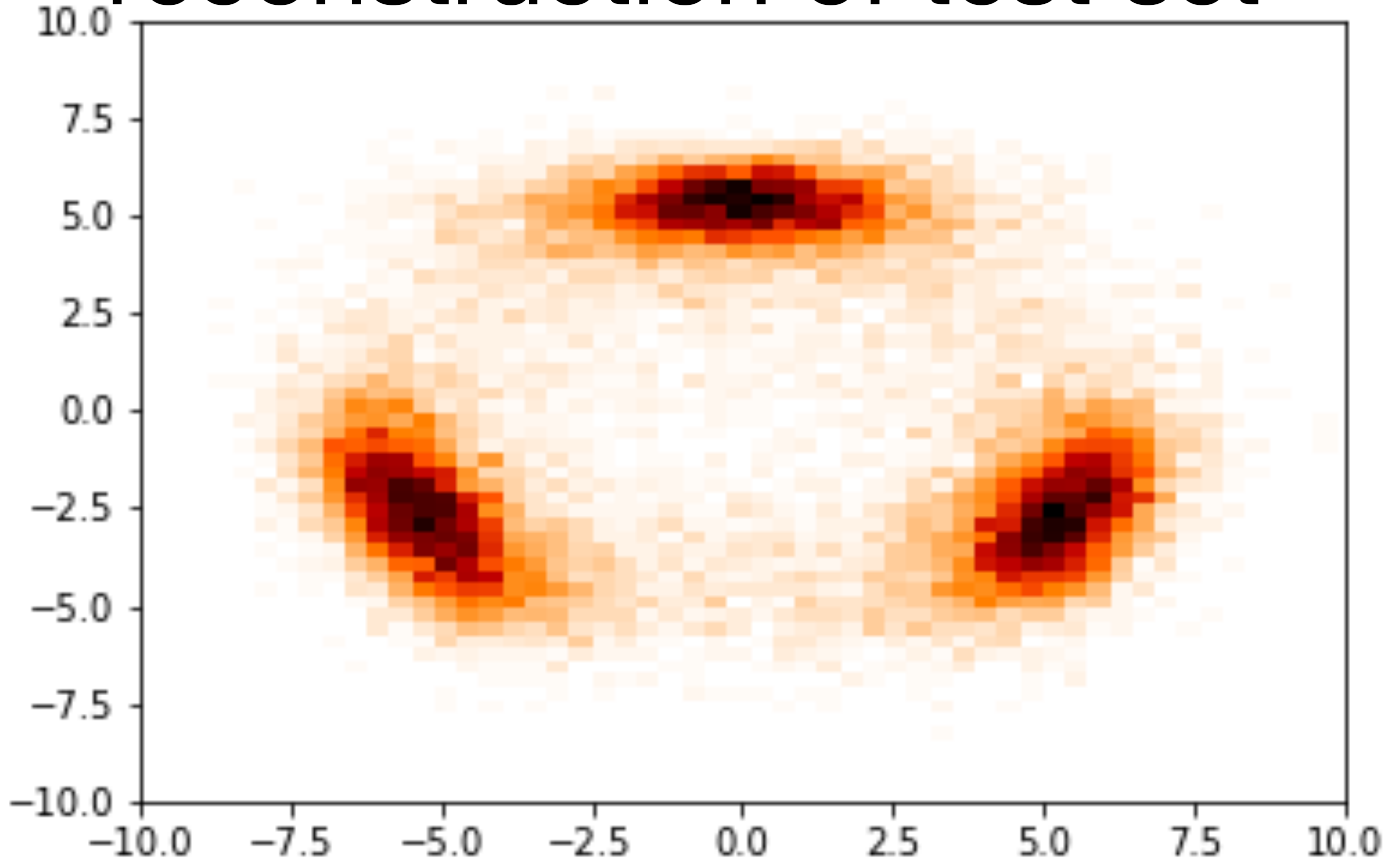


Diagonal:

generative samples

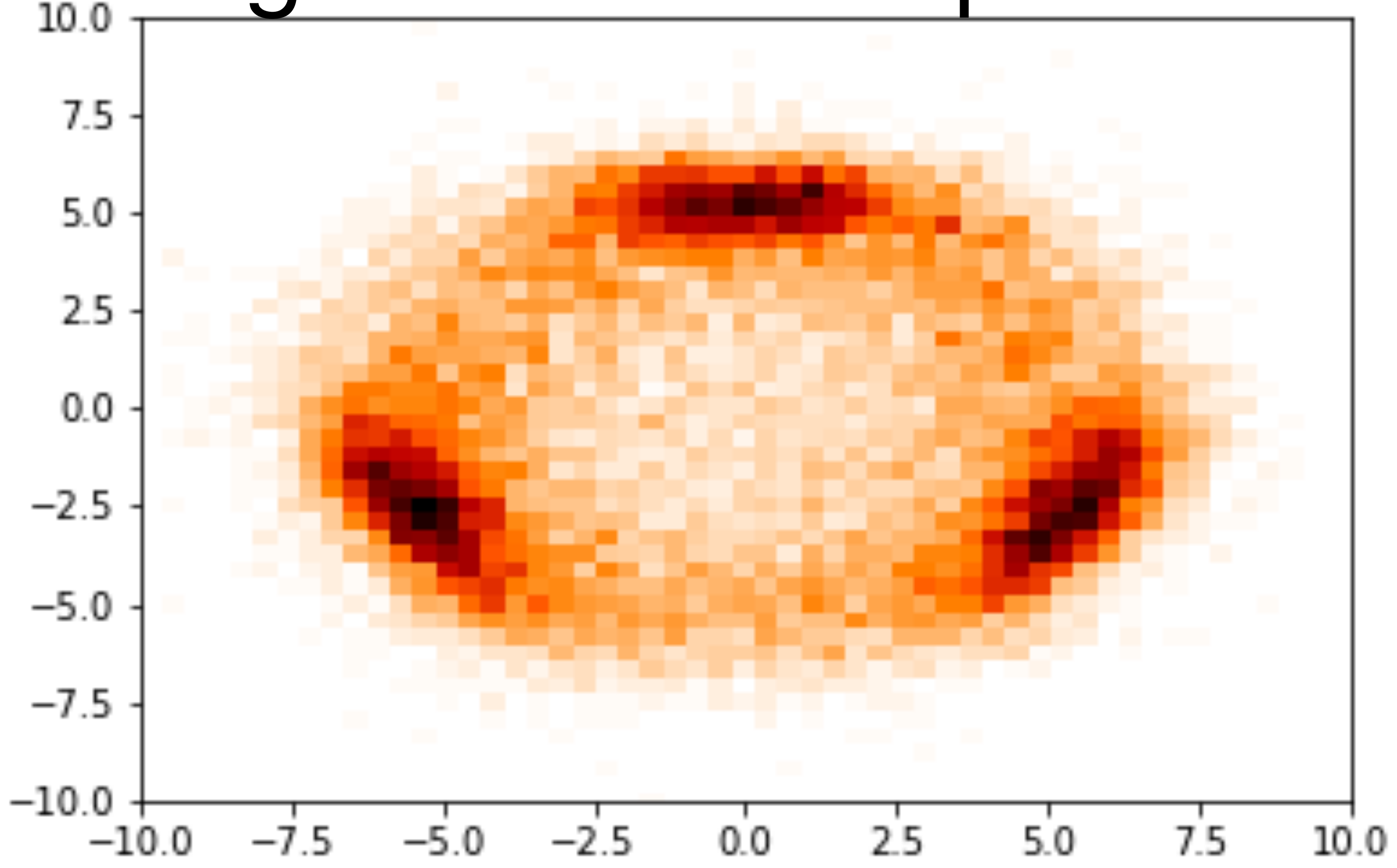


reconstruction of test set

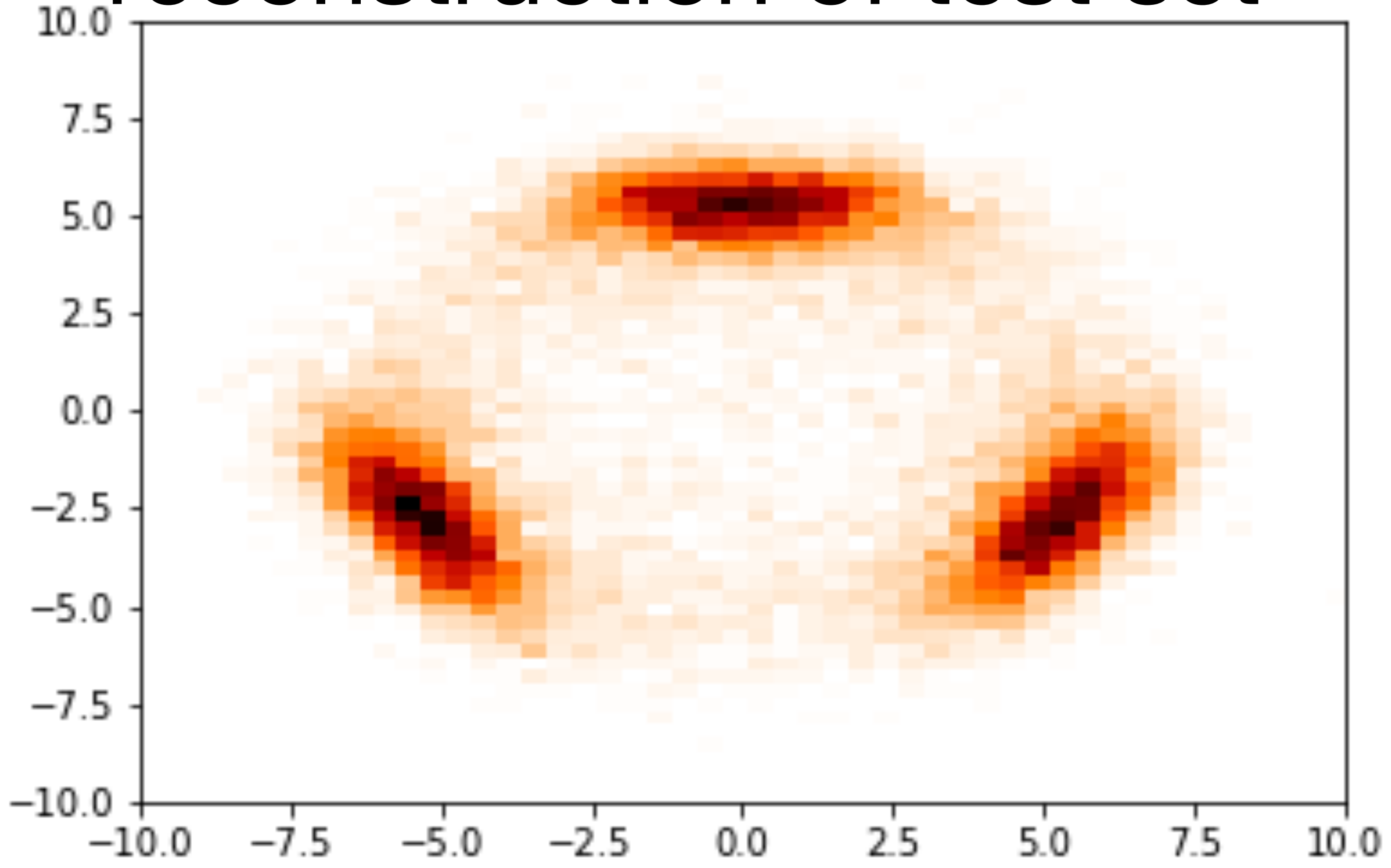


Full:

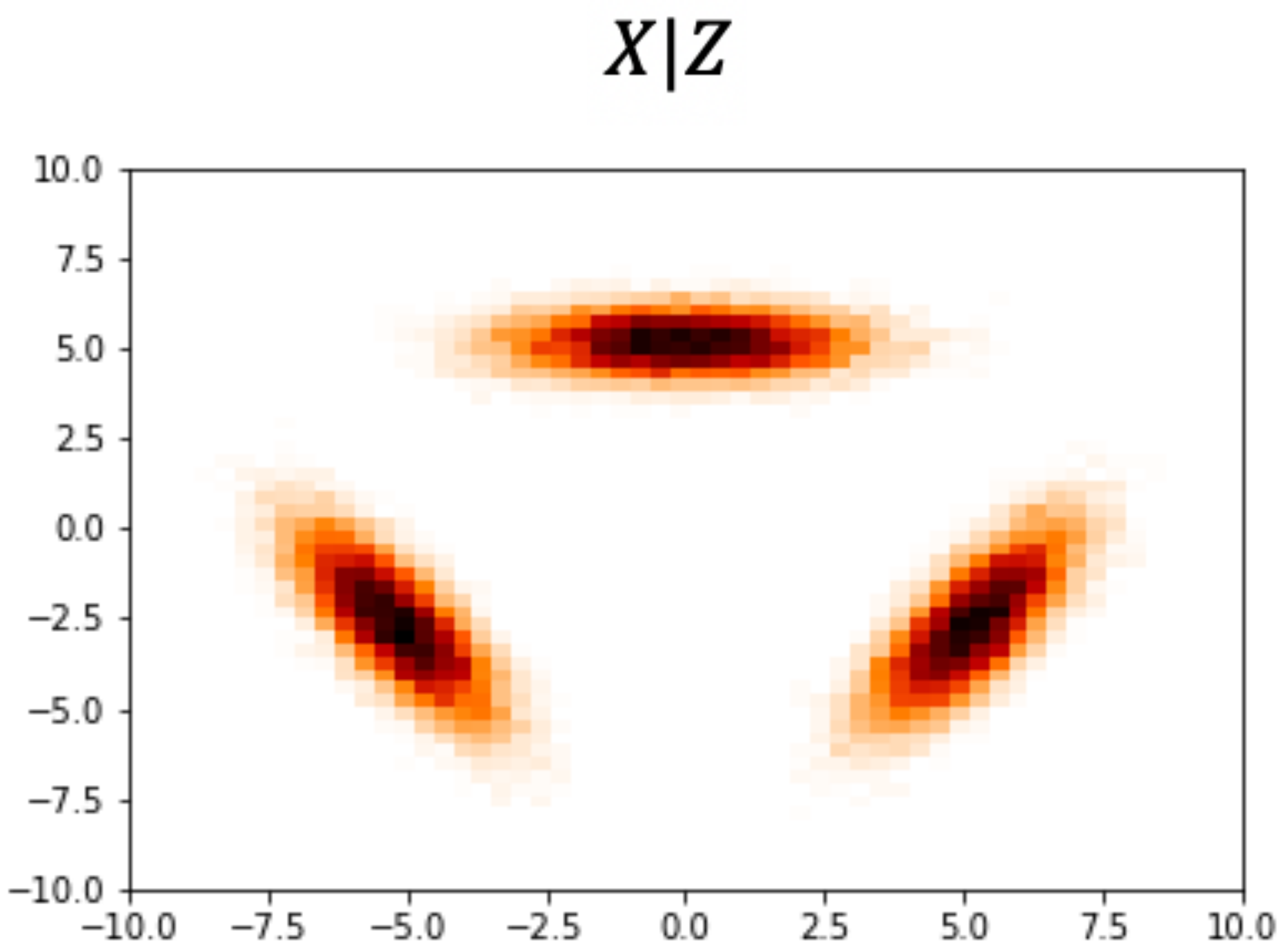
generative samples



reconstruction of test set

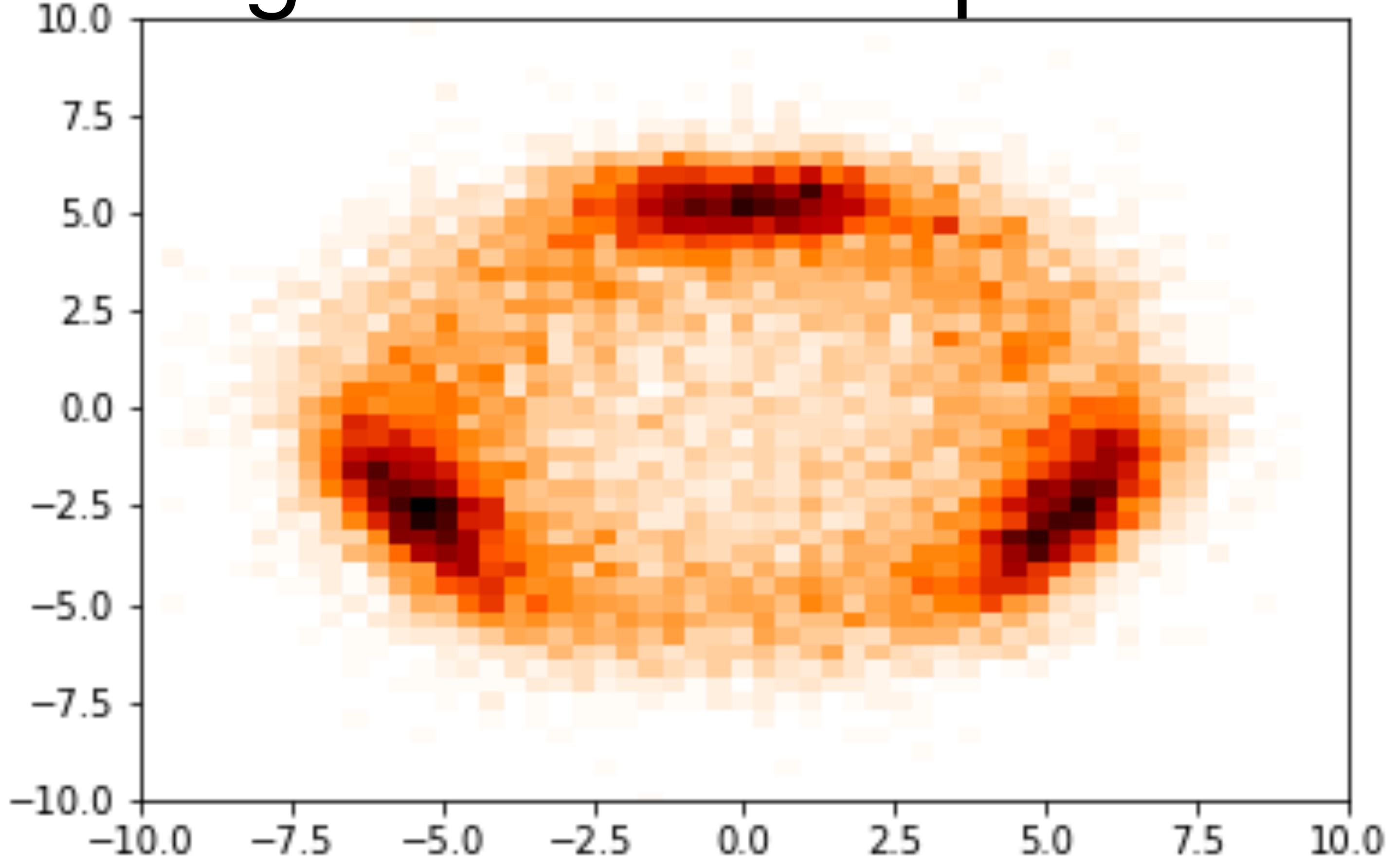


Experiment: mixture of Gaussians, latent d?

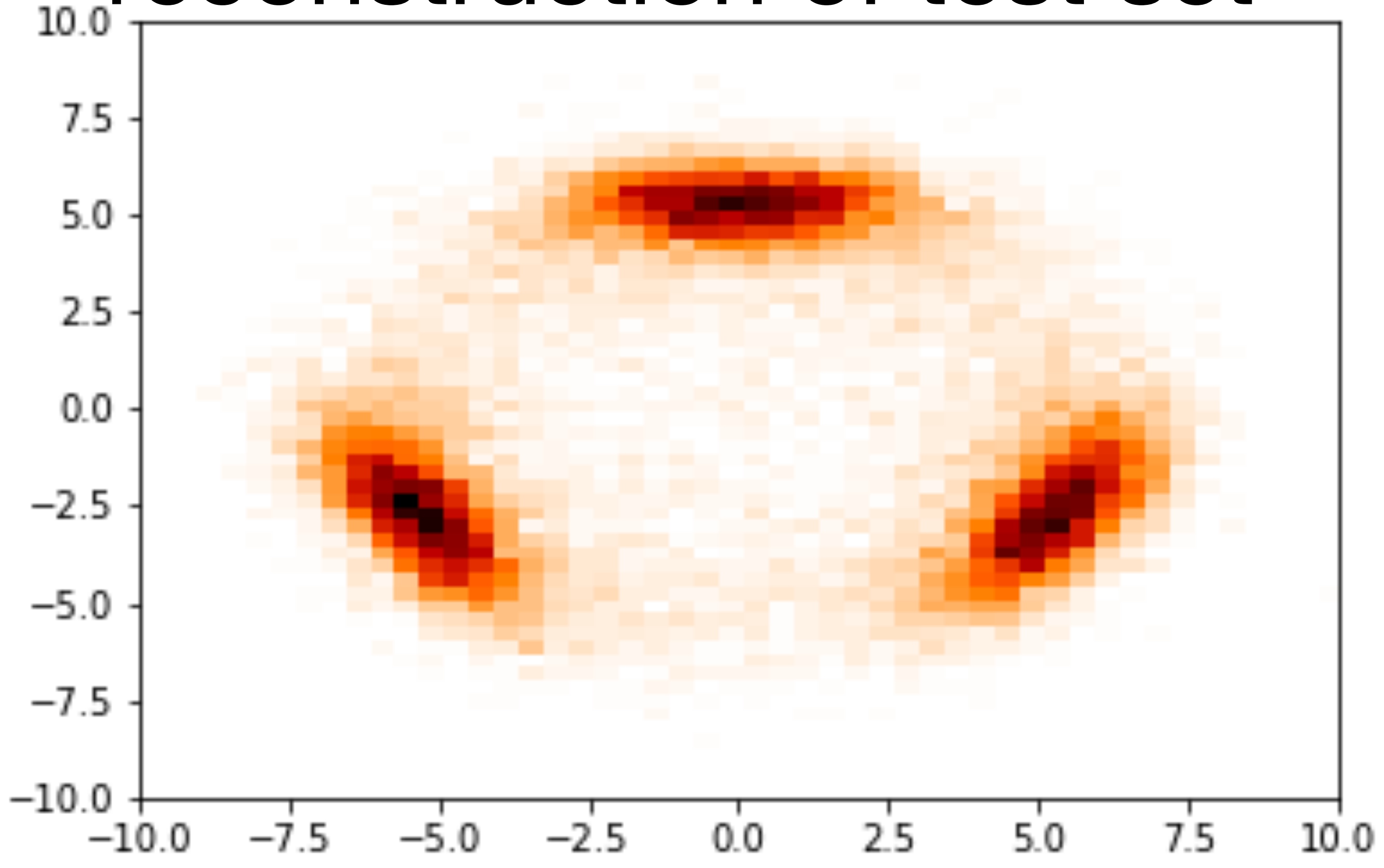


dim = 2:

generative samples

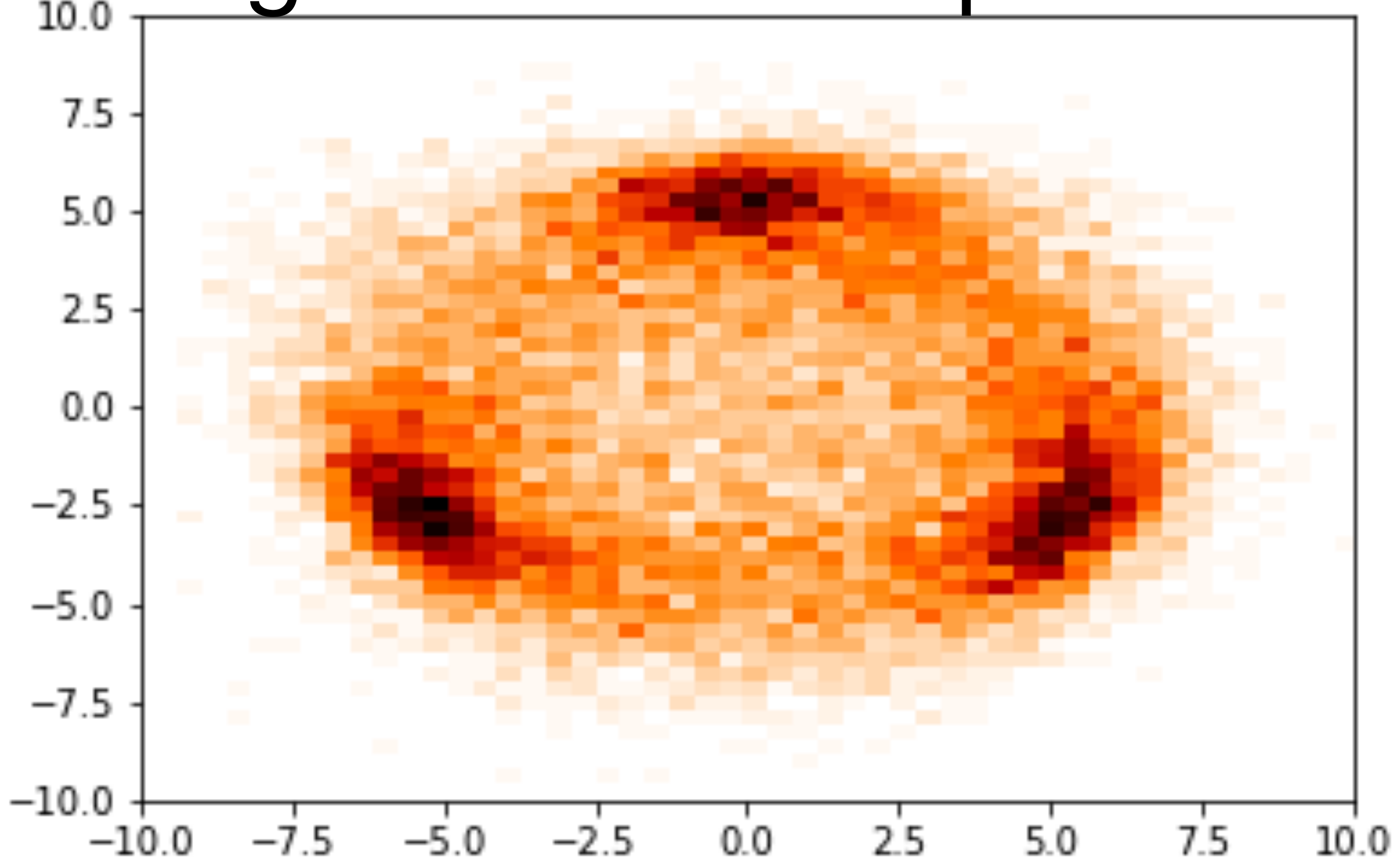


reconstruction of test set

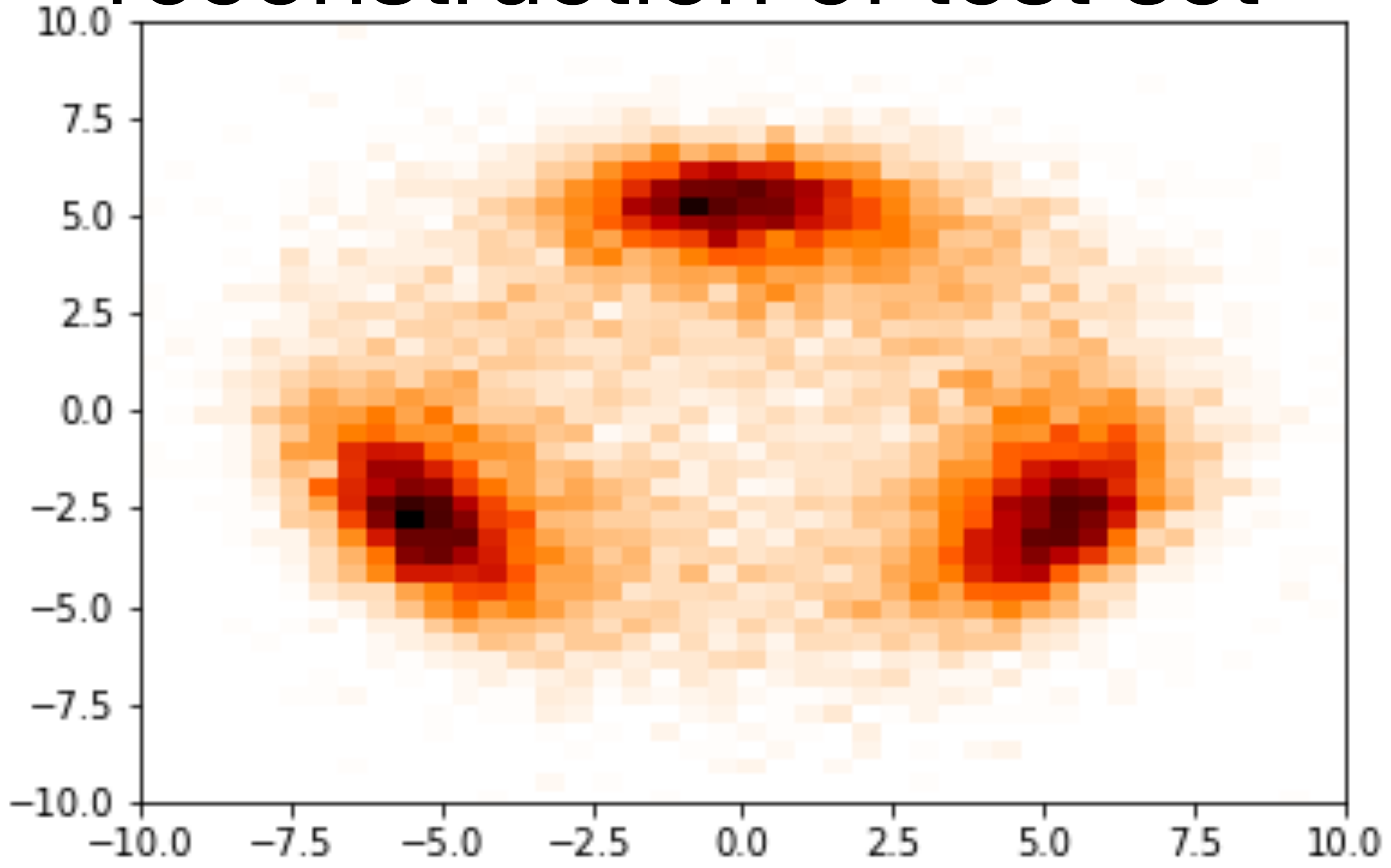


dim = 4:

generative samples

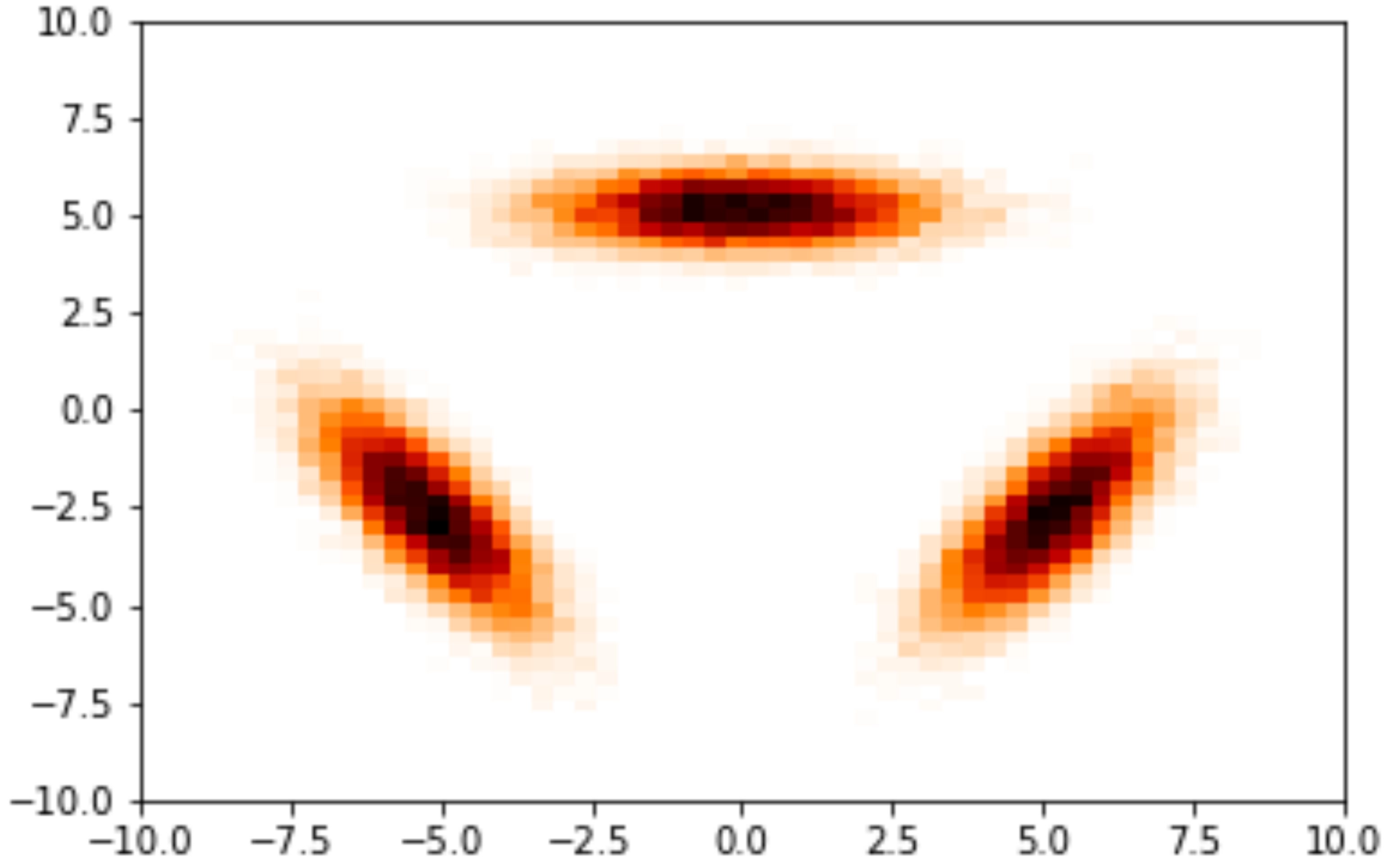


reconstruction of test set



Experiment: mixture of Gaussians

$X|Z$



There are times when it fails...