10-315
Introduction to ML

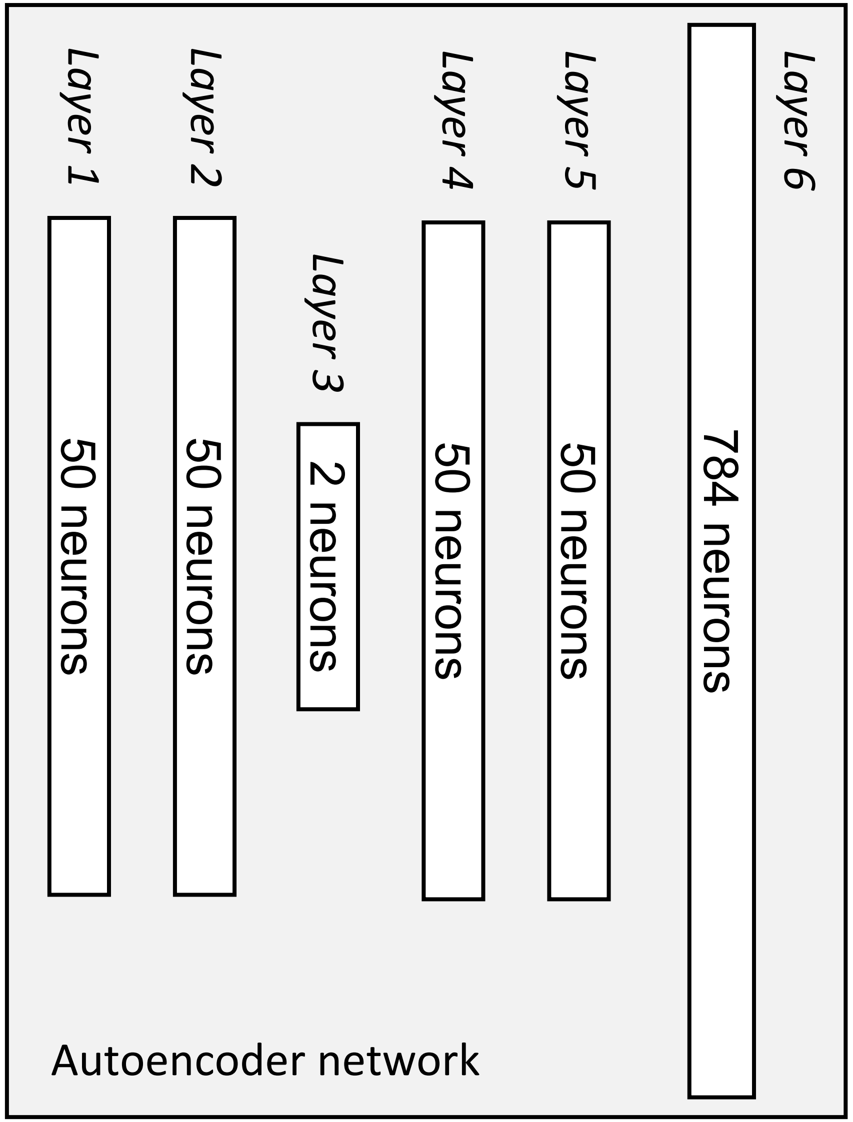
Variational Autoencoders

Instructor: Pat Virtue

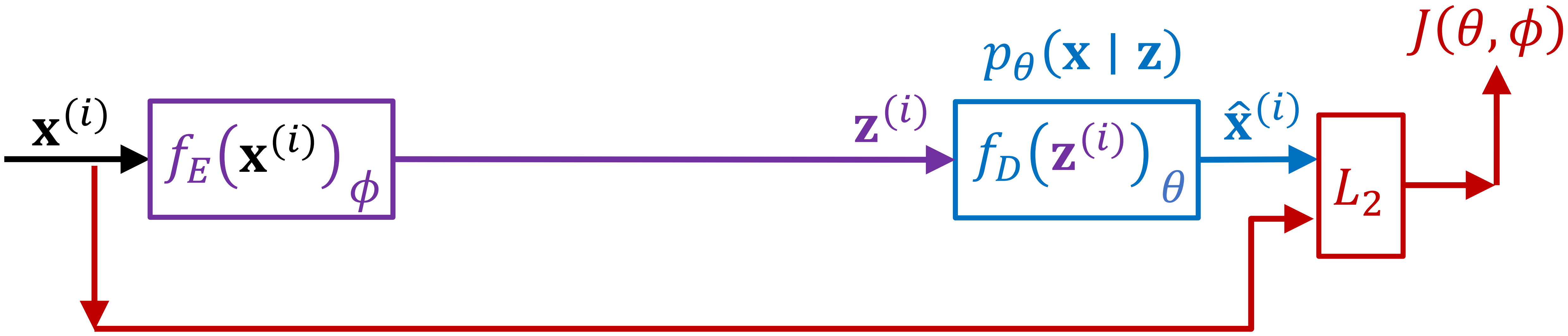
Digit Autoencoder

https://cs.stanford.edu/people/karpathy/convnetjs/demo/autoencoder.html

Input
28x28 = 784

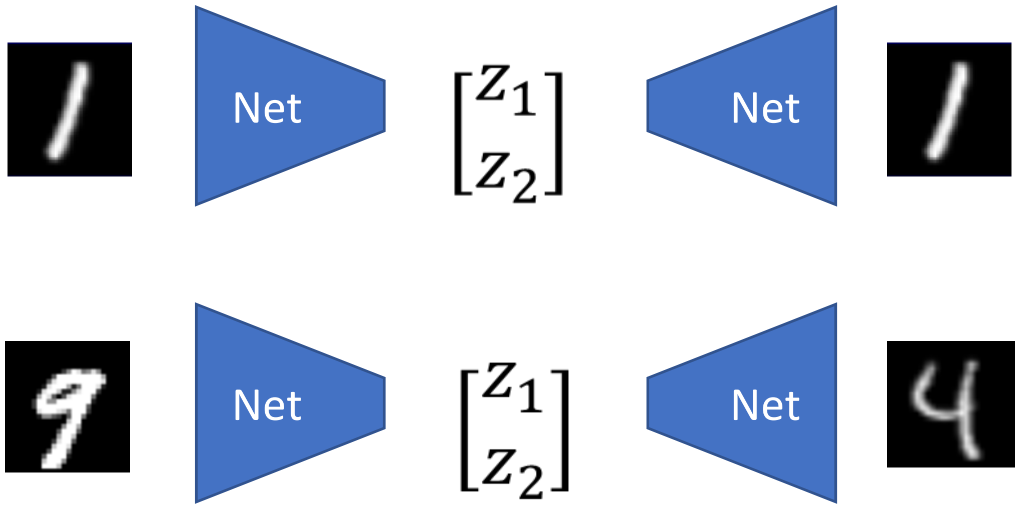


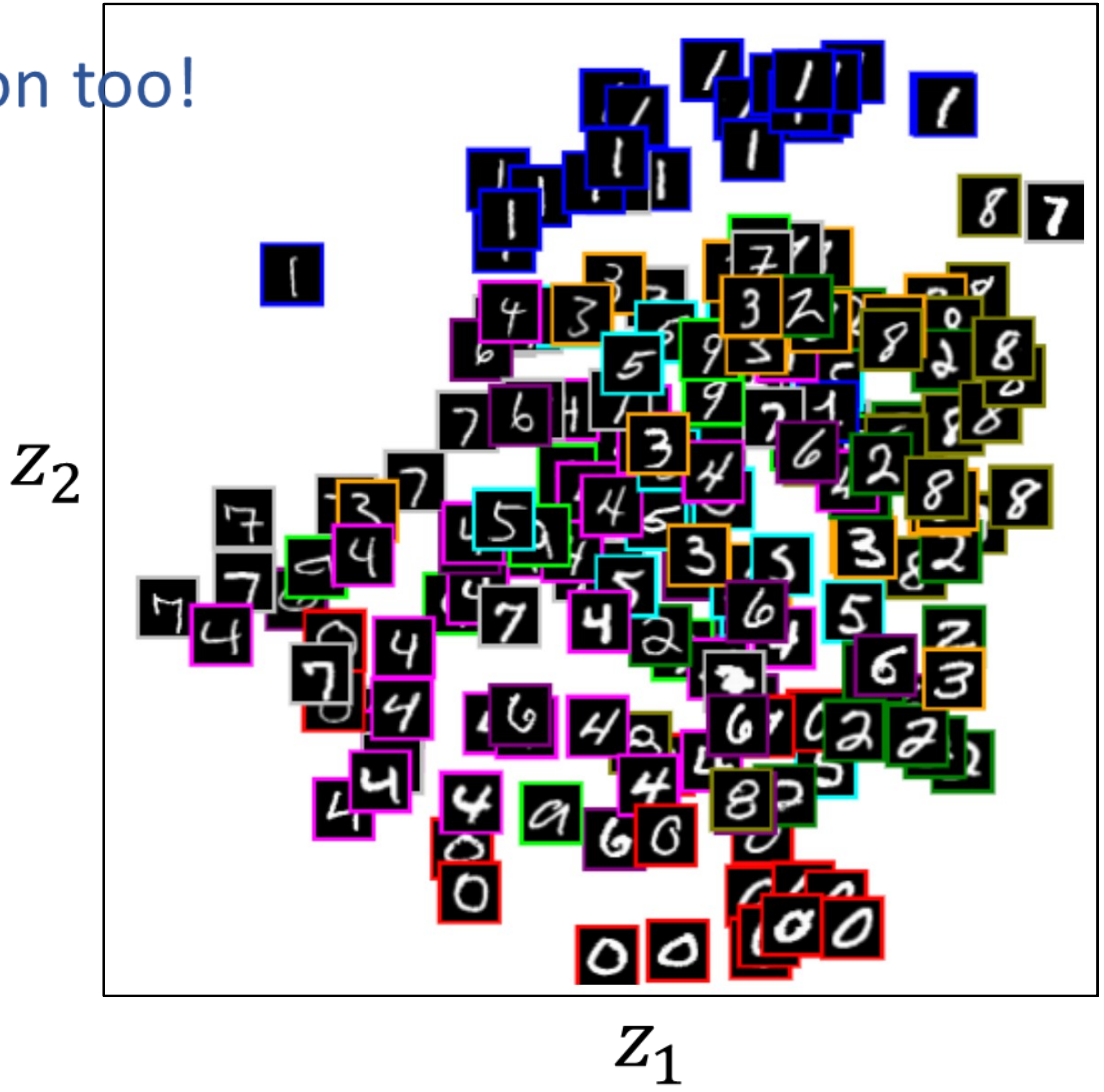
Autoencoder



Autoencoder

Neural networks can learn to organization too!



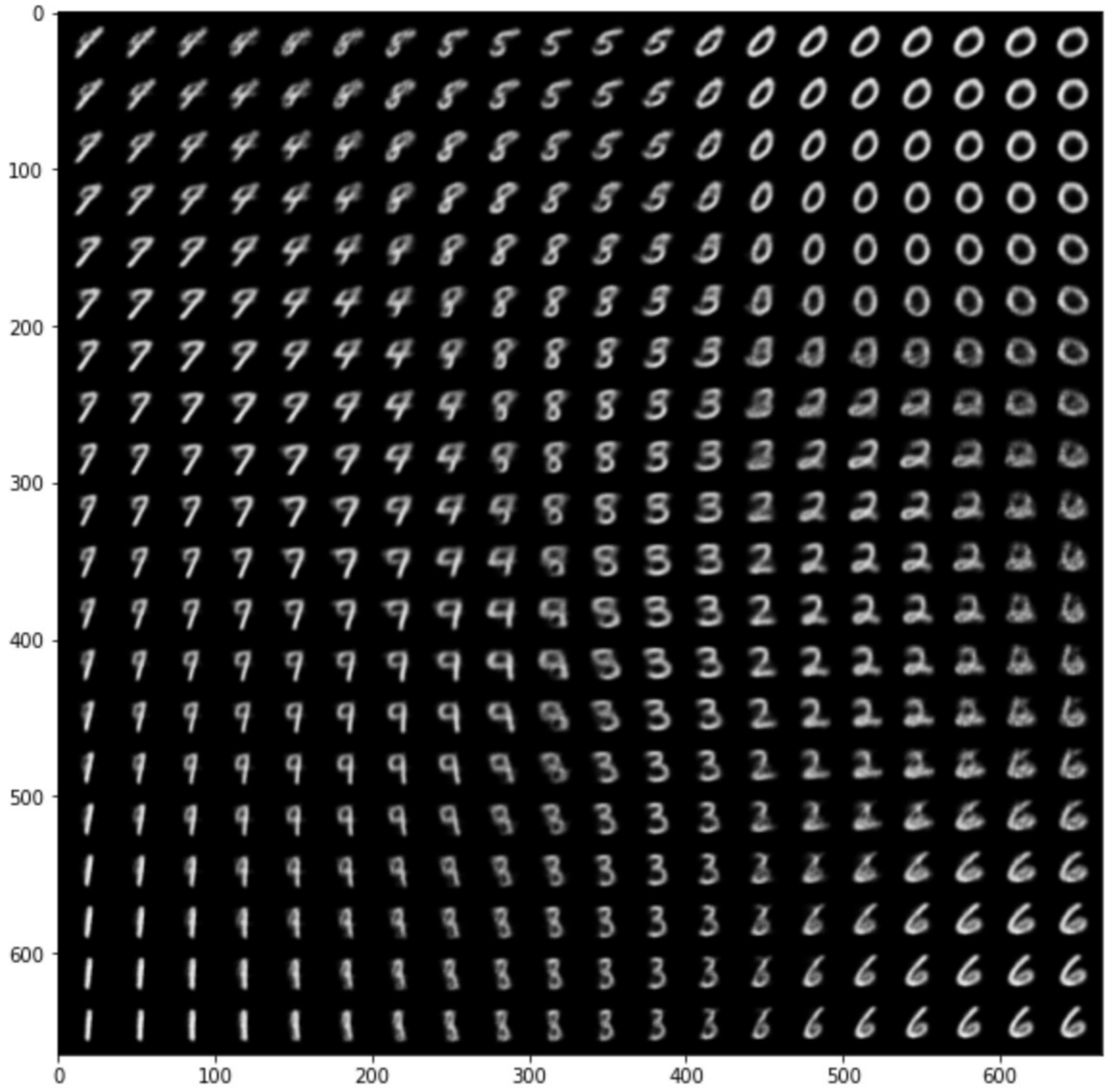


https://cs.stanford.edu/people/karpathy/convnetjs/demo/autoencoder.html

Variational Autoencoder Demo

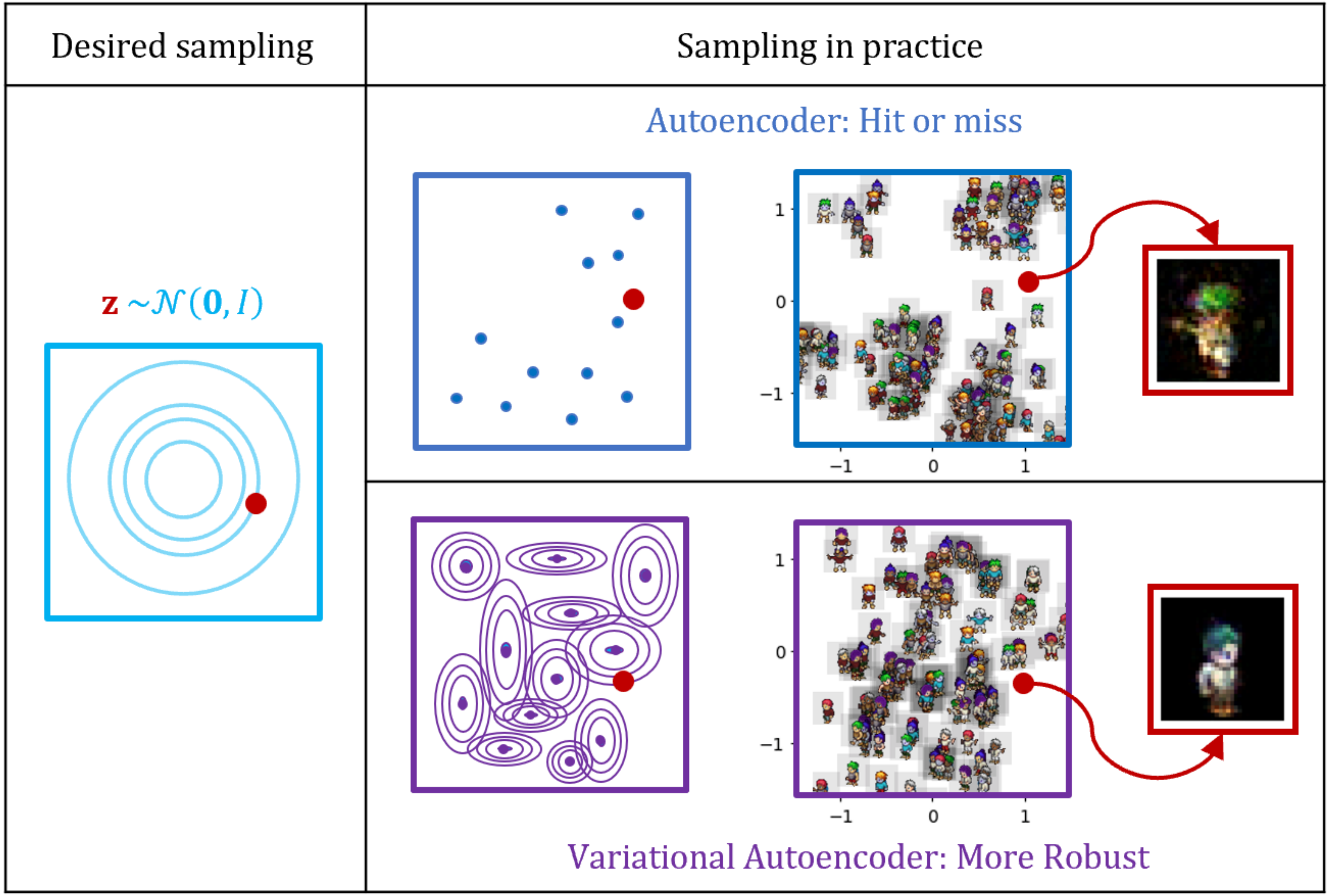
Zhuoyue Lyu, Safinah Ali, and Cynthia Breazeal. EAAI 2022.

https://colab.research.google.com/gist/ZhuoyueLyu/5046225a9ae3675cf633c1df5f63be06/digits-interpolation-notebook-eaai.ipynb



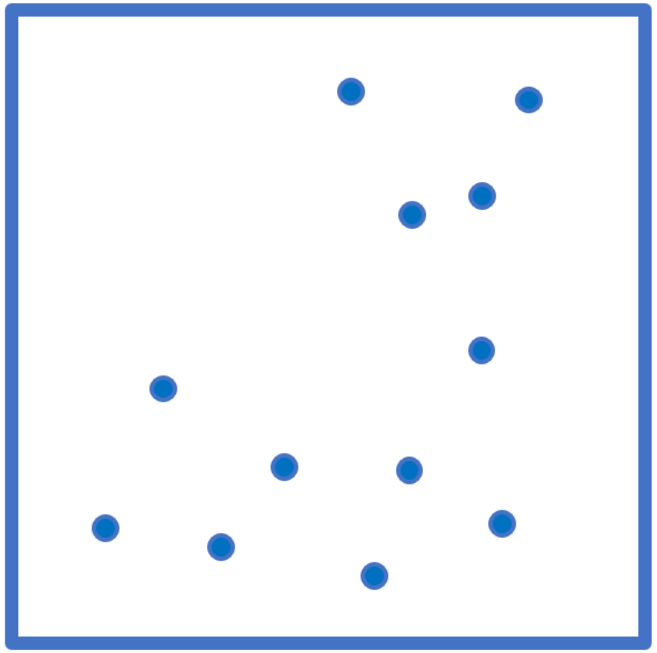
Sampling $p(x)$ from Autoencoder

How do we sample a new image?



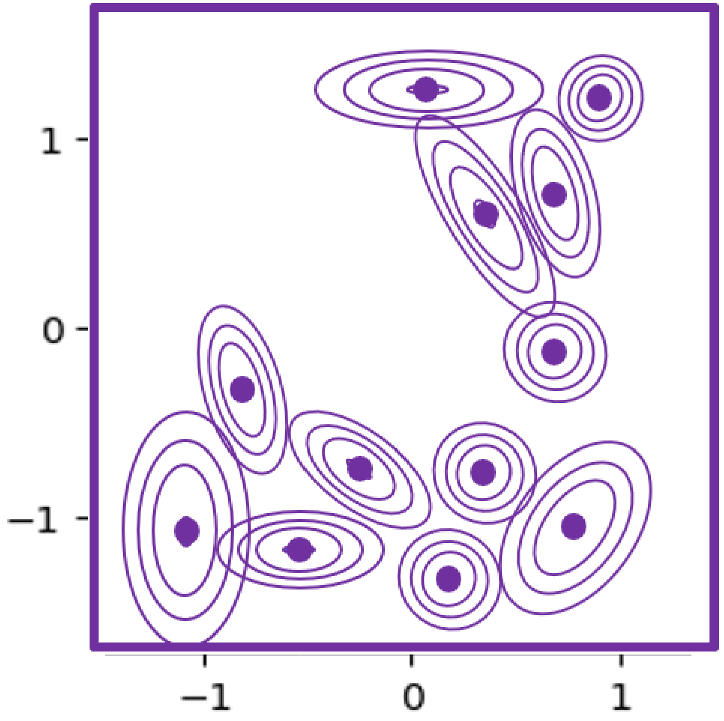
Dataset source: https://github.com/YingzhenLi/Sprites

From Autoencoders to VAE



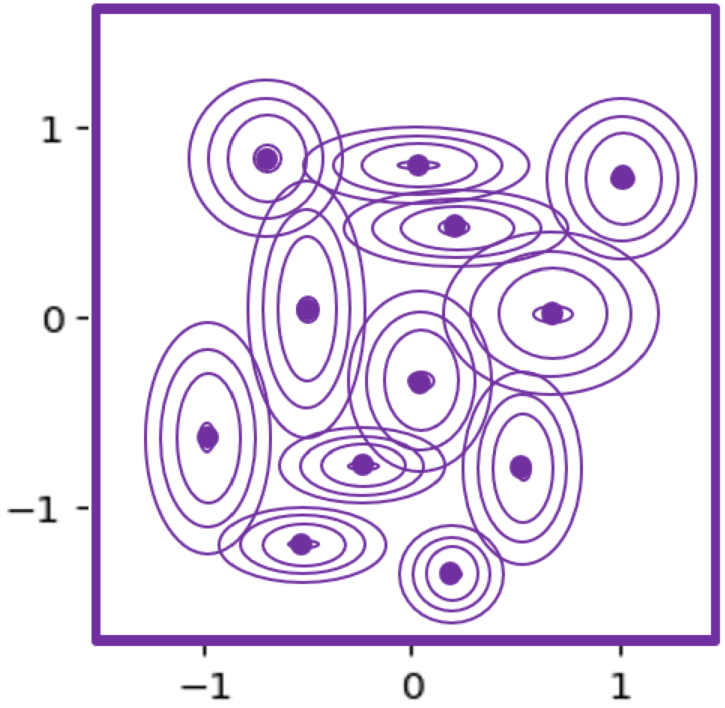
Higher KL divergence

$D_{KL}\left(\mathcal{N}\left(\boldsymbol{\mu}^{(i)}, \Sigma^{(i)}\right) \parallel \mathcal{N}(\mathbf{0}, \mathrm{I})\right)$

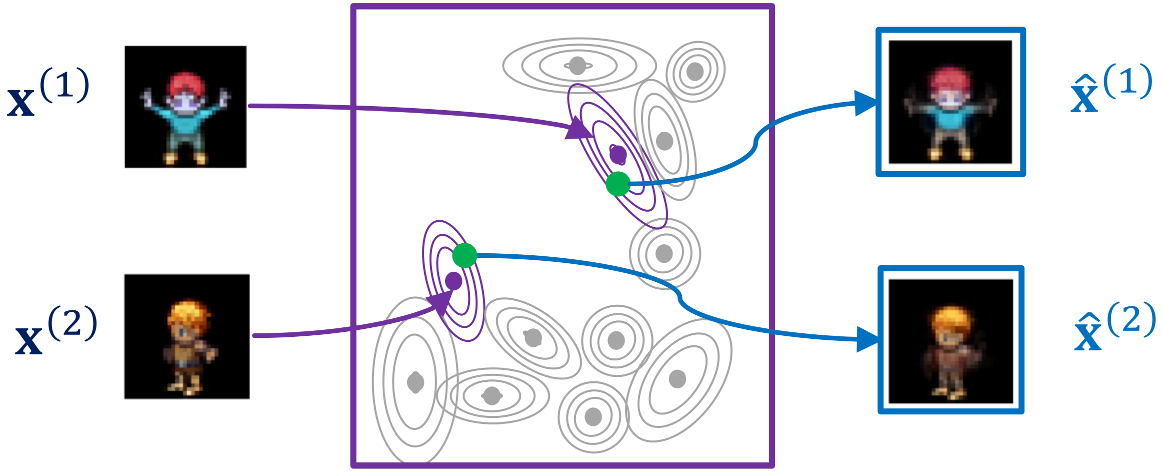


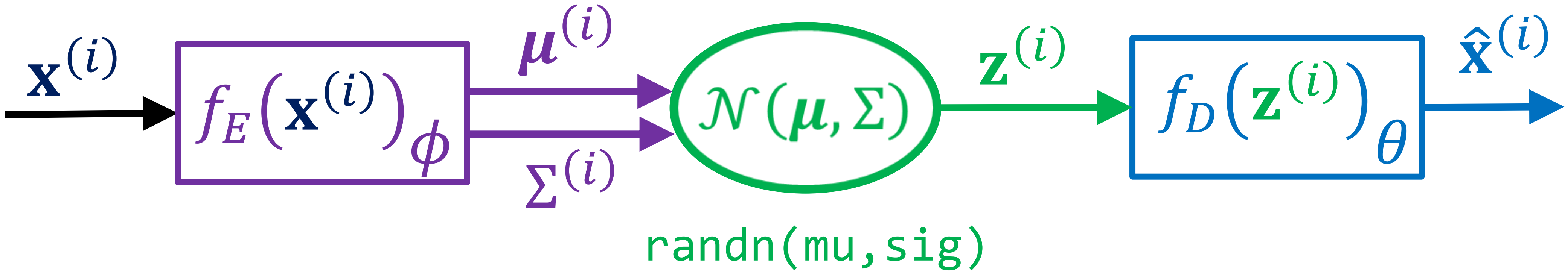
Lower KL divergence

$D_{KL}\left(\mathcal{N}\left(\boldsymbol{\mu}^{(i)}, \Sigma^{(i)}\right) \parallel \mathcal{N}(\mathbf{0}, \mathrm{I})\right)$



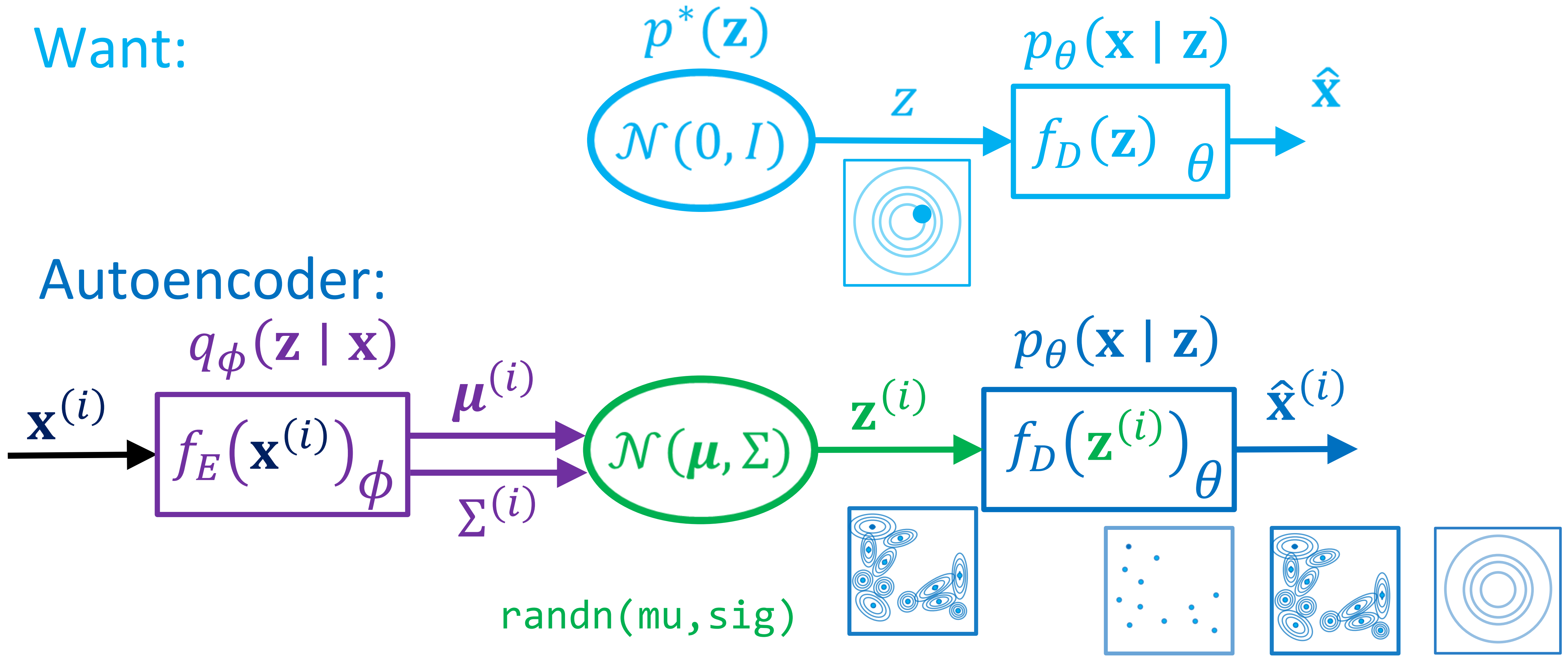
From Autoencoders to VAE





Dataset source: https://github.com/YingzhenLi/Sprites

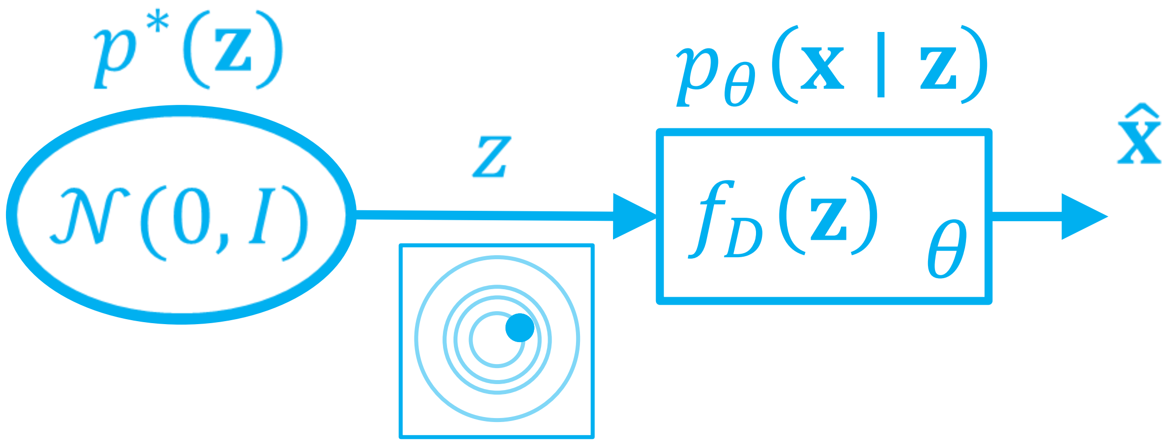
From Autoencoders to VAE



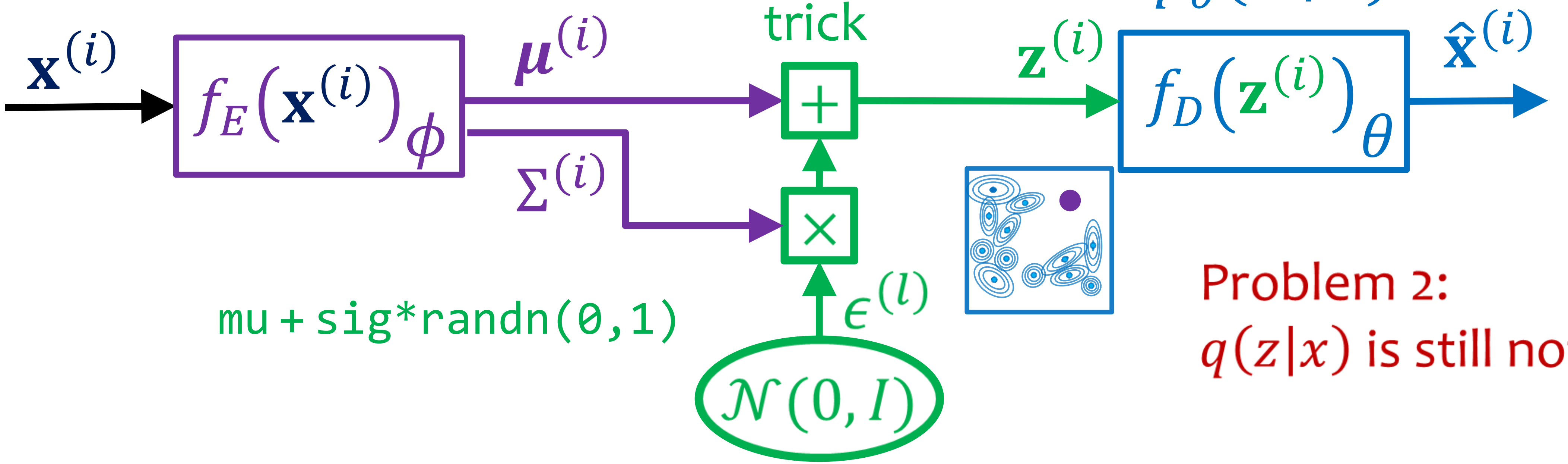
Problem 1: Random sampling function on computation path from ϕ to loss

From Autoencoders to VAE

Want:

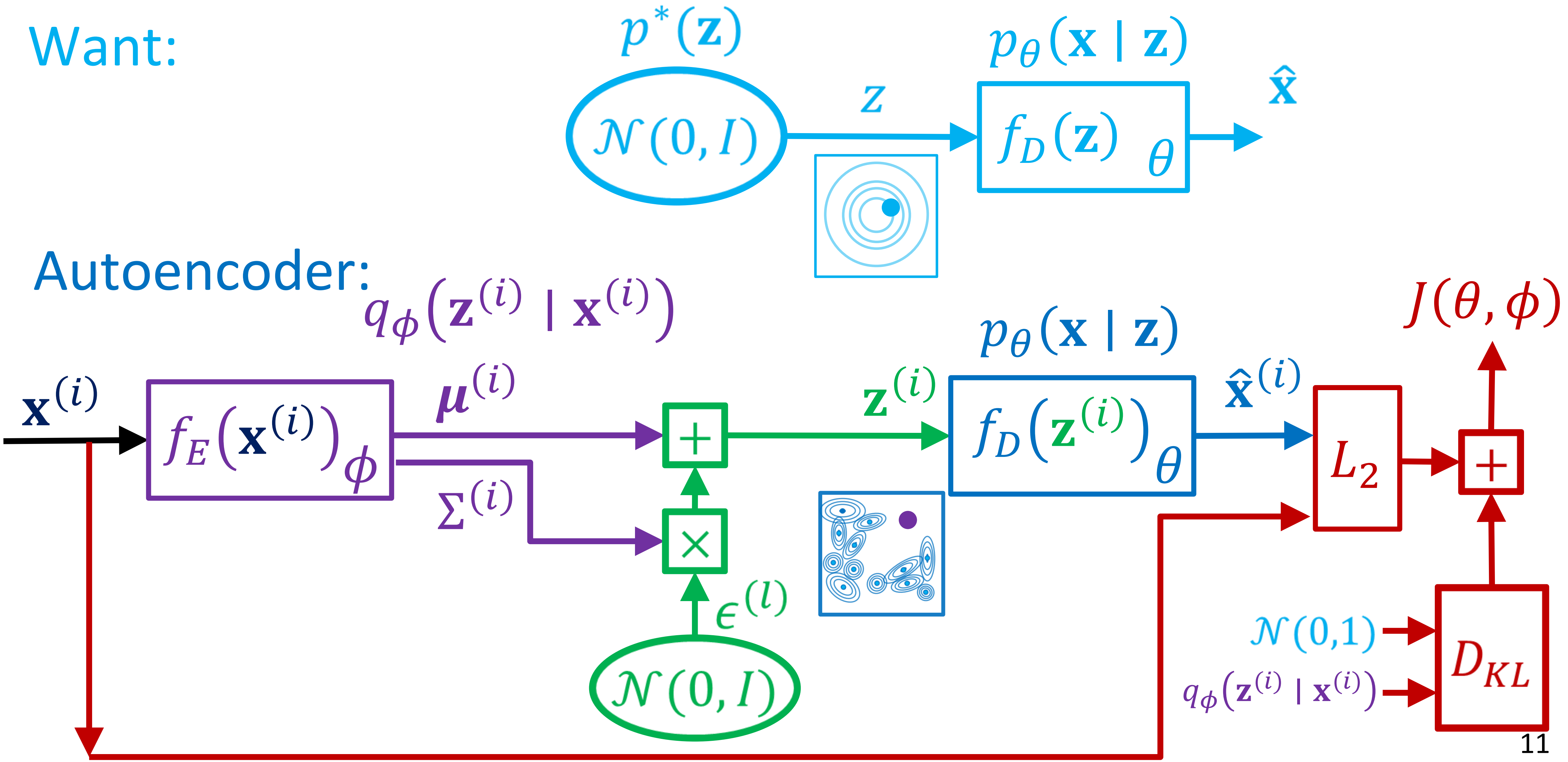


Autoencoder:



Problem 2:
$q(z|x)$ is still not $\mathcal{N}(0, I)$

From Autoencoders to VAE



VAE Objective

MLE:

$$\operatorname*{argmax}_{\theta,\phi} \ \log p(x)$$

$$\log \int_z p_\theta(x \mid z)\, p(z)\, dz$$

Multiply by 1

$$\log \int_z p_\theta(x \mid z)\, p(z) \frac{q_\phi(z \mid x)}{q_\phi(z \mid x)} dz$$

$$\log \mathbb{E}_{z \sim q(z|x)} \left[p_\theta(x \mid z) \frac{p(z)}{q_\phi(z \mid x)} \right]$$

$$\mathbb{E}_{z \sim q(z|x)} \left[\log \left(p_\theta(x \mid z) \frac{p(z)}{q_\phi(z \mid x)} \right) \right]$$

Evidence Lower bound (ELBO)

$$\mathbb{E}_{z \sim q(z|x)}[\log p_\theta(x \mid z)] + \mathbb{E}_{z \sim q(z|x)} \left[\log \frac{p(z)}{q_\phi(z \mid x)} \right]$$

$$\mathbb{E}_{z \sim q(z|x)}[\log p_\theta(x \mid z)] - D_{KL}(q_\phi(z \mid x) \parallel p(z))$$

VAE Objective

MLE:

$$\underset{\theta,\phi}{\operatorname{argmax}} \ \log p(x)$$

$$\log \int_z p_\theta(x \mid z)\, p(z)\, dz$$

Multiply by 1

$$\log \int_z p_\theta(x \mid z)\, p(z) \frac{q_\phi(z \mid x)}{q_\phi(z \mid x)} dz$$

$$\log \mathbb{E}_{z\sim q(z|x)}\left[p_\theta(x \mid z) \frac{p(z)}{q_\phi(z \mid x)}\right]$$

Jensen's inequality:

For convex $f(x)$:

$$f(a+b) \leq f(a) + f(b)$$

$$-\log(a+b) \leq -\log a - \log b$$

$$\log(a+b) \geq \log a + \log b$$

$$\mathbb{E}_{z\sim q(z|x)}\left[\log\left(p_\theta(x \mid z) \frac{p(z)}{q_\phi(z \mid x)}\right)\right]$$

Evidence Lower bound (ELBO)

$$\mathbb{E}_{z\sim q(z|x)}[\log p_\theta(x \mid z)] + \mathbb{E}_{z\sim q(z|x)}\left[\log \frac{p(z)}{q_\phi(z \mid x)}\right]$$

$$\mathbb{E}_{z\sim q(z|x)}[\log p_\theta(x \mid z)] - D_{KL}\big(q_\phi(z \mid x) \,\|\, p(z)\big)$$

Regular Autoencoder Objective

MLE:

$$\underset{\theta,\phi}{\operatorname{argmin}} - \log p(\mathcal{D} \mid \theta, \phi)$$

$$-\sum_{i=1}^{N} \log p\left(\mathbf{x}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$$-\sum_{i=1}^{N} \sum_{j}^{M} \log p\left(x_j^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$$-\sum_{i=1}^{N} \sum_{j}^{M} \log \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x_j - \hat{x}_j)^2}{2\sigma^2}}$$

$$\sum_{i=1}^{N} \sum_{j}^{M} \left(x_j - \hat{x}_j\right)^2$$

$$\hat{\mathbf{x}} = h_{\theta,\phi}(\mathbf{x}) = f_D(\mathbf{z}; \theta)$$

$$\mathbf{z} = f_E(\mathbf{x}; \phi)$$

Regression: model each pixel as noisy samples of $h(\mathbf{x})$

$$p(\mathbf{x} \mid \mathbf{x}) = \prod_j p(x_j \mid \mathbf{x})$$

$$p(x_j \mid \mathbf{x}) \sim N(x_j, \sigma^2)$$

$$loss(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 = \|\mathbf{x} - f_D(\mathrm{f}_E(\mathbf{x}))\|_2^2$$

VAE Objective

MLE:

$$\underset{\theta,\phi}{\operatorname{argmin}} - \log\, p(\mathcal{D} \mid \theta, \phi)$$

$$-\sum_{i=1}^{N} \left(\mathbb{E}_{\mathbf{z}\sim q(\mathbf{z}|\mathbf{x})}\left[\log\, p_\theta\left(\mathbf{x}^{(i)} \mid \mathbf{z}^{(i)}\right)\right] - D_{KL}\left(q_\phi\left(\mathbf{z}^{(i)} \mid \mathbf{x}^{(i)}\right) \parallel \mathcal{N}(0,1)\right)\right)$$

$$\sum_{i=1}^{N} \left(\frac{1}{L} \sum_{l=1}^{L} -\log\, p_\theta\left(\mathbf{x}^{(i)} \mid \mathbf{z}^{(i,l)}\right) \quad + D_{KL}\left(q_\phi\left(\mathbf{z}^{(i)} \mid \mathbf{x}^{(i)}\right) \parallel \mathcal{N}(0,1)\right)\right)$$

$$\sum_{i=1}^{N} \left(\frac{1}{L} \sum_{l=1}^{L} \left\| \mathbf{x}^{(i)} - f_G\left(\mathbf{z}^{(i,l)}; \theta\right)\right\|_2^2 \quad + D_{KL}\left(q_\phi\left(\mathbf{z}^{(i)} \mid \mathbf{x}^{(i)}\right) \parallel \mathcal{N}(0,1)\right)\right)$$

where we sample $\mathbf{z}^{(i,l)} \sim \mathcal{N}\left(\boldsymbol{\mu}^{(\mathrm{i})}, \Sigma^{(i)}\right)$ and $\boldsymbol{\mu}^{(\mathrm{i})}, \Sigma^{(i)} = f_E\left(\mathbf{x}^{(i)}; \phi\right)$