

Siddhant Mishra-Sharma (MIT/AI FI) Summer School

166

3

3

Towards diffusion: (Markovian) hierarchical VAEs

$$p\left(z_{T-1} \mid z_T\right) \quad p\left(z_3 \mid z_4\right) \quad p\left(z_2 \mid z_3\right) \quad p\left(z_1 \mid z_2\right) \quad p\left(x \mid z_1\right)$$

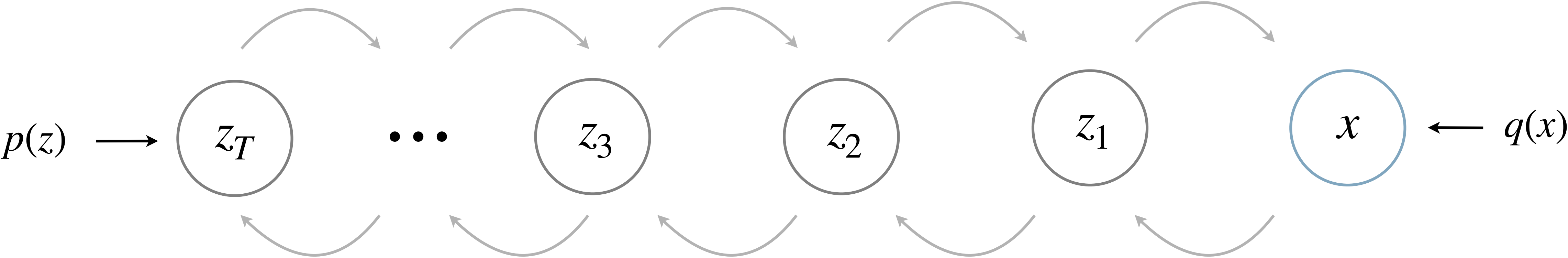
$$q\left(z_T \mid z_{T-1}\right) \quad q\left(z_4 \mid z_3\right) \quad q\left(z_3 \mid z_2\right) \quad q\left(z_2 \mid z_1\right) \quad q\left(z_1 \mid x\right)$$

Reverse process

$$p(x, z_1, z_2, \dots, z_T) = p(z_T) p(z_{T-1} | z_T) \cdots p(z_1 | z_2) p(x | z_1)$$

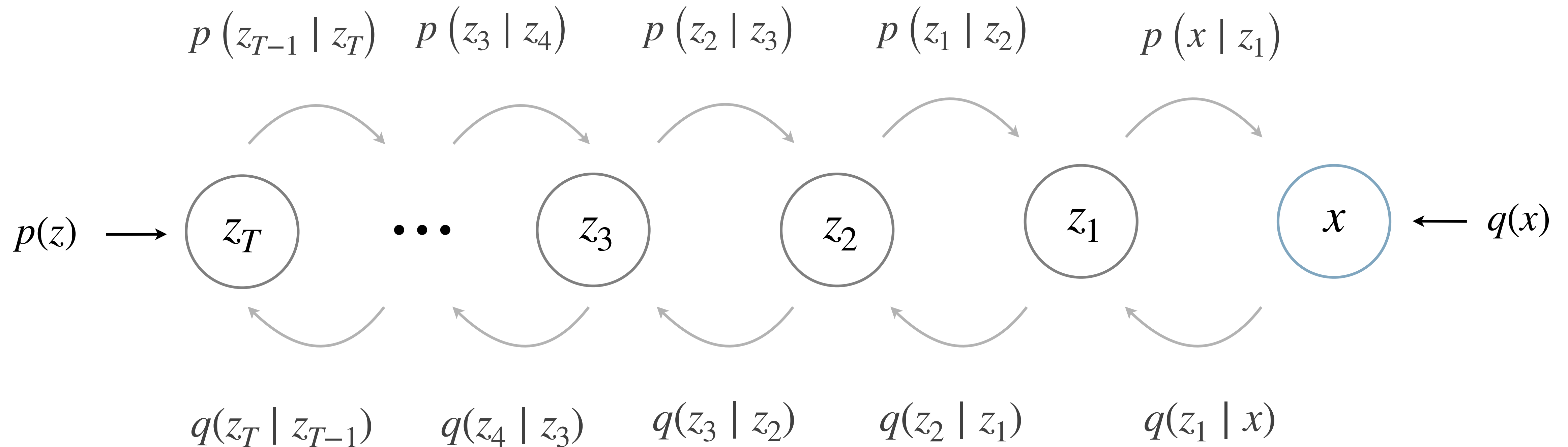
Forward process

$$q(x, z_1, z_2, \dots, z_T) = q(x) q(z_1 | x) q(z_2 | z_1) \cdots q(z_T | z_{T-1})$$



Towards diffusion: (Markovian) hierarchical VAEs

Reverse process $p(x, z_1, z_2, \dots, z_T) = p(z_T) p(z_{T-1} | z_T) \cdots p(z_1 | z_2) p(x | z_1)$



Forward process $q(x, z_1, z_2, \dots, z_T) = q(x) q(z_1 | x) q(z_2 | z_1) \cdots q(z_T | z_{T-1})$

Towards diffusion: (Markovian) hierarchical VAEs

Diffusion models can be seen as hierarchical VAEs with a few restrictions:

- The forward (*encoding*) distribution prescribed as a Markov chain of Gaussians; it is not learned

$$q(z_t | z_{t-1}) = \mathcal{N}(z_t; \alpha_t z_{t-1}, \beta_t)$$

- Distributions of latents at the final timestep T is a standard (unit) Gaussian

$$q(z_T | z_{T-1}, \dots, x) = \mathcal{N}(z_T; 0, \mathbb{I})$$

- The dimensionality of latents is the same as the data dimensionality

$$\dim(z_t) = \dim(x)$$

