/70

35

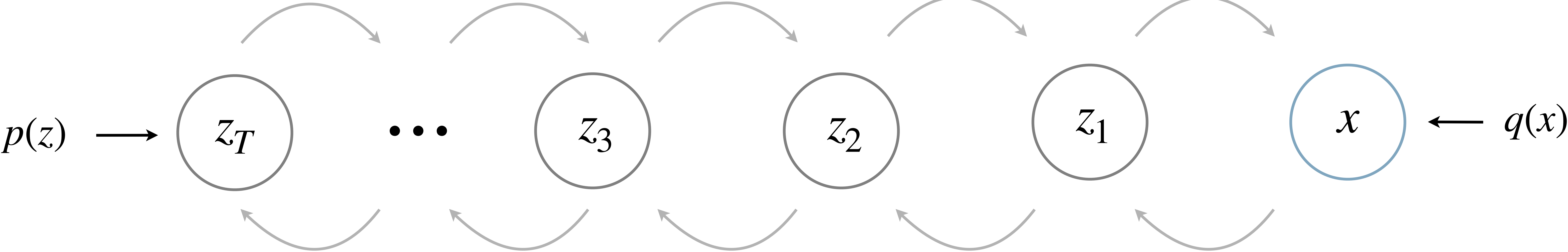
Towards diffusion: (Markovian) hierarchical VAEs

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

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

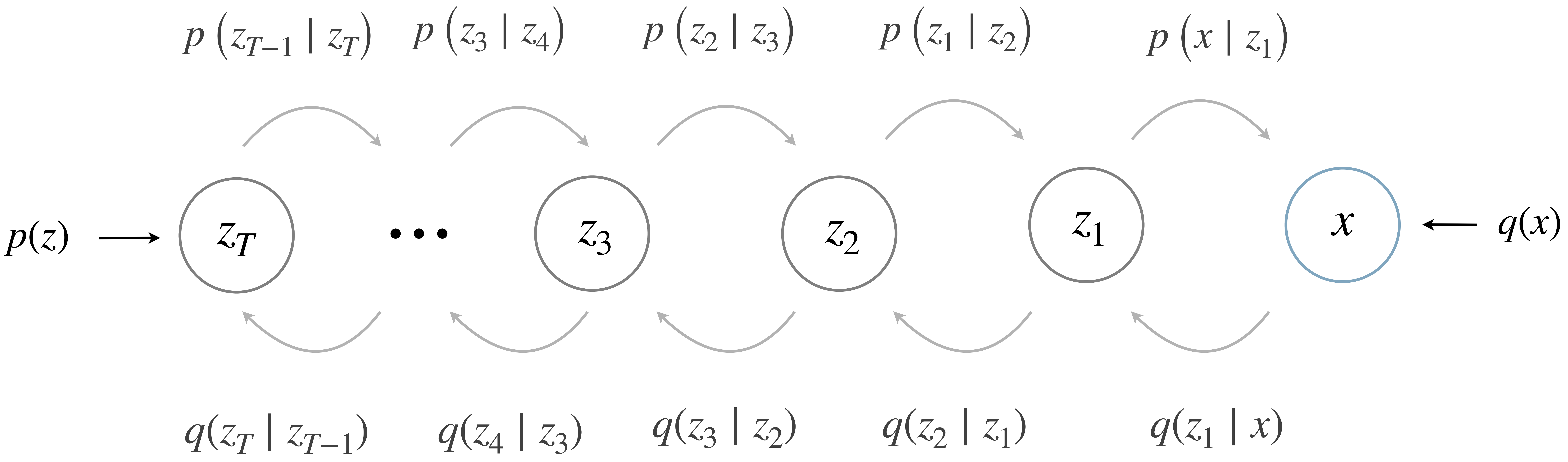
Reverse process $$p\left(x, z_1, z_2, \cdots, z_T\right) = p\left(z_T\right) p\left(z_{T-1} \mid z_T\right) \cdots p\left(z_1 \mid z_2\right) p\left(x \mid z_1\right)$$

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


$p(z)$
z_T
$\cdots$
z_3
z_2
z_1
x
$q(x)$

Towards diffusion: (Markovian) hierarchical VAEs

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



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

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

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

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

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

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

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