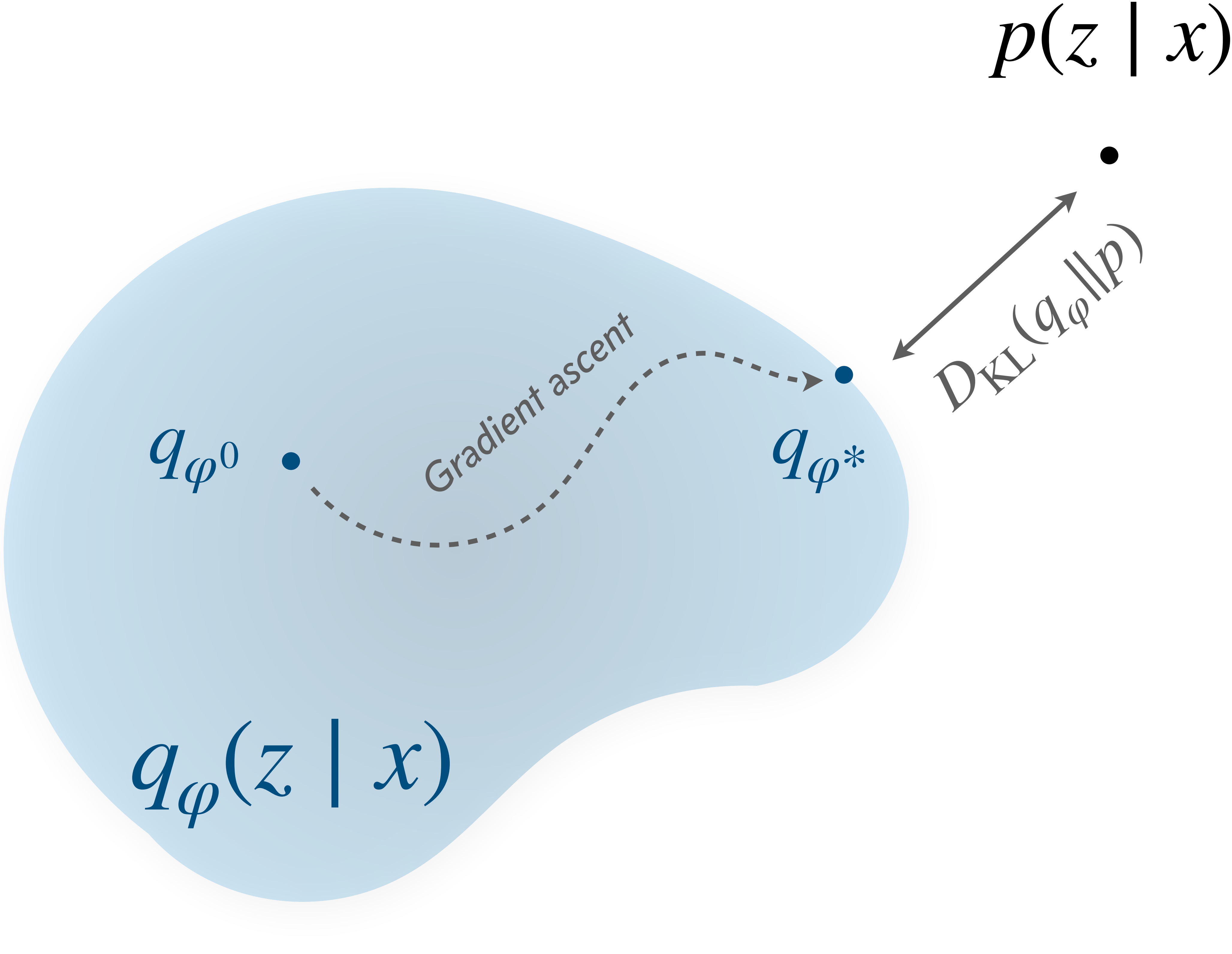
18

Variational inference

The intractability of $p(x)$ is closely related to the intractability of the *posterior* $p(z \mid x)$

$$p(z \mid x) = \frac{p(x, z)}{p(x)}$$

(Bayes' theorem)


$p(z \mid x)$
$D_{KL}(q_\varphi || p)$
Gradient ascent
q_{φ^0}
q_{φ^*}
$q_\varphi(z \mid x)$

$$\overbrace{D_{\mathrm{KL}}\left(q_\varphi(z)\|p(z\mid x)\right)}^{\geq 0} = \overbrace{\log p(x)}^{\text{Evidence}} - \overbrace{\left\langle \log p_\vartheta(x,z) - \log q_\varphi(z)\right\rangle_{q_\varphi(z)}}^{\text{Evidence Lower BOund (ELBO)}}$$

A two-for-one!

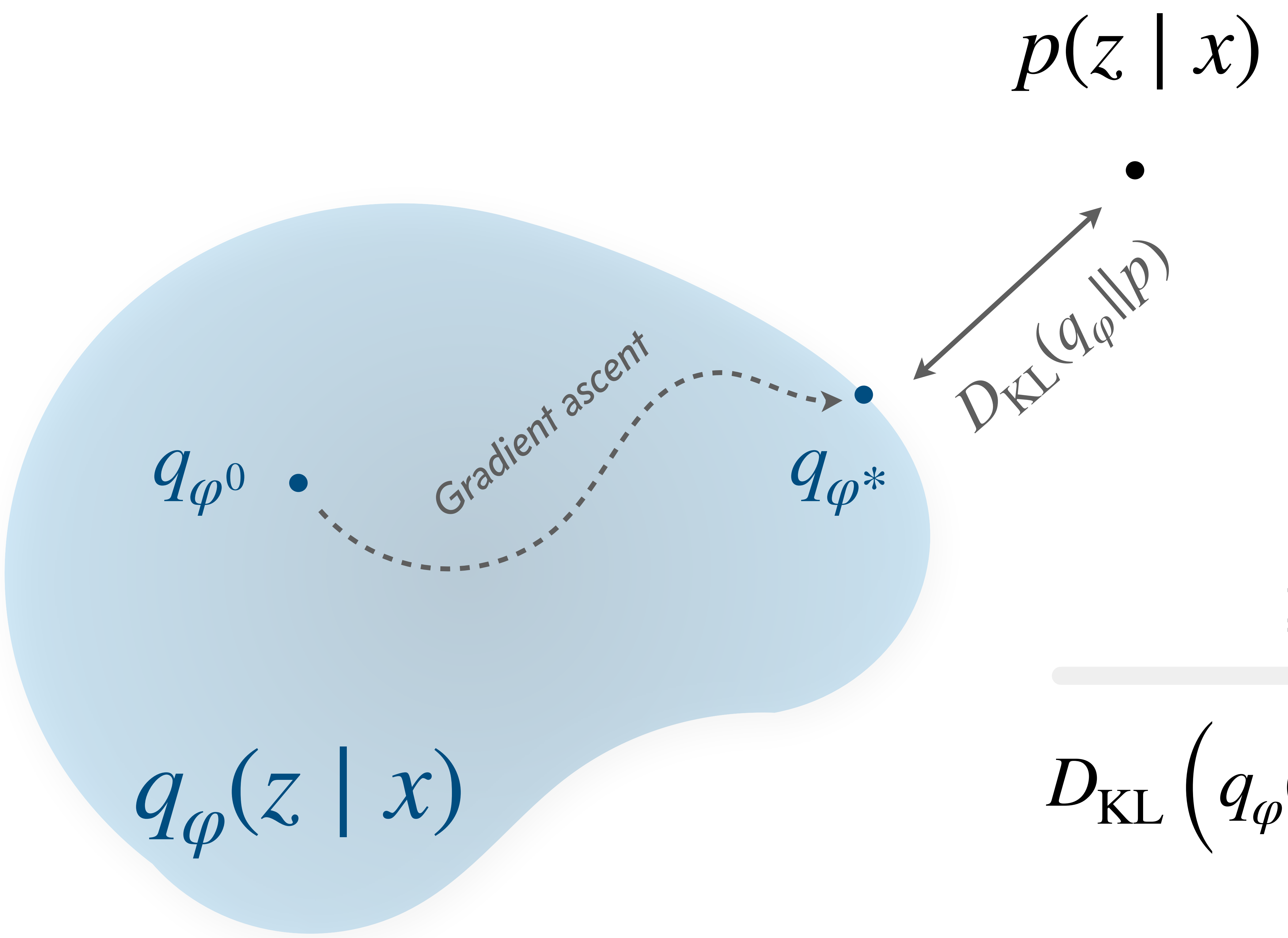
- ☑ Estimate approximate posterior $q_\phi(z) \approx p(z \mid x)$
- ☑ Estimate likelihood/evidence $\mathrm{ELBO} \approx p(x)$

Variational inference

The intractability of $p(x)$ is closely related to the intractability of the *posterior* $p(z \mid x)$

$$p(z \mid x) = \frac{p(x, z)}{p(x)}$$

(Bayes' theorem)



A two-for-one!

- ☑ Estimate approximate posterior $q_\phi(z) \approx p(z \mid x)$
- ☑ Estimate likelihood/evidence $\mathrm{ELBO} \approx p(x)$

$$\underbrace{D_{\mathrm{KL}}\Big(q_\varphi(z) \| p(z \mid x)\Big)}_{\geq 0} = \underbrace{\log p(x)}_{\text{Evidence}} - \underbrace{\Big\langle \log p_\vartheta(x, z) - \log q_\varphi(z) \Big\rangle_{q_\varphi(z)}}_{\text{Evidence Lower BOund (ELBO)}}$$

Variational inference

A general-purpose technique for posterior estimation

$$\underbrace{D_{\mathrm{KL}}\left(q_\varphi(z) \| p(z \mid x)\right)}_{\geq 0} = \underbrace{\log p(x)}_{\text{Evidence}} - \underbrace{\left\langle \log p_\vartheta(x, z) - \log q_\varphi(z) \right\rangle_{q_\varphi(z)}}_{\text{Evidence Lower BOund (ELBO)}}$$