/70

41

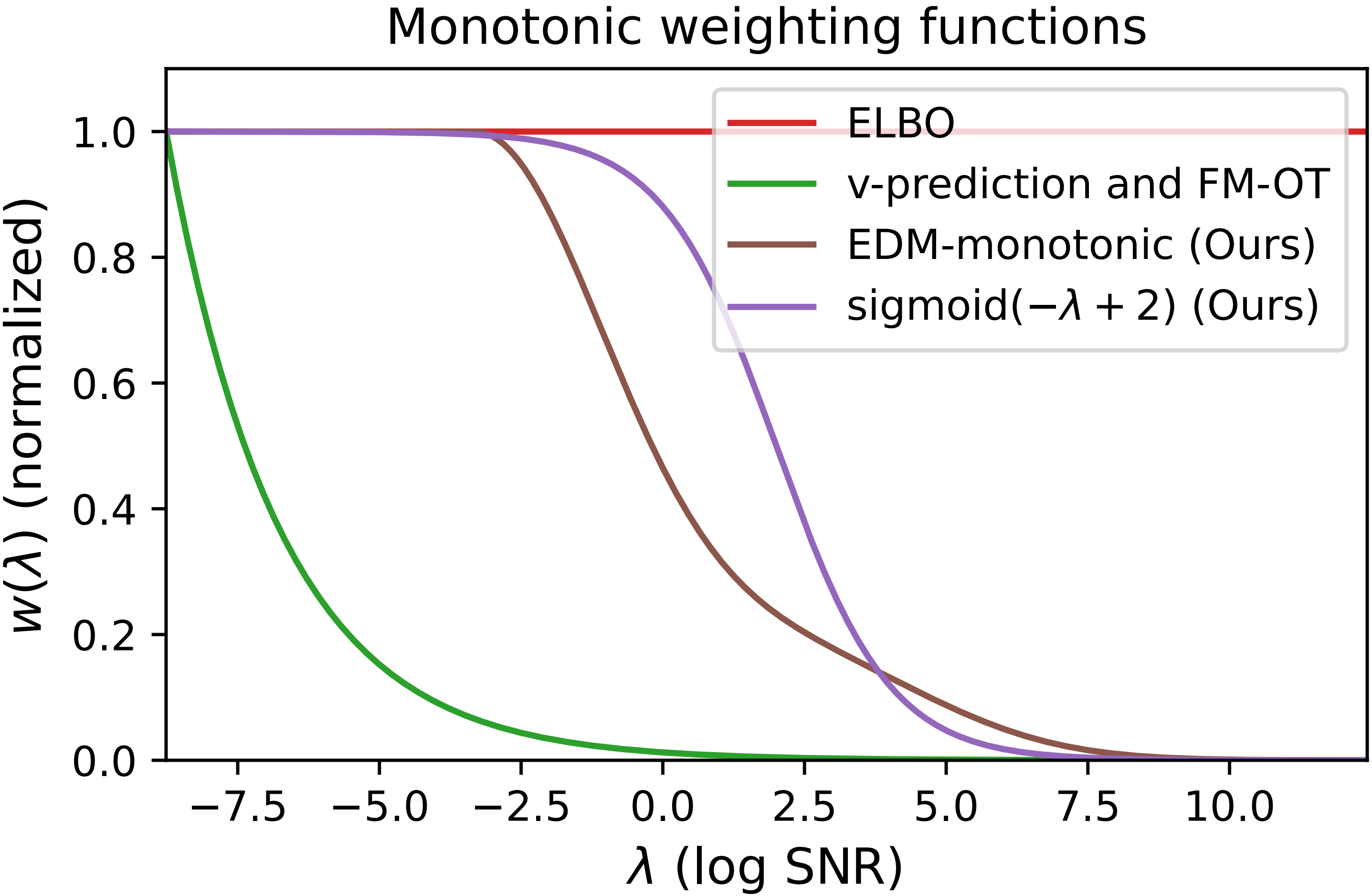
Simple objectives as a weighted sum of ELBOs

Kingma and Gao (2023) showed that common objectives can be written as a weighted sum (across different noise levels) of ELBOs

$$L_w(x) = \left\langle w(t) \cdot w_{\mathrm{ML}}(t) \left\| \epsilon - \hat{\epsilon}_\theta \left(z_t, t \right) \right\|^2 \right\rangle_{t,\epsilon}$$

Additional weighting (w_{ML}^{-1} for ϵ-prediction)

Weighting for ELBO/ ML objective

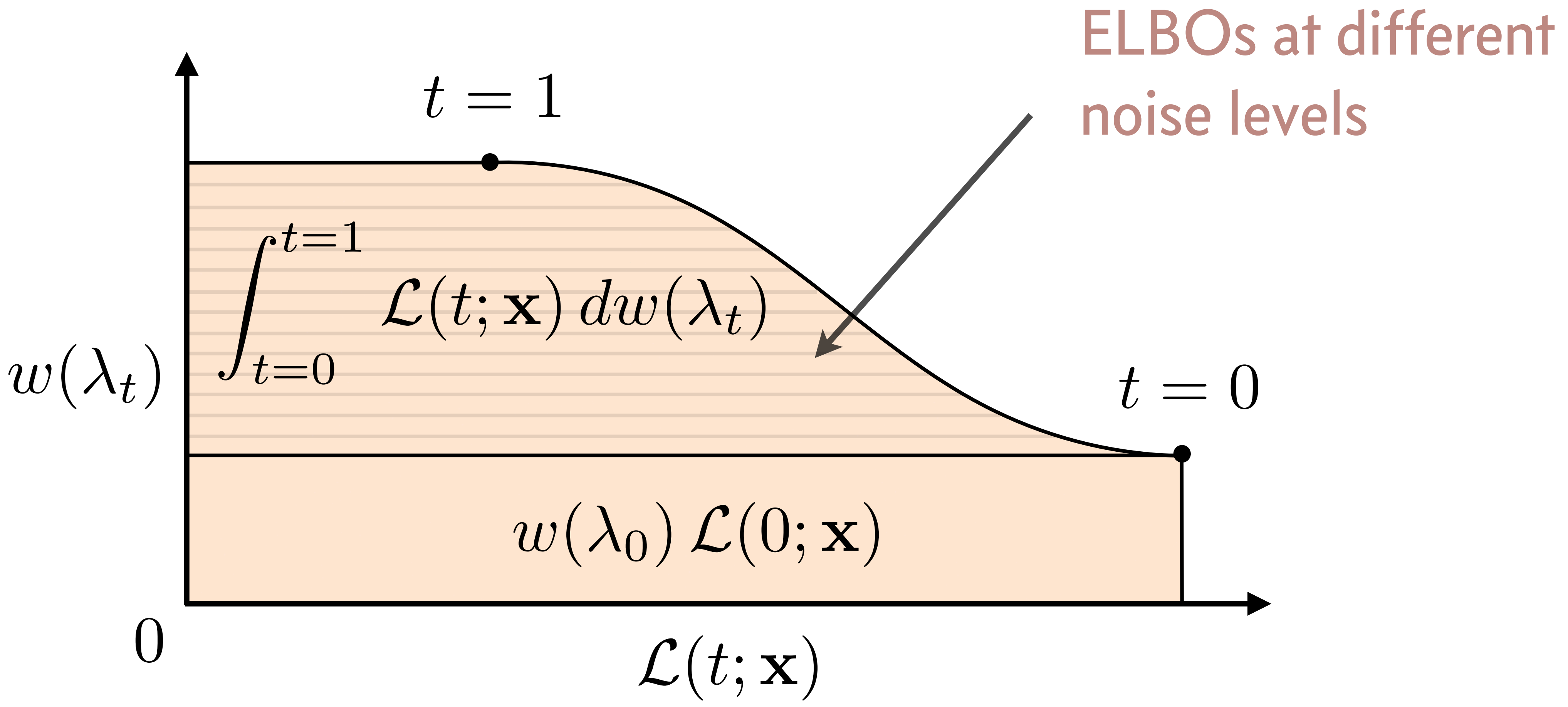

Monotonic weighting functions
ELBO
v-prediction and FM-OT
EDM-monotonic (Ours)
sigmoid($-\lambda + 2$) (Ours)
$w(\lambda)$ (normalized)
λ (log SNR)
1.0
0.8
0.6
0.4
0.2
0.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0

More important
for perceptual
quality?

$$L(t;x) \equiv D_{\mathrm{KL}}\left(q\left(z_{t:1} \mid x\right) \| p\left(z_{t:1}\right)\right)$$

$$L_w(x) \propto \int_0^1 \frac{d}{dt} w(t)\, L(t;x)\, \mathrm{d}t + w(t_0)\, L(0;x)$$

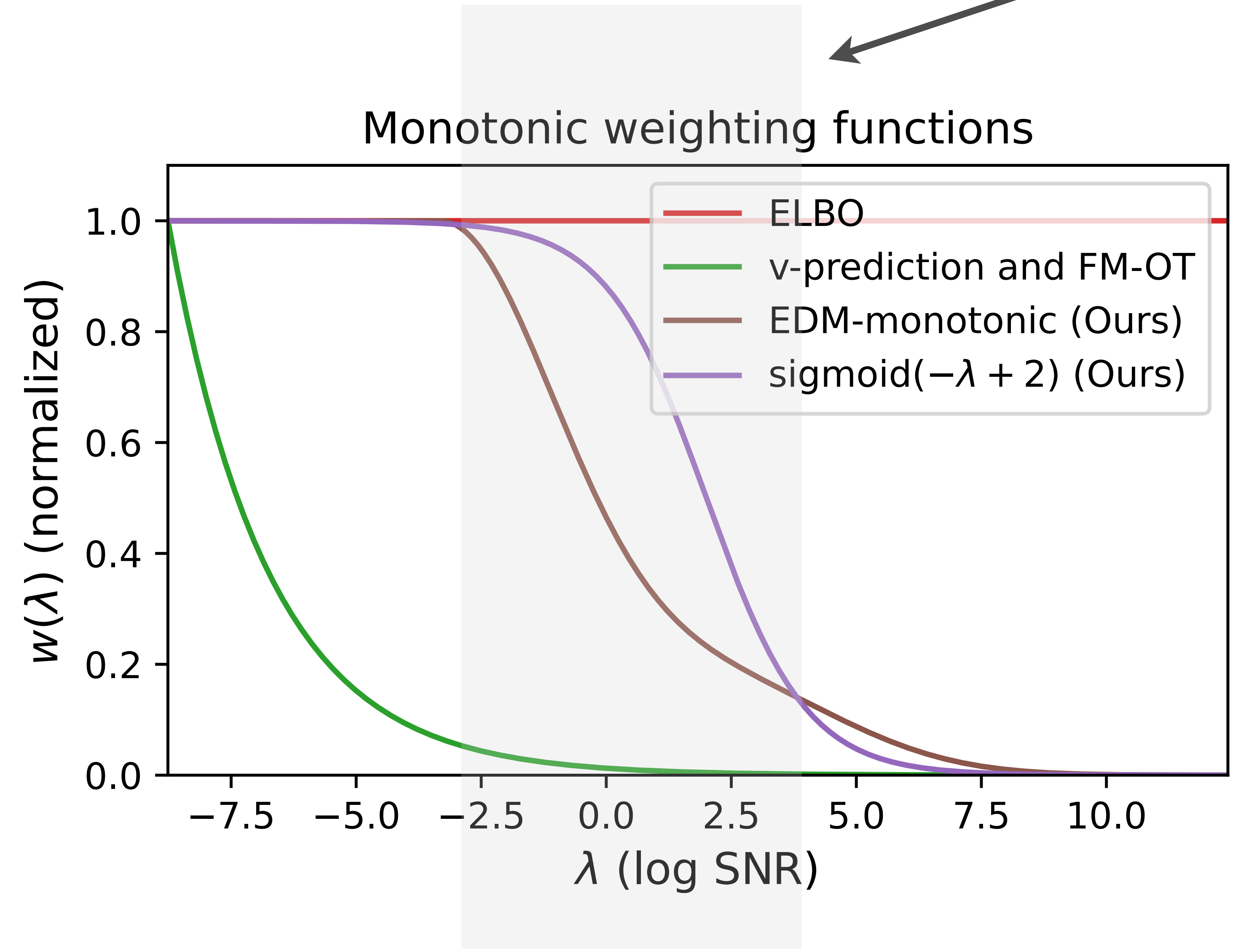
Interpretation: *data augmentation with additive Gaussian noise / data-distribution smoothing*



Simple objectives as a weighted sum of ELBOs

Kingma and Gao (2023) showed that common objectives can be written as a weighted sum (across different noise levels) of ELBOs

More important for perceptual quality?



$$L_w(x) = \left\langle w(t) \cdot w_{\mathrm{ML}}(t) \left\| \epsilon - \hat{\epsilon}_\theta\left(z_t, t\right) \right\|^2 \right\rangle_{t,\epsilon}$$

Additional weighting (w_{ML}^{-1} for ϵ-prediction)

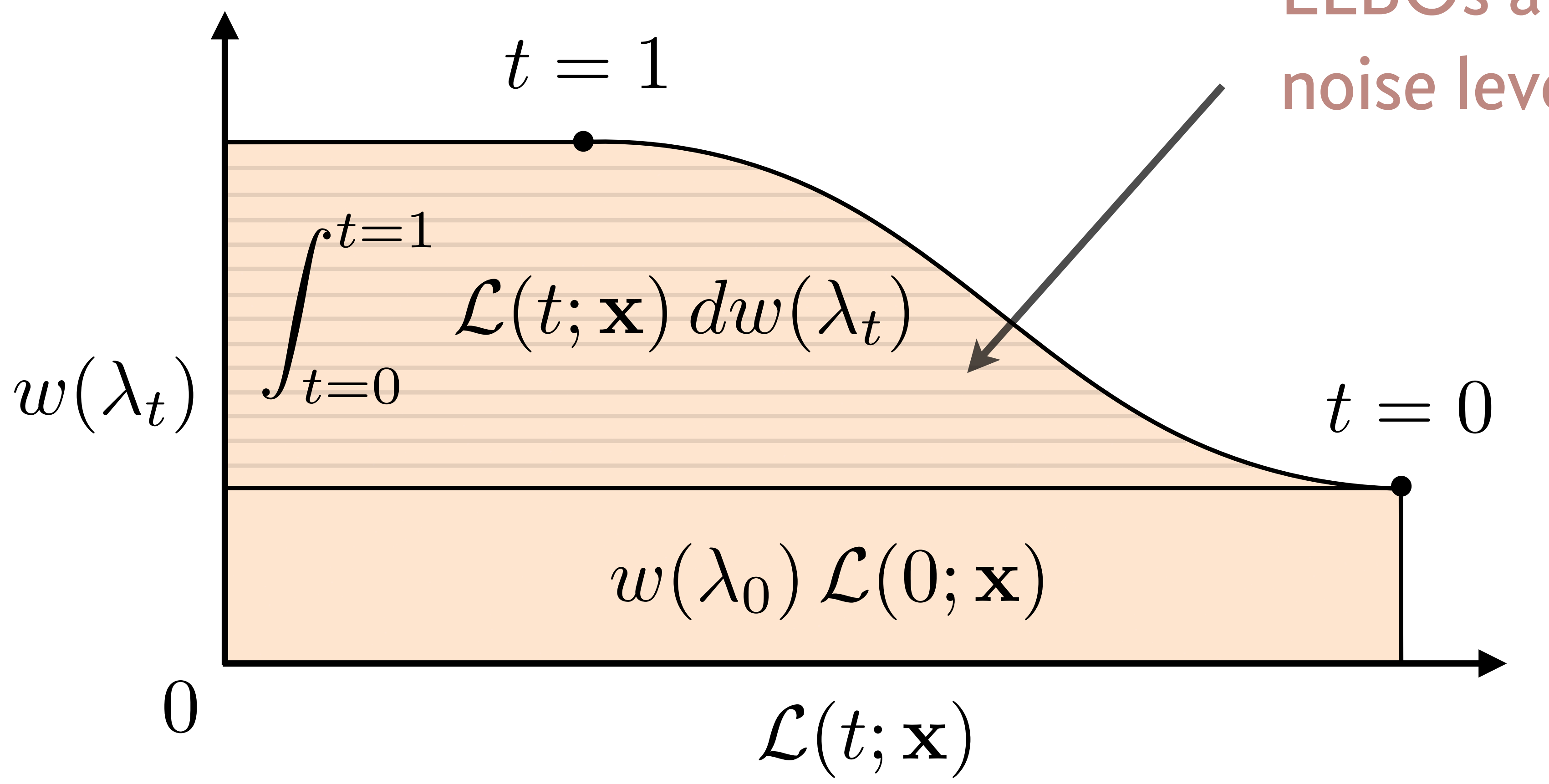
Weighting for ELBO/ ML objective

$$L(t; x) \equiv D_{\mathrm{KL}}\left(q\left(z_{t:1} \mid x\right) \| p\left(z_{t:1}\right)\right)$$

$$L_w(x) \propto \int_0^1 \frac{d}{dt} w(t)\, L(t; x)\, \mathrm{d}t + w(t_0)\, L(0; x)$$

Interpretation: *data augmentation with additive Gaussian noise / data-distribution smoothing*

ELBOs at different noise levels



Continuous-time/SDE formulation

$$x_t = \sqrt{1 - \beta_t} \cdot x_{t-1} + \sqrt{\beta_t} \cdot \varepsilon$$

In the limit of infinite time steps, $\Delta_t \to 0$ and the forward diffusion process can be written as

$$\begin{aligned} x_t &= \sqrt{1 - \beta(t)\Delta_t}x_{t-1} + \sqrt{\beta(t)\Delta_t}\mathcal{N}(0,\mathbb{I}) \\ &\approx x_{t-1} - \frac{\beta(t)\Delta_t}{2}x_{t-1} + \sqrt{\beta(t)\Delta_t}\mathcal{N}(0,\mathbb{I}) \end{aligned}$$