Siddharth Mishra-Sharma (MIT/IAIFI) | IAIFI Summer School

/66

39

Simple objectives as a weighted sum of ELBOs

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

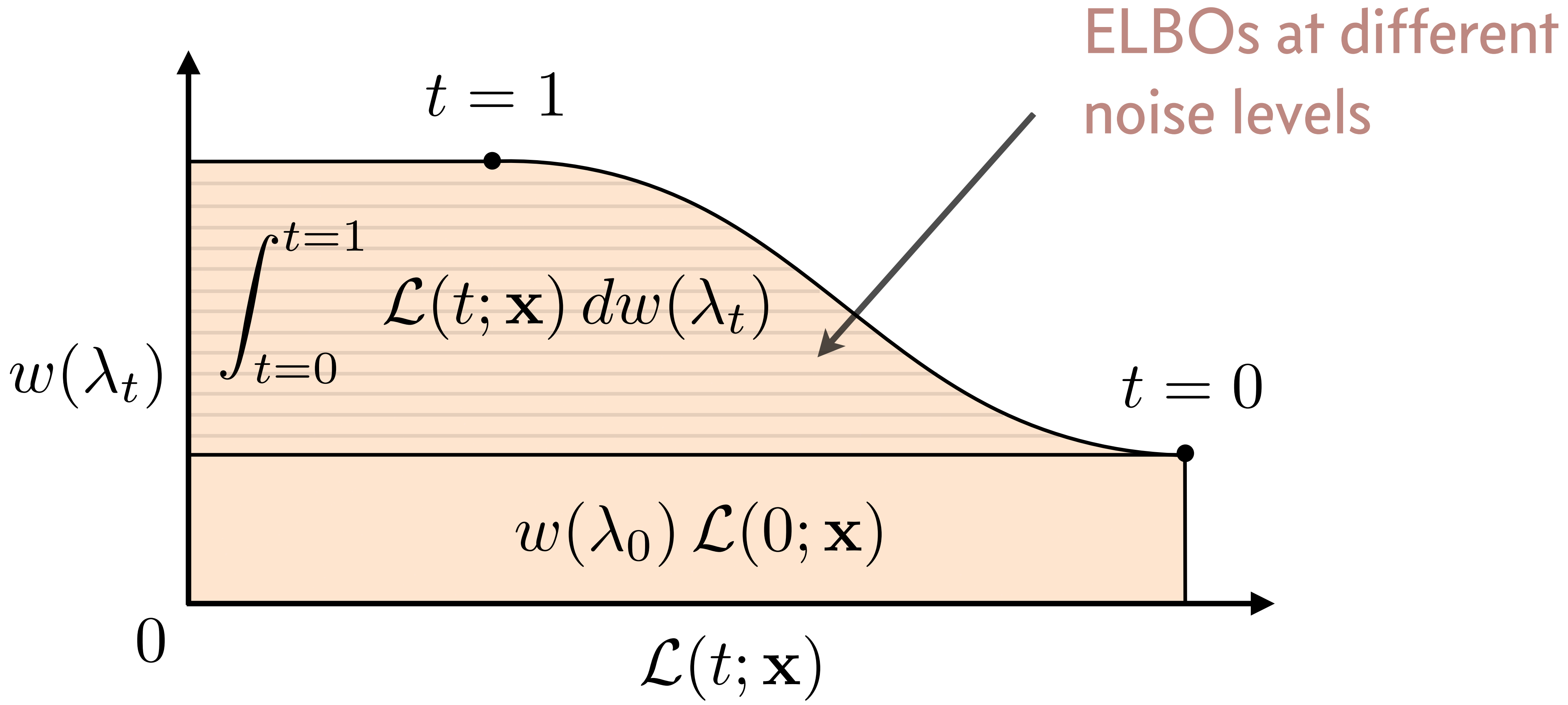
$$L_w(x) = \left\langle w(t) \cdot w_{\mathrm{ML}}(t) \left\| \epsilon - \hat{\epsilon}_\theta\left(x_t, t\right) \right\|^2 \right\rangle$$

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

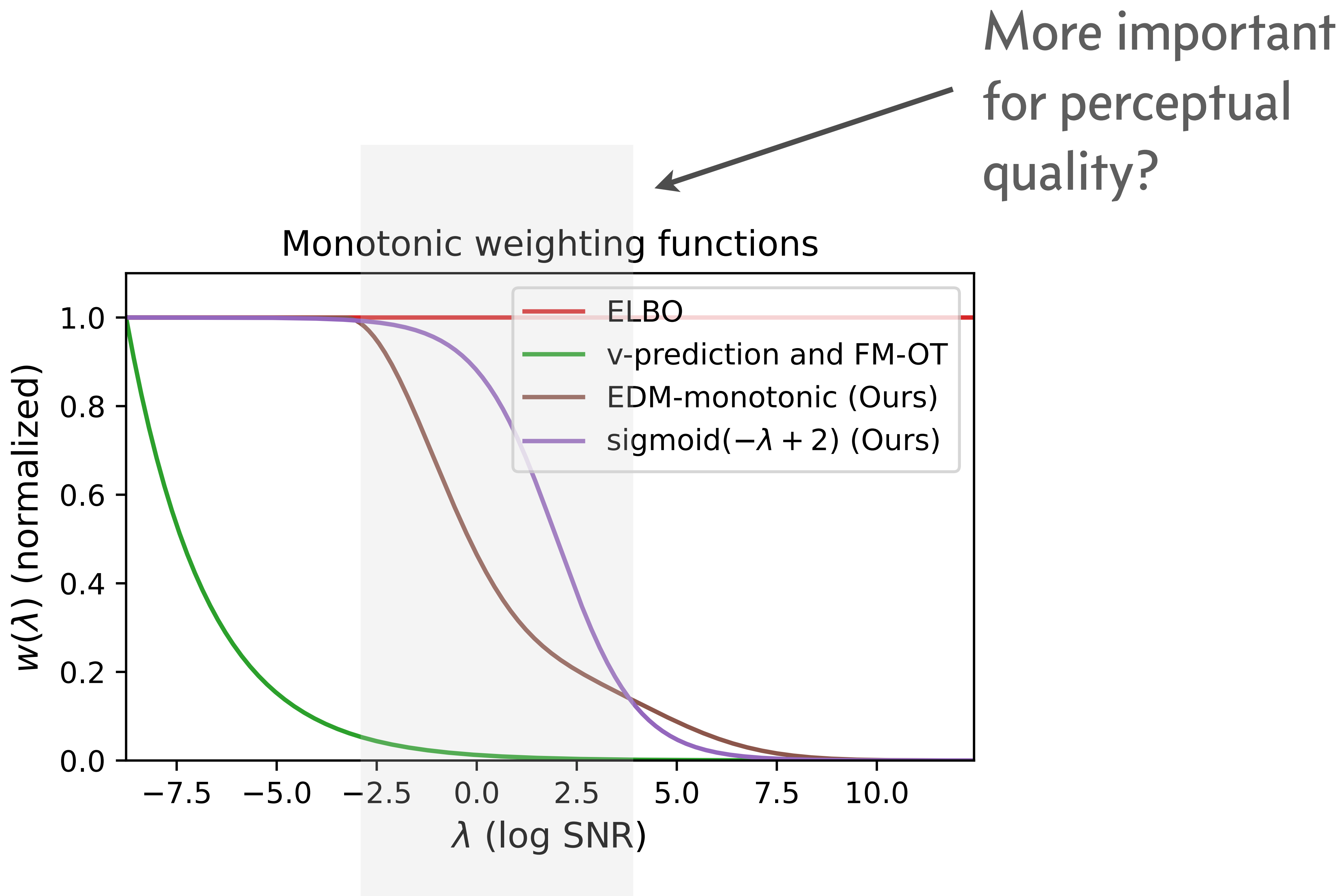
Weighting for ELBO/ ML objective

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

Importance weighting
of different noise levels


ELBOs at different noise levels
$t = 1$
$t = 0$
$\int_{t=0}^{t=1} \mathcal{L}(t; \mathbf{x})\, dw(\lambda_t)$
$w(\lambda_t)$
$w(\lambda_0)\, \mathcal{L}(0; \mathbf{x})$
0
$\mathcal{L}(t; \mathbf{x})$

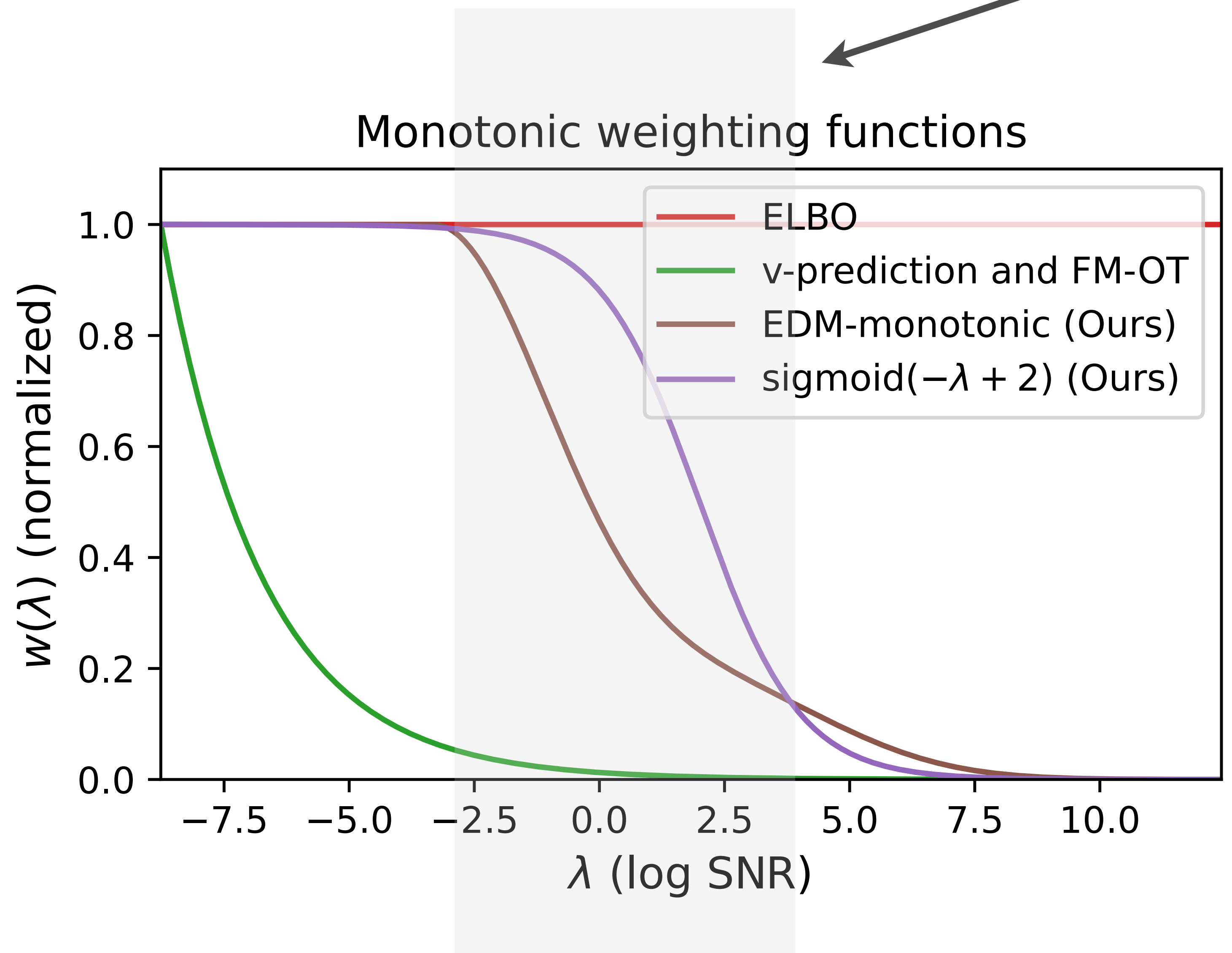
Interpretation: *data augmentation with additive Gaussian noise / data-distribution smoothing*


More important for perceptual quality?
Monotonic weighting functions
ELBO
v-prediction and FM-OT
EDM-monotonic (Ours)
sigmoid(−λ + 2) (Ours)
w(λ) (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

Simple objectives as a weighted sum of ELBOs

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

More important for perceptual quality?



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

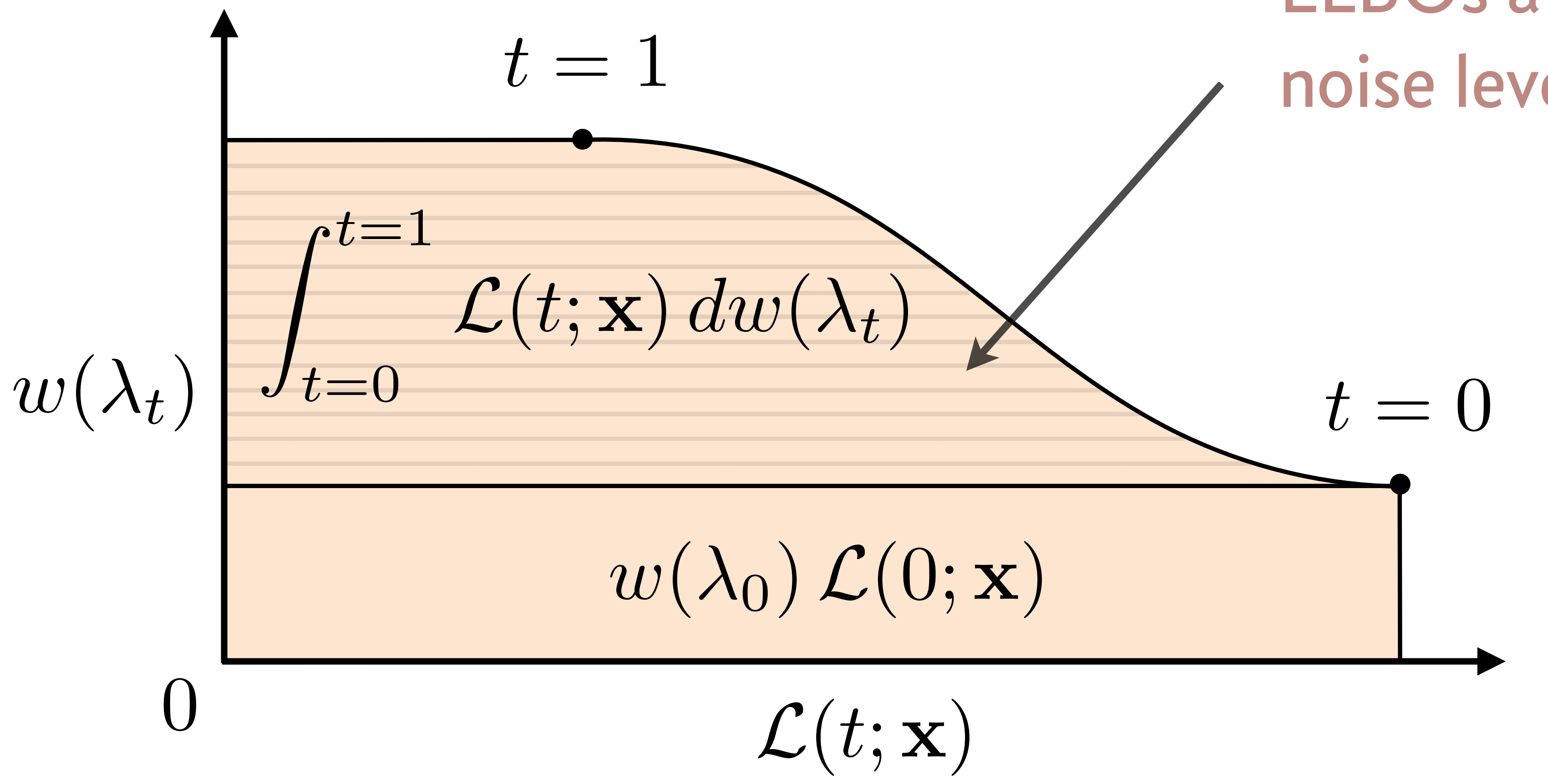
Additional weighting (w_{ML}^{-1} for ϵ-prediction)

Weighting for ELBO/ ML objective

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

Importance weighting of different noise levels

ELBOs at different noise levels



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

Continuous-time/SDE formulation

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

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