



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



166

4

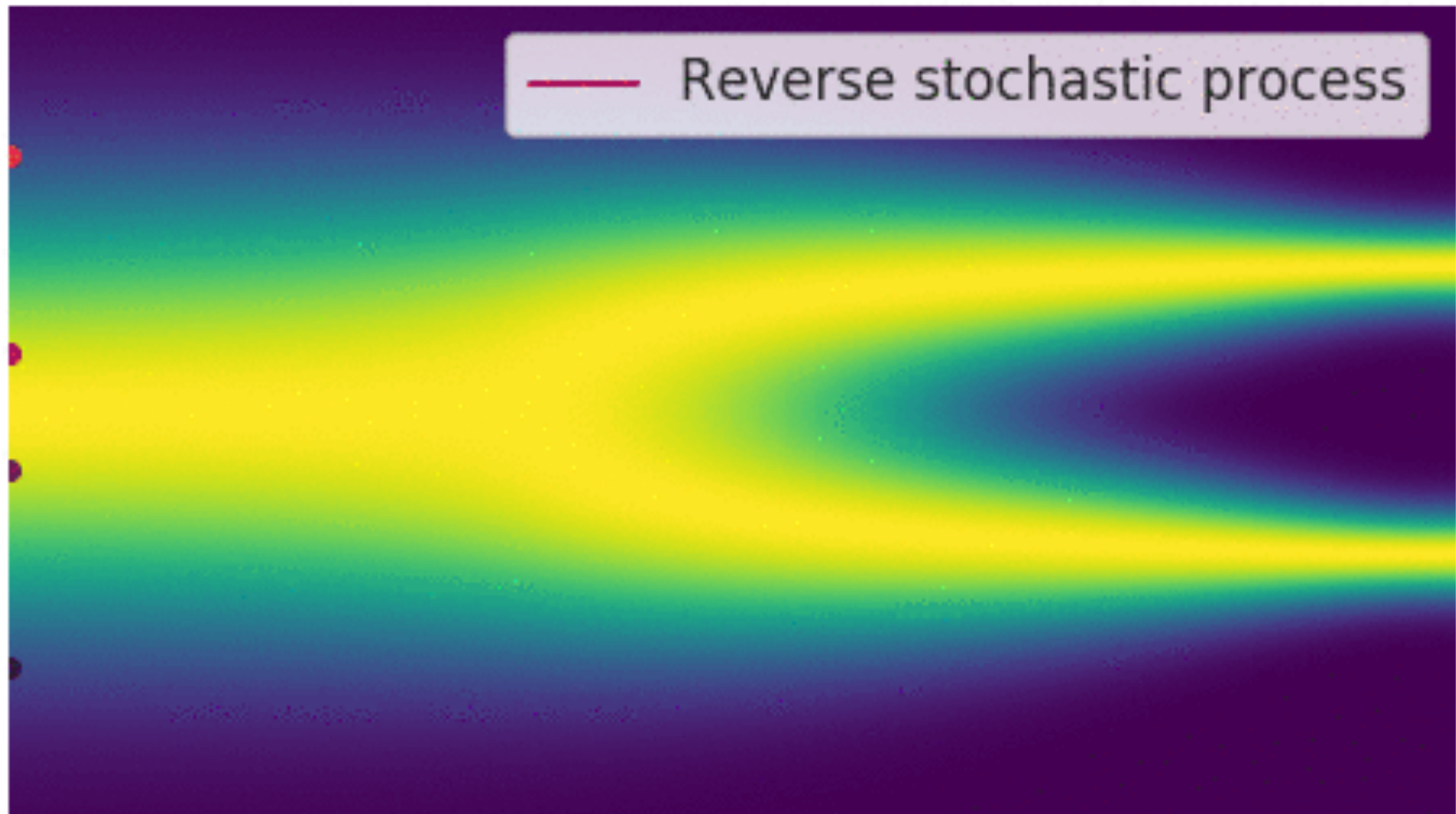
2

ThereversesDE

The reverse process satisfies a reverse-time SDE that can be derived from the forward SDE and the score of the marginal distribution,  $\nabla_{x_t} \log q(x_t)$

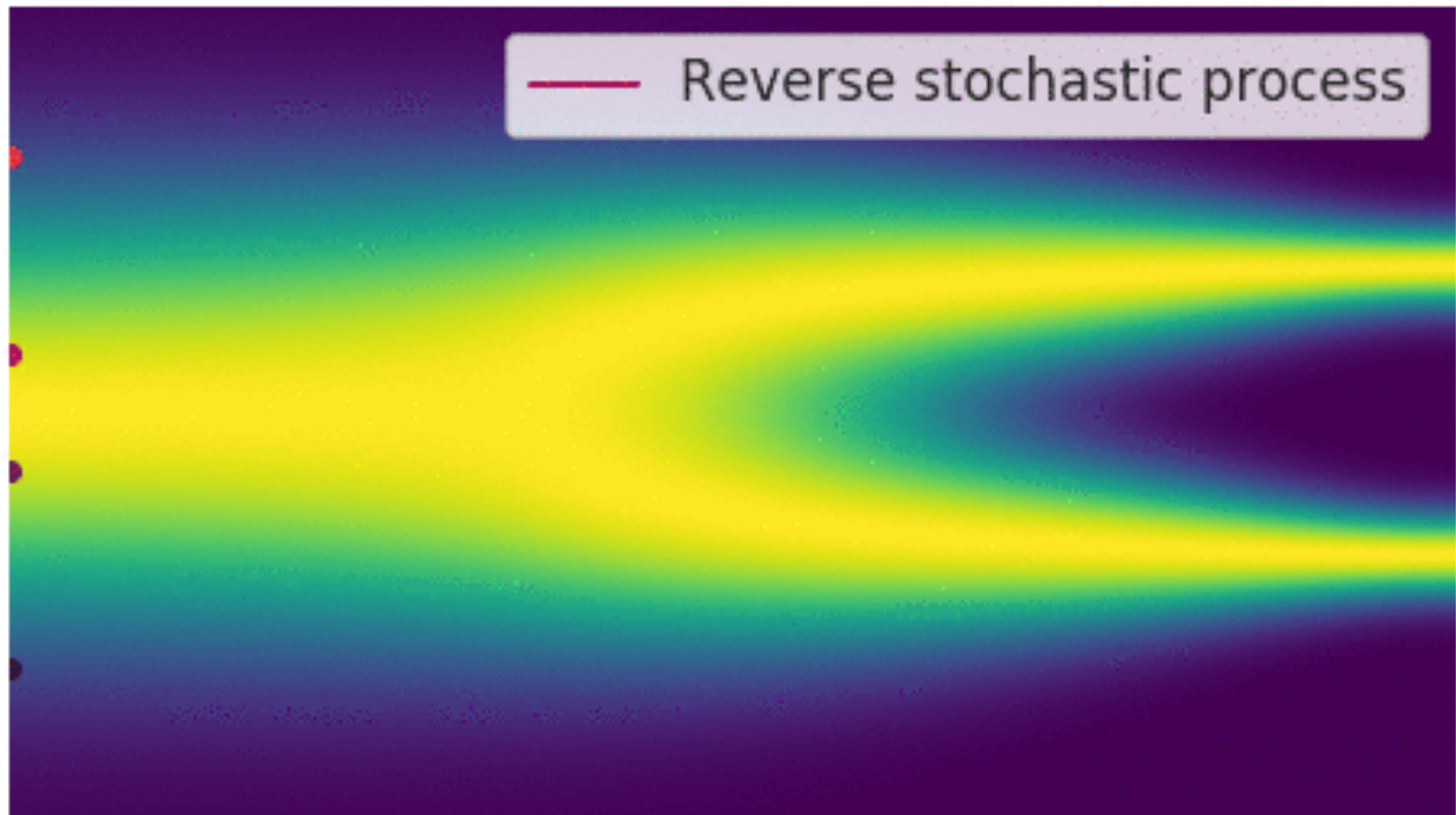
$$\mathrm{d}x_t = \left[ -\frac{1}{2}\beta(t)x_t - \beta(t)\nabla_{x_t}\log q\left(x_t\right) \right] \mathrm{d}t + \sqrt{\beta(t)}\mathrm{d}w_t$$





[song et al 2021]



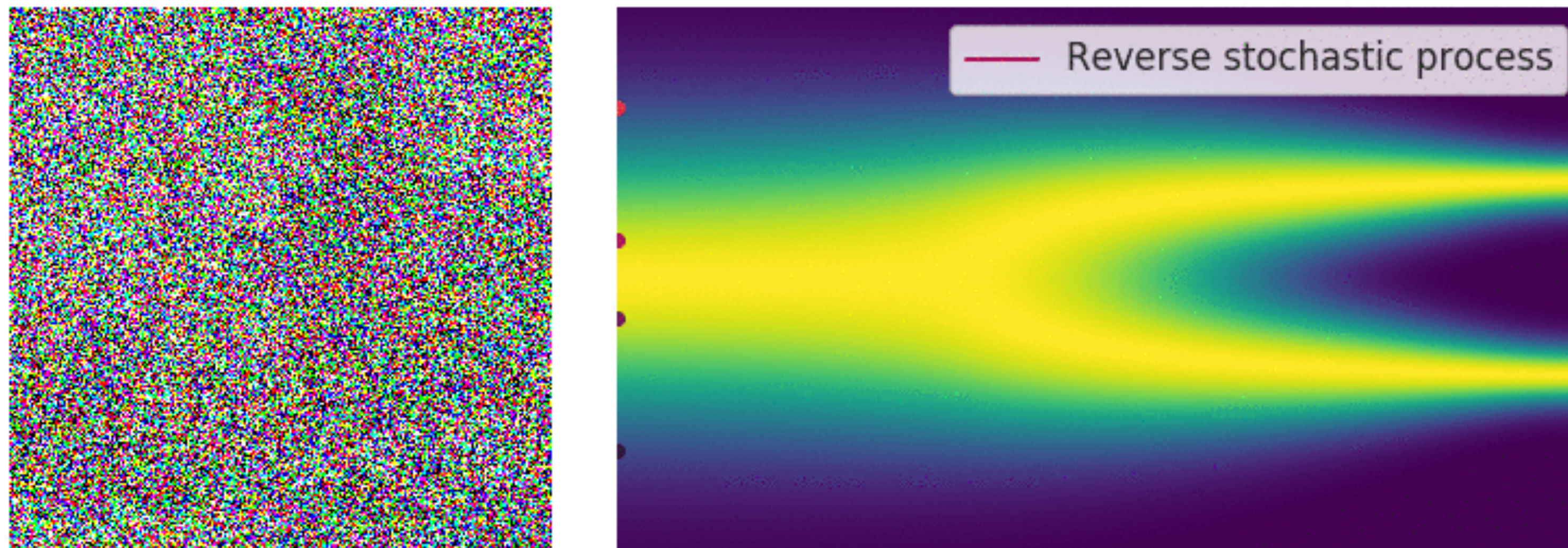




# The reverse SDE

The reverse process satisfies a reverse-time SDE that can be derived from the forward SDE and the score of the marginal distribution,  $\nabla_{x_t} \log q(x_t)$

[Song et al 2021]



$$dx_t = \left[ -\frac{1}{2}\beta(t)x_t - \beta(t) \nabla_{x_t} \log q(x_t) \right] dt + \sqrt{\beta(t)}dw_t$$

# Denoising score matching

Need to compute the score  $\nabla_{x_t} \log q(x_t)$

The *conditional* score  $\nabla_{x_t} \log q(x_t | x)$  can be computed using the diffusion kernel

$$\nabla_{x_t} \log q(x_t | x) = -\frac{(x_t - x)}{\sigma_t^2} = -\frac{\epsilon}{\sigma_t}$$