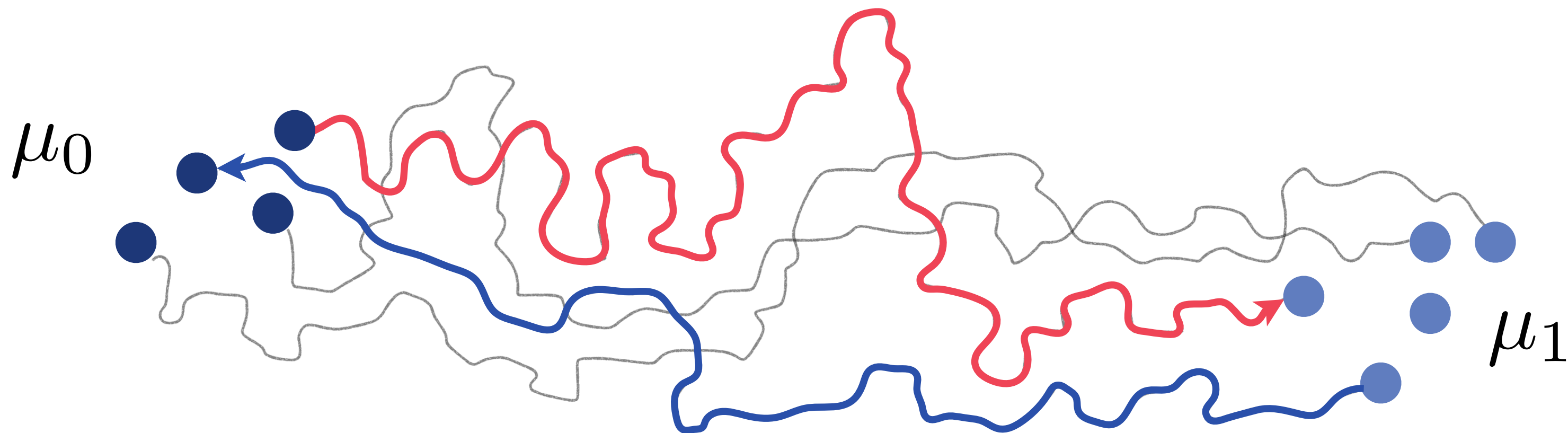


forward policy Z_t^θ

$$X_0 \sim \mu_0 \quad \mathrm{d}X_t = \left[f + g^2 \nabla \log \Phi(t, X_t) \right] \mathrm{d}t + g \, \mathrm{d}\mathbb{W}_t,$$



$$\textit{reverse SDE} \quad \mathrm{d}X_t = \left[f - g^2 \nabla \log \hat{\Phi}(t, X_t) \right] \mathrm{d}t + g \, \mathrm{d}\mathbb{W}_t \quad X_1 \sim \mu_1$$

backward policy \hat{Z}_t^ϕ