

# Stochastic control view

- Equivalently, a variational problem

$$\min_{b_t} \mathbb{E} \left[ \frac{1}{2} \int_0^1 \|b(t, X_t)\|^2 dt \right], \quad \text{s.t.} \quad \begin{cases} dX_t = [f_t + g_t b(t, X_t)] dt + g_t dW_t, \\ X_0 \sim \rho_0, X_1 \sim \rho_1. \end{cases}$$

- Initially formulated for classical case [Dai Pro 1991; Pavon 1991; Caluya & Halder 2021]
- Optimal TSB follows a forward-backward SDE system

Forward:  $dX_t = dY_t + Z_t dt, X_0 \sim \rho_0$

Backward:  $dX_t = dY_t - \tilde{Z}_t dt, X_1 \sim \rho_1$

Enables  
learning!!!

- $Z_t \equiv g_t \nabla \log \varphi_t(X_t), \tilde{Z}_t \equiv g_t \nabla \log \tilde{\varphi}_t(X_t)$ , depend on a system of PDE pairs (FPE, HJB)
- Nonlinear Feynman-Kac formula: gives us a likelihood [Chen et al. 2022]

# TSB-based learning model

- Under the recent SB-based learning framework [Vargas 2021, De Bortoli 2021, Chen 2022]

- **Learnable** models  $(Z_t(\theta), \hat{Z}_t(\hat{\theta}))$  for optimal policies  $(Z_t, \hat{Z}_t)$

- NNs, graph/simplicial NNs

- **Trainable** objective relating the TSBP objective and the models

$$\mathcal{L}_{TSB}(x_0) = \mathbb{E} [\log \nu_1(X_1)] - \int_0^1 \mathbb{E} \left[ \frac{1}{2} \|Z_t\|^2 + \frac{1}{2} \|\hat{Z}_t\|^2 + \nabla \cdot (g_t \hat{Z}_t - f_t) + \hat{Z}_t^\top Z_t \right] dt$$

- Particular choices of models give topological variants

- diffusion models using score-matching [Song et al. 2021]

$$Z_t = 0, \quad \hat{Z}_t = g_t \nabla \log p_{t|0}$$

- Diffusion bridge models based on Doob's  $h$ -transform for a particular final distri.

- Probability flow ODE: flow-matching [Lipman et al. 2022]

## TSB-learning model

$$Z_t \approx Z_t(\theta) \quad l(x_0; \phi)$$

$$\tilde{Z}_t \approx \tilde{Z}_t(\phi) \quad l(x_1; \theta)$$

**Learnable**

**Trainable**