Homework III
Generative Modeling by Transport: Mathematical Foundations

Ziseok Lee

May 13, 2026

Problem 1. KL Evolution With Unequal Diffusion Coefficients Let $\rho, \hat{\rho}$ solve

$$\partial_t \rho + \nabla \cdot (b\rho) = \varepsilon \Delta \rho,$$

$$\partial_t \hat{\rho} + \nabla \cdot (\hat{b}\hat{\rho}) = \hat{\varepsilon} \Delta \hat{\rho},$$

with the same initial density ρ_0. Assume $\rho, \hat{\rho} > 0$ are smooth and decaying.

(a) Derive an exact formula for

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{KL}(\rho(t) \| \hat{\rho}(t)).$$

(b) Show that when $\varepsilon = \hat{\varepsilon}$, your formula reduces to the standard Fokker–Planck KL identity.

(c) Identify the additional term when $\varepsilon \neq \hat{\varepsilon}$, and explain why it obstructs a clean drift-error-only bound.

Problem 2. Optimizing a Time-Dependent Diffusion Schedule Suppose a learned model satisfies pointwise-in-time excess-loss densities

$$A(t) = \frac{1}{2}\int |\hat{b} - b|^2 \rho(t,x)\,\mathrm{d}x, \qquad C(t) = \frac{1}{2}\int |\hat{s} - s|^2 \rho(t,x)\,\mathrm{d}x.$$

For a time-dependent diffusion coefficient $\varepsilon(t) > 0$, a formal KL upper bound is

$$B[\varepsilon] = \int_0^1 \left(\frac{A(t)}{2\varepsilon(t)} + \frac{\varepsilon(t)C(t)}{2} \right) \mathrm{d}t.$$

(a) Find the pointwise optimal $\varepsilon^\star(t)$.

(b) Compute the minimized bound.

(c) Suppose $\varepsilon(t) \leq E$ is constrained. Find the constrained minimizer.

(d) Interpret the result in terms of whether velocity or score is learned better at time t.

Problem 3. A Density Estimator With Killing and Resurrection Let $\hat{\rho}$ solve

$$\partial_t \hat{\rho} + \nabla \cdot (\hat{b}\hat{\rho}) = \varepsilon \Delta \hat{\rho} - r(t,x)\hat{\rho}, \qquad \hat{\rho}(0) = \rho_0,$$

where r is a smooth killing rate. Let Y_t^B solve backward

$$\mathrm{d}Y_t^B = \hat{b}(t, Y_t^B)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t^B, \qquad Y_1^B = x.$$

(a) Guess a Feynman–Kac-type representation for $\hat{\rho}(1,x)$.

(b) Prove your formula using backward Itô calculus.

(c) What changes if $r < 0$ on some region?

Problem 4. The Jensen Gap in SDE Cross-Entropy Let

$$\hat{\rho}(1,x) = \mathbb{E}\left[e^{-A}R \mid X = x\right],$$

where $R > 0$ and A are random variables under an auxiliary bridge law. Define

$$H = -\mathbb{E}_{x\sim\rho_1} \log \hat{\rho}(1,x).$$

(a) Prove the Jensen upper bound

$$-\log \mathbb{E}[e^{-A}R] \le \mathbb{E}[A] - \mathbb{E}[\log R].$$

(b) Show that equality holds if and only if $A + \log R$ is almost surely constant under the auxiliary conditional law.

(c) Interpret this equality condition for the SDE density estimator.

Problem 5. Why ODE Drift Error Alone Cannot Control KL Work on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Let $\rho_0 \equiv (2\pi)^{-1}$, let the true velocity be $b = 0$, and let the approximate velocity be

$$\hat{b}_n(x) = \frac{a}{n}\sin(nx), \qquad 0 < a < 1.$$

Let $\hat{\rho}_n(1)$ be the density at time 1 under the ODE $\dot{X}_t = \hat{b}_n(X_t)$.

(a) Prove that

$$\int_0^1 \int_{\mathbb{T}} |\hat{b}_n - b|^2 \rho_0 \, \mathrm{d}x \, \mathrm{d}t \to 0.$$

(b) Prove that the terminal density $\hat{\rho}_n(1)$ does not converge to ρ_0 in KL.

(c) Conclude that no universal inequality of the form

$$\mathrm{KL}(\rho_0 \| \hat{\rho}_n(1)) \leq C \int_0^1 \int |\hat{b}_n - b|^2 \rho_0$$

can hold for deterministic flows.