

Stochastic Interpolants

Based on “Stochastic Interpolations: A Unifying Framework for Flows and Diffusions”
(2025) by Albergo, Boffi, Vanden-Eijnden

Garrett Mulcahy
gmulcahy@uw.edu
<https://garrett-mulcahy.github.io/>

February 10 2026



UNIVERSITY *of* WASHINGTON

1 Continuity Equation Review

2 Stochastic Interpolations

Continuity Equation Review

Continuity Equation

Continuity Equation

Let $(v_t, t \geq 0)$ be a time dependent velocity field, then

$$\partial_t \rho_t + \nabla \cdot (v_t \rho_t) = 0.$$

A **weak** solution means for all $f : \mathbb{R}^d \rightarrow \mathbb{R}$ smooth, compactly supported

$$\frac{d}{dt} \int_{\mathbb{R}^d} f d\rho_t = \int_{\mathbb{R}^d} \langle \nabla f, v_t \rangle d\rho_t$$

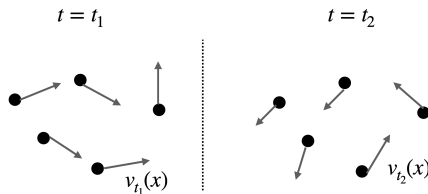


Figure: Eulerian picture

Continuity Equation

Let $\Phi(t, x)$ be such that $\frac{d}{dt}\Phi(t, x) = v_t(\Phi(t, x))$, then $\rho_t = (\Phi_t)_\# \rho_0$ [i.e., if $X_0 \sim \rho_0$, then $\Phi_t(X_0) \sim \rho_t$] is a weak solution of the continuity equation

$$\begin{aligned}
 \frac{d}{dt} \int_{\mathbb{R}^d} f d\rho_t &= \frac{d}{dt} \int_{\mathbb{R}^d} f(\Phi(t, x)) d\rho_0 \\
 &= \int_{\mathbb{R}^d} \frac{d}{dt} f(\Phi(t, x)) d\rho_0 \\
 &= \int_{\mathbb{R}^d} \langle \nabla f(\Phi(t, x)), \partial_t \Phi(t, x) \rangle d\rho_0 \\
 &= \int_{\mathbb{R}^d} \langle \nabla f(\Phi(t, x)), v_t(\Phi(t, x)) \rangle d\rho_0 \\
 &= \int_{\mathbb{R}^d} \langle \nabla f, v_t \rangle d\rho_t.
 \end{aligned}$$

Fokker Planck

Fokker Planck PDE

Let $(b_t, t \geq 0)$ be smoothly varying vector valued function, and recall $\Delta = \sum_{j=1}^d \partial_{x_j}^2$, then

$$\partial_t \rho_t + \nabla \cdot (b_t \rho_t) = \frac{\varepsilon}{2} \Delta \rho_t.$$

Key facts:

- $\Delta \rho_t = \nabla \cdot (\nabla \rho_t) = \nabla \cdot ((\nabla \log \rho_t) \rho_t)$, so get

$$\partial_t \rho_t + \nabla \cdot \left(\left(b_t - \frac{\varepsilon}{2} \nabla \log \rho_t \right) \rho_t \right) = 0.$$

- Gives time marginal evolution of **diffusion processes**

Stochastic Interpolations

Stochastic Interpolation

Stochastic Interpolation [ABVE25]

Let ρ_0, ρ_1 be two probability densities, a stochastic interpolant between ρ_0 and ρ_1 is the stochastic process defined as

$$x_t := I(t, x_0, x_1) + \gamma(t)z, \quad t \in [0, 1],$$

where

- $I(0, x_0, x_1) = x_0$, $I(1, x_0, x_1) = x_1$, $\gamma(0) = \gamma(1) = 0$,
- $(x_0, x_1) \sim \nu$ that is coupling of ρ_0, ρ_1
- $z \sim N(0, \text{Id})$ is independent of (x_0, x_1)

Common choice: linearly spaced interpolants [ABVE25, Section 4]

$$x_t^{\text{lin}} = \alpha(t)x_0 + \beta(t)x_1 + \gamma(t)z,$$

where $\alpha(0) = \beta(1) = 1$ and $\alpha(1) = \beta(0) = \gamma(0) = \gamma(1) = 0$

Stochastic Interpolation Examples

Introduces an element of **design choice**— what **theoretical basis** do we have for certain choices?

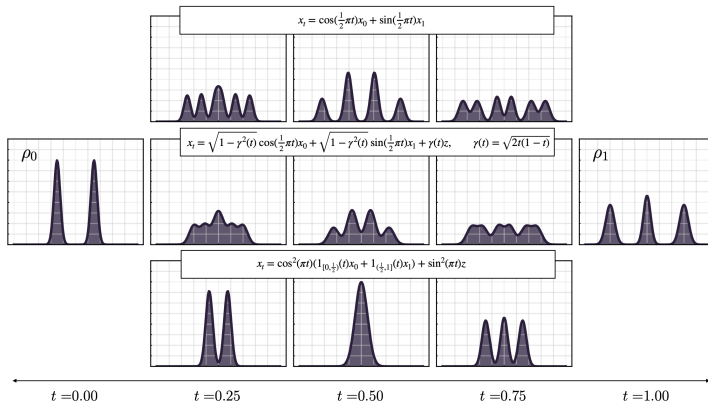


Figure: Different choices of stochastic interpolation [ABVE25, Figure 4]

Stochastic Interpolation

Given $x_t = I(t, x_0, x_1) + \gamma(t)z$, let $x_t \sim \rho_t$

Transport Equation

The PDE $\partial_t \rho_t + \nabla \cdot (b_t \rho_t) = 0$ is satisfied with

$$b(t, x) = \mathbb{E}[\dot{x}_t | x_t = x] = \mathbb{E}[\partial_t I(t, x_0, x_1) + \dot{\gamma}(t)z | x_t = x]$$

Fokker Planck Equation

The PDE $\partial_t \rho_t + \nabla \cdot (b_t^{FP} \rho_t) = \frac{\varepsilon}{2} \Delta \rho_t$ is satisfied with

$$s(t, x) = \nabla \log \rho(t, x) = -\gamma^{-1}(t) \mathbb{E}[z | x_t = x]$$

$$b^{FP} = b + \frac{\varepsilon}{2} s$$

Stochastic Interpolation

[ABVE25, Theorem 6]

Given $x_t = I(t, x_0, x_1) + \gamma(t)z$ and $\rho_t = \text{Law}(x_t)$, the PDE $\partial_t \rho_t + \nabla \cdot (b_t \rho_t) = 0$ is satisfied with

$$b(t, x) = \mathbb{E}[\dot{x}_t | x_t = x] = \mathbb{E}[\partial_t I(t, x_0, x_1) + \dot{\gamma}(t)z | x_t = x]$$

This derivation relies on computing the **characteristic function** (Fourier transform)

- Define $g(t, k)$ as the characteristic function of x_t (spatial variable)

$$g(t, k) = \mathbb{E}[\exp(i\langle k, x_t \rangle)] = \mathbb{E}[\exp(i\langle k, I(t, x_0, x_1) + \gamma(t)z \rangle)]$$

- We will compute $\partial_t g(t, k)$ and use identity

$$\int_{\mathbb{R}^d} \partial_{x_j} f(x) \exp(i\langle \xi, x \rangle) dx = -i\xi_j \int_{\mathbb{R}^d} f(x) \exp(i\langle \xi, x \rangle) dx$$

Stochastic Interpolation

Recall that $(x_0, x_1) \sim \nu$ and $z \sim N(0, \text{Id})$ is independent

$$\begin{aligned}
 \partial_t g(t, k) &= \frac{\partial}{\partial t} \mathbb{E} \left[\exp \left(i \langle k, I(t, x_0, x_1) + \gamma(t)z \rangle \right) \right] \\
 &= \mathbb{E} \left[i \langle k, \partial_t I(t, x_0, x_1) + \dot{\gamma}(t)z \rangle \exp \left(i \langle k, x_t \rangle \right) \right] \\
 &= i \left\langle k, \int_{\mathbb{R}^d} \mathbb{E} \left[\partial_t I(t, x_0, x_1) + \dot{\gamma}(t)z \mid x_t = x \right] \exp \left(i \langle k, x \rangle \right) \rho(t, x) dx \right\rangle \\
 &= i \left\langle k, \int_{\mathbb{R}^d} b(t, x) \rho(t, x) \exp \left(i \langle k, x \rangle \right) dx \right\rangle \\
 &= - \int_{\mathbb{R}^d} \nabla \cdot (b(t, x) \rho(t, x)) \exp \left(i \langle \xi, x \rangle \right) dx
 \end{aligned}$$

Apply Fourier inversion to then get

$$\partial_t \rho(t, x) + \nabla \cdot (b(t, x) \rho(t, x)) = 0.$$

Stochastic Interpolation

Recall $x_t = I(t, x_0, x_1) + \gamma(t)z$, **why set** $\gamma \neq 0$?

- [AVE23] considers $\gamma \equiv 0$
- Key point is **more regularity** in the interpolation

$$s(t, x) = \nabla \log \rho(t, x)$$

$$b(t, x) = \mathbb{E}[\partial_t I(t, x_0, x_1) | x_t = x]$$

- Example: encoder-decoder interpolation

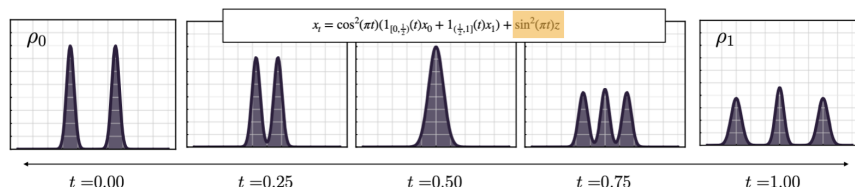


Figure: [ABVE25, Figure 4]

Stochastic Interpolation

What's the point of choosing $(X_0, X_1) \sim \nu$? Another **design choice**

- Diffusion models [SSDK⁺20], one marginal is noise and assume independent
- [AGB⁺24] analyzes theoretical and practical advantages of ν selection
- Another meaningful coupling is Schrödinger bridge [SDBCD23, DBKMD24]

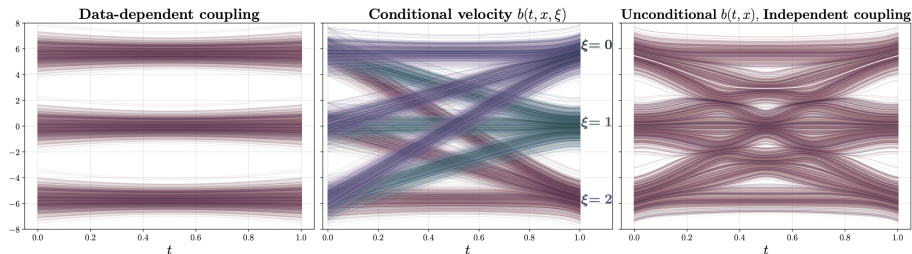


Figure: [AGB⁺24, Figure 3], Gaussian mixture with 3 modes, different couplings of endpoints

- [ABVE25] Michael Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.
- [AGB⁺24] Michael S Albergo, Mark Goldstein, Nicholas M Boffi, Rajesh Ranganath, and Eric Vanden-Eijnden. Stochastic interpolants with data-dependent couplings. In *Proceedings of the 41st International Conference on Machine Learning*, pages 921–937, 2024.
- [AVE23] Michael Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *ICLR 2023 Conference*, 2023.
- [DBKMD24] Valentin De Bortoli, Iryna Korshunova, Andriy Mnih, and Arnaud Doucet. Schrodinger bridge flow for unpaired data translation. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 103384–103441. Curran Associates, Inc., 2024.
- [SDBCD23] Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion schrödinger bridge matching. *Advances in Neural Information Processing Systems*, 36:62183–62223, 2023.

- [SSDK⁺20] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2020. ArXiv version 2011.13456.