Schrödinger-Bass Bridge and Generative Diffusion Models

Nizar Touzi
New York University

Work in progress with
A. Alouadi, P. Henry-Labordère, G. Loeper, O. Mazhar, H. Pham

Advances in Financial Mathematics

Paris, January 27, 2026

Generative models: simulation in high dimension

- Given a "training set" $(X_i)_{i\leq n}$ iid $\mu \in \mathcal{P}(\mathbb{R}^D)$
e.g. $D = 1920 \times 1080 \times 3$
- Generate new data $Y \rightsquigarrow \mu$

- μ unknown (as usual in statistics)

Generative models: simulation in high dimension

- Given a "training set" $(X_i)_{i\leq n}$ iid $\mu \in \mathcal{P}(\mathbb{R}^D)$
  e.g. $D = 1920 \times 1080 \times 3$
- Generate new data $Y \rightsquigarrow \mu$

- μ unknown (as usual in statistics)
- can be approximated by empirical measure $\widehat{\mu}_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$

Generative models: simulation in high dimension

• Given a "training set" $(X_i)_{i \leq n}$ iid $\mu \in \mathcal{P}(\mathbb{R}^D)$
e.g. $D = 1920 \times 1080 \times 3$
• Generate new data $Y \rightsquigarrow \mu$

- μ unknown (as usual in statistics)
- can be approximated by empirical measure $\widehat{\mu}_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$
- simulate from $\widehat{\mu}_n$... useless because produces one of the X_i's

- Given a "training set" $(X_i)_{i\leq n}$ iid $\mu \in \mathcal{P}(\mathbb{R}^D)$ e.g. $D = 1920 \times 1080 \times 3$
- Generate new data $Y \rightsquigarrow \mu$

- μ unknown (as usual in statistics)
- can be approximated by empirical measure $\widehat{\mu}_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$
- simulate from $\widehat{\mu}_n$... useless because produces one of the X_i's
- interpolation in high dimension is very costly

Generative models: simulation in high dimension

• Given a "training set" $(X_i)_{i\leq n}$ iid $\mu \in \mathcal{P}(\mathbb{R}^D)$
e.g. $D = 1920 \times 1080 \times 3$
• Generate new data $Y \rightsquigarrow \mu$

- μ unknown (as usual in statistics)
- can be approximated by empirical measure $\widehat{\mu}_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$
- simulate from $\widehat{\mu}_n$... useless because produces one of the X_i's
- interpolation in high dimension is very costly
- and in addition simulation from continuous distribution in high dimension is not trivial !

Example: using Langevin SDEs

Suppose $\mu(dx) = e^{-U(x)}dx$ and consider the Langevin SDE

$$dX_t = -\nabla U(X_t)dt + \sqrt{2}dW_t$$

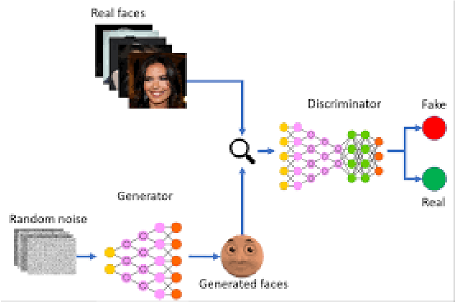
- Ergodic distribution - if exists - solves the homogeneous Fokker-Planck equation

$$\cancel{\partial_t m +} \quad \nabla \cdot (\nabla U\, m) + \Delta m = 0$$

so $m(x) = e^{-U(x)}$ is a solution and $dX_t = \boxed{\nabla \log m(X_t)}\, dt + \sqrt{2}dW_t$

- i.e. $\text{Law}(X_t) \longrightarrow \mu$ as $t \to \infty$
- Simulate X by Euler scheme, say, and consider large t

Generative Adverserial Networks



Game opposing Generator and Discriminator

What's behind Adverserial Networks?

Noise: $(U_i)_{1\le i\le n}$ iid $\mathcal{N}(0, I_d)$ (or any other simple model), $d \ll D$

Generator $G : \Theta \times \mathbb{R}^d \longrightarrow \mathbb{R}^D$ (e.g. NN), $G(\theta, U_i) \rightsquigarrow \mu_\theta \in \mathcal{P}(\mathbb{R}^D)$

Choose θ by solving: $\inf_\theta \mathbf{d}(\mu_\theta, \mu)$

Under this form, we see no game problem between the generator and some discriminator...

The game problem in Adverserial Networks

- In a context where **d** satisfies some convexity property $\Longrightarrow$ Duality:

$$\mathbf{d}(\mu_\theta, \mu) \quad = \quad \sup_{\varphi,\psi} \int \varphi d\mu + \int \psi d\mu_\theta - \mathbf{d}^*(\varphi, \psi)$$

- Approximate from data $\mu \approx \hat{\mu}^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and $\mu_\theta^N = \frac{1}{n}\sum_{i=1}^n \delta_{G(\theta,U_i)}$

$$\Longrightarrow \sup_{\varphi,\psi} \inf_\theta \frac{1}{n}\sum_{i=1}^n \varphi(X_i) + \frac{1}{N}\sum_{j=1}^N \psi(G(\theta, U_j)) - \mathbf{d}^*(\varphi, \psi)$$

- Dual variables (φ, ψ) called **Discriminator**

Method commonly presented as
Zero Sum Game opposing Generator and Discriminator

Famous examples: GAN and WGAN models

GAN $\mathbf{d}(\mu,\nu) := \int f\big(\frac{d\nu}{d\mu}\big) d\mu$ for some divergence distance defined through some convex map f:

$$\inf_\theta \int f\Big(\frac{d\mu_\theta}{d\mu}(x)\Big)\mu(dx) = \inf_\theta \sup_T \int \big[T(x)\frac{d\mu_\theta}{d\mu}(x) - f^*(T(x))\big]\mu(dx)$$
$$= \sup_T \inf_\theta \int T d\mu_\theta - \int f^*(T) d\mu$$

e.g. entropy $\equiv$ KL divergence: $f = -\ln$ and $\mathbf{d}(\mu,\nu) = -\mathrm{H}(\mu|\nu)$

Famous examples: GAN and WGAN models

GAN $\mathbf{d}(\mu,\nu) := \int f\big(\frac{d\nu}{d\mu}\big)d\mu$ for some divergence distance defined through some convex map f:

$$\inf_\theta \int f\Big(\frac{d\mu_\theta}{d\mu}(x)\Big)\mu(dx) = \inf_\theta \sup_T \int \Big[T(x)\frac{d\mu_\theta}{d\mu}(x) - f^*(T(x))\Big]\mu(dx)$$
$$= \sup_T \inf_\theta \int T d\mu_\theta - \int f^*(T)d\mu$$

e.g. entropy $\equiv$ KL divergence: $f = -\ln$ and $\mathbf{d}(\mu,\nu) = -\mathrm{H}(\mu|\nu)$

WGAN $\mathbf{d}$: Wasserstein distance with coupling cost $c(x,y) = |x-y|^p$, we obtain by the Kantorovitch OT duality

$$\inf_\theta \sup_{\varphi\oplus\psi\le c} \int \varphi d\mu + \int \psi d\mu_\theta$$

Famous examples: GAN and WGAN models

GAN $\mathbf{d}(\mu,\nu) := \int f\left(\frac{d\nu}{d\mu}\right)d\mu$ for some divergence distance defined through some convex map f:

$$\begin{aligned}\inf_\theta \int f\Big(\frac{d\mu_\theta}{d\mu}(x)\Big)\mu(dx) &= \inf_\theta \sup_T \int \Big[T(x)\frac{d\mu_\theta}{d\mu}(x) - f^*(T(x))\Big]\mu(dx)\\ &= \sup_T \inf_\theta \int Td\mu_\theta - \int f^*(T)d\mu\end{aligned}$$

e.g. entropy $\equiv$ KL divergence: $f = -\ln$ and $\mathbf{d}(\mu,\nu) = -\mathrm{H}(\mu|\nu)$

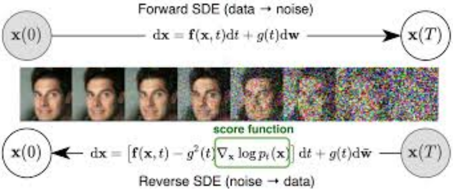
WGAN $\mathbf{d}$: Wasserstein distance with coupling cost $c(x,y) = |x-y|^p$, we obtain by the Kantorovitch OT duality

$$\inf_\theta \sup_{\varphi\oplus\psi\le c} \int \varphi d\mu + \int \psi d\mu_\theta$$

In the special case of $\mathcal{W}_1$:

$$\inf_\theta \sup_{|\nabla\varphi|_\infty} \int \varphi d\mu - \int \varphi d\mu_\theta \approx \inf_\theta \underbrace{\sup_{|\nabla\varphi|_\infty}}_{\approx \mathrm{NN}_{\theta'}} \frac{1}{n}\sum_{i=1}^n \varphi(X_i) - \frac{1}{N}\sum_{j=1}^N \varphi(G(\theta, U_j))$$

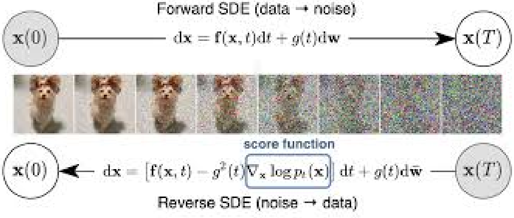
$G(\theta, U_i)$ generated by using time reversal Markov processes



- discrete time Markov chains
- or with continuous time Markov Diffusion
- $\mu_\theta = \mu$!

Time reversal generative models

- Time reversal properties $\Longrightarrow$ Score based models

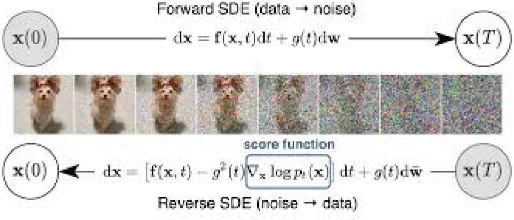


Andersen '82
Hausmann-Pardoux '85
Föllmer '85...

- exact simulation method for known μ...
  in particular, avoids $t \nearrow \infty$ and ergodicity issues in Langevin SDE

Time reversal generative models

- Time reversal properties $\Longrightarrow$ Score based models

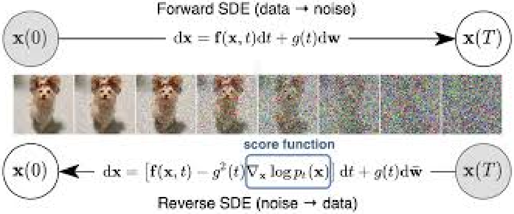


Andersen '82
Hausmann-Pardoux '85
Föllmer '85...

- exact simulation method for known μ...
  in particular, avoids $t \nearrow \infty$ and ergodicity issues in Langevin SDE
- clever usage of the data X_i to bypass μ...
  many techniques for the approximation of the score function
  ... Markovian projection, well-known in volatility modeling !

Time reversal generative models

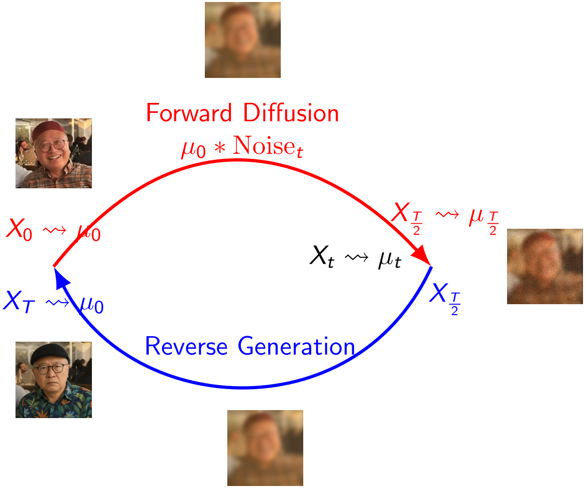
- Time reversal properties $\Longrightarrow$ Score based models



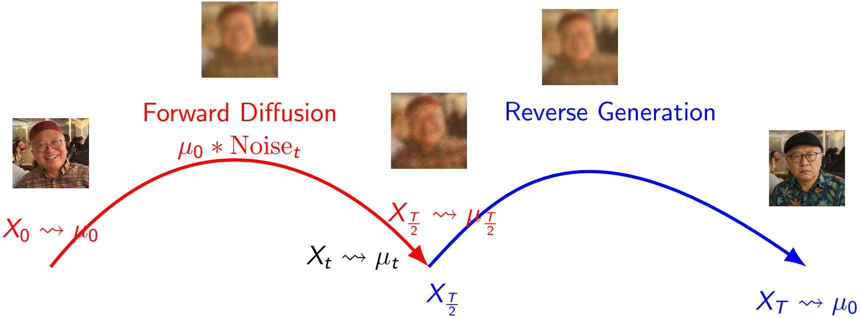
Andersen '82
Hausmann-Pardoux '85
Föllmer '85...

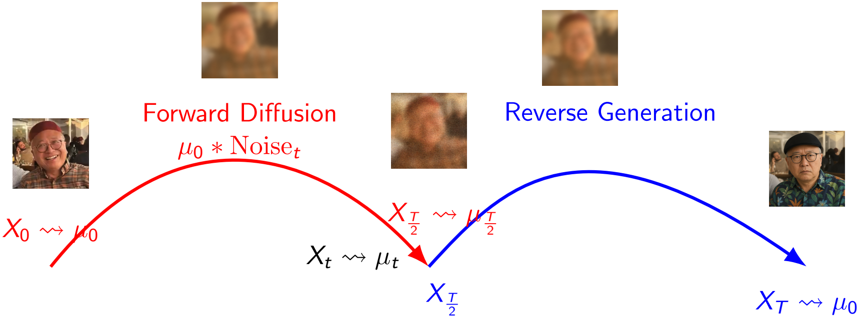
- exact simulation method for known μ...
  in particular, avoids $t \nearrow \infty$ and ergodicity issues in Langevin SDE
- clever usage of the data X_i to bypass μ...
  many techniques for the approximation of the score function
  ... Markovian projection, well-known in volatility modeling !
- Under OU dynamics, choice of $g(t)$
  - ☐ **Variance Preserving**: g bounded
  - ☐ **Exploding Variance**: g unbounded

Score based diffusion generative models



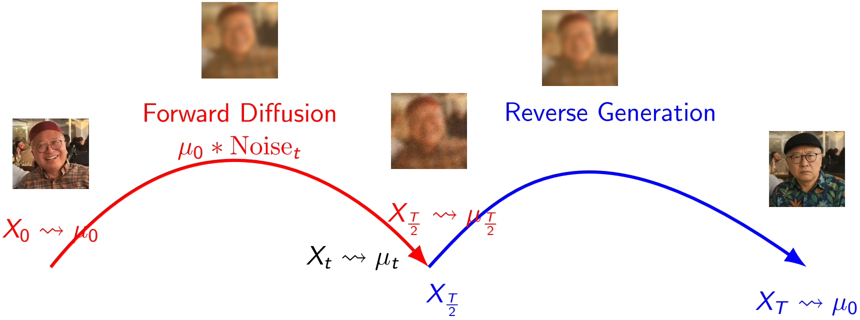
Score based diffusion generative models





Reverse generation minimizes entropy

Score based diffusion generative models



Reverse generation minimizes entropy

Score based generation take away

Given the BM $\{W_t, t \leq T\}$ (noise), Let $G(\theta, W) = X_T$ where

- on $[0, \frac{T}{2}]$: $X_0 \rightsquigarrow \mu_0$ and OU process
- on $[\frac{T}{2}, T]$: **Minimum entropy** reverse time diffusion $X_T \rightsquigarrow \mu_0$

From score based generation to diffusion optimal transport

Why happy with **"half time" optimality**?

Optimal generative diffusion

Given the BM $\{W_t, t \leq T\}$ (noise), Let $G(\theta, W) = X_T$ where

- $X_0 \rightsquigarrow \mu_0$ and $dX_t = \alpha_t dt + \sigma_t dW_t$ on $[0, T]$
- $X_T \rightsquigarrow \mu_0$ or more generally $X_T \rightsquigarrow \mu_T$
- α, σ optimal wrt to some optimal transport cost
- Our criterion: **minimize entropy** and **stabilize variance**

From score based generation to diffusion optimal transport

Why happy with **“half time” optimality**?

Optimal generative diffusion

Given the BM $\{W_t, t \le T\}$ (noise), Let $G(\theta, W) = X_T$ where

- $X_0 \rightsquigarrow \mu_0$ and $dX_t = \alpha_t dt + \sigma_t dW_t$ on $[0, T]$
- $X_T \rightsquigarrow \mu_0$ or more generally $X_T \rightsquigarrow \mu_T$
- α, σ optimal wrt to some optimal transport cost
- Our criterion: **minimize entropy** and **stabilize variance**

The semimartingale optimal transport problem

Find optimal **transport plan** $\mathbb{P}$ on $\Omega = C^0([0,T],\mathbb{R}^d)$:

$$\mathbf{P}(\mu_0,\mu_T) := \inf_{\mathbb{P}\in\mathcal{P}(\mu_0,\mu_T)} \mathbb{E}^{\mathbb{P}}\Big[\int_0^T \underbrace{\frac{1}{2}(|\alpha_t^{\mathbb{P}}|^2}_{\text{entropy}} + \underbrace{\beta|\sigma_t^{\mathbb{P}} - I_d|^2}_{\text{var. preserving}}\;)dt\Big]$$

where $\mathcal{P}(\mu_0,\mu_T)$ subset of Prob(Ω) s.t.

- $X_0 \rightsquigarrow \mu_0$ and $X_T \rightsquigarrow \mu_T$
- Canonical process $X_t(\omega) := \omega(t)$ has the decomposition

$$dX_t = \alpha_t^{\mathbb{P}} dt + \sigma_t^{\mathbb{P}} dW^{\mathbb{P}},\ \mathbb{P}-\text{a.s. for some } \mathbb{P}-\text{BM } W^{\mathbb{P}}$$

Connection with existing literature

Find optimal **transport plan** $\mathbb{P}$ on $\Omega = C^0([0,T], \mathbb{R}^d)$:

$$\mathbf{P}(\mu_0, \mu_T) := \inf_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_T)} \mathbb{E}^{\mathbb{P}} \Big[\int_0^T \underbrace{\frac{1}{2}(|\alpha_t^{\mathbb{P}}|^2}_{\text{entropy}} + \underbrace{\beta |\sigma_t^{\mathbb{P}} - I_d|^2}_{\text{var. preserving}}) dt \Big]$$

- Dynamic version of Optimal Transport [Benamou-Brenier]...
- Semimartingale Optimal Transport [Tan-NT '13, Guo-Loeper-Wang '21]...
- Martingale Optimal Transport [Beiglböck, Henry-Labordère & Penkner '13, Galichon, Henry-Labordère & NT '13]
    - $\beta = \infty$: Schrödinger Bridge/Entropic Optimal Transport (Sinkhorn)
    - $\beta = 0$: Stretched Brownian motion [Beiglböck & Backhof]
    - later identified with the Bass solution of the Skorohod Embedding Problem [Beiglböck-Backhof-Schachermayer-Schilderer]
    - and thus to the Bass martingale with arbitrary starting measure [Henry-Labordère & Conze '22, Acciaio, Marini & Pammer '25]

Dual formulation by standard LP

Recall $X_0 \sim \mu_0$, $dX_t = \alpha_t^{\mathbb{P}} dt + \sigma_t^{\mathbb{P}} dW^{\mathbb{P}}$, $\mathbb{P}-$a.s. and

$$\mathbf{P}(\mu_0, \mu_T) := \inf_{\mathbb{P}\circ X_T^{-1} = \mu_T} \mathbb{E}^{\mathbb{P}}\Big[\int_0^T c(\alpha_t^{\mathbb{P}}, \sigma_t^{\mathbb{P}}) dt\Big], \; c(\alpha, \sigma) = \frac{1}{2}(|\alpha|^2 + \beta|\sigma - I_d|^2)$$

$$= \inf_{\mathbb{E}^{\mathbb{P}}\psi(X_T) = \int \psi d\mu_T, \forall \psi} \mathbb{E}^{\mathbb{P}}\Big[\int_0^T c(\alpha_t^{\mathbb{P}}, \sigma_t^{\mathbb{P}}) dt\Big]$$

$$= \inf_{\mathbb{P}} \sup_{\psi} \int \psi d\mu_T - \mathbb{E}^{\mathbb{P}}\Big[\psi(X_T) - \int_0^T c(\alpha_t^{\mathbb{P}}, \sigma_t^{\mathbb{P}}) dt\Big]$$

$$\stackrel{?}{=} \sup_{\psi} \inf_{\mathbb{P}} \int \psi d\mu_T - \mathbb{E}^{\mathbb{P}}\Big[\psi(X_T) - \int_0^T c(\alpha_t^{\mathbb{P}}, \sigma_t^{\mathbb{P}}) dt\Big] := \mathbf{D}(\mu_0, \mu_T)$$

By the standard Kuhn-Tucker constrained optimization...

Compare with Tan & NT '13 and Guo, Loeper & Wang '21

Dual formulation

Find optimal potential map ψ:

$$\mathbf{D}(\mu_0, \mu_T) := \sup_{\psi} \mu_T(\psi) - \mu_0(v_0^\psi)$$

where $v_0^\psi(x) := \sup_{\mathbb{P} \in \mathcal{P}(\delta_x, .)} \mathbb{E}_{0,x}^{\mathbb{P}} \Big[\psi(X_T) - \int_0^T \frac{1}{2} (|\alpha_t^{\mathbb{P}}|^2 + \beta |\sigma_t^{\mathbb{P}} - I_d|^2) dt \Big]$

Dual formulation

Find optimal potential map ψ:

$$\mathbf{D}(\mu_0, \mu_T) := \sup_{\psi} \mu_T(\psi) - \mu_0(v_0^{\psi})$$

where $v_0^{\psi}(x) := \sup_{\mathbb{P}\in\mathcal{P}(\delta_x,.)} \mathbb{E}_{0,x}^{\mathbb{P}}\Big[\psi(X_T) - \int_0^T \frac{1}{2}(|\alpha_t^{\mathbb{P}}|^2 + \beta|\sigma_t^{\mathbb{P}} - I_d|^2)dt\Big]$

Proposition

(i) The value function of the last control problem is given by

$$v_0^{\psi} = \mathbf{T}_\beta^+\big[\log(\mathcal{N}_T * \underbrace{e^{\mathbf{T}_\beta^-[\psi]}}_{=:e^{\phi}})\big] \text{ with } \mathbf{T}_\beta^+[f](x) := \inf_{y\in\mathbb{R}^d}\Big\{f(y) + \frac{\beta}{2}|x-y|^2\Big\}$$

$$\mathbf{T}_\beta^-[f] := -\mathbf{T}_\beta^+[-f] \text{ and } \mathcal{N}_T := \frac{e^{-\frac{1}{2T}|\cdot|^2}}{(2\pi T)^{\frac{d}{2}}}$$

(ii) $\mathbf{D}(\mu_0, \mu_T) = \sup_{\phi+\frac{\beta}{2}|.|^2 \ \mathrm{conv}} \mu_T\big(\mathbf{T}_\beta^+[\phi]\big) - \mu_0\big(\mathbf{T}_\beta^+\big[\log(\mathcal{N}_T * e^{\phi})\big]\big)$

Main results

Theorem

Assume $\mu_0, \mu_T \in \mathcal{P}_2(\mathbb{R}^d)$. Then

(i) $\mathbf{P}(\mu_0, \mu_T) = \mathbf{D}(\mu_0, \mu_T) < \infty$ and there exists a solution $\hat{\mathbb{P}}$ for $\mathbf{P}$

(ii) If $\beta T > 1$, the dual problem $\mathbf{D}$ has a unique (up to a constant) solution $\hat{\phi} \in \mathrm{Conv}_\beta$ characterized by the SBB system

$$\begin{aligned} \mathcal{Y}_0 \# \mu_0 &= (\mathcal{N}_T * e^{\hat{\phi}})\nu_0 \\ \mathcal{Y}_T \# \mu_T &= e^{\hat{\phi}} \nu_T \end{aligned} \quad \text{where} \quad \begin{aligned} \nu_T &= \mathcal{N}_T * \nu_0, \\ \mathcal{Y}_t &:= \mathrm{id} - \tfrac{1}{\beta} \nabla \log(\mathcal{N}_{T-t} * e^{\hat{\phi}}) \end{aligned}$$

(iii) The process $Y_t := \mathcal{Y}_t(X_t)$ is a $\hat{\mathbb{Q}}-$BM with $\frac{d\hat{\mathbb{Q}}}{d\hat{\mathbb{P}}} := (\mathcal{N}_T * e^{\hat{\phi}})(Y_T)$

Optimal $\hat{\alpha}$ and $\hat{\sigma}$ follow from (iii)

Schrödinger-Bass Bridge

Find the optimal potential $\hat{\phi}$:

$$J(\hat{\phi}) = \min_{\mathrm{Conv}_\beta} J(\phi) := \int \mathbf{T}_\beta^+ \big[\log(\mathcal{N}_T * e^\phi)\big] d\mu_0 - \int \mathbf{T}_\beta^+[\phi] d\mu_T$$

Then, with $\mathcal{Y}_t := \mathrm{id} - \frac{1}{\beta}\nabla \log(\underbrace{\mathcal{N}_{T-t} * e^{\hat{\phi}}}_{=:h_t(.)})$, the SBB system is

$$\begin{array}{ccccccc}
\mu_T & \longrightarrow & \mathcal{Y}_T \# \mu_T & \longrightarrow & \frac{d\mathcal{Y}_T \# \mu_T}{d\nu_T} = h_T & \longrightarrow & \nu_T := \mathcal{N}_T * \nu_0 \\
\Big\uparrow & & \text{Bass} & & \text{Shrödinger} & & \Big\uparrow \text{ FP equation} \\
\mu_0 & \longrightarrow & \mathcal{Y}_0 \# \mu_0 & \longrightarrow & \frac{d\mathcal{Y}_0 \# \mu_0}{d\nu_0} = h_0 & \longrightarrow & \nu_0
\end{array}$$

and $Y_t := \mathcal{Y}_t(X_t)$ is a $\hat{\mathbb{Q}}$−BM, where $\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = h_T(Y_T)$

Bass martingale: $\hat{\phi} = 0$

Shrödinger bridge: $\mathcal{Y}_t = id$, $t \in [0, T]$

Application to image generation

Algorithm: Light SBB Matching

SBB construction. The SBB process $\mathbb{P}^{\mathrm{SBB}}$ is the *Bass transport* of a Schrödinger Bridge Y between $\mathcal{Y}_0\#\mu_0$, $\mathcal{Y}_T\#\mu_T$

$$X_t = \mathcal{Y}_t^{-1}(Y_t) = Y_t + \tfrac{1}{\beta}\, S_t(Y_t), \quad S_t(y) := \nabla_y \log h_t(y)$$

Iterative light SBB matching. Start from $S^{(0)} \equiv 0$. For $k = 0, 1, \ldots$:

1. **Sample endpoints** (y_0, y_T) from $(\mathcal{Y}_0^{(k)}\#\mu_0,\ \mathcal{Y}_T^{(k)}\#\mu_T)$, where

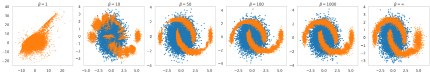
$$\mathcal{Y}_0^{(k)} = \big(\mathrm{Id} + \tfrac{1}{\beta} S_0^{(k)}\big)^{-1}, \quad \mathcal{Y}_T^{(k)} = \big(\mathrm{Id} + \tfrac{1}{\beta} S_T^{(k)}\big)^{-1}$$

2. **Update score** of Y using *Light SB matching*, yielding $S^{(k+1)}$
3. **Generation at final iteration** K of samples of μ_T via

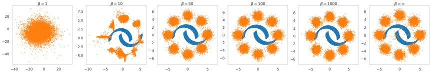
$$x_T = y_T + \tfrac{1}{\beta}\, S_T^{(K)}(y_T),$$

where y_T are samples from the SB with score-drift $S^{(K)}$.

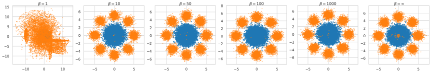
2-d benchmarks (Mila group of Y. Bengio) by varying β



(a) $\mu_0 \equiv \mathcal{N} \to \mu_T \equiv$ moons for differents values of β, with μ_0 in blue and μ_T in orange.



(b) $\mu_0 \equiv$ moons $\to \mu_T \equiv$ 8gaussians for differents values of β, with μ_0 in blue and μ_T in orange.



(c) $\mu_0 \equiv \mathcal{N} \to \mu_T \equiv$ 8gaussiansst for differents values of β, with μ_0 in blue and μ_T in orange.

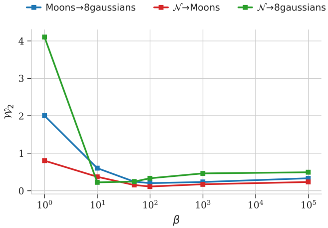


Figure: $\mathcal{W}_2$ distance computed for different values of β in SBB.

Quantitative State of the Art Comparison

Metric $\rightarrow$	$\mathcal{W}_2$ $(\downarrow)$		
Algorithm $\downarrow$	$\mathcal{N} \rightarrow$ 8gaussians	moons $\rightarrow$ 8gaussians	$\mathcal{N} \rightarrow$ moons
$[SF]^2$M-Exact	0.275±0.058	0.726±0.137	0.124±0.023
$[SF]^2$M-I	0.393±0.054	1.482±0.151	0.185±0.028
DSBM-IPF	0.315±0.079	0.812±0.092	0.140±0.006
DSBM-IMF	0.338±0.091	0.838±0.098	0.144±0.024
DSB	0.411±0.084	0.987±0.324	0.190±0.049
Light-SBM	0.339±0.099	0.330±0.081	0.231±0.012
SBB	**0.221±0.023**	**0.201±0.014**	**0.110±0.010**
OT-CFM	0.303±0.043	0.601±0.027	0.130±0.016
SB-CFM	2.314±2.112	—	0.434±0.594
RF	0.421±0.071	1.525±0.330	0.283±0.045
I-CFM	0.373±0.103	1.557±0.407	0.178±0.014
FM	0.343±0.058	—	0.209±0.055

Adult to Child with light SBB matching (through light SB)

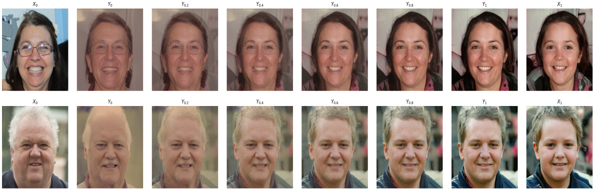


Figure: Trajectory from $X_0 \sim \mu_0$ to $X_T \sim \mu_T$ with the underlying Y process.
$\beta = 5$ for female and $\beta = 10$ for Male

Noise to Child with light SBB matching (through light SB)

Figure: Input noise μ_0 on the left column, and three representative samples of μ_T. $\beta = 5$