Exponential Convergence Guarantees for Iterative Markovian Fitting

Marta Gentiloni Silveri, Giovanni Conforti, Alain Durmus

39th Conference on Neural Information Processing Systems

Why Study IMF?

- Generative Models, e.g., Score-Based Generative Models (SGMs), often suffer from long sampling times.

Why Study IMF?

- Generative Models, e.g., Score-Based Generative Models (SGMs), often suffer from long sampling times.
- Schrödinger Bridge (SB) offers a principled alternative to SGMs.

Why Study IMF?

- Generative Models, e.g., Score-Based Generative Models (SGMs), often suffer from long sampling times.
- Schrödinger Bridge (SB) offers a principled alternative to SGMs.
- Iterative Markovian Fitting (IMF): a key algorithm to solve SB efficiently.

Why Study IMF?

- Generative Models, e.g., Score-Based Generative Models (SGMs), often suffer from long sampling times.
- Schrödinger Bridge (SB) offers a principled alternative to SGMs.
- Iterative Markovian Fitting (IMF): a key algorithm to solve SB efficiently.
- Until now: **only asymptotic convergence results**.

Why Study IMF?

- Generative Models, e.g., Score-Based Generative Models (SGMs), often suffer from long sampling times.
- Schrödinger Bridge (SB) offers a principled alternative to SGMs.
- Iterative Markovian Fitting (IMF): a key algorithm to solve SB efficiently.
- Until now: **only asymptotic convergence results**.
- This work: **first quantitative, exponential convergence bounds**.

Schrödinger Bridge (SB) Problem: Setup

Goal: Seek the most likely evolution of a cloud of independent Brownian particles that interpolates between two observed empirical distributions $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$.

Schrödinger Bridge (SB) Problem: Setup

Goal: Seek the most likely evolution of a cloud of independent Brownian particles that interpolates between two observed empirical distributions $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$.

Dynamical Formulation

Minimize $\mathrm{KL}(P|\mathrm{R})$ over $P \in \mathcal{P}(C([0,T]))$ s.t. $P_0 = \mu$, $P_T = \nu$.

R: Reference Diffusion, $i.e.$, Brownian Motion.

Schrödinger Bridge (SB) Problem: Setup

Goal: Seek the most likely evolution of a cloud of independent Brownian particles that interpolates between two observed empirical distributions $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$.

Dynamical Formulation

Minimize $\mathrm{KL}(P|\mathrm{R})$ over $P \in \mathcal{P}(C([0,T]))$ s.t. $P_0 = \mu$, $P_T = \nu$.

R: Reference Diffusion, $i.e.$, Brownian Motion.

Static Formulation

Minimize $\mathrm{KL}(\pi|\mathrm{R}_{0,T})$ over couplings $\pi \in \Pi(\mu, \nu)$.

$\mathrm{R}_{0,T}$: Joint law at time $t = 0$ and $t = T$ of R.

Schrödinger Bridge (SB) Problem: Setup

Goal: Seek the most likely evolution of a cloud of independent Brownian particles that interpolates between two observed empirical distributions $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$.

Dynamical Formulation

Minimize $\mathrm{KL}(P|\mathrm{R})$ over $P \in \mathcal{P}(C([0,T]))$ s.t. $P_0 = \mu$, $P_T = \nu$.

R: Reference Diffusion, $i.e.$, Brownian Motion.

Static Formulation

Minimize $\mathrm{KL}(\pi|\mathrm{R}_{0,T})$ over couplings $\pi \in \Pi(\mu, \nu)$.

$\mathrm{R}_{0,T}$: Joint law at time $t = 0$ and $t = T$ of R.

Theorem: Under mild assumptions, there exists a unique solution $\pi^\star$, called Schrödinger Bridge, which can be expressed as

$$\pi^\star(\mathrm{d}x, \mathrm{d}y) = \exp\left(-\varphi(x) - \psi(y)\right) \mathrm{R}_{0,T}(\mathrm{d}x, \mathrm{d}y) \ ,$$

for some $\varphi, \psi : \mathbb{R}^d \to (-\infty, +\infty]$ known as Sinkhorn potentials.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi^{(0)}_{0,T} \in \Pi(\mu, \nu)$.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi_{0,T}^{(0)} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi^{(0)}_{0,T} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$
  - **Constraint 1: Having** bR **as bridge.**
    Construct the stochastic interpolant $Y_t^{(k+1)}$ based on $\pi^{(k)}_{0,T}$.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi_{0,T}^{(0)} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$
  - **Constraint 1: Having** bR **as bridge.**
    Construct the stochastic interpolant $Y_t^{(k+1)}$ based on $\pi_{0,T}^{(k)}$.
  - **Constraint 2: Being Markov.**
    Compute its Markovian projection $X_t^{(k+1)}$.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi_{0,T}^{(0)} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$
  - **Constraint 1: Having** bR **as bridge.**
    Construct the stochastic interpolant $Y_t^{(k+1)}$ based on $\pi_{0,T}^{(k)}$.
  - **Constraint 2: Being Markov.**
    Compute its Markovian projection $X_t^{(k+1)}$.
  - Set $\pi_{0,T}^{(k+1)} = \text{Law}(X_0^{(k+1)}, X_T^{(k+1)})$.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi_{0,T}^{(0)} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$
  - **Constraint 1: Having** bR **as bridge.**
    Construct the stochastic interpolant $Y_t^{(k+1)}$ based on $\pi_{0,T}^{(k)}$.
  - **Constraint 2: Being Markov.**
    Compute its Markovian projection $X_t^{(k+1)}$.
  - Set $\pi_{0,T}^{(k+1)} = \mathrm{Law}(X_0^{(k+1)}, X_T^{(k+1)})$.
- Output: $\pi_{0,T}^{(N)}$.

Iterative Markovian Fitting (IMF): Algorithm

Idea: Alternate between Stochastic Interpolants and Markovian Projections.

- Input: $\pi_{0,T}^{(0)} \in \Pi(\mu, \nu)$.
- For $k = 0, ..., N-1$
  - **Constraint 1: Having** bR **as bridge.**
    Construct the stochastic interpolant $Y_t^{(k+1)}$ based on $\pi_{0,T}^{(k)}$.
  - **Constraint 2: Being Markov.**
    Compute its Markovian projection $X_t^{(k+1)}$.
  - Set $\pi_{0,T}^{(k+1)} = \mathrm{Law}(X_0^{(k+1)}, X_T^{(k+1)})$.
- Output: $\pi_{0,T}^{(N)}$.

Fixed point of IMF

The Schrödinger Bridge $\pi^\star$ is the **unique measure** consistent with both constraints.

Exponential Convergence: Strongly Log-Concave Case

Assumptions:

- μ, ν strongly log-concave: $\nabla^2(-\log \mu) \succeq \alpha_\mu I$, idem for ν.

Exponential Convergence: Strongly Log-Concave Case

Assumptions:

- μ, ν strongly log-concave: $\nabla^2(-\log \mu) \succeq \alpha_\mu I$, idem for ν.

Theorem: For $T > \max\{\alpha_\mu^{-1}, \alpha_\nu^{-1}\}$:

$$\mathrm{KL}(\pi_{0,T}^{(n)}|\pi^\star) \leq \left(\frac{1}{T(\alpha_\mu + \alpha_\nu) - 2}\right)^n \mathrm{KL}(\pi_{0,T}^{(0)}|\pi^\star)$$

Interpretation: Convergence for strongly convex inputs and large T.

Exponential Convergence: Strongly Log-Concave Case

Assumptions:

- μ, ν strongly log-concave: $\nabla^2(-\log \mu) \succeq \alpha_\mu I$, idem for ν.

Theorem: For $T > \max\{\alpha_\mu^{-1}, \alpha_\nu^{-1}\}$:

$$\mathrm{KL}(\pi_{0,T}^{(n)} | \pi^\star) \leq \left(\frac{1}{T(\alpha_\mu + \alpha_\nu) - 2} \right)^n \mathrm{KL}(\pi_{0,T}^{(0)} | \pi^\star)$$

Interpretation: Convergence for strongly convex inputs and large T.

Remarks:

- Exponential convergence still holds under a weaker convexity assumption, but at a slower rate;
- Exponential convergence still holds when replacing R with a Langevin-type reference measure under mild structural assumptions on it.

Main Tools

1. Contraction for Markovian projection operator;

Main Tools

1. Contraction for Markovian projection operator;

2. Propagation of assumptions under Sinkhorn.