A second order equation for Schrödinger Bridges and applications

Giovanni Conforti

CEMPI Lille, Laboratoire Paul Painlevé

05/07/2017,Warsaw

Back in 1932, Schrödinger described the following experiment

• At time t = 0 consider N Brownian particles on the sphere $(X_t^i)_{t \le 1, i \le N}$ arranged according to μ ,

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_0^i}.$$

 Let each particle move independently for a unit of time and look at their configuration ν at time 1,

$$\nu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_1^i}$$

Assume that ν is an unexpected configuration, i.e. ν ≉ vol.
 What was the most likely evolution of the particle system towards ν?

Back in 1932, Schrödinger described the following experiment

• At time t = 0 consider N Brownian particles on the sphere $(X_t^i)_{t \le 1, i \le N}$ arranged according to μ ,

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_0^i}.$$

 Let each particle move independently for a unit of time and look at their configuration ν at time 1,

$$\nu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_1^i}$$

Assume that ν is an unexpected configuration, i.e. ν ≈ vol.
 What was the most likely evolution of the particle system towards ν?

Back in 1932, Schrödinger described the following experiment

• At time t = 0 consider N Brownian particles on the sphere $(X_t^i)_{t \le 1, i \le N}$ arranged according to μ ,

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_0^i}.$$

 Let each particle move independently for a unit of time and look at their configuration ν at time 1,

$$\nu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_1^i}$$

Assume that ν is an unexpected configuration, i.e. ν ≈ vol.
 What was the most likely evolution of the particle system towards ν?

Back in 1932, Schrödinger described the following experiment

• At time t = 0 consider N Brownian particles on the sphere $(X_t^i)_{t \le 1, i \le N}$ arranged according to μ ,

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_0^i}.$$

 Let each particle move independently for a unit of time and look at their configuration ν at time 1,

$$\nu = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_1^i}$$

Assume that ν is an unexpected configuration, i.e. ν ≈ vol.
 What was the most likely evolution of the particle system towards ν?

• $\mathbf{P} = \mathrm{stationary}\ \mathrm{process}\ \mathrm{for}\ \mathrm{generator}\ \mathscr{L}$

$$\mathscr{L} = \frac{1}{2}\Delta - \nabla U \cdot \nabla$$

- μ, ν initial and final configurations
- $\mathscr{H}(\cdot|\mathbf{P})=\mathsf{Relative}$ entropy on path space

Schrödinger problem

$$\min \mathcal{H}(\mathbf{Q}|\mathbf{P}), \quad X_0 \# \mathbf{Q} = \mu, X_1 \# \mathbf{Q} = \nu$$

- The optimal value $\mathcal{T}_{\mathscr{H}}(\mu, \nu)$ of the Schrödinger problem is called **entropic transportation cost**.
- The optimal solution of SP is the Schrödinger bridge (SB) between μ and ν.

• $\mathbf{P}=\text{stationary process for generator }\mathcal{L}$

$$\mathscr{L} = \frac{1}{2}\Delta - \nabla U \cdot \nabla$$

- μ, ν initial and final configurations
- $\mathscr{H}(\cdot|\mathbf{P}) =$ Relative entropy on path space

Schrödinger problem (SP)

$$\min \mathscr{H}(\mathbf{Q}|\mathbf{P}), \quad X_0 \# \mathbf{Q} = \mu, X_1 \# \mathbf{Q} = \nu$$

- The optimal value $\mathcal{T}_{\mathscr{H}}(\mu, \nu)$ of the Schrödinger problem is called **entropic transportation cost**.
- The optimal solution of SP is the Schrödinger bridge (SB) between μ and ν.

• $\mathbf{P} = \mathrm{stationary}\ \mathrm{process}\ \mathrm{for}\ \mathrm{generator}\ \mathscr{L}$

$$\mathscr{L} = \frac{1}{2}\Delta - \nabla U \cdot \nabla$$

- μ, ν initial and final configurations
- $\mathscr{H}(\cdot|\mathbf{P}) =$ Relative entropy on path space

Schrödinger problem (SP)

$$\min \mathscr{H}(\mathbf{Q}|\mathbf{P}), \quad X_0 \# \mathbf{Q} = \mu, X_1 \# \mathbf{Q} = \nu$$

- The optimal value $\mathcal{T}_{\mathscr{H}}(\mu,\nu)$ of the Schrödinger problem is called **entropic transportation cost**.
- The optimal solution of SP is the Schrödinger bridge (SB) between μ and ν.

• $\mathbf{P}=\text{stationary process for generator }\mathcal{L}$

$$\mathscr{L} = \frac{1}{2}\Delta - \nabla U \cdot \nabla$$

- μ, ν initial and final configurations
- $\mathscr{H}(\cdot|\mathbf{P})=\mathsf{Relative}$ entropy on path space

Schrödinger problem (SP)

$$\min \mathscr{H}(\mathbf{Q}|\mathbf{P}), \quad X_0 \# \mathbf{Q} = \mu, X_1 \# \mathbf{Q} = \nu$$

- The optimal value $\mathcal{T}_{\mathscr{H}}(\mu,\nu)$ of the Schrödinger problem is called **entropic transportation cost**.
- The optimal solution of SP is the Schrödinger bridge (SB) between μ and ν .

- has to minimize entropy ⇒ particles are willing to arrange theirselves according to the equilibrium configuration m.
- has to reach un **unexpected** final configuration, which looks very different from **m**
- Thus we expect the dynamics to be divided into two phases.
 - Intropy minimization dominates: SB relaxes to m
 - (2) the influence of the final configuration prevails: particles start arranging according to ν and drifts away from **m**.

At any given time t, how far is μ_t from **m** ?

- has to minimize entropy ⇒ particles are willing to arrange theirselves according to the equilibrium configuration m.
- has to reach un unexpected final configuration, which looks very different from m

Thus we expect the dynamics to be divided into two phases.

- Intropy minimization dominates: SB relaxes to m
- e the influence of the final configuration prevails: particles start arranging according to ν and drifts away from m.

At any given time t, how far is μ_t from **m** ?

- has to minimize entropy ⇒ particles are willing to arrange theirselves according to the equilibrium configuration m.
- has to reach un unexpected final configuration, which looks very different from m
- Thus we expect the dynamics to be divided into two phases.
 - Intropy minimization dominates: SB relaxes to m
 - 2 the influence of the final configuration prevails: particles start arranging according to ν and drifts away from m.

At any given time t, how far is μ_t from **m** ?

- has to minimize entropy ⇒ particles are willing to arrange theirselves according to the equilibrium configuration m.
- has to reach un unexpected final configuration, which looks very different from m
- Thus we expect the dynamics to be divided into two phases.
 - Intropy minimization dominates: SB relaxes to m
 - **②** the influence of the final configuration prevails: particles start arranging according to ν and drifts away from **m**.

At any given time t, how far is μ_t from **m**?

- has to minimize entropy ⇒ particles are willing to arrange theirselves according to the equilibrium configuration m.
- has to reach un **unexpected** final configuration, which looks very different from **m**
- Thus we expect the dynamics to be divided into two phases.
 - Intropy minimization dominates: SB relaxes to m
 - **②** the influence of the final configuration prevails: particles start arranging according to ν and drifts away from **m**.

At any given time t, how far is μ_t from **m**?

We call (μ_t) the marginal flow of SB, and (v_t) its "speed"

$$v_t = -\nabla \mathscr{H}(\mu_t)$$

- But if particles go along the gradient flow they don't reach the "unexpected configuration" ν
- Cheap trick: differentiate in time the gradient flows

$$\frac{\mathsf{D}}{dt}v_t = -\nabla^2 \mathscr{H}(\mu_t)v_t = \nabla^2 \mathscr{H}(\mu_t)\nabla \mathscr{H}(\mu_t)$$

We call (μ_t) the marginal flow of SB, and (v_t) its "speed"

$$v_t = -\nabla \mathscr{H}(\mu_t)$$

- But if particles go along the gradient flow they don't reach the "unexpected configuration" ν
- Cheap trick: differentiate in time the gradient flows

$$\frac{\mathbf{D}}{dt}\mathbf{v}_t = -\nabla^2 \mathscr{H}(\mu_t)\mathbf{v}_t = \nabla^2 \mathscr{H}(\mu_t)\nabla \mathscr{H}(\mu_t)$$

We call (μ_t) the marginal flow of SB, and (v_t) its "speed"

$$\mathbf{v}_t = -\nabla \mathscr{H}(\mu_t)$$

- But if particles go along the gradient flow they don't reach the "unexpected configuration" ν
- Cheap trick: differentiate in time the gradient flows

$$\frac{\mathsf{D}}{dt}\mathbf{v}_t = -\nabla^2 \mathscr{H}(\mu_t)\mathbf{v}_t = \nabla^2 \mathscr{H}(\mu_t)\nabla \mathscr{H}(\mu_t)$$

We call (μ_t) the marginal flow of SB, and (v_t) its "speed"

$$\mathbf{v}_t = -\nabla \mathscr{H}(\mu_t)$$

- But if particles go along the gradient flow they don't reach the "unexpected configuration" ν
- Cheap trick: differentiate in time the gradient flows

$$\frac{\mathsf{D}}{dt}\mathsf{v}_t = -\nabla^2 \mathscr{H}(\mu_t)\mathsf{v}_t = \nabla^2 \mathscr{H}(\mu_t)\nabla \mathscr{H}(\mu_t)$$

- To make sense of the calculations above we need to choose some **metric structure** on the space of probability measures
- The theory of **optimal transport** can be used to construct one such structure.
- The continuity equation gives the notion of speed for a curve

$$\partial_t \mu + \nabla \cdot (\mu_t v_t) = 0$$

• The convective derivative can be used to make sense of the covariant derivative.

$$\partial_t v_t + \frac{1}{2} \nabla |v_t|^2$$

- To make sense of the calculations above we need to choose some **metric structure** on the space of probability measures
- The theory of **optimal transport** can be used to construct one such structure.
- The continuity equation gives the notion of speed for a curve

$$\partial_t \mu + \nabla \cdot (\mu_t v_t) = 0$$

• The convective derivative can be used to make sense of the covariant derivative.

$$\partial_t v_t + \frac{1}{2} \nabla |v_t|^2$$

- To make sense of the calculations above we need to choose some **metric structure** on the space of probability measures
- The theory of **optimal transport** can be used to construct one such structure.
- The continuity equation gives the notion of speed for a curve

$$\partial_t \mu + \nabla \cdot (\mu_t v_t) = 0$$

• The convective derivative can be used to make sense of the covariant derivative.

$$\partial_t v_t + \frac{1}{2} \nabla |v_t|^2$$

- To make sense of the calculations above we need to choose some **metric structure** on the space of probability measures
- The theory of **optimal transport** can be used to construct one such structure.
- The continuity equation gives the notion of speed for a curve

$$\partial_t \mu + \nabla \cdot (\mu_t v_t) = 0$$

• The convective derivative can be used to make sense of the covariant derivative.

$$\partial_t v_t + \frac{1}{2} \nabla |v_t|^2$$

- To make sense of the calculations above we need to choose some **metric structure** on the space of probability measures
- The theory of **optimal transport** can be used to construct one such structure.
- The continuity equation gives the notion of speed for a curve

$$\partial_t \mu + \nabla \cdot (\mu_t v_t) = 0$$

• The convective derivative can be used to make sense of the covariant derivative.

$$\partial_t v_t + \frac{1}{2} \nabla |v_t|^2$$

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathbf{D}}{dt}v_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathcal{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathsf{D}}{dt}\mathsf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathscr{I}_U is the modified Fisher information.
- $abla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathsf{D}}{dt}\mathsf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathcal{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathsf{D}}{dt}\mathsf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathcal{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathsf{D}}{dt}\mathbf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathcal{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathbf{D}}{dt}\mathbf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}} \mathscr{I}_{\boldsymbol{U}}(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathcal{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

Theorem (C.' 17)

The marginal flow (μ_t) of the Schrödinger bridge solves

$$\frac{\mathbf{D}}{dt}\mathbf{v}_t = \frac{1}{8}\nabla^{\mathcal{W}}\mathscr{I}_U(\mu_t)$$

- $\frac{D}{dt}$ is the covariant derivative
- v_t is the velocity field of (μ_t)
- \mathscr{I}_U is the modified Fisher information.
- $\nabla^{\mathcal{W}}$ is the gradient in the OT sense.

At any given time t, how far μ_t from **m** ?

Theorem (C. '17)

Let M be compact and (mu_t) be the entropic interpolation between μ and ν . The Bakry Émery condition

 $\mathfrak{Ric} + \nabla \nabla U \geq \lambda$

implies the following bound for all $\mu, \nu \in \mathcal{P}(M)$

$$\mathscr{H}_{\mathcal{U}}(\mu_{t}) \leq \frac{1 - \exp(-\alpha(1-t))}{1 - \exp(-\alpha)} \mathscr{H}_{\mathcal{U}}(\mu) + \frac{1 - \exp(-\alpha t)}{1 - \exp(-\alpha)} \mathscr{H}_{\mathcal{U}}(\nu) \\ - \frac{\cosh(\frac{\alpha}{2}) - \cosh(\alpha(t-\frac{1}{2}))}{\sinh(\frac{\alpha}{2})} \mathcal{T}_{\mathscr{H}}(\mu, \nu)$$



Let m be the invariant measure associated with $\mathscr{L}.$ Then

$$\mathcal{T}_{\mathscr{H}_U}(\mu, \mathbf{m}) \leq rac{1}{1 - \exp(-rac{lpha}{2})} \mathscr{H}_U(\mu) \; ,$$

• A priori bound on the entropic transportation cost

- The entropic cost grows at most linearly with the marginal entropies
- Upper bound for the distance between the Schrödinger bridge and the target law **P**
- Partial converse to the fact that joint entropies dominate marginal entropies

Let m be the invariant measure associated with $\mathscr{L}.$ Then

$$\mathcal{T}_{\mathscr{H}_U}(\mu, \mathbf{m}) \leq rac{1}{1 - \exp(-rac{lpha}{2})} \mathscr{H}_U(\mu) \; ,$$

- A priori bound on the entropic transportation cost
- The entropic cost grows at most linearly with the marginal entropies
- Upper bound for the distance between the Schrödinger bridge and the target law **P**
- Partial converse to the fact that joint entropies dominate marginal entropies

Let m be the invariant measure associated with $\mathscr{L}.$ Then

$$\mathcal{T}_{\mathscr{H}_U}(\mu, \mathbf{m}) \leq rac{1}{1 - \exp(-rac{lpha}{2})} \mathscr{H}_U(\mu) \; ,$$

- A priori bound on the entropic transportation cost
- The entropic cost grows at most linearly with the marginal entropies
- Upper bound for the distance between the Schrödinger bridge and the target law **P**
- Partial converse to the fact that joint entropies dominate marginal entropies

Let m be the invariant measure associated with $\mathscr{L}.$ Then

$$\mathcal{T}_{\mathscr{H}_U}(\mu, \mathbf{m}) \leq rac{1}{1 - \exp(-rac{lpha}{2})} \mathscr{H}_U(\mu) \; ,$$

- A priori bound on the entropic transportation cost
- The entropic cost grows at most linearly with the marginal entropies
- Upper bound for the distance between the Schrödinger bridge and the target law **P**
- Partial converse to the fact that joint entropies dominate marginal entropies

Connections with OT

If we consider the SB constructed with

$$\mathscr{L}^{\varepsilon} = \frac{\varepsilon}{2} \Delta$$

Theorem (Mikami '04, Léonard '12

The following limits hold as $\varepsilon \downarrow 0$.

Schrödinger bridge ightarrow displacement interpolation

and

$$\varepsilon \mathcal{T}^{\varepsilon}_{\mathscr{H}_{U}}(\mu,\nu) \to \frac{1}{2}W_{2}^{2}(\mu,\nu)$$

Using this, we can recover **displacement convexity of the entropy Talagrand's entropy transportation inequality** from our results

Connections with OT

If we consider the SB constructed with

$$\mathscr{L}^{\varepsilon} = \frac{\varepsilon}{2} \Delta$$

Theorem (Mikami '04, Léonard '12)

The following limits hold as $\varepsilon \downarrow 0$.

Schrödinger bridge \rightarrow displacement interpolation

and

$$\varepsilon \mathcal{T}^{\varepsilon}_{\mathscr{H}_{U}}(\mu,\nu) \to \frac{1}{2}W_{2}^{2}(\mu,\nu)$$

Using this, we can recover **displacement convexity of the entropy Talagrand's entropy transportation inequality** from our results

Connections with OT

If we consider the SB constructed with

$$\mathscr{L}^{\varepsilon} = \frac{\varepsilon}{2} \Delta$$

Theorem (Mikami '04, Léonard '12)

The following limits hold as $\varepsilon \downarrow 0$.

Schrödinger bridge \rightarrow displacement interpolation

and

$$\varepsilon \mathcal{T}^{\varepsilon}_{\mathscr{H}_{U}}(\mu,\nu) \to \frac{1}{2}W_{2}^{2}(\mu,\nu)$$

Using this, we can recover **displacement convexity of the entropy Talagrand's entropy transportation inequality** from our results

Reciprocal characteristic

For a given potential U, it is the vector field

$$abla \mathscr{U}, \quad ext{where} \ \mathscr{U} = rac{1}{2} |
abla U|^2 - rac{1}{2} \Delta U$$

In the flat case $M = \mathbb{R}^d$ we obtained the following

Theorem (C.'17)

Let (μ_t) be the marginal flow of (SB). If \mathscr{U} is convex and μ_t is a log-concave law for all t, then

 $t \mapsto \mathscr{I}_U(\mu_t)$ is convex

• Convexity points of $\mathscr{I}_U = \log$ concave measures?

Grazie! beginframeSchrödinger '33

"Imaginez que vous observez un système de particules en diffusion, qui soient en équilibre thermodynamique. Admettons qu' à un instant donné 0 vous les ayez trouvées en répartition à peu près uniforme et que à 1 vous ayez trouvé un écart spontané et considérable par rapport à cette uniformité. On vous demande de quelle manière cet écart sest produit. Quelle en est la manière la plus probable ? "

Minimal references I



G. Conforti.

A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost.

preprint arXiv:1704.04821, 2017.

G. Conforti and M. Von Renesse. Couplings,gradient estimates and logarithmic Sobolev inequality for Langevin bridges. preprint, available at https://arxiv.org/abs/1612.08546.



A.J. Krener.

Reciprocal diffusions in flat space.

Probability Theory and Related Fields, 107(2):243–281, 1997.

Minimal references II

C. Léonard.

From the Schrödinger problem to the Monge–Kantorovich problem.

Journal of Functional Analysis, 262(4):1879–1920, 2012.

T. Mikami.

Monge s problem with a quadratic cost by the zero-noise limit of h-path processes.

Probability Theory and Related Fields, 129(2):245–260, 2004.

E. Schrödinger.

La théorie relativiste de l'électron et l'interprétation de la mécanique quantique.

Ann. Inst Henri Poincaré, (2):269 - 310, 1932.