Approximations of displacement interpolations by entropic interpolations

Christian Léonard

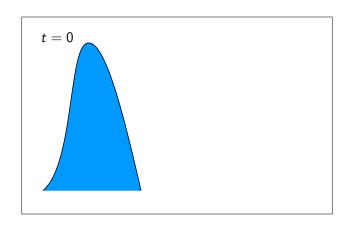
Université Paris Ouest

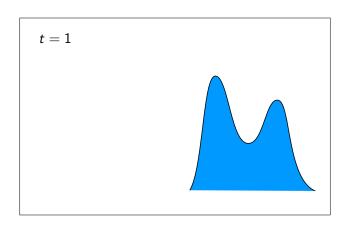
1ère journée MAS-MODE IHP - 9 janvier 2017

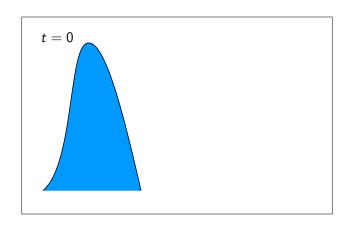
- ullet ${\cal X}$: Riemannian manifold (state space)
- ullet $\mathcal{P}(\mathcal{X})$: set of all probability measures on \mathcal{X}
- $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$
- ullet interpolate between μ_0 and μ_1

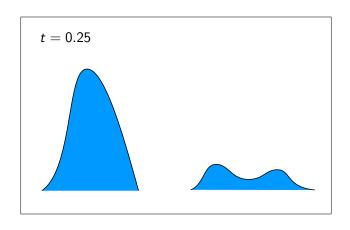
• Standard affine interpolation between
$$\mu_0$$
 and μ_1 $\mu_t^{\mathrm{aff}} := (1-t)\mu_0 + t\mu_1 \in \mathcal{P}(\mathcal{X}), \ 0 \leq t \leq 1$

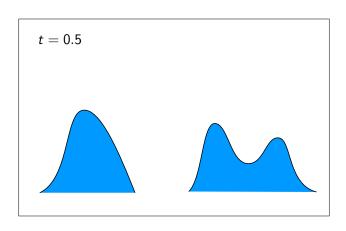


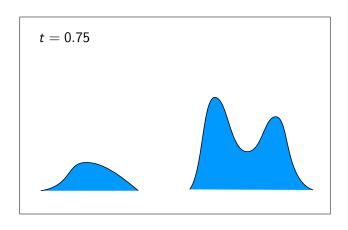


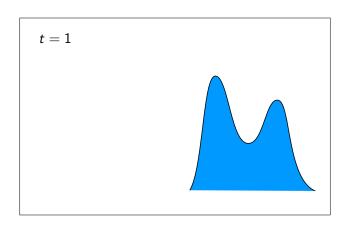










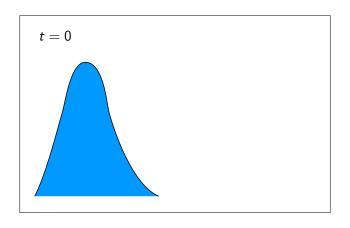


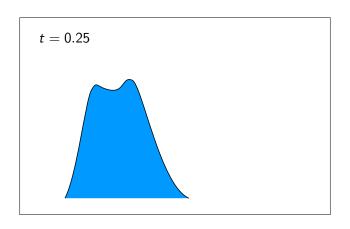
Affine interpolations require mass transference with infinite speed

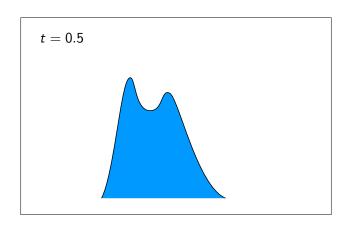


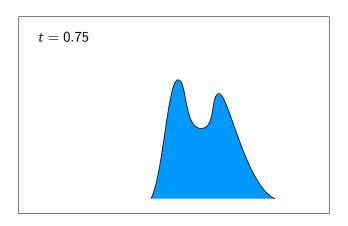
- ullet Denial of the geometry of ${\mathcal X}$
- We need interpolations built upon trans-portation, not tele-portation

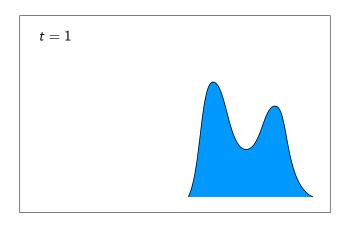
• We seek interpolations of this type	

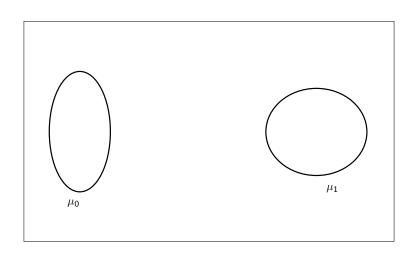


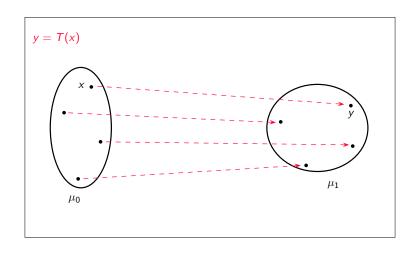


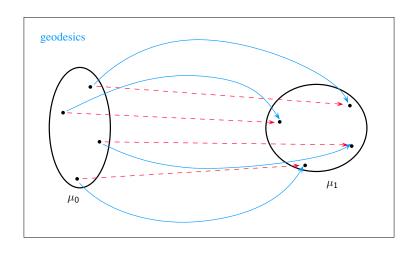


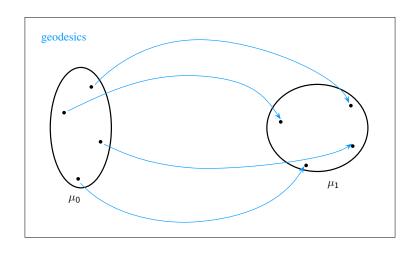


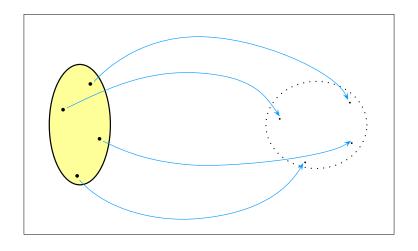


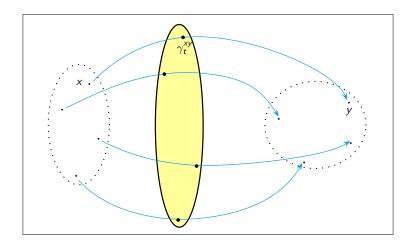


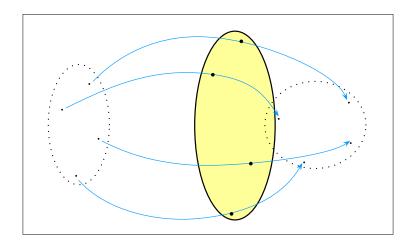


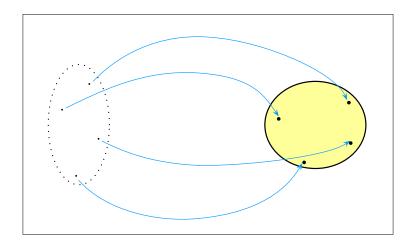


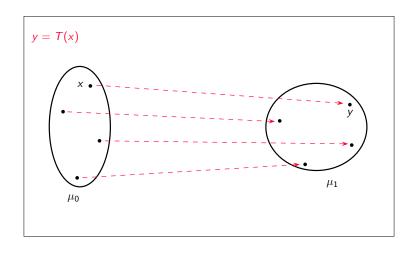












Respect geometry

- we have already used geodesics
- how to choose y = T(x) such that interpolations encrypt curvature as best as possible?
- no shock
- perform optimal transport

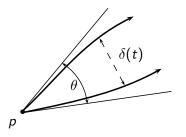
Monge's problem

$$\int_{\mathcal{X}} d^2(x, T(x)) \, \mu_0(dx) \mapsto \min; \qquad T: \quad T_\# \mu_0 = \mu_1$$

• d : Riemannian distance

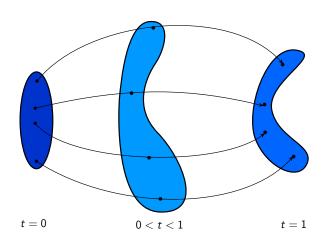
Curvature

- geodesics and curvature are intimately linked
- several geodesics give information on the curvature



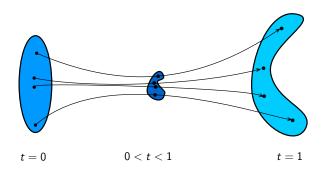
$$\delta(t) = \sqrt{2(1-\cos\theta)} \ t \left(1 - \frac{\sigma_{p}(\mathsf{S})\cos^{2}(\theta/2)}{6}t^{2} + O(t^{4})\right)$$

Lazy gas experiment



Positive curvature

Lazy gas experiment



Negative curvature

Curvature and displacement interpolations

Relative entropy

$$H(p|r) := \int \log(dp/dr) dp$$
, p, r : probability measures

Convexity of the entropy along displacement interpolations

The following assertions are equivalent

- Ric $\geq K$
- along any $[\mu_0, \mu_1]^{\mathrm{disp}} = (\mu_t)_{0 \le t \le 1}, \; \frac{d^2}{dt^2} H(\mu_t | \mathrm{vol}) \ge KW_2^2(\mu_0, \mu_1)$
- von Renesse-Sturm (04)
- W₂ is the Wasserstein distance
- starting point of the Lott-Sturm-Villani theory

Schrödinger's thought experiment

Consider a huge collection of non-interacting identical Brownian particles. If the density profile of the system at time t=0 is approximately $\mu_0 \in \mathcal{P}(\mathbb{R}^3)$, you expect it to evolve along the heat flow:

$$\begin{cases} \nu_t = \nu_0 e^{t\Delta/2}, & 0 \le t \le 1 \\ \nu_0 = \mu_0 \end{cases}$$

where Δ is the Laplace operator.

Suppose that you observe the density profile of the system at time t=1 to be approximately $\mu_1 \in \mathcal{P}(\mathbb{R}^3)$ with μ_1 different from the expected ν_1 . Probability of this rare event $\simeq \exp(-\mathit{CN}_{\mathrm{Avogadro}})$.

Schrödinger's question (1931)

Conditionally on this very rare event, what is the *most likely path* $(\mu_t)_{0 \le t \le 1} \in \mathcal{P}(\mathbb{R}^3)^{[0,1]}$ of the evolving profile of the particle system?

Schrödinger's problem

- ullet : Riemannian manifold
- $\Omega := \{ \text{paths} \} \subset \mathcal{X}^{[0,1]}$
- $R \in \mathcal{P}(\Omega)$: Wiener measure (Brownian motion)
- ullet $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ are the initial and final prescribed profiles

Schrödinger's problem

$$H(P|R) \rightarrow \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (S_{dyn})

Sanov's theorem

Schrödinger's problem

- $\omega = (\omega_t)_{0 \le t \le 1} \in \Omega$
- $\bullet \ \, X_t:\omega\in\Omega\mapsto\omega_t\in\mathcal{X},\ \, 0\leq t\leq 1\qquad \text{(canonical process)}$
- $P_t(dz) := [(X_t)_{\#}P](dz) = P(X_t \in dz) \in \mathcal{P}(\mathcal{X}), \quad P \in \mathcal{P}(\Omega)$

Definition. R-entropic interpolation

$$[\mu_0, \mu_1]^R := (P_t)_{0 \le t \le 1}$$
 with P the unique solution of (S_{dyn}) .

It is the answer to Schrödinger's question

Monge-Kantorovich problem

• $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ are the initial and final prescribed profiles

Monge problem

$$\int_{\mathcal{X}} d^2(x, \mathbf{T}(x)) \, \mu_0(dx) \mapsto \min; \qquad \mathbf{T} : \ T_{\#} \mu_0 = \mu_1$$

Monge-Kantorovich problem

$$\int_{\mathcal{X}^2} d^2(x,y) \, \pi(dxdy) \to \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi = \mu_1$$

•
$$\pi^T(dxdy) = \mu_0(dx)\delta_{T(x)}(dy)$$

Dynamical Monge-Kantorovich problem

Monge-Kantorovich problem

$$\int_{\mathcal{X}^2} d^2(x, y) \, \pi(dxdy) \to \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi = \mu_1 \qquad \text{(MK)}$$

Dynamical Monge-Kantorovich problem

$$E_PA \to \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (MK_{dyn})

where
$$A = \int_0^1 |\dot{X}_t|^2 dt \in [0, \infty],$$

- $\min \{A(\omega); \omega_0 = x, \omega_1 = y\} = d^2(x, y)$
- $\inf(MK) = \inf MK_{dyn} =: W_2^2(\mu_0, \mu_1)$

Dynamical Monge-Kantorovich and Schrödinger problems

Dynamical Monge-Kantorovich problem

$$E_PA \to \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (MK_{dyn})

where $A:=\int_0^1 |\dot{X}_t|^2 dt \in [0,\infty]$

Schrödinger's problem

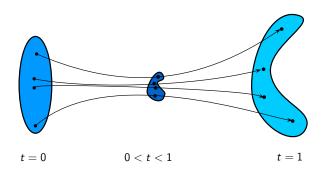
$$H(P|R) \to \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (S_{dyn})

where $H(P|R) := \int_{\Omega} \log(dP/dR) dP$

Lazy gas experiments

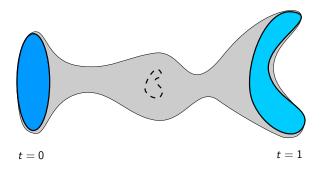
- Lazy gas experiment at zero temperature (Monge)
 - Zero temperature
 - Displacement interpolations
 - Optimal transport
- Lazy gas experiment at positive temperature (Schrödinger)
 - Positive temperature
 - Entropic interpolations
 - Minimal entropy

Lazy gas experiments



Negative curvature Zero temperature

Lazy gas experiments



Negative curvature Positive temperature

Cooling down

Aim

Drifting from Schrödinger problem to an optimal transport problem

To decrease temperature:

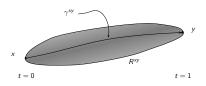
- slow down the particles of the heat bath
- more generally, decrease fluctuations

Slowed down reference measures

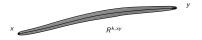
- ullet $(B_t)_{t\geq 0}$: Brownian motion on the Riemannian manifold ${\mathcal X}$
- R: law of $(B_t)_{0 \le t \le 1}$
- R^k : law of $(B_{t/k})_{0 \le t \le 1}$
- $k \to \infty$

Cooling down

k = 1:



k = 10:



 $k = \infty$:



Cooling down

- $N \to \infty, \ k=1$: the whole particle system performs a rare event to travel from μ_0 to μ_1
 - cooperative behavior
 - lacktriangleright Gibbs conditioning principle (thermodynamical limit: $N o\infty$)
- $N=1,\ k\to\infty$: each individual particle faces a hard task and must travel along an approximate geodesic
 - individual behavior
 - ▶ large deviation principle (cooling down limit: $k \to \infty$)

Cooling down principle

The cooled down sequence $(R^k)_{k\geq 1}$ encodes some geometry

• $N \to \infty, k \to \infty$: these two behaviors superpose

Results

Results 1

- displacement interpolations feel curvature
- entropic interpolations also feel curvature

Results 2

- entropic interpolations converge to displacement interpolations
- entropic interpolations regularize displacement interpolations

I-convergence

Results

Results 2 (continued)

The same kind of results hold in other settings

- (a) discrete graphs
- (b) Finsler manifolds
- (c) interpolations with varying mass
- (a) graphs: random walk
- (b) Finsler: jump process in a manifold, (work in progress)
- (c) varying mass: branching process, (work in progress)

Results

Results 3

Schrödinger's problem is an analogue of Hamilton's least action principle. It allows for *dynamical* theories of

- diffusion processes
- random walks on graphs
- stochastic Newton equation
- acceleration is related to curvature

Remainder of the talk

Let us give some details about Results 2:

entropic interpolations converge to displacement interpolations

Schrödinger's problem

- $P_{01} = (X_0, X_1)_{\#} P \in \mathcal{P}(\mathcal{X}^2)$
- $P^{xy} = P(\cdot \mid X_0 = x, X_1 = y) \in \mathcal{P}(\Omega)$: bridge
- $P(\cdot) = \int_{\mathcal{X}^2} P^{xy}(\cdot) P_{01}(dxdy) \in \mathcal{P}(\Omega)$

Result

If it exists, the unique solution P of (S_{dyn}) satisfies

- \bullet $P^{xy} = R^{xy}, \forall x, y$
- $P(\cdot) = \int_{\mathcal{X}^2} R^{xy}(\cdot) \pi(dxdy)$

where $P_{01} = \pi \in \mathcal{P}(\mathcal{X}^2)$ is the unique solution of (S) below

• $\inf(S) = H(P|R) = H(P_{01}|R_{01})$

Schrödinger's problem

$$H(\pi|R_{01}) \to \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi_1 = \mu_1$$
 (S)

•
$$H(P|R) = H(P_{01}|R_{01}) + \int_{\mathcal{X}^2} H(P^{xy}|R^{xy}) P_{01}(dxdy)$$

Schrödinger's problem

Result

Assume: R is m-stationary Markov, $m \otimes m \ll R_{01} \ll m \otimes m$, $H(\mu_0|m), H(\mu_1|m) < \infty, \dots$

Then, (S_{dyn}) and (S) admit a solution.

- long history: Schrödinger, Bernstein, Fortet, Beurling, Csiszár,
 Rüschendorf & Thomsen, Föllmer & Gantert, L.
- and also: Jamison, Zambrini, Dai Pra, Wakolbinger, Pavon, Mikami, Rœlly, Thieullen, . . .

- suppose that in addition the heat bath is pretty cold
- cooling down is (mostly) slowing down the particles
- $R = \text{Law}((B_t)_{0 \le t \le 1}), \quad R^k := \text{Law}((B_{t/k})_{0 \le t \le 1}), \quad k \to \infty$

$$H(P|R^{k})/k \to \min; \qquad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (S_{dyn})

- $\bullet \ dX_t = dB_t, \qquad R-a.s.$
- $dX_t = \sqrt{1/k} dB_t$, R^k -a.s.
- $R_{01}^k(dxdy) = (2\pi/k)^{-d/2} \exp(-k|y-x|^2/2) dxdy$
- with P^k solution of (S_{dyn}^k)

$$\inf(S_{\mathrm{dyn}}^k) = \inf(S^k) = \int_{\mathcal{X}^2} \frac{1}{2} |y - x|^2 P_{01}^k(dxdy) + o_{k \to \infty}(1)$$

this suggests that

$$"\lim_{k\to\infty}(\mathsf{S}^k)=(\mathsf{MK})"$$

Cooled down Schrödinger problem

$$H(\pi|R_{01}^k)/k \to \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi_1 = \mu_1$$
 (S^k)

Monge-Kantorovich problem

$$\int_{\mathcal{X}^2} \frac{1}{2} |y - x|^2 \, \pi(\mathbf{d}x \mathbf{d}y) \to \min; \quad \pi \in \mathcal{P}(\mathcal{X}^2) : \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (\mathsf{MK})$$

Theorem

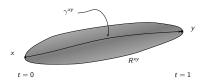
- Γ $\lim_{k\to\infty} (\mathsf{S}^k) = (\mathsf{MK})$
- $\lim_{k\to\infty}\inf(\mathsf{S}^k)=\inf(\mathsf{MK}):=W_2^2(\mu_0,\mu_1)/2$
- " $\lim_{k\to\infty} \pi^k = \pi$ ": solution of (MK)

Mikami (2004), L. (2012)

Cooling particles down brings geometry (Brownian case)

 $\lim_{k \to \infty} R^{k,xy} = \delta_{\gamma^{xy}}, \quad \gamma^{xy}$: constant speed geodesic

• k = 1:



• k = 10:



• $k = \infty$:



Convergence schema

$$P^{k}(d\omega) = \int_{\mathcal{X}^{2}} R^{k,xy}(d\omega) \quad \pi^{k}(dxdy)$$

$$\downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$$

$$P(d\omega) = \int_{\mathcal{X}^{2}} \delta_{\gamma^{xy}}(d\omega) \quad \pi(dxdy)$$

Entropic interpolations converge to displacement interpolations

$$\begin{array}{rcl} \mu_t^k(dz) & = & \int_{\mathcal{X}^2} & R_t^{k,xy}(dz) & \pi^k(dxdy), & & 0 \leq t \leq 1 \\ \downarrow & & \downarrow & & \downarrow \\ \mu_t(dz) & = & \int_{\mathcal{X}^2} & \delta_{\gamma_t^{xy}}(dz) & \pi(dxdy), & & 0 \leq t \leq 1 \end{array}$$

McCann's displacement interpolation

$$[\mu_0, \mu_1]^{\text{disp}} := (\mu_t)_{0 \le t \le 1}$$

Doubly indexed large deviation principle

• choose $(R^k)_{k>1}$ such that it satisfies some LDP

LDP for $(R^k)_{k>1}$

- $R^{k,x} \underset{k \to \infty}{\approx} \exp(-a_k[C + \iota_{\{X_0 = x\}}]), \quad \forall x$
- $a_k \to \infty$, $C: \Omega \to [0, \infty]$, coercive
- $\{C = 0\}$ is the limiting support of "all geodesics"

Example (slow Brownian motion)

$$a_k = k$$
, $C(\omega) = \int_0^1 \frac{1}{2} |\dot{\omega}_t|^2 dt$, $\lim_{k \to \infty} R^{k,xy} = \delta_{\gamma^{xy}}$

Γ-convergence

Dynamical cooled down Schrödinger problem

$$H(P|R^k)/a_k \to \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1^k$$
 (S_{dyn})

Dynamical Monge-Kantorovich problem

$$E_PC o \min; \quad P \in \mathcal{P}(\Omega) : P_0 = \mu_0, P_1 = \mu_1$$
 (MK_{dyn})

Theorem

- there exists $\mu_1^k o \mu_1$ such that $\Gamma ext{-lim}_{k o\infty}(\mathsf{S}^k_{ ext{dyn}})=(\mathsf{MK}_{ ext{dyn}})$
 - $\lim_{k\to\infty}\inf(\mathsf{S}_{\mathrm{dyn}}^k)=\inf(\mathsf{MK}_{\mathrm{dyn}})$
 - " $\lim_{k\to\infty} P^k = P$:" solution of (MK_{dyn})

Γ-convergence

ullet if $P^k o P$ we get the following schema

Convergence schema

$$P^{k}(d\omega) = \int_{\mathcal{X}^{2}} R^{k,xy}(d\omega) \pi^{k}(dxdy)$$

$$\downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$$

$$P(d\omega) = \int_{\mathcal{X}^{2}} G^{xy}(d\omega) \pi(dxdy)$$

Entropic interpolations converge to displacement interpolation

$$egin{array}{lll} \mu_t^k(extit{d}z) &=& \int_{\mathcal{X}^2} & R_t^{k, ext{xy}}(extit{d}z) & \pi^k(extit{d}x extit{d}y), & 0 \leq t \leq 1 \ \downarrow & & \downarrow & \downarrow \ \mu_t(extit{d}z) &=& \int_{\mathcal{X}^2} & G_t^{ extit{xy}}(extit{d}z) & \pi(extit{d}x extit{d}y), & 0 \leq t \leq 1 \end{array}$$

Definition (displacement interpolation)

$$[\mu_0, \mu_1]^{\text{disp}} := (\mu_t)_{0 \le t \le 1}$$

L^2 -type displacement interpolations on a vector space

- $\mathcal{X} = \mathbb{R}^n$
- ullet $R^{k,x}: \quad Z^{k,x}_t = x + k^{-1} \sum_{j=1}^{\lfloor kt
 floor} V_j, \quad (V_j)_{j \geq 1} \ ext{i.i.d.} \ , \ \mathbb{E} V = 0$

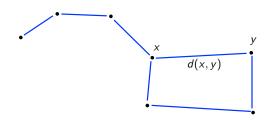
Mogulskii's theorem

$$C(\omega) = \int_0^1 L_V(\dot{\omega}_t) dt, \qquad c(x,y) = L_V(y-x)$$

- $L_V(v) = \sup_{p} \{p \cdot v H_V(p)\}; \quad H_V(p) = \log \mathbb{E} \exp(p \cdot V)$
- $L_V(v) = O_{v\to 0}(|v|^2)$
- $c(x,y) = \inf\{C(\omega); \omega \in \Omega : \omega_0 = x, \omega_1 = y\}$

Interpolations on a discrete graph

• metric graph (\mathcal{X},\sim,d)



 $x \sim y$ means that (x, y) is an edge

- ullet length $\ell(\omega) := \sum_{0 \leq t \leq 1} \mathbf{1}_{\{\omega_t =
 eq \omega_t\}} d(\omega_{t^-}, \omega_t)$
- intrinsic distance $d(x,y) = \inf \{\ell(\omega) : \omega \in \Omega, \omega_0 = x, \omega_1 = y\}$

Interpolations on a discrete graph

- to recover d:
 - ► slow down the walk
 - condition at t = 0 and t = 1
- reference walk: $R \in \mathcal{P}(\Omega)$ with jump kernel $J_x(dy) = \sum_{y:y \sim x} J_x(y) \, \delta_y$

Lazy random walks
$$R^k$$

$$J_x^k(dy) = \sum_{y:y \sim x} k^{-d(x,y)} J_x(y) \delta_y$$

Interpolations on a discrete graph

Geodesics

$$\Gamma^{xy} := \{ \omega \in \Omega; \omega_0 = x, \omega_1 = y, \ell(\omega) = d(x, y) \}$$

$$\Gamma := \cup_{x,y} \Gamma^{xy}$$

Convergence of bridges

$$\lim_{k \to \infty} R^{k,xy} = G^{xy} \in \mathcal{P}(\Gamma^{xy})$$

•
$$G := \mathbf{1}_{\Gamma} e^{\int_0^1 J_{X_t}(\mathcal{X}) dt} R$$

Convergence of the interpolations

$$a_k = \log k$$
, $C = \ell$, $c = d$

•
$$\lim_{k\to\infty}\inf(\mathsf{S}^k)=W_1(\mu_0,\mu_1)$$

L¹-type interpolations on a diffuse length space

ullet (\mathcal{X},d): diffuse metric space

Definition (diffuse metric space)

 (\mathcal{X},d) is diffuse if there exists a Borel measure m on \mathcal{X} such that

- $\sup_{x} m(B_{x}^{1}) < \infty$
- $m(B_x^{\epsilon}) > 0$, $\forall x \in \mathcal{X}, \forall \epsilon > 0$
- $\bullet \ B_{\mathsf{x}}^{\epsilon} := \{ y \in \mathcal{X} : d(\mathsf{x}, \mathsf{y}) < \epsilon \}$

L¹-type interpolations on a diffuse length space

Reference processes

- $R \leftrightarrow J_{\times}(dy) = \mathbf{1}_{B^1_{\circ}}(y) \, m(dy)$
- $\bullet \ R^k \leftrightarrow J_x^k(dy) = e^{-1} \mathbf{1}_{S_x^{1/k}}(y) \ m(dy)$
- $S_X^{\epsilon} := B_X^{\epsilon} \setminus B_X^{\epsilon \epsilon^2}$

Convergence of the entropic interpolations

$$a_k = k$$
, $C = \text{length}$, $c = d$

- $H(P|R^k) = H(P|R) E_P \log(dR^k/dR)$
- $-\log(dR^k/dR)/k = \#\{t: X_{t^-} \neq X_t\}/k + O_{k\to\infty}(1/k)$ = $\operatorname{length}(X) + O_{k\to\infty}(1/k)$
- work in progress with Luca Tamanini

References about entropic interpolations

- **1** T. Mikami. Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. PTRF, (2004)
- 2 L. From the Schrödinger problem to the Monge-Kantorovich problem. JFA, (2012)
- **1** L. A survey of the Schrödinger problem and some of its connections with optimal transport. DCDS (A), (2014)
- L. Lazy random walks and optimal transport on graphs. AoP, (2016)

