Particle Approximation to the Wasserstein Diffusion

Karl-Theodor Sturm

Universität Bonn

Outline

- **The Entropic Measure in $d = 1$**
- **The Entropic Measure on Multidimensional Spaces**
- **The Wasserstein Diffusion**
- **Particle Approximations**

Part 1

The Entropic Measure in $d = 1$

Riemannian Structure of $\mathcal{P}_2(M)$ for Riemannian M

L^2-Wasserstein distance

$$d_W(\mu_0, \mu_1) = \inf \left\{ \left[\int_{M \times M} d^2(x, y)\, d\, q(x, y) \right]^{1/2} : \quad q \text{ has marginals } \mu_0, \mu_1 \right\}$$

induces Riemmanian structure on the "Wasserstein space" $\mathcal{P}_2(M)$.

Consequence:

First order calculus, gradient flows on $\mathcal{P}_2(M)$

The gradient flow $\frac{\partial \nu}{\partial t} = -\nabla S(\nu)$ on $\mathcal{P}_2(M)$ for the relative entropy relative entropy $S : \mathcal{P}_2(M) \to [-\infty, \infty]$ with

$$S(\nu) = \begin{cases} \int \rho \cdot \log \rho \, dx, & \text{if } d\nu = \rho \, dx \\ +\infty, & \text{if } d\nu \not\ll dx \end{cases}$$

is given by $\nu_t(dx) = \rho_t(x) dx$ where ρ solves the heat equation $\frac{\partial}{\partial t} \rho = \triangle \rho$ on M.

Challenge:

- Second order calculus
- stochastic differential equations on $\mathcal{P}_2(M)$

Stochastic Dynamics on $\mathcal{P}_2(M)$

Stochastic process $\mu_t(\omega)$ on $\mathcal{P}_2(M)$ with invariant distribution $\mathbb{P}^\beta$

$$d\mu_t = -\beta \cdot \nabla S(\mu_t) dt \quad + \quad \text{noise}$$

Dirichlet form

$$\mathcal{E}(u,u) = \int_{\mathcal{P}_2(M)} \|\nabla u\|^2(\mu)\, d\mathbb{P}^\beta(\mu)$$

where ∇u denotes the gradient w.r.t. the Riemannian structure of $\mathcal{P}_2(M)$

Canonical measure on $\mathcal{P}_2(M)$

$$d\mathbb{P}^\beta(\mu) = \frac{1}{Z} \exp\left(-\beta \cdot S(\mu)\right) \cdot d\mathbb{P}^0(\mu)$$

Particular Case: d=1, say $M = [0,1]$

Continuous, one-to-one correspondence:

$$\mu \quad \overset{distrib.\,function}{\longleftrightarrow} \quad f \quad \overset{right\ \ inverse}{\longleftrightarrow} \quad g \quad \overset{distrib.\,function}{\longleftrightarrow} \quad \nu$$

between $\mathcal{P}_2(M)$ and $\mathcal{G} = \{\text{incr., right cont. } g : [0,1] \to [0,1]\}$.

New ansatz:

$$d\mathbb{Q}^\beta(g) = \frac{1}{Z} \exp\left(-\beta \cdot S(g)\right) \cdot d\mathbb{Q}^0(g)$$

with $S(g) = -\int_0^1 \log g_t' \, dt$ if $S(\mu) =$ relative entropy.

(Cf. Construction of Wiener measure with $H(g) = \frac{1}{2}\int_0^1 g_t'^2 dt$.)

Heuristic Derivation of the Wiener Measure

Recall heuristic construction of Wiener measure as

$$d\mathbf{P}^{\beta}(g) = \frac{1}{Z} \exp\left(-\beta \cdot H(g)\right) \cdot d\mathbf{P}^{0}(g)$$

with $H(g) = \frac{1}{2} \int_0^1 g_t'^2 dt$.

Finite dimensional approximation yields

$$\begin{aligned} &\mathbf{P}^{\beta}\left(g_{t_1} \in dx_1, \ldots, g_{t_n} \in dx_n\right) \\ = \quad & \frac{1}{Z'} \exp\left(-\beta \sum_{i=1}^{n} \frac{|x_i - x_{i-1}|^2}{2(t_i - t_{i-1})}\right) \quad dx_1 \ldots dx_n \end{aligned}$$

$\rightsquigarrow$ rigorous construction via Kolmogorov

$\rightsquigarrow$ law of Brownian motion

Heuristic Derivation of the Entropic Measure

Similarly, the ansatz

$$d\mathbb{Q}^{\beta}(g) = \frac{1}{Z} \exp\left(-\beta \cdot S(g)\right) \cdot d\mathbb{Q}^{0}(g)$$

with $S(g) = -\int_0^1 \log g_t' \, dt$ leads to

$$\begin{aligned} &\mathbb{Q}^{\beta}\left(g_{t_1} \in dx_1, \ldots, g_{t_n} \in dx_n\right) \\ &= \quad \frac{1}{Z'} \exp\left(\beta \sum_{i=1}^{n+1} \log \frac{x_i - x_{i-1}}{t_i - t_{i-1}} \cdot (t_i - t_{i-1})\right) \ \mathbb{Q}^{0}(dx_1 \ldots dx_n) \\ &= \quad \frac{1}{Z''} \prod_{i=1}^{n+1} (x_i - x_{i-1})^{\beta(t_i - t_{i_1})} \ \frac{dx_1 \ldots dx_n}{x_1 \cdot (x_2 - x_1) \cdot \ldots \cdot (1 - x_n)} \end{aligned}$$

$\rightsquigarrow$ consistent family, projective limit

$\rightsquigarrow$ $(g_t)_{0 \leq t \leq 1}$ is Dirichlet-Ferguson process (normalized Gamma process), i.e.

$$g_t \overset{(d)}{=} \frac{X_{\beta t}}{X_{\beta}}.$$

Part 2

The Entropic Measure on Multidimensional Spaces

General Case: M compact Riemannian, m=vol

Correspondence:

$$\mu \quad \overset{distrib.function}{\longleftrightarrow} \quad f \quad \overset{right\ \ inverse}{\longleftrightarrow} \quad g \quad \overset{distrib.function}{\longleftrightarrow} \quad \nu$$

should be re-interpreted in terms of Brenier maps:

$$\mu = g_* m, \qquad f_* m = \nu.$$

They are given as

$$g = \exp(\nabla \varphi), \qquad f = \exp(\nabla \psi)$$

where φ and ψ are conjugate $d^2/2$-convex functions. That is,

$$\varphi(x) = -\inf_{y \in M} \left[\frac{1}{2} d^2(x, y) + \psi(y) \right], \qquad \psi(y) = -\inf_{x \in M} \left[\frac{1}{2} d^2(x, y) + \varphi(x) \right].$$

If $\mu \ll m$ then

$$f = g^{-1}.$$

Theorem. The conjugation map $\mathfrak{C} : \mu \mapsto \nu$ is continuous and involutive on $\mathcal{P}(M)$.

Entropic Measure on Multidimensional Spaces: Heuristics

Ansatz for distribution of μ

$$d\mathbb{P}^{\beta}(\mu) = \frac{1}{Z} \exp\left(-\beta \cdot \mathrm{Ent}(\mu|m)\right) \cdot d\mathbb{P}^{0}(\mu)$$

leads to ansatz for distribution of $\nu = \mu^{\mathfrak{c}}$

$$d\mathbb{Q}^{\beta}(\nu) = \frac{1}{Z} \exp\left(-\beta \cdot \mathrm{Ent}(m|\nu)\right) \cdot d\mathbb{Q}^{0}(\nu)$$

since $\mathrm{Ent}(\mu|m) = \mathrm{Ent}(m|\mu^{\mathfrak{c}})$.

Given a partition $M = \bigcup_{i=1}^{N} M_i$, replace the conditional expectation of $\mathrm{Ent}(m \mid \cdot\,)$ under the constraint $\nu(M_1) = x_1, \ldots, \nu(M_N) = x_N$ by the minimum of $\nu \mapsto \mathrm{Ent}(m \mid \nu)$ under the constraint $\nu(M_1) = x_1, \ldots, \nu(M_N) = x_N$.
Obviously, this minimum is attained at a measure with constant density on each of the sets M_i of the partition. Hence,

$$\mathbb{Q}^{\beta}\left((\nu(M_1), \ldots, \nu(M_N)) \in dx\right) = c \cdot x_1^{\beta \cdot m(M_1)} \cdot \ldots \cdot x_N^{\beta \cdot m(M_N)} \, q_N(dx).$$

Choosing $q_N(dx) = c_N \cdot \delta_{\{\sum_{i=1}^{N} x_i = 1\}} \frac{dx_1 \ldots dx_N}{x_1 \cdot x_2 \cdot \ldots \cdot x_N}$ yields the finite dimensional distributions of the

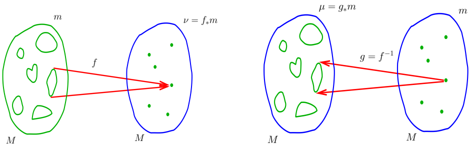
Dirichlet-Ferguson process.

Entropic Measure on Multidimensional Spaces: Rigorous

$\mathbb{Q}^\beta$ = distrib. of Dirichlet-Ferguson process $\nu = \sum_{k=1}^\infty \lambda_k \cdot \delta_{x_k}$

Random probability measure with finite dimensional distributions

$$\mathbb{Q}^\beta\left((\nu(M_1), \ldots, \nu(M_N)) \in dx\right) = \frac{\Gamma(\beta)}{\prod_{i=1}^N \Gamma(\beta m(M_i))} \cdot x_1^{\beta \cdot m(M_1)-1} \cdot \ldots \cdot x_N^{\beta \cdot m(M_N)-1} \, \delta_{\{\sum_{i=1}^N x_i = 1\}} dx$$



Entropic measure $\mathbb{P}^\beta = \mathfrak{C}_* \mathbb{Q}^\beta$

Almost every μ has no density and no atoms ('Cantor like')

Part 3

The Wasserstein Diffusion

Change of Variables Formula, until now only in $d = 1$

Change of variables formula for $\mathbb{P}^{\beta}(d\mu)$ under transformation $\mu \mapsto h_* \mu$

$$\Updownarrow$$

Change of variables formula for $\mathbb{Q}^{\beta}(dg)$ under transformation $g \mapsto h \circ g$

"Girsanov" theorem for Dirichlet process:

$$\mathrm{Law}\,(h(g)) \;=\; Y_h(g) \cdot \mathrm{Law}(g)$$

with some explicitly known density Y_h ("Quasi-Invariance")

vRenesse/Sturm, vRenesse/Yor/Zambotti

Quasi-Invariance

Theorem. For smooth, strictly increasing $h \in \mathcal{G}$:

$$d\mathbb{Q}^\beta(h \circ g) = Y_h^\beta(g) \cdot d\mathbb{Q}^\beta(g)$$

with

$$Y_h^\beta(g) = \underbrace{\exp\left(+\beta \int_0^1 \log h'(g_t)dt\right)}_{\exp\left(-\beta\left[S(h\circ g) - S(g)\right]\right)} \cdot \prod_{t \in J(g)} \frac{\sqrt{h'(g_{t-}) \cdot h'(g_{t+})}}{\frac{h(g_{t+})-h(g_{t-})}{g_{t+}-g_{t-}}}$$

Consequences of Quasi-Invariance

- Integration by parts for directional derivative

$$D_\varphi u(\mu) = \lim_{t\longrightarrow 0} \frac{1}{t}\left[u((I+t\varphi)_*\mu) - u(\mu)\right]$$

- Closability of Dirichlet form on $L^2(\mathcal{P}, \mathbb{P}^\beta)$

$$\mathcal{E}(u,u) \quad = \quad \int_{\mathcal{P}} \|Du\|^2(\mu)\, d\mathbb{P}^\beta(\mu)$$

- $\exists$ Wasserstein diffusion: reversible stoch. process $(\mu_t)_{t\geq 0}$ on $\mathcal{P}_2$

- Laplacian on the Wasserstein space

$$\begin{aligned} Lu(\mu) \quad &= \quad U''(.)\cdot \int \alpha'^2 d\mu \\ &+ \quad U'(.)\cdot \beta \int \alpha'' d\mu \\ &+ \quad U'(.)\cdot \sum_{I\in \mathrm{gaps}(\mu)} \left[\frac{\alpha''(I_+)+\alpha''(I_-)}{2} - \frac{\alpha'(I_+)-\alpha'(I_-)}{|I|}\right] \end{aligned}$$

for cylinder functions $u(\mu) = U(\int_0^1 \alpha d\mu)$, directional derivative $D_\varphi u(\mu) = U'(.)\cdot \int \alpha'\varphi d\mu$

Wasserstein Diffusion

Square field operator on $L^2(\mathcal{P}_2(M), \mathbb{P}^\beta)$

$$\Gamma(u,u)(\mu) = \|\nabla u\|^2(\mu)$$

Intrinsic distance = L^2-Wasserstein distance, Rademacher theorem:

- Every 1-Lipschitz function u on the L^2-Wasserstein space $\mathcal{P}_2$ belongs to the domain of the Dirichlet form and $\Gamma(u,u) \leq 1$ a.e.
- Every continuous function u in the domain of the Dirichlet form with $\Gamma(u,u) \leq 1$ a.e. is 1-Lipschitz on $\mathcal{P}_2$.

Gaussian short time asymptotic (Hino/Ramirez)

$$\lim_{t \to 0} t \log p_t(A,B) = -\frac{d_W(A,B)^2}{2}$$

Logarithmic Sobolev inequality, Poincaré inequality with constant $\frac{1}{\beta}$

(Döring/Stannat)

Part 4

Particle Approximation

The Particle System

Consider the interacting system of stochastic differential equations

$$dX_t^i = \frac{\partial \log \rho_k^\beta}{\partial x_i}(X_t)\,dt + \sqrt{2}\,dW_t^i, \qquad i = 1, \ldots, k$$

on the simplex $\Sigma_k := \{(x_1, \ldots, x_k) : 0 \le x_1 \le \ldots \le x_k \le 1\} \subset \mathbb{R}^k$ with some k-dimensional Brownian motion $(W_t)_{t \ge 0}$ and with the weight

$$\rho_k^\beta(x_1, \ldots, x_k) =$$
$$= \int_{x_{k-1}}^{x_k} \ldots \int_{x_1}^{x_2} \prod_{i=1}^{k} \left[\int_0^1 \left(\frac{x_i - y_{i-1}}{y_i - y_{i-1}} - z_i \right)_+^{\beta/k-1} \cdot z_i^{-z_i\beta/k} \cdot (1 - z_i)^{-(1-z_i)\beta/k} \cdot \right.$$
$$\left. \cdot (y_i - y_{i-1})^{\beta/k-2} \cdot \left(\cos(\pi z_i \beta/k) - \frac{1}{\pi} \sin(\pi z_i \beta/k) \cdot \log \frac{z_i}{1 - z_i} \right) dz_i \right] dy_1 \ldots dy_{k-1}$$

(where $y_0 := 0$, $y_k := 1$). Then for $k = 2^n \to \infty$ the empirical distributions

$$\mu_t^k(\omega) = \frac{1}{k} \sum_{i=1}^{k} \delta_{X_{kt}^i(\omega)}$$

converge to the Wasserstein diffusion $(\mu_t)_{t \ge 0}$ on $\mathcal{P}([0, 1])$.

The Particle System

The density ρ_k^β is continuous, positive and bounded from above by

$$\tilde{\rho}_k^\beta(x) = C \cdot [x_1(1-x_k)]^{\beta/(2k)-1} \cdot \prod_{i=2}^{k} \left(x_i - x_{i-1}\right)^{\beta/k-1}$$

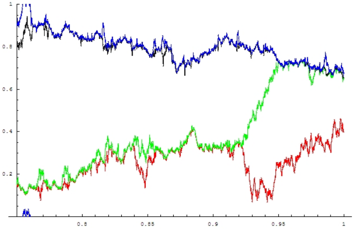
in the interior of the simplex Σ_k and it vanishes on $\mathbb{R}^k \setminus \Sigma_k$.
Replacing ρ by $\tilde{\rho}$ leads to the particle system studied by S.Andres and M.-K. von Renesse

$$dX_t^i = \left(\frac{\beta}{k} - 1\right) \left[\frac{1}{X_t^{i-1} - X_t^i} - \frac{1}{X_t^i - X_t^{i+1}}\right] dt + \sqrt{2}\, dW_t^i$$

$i = 1, \ldots, k$. In this case, however, until now the convergence

$$\frac{1}{k} \sum_{i=1}^{k} \delta_{X_{kt}^i} \longrightarrow \mu_t$$

could not be proven rigorously.



Particle systems for N=4 and β = 1

Key argument in our approach: monotone approximation of the energy

Recall Dirichlet form on the Wasserstein space $\mathcal{P}_2$

$$\mathcal{E}(u,u) \quad = \quad \int_{\mathcal{P}} \|Du(\mu)\|^2_{L^2(\mu)} \, d\mathbb{P}^\beta(\mu)$$

for $u \in \mathcal{D}(\mathcal{E}) \subset L^2(\mathcal{P}, \mathbb{P}^\beta)$.

Isomorphism $\chi : \mathcal{G} \longrightarrow \mathcal{P}_2, g \mapsto g_* \mathrm{Leb}|_{[0,1]}$ induces Dirichlet form on $\mathcal{G}$, regarded as a convex subset of the Hilbert space $L^2([0,1], \mathrm{Leb})$:

$$\mathbf{E}(u,v) = \int_{\mathcal{G}} \langle \mathbf{D}u(g), \mathbf{D}v(g) \rangle \, d\mathbb{Q}^\beta(g)$$

where $\mathbf{D}u$ denotes the Frechet derivative for "smooth" functions $u : \mathcal{G} \longrightarrow \mathbb{R}$.

Let $\mathcal{C}^1(\mathcal{G})$ denote the set of all ('cylinder') functions $u : \mathcal{G} \longrightarrow \mathbb{R}$ which can be written as $u(g) = U(\langle g, \psi_1 \rangle, \ldots, \langle g, \psi_n \rangle)$ with $n \in \mathbb{N}$, $U \in \mathcal{C}^1(\mathbb{R}^n)$ and $\psi_1, \ldots, \psi_n \in L^2([0,1], \mathrm{Leb})$. For u of this form the gradient

$$\mathbf{D}u(g) = \sum_{i=1}^n \partial_i U(\langle g, \psi_1 \rangle, \ldots, \langle g, \psi_n \rangle) \cdot \psi_i(.)$$

exists in $L^2([0,1], \mathrm{Leb})$ and

$$\|\mathbf{D}u(g)\|^2 = \int_0^1 \left| \sum_{i=1}^n \partial_i U(\langle g, \psi_1 \rangle, \ldots, \langle g, \psi_n \rangle) \cdot \psi_i(s) \right|^2 ds.$$

$\mathcal{D}(\mathbf{E})$ is the closure of $\mathcal{C}^1(\mathcal{G})$.

Key argument in our approach: monotone approximation of the energy

For $k \in \mathbb{N}$ put $\varphi_k^{(i)}(t) := k \cdot 1_{(\frac{i-1}{k}, \frac{i}{k}]}(t)$ and let $\mathcal{C}_k^1(\mathcal{G})$ denote the set of all functions $u : \mathcal{G} \longrightarrow \mathbb{R}$ which can be written as $u(g) = U\left(\langle g, \varphi_k^{(1)}\rangle, \ldots, \langle g, \varphi_k^{(k)}\rangle\right)$ with $U \in \mathcal{C}^1(\mathbb{R}^k)$.

If we define $\mathbf{E}_k$ on $L^2(\mathcal{G}, \mathbb{Q}^\beta)$ as before, – but now with domain being the closure of $\mathcal{C}_k^1(\mathcal{G})$ then

$$\mathbf{E}_{2^k} \searrow \mathbf{E}.$$

For $u, v \in \mathcal{C}_k^1(\mathcal{G})$

$$\mathbf{E}_k(u, v) = k \int_{\mathbb{R}^k} \nabla U(x) \cdot \nabla V(x)\, dm_k^\beta(x)$$

with ∇U denoting the usual gradient of U on $\mathbb{R}^k$ and with $m_k^\beta = (J_k)_* \mathbb{Q}^\beta$ where

$$J_k : \quad \mathcal{G} \mapsto \Sigma_k, \quad g \mapsto \left(\langle g, \varphi_k^{(1)}\rangle, \ldots, \langle g, \varphi_k^{(k)}\rangle\right).$$

That is, the Dirichlet form $\mathbf{E}_k$ on $\mathcal{G}$ is isomorphic to the Dirichlet form

$$\mathcal{E}_k(U, V) = k \int_{\mathbb{R}^k} \nabla U(x) \cdot \nabla V(x)\, dm_k^\beta(x)$$

on $\Sigma_k \subset \mathbb{R}^k$.

It remains to identify the measure m_k^β.

Identification of the Measure m_k^β

Put

$$\Phi_k^{(i)}(t) := \begin{cases} 1, & \text{for } t \in [0, \frac{i-1}{k}] \\ i - kt, & \text{for } t \in [\frac{i-1}{k}, \frac{i}{k}] \\ 0, & \text{for } t \in [\frac{i}{k}, 1]. \end{cases}$$

Then integration by parts yields

$$\int_0^1 \Phi_k^{(i)}(t) dg(t) = k \int_{\frac{i-1}{k}}^{\frac{i}{k}} g(t)dt = \langle g, \varphi_k^{(i)} \rangle$$

for all $i = 1, \ldots, k$ and all $g \in \mathcal{G}$.

Recall that $m_k^\beta := (J_k)_* \, \mathbb{Q}^\beta$ with $J_k(g) = \left(\int_0^1 \Phi_k^{(1)} dg, \ldots, \int_0^1 \Phi_k^{(k)} dg \right)$.

For $k = 1$ the measure m_1^β coincides with the distribution of the "random means" $\int_0^1 t \, dg(t)$ of the Dirichlet-Ferguson process. It is absolutely continuous with density

$$\vartheta_\beta(x) = \beta e^\beta \int_0^x (x-y)^{\beta-1} \cdot y^{-\beta y} \cdot (1-y)^{-\beta(1-y)} \cdot \left[\cos(\pi\beta y) - \frac{1}{\pi} \sin(\pi\beta y) \cdot \log \frac{y}{1-y} \right] dy$$

on $[0, 1]$.

Identification of the Measure m_k^β

In the case $k > 1$ note that

$$\int_0^1 \Phi_k^{(i)} dg = g\left(\frac{i-1}{k}\right) + \left[g\left(\frac{i}{k}\right) - g\left(\frac{i-1}{k}\right)\right] \cdot \int_0^1 (1-t)d\tilde{g}_i(t)$$

with

$$\tilde{g}_i(t) := \frac{g\left(\frac{t+i-1}{k}\right) - g\left(\frac{i-1}{k}\right)}{g\left(\frac{i}{k}\right) - g\left(\frac{i-1}{k}\right)}.$$

Now the crucial fact is that, conditioned on $\left(g\left(\frac{1}{k}\right), \ldots, g\left(\frac{k-1}{k}\right)\right)$, the processes $(\tilde{g}_i(t))_{t\in[0,1]}$ for $i = 1, \ldots, k$ are independent and distributed according to $\mathbb{Q}^{\beta/k}$.

Hence,

$$\rho_k^\beta(x_1, \ldots, x_k) = \frac{\Gamma(\beta)}{\Gamma(\beta/k)^k} \int_{x_{k-1}}^{x_k} \ldots \int_{x_1}^{x_2} \prod_{i=1}^{k} \left[\vartheta_{\beta/k}\left(\frac{x_i - y_{i-1}}{y_i - y_{i-1}}\right) \cdot (y_i - y_{i-1})^{\beta/k-2}\right] dy_1 \ldots dy_{k-1}$$

(where $y_0 := 0$, $y_k := 1$) with ϑ_β as above.