# Mouvement de foules et dynamique de populations sous contraintes de densité

#### Filippo Santambrogio

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud http://www.math.u-psud.fr/~filippo/

Congrès SMAI 2017, Ronce-les-Bains

A B M A B M

- Micro and macro models for crowd motions with constraints
  - Disks with no overlapping
  - Density  $ho \leq 1$
  - Continuity equation, nonlinear diffusion, Hele-Shaw
- **②** Existence and approximation : the role of optimal transport
  - A splitting method
  - Few words about optimal transport and Wasserstein distance
  - Diffusive variants
  - The projection operator
  - Gradient flows and the JKO scheme
- Oniqueness
  - Contractivity in W<sub>2</sub>
  - Contractivity in L<sup>1</sup>
  - Very rough vector fields
- Output State Numerical methods
  - Augmented Lagrangian for the JKO scheme
  - Optimization among convex functions
  - Stochastic approximation of the projection

3

# Modeling

### Micro and macro models with constraints

Filippo Santambrogio Foules et contraintes de densité

æ

A B + A B +

A >

# Non-overlapping disks

A particle population moves, and each particle - if alone - would follow its own velocity u (depending on time, position... a typical case is :  $u = -\nabla V$ , where  $V(x) = \text{dist}(x, \Gamma)$ ,  $\Gamma$  being an exit). Yet, particles are (modeled by) rigid disks that cannot overlap, hence, the

Yet, particles are (modeled by) rigid disks that cannot overlap, hence, the actual velocity v will not be u if u is too concentrating.

If q is the particle configuration, we define adm(q) the set of velocities that do not induce overlapping : if every particle is a disk with radius R, located at  $q_i$ , we have

$$q \in \mathcal{K} := \{q = (q_i)_i \in \Omega^N : |q_i - q_j| \ge 2R\}$$
  
$$adm(q) = \{v = (v_i)_i : (v_i - v_j) \cdot (q_i - q_j) \ge 0 \ \forall (i,j) : |q_i - q_j| = 2R\}$$



# A general model

How to handle  $v \approx u$  and  $q \in K$  at the same time? We will assume  $v = P_{adm(q)}[u]$  and solve  $q'(t) = P_{adm(q(t))}[u_t]$  (with q(0) given). This can be discretized (*catching-up algorithm*) as follows

$$\tilde{q}_{n+1}^{\tau} = q_n^{\tau} + \tau u_{n\tau}, \quad q_{n+1}^{\tau} = P_{\mathcal{K}}[\tilde{q}_{n+1}^{\tau}]$$

(for a small time step au > 0) and is the same as the differential inclusion

$$q'(t) \in u_t - N_K(q(t))$$

where  $N_K$  is the normal cone to the set K

$$N_{\mathcal{K}}(q_0)=\{ v \ : \ q_1\in \mathcal{K} \Rightarrow v\cdot (q_1-q_0)\leq o(|q_1-q_0|) \}.$$

It is important here that K, even if not convex in  $\Omega^N$ , is as at least *prox-regular* (the projection on K is well defined on a neighborhood of K).

B. MAURY, J. VENEL, Handling of contacts in crowd motion simulations, *Traffic and Granular Flow*, 2007.

# From the RER to the escalator

Filippo Santambrogio Foules et contraintes de densité

2

個 と く ヨ と く ヨ と

### Continuous formulation

- The particles population will be described by a density ρ ∈ P(Ω),
- the constraint by  $\mathcal{K} = \{ \rho \in \mathcal{P}(\Omega) \, : \, \rho \leq 1 \}$ ,
- $u: \Omega \to \mathbb{R}^d$  will be a vector field, possibly depending on time or  $\rho$ ,
- $adm(\rho) = \{ \mathbf{v} : \Omega \to \mathbb{R}^d : \nabla \cdot \mathbf{v} \ge 0 \text{ on } \{ \rho = 1 \} \},$
- P is the projection in  $L^2(dx)$  or (which is the same) in  $L^2(\rho)$ ,
- we'll solve  $\partial_t \rho_t + \nabla \cdot \left( \rho_t \left( P_{adm(\rho_t)} u_t \right) \right) = 0.$

**Difficulty :**  $v = P_{adm(\rho_t)}u_t$  is not regular, neither depends regularly on  $\rho$ .

(国) (日) (日) (日)

#### Pressures and duality

The set  $adm(\rho)$  may be better described by duality :

$$adm(
ho) = \left\{ v \in L^2(
ho) \ : \ \int v \cdot \nabla p \leq 0 \quad \forall p \ : \ p \geq 0, \ p(1-
ho) = 0 
ight\}.$$

We can characterize  $v = P_{adm(\rho)}(u)$  through

$$egin{aligned} &u=v+
abla p, \quad v\in \mathit{adm}(
ho), \quad \int v\cdot 
abla p=0, \ &p\in \mathrm{press}(
ho):=\{p\in H^1(\Omega),\ p\geq 0,\ p(1-
ho)=0\} \end{aligned}$$

This function p plays the role of a pressure affecting the movement.



#### Pressures and duality

The set  $adm(\rho)$  may be better described by duality :

$$adm(\rho) = \left\{ v \in L^2(\rho) : \int v \cdot \nabla p \leq 0 \quad \forall p : p \geq 0, \ p(1-\rho) = 0 \right\}.$$

We can characterize  $v = P_{adm(\rho)}(u)$  through

$$egin{aligned} &u=v+
abla p, \quad v\in \mathit{adm}(
ho), \quad \int v\cdot 
abla p=0, \ &p\in \mathrm{press}(
ho):=\{p\in H^1(\Omega),\ p\geq 0,\ p(1-
ho)=0\} \end{aligned}$$

This function p plays the role of a pressure affecting the movement.



#### ... very similar to the Hele-Shaw flow

The Hele-Shaw flow in few words :

$$\partial_t \rho_t - \Delta p_t = G \ (= \text{ reaction terms})$$

where  $p_t \in H(\rho_t)$  and H is a monotone graph. When  $H(s) = s^m$  we have a **porous-medium equation** (non-linear diffusion), when  $H = \partial I_{[0,1]}$  (hence  $H(s) = [0, +\infty[$  for  $s = 1, H(s) = \{0\}$  for 0 < s < 1), the pressure  $p \ge 0$  is arbitrary, but satisfies  $p(1 - \rho) = 0$ . (density dynamics) When  $G \ge 0$  and  $\rho_0 = \mathbb{1}_{\Omega_0}$  is a patch, the evolution is  $\rho_t = \mathbb{1}_{\Omega_t}$  with  $\Omega_t$  evolving with normal velocity (free boundary geometric evolution)

$$v_t = -\partial p_t / \partial n, \quad -\Delta p_t = G \text{ in } \Omega_t, \ p_t = 0 \text{ on } \partial \Omega_t.$$

Our equation has the same form, but with an advection term instead of reactions.

M.G. CRANDALL, An introduction to evolution governed by accretive operators, 1976 PH. BÉNILAN, L. BOCCARDO, M. HERRERO, On the limit of solution of  $u_t = \Delta u^m$  as  $m \to \infty$ , 1989

B. PERTHAME, F. QUIRÓS, J.-L. VÁZQUEZ, The Hele-Shaw Asymptotics for Mechanical Models of Tumor Growth, *ARMA*, 2014.  $(\Box \Rightarrow \langle \Box \Rightarrow \langle \Xi \Rightarrow \langle \Xi \Rightarrow \langle \Xi \Rightarrow \rangle \equiv$ 

## An example

Closed door (no-flux boundary conditions on  $\partial \Omega$ )



Open door (free flux on the bottom, i.e. mixed Dirichlet-Neumann for p)



# Existence and approximation

# The role of optimal transport

(\* ) \* ) \* ) \* )

\_\_\_\_

3

# A splitting (catching-up) scheme for the PDE

Fix a time step  $\tau > 0$ . We look for a sequence  $(\rho_n^{\tau})_n$  where  $\rho_n^{\tau}$  stands for  $\rho$  at time  $n\tau$ . We first define

$$\tilde{\rho}_{n+1}^{\tau} = (id + \tau u_{n\tau})_{\#} \rho_n^{\tau}; \quad \rho_{n+1}^{\tau} = P_{\mathcal{K}}(\tilde{\rho}_{n+1}^{\tau})$$

where the projection  $P_{K}$  is in the sense of the Wasserstein distance, induced by optimal transport.

The key point is actually using the  $W_2$  projection (instead of  $L^2$  or other projections). It corresponds to the  $L^2$  projection of velocity fields and of (Lagrangian) positions.

B. MAURY, A. ROUDNEFF-CHUPIN AND F. SANTAMBROGIO, A macroscopic crowd motion model of gradient flow type, *Math. Mod. Meth. Appl. Sci.*, 2010

B. MAURY, A. ROUDNEFF-CHUPIN, F. SANTAMBROGIO AND J. VENEL, Handling congestion in crowd motion modeling *Net. Het. Media*, 2011

- ( 同 ) ( 目 ) ( 目 ) ( 目 )

#### Optimal transport and Wasserstein distances

If two probabilities  $\mu,\nu\in\mathcal{P}(\Omega)$  are given on a compact domain, the Monge-Kantorovitch problem reads

$$\begin{split} W_2^2(\mu,\nu) &= \inf \left\{ \int |x - T(x)|^2 d\mu \ : \ T : \Omega \to \Omega, \ T_{\#}\mu = \nu \right\} \\ &= \inf \left\{ \int |x - y|^2 d\gamma \ : \ \gamma \in \mathcal{P}(\Omega^2), \ (\pi_x)_{\#}\gamma = \mu, \ (\pi_y)_{\#}\gamma = \nu \right\} \\ &= 2 \sup \left\{ \int \phi \ d\mu + \int \psi \ d\nu \ : \ \phi(x) + \psi(y) \le \frac{1}{2} |x - y|^2 \right\}. \end{split}$$

Under suitable assumptions, there exist an optimal transport T and an optimal  $\phi$ , called **Kantorovich potential**, which is Lipschitz continuous. They are linked by  $T(x) = x - \nabla \phi(x)$  (also,  $T = \nabla u$  with  $u(x) = |x|^2/2 - \phi(x)$  convex). Moreover,  $W_2(\mu, \nu)$  is a distance on  $\mathcal{P}(\Omega)$  which metrizes the weak-\* convergence of probabilities (on compact domains). G. MONGE, Mémoire sur la théorie des déblais et des remblais, 1781 L. KANTOROVICH, On the transfer of masses, *Dokl. Acad. Nauk. USSR*, 1942. Y. BRENIER, Décomposition polaire et réarrangement monotone des champs de vecteurs, *CRAS*, 1987.

### Projections and pressures

Fix a measure  $\nu \in \mathcal{P}(\Omega)$  and solve

$$\min\left\{\frac{1}{2}W_2^2(\rho,\nu) \ : \ \rho \in \mathcal{K}\right\}$$



$$= \min_{\rho \leq 1} \sup_{\phi,\psi} \int \phi \, d\rho + \int \psi \, d\nu.$$

By duality and inf-sup exchange, the optimal  $\rho$  must also solve

$$\min\int \phi \, d\rho \ : \ \rho \leq 1,$$

where  $\phi$  is the Kantorovich potential in the transport from  $\rho$  to  $\nu.$  This implies

 $\exists \ell \ : \ \rho = \begin{cases} 1 & \text{on } \phi < \ell, \\ 0 & \text{on } \phi > \ell, \Rightarrow p := (\ell - \phi)_+ \ge 0, \ p(1 - \rho) = 0. \\ \in [0, 1] & \text{on } \phi = \ell \end{cases}$ Hence,  $p \in \text{press}(\rho)$  and, passing to gradients, we have

=

# Getting back to the PDE

 $T(x) = x + \nabla p(x)$  is the optimal transport from  $\rho_{n+1}^{\tau}$  to  $\tilde{\rho}_{n+1}^{\tau}$ . Notice

$$|\nabla p||_{L^{2}(\rho_{n+1}^{\tau})} = W_{2}(\rho_{n+1}^{\tau}, \tilde{\rho}_{n+1}^{\tau}) \leq W_{2}(\rho_{n}^{\tau}, \tilde{\rho}_{n+1}^{\tau}) \leq \tau ||u_{n\tau}||_{L^{2}(\rho_{n}^{\tau})}.$$

This suggest to scale the pressure (we call it now  $\tau p$ ) and get the following situation



Formally, we have  $(id + \tau u_{n\tau})^{-1}(id + \tau \nabla p) = id - \tau (u_{(n+1)\tau} - \nabla p) + o(\tau)$ provided u is regular enough. This allows to get, in the limit  $\tau \to 0$ , the vector field  $v_t = P_{adm(\rho_t)}[u_t]$  and get a solution of the PDE.

# Diffusive variants, different splitting methods

Taking  $\tilde{\rho}_{n+1}^{\tau} = (id + \tau u_{n\tau})_{\#} \rho_n^{\tau}$  is just a possible choice. When u is not regular enough or depends on  $\rho$  there are better options. Consider

$$\partial_t \rho_t + \nabla \cdot (\rho_t u_t) - \sigma \Delta \rho_t - \Delta \rho_t = 0$$

where  $\sigma \geq 0$  is a volatility. Take the solution of the Fokker-Planck equation

$$\begin{cases} \partial_{s}\rho_{s} + \nabla \cdot (u_{s}\rho_{s}) - \sigma \Delta \rho_{s} = 0, \\ \rho_{0} = \rho_{n}^{\tau}; \end{cases}$$

then, define  $\tilde{\rho}_{n+1}^{\tau} = \rho_{\tau}$ .

The method works and converges under the same assumptions for the FP equation to be well-posed :  $u \in L^{\infty}$  if  $\sigma > 0$ , u satisfying DiPerna-Lions (or Ambrosio) theory for  $\sigma = 0$  ( $u \in W^{1,1}$  or  $u \in BV$  + bounds on  $\nabla \cdot u$ ).

R. J. DIPERNA, P. L. LIONS, Ordinary differential equations, transport theory and Sobolev spaces, *Inv. Math.*, 1989 L. AMBROSIO Transport equation and Cauchy problem for BV vector fields, *Inv. Math.* 2003 A. R. MÉSZÁROS, F. SANTAMBROGIO Advection-diffusion equations with density con-

straints, An. PDEs, 2016.

### The projection operator

A key tool is the projection operator  $P_{K(f)}$  (even if we mainly use it for f = 1):  $P_{K(f)}[\nu] := \operatorname{argmin}\{W_2^2(\rho,\nu), \rho \leq f\}$ . Its properties are essential for proving convergence. What we know :

- W<sub>2</sub><sup>2</sup>(·, ν) is strictly convex as soon as ν ≪ L<sup>d</sup>. This provides uniqueness and hence continuity in this case.
- Uniqueness actually holds for every  $\nu$ , in the case  $f \ll \mathcal{L}^d$ .
- For f = 1, the geodesic convexity of  $\{\rho : \rho \leq 1\}$  (w.r.t. Wasserstein geodesics) also gives uniqueness, and Hölder continuity w.r.t.  $W_2$ .
- (1-)Lipschitz continuity of  $P_{K(f)}$  is an open question !
- The projection preserves ordering and decreases the *L*<sup>1</sup> distance between densities.
- Estimates (order 0) : for every convex U, ρ → ∫ U(ρ(x))dx decreases under projection.
- Estimates (order 1) : the BV norm decreases under projection.

A. ROUDNEFF-CHUPIN, Modélisation macroscopique de mouvements de foule, PhD thesis, Orsay, 2011 G. DE PHILIPPIS, A. R. MÉSZÁROS, F. SANTAMBROGIO, B. VELICHKOV BV estimates in optimal transportation and applications, *ARMA*, 2016. < D > ( D >

# Gradient flows

When *u* has a suitable gradient structure, it is possible to do the two steps of the splitting algorithm at once, thanks to the theory of gradient flows. **Gradient flows in few words :** consider an evolution equation of the kind

$$x'(t) = -\nabla F(x(t))$$

(we follow the steepest descent lines of a function  $F : \mathbb{R}^n \to \mathbb{R}$ ). We can discretize in time such an equation by solving

$$x_{k+1}^{ au} \in \operatorname{argmin}_{x} F(x) + rac{1}{2 au} |x - x_{k}^{ au}|^{2}, \quad au > 0 ext{ fixed}.$$

The optimal  $x_{k+1}^{\tau}$  satisfies

$$\frac{x_{k+1}^{\tau}-x_{k}^{\tau}}{\tau}+\nabla F(x_{k+1}^{\tau})=0$$

which corresponds to an implicit Euler scheme for  $x' = -\nabla F(x)$ , the solution being found as a limit  $\tau \to 0$ .

This formulation may easily be adapted to a general metric space...

E. DE GIORGI, New problems on minimizing movements, *Boundary Value Problems for PDE and Applications*, 1993

### Gradient flows in $W_2$

Let *F* be a functional over  $(\mathcal{P}(\Omega), W_2)$ , and let us follow the so-called JKO scheme

$$ho_{k+1}^ au \in \operatorname{argmin}_
ho {\mathcal F}(
ho) + rac{W_2^2(
ho, 
ho_k^ au)}{2 au}$$

Discrete optimality conditions :

$$\frac{\delta F}{\delta \rho}(\rho_{k+1}^{\tau}) + \frac{\phi}{\tau} = const$$

which implies

$$v(x) := rac{x - T(x)}{\tau} = rac{
abla \phi(x)}{\tau} = -
abla \Big( rac{\delta F}{\delta 
ho}(
ho) \Big)$$

and, since v represents the discrete velocity (displacement / time step), at the limit  $\tau \to 0$  the continuity equation  $\partial_t \rho + \nabla \cdot (\rho v) = 0$  gives

$$\partial_t \rho - \nabla \cdot \left( \rho \, \nabla \left( \frac{\delta F}{\delta \rho}(\rho) \right) \right) = 0.$$

R. JORDAN, D. KINDERLEHRER, F. OTTO, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.*, 1998.

同下 イヨト イヨト

3

Modeling Existence and approximation Uniqueness Numerics

#### Examples

Take 
$$F(\rho) = \int U(\rho(x))dx$$
. Then  $\frac{\delta F}{\delta \rho}(\rho) = U'(\rho)$ . The equation becomes  
 $\partial_t \rho - \nabla \cdot (\rho \nabla U'(\rho)) = 0.$ 

For instance, for  $U(t) = t \log t$  we get  $\nabla f'(\rho) = \frac{\nabla \rho}{\rho}$ , which gives the heat equation  $\partial_t \rho - \Delta \rho = 0$ .

For  $F(\rho) = \int V(x)d\rho$  we get  $\frac{\delta F}{\delta \rho}(\rho) = V$ . We can obtain the Fokker-Planck equation in the case  $F(\rho) = \int V(x)d\rho + \int \rho \log \rho \dots$ 

The equation  $\partial_t \rho - \nabla \cdot (\rho \nabla V) - \Delta p = 0$  (with  $p(1 - \rho) = 0$ ) is the gradient flow of the functional

$$\mathcal{F}(
ho) = egin{cases} \int \mathcal{V}(x) d
ho & ext{if } 
ho \in \mathcal{K} \ +\infty & ext{if not,} \end{cases}$$

which is the limit as  $m \to \infty$  of the functional  $\int (\frac{1}{m}\rho(x)^m + V(x)\rho(x))dx$ . For the diffusive variant, just add  $\sigma \int \rho(x) \log \rho(x)dx$ .

L. AMBROSIO, N. GIGLI, G. SAVARÉ *Gradient Flows*, Birkäuser, 2005 F. SANTAMBROGIO {Euclidean, Metric, and Wasserstein} Gradient Flows : an overview, *Bull. Math. Sci.*, 2017.

# Uniqueness

#### Much harder than existence

æ

個 と く ヨ と く ヨ と

# $\lambda$ -convexity

The general theory of gradient flows ensures uniqueness whenever the functional F is geodesically semi-convex, i.e. for some  $\lambda \in \mathbb{R}$  the function  $s \mapsto g(s) := F(\rho_s)$  satisfies  $g'' \ge \lambda W_2^2(\rho_0, \rho_1)$  whenever  $\rho_s$  is a constant-speed geodesic connecting  $\rho_0$  to  $\rho_1$  (in  $W_2$ , these curves have the form  $\rho_s = (id - s\nabla\phi)_{\#}\rho_0 = ((1-s)id + sT)_{\#}\rho_0)$ .

In this case whenever  $\rho_t^i$  are two gradient flows of F we also have

$$\partial_t W_2^2(\rho_t^1,\rho_t^2) \leq -2\lambda W_2^2(\rho_t^1,\rho_t^2).$$

The constraint  $\rho \in K$  is geodesically convex, but this only allows to deal with the case  $u = -\nabla V$  with  $D^2 V \ge \lambda I$ .

- 「同 ) - ( 三 ) - ( 三 ) - 三 三

# Techniques for contractivity – $W_2$

Take two solutions  $(\rho_t^i, \rho_t^i)$  and let  $\phi$  and  $\psi$  be the Kantorovich potentials between  $\rho_t^1$  and  $\rho_t^2$ , with  $T(x) = x - \nabla \phi(x)$ . We can compute

$$\partial_t \frac{1}{2} W_2^2(\rho_t^1, \rho_t^2) = \int \rho_t^1 \nabla \phi \cdot (u - \nabla p_t^1) + \int \rho_t^2 \nabla \psi \cdot (u - \nabla p_t^2)$$
  
=  $\int (x - T(x)) \cdot (u(x) - u(T(x))) d\rho_t^1 + \int \rho_t^1 \Delta \phi + \rho_t^2 \Delta \psi.$ 

If u satisfies  $(u(x) - u(y)) \cdot (x - y) \leq -\lambda |x - y|^2$ , the first term gives  $-\lambda W_2^2(\rho_t^1, \rho_t^2)$ ; for the second, we have

$$p_t^1 > 0 \Rightarrow 
ho_t^1 = 1, 
ho_t^2 \le 1 \Rightarrow \det(I - D^2 \phi) = \det(DT) \ge 1 \Rightarrow \Delta \phi \le 0$$

and this rest is negative.

Again, the pressure term (or the density constraint) only improves the contractivity that we would have with u only.

・ 同 ト ・ ヨ ト ・ ヨ ト … ヨ

# Techniques for contractivity $-L^1$

**Question** Can we say that we have uniqueness as soon as u is such to guarantee uniqueness for  $\partial_t \rho + \nabla \cdot (\rho u) = 0$ ?

For non-smooth u, the technique with no pressure is the DiPerna-Lions one, not based on  $W_2$ . It recalls much more an estimate for  $\int |\rho_t^1 - \rho_t^2| \dots$ How to combine the two? What can we say about  $L^1$  contraction? Following standard methods for Hele-Shaw, uniqueness comes from existence and estimates results on the adjoint equation... In the diffusive case ( $\sigma > 0$ ) : the solution of

$$\partial_t \rho - \nabla \cdot (\rho u) - \sigma \Delta \rho - \Delta p = 0$$

is unique as soon as  $u \in L^{\infty}$ , and we also have  $\partial_t \int |\rho_t^1 - \rho_t^2| \leq 0$ . Without diffusion this is open, even if the discrete JKO steps are indeed an  $L^1$  contraction (but, to pass this to the limit on arbitrary solutions, one already needs uniqueness).

S. DI MARINO, A. R. MÉSZÁROS, Uniqueness issues for evolution equations with density constraints, *Math. Models Methods Appl. Sci.*, 2016.

### Numerics

# Optimal transport methods for JKO or splitting schemes

回 と く ヨ と く ヨ と

э

#### Optimization of transport costs

In a splitting method, we first have to compute  $(id + \tau u)_{\#}\rho_n$  (or to solve a PDE without density constraints), then to compute a projection, i.e. minimize a transport cost with a constraint on the final density; in the JKO scheme, we directly minimize with a constraint and a penalization on the final density.



Yet, given  $\nu$ , how to solve

 $\min\{W_2^2(\rho,\nu),\,\rho\leq 1\}$ 

or, more generally

$$\min\{W_2^2(\rho,\nu) + F(\rho)\}$$
 ?

# Optimization methods for JKO - Augmented Lagrangian

Use the so-called Benamou-Brenier formula

$$W_2^2(\nu,\mu) = \min\left\{\int\int\rho|\nu|^2: \partial_t\rho + \nabla\cdot(\rho\nu) = 0, \rho_0 = \nu, \rho_1 = \mu\right\}.$$

Write  $E = \rho v$ , so that  $\rho |v|^2 = |E|^2 / \rho$  which is convex in  $(\rho, E)$ . The minimization can be written (by duality) as a saddle point for a Lagrangian

$$L(m,(A,\phi)) := m \cdot (A - \nabla_{t,x}\phi) - K^*(A) + G(\phi),$$

where  $m := (\rho, E)$ , A is the dual variable to m, and  $\nabla_{t,x}\phi := (\partial_t \phi, \nabla \phi)$  involves the test function for the constraint.

**Augmented Lagrangian :** use  $\tilde{L}(m, (A, \phi)) := L(m, (A, \phi)) - \frac{r}{2} ||A - \nabla_{t,x}\phi||^2$  (same saddle points as *L*, but more strictly convex). Saddle points can be approximated by alternate minimization.

When F is convex in  $\mu$ , this can be adapted to solve min{ $W_2^2(\rho, \nu) + F(\rho)$ }.

Example 1 — Example 2 — Example 3 J.-D. BENAMOU, Y. BRENIER A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 2000. J.-D. BENAMOU, G. CARLIER, M. LABORDE An augmented Lagrangian approach to Wasserstein gradient flows and applications, *ESAIM* : *Proc.*<sub>2</sub>2016

# Optimization methods for JKO - discrete convex functions

We can rewrite the problem as

$$\min_{u \text{ convex}: \nabla u \in \Omega} \quad \frac{1}{2} \int_{\Omega} |\nabla u(x) - x|^2 d\nu + F((\nabla u)_{\#}\nu).$$

This problem is convex in u essentially when F is geodesically convex. Suppose that  $\nu$  is discrete,  $\nu = \sum_{j} a_{j} \delta_{x_{j}}$ . A convex function defined on  $\{x_{j}\}_{j}$  is a function  $u: S \to \mathbb{R}$  such that for every  $x \in S$  we have

$$\partial u(x) := \{ p \in \mathbb{R}^d : u(x) + p \cdot (y - x) \le u(y) \text{ for all } y \in S \} \neq \emptyset.$$

When *F* has the form  $F(\rho) = \int U(\rho(x))dx$  we need to associate with  $(\nabla u)_{\#}\nu$  a diffuse measure : let us spread the mass  $a_j$  uniformly on  $\partial u(x_j)$ . Then, it is possible to optimize, using a Newton algorithm, the functional by means of **computational geometry tools** which compute and differentiate the volumes of the subdifferential cells. The new measure  $\rho$  can be defined as  $\rho = \sum_j a_j \delta_{y_j}$  where  $y_j$  is the barycenter of  $\partial u(x_j)$ . Example1 Example2 J.-D. BENAMOU, G. CARLIER, Q. MÉRIGOT AND É. OUDET Discretization of functionals involving the Monge-Ampère operator, *Num. Math.*, 2016.

#### A stochastic approach to the projection operator

Remember that  $\rho = P_K(\nu)$  means  $\nu = (id + \nabla p)_{\#}\rho$ , with  $p \in \operatorname{press}(\rho)$ , and  $\rho = 1$  on  $\{\rho \neq \nu\}$ . Hence, infinitesimally (i.e. when  $(\nu - 1)_+$  is small), setting  $A := \{\nu > 1\}$ , we can find  $\rho$  just by letting Aevolve into A' with normal velocity equal to  $-\partial p/\partial n$ (with  $-\Delta p = \nu - 1$  on A and p = 0 on  $\partial A$ ) and setting  $\rho = 1$  on A' and  $\rho = \nu$  elsewhere.

**A useful probabilistic fact :** The law of the first exit point through  $\partial A$  of  $(X + B_t)$ , with  $X \sim \mu$  on A and  $(B_t)$  a Brownian motion, is the measure  $(-\partial p/\partial n) \cdot \mathcal{H}^1_{|\partial A}$ , where  $-\Delta p = \mu$  on A and p = 0 on  $\partial A$ .

**Algorithm :** pick a random pixel among sorted according to  $(\nu - 1)_+$  and start a random walk from there; as soon as it meets a pixel with  $\nu < 1$ , leave there as much mass as you can and go on. Repeat till there is some excess. Use the obtained measure as  $\rho$ .

Warning : it works well in practice, but nothing is proven on the convergence of this approximation.

# Micro vs Macro

#### 5 obstacles, micro — 5 obstacles, macro





▲圖 ▶ ▲ 臣 ▶ ▲ 臣 ▶

э

#### The End

#### Thanks for your attention

・ロン ・回と ・ヨン ・ ヨン

э.

Filippo Santambrogio Foules et contraintes de densité