Computation of optimal transport with finite volumes

Journées MAGA, 03/02/2022

Gabriele Todeschi

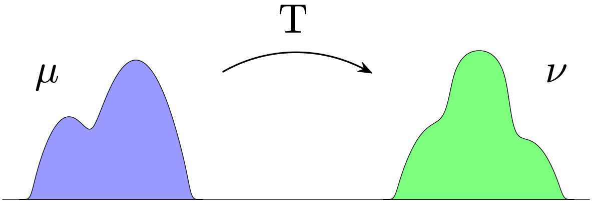
Ceremade-Inria MOKAPLAN team

in collaboration with Andrea Natale, Inria Lille

Outline of the talk

- ◇ Quadratic optimal transport problem in dynamical form
- ◇ Finite volume discretization
- ◇ Stability issues
- ◇ Convergence results
- ◇ Interior point strategy

Quadratic optimal transport problem

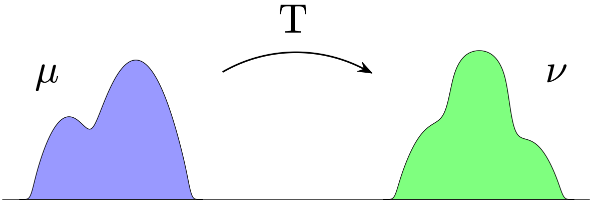


$\Omega \subset \mathbb{R}^d, \rho^{in}, \rho^f \in \mathcal{P}(\Omega)$

$\Pi(\rho^{in}, \rho^f) = \{\gamma \in \mathcal{P}(\Omega \times \Omega), (\pi_1)_\# \gamma = \rho^{in}, (\pi_2)_\# \gamma = \rho^f\}$

$$\boxed{\inf_{\gamma \in \Pi(\rho^{in}, \rho^f)} \int_{\Omega \times \Omega} \frac{1}{2} |\mathbf{x} - \mathbf{y}|^2 \, d\gamma(\mathbf{x}, \mathbf{y})}$$

Quadratic optimal transport problem



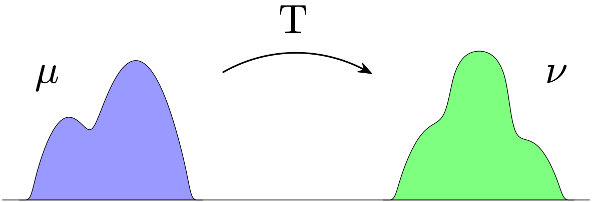
$\Omega \subset \mathbb{R}^d, \rho^{in}, \rho^f \in \mathcal{P}(\Omega)$

$\Pi(\rho^{in}, \rho^f) = \{\gamma \in \mathcal{P}(\Omega \times \Omega), (\pi_1)_\# \gamma = \rho^{in}, (\pi_2)_\# \gamma = \rho^f\}$

$$\mathcal{W}_2^2(\rho^{in}, \rho^f) := \inf_{\gamma \in \Pi(\rho^{in}, \rho^f)} \int_{\Omega \times \Omega} \frac{1}{2} |\mathbf{x} - \mathbf{y}|^2 \, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y})$$

$\mathcal{W}_2 : \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) \to \mathbb{R}_+$ is a distance

McCann's displacement interpolation



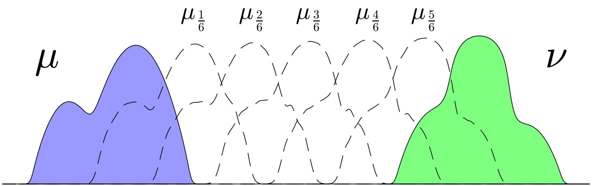
Assume ρ^{in} a.c.

$\exists$T such that

$$\mathcal{W}_2^2(\rho^{in}, \rho^f) = \int_\Omega \frac{1}{2}|\mathbf{x} - \mathrm{T}(x)|^2 \mathrm{d}\rho^{in} = \inf_{\mathrm{T}|\mathrm{T}_\#\rho^{in}=\rho^f} \int_\Omega \frac{1}{2}|\mathbf{x} - \mathrm{T}(x)|^2 \mathrm{d}\rho^{in}$$

$\gamma = (\mathrm{Id}, \mathrm{T})_\# \rho^{in}$

McCann's displacement interpolation



Assume ρ^{in} a.c.

$\exists$T such that

$$\mathcal{W}_2^2(\rho^{in}, \rho^f) = \int_\Omega \frac{1}{2}|\mathbf{x} - \mathsf{T}(x)|^2 d\rho^{in} = \inf_{\mathsf{T}|\mathsf{T}_\#\rho^{in}=\rho^f} \int_\Omega \frac{1}{2}|\mathbf{x} - \mathsf{T}(x)|^2 d\rho^{in}$$

$\gamma = (\mathrm{Id}, \mathsf{T})_\# \rho^{in}$

Interpolation: $\rho_t = (\mathsf{T}_t)_\# \rho^{in}$ where $\mathsf{T}_t = (1-t)\mathrm{Id} + t\mathsf{T}$

Benamou-Brenier dynamical formulation[1]

$$\mathcal{W}_2^2(\rho^{in}, \rho^f) := \inf_{(\rho, \boldsymbol{m}) \in \mathcal{C}} \int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t, \mathbf{x}))|^2}{2\rho(t, \mathbf{x})} \,\mathrm{d}\mathbf{x}\mathrm{d}t$$

where $\mathcal{C}$ is the convex subset of $(\rho, \boldsymbol{m})$ such that

$$\begin{cases} \partial_t \rho + \nabla \cdot \boldsymbol{m} = 0 & \text{in } [0,1] \times \Omega \\ \boldsymbol{m} \cdot \boldsymbol{n} = 0 & \text{on } [0,1] \times \partial\Omega \end{cases} \quad \text{with} \quad \begin{cases} \rho(0, \cdot) = \rho^{in} \\ \rho(1, \cdot) = \rho^f \end{cases}$$

[1] Benamou and Brenier, 2000

Benamou-Brenier dynamical formulation[1]

$$\mathcal{W}_2^2(\rho^{in}, \rho^f) := \inf_{(\rho, \boldsymbol{m}) \in \mathcal{C}} \int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t, \mathbf{x}))|^2}{2\rho(t, \mathbf{x})} \,\mathrm{d}\mathbf{x}\mathrm{d}t$$

where $\mathcal{C}$ is the convex subset of $(\rho, \boldsymbol{m})$ such that

$$\begin{cases} \partial_t \rho + \nabla \cdot \boldsymbol{m} = 0 & \text{in } [0,1] \times \Omega \\ \boldsymbol{m} \cdot \boldsymbol{n} = 0 & \text{on } [0,1] \times \partial\Omega \end{cases} \quad \text{with} \quad \begin{cases} \rho(0, \cdot) = \rho^{in} \\ \rho(1, \cdot) = \rho^f \end{cases}$$

$$\frac{|\boldsymbol{b}|^2}{2a} := \begin{cases} \frac{|\boldsymbol{b}|^2}{2a} & \text{if } a > 0 \\ 0 & \text{if } a = 0,\ \boldsymbol{b} = 0 \\ +\infty & \text{else} \end{cases}$$

Convex optimization problem with linear constraints

Non-smooth

[1] Benamou and Brenier, 2000

Benamou-Brenier dynamical formulation

Strong duality $\longrightarrow$ infsup optimization problem

Optimality conditions: continuity + HJ equation

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla \phi) = 0 \\ \partial_t \phi - \frac{1}{2} |\nabla \phi|^2 \leq 0 \end{cases} \quad \text{with} \quad \begin{cases} \rho(0, \cdot) = \rho^{in} \\ \rho(1, \cdot) = \rho^{f} \end{cases}$$

and $\boldsymbol{m} = -\rho \nabla \phi$, $\rho \nabla \phi \cdot \boldsymbol{n} = 0$ on $\partial \Omega$

HJ equation $\rightarrow$ conservation of momentum $\Longrightarrow$ zero acceleration

BB interpolation coincides with McCann's: Eulerian formulation vs Lagrangian

Bibliography

Linear programming:

Oberman, Ruan, 2015
Schmitzer, 2016

Semi-discrete optimal transport:

Merigot, 2011
Gallouët, Mérigot, 2018
Lévy, Schwindt, 2018
Mérigot, Mayron, Thibert, 2018

Entropic regularization:

Cuturi, 2013
Peyré, 2015

Monge-Ampère equation:

Benamou, Collino, Mirebeau, 2016
Bonnet, Mirebeau, 2021

Compute the transport map
$\rightarrow$ reconstruct trajectories
of particles

Eulerian schemes:

Finite elements:

Benamou, Carlier, 2015
Lavenant, Claici, Chien, Solomon, 2018

Finite difference:

Papadakis, Peyré, Oudet, 2014
Carrilo, Kraig, Wang, Wei, 2021

Finite volumes:

Erbar, Rumpf, Schmitzer, Simon, 2020
Gladbach, Kopfer, Maas, 2020

Compute directly the interpolation
$\rightarrow$ reconstruct density and
velocity fields

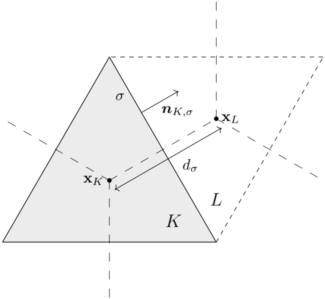
Objectives

AIM: Solve the quadratic OT problem and compute the related interpolation with the perspective of physics based applications

- **BB formulation**:
  - Continuum mechanics form
  - Easy to generalize: penalization of the density curve, non-convex domains, anisotropy, obstacles,...
- **Finite Volumes**:
  - Preserve the conservative structure
  - Handle complex domains
- **Interior Point Method**: Accuracy and efficiency

Discretization of $[0,1] \times \Omega$

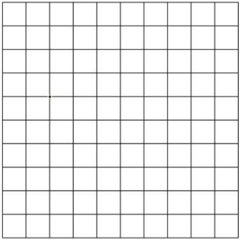
$N+1$ subintervals of length $\Delta t = \frac{1}{N+1}$

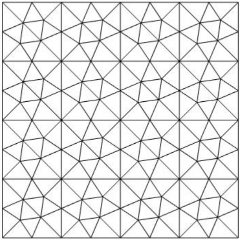


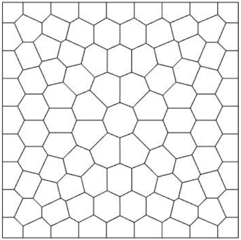
Admissible mesh for TPFA scheme:

- $\mathcal{T}$ set of control volumes K
- Σ set of edges σ
- $(\mathbf{x}_K)_{K\in\mathcal{T}}$ set of cell centers

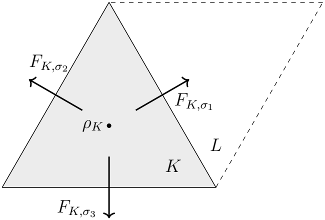
Main assumption: $\mathbf{x}_K - \mathbf{x}_L \perp \sigma$ for $\sigma = K|L \in \Sigma$







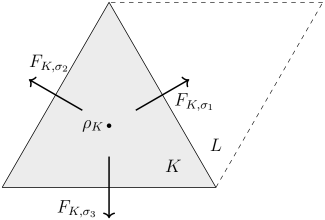
Discrete continuity equation



$m_K = |K|, m_\sigma = |\sigma|$

$$\partial_t \rho + \nabla \cdot \boldsymbol{m} = 0 \quad \longrightarrow \quad \frac{\rho_K^i - \rho_K^{i-1}}{\Delta t} m_K + \sum_{\sigma \in \Sigma_K} F_{K,\sigma}^{i-\frac{1}{2}} m_\sigma = 0, \quad \forall i, K$$

Discrete continuity equation

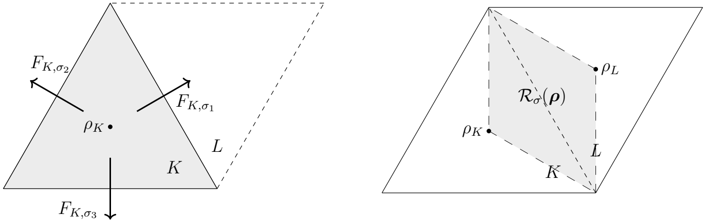


$m_K = |K|, m_\sigma = |\sigma|$

$$\partial_t \rho + \nabla \cdot \boldsymbol{m} = 0 \quad \longrightarrow \quad \frac{\rho_K^i - \rho_K^{i-1}}{\Delta t} m_K + \sum_{\sigma \in \Sigma_K} F_{K,\sigma}^{i-\frac{1}{2}} m_\sigma = 0, \quad \forall i, K$$

$$\begin{aligned} F_{K,\sigma}^{i-\frac{1}{2}} + F_{L,\sigma}^{i-\frac{1}{2}} &= 0, \qquad \text{if } \sigma \text{ internal} \\ F_{K,\sigma}^{i-\frac{1}{2}} &= 0, \qquad \text{if } \sigma \text{ external} \end{aligned} \quad \Longrightarrow \quad \sum_K \rho_K^i m_K = \sum_K \rho_K^{i-1} m_K$$

Discrete kinetic energy



$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})} \,\mathsf{d}\mathbf{x}\mathsf{d}t \approx \quad ?$$

Reconstruction in time

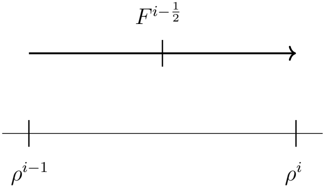
Reconstruction in space

Compensation of one directional discretization of $\boldsymbol{m}$

Time average

$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})}\,\mathrm{dxdt} \approx \sum_{i=1}^{N+1} \Delta t \int_\Omega \frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{2\rho^{i-\frac{1}{2}}}$$

$\frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{\rho^{i-\frac{1}{2}}}$ finite $\implies \boldsymbol{m}^{i-\frac{1}{2}} = \rho^{i-\frac{1}{2}}\boldsymbol{v}$



Time average

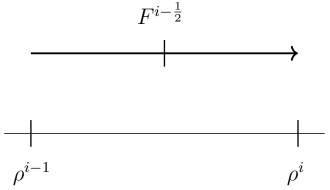
$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})}\,\mathrm{dxdt} \approx \sum_{i=1}^{N+1} \Delta t \int_\Omega \frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{2\rho^{i-\frac{1}{2}}}$$

$\frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{\rho^{i-\frac{1}{2}}}$ finite $\implies$ $\boldsymbol{m}^{i-\frac{1}{2}} = \rho^{i-\frac{1}{2}}\boldsymbol{v}$

If e.g. $\rho^{i-\frac{1}{2}} = \rho^{i-1}$:

$$\frac{\rho^i - \rho^{i-1}}{\Delta t} + \nabla \cdot \rho^{i-1}\boldsymbol{v}^{i-\frac{1}{2}} = 0, \quad \forall i$$

$$\frac{\rho^1 - \rho^{in}}{\Delta t} + \nabla \cdot \rho^{in}\boldsymbol{v}^{1-\frac{1}{2}} = 0$$



$\nexists$ a (finite) solution if $\mathrm{supp}(\rho^f) \not\subseteq \mathrm{supp}(\rho^{in})$

Time average

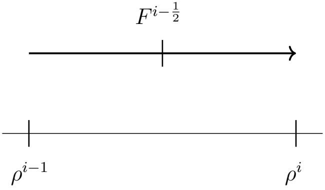
$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})}\,\mathrm{dxdt} \approx \sum_{i=1}^{N+1} \Delta t \int_\Omega \frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{2(\frac{\rho^i+\rho^{i-1}}{2})}$$

$\frac{|\boldsymbol{m}^{i-\frac{1}{2}}|^2}{\rho^{i-\frac{1}{2}}}$ finite $\implies$ $\boldsymbol{m}^{i-\frac{1}{2}} = \rho^{i-\frac{1}{2}}\boldsymbol{v}$

If e.g. $\rho^{i-\frac{1}{2}} = \rho^{i-1}$:

$$\frac{\rho^i - \rho^{i-1}}{\Delta t} + \nabla\cdot\rho^{i-1}\boldsymbol{v}^{i-\frac{1}{2}} = 0, \quad \forall i$$

$$\frac{\rho^1 - \rho^{in}}{\Delta t} + \nabla\cdot\rho^{in}\boldsymbol{v}^{1-\frac{1}{2}} = 0$$



$\nexists$ a (finite) solution if $\mathrm{supp}(\rho^f) \nsubseteq \mathrm{supp}(\rho^{in})$

Arithmetic average: $\rho^{i-\frac{1}{2}} = \frac{\rho^i+\rho^{i-1}}{2}$

Harmonic, logarithmic or geometric averages are NOT suited

$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})} \mathrm{dxdt} \approx \sum_{i=1}^{N+1} \Delta t \sum_{\sigma \in \Sigma} \frac{(F_\sigma^{i-\frac{1}{2}})^2}{2\mathcal{R}_\sigma(\frac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2})} m_\sigma d_\sigma$$

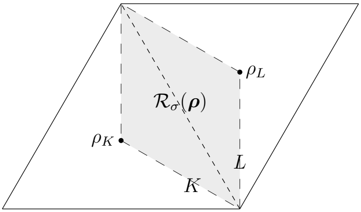
Averages of neighboring cell values
$\mathcal{R}_\sigma(\boldsymbol{\rho}) = f(\rho_K, \rho_L)$

Component-wise convex, positive

Examples: weighted arithmetic and harmonic averages

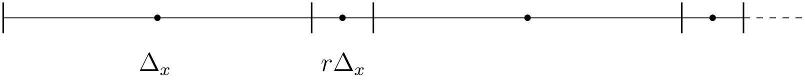
$$\mathcal{R}_\sigma(\boldsymbol{\rho}) = \lambda_{K,\sigma}\rho_K + \lambda_{L,\sigma}\rho_L$$
$$\mathcal{R}_\sigma(\boldsymbol{\rho}) = \frac{\rho_K \rho_L}{\lambda_{L,\sigma}\rho_K + \lambda_{K,\sigma}\rho_L}$$

$\forall \sigma, \lambda_{K,\sigma} + \lambda_{L,\sigma} = 1$



Counterexample[1]

$\Delta_x \in \mathbb{R}_+,\ r \in (0,1)$

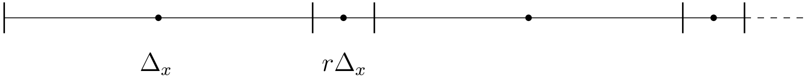


$(\lambda_{K,\sigma}, \lambda_{L,\sigma}) = (\frac{1}{2}, \frac{1}{2})$

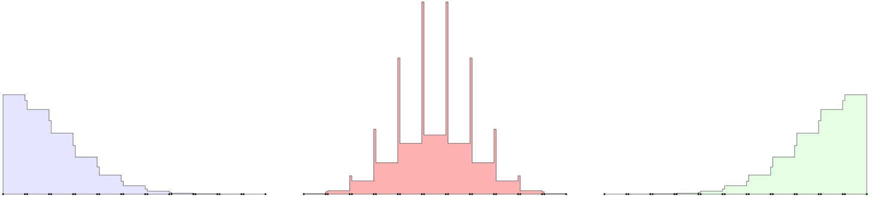
[1] Gladbach, Kopfer, Maas, *Scaling limits of discrete optimal transport*, 2020

Counterexample[1]

$\Delta_x \in \mathbb{R}_+,\ r \in (0,1)$



$(\lambda_{K,\sigma}, \lambda_{L,\sigma}) = (\frac{1}{2}, \frac{1}{2})$



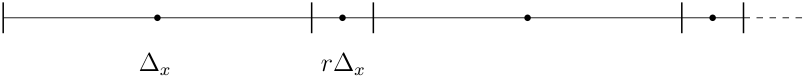
$t = 0$ $t = 0.5$ $t = 1$

The discrete solution converges to something cheaper!

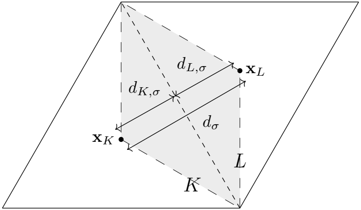
[1] Gladbach, Kopfer, Maas, *Scaling limits of discrete optimal transport*, 2020

Counterexample[1]

$\Delta_x \in \mathbb{R}_+, \ r \in (0,1)$



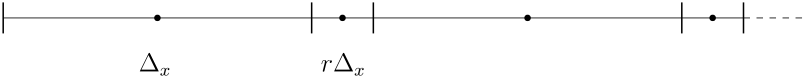
$(\lambda_{K,\sigma}, \lambda_{L,\sigma}) = (\frac{d_{K,\sigma}}{d_\sigma}, \frac{d_{L,\sigma}}{d_\sigma})$



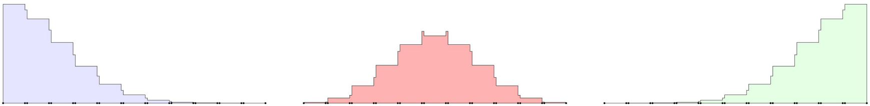
[1] Gladbach, Kopfer, Maas, *Scaling limits of discrete optimal transport*, 2020

Counterexample[1]

$\Delta_x \in \mathbb{R}_+, \ r \in (0,1)$



$(\lambda_{K,\sigma}, \lambda_{L,\sigma}) = (\frac{d_{K,\sigma}}{d_\sigma}, \frac{d_{L,\sigma}}{d_\sigma})$



$t = 0$ $\quad$ $t = 0.5$ $\quad$ $t = 1$

[1] Gladbach, Kopfer, Maas, *Scaling limits of discrete optimal transport*, 2020

Necessary condition

Asymptotic anisotropy condition

Given a (admissible) mesh and the weights $(\lambda_{K,\sigma})_{(K,\sigma)\in\mathcal{T}\times\Sigma}$, there exists $\eta, \eta \to 0$ with $h = \max(diam(K)) \to 0$, such that

$$\sum_{\sigma\in\Sigma_K} (\lambda_{K,\sigma} m_\sigma d_\sigma)\boldsymbol{n}_{K,\sigma} \otimes \boldsymbol{n}_{K,\sigma} \leq m_K(1+\eta)\mathrm{Id}, \quad \forall K \in \mathcal{T}$$

If cell centers are circumcenters:

$$(\lambda_{K,\sigma}, \lambda_{L,\sigma}) = (\frac{d_{K,\sigma}}{d_\sigma}, \frac{d_{L,\sigma}}{d_\sigma}), \quad \forall\sigma$$

$\Longrightarrow$ **asymptotic anisotropy guaranteed** with $\eta = 0$

$$\int_0^1 \int_\Omega \frac{|\boldsymbol{m}(t,\mathbf{x}))|^2}{2\rho(t,\mathbf{x})} \mathrm{d}\mathbf{x}\mathrm{d}t \approx \sum_{i=1}^{N+1} \Delta t \sum_{\sigma\in\Sigma} \frac{(F_\sigma^{i-\frac{1}{2}})^2}{2\mathcal{R}_\sigma(\frac{\rho^i+\rho^{i-1}}{2})} m_\sigma d_\sigma$$

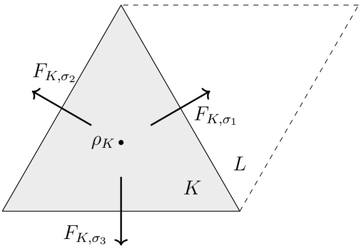
$(F_\sigma^{i-\frac{1}{2}})^2 \approx |\boldsymbol{m}^{i-\frac{1}{2}} \cdot \boldsymbol{n}_{K,\sigma}|^2$

$\boldsymbol{m}$ is approximated along only one direction

We need to compensate for the other $d-1$

We increase the measure by d times:

$$dm_{\Delta_\sigma} = m_\sigma d_\sigma$$



Discrete kinetic energy

$$\mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F}) = \begin{cases} \displaystyle\sum_{i=1}^{N+1} \Delta t \sum_{\sigma \in \Sigma} \frac{(F_\sigma^{i-\frac{1}{2}})^2}{2\mathcal{R}_\sigma(\frac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2})} m_\sigma d_\sigma & \text{if } \rho_K^i \geq 0 \\ +\infty & \text{else} \end{cases}$$

Convex and lower semi-continuous

$\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f} \in \mathbb{R}_+^{\mathcal{T}}$ with the same mass, $\sum_K \boldsymbol{\rho}^{in} m_K = \sum_K \boldsymbol{\rho}^{f} m_K$

Discrete optimal transport problem:

$$\boxed{\inf_{(\boldsymbol{\rho}, \boldsymbol{F}) \in \mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F})}$$

$\mathcal{C}_{N,\mathcal{T}}$: $(\boldsymbol{\rho}, \boldsymbol{F})$ satisfying the discrete continuity equation with $\boldsymbol{\rho}^0 = \boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{N+1} = \boldsymbol{\rho}^{f}$

Discrete optimal transport problem

$\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f} \in \mathbb{R}_+^{\mathcal{T}}$ with the same mass, $\sum_K \boldsymbol{\rho}^{in} m_K = \sum_K \boldsymbol{\rho}^{f} m_K$

Discrete optimal transport problem:

$$\inf_{(\boldsymbol{\rho}, \boldsymbol{F}) \in \mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F})$$

$\mathcal{C}_{N,\mathcal{T}}$: $(\boldsymbol{\rho}, \boldsymbol{F})$ satisfying the discrete continuity equation with $\boldsymbol{\rho}^0 = \boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{N+1} = \boldsymbol{\rho}^{f}$

Well-posed convex optimization problem

Discrete optimal transport problem

$\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f} \in \mathbb{R}_+^{\mathcal{T}}$ with the same mass, $\sum_K \boldsymbol{\rho}^{in} m_K = \sum_K \boldsymbol{\rho}^f m_K$

Discrete optimal transport problem:

$$W_{N,\mathcal{T}}^2(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^f) := \inf_{(\boldsymbol{\rho}, \boldsymbol{F}) \in \mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F})$$

$\mathcal{C}_{N,\mathcal{T}}$: $(\boldsymbol{\rho}, \boldsymbol{F})$ satisfying the discrete continuity equation with $\boldsymbol{\rho}^0 = \boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{N+1} = \boldsymbol{\rho}^f$

Well-posed convex optimization problem

$\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f} \in \mathbb{R}_+^{\mathcal{T}}$ with the same mass, $\sum_K \boldsymbol{\rho}^{in} m_K = \sum_K \boldsymbol{\rho}^{f} m_K$

Discrete optimal transport problem:

$$\boxed{W_{N,\mathcal{T}}^2(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f}) := \inf_{(\boldsymbol{\rho}, \boldsymbol{F}) \in \mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F})}$$

$\mathcal{C}_{N,\mathcal{T}}$: $(\boldsymbol{\rho}, \boldsymbol{F})$ satisfying the discrete continuity equation with $\boldsymbol{\rho}^0 = \boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{N+1} = \boldsymbol{\rho}^{f}$

Well-posed convex optimization problem

Strong duality $\implies$ saddle point in $\boldsymbol{\rho}, \boldsymbol{\phi} \in [\mathbb{R}^{\mathcal{T}}]^N \times [\mathbb{R}^{\mathcal{T}}]^{N+1}$ with

$$\boldsymbol{F}^{i-\frac{1}{2}} = -\mathcal{R}_\Sigma\left(\frac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2}\right) \odot \nabla_\Sigma \boldsymbol{\phi}^{i-\frac{1}{2}}$$

$\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f} \in \mathbb{R}_+^{\mathcal{T}}$ with the same mass, $\sum_K \boldsymbol{\rho}^{in} m_K = \sum_K \boldsymbol{\rho}^{f} m_K$

Discrete optimal transport problem:

$$\boxed{W_{N,\mathcal{T}}^2(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f}) := \inf_{(\boldsymbol{\rho}, \boldsymbol{F}) \in \mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho}, \boldsymbol{F})}$$

$\mathcal{C}_{N,\mathcal{T}}$: $(\boldsymbol{\rho}, \boldsymbol{F})$ satisfying the discrete continuity equation with $\boldsymbol{\rho}^0 = \boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{N+1} = \boldsymbol{\rho}^{f}$

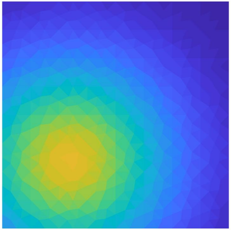
Well-posed convex optimization problem

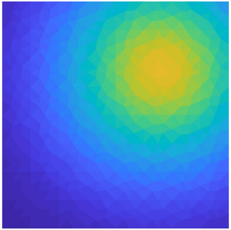
Strong duality $\implies$ saddle point in $\boldsymbol{\rho}, \boldsymbol{\phi} \in [\mathbb{R}^{\mathcal{T}}]^N \times [\mathbb{R}^{\mathcal{T}}]^{N+1}$ with

$$\boldsymbol{F}^{i-\frac{1}{2}} = -\mathcal{R}_\Sigma\left(\frac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2}\right) \odot \nabla_\Sigma \boldsymbol{\phi}^{i-\frac{1}{2}}$$

Non-smooth, $d+1$ dimensional, positivity constraint

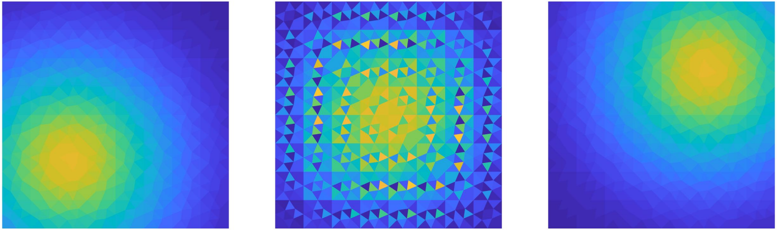
Oscillations





Oscillations

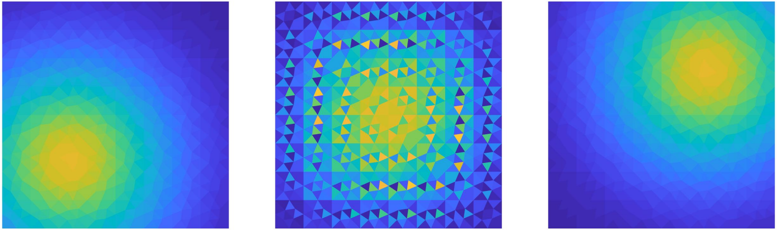
Linear average



Infsup type instabilities on the density

Oscillations

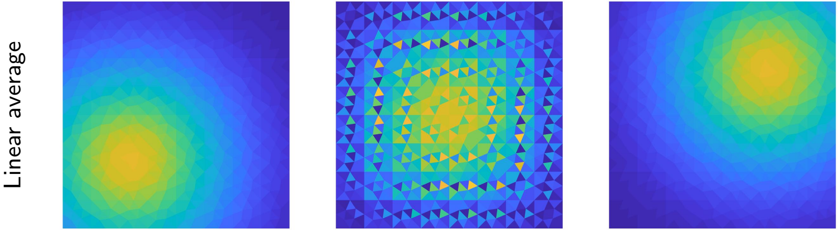
Linear average



Infsup type instabilities on the density

OT does not provide any regularity to the interpolating density

Oscillations



Infsup type instabilities on the density

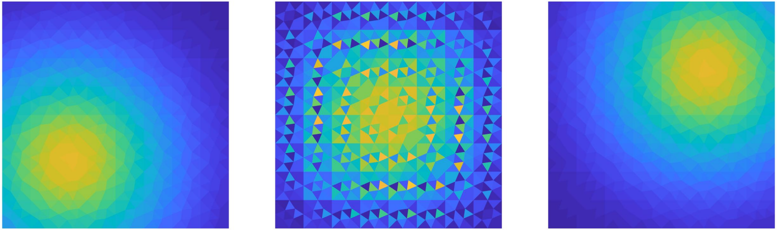
OT does not provide any regularity to the interpolating density

However, L^p norms are convex along the interpolation:

$$||\rho_t||_{L^p}^p \leq (1-t)||\rho^{in}||_{L^p}^p + t||\rho^f||_{L^p}^p$$

Oscillations

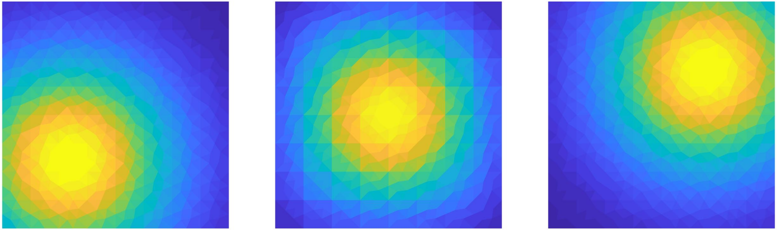
Linear average



Do not depend on the time refinement

Oscillations

Harmonic average

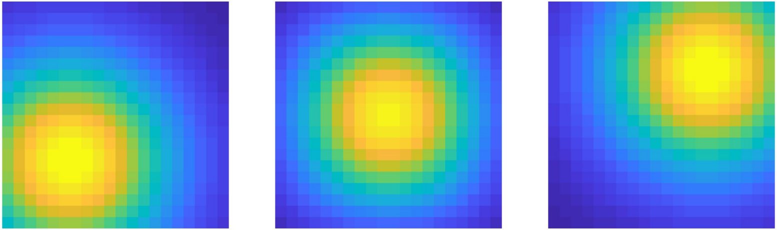


Do not depend on the time refinement

Depend on the reconstruction chosen

Oscillations

Linear average

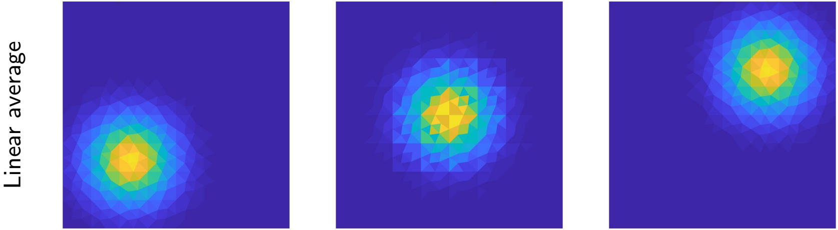


Do not depend on the time refinement

Depend on the reconstruction chosen

The grid influences the oscillations, they disappear on cartesian grids

Oscillations



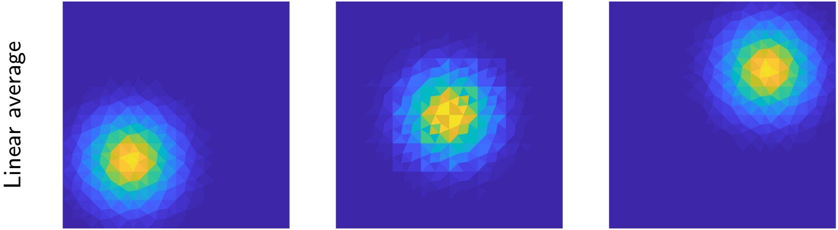
Do not depend on the time refinement

Depend on the reconstruction chosen

The grid influences the oscillations, they disappear on cartesian grids

More severe/persistent with mass compression and tend to disappear on pure translations

Oscillations



Do not depend on the time refinement

Depend on the reconstruction chosen

The grid influences the oscillations, they disappear on cartesian grids

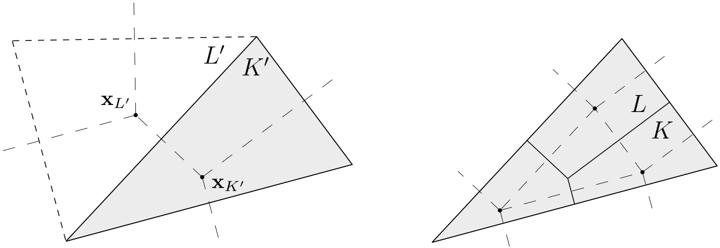
More severe/persistent with mass compression and tend to disappear on pure translations

Not limited to the FV discretization[1]

[1] *A. Natale, G. Todeschi, A mixed finite element discretization of optimal transport*, 2021

Nested discretization

We **enrich the space of discrete potentials** to overcome the problem



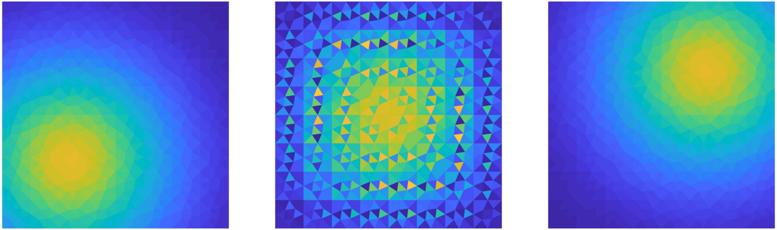
Two nested discretizations of Ω

$\mathcal{B}_{N,\mathcal{T}}$ and the continuity equation are defined on the finer grid

The density is discretized on the coarser grid and injected in the finer space

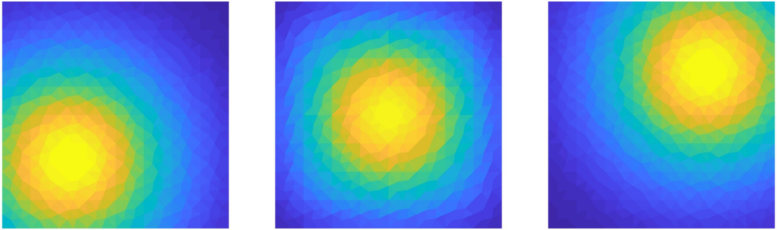
Enriched scheme

Linear average



Enriched scheme

Linear average



The oscillations are softened

Computationally the scheme is more expensive (but the perfomance of the discrete solver improves)

Convergence results

Non enriched case [Lavenant,2021]:

$$(\boldsymbol{\rho}, \boldsymbol{F}) \xrightarrow{\Delta t, h \to 0} (\rho, \boldsymbol{m}) \text{ weakly and } W^2_{N,\mathcal{T}}(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^f) \xrightarrow{\Delta t, h \to 0} \mathcal{W}^2_2(\rho^{in}, \rho^f)$$

Non enriched case [Lavenant,2021]:

$$(\boldsymbol{\rho}, \boldsymbol{F}) \xrightarrow{\Delta t, h \to 0} (\rho, \boldsymbol{m}) \text{ weakly and } W^2_{N,\mathcal{T}}(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^f) \xrightarrow{\Delta t, h \to 0} \mathcal{W}^2_2(\rho^{in}, \rho^f)$$

Theorem

Given a smooth solution (ϕ, ρ) with ρ uniformly greater than zero:

- $W^2_{N,\mathcal{T}}(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^f) \xrightarrow{\Delta t, h \to 0} \mathcal{W}^2_2(\rho^{in}, \rho^f)$ with order at least one
- $(\boldsymbol{\rho}, \boldsymbol{F}) \xrightarrow{\Delta t, h \to 0} (\rho, \boldsymbol{m})$ weakly

Convergence results

Non enriched case [Lavenant,2021]:

$$(\boldsymbol{\rho}, \boldsymbol{F}) \xrightarrow{\Delta t, h \to 0} (\rho, \boldsymbol{m}) \text{ weakly and } W^2_{N,\mathcal{T}}(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f}) \xrightarrow{\Delta t, h \to 0} \mathcal{W}^2_2(\rho^{in}, \rho^{f})$$

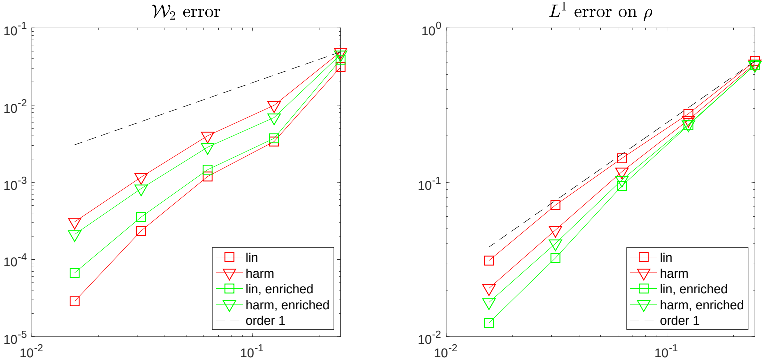
Theorem

Given a smooth solution (ϕ, ρ) with ρ uniformly greater than zero:

- $W^2_{N,\mathcal{T}}(\boldsymbol{\rho}^{in}, \boldsymbol{\rho}^{f}) \xrightarrow{\Delta t, h \to 0} \mathcal{W}^2_2(\rho^{in}, \rho^{f})$ with order at least one
- $(\boldsymbol{\rho}, \boldsymbol{F}) \xrightarrow{\Delta t, h \to 0} (\rho, \boldsymbol{m})$ weakly

- Obtained constructing competitors in the discrete problem
- Holds in both the enriched and non-enriched case

Convergence tests: translation

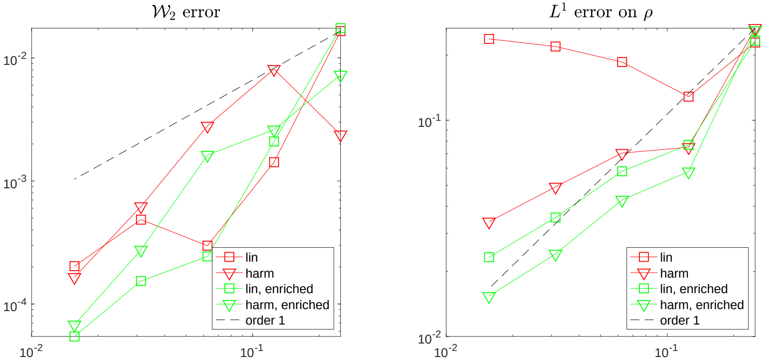


Exact solution:

$$\rho(t,x,y) = \left(1 + \cos\left(\frac{10^2\pi}{3^2}\,|\mathbf{x} - \mathbf{x}_t|^2\right)\right)\mathbb{1}_{|\mathbf{x}-\mathbf{x}_t|\le\frac{3}{10}},\ \mathbf{x}_t = \left(\frac{3}{10} + \frac{2}{5}t,\ \frac{3}{10} + \frac{2}{5}t\right)$$

$$\phi(t,x,y) = \frac{2}{5}x + \frac{2}{5}y - \frac{4}{25}t$$

Convergence tests: compression



Exact solution:

$$\rho(t,x,y) = \frac{1}{t(c-1)+1}\left(1+\cos\left(\frac{2\pi}{t(c-1)+1}\left(x-\frac{1}{2}\right)\right)\right)\mathbf{1}_{\left|x-\frac{1}{2}\right|\le\frac{t(c-1)+1}{2}}$$

$$\phi(t,x,y) = \frac{1}{2}\frac{c-1}{t(c-1)+1}\left(x-\frac{1}{2}\right)^2$$

Solution of the optimization problem

Usually solved using primal-dual/proximal splitting optimization techniques[1234]

Projection onto the parabola by rewriting the kinetic energy as

$$\frac{|\boldsymbol{m}(t,\mathbf{x})|^2}{2\rho(t,\mathbf{x})} := \sup_{a+\frac{|\boldsymbol{b}|^2}{2}\leq 0} a\rho(t,x) + \boldsymbol{b}\cdot\boldsymbol{m}(t,\mathbf{x}) = \begin{cases} \frac{|\boldsymbol{b}|^2}{2a} & \text{if } a > 0 \\ 0 & \text{if } a = 0,\ \boldsymbol{b} = 0 \\ +\infty & \text{else} \end{cases}$$

[1]Benamou, Brenier, 2000
[2]Papadakis, Peyré, Oudet, 2014
[3]Benamou, Carlier, 2015
[4]Natale, Todeschi, 2021

Solution of the optimization problem

Usually solved using primal-dual/proximal splitting optimization techniques[1234]

Projection onto the parabola by rewriting the kinetic energy as

$$\frac{|\boldsymbol{m}(t,\mathbf{x})|^2}{2\rho(t,\mathbf{x})} := \sup_{a+\frac{|\boldsymbol{b}|^2}{2}\leq 0} a\rho(t,x) + \boldsymbol{b}\cdot\boldsymbol{m}(t,\mathbf{x}) = \begin{cases} \frac{|\boldsymbol{b}|^2}{2a} & \text{if } a>0 \\ 0 & \text{if } a=0,\ \boldsymbol{b}=0 \\ +\infty & \text{else} \end{cases}$$

- NOT flexible to adapt to the discretization
- Efficient only as long as cartesian grids are used and low accuracy is required
- Suffer the lack of smoothness of the problem

[1] Benamou, Brenier, 2000
[2] Papadakis, Peyré, Oudet, 2014
[3] Benamou, Carlier, 2015
[4] Natale, Todeschi, 2021

Interior point strategy

Perturb the problem with a barrier function: $\mu \in \mathbb{R}_+$

$$\inf_{(\boldsymbol{\rho},\boldsymbol{F})\in\mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho},\boldsymbol{F}) - \mu \sum_i \Delta t \sum_{K\in\mathcal{T}} \log(\rho_K) m_K$$

The minimizer is strictly positive and the problem smooth

Interior point strategy

Perturb the problem with a barrier function: $\mu \in \mathbb{R}_+$

$$\inf_{(\boldsymbol{\rho},\boldsymbol{F})\in\mathcal{C}_{N,\mathcal{T}}} \mathcal{B}_{N,\mathcal{T}}(\boldsymbol{\rho},\boldsymbol{F}) - \mu \sum_i \Delta t \sum_{K\in\mathcal{T}} \log(\rho_K) m_K$$

The minimizer is strictly positive and the problem smooth

Theorem

The solution $(\rho^\mu, \boldsymbol{F}^\mu)$ satisfies:

- $\exists C > 0$ such that $\boldsymbol{\rho}^\mu \geq C\mu$
- $(\boldsymbol{\rho}^\mu, \boldsymbol{F}^\mu) \to (\boldsymbol{\rho}, \boldsymbol{F})$ solution of the unperturbed problem for $\mu \to 0$

The smaller μ, the more difficult is the problem

Newton method

Continuation method: solve a sequence of perturbed problems with $\mu \to 0$

Optimize from the interior of the domain

Newton method

Continuation method: solve a sequence of perturbed problems with $\mu \to 0$

Optimize from the interior of the domain

Optimality conditions: unperturbed problem

$$\begin{cases} \dfrac{\boldsymbol{\rho}^{i} - \boldsymbol{\rho}^{i-1}}{\Delta t} - \operatorname{div}_{\mathcal{T}}(\mathcal{R}_{\Sigma}(\dfrac{\boldsymbol{\rho}^{i} + \boldsymbol{\rho}^{i-1}}{2}) \odot \nabla_{\Sigma}\boldsymbol{\phi}^{k}) = 0, \\ \dfrac{\boldsymbol{\phi}^{i+1} - \boldsymbol{\phi}^{i}}{\Delta t} - \dfrac{1}{4}\mathcal{R}_{\mathcal{T}}^{i}(\nabla_{\Sigma}\boldsymbol{\phi}^{k})^2 - \dfrac{1}{4}\mathcal{R}_{\mathcal{T}}^{i+1}(\nabla_{\Sigma}\boldsymbol{\phi}^{i+1})^2 \leq 0 \end{cases}$$

Newton method

Continuation method: solve a sequence of perturbed problems with $\mu \to 0$

Optimize from the interior of the domain

Optimality conditions: unperturbed problem

$$\begin{cases} \dfrac{\boldsymbol{\rho}^i - \boldsymbol{\rho}^{i-1}}{\Delta t} - \mathrm{div}_{\mathcal{T}}(\mathcal{R}_\Sigma(\dfrac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2}) \odot \nabla_\Sigma \boldsymbol{\phi}^k) = 0, \\ \dfrac{\boldsymbol{\phi}^{i+1} - \boldsymbol{\phi}^i}{\Delta t} - \dfrac{1}{4}\mathcal{R}_{\mathcal{T}}^i(\nabla_\Sigma \boldsymbol{\phi}^k)^2 - \dfrac{1}{4}\mathcal{R}_{\mathcal{T}}^{i+1}(\nabla_\Sigma \boldsymbol{\phi}^{i+1})^2 = -\boldsymbol{s}^i, \\ \boldsymbol{\rho}^i \geq 0, \boldsymbol{s}^i \geq 0, \boldsymbol{\rho}^i \odot \boldsymbol{s}^i = 0, \end{cases}$$

Newton method

Continuation method: solve a sequence of perturbed problems with $\mu \to 0$

Optimize from the interior of the domain

Optimality conditions: perturbed problem

$$\begin{cases} \dfrac{\boldsymbol{\rho}^i - \boldsymbol{\rho}^{i-1}}{\Delta t} - \operatorname{div}_{\mathcal{T}}(\mathcal{R}_\Sigma(\dfrac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2}) \odot \nabla_\Sigma \boldsymbol{\phi}^k) = 0, \\ \dfrac{\boldsymbol{\phi}^{i+1} - \boldsymbol{\phi}^i}{\Delta t} - \dfrac{1}{4}\mathcal{R}^i_{\mathcal{T}}(\nabla_\Sigma \boldsymbol{\phi}^k)^2 - \dfrac{1}{4}\mathcal{R}^{i+1}_{\mathcal{T}}(\nabla_\Sigma \boldsymbol{\phi}^{i+1})^2 = -\boldsymbol{s}^i, \\ \boldsymbol{\rho}^i \odot \boldsymbol{s}^i = \boldsymbol{\mu}, \end{cases}$$

Continuation method: solve a sequence of perturbed problems with $\mu \to 0$

Optimize from the interior of the domain

Optimality conditions: perturbed problem

$$\begin{cases} \dfrac{\boldsymbol{\rho}^i - \boldsymbol{\rho}^{i-1}}{\Delta t} - \mathrm{div}_{\mathcal{T}}(\mathcal{R}_\Sigma(\dfrac{\boldsymbol{\rho}^i + \boldsymbol{\rho}^{i-1}}{2}) \odot \nabla_\Sigma \boldsymbol{\phi}^k) = 0, \\ \dfrac{\boldsymbol{\phi}^{i+1} - \boldsymbol{\phi}^i}{\Delta t} - \dfrac{1}{4}\mathcal{R}^i_{\mathcal{T}}(\nabla_\Sigma \boldsymbol{\phi}^k)^2 - \dfrac{1}{4}\mathcal{R}^{i+1}_{\mathcal{T}}(\nabla_\Sigma \boldsymbol{\phi}^{i+1})^2 = -\boldsymbol{s}^i, \\ \boldsymbol{\rho}^i \odot \boldsymbol{s}^i = \boldsymbol{\mu}, \end{cases}$$

Use a Newton scheme

The smoothness of the problem favors a good behavior

Algorithm: Interior point method

Given the starting point $(\phi_0, \boldsymbol{\rho}_0, \boldsymbol{s}_0)$ and the parameters $\mu_0 > 0, \theta \in (0,1), \varepsilon_0 > 0$;
while $\delta_0 > \varepsilon_0$ **do**
  $\mu = \theta\mu$;
  while $\delta_\mu > \varepsilon_\mu$ **do**
    compute Newton direction $\boldsymbol{d}$;
    compute $\alpha \in (0,1]$ such that $\boldsymbol{\rho} + \alpha \boldsymbol{d}_{\boldsymbol{\rho}} > 0$ and $\boldsymbol{s} + \alpha \boldsymbol{d}_{\boldsymbol{s}} > 0$;
    update: $(\phi, \boldsymbol{\rho}, \boldsymbol{s}) = (\phi, \boldsymbol{\rho}, \boldsymbol{s}) + \alpha(\boldsymbol{d}_{\phi}, \boldsymbol{d}_{\boldsymbol{\rho}}, \boldsymbol{d}_{\boldsymbol{s}})$;
    if $n > n_{max}$ **or** $\alpha < \alpha_{min}$ **then**
      increase μ and repeat from previous iteration;
    end
  end
end

! θ decrease ratio for μ;
! δ_0 and ε_0 error and tolerance on the real solution;
! δ_μ and ε_μ error and tolerance on the perturbed solution;

Solution of linear systems

The complexity lies in the computation of linear systems

$\boldsymbol{d^k} = -\boldsymbol{J}^k / \boldsymbol{f}^k$

$$\boldsymbol{J}^k = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$$

$A = \partial^2_{\phi\phi}\mathcal{L}, B = \partial^2_{\rho\phi}\mathcal{L}, C = \partial^2_{\rho\rho}\mathcal{L}$

Solution of linear systems

The complexity lies in the computation of linear systems

$\boldsymbol{d^k} = -\boldsymbol{J}^k / \boldsymbol{f}^k$

$$\boldsymbol{J}^k = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$$

$A = \partial^2_{\phi\phi}\mathcal{L}, B = \partial^2_{\rho\phi}\mathcal{L}, C = \partial^2_{\rho\rho}\mathcal{L}$

A becomes singular for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$

C explodes for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$

The complexity lies in the computation of linear systems

$\boldsymbol{d^k} = -\boldsymbol{J}^k / \boldsymbol{f}^k$

$$\boldsymbol{J}^k = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$$

$A = \partial^2_{\phi\phi}\mathcal{L}, B = \partial^2_{\rho\phi}\mathcal{L}, C = \partial^2_{\rho\rho}\mathcal{L}$

A becomes singular for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$
C explodes for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$ $\longrightarrow$ $\boldsymbol{J}$ becomes ill-conditioned

Preconditioned iterative methods can deal with ill-conditioning

The complexity lies in the computation of linear systems

$\boldsymbol{d^k} = -\boldsymbol{J}^k / \boldsymbol{f}^k$

$$\boldsymbol{J}^k = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$$

$A = \partial^2_{\phi\phi}\mathcal{L}, B = \partial^2_{\rho\phi}\mathcal{L}, C = \partial^2_{\rho\rho}\mathcal{L}$

A becomes singular for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$
C explodes for $\mu \to 0$ if $\boldsymbol{\rho}^\mu \to 0$ $\longrightarrow$ $\boldsymbol{J}$ becomes ill-conditioned

Preconditioned iterative methods can deal with ill-conditioning

Difficulty to find good preconditioner due to interplay of time and space discretization[1]

[1] Ongoing work with Enrico Facca, Inria Lille

Perspectives

OT can be accurately and efficiently computed using FV and IP

Perspectives:

- Better understanding of the instability issues
- Improve the solution of linear systems
- Construct more general finite volume schemes able to deal with anisotropy and less regular grids

Thank you for your attention!