# Proximal first-order algorithms for solving stationary mean field games with non-local couplings[1]

**L. M. Briceño-Arias**
Universidad Técnica Federico Santa María

[1] with F.J. Silva, J. Deride, S. López

# Second order stationary MFG system[3]

**MFG**

$$
\begin{aligned}
-\partial_t u - \nu\Delta u + \frac{|\nabla u|^2}{2} &= f(x, m(t)) \quad \text{in } \mathbb{T}^2 \times [0, T],\\
u(x,T) &= g(x, m(T))\\
\partial_t m - \nu\Delta m - \operatorname{div}(m\nabla u) &= 0\\
m(0) &= m_0
\end{aligned}
$$

The long time average of solutions $(u^T, m^T)$ of MFG as $T \to \infty$ [2] leads to...

[2] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta, *Long time average of mean field games with a nonlocal coupling*, *SIAM J. Control Optim.*, 2013

[3] Lasry & Lions (2006-7).

# Second order stationary MFG system[2]

**SMFG**

$$
\begin{aligned}
-\nu\Delta u + \frac{1}{2}|\nabla u|^2 + \lambda &= f(x,m) \quad \text{in } \mathbb{T}^2,\\
-\nu\Delta m - \operatorname{div}(m\nabla u) &= 0 \qquad\quad \text{in } \mathbb{T}^2,
\end{aligned}
$$

$$
\int_{\mathbb{T}^2} u(x)\mathrm{d}x = 0,\;\; m \geq 0,\; \int_{\mathbb{T}^2} m(x)\mathrm{d}x = 1.
$$

- $\mathbb{T}^2$: 2-dimensional torus.
- $f : \mathbb{T}^2 \times L^1(\mathbb{T}^2) \to \mathbb{R}$: coupling or interaction term ($f(x, m(x))$: local coupling).
- Existence/uniqueness for $(u, m)$: Lasry & Lions (2006, 2007), Bardi & Feleqi (2016), Meszaros & Silva (2018).

[2] Lasry & Lions (2006-7).

# Non-local couplings

We focus on non-local interactions

$$f\colon (x,m) \mapsto \int_{\mathbb{T}^2} k(x,y)m(y)\mathrm{d}y + k_0(x),$$

- $k_0 : \mathbb{T}^2 \to \mathbb{R}$ is a Lipschitz function.
- $k : \mathbb{T}^2 \times \mathbb{T}^2 \to \mathbb{R}$ is a smooth function such that[3]

**Positive Definite Symmetric (PDS) property**

for all $(x_r)_{1\leq r\leq \ell} \subset \mathbb{T}^2$ and every $m\colon \mathbb{T}^2 \to \mathbb{R}$,

$$\sum_{r=1}^{\ell}\sum_{s=1}^{\ell} k(x_r,x_s)m(x_r)m(x_s) \geq 0$$

and $k(x_r,x_s) = k(x_s,x_r)$ for every $r$ and $s$ in $\{1,\ldots,\ell\}$.

[3] Mohri, Rostamizadeh, Talkawar (2018)

# Non-local couplings

(PDS) implies that $k$ satisfies,

$$\int_{\mathbb{T}^2\times\mathbb{T}^2} k(x,y)(m_1(x)-m_2(x))(m_1(y)-m_2(y))\mathrm{d}x\mathrm{d}y \geq 0,$$

for all $m_1$ and $m_2$ in $L^1(\mathbb{T}^2)$ such that $\int_{\mathbb{T}^2} m_1(x)\mathrm{d}x = \int_{\mathbb{T}^2} m_2(x)\mathrm{d}x = 1$.

- Under the previous assumptions, the results in Lasry & Lions (2006, 2007) ensure the existence of an unique classical solution $(u,m)$ to (MFG).
- Moreover, (MFG) admits a variational formulation, i.e., it corresponds to the optimality condition of an optimization problem (Lasry & Lions, 2007).

# Variational formulation

$$(P) \quad \inf_{(m,u)} \int_{\mathbb{T}^2} \left[ \frac{1}{2} |\nabla u(x)|^2 + F(x,m) \right] dx,$$

$$\text{s.t.} \quad \begin{cases} -\nu \Delta m + \operatorname{div}(\underbrace{m \nabla u}_{w}) & = 0 \ \text{ in } \mathbb{T}^2, \\ \int_{\mathbb{T}^2} m(x) dx & = 1, \end{cases}$$

$$F(x,m) := \begin{cases} \int_0^m f(x,m') dm', & \text{if } m \geq 0, \\ +\infty, & \text{otherwise}, \end{cases}$$

# Variational formulation

## Benamou & Brenier (2000)

$$\inf_{(m,w)} \int_{\mathbb{T}^2} \left[ b_2(m(x), w(x)) + F(x,m) \right] dx,$$

$$\text{s.t.} \quad \begin{cases} -\nu \Delta m + \operatorname{div}(w) & = 0 \ \text{ in } \mathbb{T}^2, \\ \int_{\mathbb{T}^2} m(x)dx & = 1, \end{cases}$$

where

$$b_2(m,w) := \begin{cases} \frac{|w|^2}{2m}, & \text{if } m > 0, \\ 0, & \text{if } (m,w) = (0,0), \\ +\infty, & \text{otherwise} \end{cases}$$

is convex but not differentiable.

# Numerical solution: Discretization

- $N \in \mathbb{N}$, $h = 1/N$, and $\mathbb{T}_h^2 = \{x_{i,j} = (hi, hj) \mid 0 \leq i, j \leq N\}$.
- $\mathcal{M}_h = \mathbb{R}^{\mathbb{T}_h^2}$, $\mathcal{W}_h = \mathcal{M}_h^4$, $\mathcal{Y}_h = \{z \in \mathcal{M}_h \mid \sum_{i,j} z_{i,j} = 0\}$, where $z_{i,j} = z(x_{i,j})$.
- Define

$$
\begin{aligned}
&(D_1 z)_{i,j} = \tfrac{z_{i+1,j} - z_{i,j}}{h}, \quad (D_2 z)_{i,j} = \tfrac{z_{i,j+1} - z_{i,j}}{h},\\
&[D_h z]_{i,j} = ((D_1 z)_{i,j}, (D_1 z)_{i-1,j}, (D_2 z)_{i,j}, (D_2 z)_{i,j-1}),\\
&(\Delta_h z)_{i,j} = \tfrac{z_{i-1,j} + z_{i+1,j} + z_{i,j-1} + z_{i,j+1} - 4 z_{i,j}}{h^2},\\
&(\mathrm{div}_h(w))_{i,j} = (D_1 w^1)_{i-1,j} + (D_1 w^2)_{i,j} + (D_2 w^3)_{i,j-1} + (D_2 w^4)_{i,j}.
\end{aligned}
$$

- $C = [0, +\infty[ \times ]-\infty, 0] \times [0, +\infty[ \times ]-\infty, 0]$.

$$
(P_{C^{N^2}}(w))_{i,j} = \left([w^1_{i,j}]_+, [w^2_{i,j}]_-, [w^3_{i,j}]_+, [w^4_{i,j}]_-\right).
$$

- For every $\xi \in \mathbb{R}$, $[\xi]_+ = \max\{0, \xi\}$ and $[\xi]_- = \min\{0, \xi\}$.

# Numerical solution: Discretization

- $f(x_{i,j}, m) = (K_h m)_{i,j} + (K_0)_{i,j}$, where

$$(K_h m)_{i,j} = h^2 \sum_{i',j'} k(x_{i,j}, x_{i',j'}) m_{i',j'} \quad \text{and} \quad (K_0)_{i,j} = k_0(x_{i,j}).$$

- (PDS) implies that $K_h$ is positive semidefinite.

**Discrete SMFG (Achdou & Capuzzo Dolcetta, 2010)**

Find $(m, u) \in \mathcal{M}_h^2$ such that, for every $0 \leq i, j \leq N$

$$-\nu(\Delta_h u)_{i,j} + \frac{1}{2}|P_C[D_h u]_{i,j}|^2 + \lambda^h = f(x_{i,j}, m)$$
$$-\nu(\Delta_h m)_{i,j} - \left(\mathrm{div}_h(m P_{C^{N^2}}[D_h u])\right)_{i,j} = 0$$
$$m_{i,j} \geq 0, \; h^2 \sum_{i,j} m_{i,j} = 1, \; \sum_{i,j} u_{i,j} = 0.$$

# Numerical solution: Newton's alg.

- If $\nu > 0$, $f(x, \cdot)$ is increasing and we suppose that the stationary system admits a unique classical solution, in Achdou, Camilli & Capuzzo Dolcetta (2013) the convergence of DSMFG (unif-$L^2$) to the unique solution to the stationary system as $h \to 0$ is proved.
- To solve the discretized system, Newton's method can be used (Achdou & Capuzzo Dolcetta, 2010; Achdou & Perez, 2012; Cacace & Camilli, 2016) if the initial guess is close enough to the solution.
- The performance of Newton's method depends heavily on the values of $\nu$: for small values of $\nu$ the convergence is much slower and cannot be guaranteed in general since $m^h$ can become negative.

# Numerical solution: Optimization Problem $(P_h)$

**Discrete optimization problem $(P_h)$: BA-Kalise-Silva (2018)**

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2 \sum_{i,j=0}^{N-1}\left[\hat{b}(m_{i,j},w_{i,j}) + F(x_{i,j},m)\right]$$

$$\text{s.t.}\quad \begin{cases} -\nu(\Delta_h m)_{i,j} + (\operatorname{div}_h w)_{i,j} = 0, & \forall\, 0\le i,j,\le N-1 \\ h^2\sum_{i,j} m_{i,j} = 1. \end{cases}$$

- $-\nu\Delta_h\colon \mathcal{M}_h \to \mathcal{M}_h$ and $\operatorname{div}_h\colon \mathcal{W}_h \to \mathcal{M}_h$ are linear.
- $\hat{b}\colon \mathbb{R}\times\mathbb{R}^4$ is given by

$$\hat{b}\colon(\mu,\omega)\mapsto\begin{cases} \dfrac{|\omega|^2}{2\,\mu}, & \text{if } \mu>0,\ \omega\in C, \\ 0, & \text{if } (\mu,\omega)=(0,0), \\ +\infty, & \text{otherwise.}\end{cases}$$

# Numerical solution: state of the art

In the case of local couplings $(F(x_{i,j}, m) = F(x_{i,j}, m_{i,j}))$

$(P_h)$

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2 \sum_{i,j=0}^{N-1} \left[\hat{b}(m_{i,j}, w_{i,j}) + F(x_{i,j}, m_{i,j})\right]$$

$$\text{s.t.} \quad \begin{cases} -\nu(\Delta_h m)_{i,j} + (\text{div}_h w)_{i,j} = 0, & \forall\, 0 \leq i, j, \leq N-1 \\ h^2 \sum_{i,j} m_{i,j} = 1. \end{cases}$$

- MFGs local couplings: Andreev (2017), Benamou & Carlier (2015), Benamou, Carlier & Santambrogio (2017), BA, Kalise & Silva (2018), Lachapelle, Salomon & Turinici (2010).
- Methods exploit separability of $(m, w) \mapsto h^2 \sum_{i,j=0}^{N-1} \phi_{i,j}(m_{i,j}, w_{i,j})$ (lost for non-local couplings).

# Numerical solution: state of the art

- Liu, Jacobs, Li, Nurbekyan & Osher (2021) solves (SMFG) via approximated solutions to the analogous time-dependent MFG system as the time horizon goes to infinity.
- In contrast, we propose algorithms following a direct approach based on the variational formulation $(P_h)$.

# Goal of this talk

Comparison of different splitting algorithms for solving $(P_h)$ for non-local couplings.

# Existence and uniqueness

**Proposition**

Let $\nu > 0$. Then, there exists a unique solution $(\widehat{m}, \widehat{u}, \widehat{\lambda})$ to the Discrete SMFG. Moreover, $\widehat{m}$ is strictly positive, and $(\widehat{m}, \widehat{m}P_C(-[D\widehat{u}]))$ is the unique solution to $(P_h)$.

**Proof.**

- Injectivity of $\mathcal{Y}_h \times \mathbb{R} \ni (u, \lambda) \mapsto -\nu\Delta_h(u) + \lambda\mathbf{1} \in \mathcal{M}_h$ and the uniqueness of the multiplier $(\widehat{u}, \widehat{\lambda})$ associated to any $(\widehat{m}, \widehat{w})$ solution to $(P_h)$[4].
- Hence, by convexity of $(P_h)$, we deduce uniqueness of $(\widehat{u}, \widehat{\lambda})$ in Discrete SMFG.
- The uniqueness of $\widehat{m}$ follows from rank-nullity and finite dimensional arguments using the particular structure of the continuity equation.

[4] Existence and $> 0$ guaranteed in BA, Kalise, Silva (2018)

# Optim. problem structure

- Since $f(x_{i,j},m) = h^2 \sum_{i',j'} k(x_{i,j},x_{i',j'})m_{i',j'} + k_0(x_{i,j})$, we have

$$F(x_{i,j},m) = \frac{1}{2}\langle m \mid K_h m\rangle + \langle K_0 \mid m\rangle,$$

  where $\langle m \mid u\rangle = \sum_{i,j} m_{i,j}u_{i,j}$.
- (PDS) implies $F(x_{i,j},\cdot)$ is convex.

$(P_h)$

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2 \sum_{i,j=0}^{N-1} [\hat{b}(m_{i,j},w_{i,j}) + (K_0)_{i,j}m_{i,j}] + \frac{1}{2}\langle m \mid K_h m\rangle$$

$$\begin{cases} -\nu(\Delta_h m)_{i,j} + (\mathrm{div}_h w)_{i,j} = 0, & \forall\, 0 \le i,j \le N-1 \\ h^2 \sum_{i,j} m_{i,j} = 1. \end{cases}$$

# Formulation with vector subspaces

$(P_h)$

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m\mid K_h m\rangle$$

$$\begin{cases}-\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\leq i,j\leq N-1\\ h^2\sum_{i,j} m_{i,j}=1. & \end{cases}$$

- Change of variables: $\rho = m - \mathbf{1}$.
- Note that $\Delta_h\mathbf{1}=0$.
- Constraints:

$$\begin{cases}-\nu(\Delta_h \rho)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\leq i,j\leq N-1\\ h^2\sum_{i,j} \rho_{i,j}=h^2\langle\mathbf{1}\mid\rho\rangle=0. & \end{cases}$$

- $(\rho,w)\in\ker L$, where
  $L\colon(\rho,w)\mapsto(-\nu\Delta_h\rho+\mathrm{div}_h(w),h^2\langle\mathbf{1}\mid\rho\rangle)$.

# Formulation with vector subspaces

$(P_h)$

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m\mid K_h m\rangle$$

$$\begin{cases}-\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\le i,j\le N-1\\ h^2\sum_{i,j} m_{i,j}=1. \quad \leftarrow S\end{cases}$$

$$S\cap\arg\min_{(\rho,w)\in\ker L} f(\rho+\mathbf{1},w)+g(\rho+\mathbf{1},w)+h_1(L_1(\rho+\mathbf{1},w)),$$

$$\begin{cases}f\colon (m,w)\mapsto \frac{1}{2}\langle m\mid K_h m\rangle\\ g\colon (m,w)\mapsto h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]\\ h=\iota_{\{(0,1)\}}\\ L\colon (\rho,w)\mapsto(-\nu\Delta_h\rho+\mathrm{div}_h(w),h^2\langle\mathbf{1}\mid\rho\rangle).\end{cases}$$

# Convex non-differentiable optimization problem

## Problem (P)

$$\text{find} \quad x \in Z = S \cap \arg\min_{x\in V} f(x) + g(x) + h(Lx).$$

- $S \subset \mathbb{R}^N$ closed convex (a priori information).
- $V \subset \mathbb{R}^N$ closed vector subspace.
- $f\colon \mathbb{R}^N \to \mathbb{R}$ is differentiable and $\nabla f$ is $\beta^{-1}-$Lipschitz.
- $g \in \Gamma_0(\mathbb{R}^N)$ and $h \in \Gamma_0(\mathbb{R}^M)$:
  - **convex :** $(\forall x, y \in \mathbb{R}^N)(\forall \lambda \in [0,1])$ $f(x + \lambda(y-x)) \leq f(x) + \lambda(f(y) - f(x))$.
  - **lower semicontinuous (l.s.c):** $(\forall x \in \mathbb{R}^N) \quad f(x) \leq \liminf_{y\to x} f(y)$.
  - **proper:** $f$ is not always $+\infty$ and never $-\infty$.
- $L$ is a $M \times N$ real matrix.
- $Z \neq \emptyset$.

# Basic case: $S = V = \mathbb{R}^N$ and $f = h = 0$[5]

**Problem (P)**

$$\text{find} \quad x \in S \cap \arg\min_{x \in V} f(x) + g(x) + h(Lx).$$

[5] Rockafellar, 1973; Martinet, 1970

# Basic case: $S = V = \mathbb{R}^N$ and $f = h = 0$[5]

**Problem (P)**

$$\min_{x\in\mathbb{R}^N} g(x)$$

**Proximal point algorithm (PPA)**

$$x_0 \in \mathbb{R}^N, \quad x_{n+1} = \mathrm{prox}_{\tau g} x_n.$$

**Proximity operator**

$$\mathrm{prox}_{\tau g} : x \mapsto \operatorname*{argmin}_{y\in\mathbb{R}^N} \; g(y) + \frac{1}{2}\|y - x\|^2$$

- $\mathrm{prox}_{\iota_C} = P_C$.
- Differentiable case:

**PPA**

$$-\frac{x_{n+1} - x_n}{\tau} = \nabla g(x_{n+1})$$

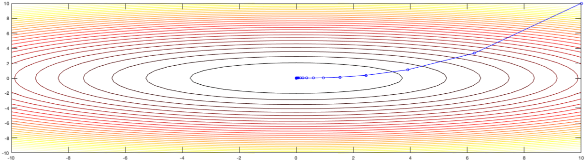
**Implicit discretization**

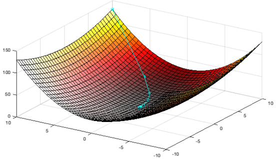
$$\begin{cases} -x'(t) = \nabla g(x(t)), & t > 0 \\ x(0) = x_0. \end{cases}$$

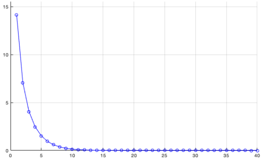
[5] Rockafellar, 1973; Martinet, 1970

# Example PPA: $\tau = 1$

- $f\colon (x,y) \mapsto 0.3\,x^2 + y^2$
- $\mathrm{prox}_{\tau f}\colon (x,y) \mapsto (x/(1+0.6\,\tau), y/(1+2\,\tau))$
- $\chi = 2$, $\beta = 0.6$

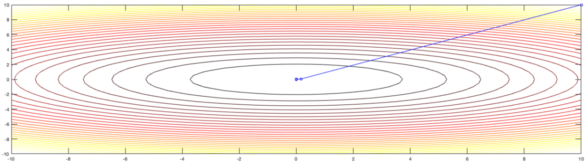


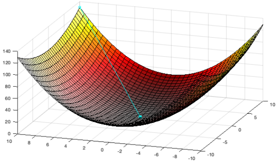


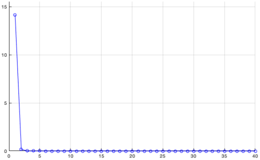


# Example PPA: $\tau = 100$

- $f\colon (x,y) \mapsto 0.3\,x^2 + y^2$
- $\mathrm{prox}_{\tau f}\colon (x,y) \mapsto (x/(1+0.6\,\tau), y/(1+2\,\tau))$
- $\chi = 2,\ \beta = 0.6$







# Algorithm[6]

## Problem (P)

$$\text{find} \quad x \in Z = S \cap \arg\min_{x \in V} f(x) + g(x) + h(Lx).$$

## Primal-dual partial inverse splitting

Set $x^0 = \overline{x}^0 \in V$, $y^0 \in V^\perp$, $u^0 \in \mathcal{G}$, $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$.

$$\left\lfloor \begin{array}{ll} u^{k+1} & = \operatorname{prox}_{\gamma h^*}\left(u^k + \gamma L\overline{x}^k\right) \\ w^{k+1} & = \operatorname{prox}_{\tau g}\left(x^k + \tau y^k - \tau P_V(L^* u^{k+1} + \nabla f(x^k))\right) \\ r^{k+1} & = P_V w^{k+1} \\ x^{k+1} & = P_V P_S r^{k+1} \\ y^{k+1} & = y^k + (r^{k+1} - w^{k+1})/\tau \\ \overline{x}^{k+1} & = x^{k+1} + r^{k+1} - x^k. \end{array} \right.$$

Then, $x^k \to x$ solution to $(P)$

[6] BA, Deride, López, and Silva, AMO, 2022.

# It generalizes...

$$
\left\lfloor
\begin{array}{ll}
u^{k+1} & = \operatorname{prox}_{\gamma h^*}\left(u^k + \gamma L\overline{x}^k\right) \\
w^{k+1} & = \operatorname{prox}_{\tau g}\left(x^k + \tau y^k - \tau P_V\left(L^* u^{k+1} + \nabla f(x^k)\right)\right) \\
r^{k+1} & = P_V w^{k+1} \\
x^{k+1} & = P_V P_S r^{k+1} \\
y^{k+1} & = y^k + (r^{k+1} - w^{k+1})/\tau \\
\overline{x}^{k+1} & = x^{k+1} + r^{k+1} - x^k.
\end{array}
\right.
$$

- Projected primal-dual (2019): $V = \mathbb{R}^N$.

# It generalizes...

$$
\left\lfloor
\begin{array}{ll}
u^{k+1} & = \mathrm{prox}_{\gamma h^*}\left(u^k + \gamma L\overline{x}^k\right) \\
w^{k+1} & = \mathrm{prox}_{\tau g}\left(x^k + \tau y^k - \tau P_V\left(L^* u^{k+1} + \nabla f(x^k)\right)\right) \\
r^{k+1} & = P_V w^{k+1} \\
x^{k+1} & = P_V P_S r^{k+1} \\
y^{k+1} & = y^k + (r^{k+1} - w^{k+1})/\tau \\
\overline{x}^{k+1} & = x^{k+1} + r^{k+1} - x^k.
\end{array}
\right.
$$

- Projected primal-dual (2019): $V = \mathbb{R}^N$.
- Forward-Partial Inverse (2015): $S = \mathbb{R}^N$ and $h = 0$ $(h^* = \iota_{\{0\}})$.
- Condat-Vũ (2013), Chambolle-Pock (2011), Proximal-gradient, PPA, etc.

# Formulations

## $(P_h)$

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m\mid K_h m\rangle$$

$$\begin{cases}-\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\leq i,j\leq N-1\\ h^2\sum_{i,j} m_{i,j}=1. & \leftarrow S\end{cases}$$

## Form. 1: Projected primal-dual (2019)

$$S\cap\arg\min_{(m,w)} f(m,w)+g(m,w)+h_1(L_1(m,w)),$$

$$\begin{cases}f\colon(m,w)\mapsto\frac{1}{2}\langle m\mid K_h m\rangle\\ g\colon(m,w)\mapsto h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]\\ h_1=\iota_{\{(0,1)\}}\\ L_1\colon(m,w)\mapsto(-\nu\Delta_h m+\mathrm{div}_h(w),h^2\langle\mathbf{1}\mid m\rangle).\end{cases}$$

# Formulations

**$(P_h)$**

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m \mid K_h m\rangle$$

$$\begin{cases} -\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\le i,j\le N-1\\ h^2\sum_{i,j} m_{i,j}=1. \end{cases}$$

**Form. 2: Condat-Vũ (2013)**

$$\min_{(m,w)} f(m,w)+g(m,w)+h_2(L_2(m,w)),$$

$$\begin{cases} f\colon (m,w)\mapsto \frac{1}{2}\langle m \mid K_h m\rangle & \leftarrow \nabla f\\ g\colon (m,w)\mapsto h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}] & \\ h_2=\iota_{\{(0,1)\}}\circ L_1 & \leftarrow \mathrm{prox}_{h_2}\\ L_2=\mathrm{Id}. & \end{cases}$$

# Formulations

**$(P_h)$**

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m \mid K_h m\rangle$$

$$\begin{cases}-\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\le i,j\le N-1\\ h^2\sum_{i,j} m_{i,j}=1.\end{cases}$$

**Form. 3: Forward-partial inverse (2015)**

$$\min_{(\rho,w)\in V} f(\rho+\mathbf{1},w)+g(\rho+\mathbf{1},w),$$

$$\begin{cases}f\colon(m,w)\mapsto \frac{1}{2}\langle m \mid K_h m\rangle \quad \to \nabla f\\ g\colon(m,w)\mapsto h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]\\ V=\ker L_1.\end{cases}$$

# Formulations

**$(P_h)$**

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]+\frac{1}{2}\langle m \mid K_h m\rangle$$

$$\begin{cases}-\nu(\Delta_h m)_{i,j}+(\mathrm{div}_h w)_{i,j}=0, & \forall\, 0\le i,j\le N-1\\ h^2\sum_{i,j} m_{i,j}=1.\end{cases}$$

**Form. 4: Primal-dual-partial inverse 1**

$$\min_{(\rho,w)\in V} f(\mathrm{Id}(\rho+\mathbf{1},w))+g(\rho+\mathbf{1},w),$$

$$\begin{cases} f\colon (m,w)\mapsto \frac{1}{2}\langle m \mid K_h m\rangle \quad \to \mathrm{prox}_f\\ g\colon (m,w)\mapsto h^2\sum_{i,j=0}^{N-1}[\hat{b}(m_{i,j},w_{i,j})+(K_0)_{i,j}m_{i,j}]\\ V=\ker L_1.\end{cases}$$

# Formulations

**$(P_h)$**

$$\inf_{(m,w)\in\mathcal{M}_h\times\mathcal{W}_h} h^2 \sum_{i,j=0}^{N-1} [\hat{b}(m_{i,j}, w_{i,j}) + (K_0)_{i,j} m_{i,j}] + \frac{1}{2}\langle m \mid K_h m\rangle$$

$$\begin{cases} -\nu(\Delta_h m)_{i,j} + (\mathrm{div}_h w)_{i,j} = 0, \quad \forall\, 0 \le i, j \le N-1 \\ h^2 \sum_{i,j} m_{i,j} = 1. \end{cases}$$

**Form. 5: Primal-dual-partial inverse 2**

$$\min_{(\rho,w)\in V} h_3(L_3(\rho + \mathbf{1}, w)) + g(\rho + \mathbf{1}, w),$$

$$\begin{cases} h_3 = \|\cdot\|^2/2 \quad \leftarrow \mathrm{prox}_{h_3} \\ L_3 \colon (m, w) \mapsto \sqrt{K_h} m \\ g \colon (m, w) \mapsto h^2 \sum_{i,j=0}^{N-1} [\hat{b}(m_{i,j}, w_{i,j}) + (K_0)_{i,j} m_{i,j}] \\ V = \ker L_1. \end{cases}$$

# $\text{prox}_{\gamma g}$ computation

- Recall that $g\colon (m,w) \mapsto \sum_{i,j} \phi_{i,j}(m_{i,j}, w_{i,j})$, where $\phi_{i,j}\colon (\mu,\omega) \mapsto \hat{b}(\mu,\omega) + (K_0)_{i,j}\mu$.
- We have $\text{prox}_{\gamma g}(m,w) = (\text{prox}_{\gamma\phi_{i,j}}(m_{i,j}, w_{i,j}))_{i,j}$.

## Prox computation

$$\text{prox}_{\gamma\phi_{i,j}} : (\mu,\omega) \mapsto \begin{cases} (0,0), & \text{if } \mu + \frac{1}{2\gamma}|P_C\omega|^2 \leq \gamma (K_0)_{i,j}; \\ (p^*, p^*\, P_C\omega/(p^*+\gamma)), & \text{otherwise }, \end{cases}$$

where $p^* \geq 0$ is the unique solution to

$$(p + \gamma(K_0)_{i,j} - \mu)(p+\gamma)^2 - \frac{\gamma}{2}|P_C\omega|^2 = 0.$$

# Example

- $h \in \{1/20, 1/40\}$.
- $\nu \in \{0.05, 0.2, 0.5\}$.
- $K_h = \mu(\mathrm{Id} - \Delta_h)^{-p}$, for $\mu = 10$ and $p = 1$.[7]

$$(\forall (i,j) \in \mathcal{I}_N^2) \quad (K_0)_{i,j} = -\sin(2\pi hj) - \sin(2\pi hi) - \cos(4\pi hi). \tag{1}$$

For every $h \in \{1/20, 1/40\}$, we choose to stop every algorithm when the $L^2$ norm of the difference between two consecutive iterations is less than $5h^3$ or the number of iterations exceeds 3000.

[7] Achdou, Capuzzo Dolceta (2010)

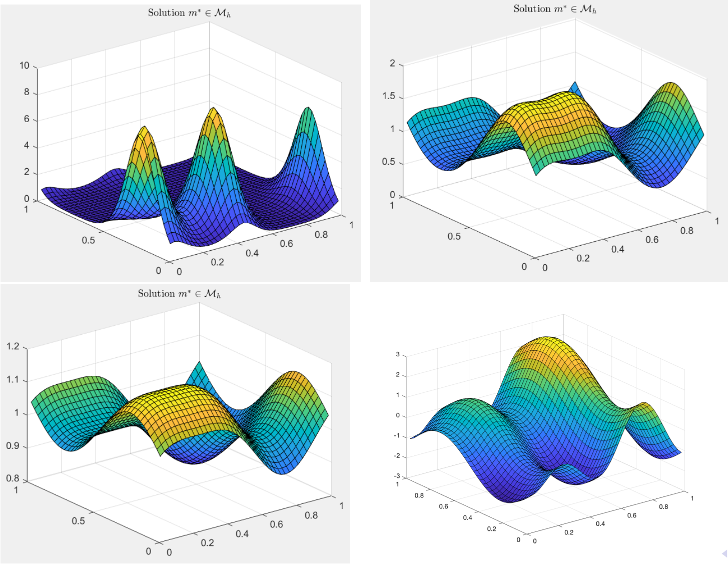
Table: Tolerance $5h^3$, $h = 1/40$

| $\nu = 0.05$ | time(s) | iter. | obj. value | $\|L_1(m^*, w^*) - (0,1)\|$ | $d_C^2(w^*)$ |
|---|---|---|---|---|---|
| Form. 1 | 22.32 | 462 | 80000 | 7.365 | 0 |
| Form. 2 | 2389.47 | 3000 | 80705 | 0.00786 | 0 |
| Form. 3 | 2776.92 | 1915 | 80710 | $3.152 \cdot 10^{-12}$ | $5.686 \cdot 10^{-11}$ |
| Form. 4 | 4189.86 | 3000 | 80717 | $4.061 \cdot 10^{-12}$ | $3.258 \cdot 10^{-6}$ |
| Form. 5 | 751.18 | 695 | 80717 | $3.757 \cdot 10^{-12}$ | $9.748 \cdot 10^{-9}$ |

| $\nu = 0.2$ | time(s) | iter. | obj. value | $\|L_1(m^*, w^*) - (0,1)\|$ | $d_C^2(w^*)$ |
|---|---|---|---|---|---|
| Form. 1 | 38.55 | 727 | 80000 | 27.414 | 0 |
| Form. 2 | 533.00 | 727 | 80369 | 0.00770 | 0 |
| Form. 3 | 636.59 | 461 | 80372 | $4.219 \cdot 10^{-12}$ | $1.298 \cdot 10^{-10}$ |
| Form. 4 | 1387.21 | 950 | 80376 | $4.708 \cdot 10^{-12}$ | $5.260 \cdot 10^{-12}$ |
| Form. 5 | 136.66 | 119 | 80375 | $4.491 \cdot 10^{-12}$ | $2.492 \cdot 10^{-9}$ |

| $\nu = 0.5$ | time(s) | iter. | obj. value | $\|L_1(m^*, w^*) - (0,1)\|$ | $d_C^2(w^*)$ |
|---|---|---|---|---|---|
| Form. 1 | 37.06 | 724 | 80000 | 70.955 | 0 |
| Form. 2 | 286.23 | 387 | 80082 | 0.0296 | 0 |
| Form. 3 | 428.79 | 251 | 80083 | $8.392 \cdot 10^{-12}$ | $2.099 \cdot 10^{-11}$ |
| Form. 4 | 2036.42 | 950 | 80085 | $8.279 \cdot 10^{-12}$ | $1.335 \cdot 10^{-12}$ |
| Form. 5 | 111.39 | 100 | 80085 | $8.408 \cdot 10^{-12}$ | $2.984 \cdot 10^{-12}$ |

# Figures ($\nu = 0.05, 0.2, 0.5$)



# Extensions

- Monotone operators in Hilbert spaces (weak convergence)
- $S \to \operatorname{Fix} T$ for $T$ averaged nonexpansive (common solutions)
- Parallel sums/inf-convolutions in primal problem
- $S = \cap_{i \in I} S_i$ and $V = \mathbb{R}^N$: Kaczmarz type extensions[8]
- $V = \mathbb{R}^N$: Acceleration under presence of strong monotonicity/convexity (Chambolle-Pock, 2011).

[8] L. M. Briceño-Arias, J. Deride, and C. Vega, Random activations in primal-dual splittings for monotone inclusions with a priori information, *J. Optim. Theory Appl.*, vol. 192, pp. 56–81, 2022.

# References


- Y. Achdou, I. Capuzzo Dolcetta, Mean field games: Numerical methods. *SIAM J. Numer. Anal.*, 48:1136–1162, 2010.
- Y. Achdou, F. Camilli, I. Capuzzo Dolcetta. Mean field games: convergence of a finite difference method. *SIAM J. Numer. Anal.*, 51:2585–2612, 2013.
- L. M. Briceño-Arias, S. López Rivera, A projected primal-dual splitting for solving constrained monotone inclusions, *J. Optim. Theory Appl.*, 180:907–924, 2019.
- L. M. Briceño-Arias, J. Deride, S. López Rivera, and F. J. Silva, A Primal-dual partial inverse splitting for constrained monotone inclusions: applications to stochastic programming and mean field games, *Appl. Math. Optim.*, to appear, 2022.
- L. M. Briceño-Arias, D. Kalise, F. J. Silva, Proximal methods for stationary Mean Field Games with local couplings, *SIAM J. Control Optim.*, 56: 801–836, 2018.
- A. Chambolle, T. Pock. A first order primal dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.*, 40:120–145, 2011.
- J.-M. Lasry, P.L. Lions, Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris* 343: 619–625 2006.
- J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.* 2:229–260, 2007.