Generalized conditional gradient for potential mean field games[1]

Pierre Lavigne, Laurent Pfeiffer

Institut Louis Bachelier

January 7, 2022

[1] ArXiv : [BLP21]

Contents

Introduction

Mean field games (MFGs): first introduced by J.-M. Lasry and P.-L. Lions [LL06] and M. Huang, R. Malhamé, and P. Caines in [HMC06], to study **interactions among a large population of rational players**.

Main features:

- **Asymptotic** models of N-rational and identical players,
- Interaction through a **mean field** effect,
- Non-cooperative games. Notion of solution: **Nash equilibrium**.

Examples of applications[2]

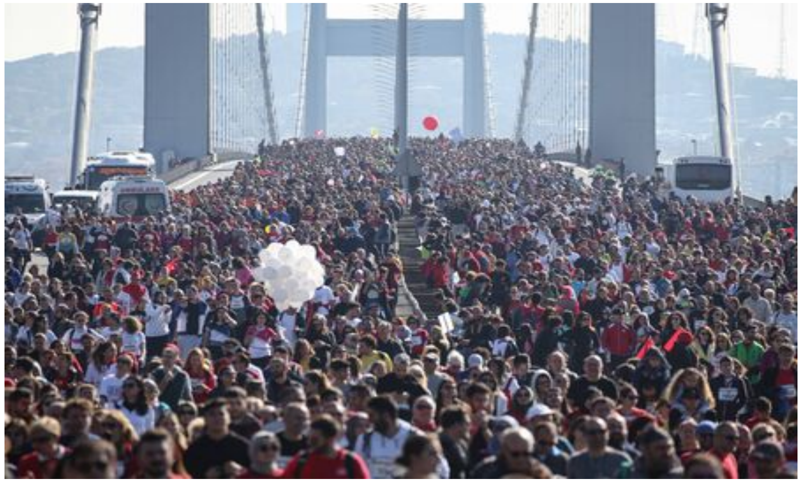


Figure: Crowd motion

Figure: Electrical systems

Figure: Finance

Figure: Flock motion

[2] Credits: evrenkalinbacak (Crowd motion) ; Viktor Yelantsev (Electrical systems) ; Sergey Nivens (Finance) ; A.G.D. Beukhof (Flock motion).

Fixed point structure

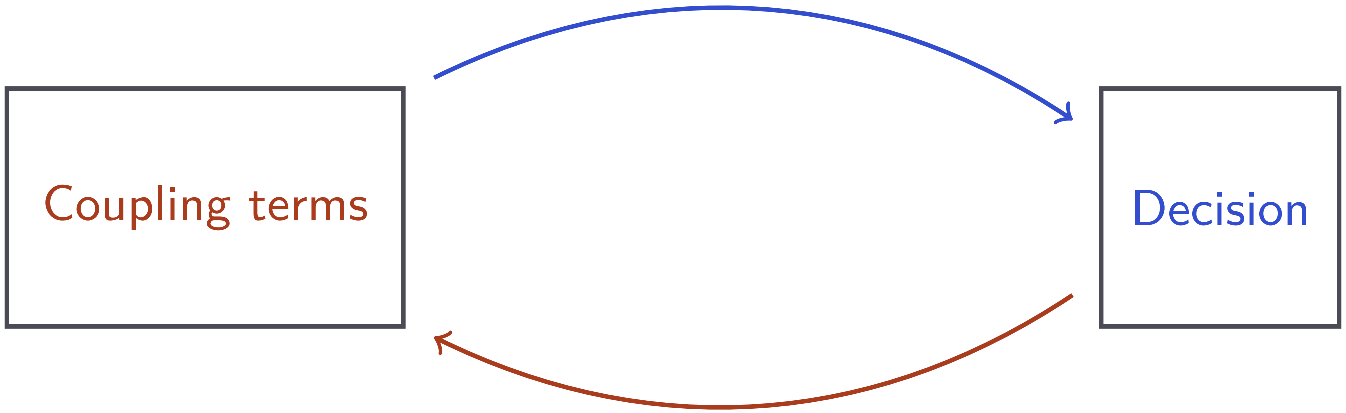
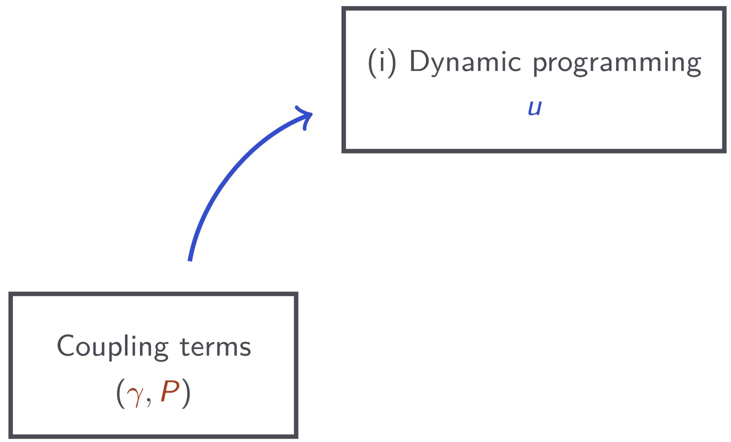


Figure: Fixed point structure of the mean field game problem.

Fixed point structure

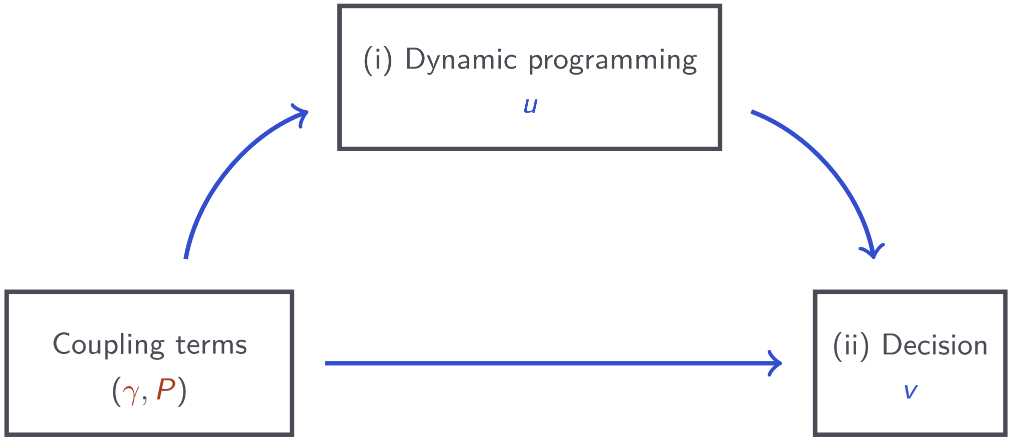
$$\left\{ \text{(i)} \quad \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, \\ u(x,T) = g(x), \end{cases} \right.$$



Unknowns:
Value fonction u

Fixed point structure

$$\begin{cases} \text{(i)} & \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, \\ u(x,T) = g(x), \end{cases} \\ \text{(ii)} & v = -H_p(\nabla u + P), \end{cases}$$

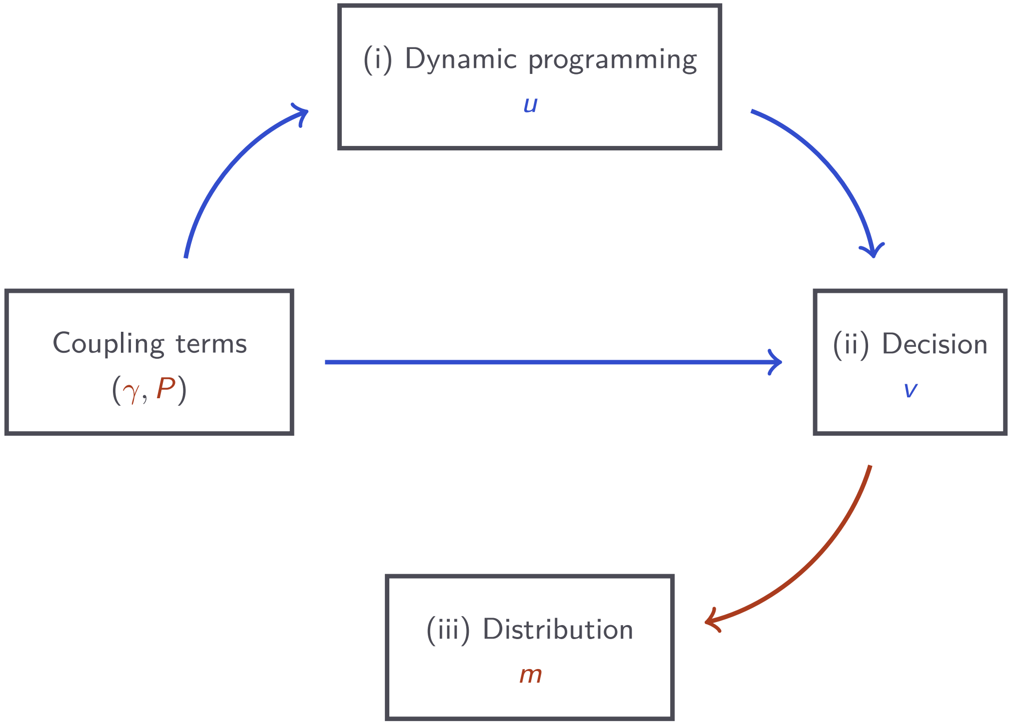


Unknowns:

Value fonction	u
Control	v

Fixed point structure

$$
\begin{cases}
\text{(i)} & \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, \\ u(x,T) = g(x), \end{cases} \\
\text{(ii)} & v = -H_p(\nabla u + P), \\
\text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, \\ m(0,x) = m_0(x), \end{cases}
\end{cases}
$$



Unknowns:

Value fonction	u
Control	v
Distribution	m

Fixed point structure

$$
\begin{cases}
\text{(i)} & \begin{cases} -\partial_t u - \Delta u + H(\nabla u + P) = \gamma, \\ u(x,T) = g(x), \end{cases} \\
\text{(ii)} & v = -H_p(\nabla u + P), \\
\text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, \\ m(0,x) = m_0(x), \end{cases} \\
\text{(iv)} & \gamma(x,t) = f(x, m(t)), \\
\text{(v)} & P(t) = \phi\left(\int_{\mathbb{T}^d} v(x,t) m(x,t)\right),
\end{cases}
$$

Unknowns:

Value fonction	u
Control	v
Distribution	m
Congestion term	γ
Price term	P

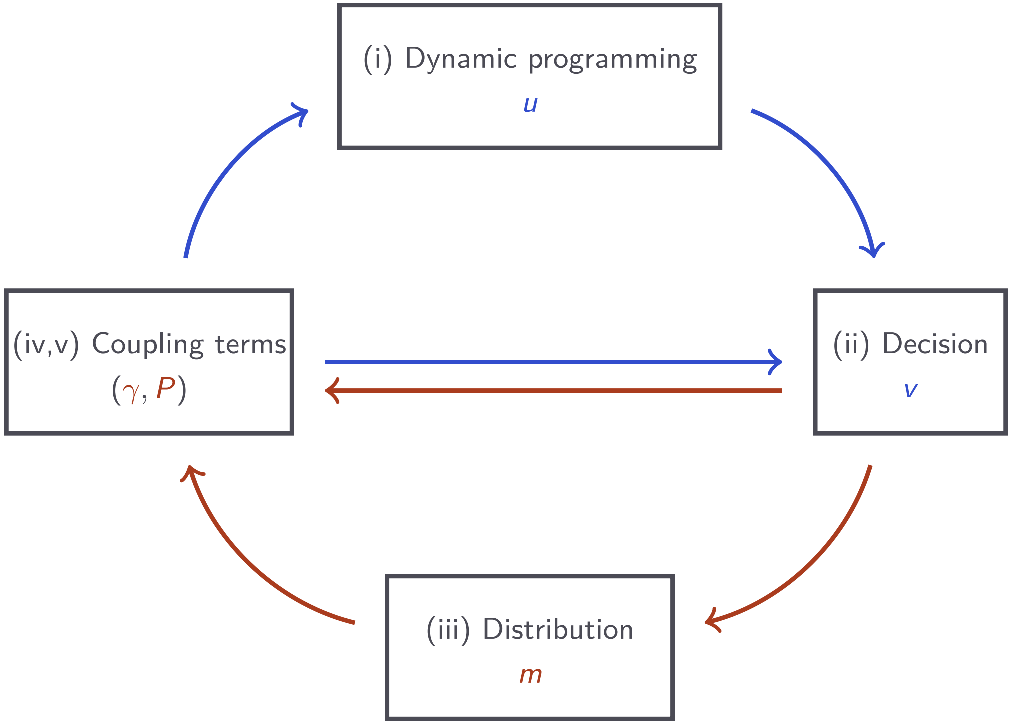


Figure: Fixed point structure of the mean field game problem.

Contents

Framework and contributions

- Present the **generalized conditional gradient** (GCG) algorithm,

[3] Idea also developed in discrete framework for the Frank-Wolfe algorithm [GPL$^+$21]: M. Geist and al. Concave utility reinforcement learning: the mean field game viewpoint, 2021.

Framework and contributions

- Present the **generalized conditional gradient** (GCG) algorithm,
- Apply this algorithm to **potential mean field games**,

[3] Idea also developed in discrete framework for the Frank-Wolfe algorithm [GPL$^+$21]: M. Geist and al. Concave utility reinforcement learning: the mean field game viewpoint, 2021.

Framework and contributions

- Present the **generalized conditional gradient** (GCG) algorithm,
- Apply this algorithm to **potential mean field games**,
- Establish a link between the GCG algorithm and the **fictitious play** algorithm[3],

[3] Idea also developed in discrete framework for the Frank-Wolfe algorithm [GPL+21]: M. Geist and al. Concave utility reinforcement learning: the mean field game viewpoint, 2021.

Framework and contributions

- Present the **generalized conditional gradient** (GCG) algorithm,
- Apply this algorithm to **potential mean field games**,
- Establish a link between the GCG algorithm and the **fictitious play** algorithm[3],
- Establish **convergence results**: convergence in potential cost $O(1/k)$ of the potential cost, and in $O(1/\sqrt{k})$ of the exploitability and the variables of the problem (distribution, congestion, price, value function and control terms).

[3] Idea also developed in discrete framework for the Frank-Wolfe algorithm [GPL+21]: M. Geist and al. Concave utility reinforcement learning: the mean field game viewpoint, 2021.

Abstract framework

Consider the optimization problem

$$\min_{x \in K} f(x) = f_1(x) + f_2(x).$$

Assumptions on the data:

- K is a convex and compact subset of $\mathbb{R}^n$ of finite diameter D,
- f_1 is a (possibly non-smooth) convex function,
- f_2 a continuous differentiable function with L-Lipschitz gradient.

Abstract generalized conditional gradient algorithm

Consider the mapping $h\colon K \times K \to \mathbb{R}$ defined by

$$h(x,y) = f_1(y) - f_1(x) + \langle \nabla f_2(x), y - x \rangle.$$

Algorithm 1 Generalized conditional gradient

Choose $\bar{x}_0 \in K$ and choose a sequence $(\delta_k)_{k\in\mathbb{N}} \in [0,1]$.
for $0 \leq k < N$ **do**
  Find $x_k \in \arg\min_{y \in K} h(\bar{x}_k, y)$
  Update $\bar{x}_{k+1} = (1-\delta_k)\bar{x}_k + \delta_k x_k$
end for
return $\bar{x}_N$.

Let $\bar{x} = \arg\min_{x\in K} f(x)$. Following standard arguments[4] we have for $\delta_k = 2/(k+2)$,

$$f(\bar{x}_k) - f(\bar{x}) \leq C/(k+2).$$

[4][Jag13]: Martin Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization, 2013.

Mean field game system

$$
\begin{cases}
\text{(i)} & \begin{cases} -\partial_t u - \Delta u + \boldsymbol{H}[\nabla u + A^\star P] = \gamma, \\ u(x,T) = g(x), \end{cases} \\
\text{(ii)} & v = -\boldsymbol{H}_p[\nabla u + A^\star P], \\
\text{(iii)} & \begin{cases} \partial_t m - \Delta m + \nabla \cdot (vm) = 0, \\ m(0,x) = m_0(x), \end{cases} \\
\text{(iv)} & \gamma(x,t) = f(x,t,m(t)), \\
\text{(v)} & P(t) = \phi[A[vm]](t).
\end{cases}
$$

Unknowns:

Value fonction	u
Control	v
Distribution of states	m
Congestion	γ
Price	P

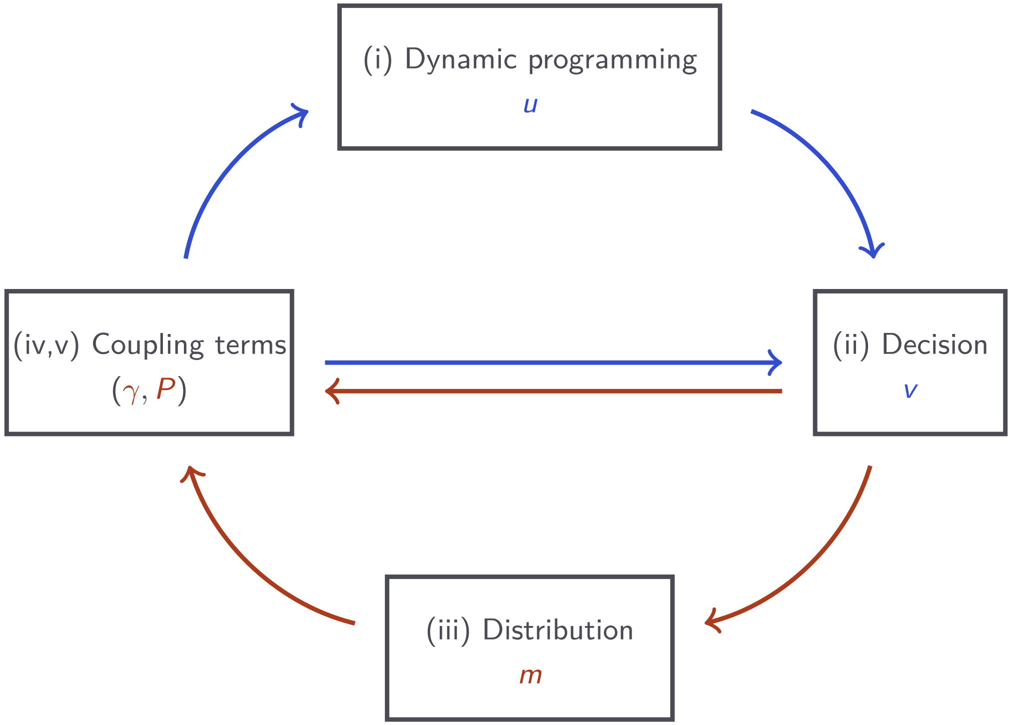


Figure: Fixed point structure of the mean field game problem.

Assumptions

Nature of the assumptions on the data:

- $m_0 \in \mathcal{C}^{2+\alpha_0}(\mathbb{T}^d), \quad g \in \mathcal{C}^{2+\alpha_0}(\mathbb{T}^d)$.
- Regularity assumptions on $\boldsymbol{L}$, f and ϕ and boundedness of f and ϕ.
- **There exists convex** maps $F\colon [0,T] \times \mathcal{D}_1(\mathbb{T}^d)$ and $\Phi : [0,T] \times \mathbb{R}^k \to \mathbb{R}$ such that

$$F(t, m_2) - F(t, m_1) = \int_0^1 \int_{\mathbb{T}^d} f(x, t, sm_2 + (1-s)m_1)(m_2(x) - m_1(x)) \mathrm{d}x \mathrm{d}s,$$

$$\phi(t,z) = \nabla_z \Phi(t,z).$$

Theorem

There exists[a] $\alpha \in (0,1)$ *such that (MFG) has a unique classical solution* $(\bar{m}, \bar{v}, \bar{u}, \bar{\gamma}, \bar{P})$, *with*

$$\begin{cases} \bar{m} \in & \mathcal{C}^{2+\alpha,1+\alpha/2}(Q), \\ \bar{v} \in & \mathcal{C}^{1+\alpha,\alpha}(Q;\mathbb{R}^d), \\ \bar{u} \in & \mathcal{C}^{2+\alpha,1+\alpha/2}(Q), \\ \bar{\gamma} \in & \mathcal{C}^{\alpha}(Q), \\ \bar{P} \in & \mathcal{C}^{\alpha}(0,T;\mathbb{R}^k). \end{cases}$$

[a] [BHP21]: J. F. Bonnans, S. Hadikanloo, L. Pfeiffer, Schauder Estimates for a Class of Potential Mean Field Games of Controls, 2019.

Question : **how to compute this solution ?**

Approach:

- Use the variational form of the mean field game,
- Fit the framework of the GCG,
- Show the convergence of the method.

A potential mean field game

Let $\mathcal{B}^p := W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$. We define the following primal problem

$$\begin{cases} \inf_{(m,v)\in\mathcal{B}^p} \mathcal{J}(m,v) := \int_Q \boldsymbol{L}[v] m \mathrm{d}x\mathrm{d}t + \int_0^T \left(\boldsymbol{F}[m] + \boldsymbol{\Phi}[A[mv]]\right)\mathrm{d}t + \int_{\mathbb{T}^d} g m(T)\mathrm{d}x, \\ s.t: \quad \begin{aligned} \partial_t m - \Delta m + \nabla\cdot(vm) &= \quad 0, & (x,t)\in Q, \\ m(x,0) &= \quad m_0(x), & x \in \mathbb{T}^d. \end{aligned} \end{cases}$$

The above problem is **not convex**.

A potential mean field game

Let $\mathcal{B}^p := W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$. We define the following primal problem

$$\begin{cases} \inf_{(m,v)\in\mathcal{B}^p} \mathcal{J}(m,v) := \int_Q \boldsymbol{L}[v] m \mathrm{d}x\mathrm{d}t + \int_0^T (\boldsymbol{F}[m] + \boldsymbol{\Phi}[A[mv]])\, \mathrm{d}t + \int_{\mathbb{T}^d} g m(T) \mathrm{d}x, \\ s.t: \quad \partial_t m - \Delta m + \nabla \cdot (vm) = 0, \qquad (x,t) \in Q, \\ \qquad\qquad m(x,0) = m_0(x), \qquad x \in \mathbb{T}^d. \end{cases}$$

The above problem is **not convex**. Using the change of variable “à la Benamou-Brenier” $w = mv$ we define the following **convex** problem

$$\inf_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m,w) := \int_Q \tilde{\boldsymbol{L}}[m,w] \mathrm{d}x\mathrm{d}t + \int_0^T (\boldsymbol{F}[m] + \boldsymbol{\Phi}[Aw])\, \mathrm{d}t + \int_{\mathbb{T}^d} g m(T) \mathrm{d}x, \quad (\tilde{\mathrm{P}})$$

$$\tilde{\mathcal{R}} := \{(m,w) \in \mathcal{B}^p,\ \partial_t m - \Delta m + \nabla \cdot w = 0,\ m(0) = m_0,\ m(x,t) > 0,\ (x,t) \in Q\}.$$

Application to potential mean field games

- We have the following convex problem

$$\inf_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m,w) := \int_Q \tilde{\boldsymbol{L}}[m,w]\mathrm{d}x\mathrm{d}t + \int_0^T (\boldsymbol{F}[m] + \boldsymbol{\Phi}[Aw])\,\mathrm{d}t + \int_{\mathbb{T}^d} gm(T)\mathrm{d}x, \tag{$\tilde{\mathrm{P}}$}$$

- Define the following semi-linearization mapping $h : \tilde{\mathcal{R}} \times \tilde{\mathcal{R}} \to \mathbb{R}$ of the potential cost,

$$h(\underbrace{(m,w)}_{x},\underbrace{(m',w')}_{y}) = \underbrace{\int_Q \left(\tilde{\boldsymbol{L}}[m',w'] - \tilde{\boldsymbol{L}}[m,w]\right)\mathrm{d}x\mathrm{d}t + \int_{\mathbb{T}^d} g(m'-m)(T)\mathrm{d}x}_{f_1(y)-f_1(x)}$$

$$+ \underbrace{\int_Q \gamma(m'-m)\mathrm{d}x\mathrm{d}t + \int_0^T \langle A[w'-w], P\rangle \mathrm{d}t}_{\langle \nabla f_2(x), y-x\rangle},$$

where $\gamma(x,t) = f(x,t,m(t))$ and $P(t) = \phi(t, Aw(t))$ for any $(x,t) \in Q$.

Application to potential mean field games

- We have the following convex problem

$$\inf_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m,w) := \int_Q \tilde{\boldsymbol{L}}[m,w]\mathrm{d}x\mathrm{d}t + \int_0^T (\boldsymbol{F}[m] + \boldsymbol{\Phi}[Aw])\,\mathrm{d}t + \int_{\mathbb{T}^d} gm(T)\mathrm{d}x, \tag{$\tilde{\mathrm{P}}$}$$

- Define the following semi-linearization mapping $h : \tilde{\mathcal{R}} \times \tilde{\mathcal{R}} \to \mathbb{R}$ of the potential cost,

$$h(\underbrace{(m,w)}_{x},\underbrace{(m',w')}_{y}) = \underbrace{\int_Q \left(\tilde{\boldsymbol{L}}[m',w'] - \tilde{\boldsymbol{L}}[m,w]\right)\mathrm{d}x\mathrm{d}t + \int_{\mathbb{T}^d} (gm' - gm)(T)\mathrm{d}x}_{f_1(y)-f_1(x)}$$

$$+ \underbrace{\int_Q (\gamma m' - \gamma m)\mathrm{d}x\mathrm{d}t + \int_0^T \langle A[w'], P\rangle - \langle A[w], P\rangle \mathrm{d}t}_{\langle \nabla f_2(x), y-x\rangle},$$

$$= \tilde{\mathcal{Z}}_{\gamma,P}(m',w') - \tilde{\mathcal{Z}}_{\gamma,P}(m,w)$$

Algorithm

Algorithm 2 Generalized conditional gradient

Choose $(\bar{m}_0, \bar{w}_0) \in \mathcal{C}^{2+\alpha,1+\alpha/2}(Q) \times \mathcal{C}^{1+\alpha,\alpha}(Q;\mathbb{R}^d)$ with $\bar{m}_0(x,t) > 0$ for any $(x,t) \in Q$ and choose a sequence $(\delta_k)_{k\in\mathbb{N}} \in [0,1]$.
for $0 \leq k < N$ **do**
  Find $(m_k, w_k) = \arg\min_{(m,w)\in\tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k),(m,w))$
  Update $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1-\delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$
end for
return $(\bar{m}_N, \bar{w}_N)$.

Using that $h((\bar{m}_k, \bar{w}_k),(m,w)) = \tilde{\mathcal{Z}}_{\gamma_k,P_k}(m,w) - \tilde{\mathcal{Z}}_{\gamma_k,P_k}(\bar{m}_k, \bar{w}_k)$,

Algorithm

Algorithm 3 Generalized conditional gradient

Choose $(\bar{m}_0, \bar{w}_0) \in \mathcal{C}^{2+\alpha,1+\alpha/2}(Q) \times \mathcal{C}^{1+\alpha,\alpha}(Q;\mathbb{R}^d)$ with $\bar{m}_0(x,t) > 0$ for any $(x,t) \in Q$ and choose a sequence $(\delta_k)_{k\in\mathbb{N}} \in [0,1]$.
for $0 \leq k < N$ **do**
  Find $(m_k, w_k) = \arg\min_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma_k,P_k}(m,w)$
  Update $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1-\delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$
end for
return $(\bar{m}_N, \bar{w}_N)$.

Using that $h((\bar{m}_k, \bar{w}_k),(m,w)) = \tilde{\mathcal{Z}}_{\gamma_k,P_k}(m,w) - \tilde{\mathcal{Z}}_{\gamma_k,P_k}(\bar{m}_k, \bar{w}_k)$,

Generalized conditional gradient interpretation

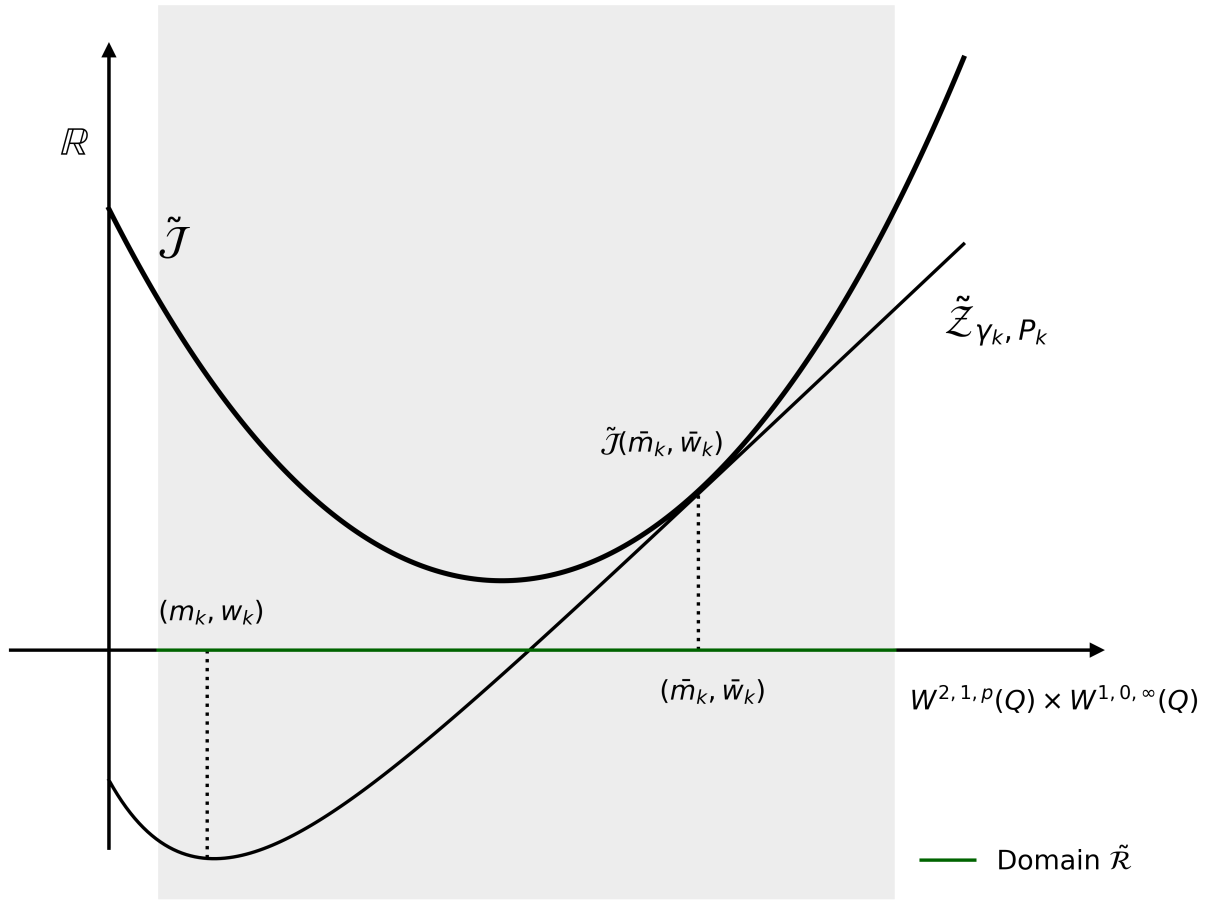


Figure: Illustration of the potential cost $\tilde{\mathcal{J}}$, the individual cost $\tilde{\mathcal{Z}}_{\gamma,P}$ and the exploitability σ.

Game theory interpretation: fictitious play

1. Given $(\bar{m}_k, \bar{w}_k)$ compute $P_k = \phi(A\bar{w}_k)$ and $\gamma_k = f(\bar{m}_k)$.

2. Find u_k solution to,

$$\begin{aligned} -\partial_t u - \Delta u + \boldsymbol{H}[\nabla u + A^\star P_k] &= \gamma_k, \\ u(T) &= g \end{aligned}$$

3. Find the associated optimal control $v_k = -\boldsymbol{H}_p[\nabla u_k + A^\star P_k]$.

4. Find the solution m_k to,

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (v_k m) &= 0, \\ m(0) &= m_0. \end{aligned}$$

5. Compute $w_k = m_k v_k$ and actualize $(\bar{m}_{k+1}, \bar{w}_{k+1}) = (1 - \delta_k)(\bar{m}_k, \bar{w}_k) + \delta_k(m_k, w_k)$.

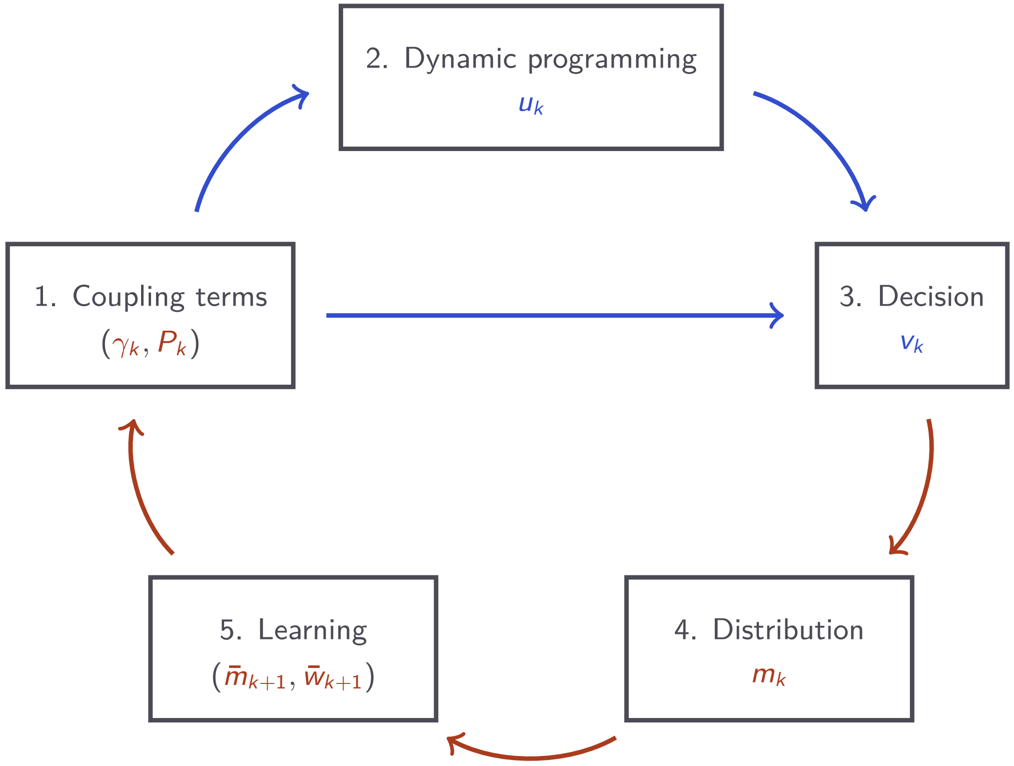


Figure: Fictitious play: a fixed point approach.

Convergence analysis

- **Primal gaps** let $(\bar{m}, \bar{w}) = \arg\min_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{J}}(m, w)$,

$$\epsilon_k = \epsilon(\bar{m}_k, \bar{w}_k) = \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \tilde{\mathcal{J}}(\bar{m}, \bar{w}),$$

- **Exploitability:** largest decrease in cost that a representative agent can reach by playing its best response, assuming that all other agents use the feedback

$$\begin{aligned}\sigma_k = \sigma(\bar{m}_k, \bar{w}_k) = &- \min_{(m,w)\in\tilde{\mathcal{R}}} h((\bar{m}_k, \bar{w}_k), (m, w)),\\ &= \tilde{\mathcal{Z}}_{\gamma_k, P_k}(\bar{m}_k, \bar{w}_k) - \min_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma_k, P_k}(m, w).\end{aligned}$$

Lemma

We have that $\epsilon_k \leq \sigma_k$.

Direct consequence of the GCG framework: visual proof.

Convergence analysis

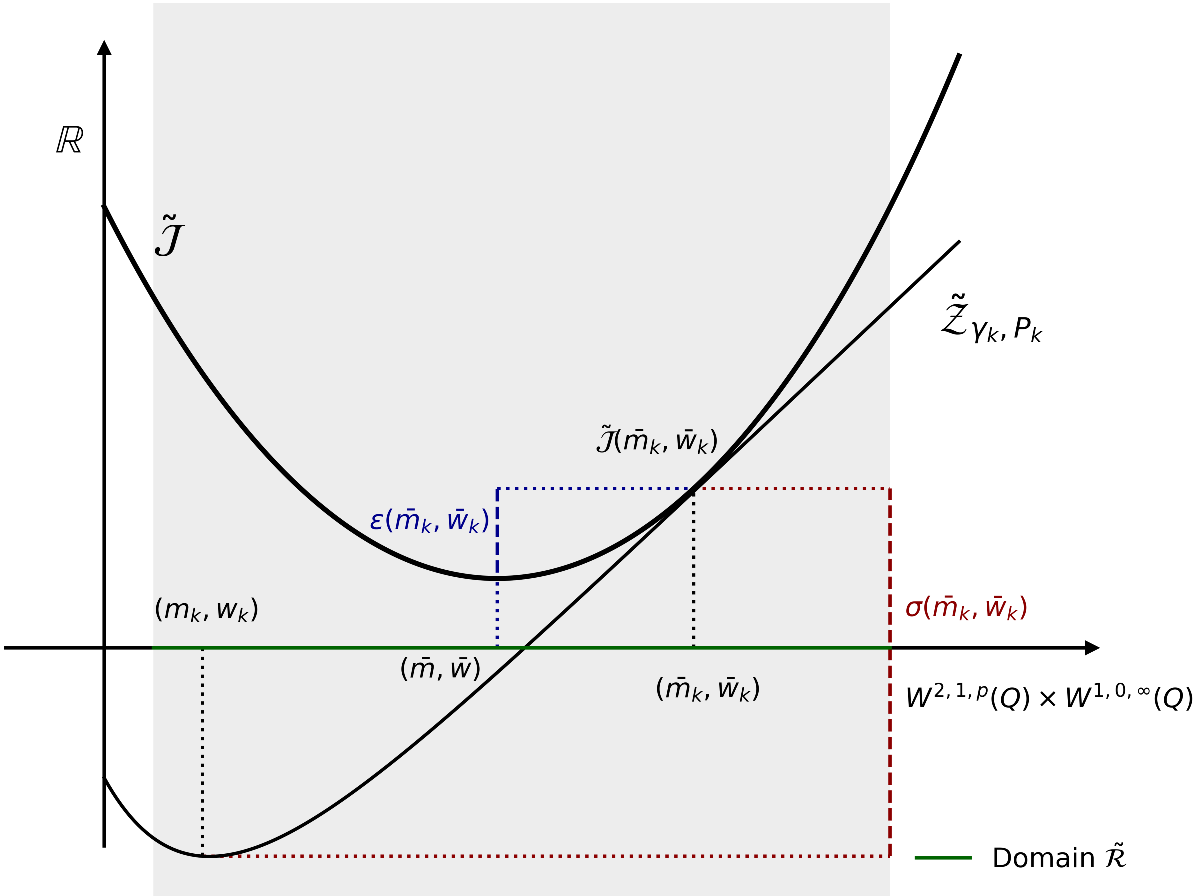


Figure: Illustration of the potential cost $\tilde{\mathcal{J}}$, the individual cost $\tilde{\mathcal{Z}}_{\gamma,P}$ and the exploitability σ.

Main result

Lemma

- *Frank-Wolfe learning rate :* $\delta_k = \frac{2}{k+2}$ *implies* $\epsilon_k \leq \frac{4L_0}{k+2}$, *for some* $L_0 > 0$.
- *Fictitious play learning rate :* $\delta_k = \frac{1}{k+1}$ *implies* $\epsilon_k \leq \frac{\ln(k+1)L_1}{k+1}$, *for some* $L_1 > 0$.

Additional convergence results, based on a quadratic growth property of the potential cost, have been obtained. For any $k \in \mathbb{N}$ we denote

$$\begin{aligned} \delta \bar{m}_k &= \bar{m}_k - \bar{m}, & \delta \bar{w}_k &= \bar{w}_k - \bar{w}, & \delta \bar{v}_k &= \bar{v}_k - \bar{v}, \\ \delta m_k &= m_k - \bar{m}, & \delta w_k &= w_k - \bar{w}, & \delta \bar{v}_k &= v_k - \bar{v}, \\ \delta P_k &= P_k - \bar{P}, & \delta \gamma_k &= \gamma_k - \bar{\gamma}, & \delta u_k &= u_k - \bar{u}. \end{aligned}$$

Theorem

There exists $C > 0$ *such that for all* $k \in \mathbb{N}$, $\sigma_k \leq C\epsilon_k^{1/2}$ *and*

$$\begin{aligned} & \|\delta \bar{v}_k\|_{L^2(Q;\mathbb{R}^d)} &+& \|\delta \bar{m}_k\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} &+& \|\delta \bar{w}_k\|_{L^2(Q;\mathbb{R}^d)} \\ + & \|\delta v_k\|_{L^2(Q;\mathbb{R}^d)} &+& \|\delta m_k\|_{L^\infty(0,T;L^2(\mathbb{T}^d))} &+& \|\delta w_k\|_{L^2(Q;\mathbb{R}^d)} \\ + & \|\delta P_k\|_{L^2(0,T;\mathbb{R}^k)} &+& \|\delta \gamma_k\|_{L^\infty(Q)} &+& \|\delta u_k\|_{L^\infty(Q)} \quad \leq C\epsilon_k^{1/2}. \end{aligned}$$

Contents

Numerical experiments

- 2 examples:
  - **Congestion** model,
  - **Cournot competition** model.

Numerical experiments

- 2 examples:
  - **Congestion** model,
  - **Cournot competition** model.
- 4 learning rules:
  - 2 open loop learning rules:
    1. **Fictitious Play**: $\delta_k = 1/(k+1)$,
    2. **Frank Wolfe**: $\delta_k = 2/(k+2)$,
  - 2 closed loop learning rules:
    1. **Golden-section** rule,
    2. **Armijo** like rule.

Numerical experiments

- 2 examples:
  - **Congestion** model,
  - **Cournot competition** model.
- 4 learning rules:
  - 2 open loop learning rules:
    1. **Fictitious Play**: $\delta_k = 1/(k+1)$,
    2. **Frank Wolfe**: $\delta_k = 2/(k+2)$,
  - 2 closed loop learning rules:
    1. **Golden-section** rule,
    2. **Armijo** like rule.

Objective of the **closed loop** rules: at each step $k \in \mathbb{N}$, find δ_k such that

$$\delta_k = \min_{\delta \in [0,1]} \tilde{\mathcal{J}}(\bar{m}_k^\delta, \bar{w}_k^\delta),$$

where $(\bar{m}_k^\delta, \bar{w}_k^\delta) = \delta(m_k, w_k) + (1-\delta)(\bar{m}_k, \bar{w}_k)$.

Golden-section search

Algorithm 4 Golden-section search

$a = 0$ and $d = 1$.
for $i \leq I \in \mathbb{N}$ **do**
  Set $b = d - (d-a)/\varphi$ and $c = a + (d-a)/\varphi$.
  Find $\delta^i = \arg\min_{\delta \in \{a,b,c,d\}} \tilde{\mathcal{J}}(\bar{m}_k^\delta, \bar{w}_k^\delta)$
  if $\delta^i = a$ **then**
    Set $d = b$
  else if $\delta^i = b$ **then**
    Set $d = c$
  else if $\delta^i = c$ **then**
    Set $a = b$
  else if $\delta^i = d$ **then**
    Set $a = c$
  end if
end for
return δ^i.

where $\varphi := (\sqrt{5}+1)/2$ is the golden number.

Golden-section search

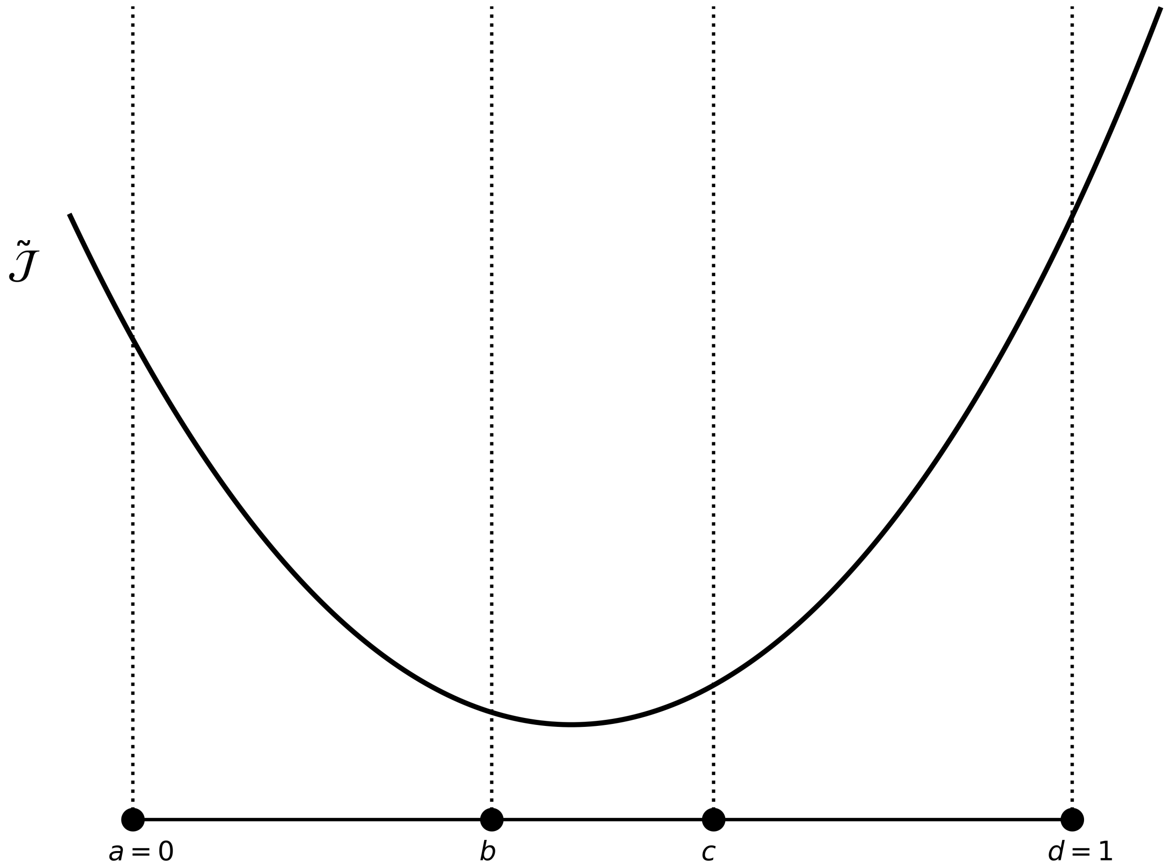


Figure: Illustration: golden section search

Golden-section search

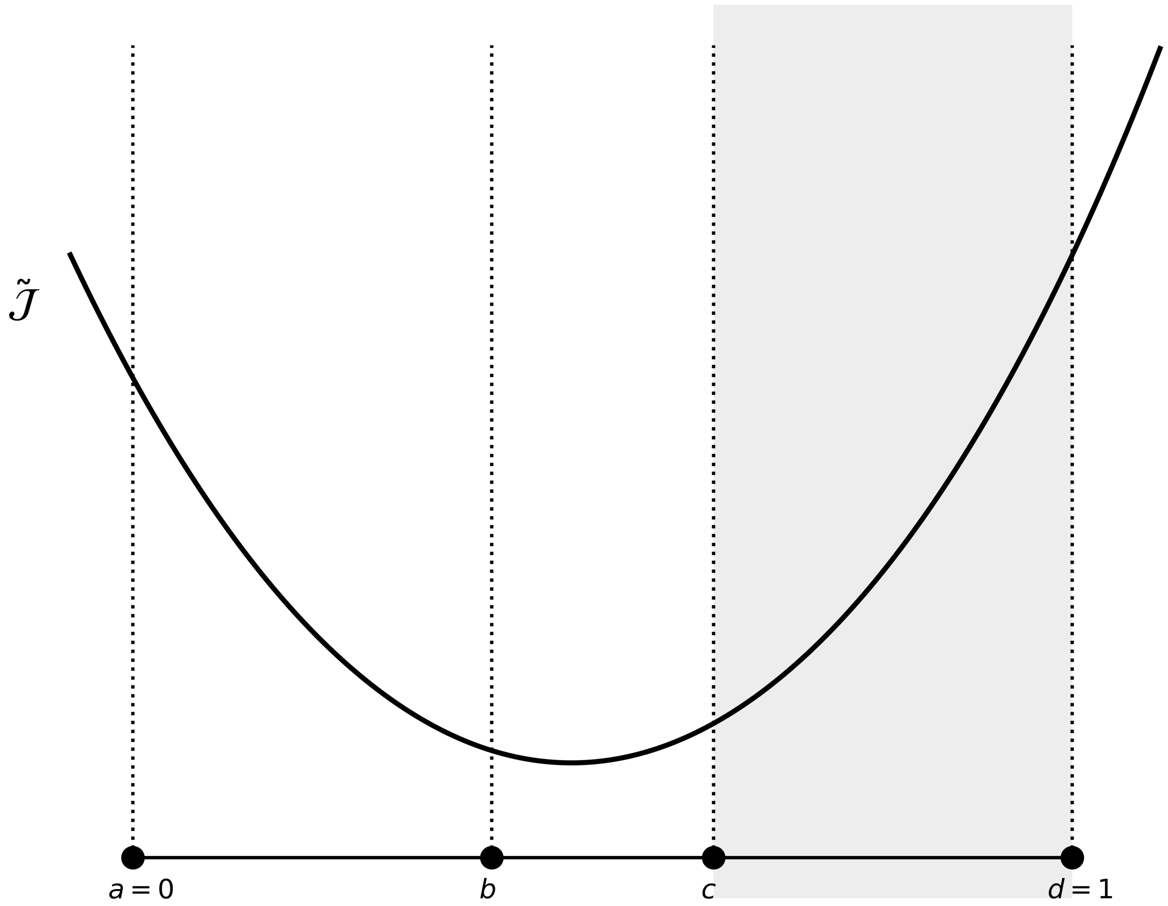


Figure: Illustration: golden section search

Golden-section search

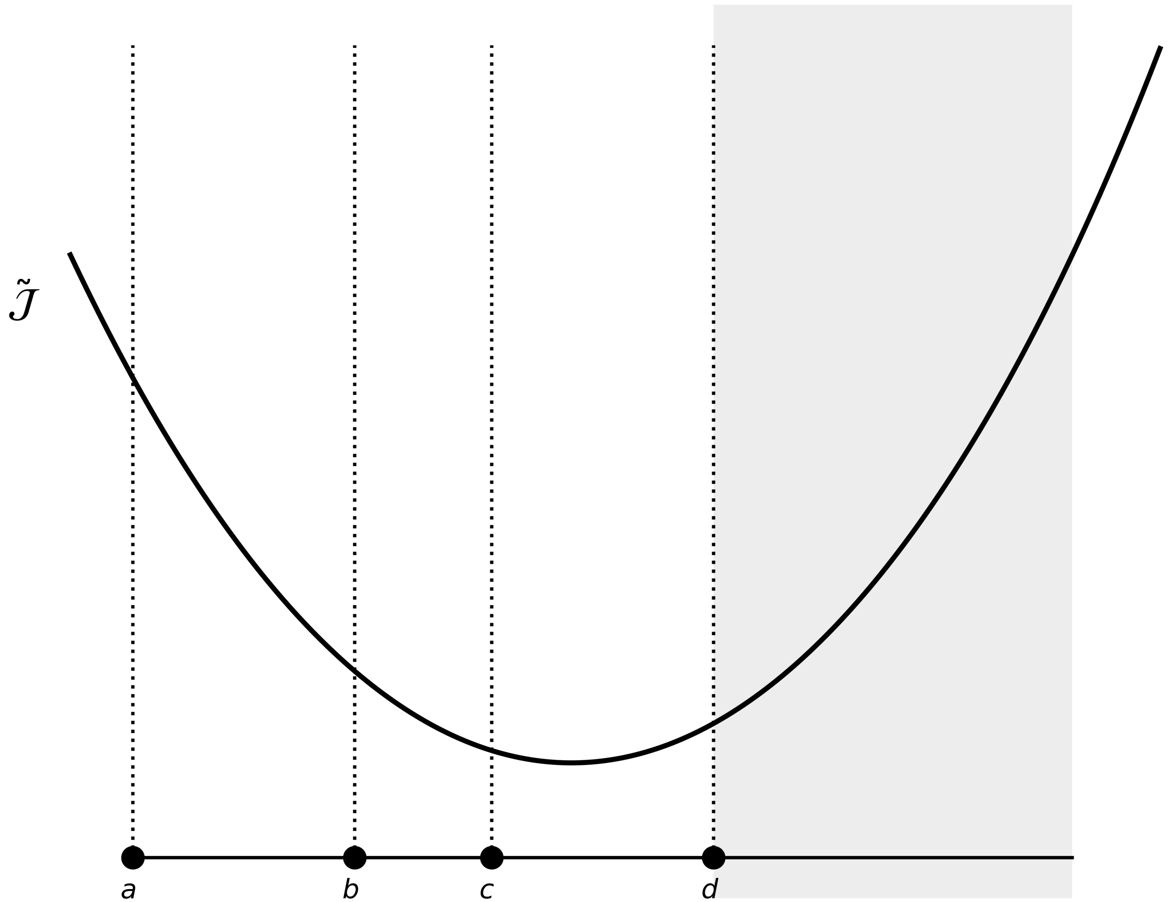


Figure: Illustration: golden section search

Golden-section search

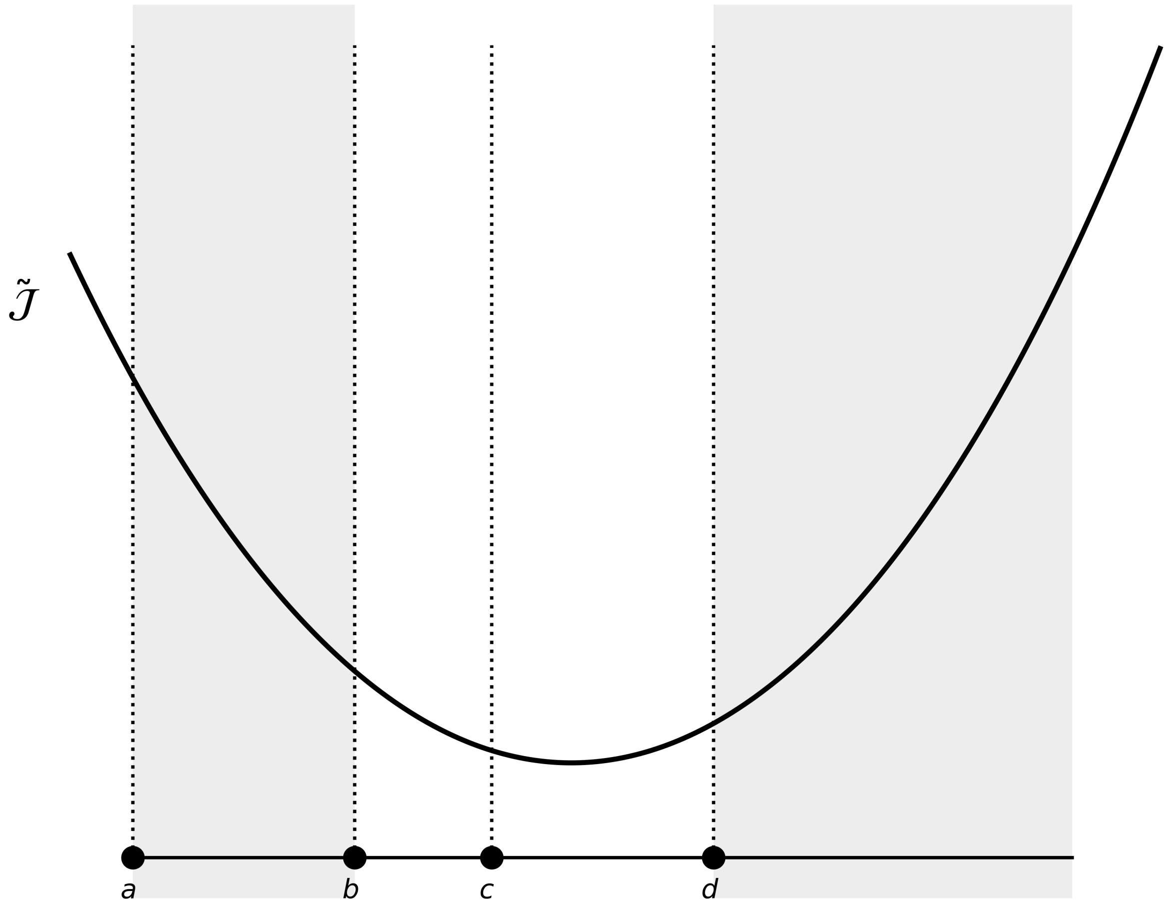


Figure: Illustration: golden section search

Armijo rule

Lemma

There exists $C > 0$ such that for any $\delta \in [0,1]$, it holds:

$$\tilde{\mathcal{J}}(\bar{m}_k^\delta, \bar{w}_k^\delta) \leq \tilde{\mathcal{J}}(\bar{m}_k, \bar{w}_k) - \delta\sigma_k + \delta^2 C,$$

where $(\bar{m}_k^\delta, \bar{w}_k^\delta) = \delta(m_k, w_k) + (1-\delta)(\bar{m}_k, \bar{w}_k)$.

Algorithm 5 Quasi-Armijo-Goldstein

Choose $c \in (0, 1/2]$ and $\tau \in (0,1)$. Initialize $i = 0$ and $\delta^i = 1$.
while $\tilde{\mathcal{J}}(\bar{m}^{\delta^i}, \bar{w}^{\delta^i}) \geq \tilde{\mathcal{J}}(\bar{m}, \bar{w}) - c\delta^i\sigma_k$ **do**
  Update $\delta^{i+1} = \tau\delta^i$
  Update $i = i + 1$
end while
return δ^i.

Numerical experiments

A congestion model ($\phi = 0$):

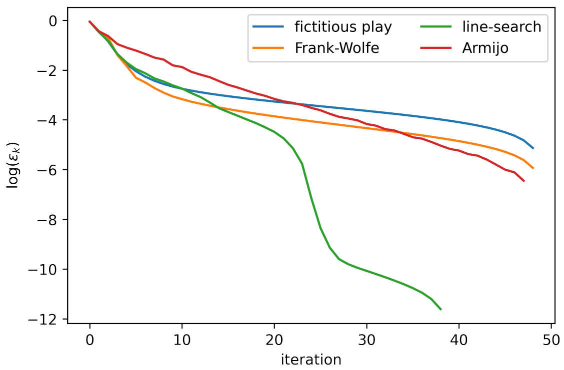


Figure: $\log(\varepsilon_k)$

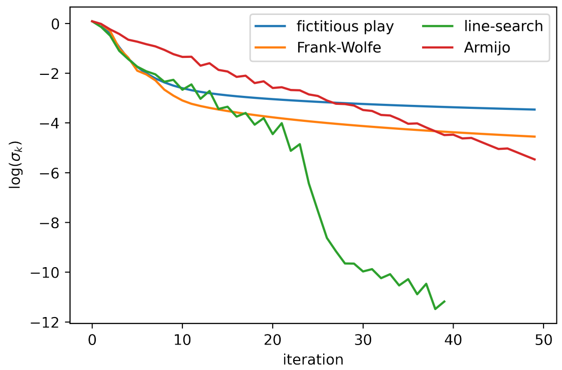


Figure: $\log(\sigma_k)$

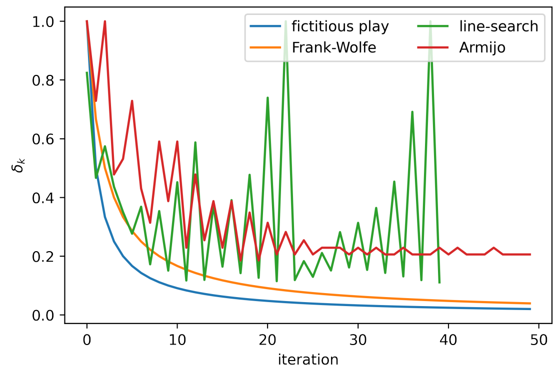


Figure: Learning rate δ_k

A Cournot model ($f = 0$):

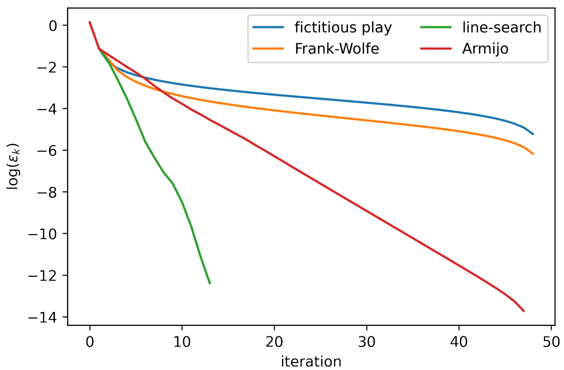


Figure: $\log(\varepsilon_k)$

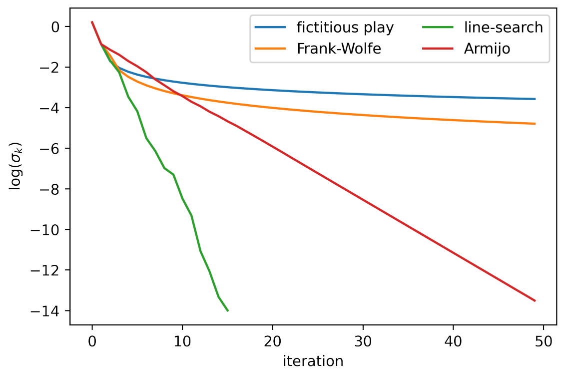


Figure: $\log(\sigma_k)$

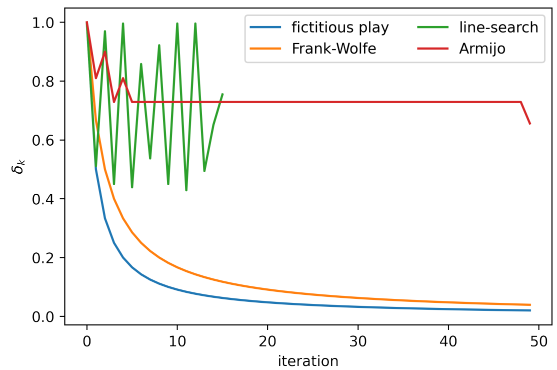


Figure: Learning rate δ_k

Congestion model

Data of the problem :

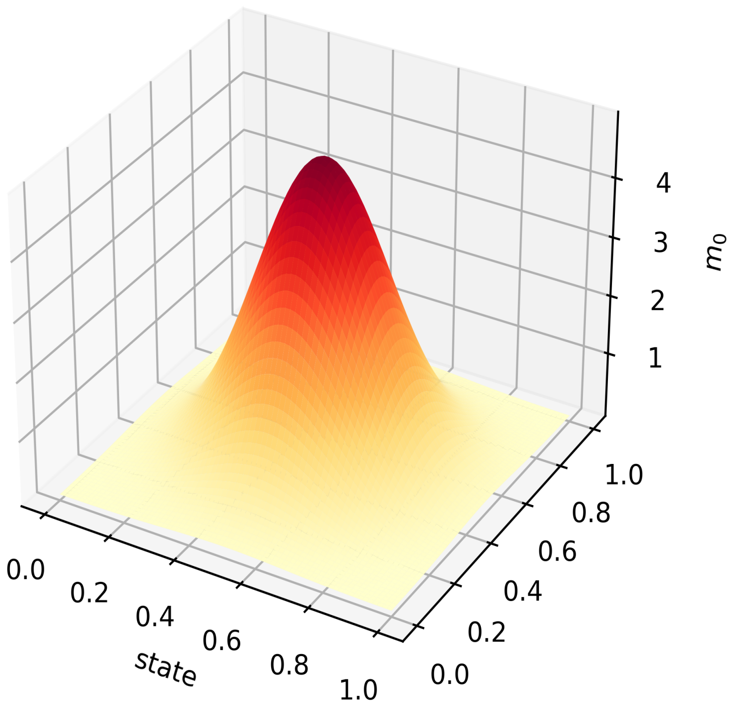


Figure: Initial measure m_0

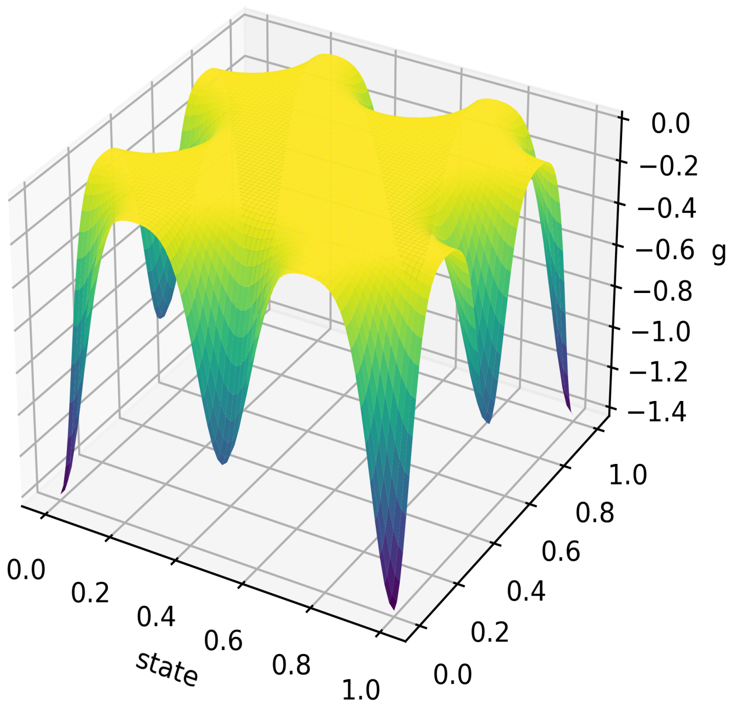
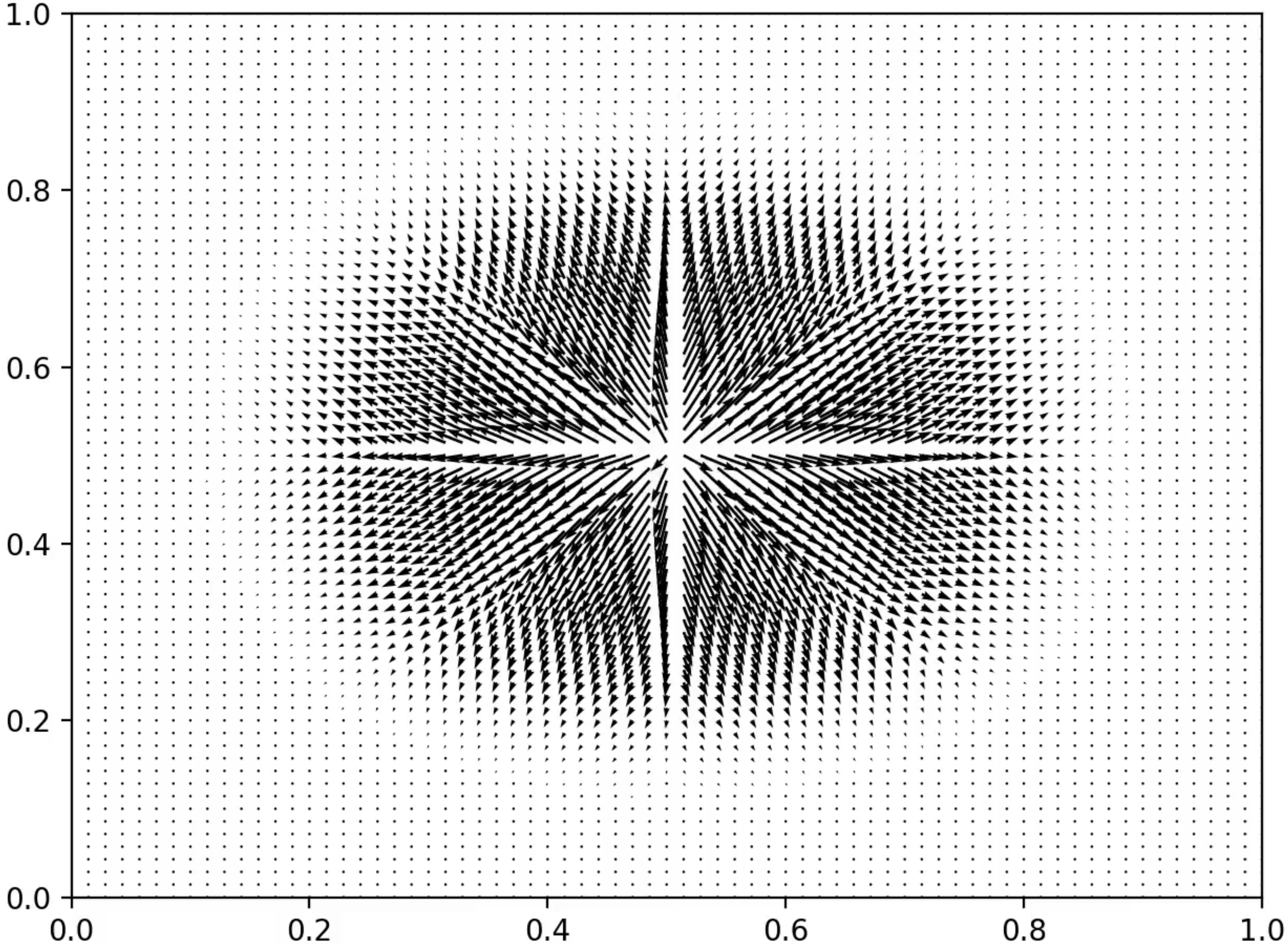


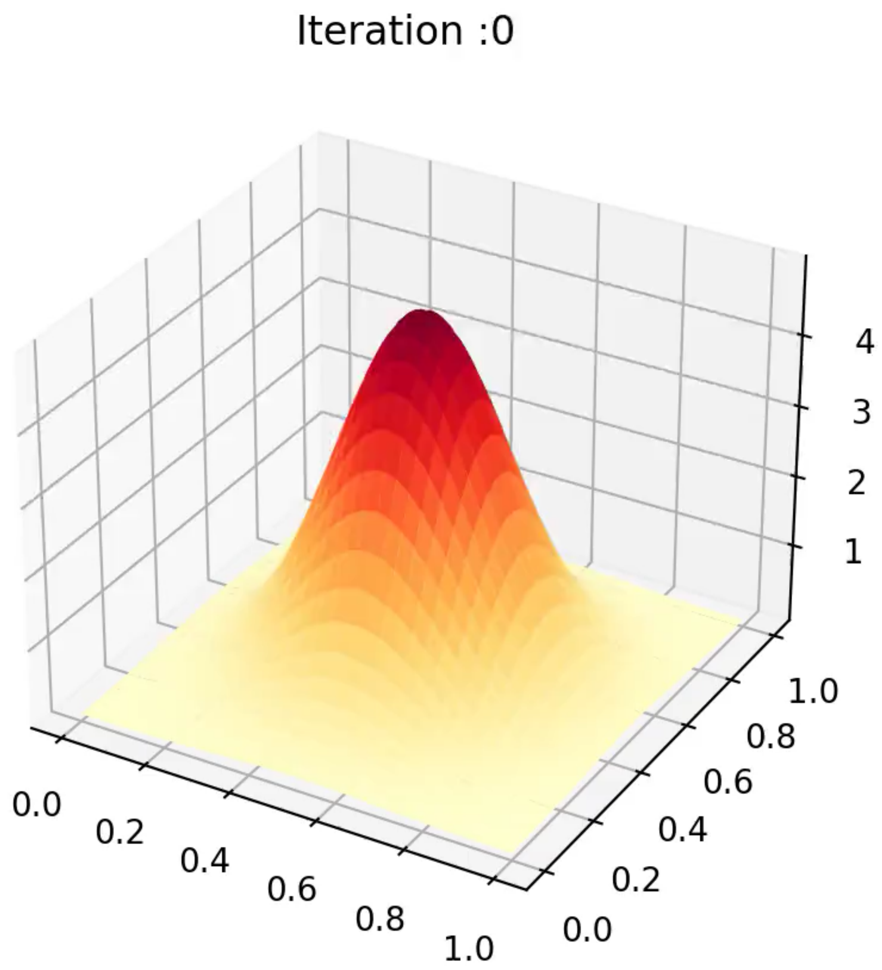
Figure: Terminal condition g

Linear congestion of the form $f(x, t, m(t)) = cm(x, t)$ for any $(x, t) \in Q$.

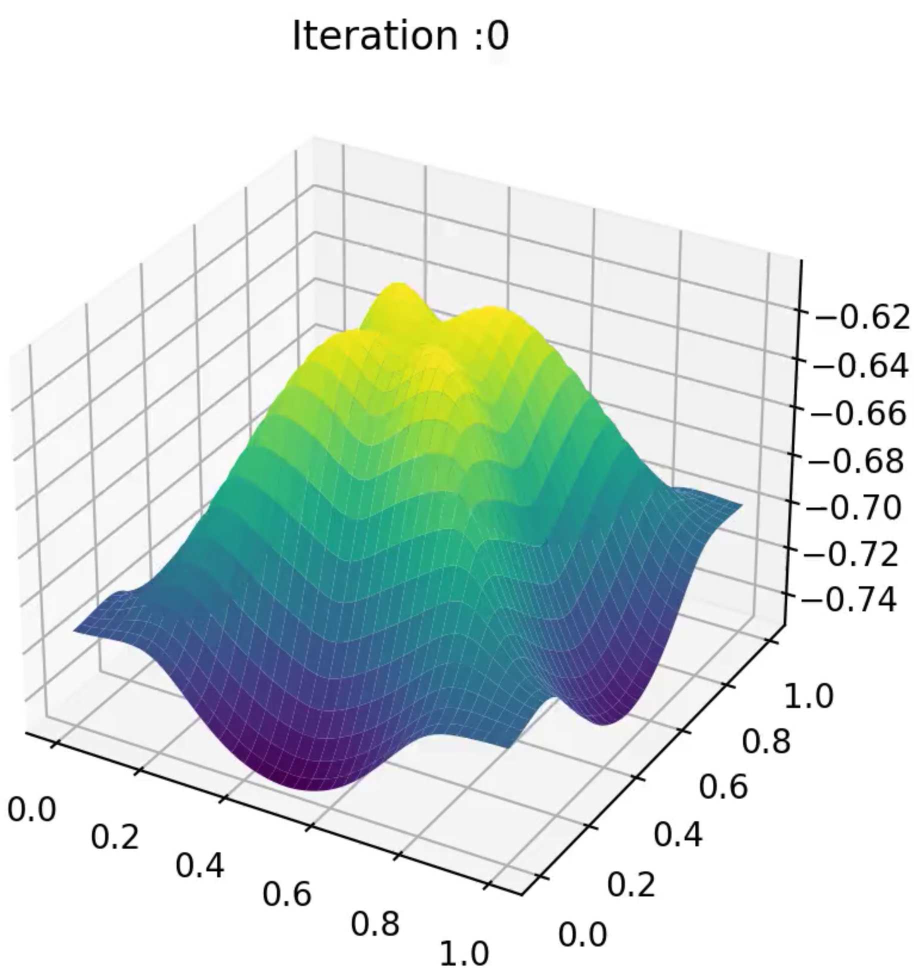
Equilibrium control $\bar{v}$



Equilibrium measure $\bar{m}$



Equilibrium value function $\bar{u}$



Price model

Data of the problem (terminal condition $g = 0$):

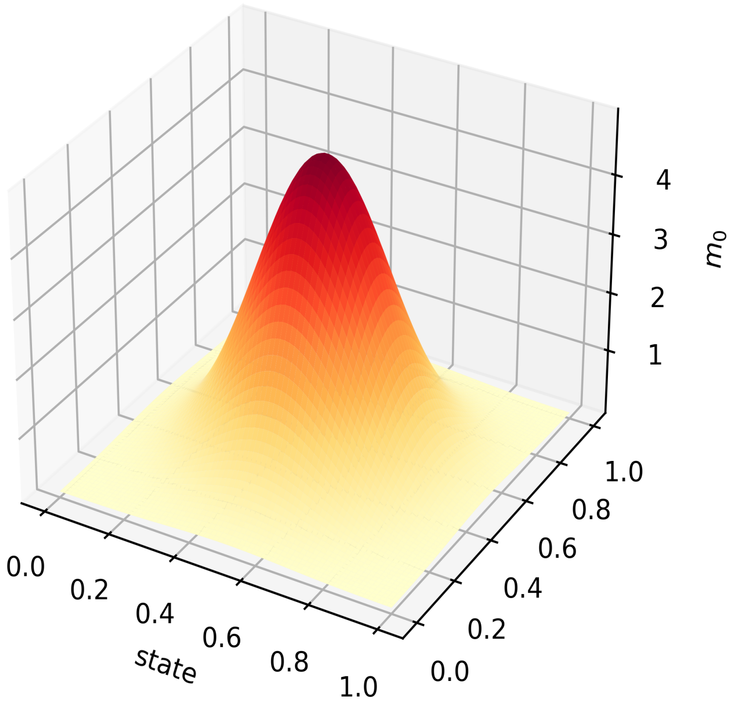


Figure: Initial measure m_0

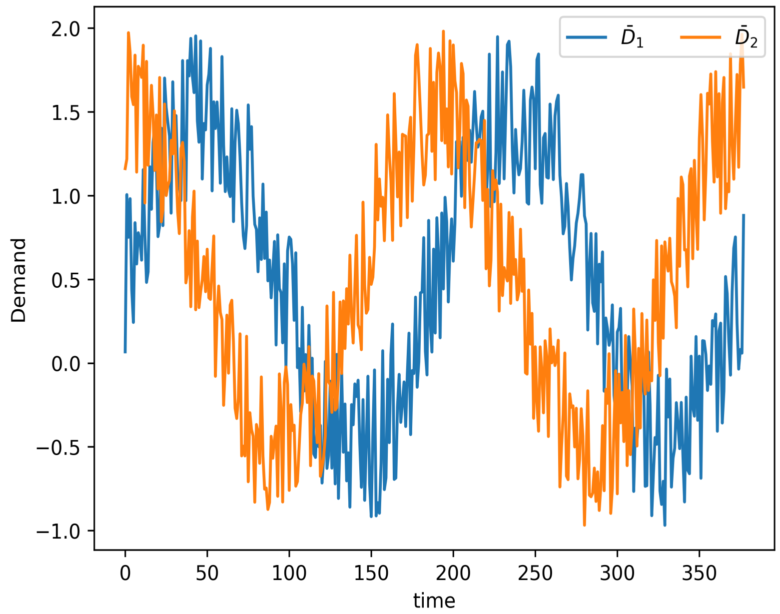


Figure: Reference demand $\bar{D}$

Equilibrium prices

Pricing function $P(t, D(t) + \bar{D}(t)) = c(D(t) + \bar{D}(t))$ for any $t \in [0, T]$.

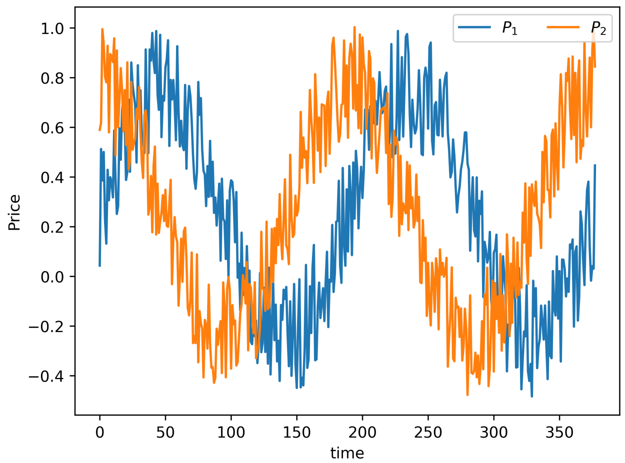


Figure: Equilibrium prices P

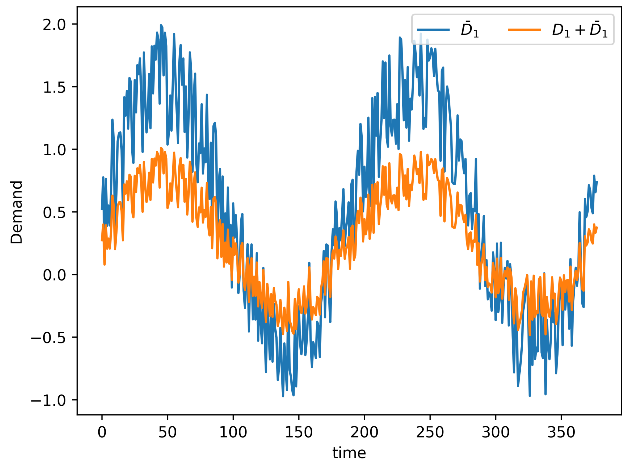


Figure: Remaining demand $D_1 + \bar{D}_1$

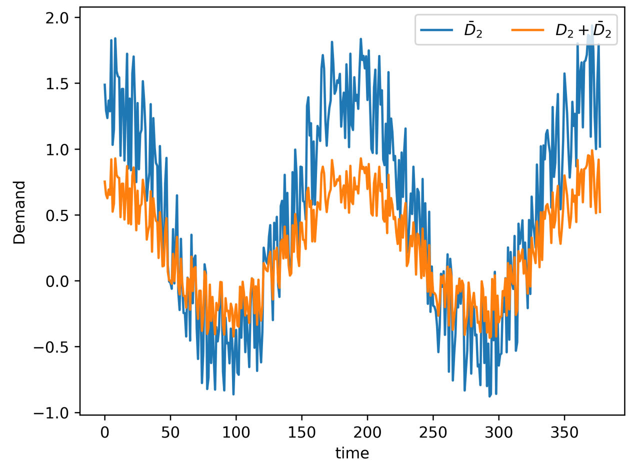
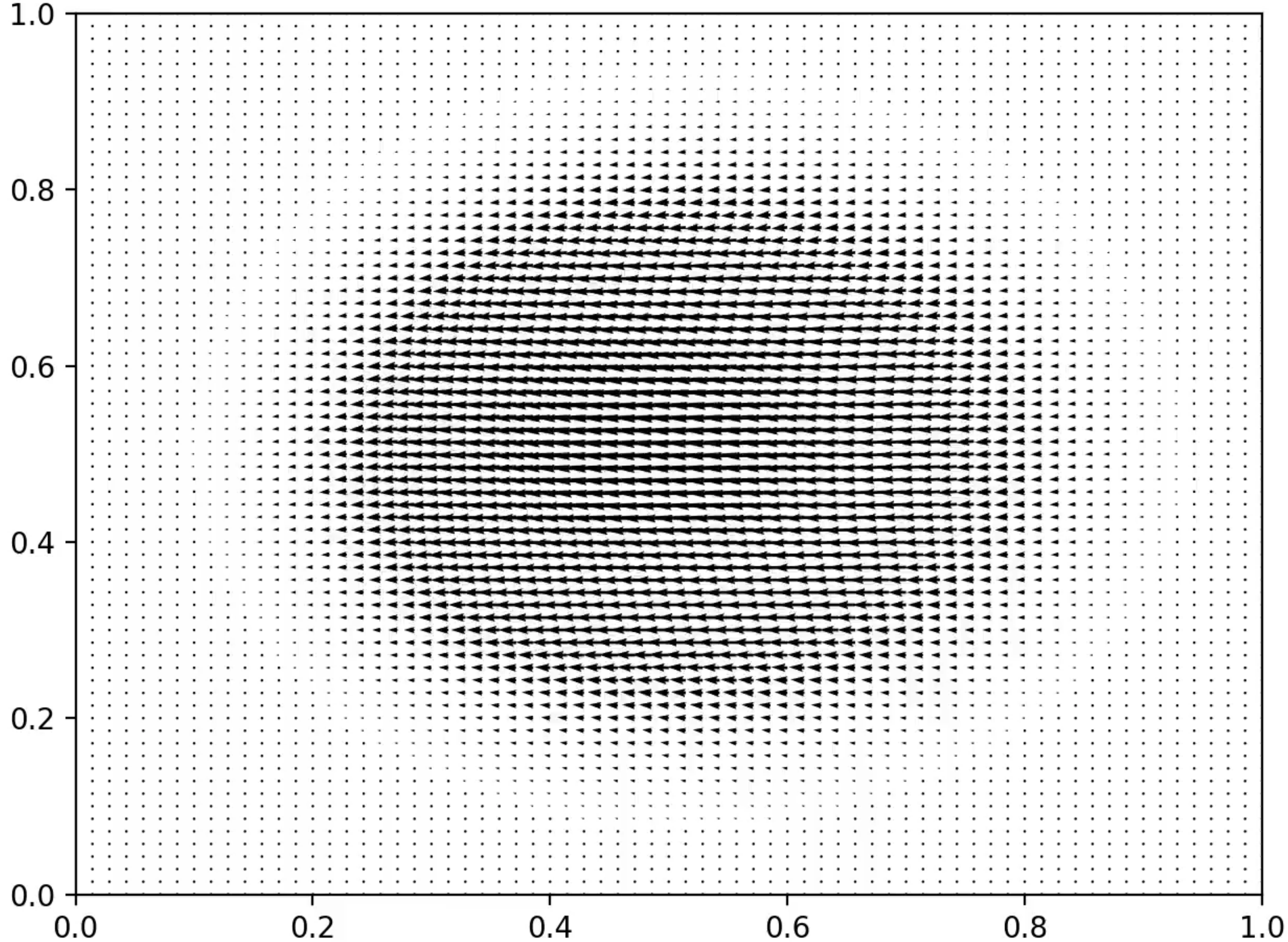
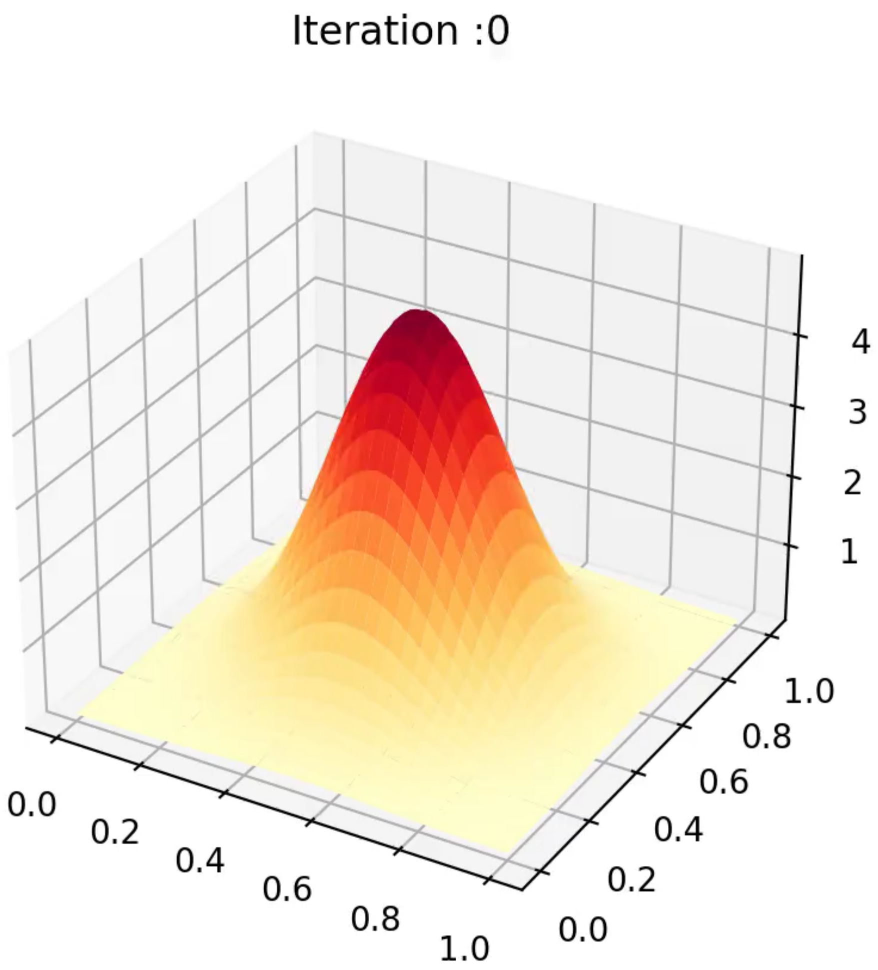


Figure: Remaining demand $D_2 + \bar{D}_2$

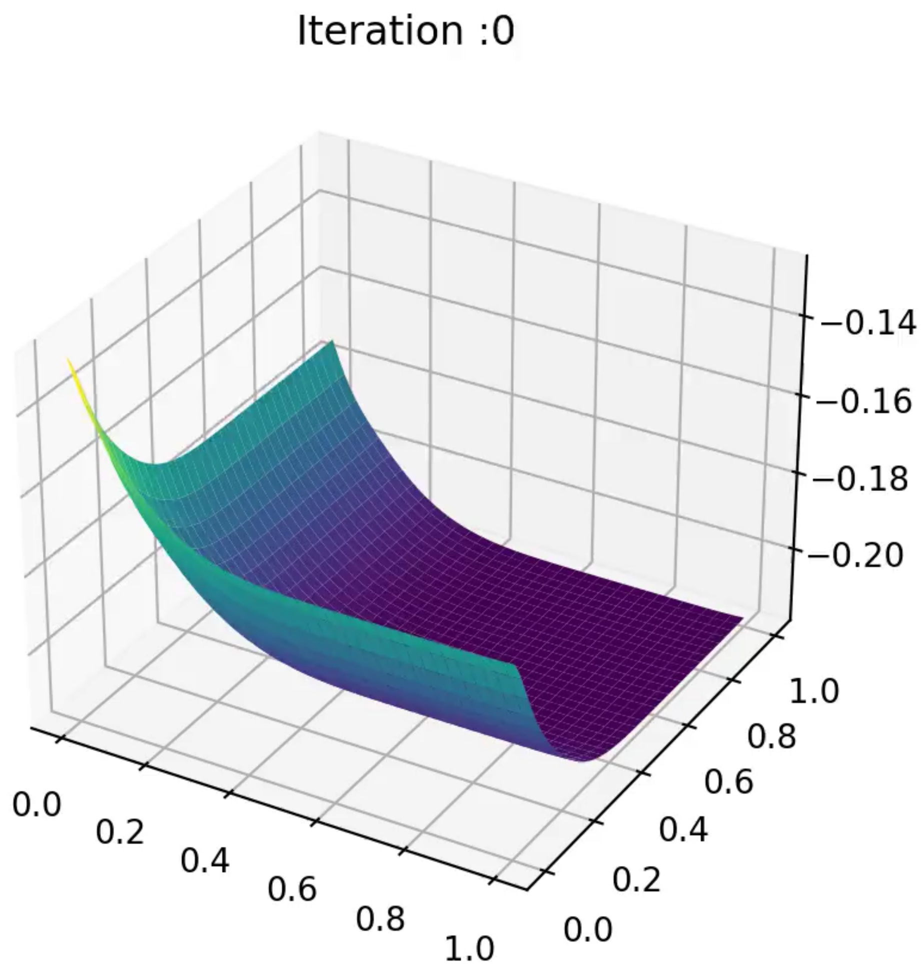
Equilibrium control $\bar{v}$



Equilibrium measure $\bar{m}$



Equilibrium value function $\bar{u}$



Contents

Summary

- GCG algorithm **applies** to potential MFGs,
- It has a game theory **interpretation**,
- **Convergence** in $O(1/k)$ of the primal gaps and $O(1/\sqrt{k})$ of the exploitability and the variables of the problem for $\delta_k = 2/(k+2)$,
- Possible accelerations via line search.

Perspectives

Line-search Investigate different line-search rules and compare their performances,

First order Apply GCG to first order MFG,

Non-convex Application to non-convex case (MFG “à la Cucker-Smale”).

Conclusion

Thank you for your attention.

References I

J. Frédéric Bonnans, Saeed Hadikhanloo, and Laurent Pfeiffer, *Schauder estimates for a class of potential mean field games of controls*, Applied Mathematics & Optimization **83** (2021), 1431–1464.

J. Frédéric Bonnans, Pierre Lavigne, and Laurent Pfeiffer, *Generalized conditional gradient and learning in potential mean field games*, 2021.

Matthieu Geist, Julien Pérolat, Mathieu Laurière, Romuald Elie, Sarah Perrin, Olivier Bachem, Rémi Munos, and Olivier Pietquin, *Concave utility reinforcement learning: the mean-field game viewpoint*, arXiv preprint arXiv:2106.03787 (2021).

Minyi Huang, Roland P. Malhamé, and Peter E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Communications in Information & Systems **6** (2006), no. 3, 221–252.

Martin Jaggi, *Revisiting Frank-Wolfe: Projection-free sparse convex optimization*, International Conference on Machine Learning, PMLR, 2013, pp. 427–435.

References II

Jean-Michel Lasry and Pierre-Louis Lions, *Jeux à champ moyen. i–le cas stationnaire*, Comptes Rendus Mathématique **343** (2006), no. 9, 619–625.

The stochastic individual control problem

For all $\nu \in L^2_{\mathbb{F}}(t, T; \mathbb{R}^d)$, we denote by $(X^\nu_s)_{s\in[0,T]}$ the solution to the stochastic differential equation

$$\mathrm{d}X_s = \nu_s \mathrm{d}s + \sqrt{2}\mathrm{d}B_s, \quad X_0 = Y.$$

We define the individual cost $Z_{\gamma,P} \colon L^2_{\mathbb{F}}(0, T; \mathbb{R}^d) \to \mathbb{R}$,

$$Z_{\gamma,P}(\nu) = \mathbb{E}\left[\int_0^T L(X^\nu_s, s, \nu_s) + \langle A^\star[P](X^\nu_s, s), \nu_s\rangle + \gamma(X^\nu_s, s)\mathrm{d}s + g(X^\nu_T)\right].$$

We consider the following stochastic individual control problem

$$\inf_{\nu\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^d)} Z_{\gamma,P}(\nu). \tag{$\mathrm{P}_{\gamma,P}$}$$

An equivalent PDE individual control problem

Consider the mapping m : $W^{1,0,\infty}(Q) \to W^{2,1,p}(Q)$ which associates to any $v \in W^{1,0,\infty}(Q)$ the solution to the Fokker-Planck equation

$$\begin{aligned} \partial_t m - \Delta m + \nabla \cdot (vm) = &\quad 0, && (x,t) \in Q, \\ m(x,0) = &\quad m_0(x), && x \in \mathbb{T}^d. \end{aligned}$$

We define $\mathcal{B}^p = W^{2,1,p}(Q) \times W^{1,0,\infty}(Q)$ (recall that $p > d+2$ is fixed) and we define

$$\mathcal{R} = \left\{(m,v) \in \mathcal{B}^p,\ \partial_t m - \Delta m + \nabla \cdot (vm) = 0,\ m(0) = m_0,\ (x,t) \in Q\right\},$$

$$\tilde{\mathcal{R}} = \left\{(m,w) \in \mathcal{B}^p,\ \partial_t m - \Delta m + \nabla \cdot w = 0,\ m(0) = m_0,\ m(x,t) > 0,\ (x,t) \in Q\right\}.$$

An equivalent PDE individual control problem

We define the individual cost $\mathcal{Z}_{\gamma,P} \colon \mathcal{R} \to \mathbb{R}$,

$$\mathcal{Z}_{\gamma,P}(m,v) = \int_Q (\boldsymbol{L}[v] + \gamma)\, m \mathrm{d}x\mathrm{d}t + \int_0^T \langle A[mv], P\rangle \mathrm{d}t + \int_{\mathbb{T}^d} gm(T)\mathrm{d}x.$$

We define the following individual control problem

$$\inf_{(m,v)\in\mathcal{R}} \mathcal{Z}_{\gamma,P}(m,v). \tag{$\mathcal{P}_{\gamma,P}$}$$

We define the individual cost $\tilde{\mathcal{Z}}_{\gamma,P} \colon \tilde{\mathcal{R}} \to \mathbb{R}$,

$$\tilde{\mathcal{Z}}_{\gamma,P}(m,w) = \int_Q \left(\tilde{\boldsymbol{L}}[m,w] + \gamma m\right) \mathrm{d}x\mathrm{d}t + \int_0^T \langle A[w], P\rangle \mathrm{d}t + \int_{\mathbb{T}^d} gm(T)\mathrm{d}x,$$

where $\tilde{\boldsymbol{L}}$ is the perspective function of $\boldsymbol{L}$. and the following control problem

$$\inf_{(m,w)\in\tilde{\mathcal{R}}} \tilde{\mathcal{Z}}_{\gamma,P}(m,w). \tag{$\tilde{\mathcal{P}}_{\gamma,P}$}$$

An equivalent PDE individual control problem

Given $v \in W^{1,0,\infty}(Q)$, we denote $(X_s^v)_{s\in[0,T]}$ the solution to the following stochastic differential equation

$$\mathrm{d}X_s = v(X_s, s)\mathrm{d}s + \sqrt{2}\mathrm{d}B_s, \quad X_0 = Y.$$

We further consider the associated control $\nu_s^v \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^d)$ defined by $\nu_s^v = v(s, X_s^v)$.

Lemma

For any $v \in W^{1,0,\infty}(Q,\mathbb{R}^d)$, we have

$$Z_{\gamma,P}(\nu^v) = \mathcal{Z}_{\gamma,P}(\mathrm{m}[v], v) = \tilde{\mathcal{Z}}_{\gamma,P} \circ \chi(\mathrm{m}[v], v).$$

An equivalent PDE individual control problem

Lemma

Let $u = \mathrm{u}[\gamma, P]$ *and let* $v = -\boldsymbol{H}_p[\nabla u + A^\star P]$. *Let* $m = \mathrm{m}[v]$ *and let* $(m, w) = \chi(m, v)$.

1. *There exists* $\alpha \in (0,1)$, *depending on* γ *and* P, *such that*
$$v \in \mathcal{C}^{1+\alpha,\alpha}(Q;\mathbb{R}^d), \quad m \in \mathcal{C}^{2+\alpha,1+\alpha/2}(Q), \quad w \in \mathcal{C}^{1+\alpha,\alpha}(Q;\mathbb{R}^d).$$
2. *There exists* $C > 0$, *depending only on* R, *such that*
$$\|v\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C, \quad \|m\|_{W^{2,1,p}(Q)} \leq C, \quad \|w\|_{W^{1,0,\infty}(Q;\mathbb{R}^d)} \leq C.$$
3. *The stochastic process* $(\nu_s^v)_{s\in[0,T]}$ *is a solution to* (**??**).
4. *The pair* (m, v) *is a solution to* $(\mathcal{P}_{\gamma,P})$ *and* (m, w) *is a solution to* $(\tilde{\mathcal{P}}_{\gamma,P})$.

Game theory interpretation: exploitability

We denote $(X_s^v)_{s\in[0,T]}$ the solution to the following stochastic differential equation

$$\mathrm{d}X_s = v(X_s, s)\mathrm{d}s + \sqrt{2}\mathrm{d}B_s, \quad X_0 = Y.$$

We further consider the associated control $\nu_s^v \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^d)$ defined by $\nu_s^v = v(s, X_s^v)$. Defining the individual stochastic control problem,

$$\inf_{\nu\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^d)} Z_{\gamma,P}(\nu) = \mathbb{E}\left[\int_0^T L(X_s^\nu, s, \nu_s) + \langle A^\star[P](X_s^\nu, s), \nu_s\rangle + \gamma(X_s^\nu, s)\mathrm{d}s + g(X_T^\nu)\right],$$

we have that the **primal-dual gap** coincides with the notion of **exploitability** for $\bar{v}_k = \bar{w}_k/\bar{m}_k$,

$$\boxed{\sigma_k = Z_{\gamma_k,P_k}(\nu^{\bar{v}_k}) - \inf_{\nu\in L^2_{\mathbb{F}}(t,T;\mathbb{R}^d)} Z_{\gamma_k,P_k}(\nu).}$$

Exploitability: largest decrease in cost that a representative agent can reach by playing its best response, assuming that all other agents use the feedback $\bar{v}_k = \bar{w}_k/\bar{m}_k$.