Mean field limit of controlled system:
Optimal control and non-smooth dynamics

Nicolas Gast

EPFL LIG and INRIA

November, 25

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Introduction : mean field interacting objects

The term "mean field" applies for system of interacting objects :

- communication networks, computing clusters.
- epidemic models, gossip.
- chemical reactions

Introduction : mean field interacting objects

The term "mean field" applies for system of interacting objects :

- communication networks, computing clusters.
- epidemic models, gossip.
- chemical reactions

Objectives

Analyze and improve the performance of the system :

- Characterize the dynamics of the system
- Find good (or optimal) policies to control the system.

Problem

- Systems composed by N interacting objects.
- N is large.

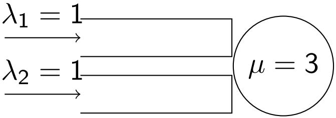
Number of states needed to represent the model explodes.

$\rightarrow$ Study the system when N grows.

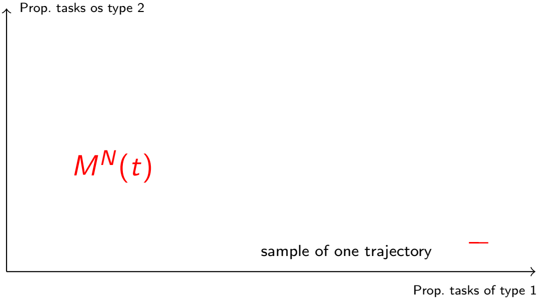
Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.



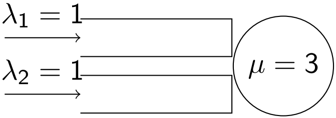
1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).



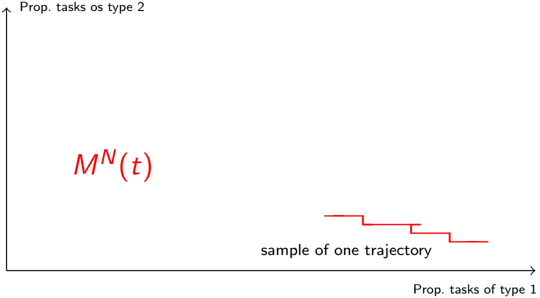
Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.



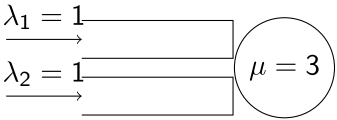
1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).



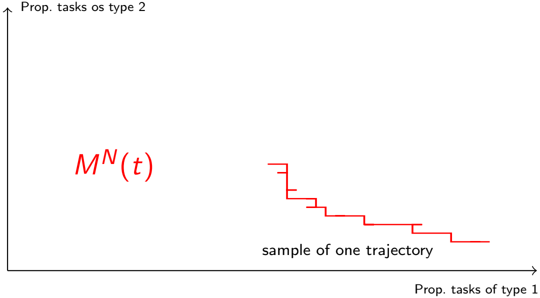
Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.



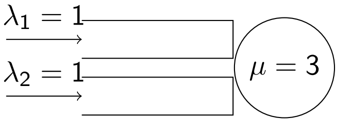
1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).



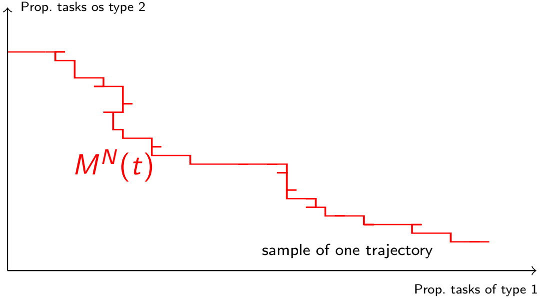
Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.



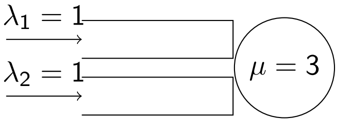
1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).



Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.

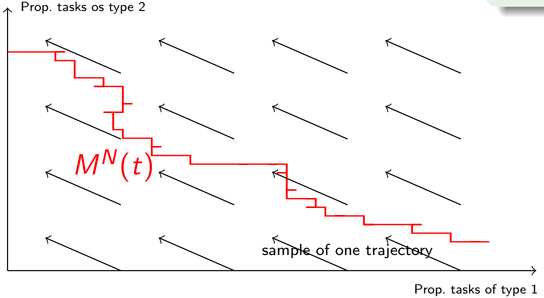


1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).

2. Drift : average difference of $M^N(t)$ between t and $t + dt$. :

- $(+1, 0)$ at rate 1 (arrival 1)
- $(0, +1)$ at rate 1 (arrival 2)
- $(-1, 0)$ at rate 3 (departure 1)

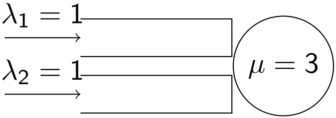
Drift : $f(M^N(t)) = (-2, +1)$.



Mean field approximation on a simple example

Example : static priority.

- N = number of tasks 1 at $t = 0$
- Tasks of type 1 have full priority on the tasks of type 2.

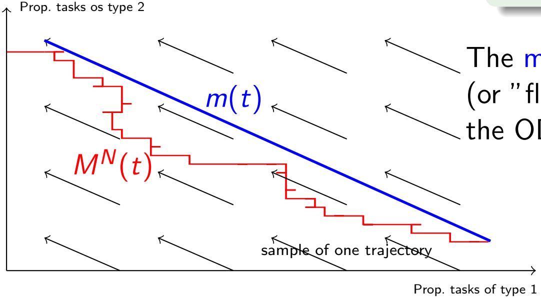


1. Objects are exchangeable :
The important quantity is the empirical measure $M^N(t)$ (*i.e.* $M_i^N(t)$ is the number of objects in state i).

2. Drift : average difference of $M^N(t)$ between t and $t + dt$. :

- $(+1, 0)$ at rate 1 (arrival 1)
- $(0, +1)$ at rate 1 (arrival 2)
- $(-1, 0)$ at rate 3 (departure 1)

Drift : $f(M^N(t)) = (-2, +1)$.



The mean field approximation $m(t)$ (or "fluid limit") is the solution of the ODE $\frac{dm}{dt} = f(m)$.

Goal : study the link between $M^N(t)$ and $m(t)$?

Mean field for performance evaluation.

Since Kurtz (70), lots of work study the approximation of $M^N(t)$ by N.

Looking for convergence results.

- Generic models (Kurtz 70, Le Boudec, Benaïm 08)
  - If f is lipschitz-continuous, then $M^N(t)$ converges to the ODE, $m(t)$.
  - Steady state behavior.
- Propagation du chaos (Snitzman 91, Graham 00) :
  - Objects are asymptotically independent.

Applications in many areas. For example :

- Chemical reactions (Gillespie 77).
- Load balancing (Mitzenmacher 98).
- TCP RED (Baccelli et al. 02).
- 802.11 (Bianchi 00, Bordenave et al. 05)

Mean field for performance evaluation.

Since Kurtz (70), lots of work study the approximation of $M^N(t)$ by N.

Looking for convergence results.

- Generic models (Kurtz 70, Le Boudec, Benaïm 08)
  - If f is lipschitz-continuous, then $M^N(t)$ converges to the ODE, $m(t)$.
  - Steady state behavior.
- Propagation du chaos (Snitzman 91, Graham 00) :
  - Objects are asymptotically independent.

Applications in many areas. For example :

- Chemical reactions (Gillespie 77).
- Load balancing (Mitzenmacher 98).
- TCP RED (Baccelli et al. 02).
- 802.11 (Bianchi 00, Bordenave et al. 05)

In this talk :
extensions to controlled system.

- Optimal control problems.
- Non-smooth dynamics.

Outline

1. Introduction

2. Optimal control of mean field models

- How to define the limiting deterministic optimization problem?
- Convergence results.

3. Application to a brokering problem and speed of convergence.

- When the approximation becomes valid?
- What is the quality of the approximation?

4. Going further: non-smooth dynamics

- if the drift f is not continuous: can we define an ODE $\frac{dm}{dt} = f(m)$?
- What is the limit of M^N?

5. Conclusion

Outline

1. Introduction

2. Optimal control of mean field models
   - How to define the limiting deterministic optimization problem?
   - Convergence results.

3. Application to a brokering problem and speed of convergence.
   - When the approximation becomes valid?
   - What is the quality of the approximation?

4. Going further: non-smooth dynamics
   - if the drift f is not continuous: can we define an ODE $\frac{dm}{dt} = f(m)$?
   - What is the limit of M^N?

5. Conclusion

High dimension stochastic optimization problem

- At time t, a controller chooses an action $a_t \in \mathcal{A}$.
- Goal of the controller : find a policy $\pi : X \to \mathcal{A}$ that minimizes the response time.

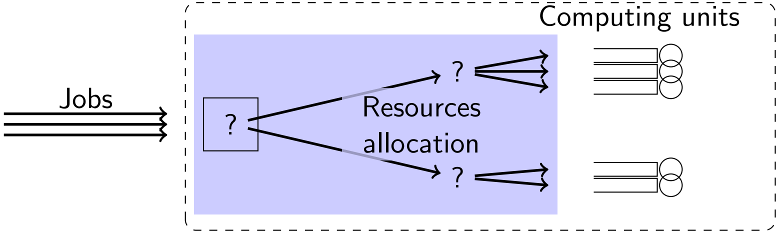


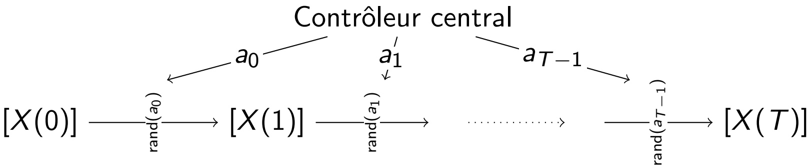
Figure: Example of optimal control

Problem :

- Many resources and applications.
- $\rightarrow$ State space of the system is huge.

A markov decision problem

The theoretical framework is well-known (MDP).

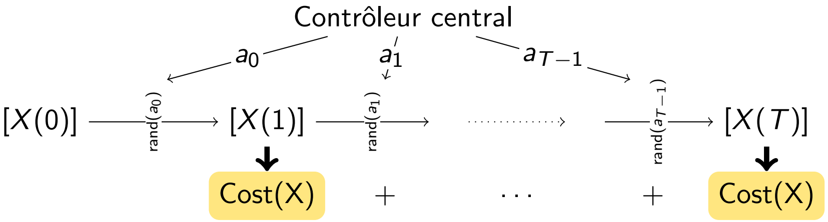


State of the system at time t is $X(t)$

1. A centralized controller modifies the dynamics of the system.

A markov decision problem

The theoretical framework is well-known (MDP).

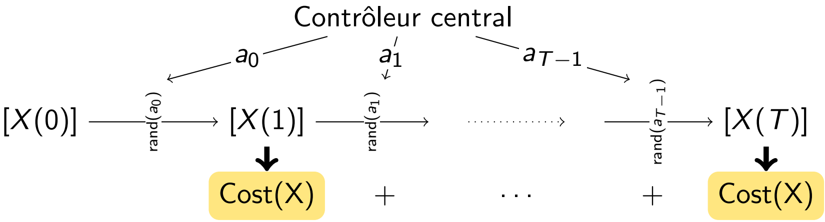


State of the system at time t is $X(t)$

1. A centralized controller modifies the dynamics of the system.
2. At state t is associated a cost $\mathrm{cost}(X(t))$.

A markov decision problem

The theoretical framework is well-known (MDP).



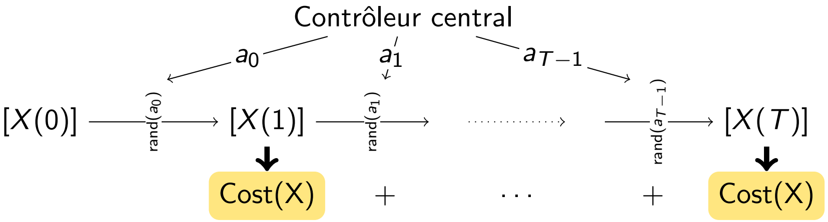
State of the system at time t is $X(t)$

1. A centralized controller modifies the dynamics of the system.
2. At state t is associated a cost $\text{cost}(X(t))$.

Goal of the controller : find the best policy $\pi^* : X \to \mathcal{A}$ to minimize the average cost (over a finite horizon of time).

A markov decision problem

The theoretical framework is well-known (MDP).



State of the system at time t is $X(t)$

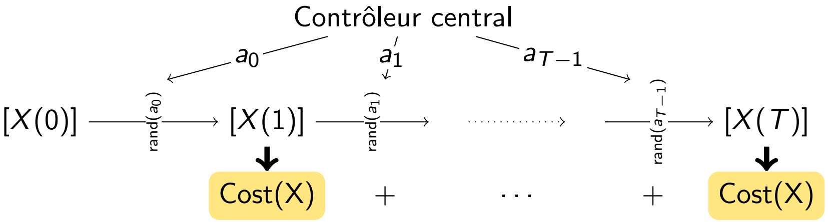
1. A centralized controller modifies the dynamics of the system.
2. At state t is associated a cost $\mathrm{cost}(X(t))$.

Goal of the controller : find the best policy $\pi^* : X \to \mathcal{A}$ to minimize the average cost (over a finite horizon of time).

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathrm{cost}(X_\pi(t))\right]$$

A markov decision problem

The theoretical framework is well-known (MDP).



State of the system at time t is $X(t)$

1. A centralized controller modifies the dynamics of the system.
2. At state t is associated a cost $\text{cost}(X(t))$.

Goal of the controller : find the best policy $\pi^* : X \to \mathcal{A}$ to minimize the average cost (over a finite horizon of time).

$$V_*^N(X(0)) = \inf_{\pi} \quad \mathbb{E}\left[\sum_{t=1}^{T} \text{cost}(X_\pi(t))\right]$$

Mean field controlled system

A mean field controlled system is described by :

- Symmetric system with N objects.
- Intensity I_N and drift $f(a, \cdot)$

$+$

- A controller : chooses an action $a \in \mathcal{A}$.
- Cost function : $\mathrm{cost}(m)$.

If $M^N(t)$ denotes the proportion of objects in each state when applying a sequence of action A. We know that :

$$M_A^N(t) \to \infty m_a(t)$$

where m is a deterministic system :

- in discrete time if $I_N = O(1)$.
- in continuous time if $I_N = o(1)$.

The limiting deterministic optimal control

When $I_N = o(1)$, the limit is in continuous time. If the sequence of action choosen is $a(t)$, we have :

	$I_N = O(1)$	$I_N = o(1)$
State $m_a(t)$	$\frac{dm_a}{dt} = f(m_a(t), a(t))$ $m_a(0) = m_0$.	$m_a(t+1) = f(m_a(t), a(t+1))$ $m_a(0) = m_0$.
Cost $v_a(m_0)$	$\int_0^T \mathrm{cost}(m_a(t))dt$.	$\sum_0^T \mathrm{cost}(m_a(t))dt$.

The optimal cost is $v_*(m_0) = \inf_{\{a \mid a \text{ piecewize lipschitz}\}} v_a(m_0)$.

Question : what is the relation between the stochastic optimal control V_*^N and the deterministic optimal control v_* ?

- Convergence of value functions ?
- Convergence of optimal policies ?

Convergence results

For both discrete $I_N = O(1)$ and continuous $I_N = o(1)$ cases, we have :

Theorem

+ The optimal cost for the stochastic system converges to the optimal cost of its deterministic limit :

$$V_*^N \xrightarrow{N\to\infty} v_*.$$

+ The optimal policy for the deterministic system α_* is asymptotically optimal :

$$\left| V_*^N - V_{a^*}^N \right| \xrightarrow{N\to\infty} 0.$$

- Convergence holds in probability with explicit bounds.
- Second order results (CLT-like) for the discrete case.

However :

– π_*^N might not converge.

– Deterministic limit might be hard.

Idea of the proof – $I_N = \frac{1}{N}$

Proof based on stochastic approximation methods.

1. First auxiliary system : for a policy π, we build a (random) sequence of actions A_π^N corresponding to π :

$$A_\pi^N(t) \stackrel{\text{def}}{=} \pi_t(M_\pi^N(t))$$

Idea of the proof – $I_N = \frac{1}{N}$

Proof based on stochastic approximation methods.

1. First auxiliary system : for a policy π, we build a (random) sequence of actions A_π^N corresponding to π :

$$A_\pi^N(t) \stackrel{\text{def}}{=} \pi_t(M_\pi^N(t))$$

By definition of the drift, $M_\pi^N(t + \frac{1}{N})$ can be written :

$$M_\pi^N(t + \frac{1}{N}) = M_\pi^N(t) + \frac{1}{N}\Big(\underbrace{f(A_\pi^N(t), M_\pi^N(t))}_{\text{Drift (deterministic)}} + \underbrace{\text{noise}}_{\text{Random, } \mathbb{E}[.]=0} \Big).$$

At time $t = k\frac{1}{N}$, $M^N(t)$ is equal to :

$$\underbrace{M_\pi^N(t) = M_0^N + \sum_{i=0}^{k-1} \frac{1}{N} f(A_\pi^N(t), M_\pi^N(\frac{i}{N}))}_{\text{Euler discretization}} \;+\; \underbrace{\frac{1}{N}\sum_{i=0}^{k} \text{noise}}_{\text{Converges to 0}}.$$

Idea of the proof – $I_N = \frac{1}{N}$

Thus, $\left\| M_\pi^N - m_{A_\pi^N} \right\|$ converges to 0 "uniformly in A_π^N" :

$$\mathcal{P}\left(\sup_{t \le T} \left\| M_\pi^N(t) - m_{A_\pi^N}(t) \right\| \ge \frac{A_T}{N} + \epsilon\right) \le \frac{B_T}{N\epsilon^2}. \quad (1)$$

where $m_{A_\pi^N}$ satisfies $\begin{cases} \frac{dm_{A_\pi^N}}{dt} = f(m_{A_\pi^N}(t), A_\pi^N(t)) \\ m_{A_\pi^N}(0) = m_0. \end{cases}$ (deterministic ODE with random parameter).

Idea of the proof – $I_N = \frac{1}{N}$

Thus, $\left\| M_\pi^N - m_{A_\pi^N} \right\|$ converges to 0 "uniformly in A_π^N" :

$$\mathcal{P}\left(\sup_{t \le T} \left\| M_\pi^N(t) - m_{A_\pi^N}(t) \right\| \ge \frac{A_T}{N} + \epsilon \right) \le \frac{B_T}{N\epsilon^2}. \quad (1)$$

where $m_{A_\pi^N}$ satisfies $\begin{cases} \frac{dm_{A_\pi^N}}{dt} = f(m_{A_\pi^N}(t), A_\pi^N(t)) \\ m_{A_\pi^N}(0) = m_0. \end{cases}$ (deterministic ODE with random parameter).

2. Second auxiliary system : for an action function α, we build a policy : $\pi_t(m) = \alpha(t)$.

$$\mathcal{P}\left(\sup_{t \le T} \left\| M_\alpha^N(t) - m_\alpha(t) \right\| \ge \frac{C_T}{N}(\alpha) + \epsilon \right) \le \frac{D_T}{N\epsilon^2}. \quad (2)$$

Idea of the proof – $I_N = \frac{1}{N}$

Thus, $\left\| M_\pi^N - m_{A_\pi^N} \right\|$ converges to 0 "uniformly in A_π^N" :

$$\mathcal{P}\left(\sup_{t \le T} \left\| M_\pi^N(t) - m_{A_\pi^N}(t) \right\| \ge \frac{A_T}{N} + \epsilon \right) \le \frac{B_T}{N\epsilon^2}. \qquad (1)$$

where $m_{A_\pi^N}$ satisfies $\begin{cases} \frac{dm_{A_\pi^N}}{dt} = f(m_{A_\pi^N}(t), A_\pi^N(t)) \\ m_{A_\pi^N}(0) = m_0. \end{cases}$ (deterministic ODE with random parameter).

2. Second auxiliary system : for an action function α, we build a policy : $\pi_t(m) = \alpha(t)$.

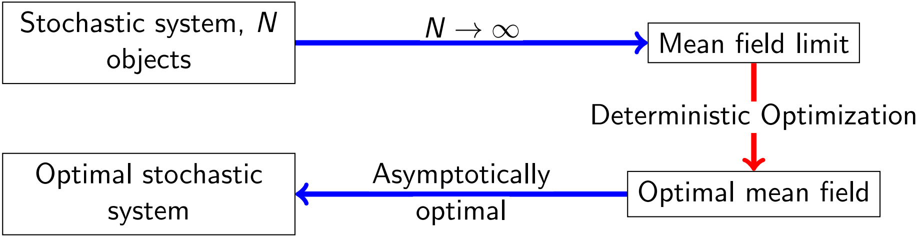
$$\mathcal{P}\left(\sup_{t \le T} \left\| M_\alpha^N(t) - m_\alpha(t) \right\| \ge \frac{C_T}{N}(\alpha) + \epsilon \right) \le \frac{D_T}{N\epsilon^2}. \qquad (2)$$

Combining (1) and (2), one has :

$$\left| V_{\alpha_*}^N - v_{\alpha_*} \right| \to 0. \quad \text{and} \quad \left| \sup_\pi V_\pi^N - \sup_\alpha v_\alpha \right| \to 0.$$

□

How to apply this in practice ?



The complexity of the method depends on the complexity of the deterministic problem :

1. If we can solve the deterministic limit :
   - Compute a^*.
   - Works well for the stochastic system.
2. Design an approximation algorithm for the deterministic system :
   - also an approximation (asymptotically) for stochastic problem.
3. Use brute force computation :
   - Compared to the random case, there is no expectation to compute.
   - Problem simpler but still hard (HJB, dynamic programming)

Outline

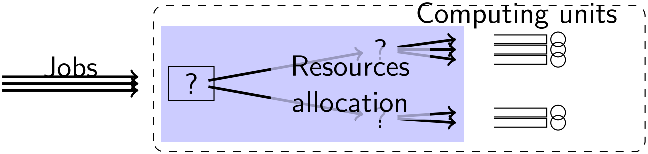
1. Introduction

2. Optimal control of mean field models
   - How to define the limiting deterministic optimization problem?
   - Convergence results.

3. Application to a brokering problem and speed of convergence.
   - When the approximation becomes valid?
   - What is the quality of the approximation?

4. Going further: non-smooth dynamics
   - if the drift f is not continuous: can we define an ODE $\frac{dm}{dt} = f(m)$?
   - What is the limit of M^N?

5. Conclusion



A simple brokering problem

Model

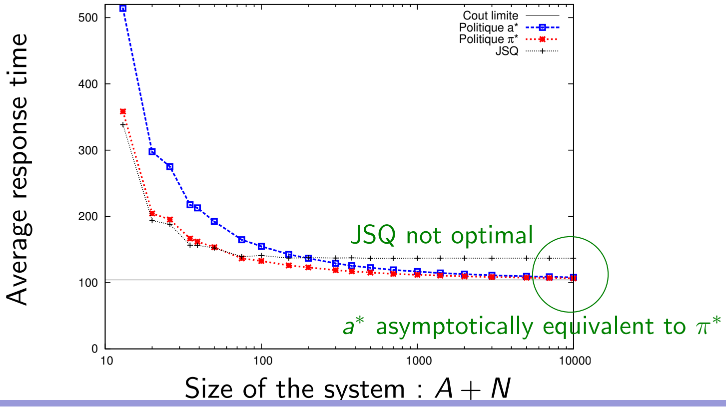
- A applications : an application sends a tasks at time t with proba. λ_t.
- N processors grouped in C clusters : each processor of cluster c completes one task per unit of time with probability μ_t^c.
- Goal : allocate tasks to clusters minimizing the average response time.

- Stochastic policy impossible to compute.
- When $A \to \infty$ and $N \to \infty$, The problem becomes :
  - Find $y_1^1 \dots y_T^d \in \mathbb{R}$ to minimize $\sum_{t=1}^{T} \sum_{i=1}^{d} e_t^i$ such that
    - $e_{t+1}^i = (e_t^i + y_t^i - x_t^i)^+$ and
    - $\sum_i y_t^i = y_t$.
  - Optimal policy can be computed by a greedy algorithm (best effort).

Comparison of the performance of π^* against JSQ

Two optimal policies for the deterministic system : π^* and a^*.

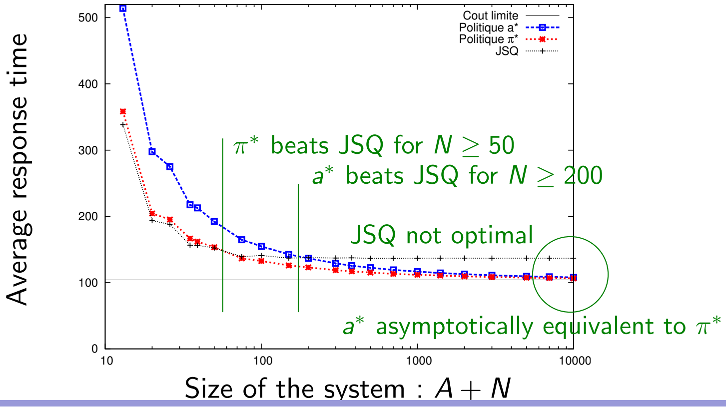
- $V_{a^*}^N$ – average response time for open-loop policy a^*.
  - Action chosen at time t is $a^*(t)$.
- $V_{\pi^*}^N$ – average response time for the closed-loop policy π^*.
  - Action chosen at time t is $\pi^*(t, M^N(t))$.



Comparison of the performance of π^* against JSQ

Two optimal policies for the deterministic system : π^* and a^*.

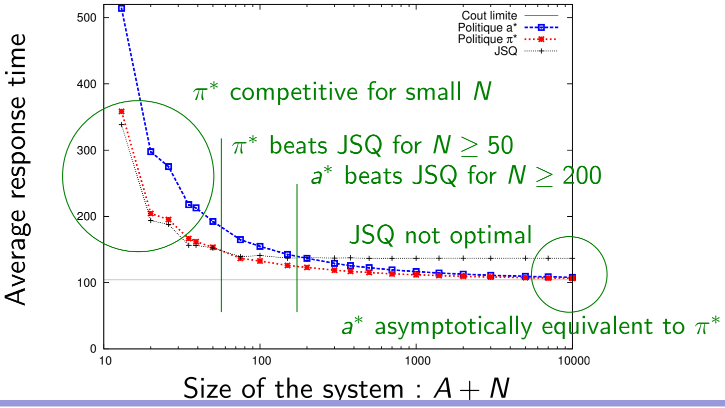
- $V_{a^*}^N$ – average response time for open-loop policy a^*.
  - Action chosen at time t is $a^*(t)$.
- $V_{\pi^*}^N$ – average response time for the closed-loop policy π^*.
  - Action chosen at time t is $\pi^*(t, M^N(t))$.



Comparison of the performance of π^* against JSQ

Two optimal policies for the deterministic system : π^* and a^*.

- $V_{a^*}^N$ – average response time for open-loop policy a^*.
  - Action chosen at time t is $a^*(t)$.
- $V_{\pi^*}^N$ – average response time for the closed-loop policy π^*.
  - Action chosen at time t is $\pi^*(t, M^N(t))$.



Outline

1. Introduction

2. Optimal control of mean field models

- How to define the limiting deterministic optimization problem?
- Convergence results.

3. Application to a brokering problem and speed of convergence.

- When the approximation becomes valid?
- What is the quality of the approximation?

4. Going further: non-smooth dynamics

- if the drift f is not continuous: can we define an ODE $\frac{dm}{dt} = f(m)$?
- What is the limit of M^N?

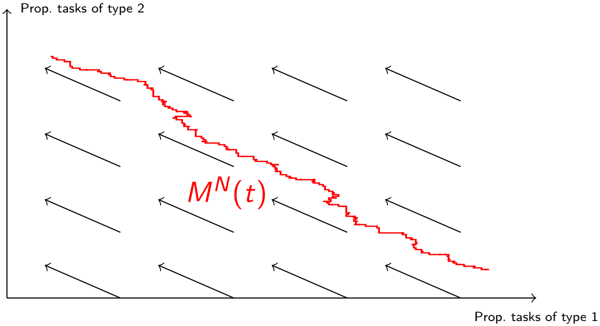
5. Conclusion

Non-smooth dynamics are common

- Boundary conditions (ex : queuing systems)
- When applying a policy to control the system (threshold effects).

Example : Static priority

- Tasks of type 1 have priority.

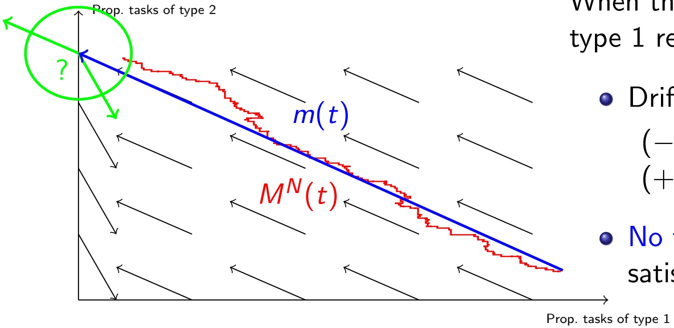


Non-smooth dynamics are common

- Boundary conditions (ex : queuing systems)
- When applying a policy to control the system (threshold effects).

Example : Static priority

- Tasks of type 1 have priority.



When the number of tasks of type 1 reaches 0 :

- Drift not continuous :

$$\begin{array}{ll} (-2,+1) & \text{if } x > 0 \\ (+1,-2) & \text{if } x = 0 \end{array}$$

- No trajectory $m(t)$ satisfying $\frac{dm}{dt} = f(m)$.

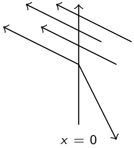
Idea : build a differential inclusion using convex closure.

- Let F be the *"convex closure"* of f :

$$F(m) \stackrel{\text{def}}{=} \bigcap_{\epsilon>0} \overline{\text{conv}}\left(\{f(z) : \|z-m\| \leq \epsilon\}\right).$$

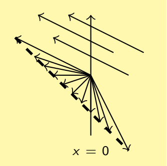
- The original ODE is replaced by :

$$\frac{dm}{dt} \in F(m) \quad \text{p.p.}$$



Drift f

becomes



Convex closure F

General convergence result

Let $D(m_0) \stackrel{\text{def}}{=}$ be the set of solutions of $\dfrac{dm}{dt} \in F(m)$ with $m(0) = m_0$.

Then, without any condition on f :

Theorem

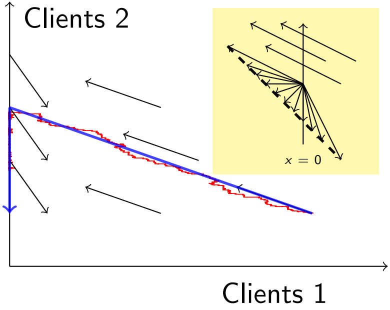
$$\forall T > 0, \inf_{m \in D(m_0)} \quad \sup_{0 \leq t \leq T} \left\| M^N(t) - m(t) \right\|_\infty \xrightarrow{\mathcal{P}} 0.$$

In particular,

Corollary

If $\dfrac{dm}{dt} \in F(m)$ has a unique solution m : $\sup_{0 \leq t \leq T} \left\| M^N(t) - m(t) \right\|_\infty \xrightarrow{\mathcal{P}} 0.$

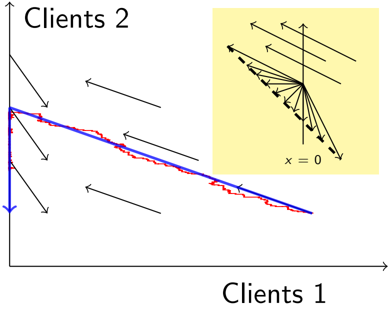
Application on the static priority example.



- The drift is easy to compute.
- The ODE has no solution.

- Computation of convex closure.
- The DI has a unique solution.

Application on the static priority example.



- The drift is easy to compute.
- The ODE has no solution.

- Computation of convex closure.
- The DI has a unique solution.

The differential inclusion approach :

- Allows one to apply mean field methods to non-smooth dynamics (boundary conditions, controlled systems).

In particular, if π is a policy, then :

- The controlled system M_π^N converges to the deterministic controlled dynamic m_π.

Outline

1. Introduction

2. Optimal control of mean field models
   - How to define the limiting deterministic optimization problem?
   - Convergence results.

3. Application to a brokering problem and speed of convergence.
   - When the approximation becomes valid?
   - What is the quality of the approximation?

4. Going further: non-smooth dynamics
   - if the drift f is not continuous: can we define an ODE $\frac{dm}{dt} = f(m)$?
   - What is the limit of M^N?

5. Conclusion

Conclusion and Perspectives

Mean field approximation for optimization problem :

- Two new heuristics for the stochastic problem : a^* and π^*.
- Asymptotically optimal as N grows but also efficient for small N.
- Differential inclusions allows to consider non-smooth dynamics.

Prove that the deterministic optimization problem is an approximation of the stochastic optimization problem.

Applications and perspectives

- Use this results to prove optimality of certain class of policies.
- Optimization of SDE *v.s.* MDP.
- Study infinite horizon results (average cost).

References corresponding to this work :

- Optimization and mean field, discrete time limit :
  - N. Gast, B. Gaujal – *A Mean Field Approach for Optimization in Particles Systems and Applications* – ValueTools 2009
  - N. Gast, B. Gaujal – *A Mean Field Approach for Optimization in Discrete Time* – to appear in DEDS
- Optimization and mean field, continuous time limit :
  - N. Gast, B. Gaujal, J.-Y. Le Boudec – *Mean field for Markov Decision Processes : from Discrete to Continuous Optimization* – Inria RR 7239
- Non-smooth mean field :
  - N. Gast, B. Gaujal – *Mean field Limit of Non-Smooth Systems and Differential Inclusions* – MAMA 2010
  - N. Gast, B. Gaujal – *Mean field limit of non-smooth systems : a differential inclusion limit.* – Inria RR 7315