A Mean Field Approach for Optimization in Particles Systems and Applications

Nicolas Gast Bruno Gaujal

Grenoble University

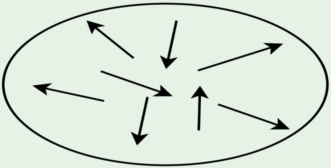
ValueTools 2009, Oct. 20-22, Pisa

Introduction

Mean field has been introduced by physicists to study systems of interacting objects. For example, the movement of particles in the air:

First solution: the microscopic description

The system is represented by the states of each particle.



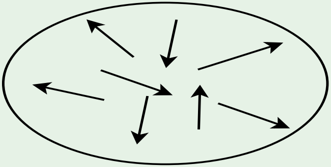
- Impossible to study.

Introduction

Mean field has been introduced by physicists to study systems of interacting objects. For example, the movement of particles in the air:

First solution: the microscopic description

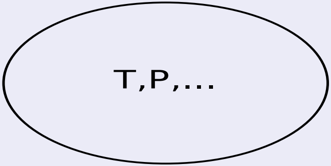
The system is represented by the states of each particle.



- Impossible to study.

Second solution: macroscopic equations

System described by temperature, pressure,...



- Much easier

- The transition from microscopic description to macroscopic equations is called the mean field approximation.

Mean Field in Computer Science

More recently, Mean field has been used to analyze performance of communication systems with one goal: prove the convergence.

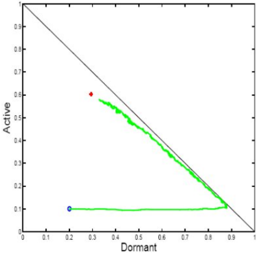
- Approximation of population processes [Kurtz, 81]
- Propagation of chaos [Sznitman, 91]
- Performance of TCP [Baccelli, McDonald, Reynier 02]
- Approximation of stochastic evolution in games [Benaïn, Weibull 03]
- Random Acces Control [Bordenave, McDonald, Proutière 05]
- Reputation Systems [Le Boudec et al. 07]
- ...

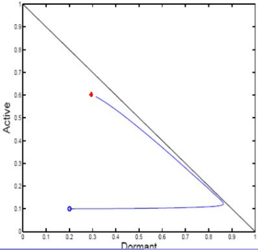
Mean Field in Computer Science

More recently, Mean field has been used to analyze performance of communication systems with one goal: prove the convergence.

- Approximation of population processes [Kurtz, 81]
- Propagation of chaos [Sznitman, 91]
- Performance of TCP [Baccelli, McDonald, Reynier 02]
- Approximation of stochastic evolution in games [Benaïn, Weibull 03]
- Random Acces Control [Bordenave, McDonald, Proutière 05]
- Reputation Systems [Le Boudec et al. 07]
- ...

In many examples, it can be shown that when the number of users grows, the average behavior of the system becomes deterministic.





Mean Field in Computer Science

More recently, Mean field has been used to analyze performance of communication systems with one goal: prove the convergence.

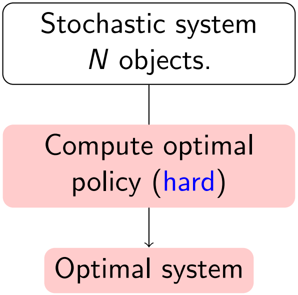
- Approximation of population processes [Kurtz, 81]
- Propagation of chaos [Sznitman, 91]
- Performance of TCP [Baccelli, McDonald, Reynier 02]
- Approximation of stochastic evolution in games [Benaïn, Weibull 03]
- Random Acces Control [Bordenave, McDonald, Proutière 05]
- Reputation Systems [Le Boudec et al. 07]
- ...

In many examples, it can be shown that when the number of users grows, the average behavior of the system becomes deterministic.

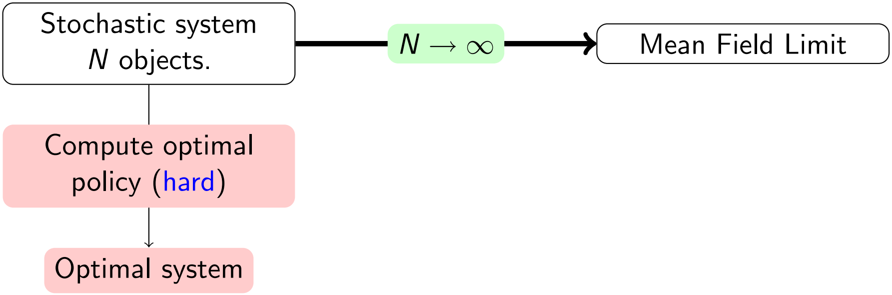
Aim of this work

Study mean field results for a large class of optimization problems.

Our Approach

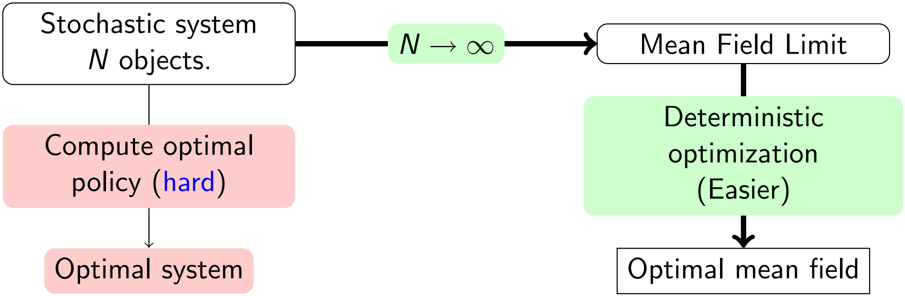


Our Approach



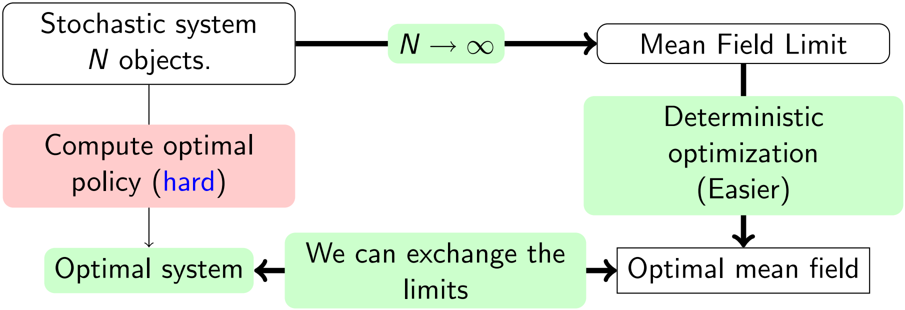
1. Compute mean field limit (constructive definition).

Our Approach



1. Compute mean field limit (constructive definition).
2. Solve the deterministic problem.

Our Approach

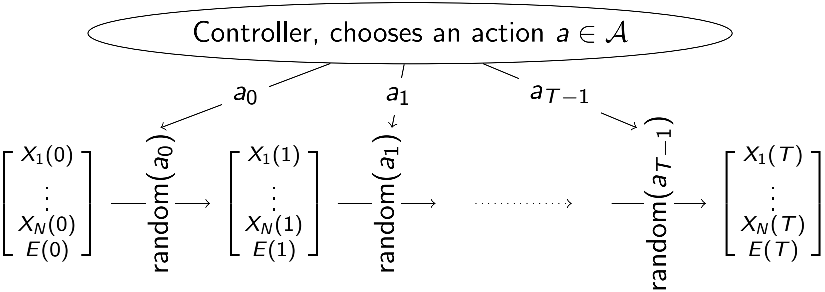


Our results

The Markov Decision Process also converges to a deterministic limit. More precisely, when N grows:

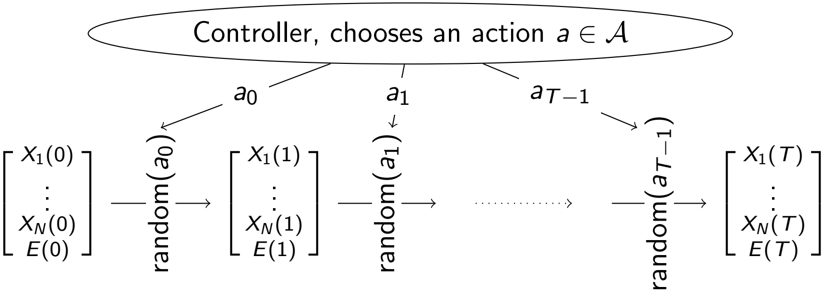
1. The optimal cost converges.
2. The optimal policy is asymptotically optimal.
3. The speed of convergence is $O(\sqrt{N})$.

A Markov Decision Process



- N objects evolving in a finite state space.
- Environment $E(t)$ at time t ($E(t) \in \mathbb{R}^d$)

A Markov Decision Process

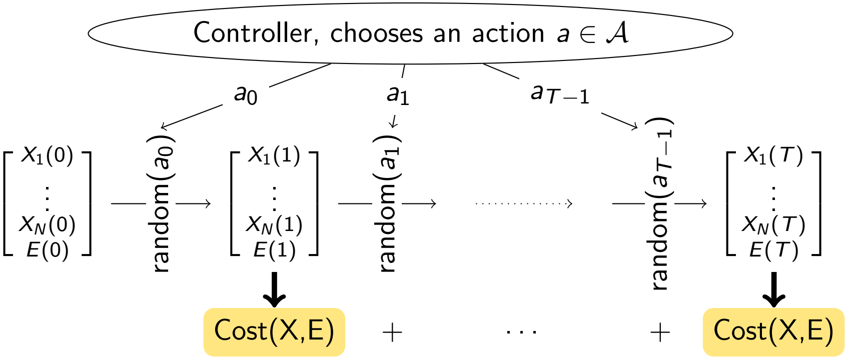


Mean field assumption

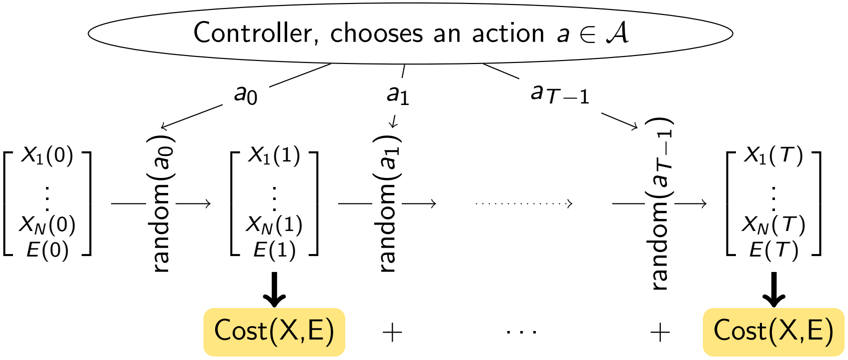
We define the Empirical measure $\mathrm{M}(t)$ – The i^{th} component $(\mathrm{M}(t))_i$ is the proportion of objects in state i at time t.

1. $E(t+1)$ only depends on the empirical measure $\mathrm{M}(t)$.
2. The evolution of an object is Markovian, depends on $E(t)$ but is independent of the other objects.

A Markov Decision Process

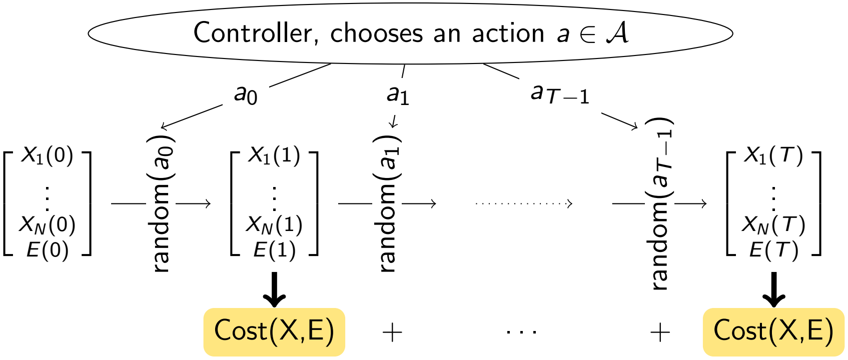


A Markov Decision Process



Goal: Find a policy to minimize finite-time expected cost (*or* infinite horizon discounted cost):

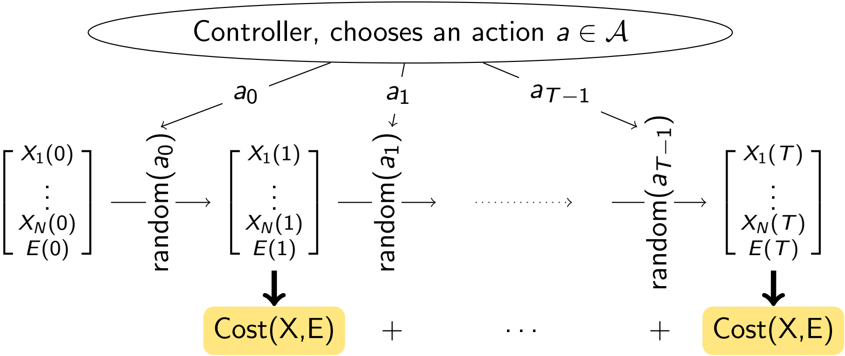
A Markov Decision Process



Goal: Find a policy to minimize finite-time expected cost (*or* infinite horizon discounted cost):

$$\sum_{t=1}^{T} \mathrm{c}(\mathrm{M}_t^N(\pi^N), \mathrm{E}_t^N(\pi^N))$$

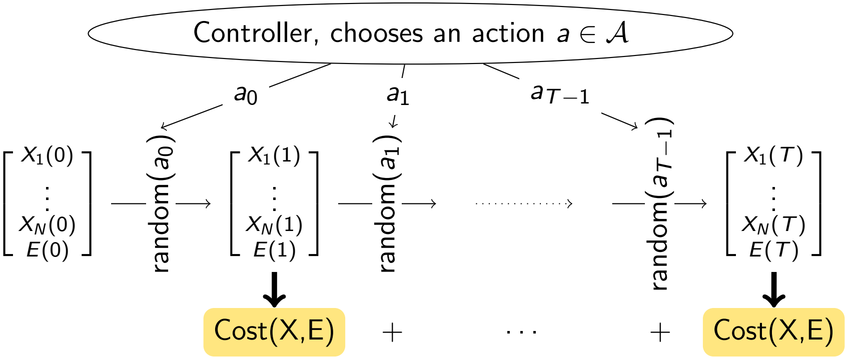
A Markov Decision Process



Goal: Find a policy to minimize finite-time expected cost (*or* infinite horizon discounted cost):

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathrm{c}(\mathrm{M}_t^N(\pi^N), \mathrm{E}_t^N(\pi^N))\right]$$

A Markov Decision Process



Goal: Find a policy to minimize finite-time expected cost (*or* infinite horizon discounted cost):

$$V_T^{*N}(\mathrm{M}_0^N, \mathrm{E}_0^N) = \inf_{\pi^N} \mathbb{E}\left[\sum_{t=1}^{T} \mathrm{c}(\mathrm{M}_t^N(\pi^N), \mathrm{E}_t^N(\pi^N))\right]$$

Construction of limit

A1 Users are only dependent through environment;

A2 Compact action set; **A3** Continuity of parameters;

A4 $\mathrm{M}_0^N, \mathrm{E}_0^N \xrightarrow{a.s} m_0, e_0$.

Construction of limit

A1 Users are only dependent through environment;

A2 Compact action set; **A3** Continuity of parameters;

A4 $\mathrm{M}_0^N, \mathrm{E}_0^N \xrightarrow{a.s} m_0, e_0$.

Theorem (Without control – Le Boudec, Mc Donald, Mudinger 07)

Let $a = a_0 a_1 \dots$ be a sequence of actions. Under (A1,A3,A4), for all t:

$$\mathrm{M}_a^N(t), \mathrm{E}_a^N(t) \xrightarrow{a.s} m_a(t), e_a(t)$$

where m_a, e_a is a discrete time dynamical system, i.e. can be written

$$m_a(t+1), e_a(t+1) = f(m_a(t), e_a(t)).$$

Construction of limit

A1 Users are only dependent through environment;
A2 Compact action set; **A3** Continuity of parameters;
A4 $\mathrm{M}_0^N, \mathrm{E}_0^N \xrightarrow{a.s} m_0, e_0$.

Theorem (Without control – Le Boudec, Mc Donald, Mudinger 07)

Let $a = a_0 a_1 \dots$ be a sequence of actions. Under (A1,A3,A4), for all t:

$$\mathrm{M}_a^N(t), \mathrm{E}_a^N(t) \xrightarrow{a.s} m_a(t), e_a(t)$$

where m_a, e_a is a discrete time dynamical system, i.e. can be written

$$m_a(t+1), e_a(t+1) = f(m_a(t), e_a(t)).$$

This defines a deterministic optimization problem:

$$\text{Find} \quad a^* = \quad \underset{a_0 \dots a_{T-1}}{\arg\min} \quad \sum_{t=0}^{T} \mathrm{c}_t(m_a(t), e_a(t))$$

Optimal cost convergence

- V^{*N} – optimal cost for the system of size N.
- v^* – optimal cost for the deterministic limit.
- $a_0^* a_1^* a_2^* \dots$ – sequence of optimal actions for the deterministic limit.

Theorem (Convergence of the optimal cost)

Under assumptions (A1,A2,A3,A4), almost surely:

$$\lim_{N\to\infty} V_T^{*N}(M_0^N, E_0^N) = v_T^*(m_0, e_0) = \lim_{N\to\infty} V^N_{a_0^* \dots a_{T-1}^*}(M_0^N, E_0^N)$$

In particular, this shows that:

- Optimal cost converges
- Static policy (a^*) is asymptotically optimal

Remark: π^* is also asymptotically optimal but not asymptotically better.

Speed of convergence: a central limit theorem

(A4-bis) $\sqrt{N}((\mathrm{M}_0^N, \mathrm{E}_0^N) - (m_0, e_0)) \xrightarrow{\mathcal{L}} G_0$. **(A5)** Parameters differentiable.

Theorem (CLT for the evolution of objects)

Under assumptions (A1,A2,A3,A4bis,A5), if the actions taken by the controller are fixed, then there exists a Gaussian variable G_t s.t:

$$\sqrt{N}\big((\mathrm{M}_t^N, \mathrm{E}_t^N) - (m_t, e_t)\big) \xrightarrow{\text{Law}} G_t$$

The covariance of G_t is given by an iterative expression.

Speed of convergence: a central limit theorem

(A4-bis) $\sqrt{N}((\mathrm{M}_0^N, \mathrm{E}_0^N) - (m_0, e_0)) \xrightarrow{\mathcal{L}} G_0$. **(A5)** Parameters differentiable.

Theorem (CLT for the evolution of objects)

Under assumptions (A1,A2,A3,A4bis,A5), if the actions taken by the controller are fixed, then there exists a Gaussian variable G_t s.t:

$$\sqrt{N}\left((\mathrm{M}_t^N, \mathrm{E}_t^N) - (m_t, e_t)\right) \xrightarrow{\text{Law}} G_t$$

The covariance of G_t is given by an iterative expression.

Theorem (Second order theorem for cost)

Under assumptions (A1,A2,A3,A4bis,A5), $\exists \beta, \gamma, \beta', \gamma'$ such that when N does to infinity:

- $\sqrt{N}\left| V_T^{*N}(\mathrm{M}_0^N, \mathrm{E}_0^N) - V_{a^*}^N(\mathrm{M}_0^N, \mathrm{E}_0^N) \right| \leq_{\mathrm{st}} \beta + \gamma \| G_0 \|_\infty$
- $\sqrt{N}\left| V_T^{*N}(\mathrm{M}_0^N, \mathrm{E}_0^N) - v_{0\dots T}^*(m_0, e_0) \right| \leq_{\mathrm{st}} \beta' + \gamma' \| G_0 \|_\infty$

Discounted case

Assumptions

A6 **Homogeneity in time** – cost and probability kernel K_t do not depend on t.

A7 **Bounded cost** – $\sup_{\mathrm{M},\mathrm{E}} |\mathrm{c}(\mathrm{M},\mathrm{E})| < \infty$.

$$V_\pi^{\delta N}(\mathrm{M}_0^N, \mathrm{E}_0^N) = \mathbb{E}_\pi\Big[\sum_{t=0}^{\infty} \delta^t \mathrm{c}(\mathrm{M}_t^N, \mathrm{E}_t^N)\Big].$$

Theorem

Under assumptions (A1,A2,A3,A4,A6,A7),

$$\lim_{N\to\infty} V_*^{\delta N}(\mathrm{M}^N, \mathrm{E}^N) = v_*^\delta(m, e) \text{ a.s.}$$

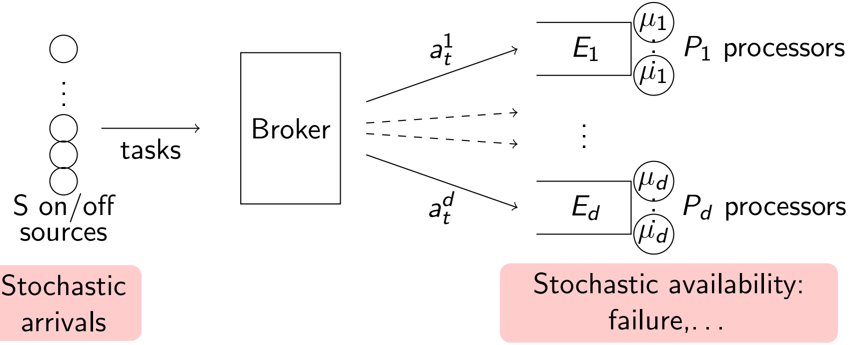
A simple resource allocation problem

Aim of the example

- The set of problem we solved is non empty!
- When does N becomes large enough for the approximation to apply?
  - *i.e.* when do we beat classical solutions?
- Show the tightness of the bounds

A simple resource allocation problem

Number of objects is $N = S + P_1 + \cdots + P_d$.



- Optimize the total completion time $= \sum_{t=0}^{T} \sum_{i=1}^{d} E_i(t)$.

Optimal policy: stochastic and limit case

The stochastic system is hard to solve

1. This problem is a multidimensional restless bandit problem
   - Known to be hard
   - Existence of heuristics (Index policies)
2. In practice in such systems [EGEE]
   - Use of heuristics (JSQ)

Optimal policy: stochastic and limit case

The stochastic system is hard to solve

1. This problem is a multidimensional restless bandit problem
   - Known to be hard
   - Existence of heuristics (Index policies)
2. In practice in such systems [EGEE]
   - Use of heuristics (JSQ)

Using our framework: compute optimal mean field

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).

Numerical example

This provides two policies for the initial stochastic system.

- a^* : we apply $a^*_t \stackrel{\mathrm{def}}{=} \pi^*_t(m_t, e_t)$ – static policy.
- π^* : at t, we apply $\pi^*_t(\mathrm{M}^N_t, \mathrm{E}^N_t)$ – adaptive policy.

- Static policy a^*
  At time t, we apply the optimal deterministic sequence of actions $a_1 \dots a_T$, regardless of the current state $\mathrm{M}^N_t, \mathrm{E}^N_t$.
- Adaptive policy π^*
  At time t, the system is in state $\mathrm{M}^N_t, \mathrm{E}^N_t$. We compute the optimal deterministic action that would be taken in this state: we apply $\pi^*(\mathrm{M}^N_t, \mathrm{E}^N_t)$

Numerical example

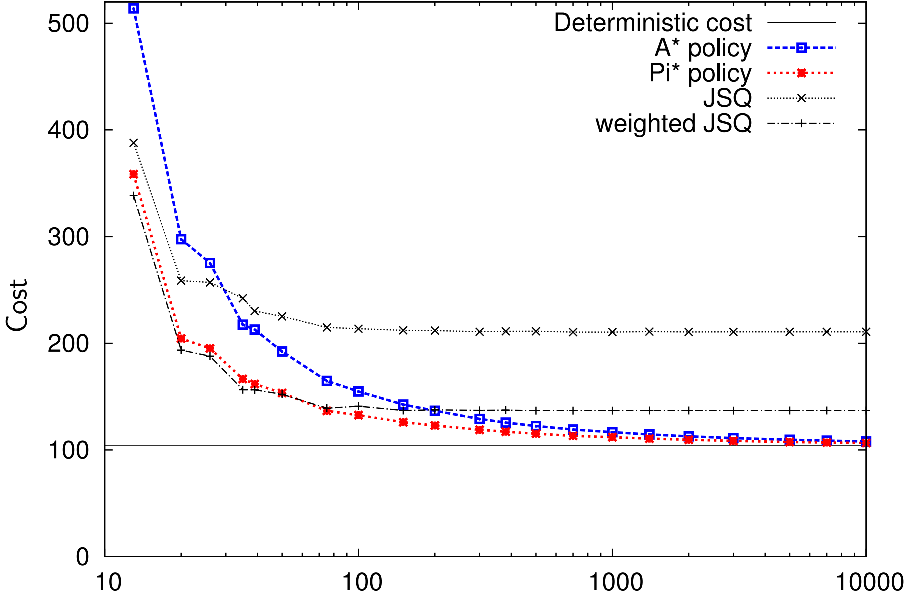
This provides two policies for the initial stochastic system.

- a^* : we apply $a_t^* \stackrel{\text{def}}{=} \pi_t^*(m_t, e_t)$ – static policy.
- π^* : at t, we apply $\pi_t^*(\mathrm{M}_t^N, \mathrm{E}_t^N)$ – adaptive policy.

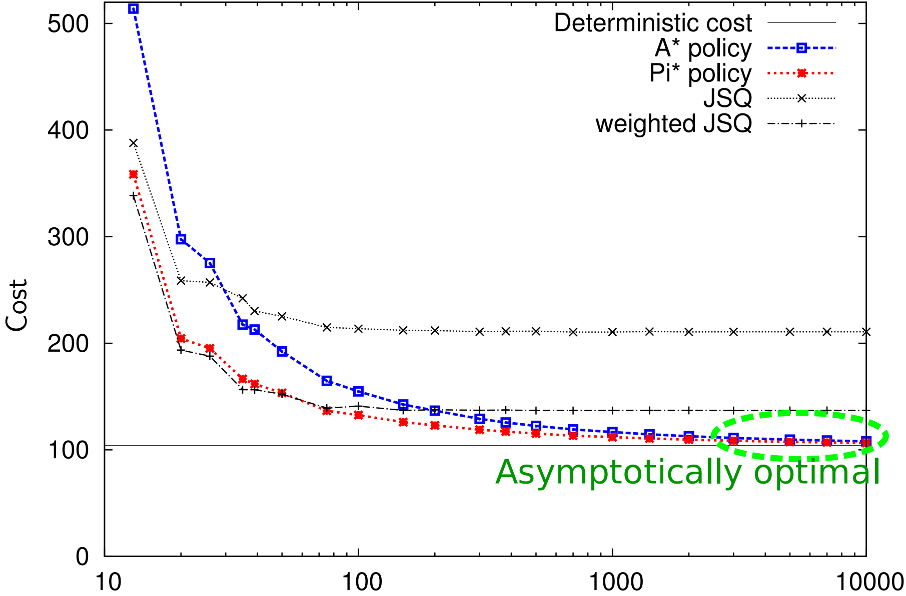
We want to compare

- $V_{a^*}^N$ – cost when applying a^*
- $V_{\pi^*}^N$ – cost when applying π^*
- V_{JSQ}^N – cost of Join Shortest Queue.
- v^* – cost of the deterministic limit.

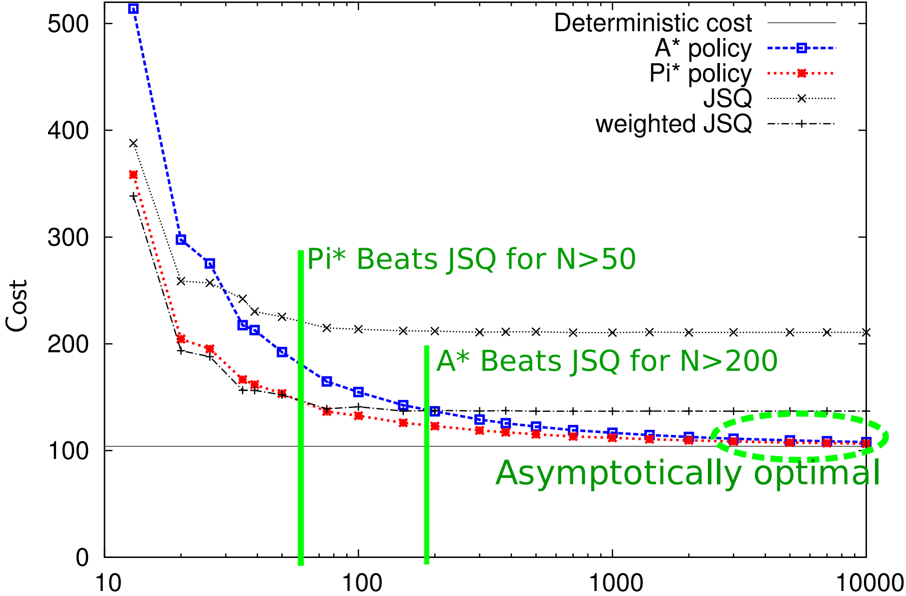
Cost convergence



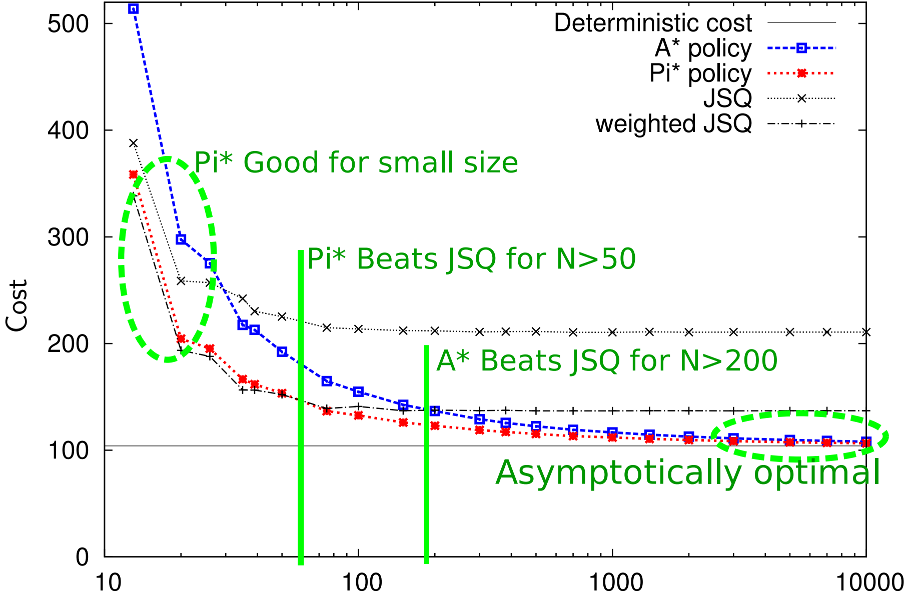
Cost convergence



Cost convergence

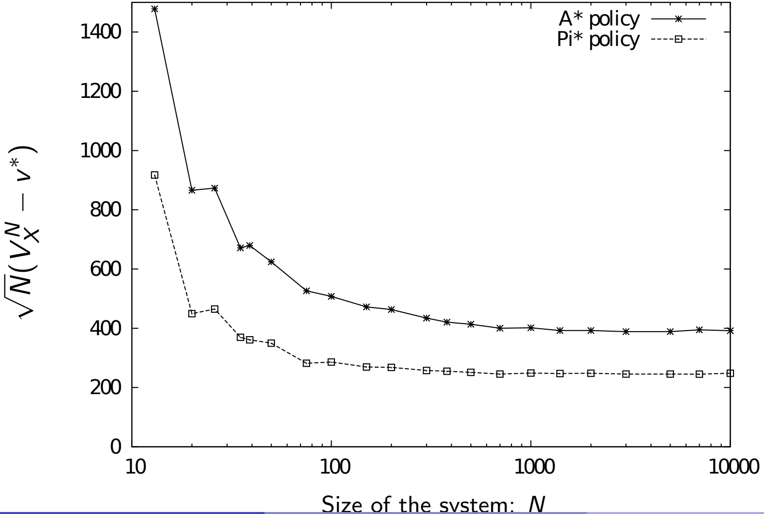


Cost convergence



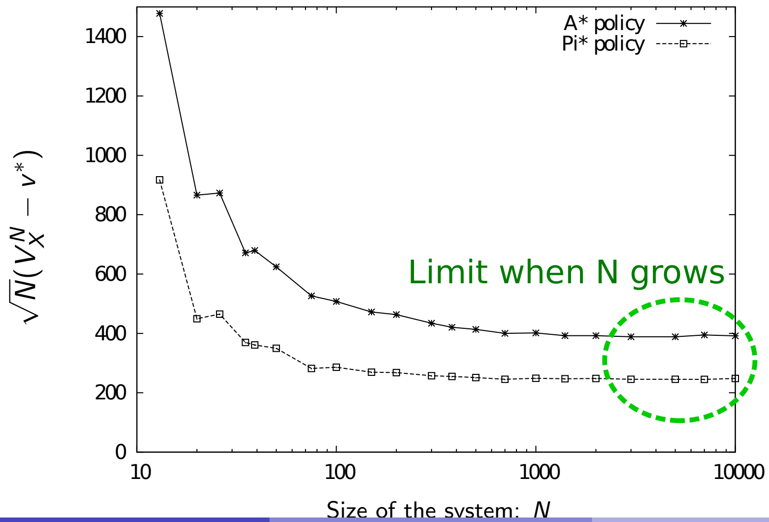
Speed of convergence – central limit theorem

Plot of $\sqrt{N}(V^N_{\pi^*}(\mathrm{M},\mathrm{E}) - v^*(m,e))$ and $\sqrt{N}(V^N_{a^*}(\mathrm{M},\mathrm{E}) - v^*(m,e))$



Speed of convergence – central limit theorem

Plot of $\sqrt{N}(V^N_{\pi^*}(\mathrm{M},\mathrm{E}) - v^*(m,e))$ and $\sqrt{N}(V^N_{a^*}(\mathrm{M},\mathrm{E}) - v^*(m,e))$



Conclusion

Theoretical results

- Optimal policy of the deterministic limit is asymptotically optimal.
- Speed of convergence in $O(\sqrt{N})$.

The example shows

- Efficient for low values of N (≈ 100 in the example).
  - Beats classical heuristics
- Two new heuristics for the stochastic problem: a^* and Π^*.
- The bounds for speed are tight: $\Omega(\sqrt{N})$.

Conclusion

How to apply this in practice?

To apply this in practice, there are three cases (from best to worse):

1. We can solve the deterministic limit:
   - apply a^* or π^*.
2. Design an approximation algorithm for the deterministic system:
   - also an approximation (asymptotically) for stochastic problem.
3. Use brute force computation:
   - $v^*_{t\ldots T}(m,e) = \mathrm{c}(m,e) + \inf_a v^*_{t+1\ldots T}(\phi_a(m,e))$
   - Compared to the random case, there is no expectation to compute.

Conclusion

How to apply this in practice?

To apply this in practice, there are three cases (from best to worse):

1. We can solve the deterministic limit:
   - apply a^* or π^*.
2. Design an approximation algorithm for the deterministic system:
   - also an approximation (asymptotically) for stochastic problem.
3. Use brute force computation:
   - $v^*_{t\dots T}(m,e) = \mathrm{c}(m,e) + \inf_a v^*_{t+1\dots T}(\phi_a(m,e))$
   - Compared to the random case, there is no expectation to compute.

Future work

- Study other limiting regimes (*e.g.* number of transitions is $o(N)$).
- Steady-state behavior
- Dependence of the users.

Thanks

Thank you for your attention.