Utilisation des méthodes champ moyen pour l'évaluation de performance

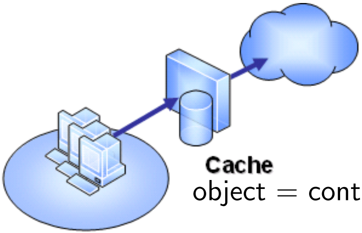
Nicolas Gast (Inria)

Inria, Grenoble, France

Séminaire de l'institut Fourier, Octobre 2016

Models of interacting objects (in computer science)

Wifi: object = device



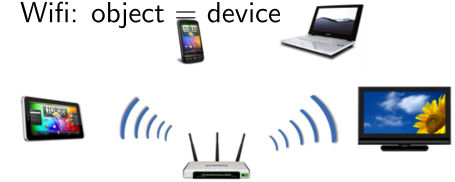
Cache

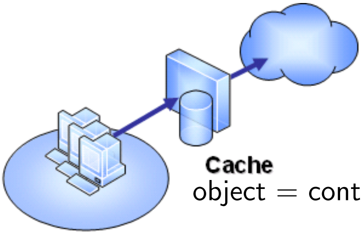
object = content

Cluster: object = server

Models of interacting objects (in computer science)

Wifi: object = device





Cache
object = content

Cluster: object = server

Problem: state space explosion
S states per object, N objects

$$\Rightarrow S^N \text{ states}$$

(and $4^{20} = 10^{12}$)

Mean-field model

Population of N objects.

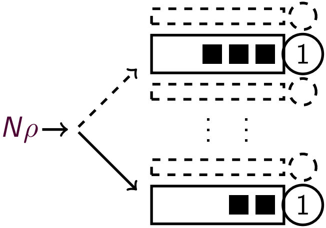
$$X_i(t) = \text{ fraction of objects in state } i$$

Mean-field model

Population of N objects.

$$X_i(t) = \text{ fraction of objects in state } i$$

Example: N servers

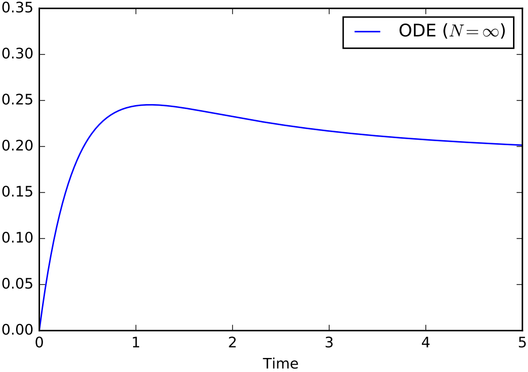


Randomly choose two, and select one

The state is $(X_0, X_1, X_2 \dots)$.

$$X_i(t) = \text{ fraction of servers with } i \text{ jobs}$$

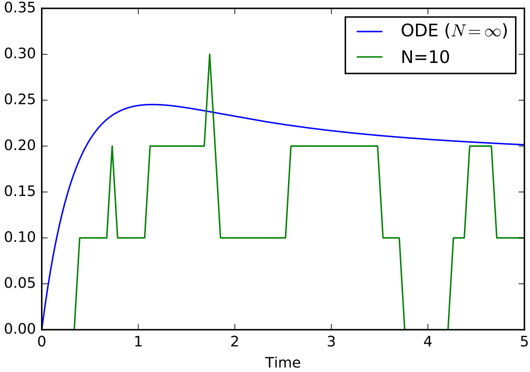
Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

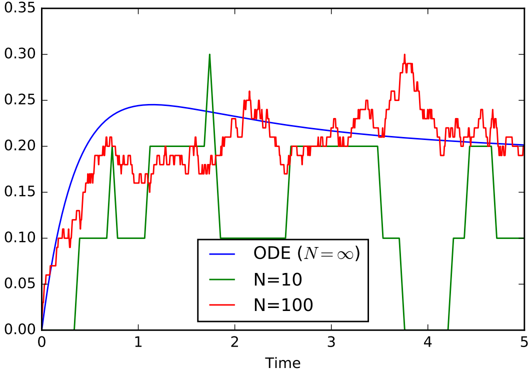
Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

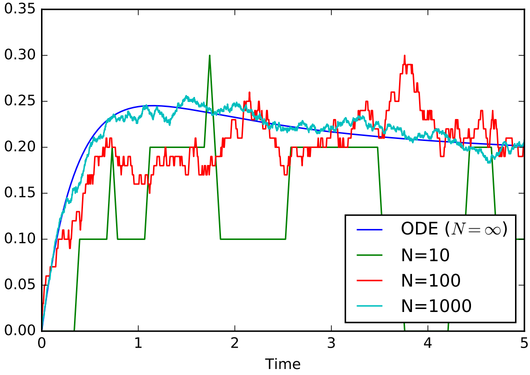
Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

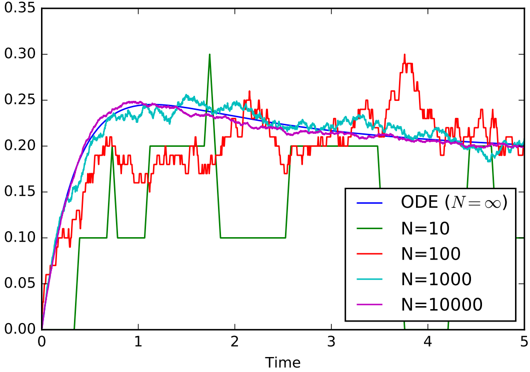
Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

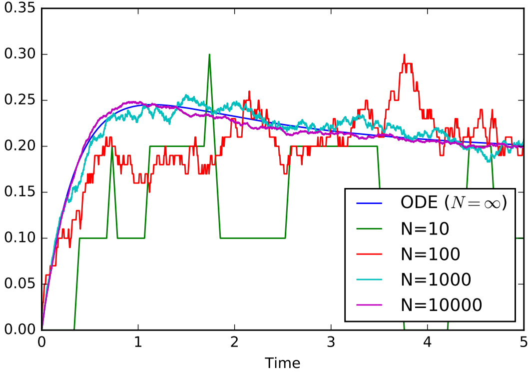
Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

Some systems simplify as N grows



Example. Two-choice model
Fraction of servers with 3 jobs

At time 0: all servers have 1 jobs.

Objective of this talk

- When is the ODE approximation valid / not valid?
- What is the accuracy?

Outline

Population CTMC

A population process is a sequence of CTMC $\mathbf{X}^N$, indexed by the population size N, with state spaces $\mathbf{E}^N \subset \mathbf{E}$ such that the transitions are (for $\ell \in \mathcal{L}$):

$$X \mapsto X + \frac{\ell}{N} \qquad \text{at rate } N\beta_\ell(X).$$

The drift is $f(x) = \sum_\ell \ell\beta_\ell(x)$.

We denote by x the solution of the associated ODE

$$\dot{x} = f(x).$$

Transient regime

Let Φ_t denotes the (unique) solution of the ODE:

$$\Phi_t x = x + \int_0^t \Phi_s x ds.$$

Theorem (Kurtz 70s)

If f is Lipschitz-continuous with constant L, then for any fixed T:

$$\lim_{N\to\infty} \sup_{t<T} \left\| X^N(t) - x(t) \right\| = 0.$$

Proof.

Martingale concentration + Gronwall. $\square$

The fixed point method

Markov chain

Transient regime $\dot{p} = pK$

$t \to \infty$

Stationary $\pi K = 0$

The fixed point method

Markov chain | Mean-field

Transient regime

Stationary

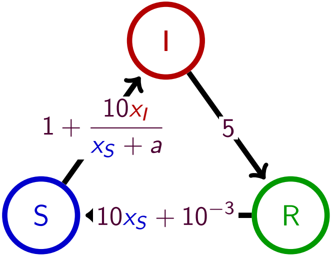
$$\begin{tikzcd} \dot{p} = pK \arrow[r, "N \to \infty"] \arrow[d, "t \to \infty"] & \dot{x} = f(x) \\ \pi K = 0 \arrow[r, "?"] & \boxed{\underset{\text{fixed points}}{f(x^*) = 0}} \end{tikzcd}$$

Method was used in many papers:

- Bianchi 00, Performance analysis of the IEEE 802.11 distributed coordination function.
- Ramaiyan et al. 08, Fixed point analys is of single cell IEEE 802.11e WLANs: Uniqueness, multistability.
- Kwak et al. 05, Performance analysis of exponenetial backoff.
- Kumar et al 08, New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs.

Does it always work?

SIRS model:

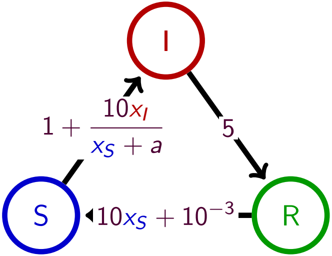


- Markov chain is irreducible.
- Unique fixed point $f(x^*) = 0$.

	Fixed point $f(x) = 0$		Stat. measure $N = 10^3, 10^4 \dots$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234

Does it always work?

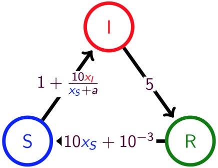
SIRS model:

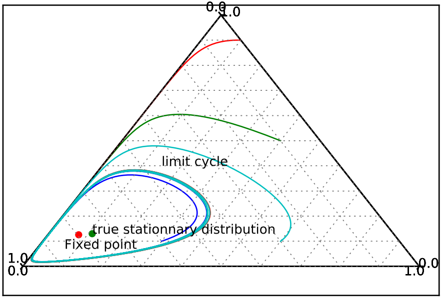


- Markov chain is irreducible.
- Unique fixed point $f(x^*) = 0$.

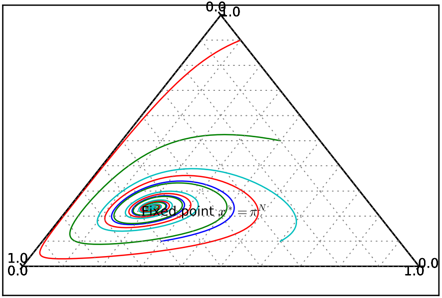
	Fixed point $f(x) = 0$		Stat. measure $N = 10^3, 10^4 \dots$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234
$a = .1$	0.078	0.126	0.11	0.13

What happened?





$a = .1$



$a = .3$

Fixed points?

$$\begin{tikzcd}[column sep=large]
 & \text{Markov chain} & & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[rr, "N \to \infty"] \arrow[d, "t \to \infty"] & & \dot{x} = f(x) \\
\text{Stationary} & \pi K = 0 \arrow[rr, "?"] & & \boxed{f(x^*) = 0} \\
 & & & \text{fixed points}
\end{tikzcd}$$

Fixed points?

$$\begin{array}{cccc} & \text{Markov chain} & & \text{Mean-field} \\ \text{Transient regime} & \dot{p} = pK & \xrightarrow{\quad N \to \infty \quad} & \dot{x} = f(x) \\ & \Big\downarrow t \to \infty & & \Big\downarrow \cancel{t \to \infty} \\ \text{Stationary} & \pi K = 0 & \xrightarrow{\quad \cancel{N \to \infty} \quad} & \boxed{\cancel{f(x^*) = 0}} \\ & & & \text{fixed points} \end{array}$$

Fixed points?

$$
\begin{tikzcd}
 & \text{Markov chain} & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[r, "N \to \infty"] \arrow[d, "t \to \infty"] & \dot{x} = f(x) \arrow[d, "\cancel{t \to \infty}"'] \arrow[d, bend left, "\text{if yes}"] \\
\text{Stationary} & \pi K = 0 \arrow[r, "\cancel{N \to \infty}"] \arrow[r, bend right, "\text{then yes}"'] & \boxed{f(x^*) = 0} \ \text{fixed points}
\end{tikzcd}
$$

Fixed points?

$$
\begin{tikzcd}
 & \text{Markov chain} & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[r, "N \to \infty"] \arrow[d, "t \to \infty"] & \dot{x} = f(x) \arrow[d, "\cancel{t \to \infty}"'] \arrow[d, bend left, "\text{if yes}"] \\
\text{Stationary} & \pi K = 0 \arrow[r, "\cancel{N \to \infty}"] \arrow[r, bend right, "\text{then yes}"'] & \boxed{\begin{array}{c} f(x^*) = 0 \\ \text{fixed points} \end{array}}
\end{tikzcd}
$$

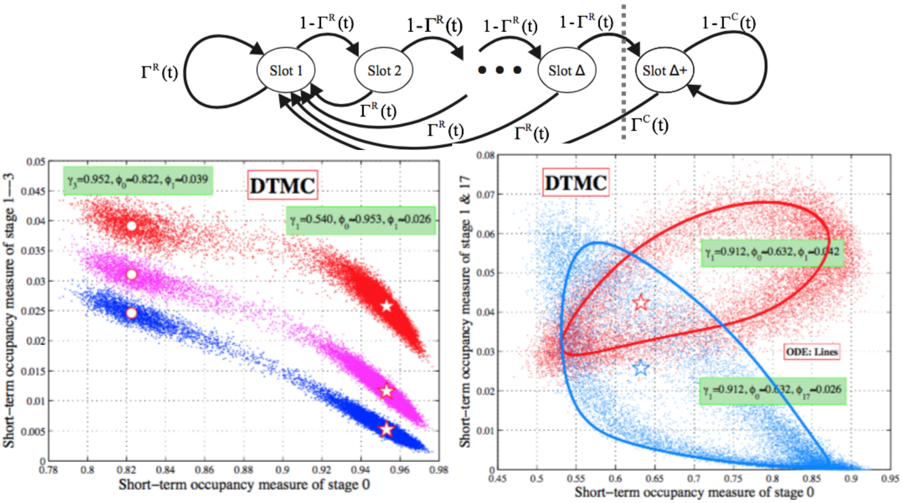
Theorem (Benaim Le Boudec 08)

If all trajectories of the ODE converges to the fixed points, the stationary distribution π^N concentrates on the fixed points

In that case, we also have:

$$\lim_{N\to\infty} \mathbf{P}\left[Z_1 = i_1 \dots Z_k = i_k\right] = x_1^* \dots x_k^*.$$

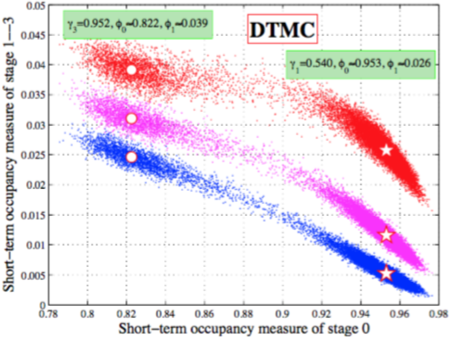
Example of 802.11[1]



[1] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

Quiz

Consider the 802.11 model:



Under the stationary distribution π^N:

(A) $P(Z_1 = 0, Z_2 = 0) \approx P(Z_1 = 0)P(Z_2 = 0)$

(B) $P(Z_1 = 0, Z_2 = 0) > P(Z_1 = 0)P(Z_2 = 0)$

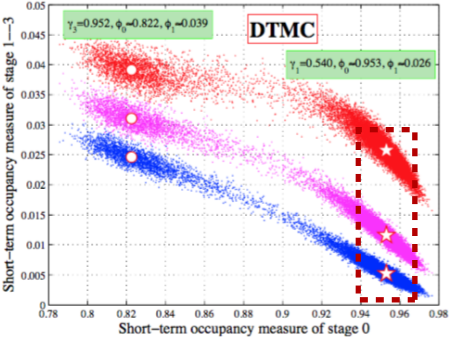
(C) $P(Z_1 = 0, Z_2 = 0) < P(Z_1 = 0)P(Z_2 = 0)$

(D) There is no stationary distribution

(E) I do not know

Quiz

Consider the 802.11 model:



Under the stationary distribution π^N:

(A) $P(Z_1 = 0, Z_2 = 0) \approx P(Z_1 = 0)P(Z_2 = 0)$

(B) $P(Z_1 = 0, Z_2 = 0) > P(Z_1 = 0)P(Z_2 = 0)$

(C) $P(Z_1 = 0, Z_2 = 0) < P(Z_1 = 0)P(Z_2 = 0)$

(D) There is no stationary distribution

(E) I do not know

Answer: B

$P(Z_1(t) = 0, Z_2(t) = 0) = x_1(t)^2$. Thus: positively correlated.

Outline

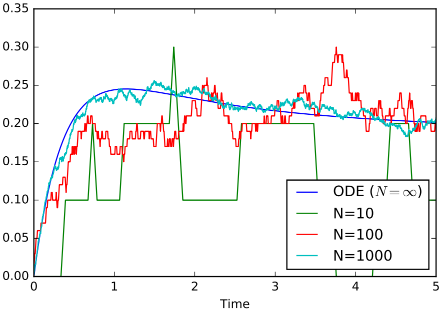
How accurate is mean-field approximation?

- $X_i^N(t)$ = fraction of objects in state i.

Theorem (Kurtz 70s')

When f is Lipschitz:

$$X^N(t) - x(t) = O\left(\frac{1}{\sqrt{N}}\right)$$



Example. Two-choice model, Fraction of servers with 3 jobs

In practice, we use mean-field for $N \geq 50$. Are we wrong?

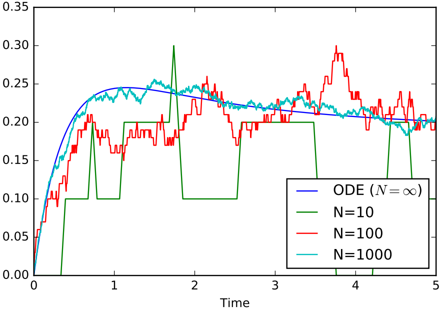
How accurate is mean-field approximation?

- $X_i^N(t)$ = fraction of objects in state i.

Theorem (Kurtz 70s')

When f is Lipschitz:

$$X^N(t) - x(t) = O\left(\frac{1}{\sqrt{N}}\right)$$



Example. Two-choice model, Fraction of servers with 3 jobs

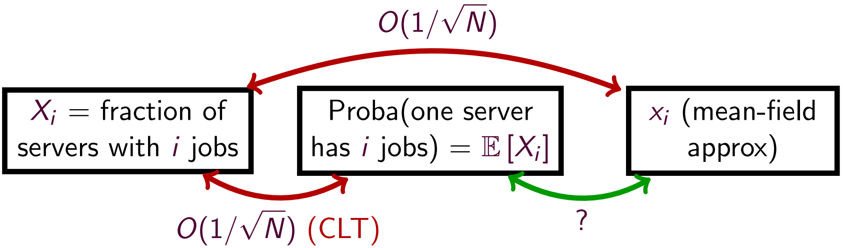
In practice, we use mean-field for $N \geq 50$. Are we wrong?

N	10	100	1000	$+\infty$
Average queue length (m^N)	3.81	3.39	3.36	3.35
Error ($m^N - m^\infty$)	0.45	0.039	0.004	0

Where is the catch?

$$\boxed{X_i = \text{fraction of servers with } i \text{ jobs}} \xleftrightarrow{O(1/\sqrt{N})} \boxed{x_i \text{ (mean-field approx)}}$$

Where is the catch?



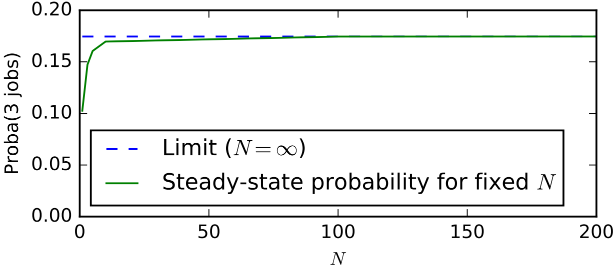
Where is the catch?

$O(1/\sqrt{N})$

X_i = fraction of servers with i jobs	Proba(one server has i jobs) = $\mathbb{E}[X_i]$	x_i (mean-field approx)

$O(1/\sqrt{N})$ (CLT)

$\boxed{O(1/N)}$



Numerical example : steady-state probability of having 3 jobs.

Transient regime

Theorem

If f differentiable and if Df is Lipschitz-continuous, then there exists a constant $C(t)$ such that:

$$\left|\mathbb{E}\left[X^N(t)\right] - x(t)\right| \leq \frac{C(t)}{N}.$$

The classical result only requires f to be Lipschitz-continuous and implies

$$\mathbb{E}\left[\left\|X^N(t) - x(t)\right\|\right] \leq \frac{C'(t)}{\sqrt{N}}.$$

Steady-state analysis

We say that $\dot{x} = f(x)$ has an exponentially stable attractor x^* if for any solution:

$$\|x(t) - x^*\| \leq Ce^{-\alpha t} \|x(0) - x^*\| .$$

Steady-state analysis

We say that $\dot{x} = f(x)$ has an exponentially stable attractor x^* if for any solution:

$$\|x(t) - x^*\| \leq Ce^{-\alpha t} \|x(0) - x^*\| .$$

Theorem

If f differentiable, Df is Lipschitz-continuous and the ODE has an exponentially stable attractor x^*, then there exists a constant C such that:

$$\left| \mathbb{E}\left[X^N\right] - x^* \right| \leq \frac{C}{N}.$$

Idea of the proof

We study:

$$\mathbb{E}\left[X^N(t)\right] - x(t) = \int_0^t \frac{d}{ds}\mathbb{E}\left[X^N(t) \mid X^N(s) = x(s)\right] ds.$$

Idea of the proof

We study:

$$\mathbb{E}\left[X^N(t)\right] - x(t) = \int_0^t \frac{d}{ds}\mathbb{E}\left[X^N(t) \mid X^N(s) = x(s)\right] ds.$$
$$= \int_0^t \frac{d}{ds}\Psi_{t-s}^{(N)}\Phi_s ds$$

where

$$\Psi_t^{(N)}h(x) = \mathbb{E}\left[h(X^N(t)) \mid X^N(0) = x\right] \qquad \Phi_t h(x) = h(\Phi_t x)$$

Idea of the proof

We study:

$$\begin{aligned}\mathbb{E}\left[X^N(t)\right] - x(t) &= \int_0^t \frac{d}{ds}\mathbb{E}\left[X^N(t) \mid X^N(s) = x(s)\right] ds.\\ &= \int_0^t \frac{d}{ds}\Psi^{(N)}_{t-s}\Phi_s ds\\ &= \int_0^t \Psi^{(N)}_{t-s}(\Lambda - L^{(N)})\Phi_s ds,\end{aligned}$$

where

$$\Psi^{(N)}_t h(x) = \mathbb{E}\left[h(X^N(t)) \mid X^N(0) = x\right] \qquad \Phi_t h(x) = h(\Phi_t x)$$

$$L^{(N)} h(x) = \sum_{\ell\in\mathcal{L}} N\beta_\ell(x)(h(x+\frac{\ell}{N}) - h(x)) \qquad \Lambda h(x) = Dh(x)\cdot f(x)$$

Idea of the proof

We study:

$$\begin{aligned}
\mathbb{E}\left[X^N(t)\right] - x(t) &= \int_0^t \frac{d}{ds}\mathbb{E}\left[X^N(t) \mid X^N(s) = x(s)\right] ds.\\
&= \int_0^t \frac{d}{ds}\Psi_{t-s}^{(N)}\Phi_s ds\\
&= \int_0^t \Psi_{t-s}^{(N)}(\Lambda - L^{(N)})\Phi_s ds,
\end{aligned}$$

where

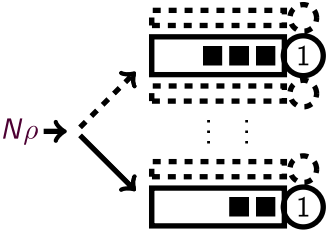
$$\Psi_t^{(N)}h(x) = \mathbb{E}\left[h(X^N(t)) \mid X^N(0) = x\right] \qquad \Phi_t h(x) = h(\Phi_t x)$$

$$L^{(N)}h(x) = \sum_{\ell\in\mathcal{L}} N\beta_\ell(x)(h(x + \frac{\ell}{N}) - h(x)) \qquad \Lambda h(x) = Dh(x)\cdot f(x)$$

We then obtain a $O(1/N)$ convergence if $\int_0^t D\Phi_s ds$ exists and is Lipschitz-continuous with respect to the initial condition (also works for steady-state).

Outline

The two choice model[2]



Randomly choose two, and select one

Infinite state-space:

$$X_0(t), X_1(t), \ldots$$

where

$X_i(t) =$ fraction with i or more jobs.

[2] This model or variants have been heavily studied (Vvedenskaya 96, Mitzenmacher 98, ... Fricker G. 2014, Tsitsiklis 2016).

Why is this called the power of two-choices?

As N goes to infinity, in steady-state,

$$\lim_{N\to\infty} X_i^N = \rho^{2^i-1}$$

The average queue length $m^N(\rho)$ satisfies:

$$\lim_{N\to\infty} m^N(\rho) = m^\infty(\rho) = \Theta_{\rho\to 1}\Big(\log\frac{1}{1-\rho}\Big)$$

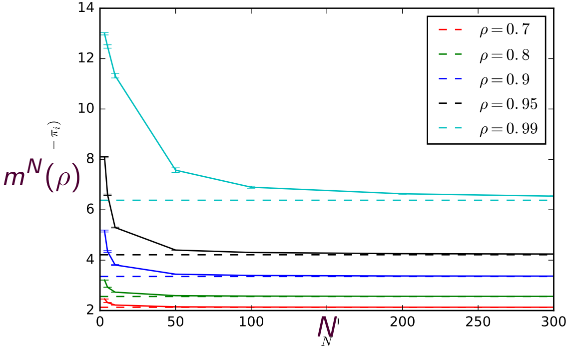
One-choice

$$\rho^i$$

$$\frac{1}{1-\rho}$$

Our result shows that $\limsup_{N\to\infty} N\left|m^N(\rho) - m^\infty(\rho)\right| < \infty$.

Can we quantify the $O(1/N)$?

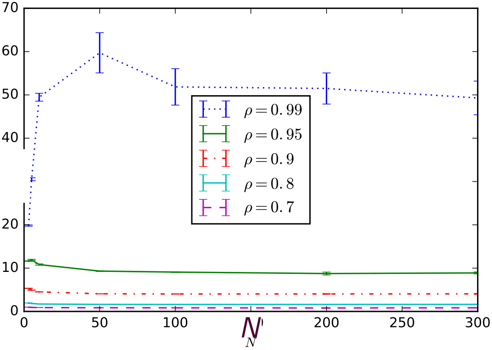


In particular, the average queue length satisfies:

$$m^N(\rho) = \Theta_{\rho \to 1}\left(\log \frac{1}{1-\rho}\right) + O(1/N)$$

Can we quantify the $O(1/N)$?

$N(m^N(\rho) - m^\infty(\rho))$



In particular, the average queue length satisfies:

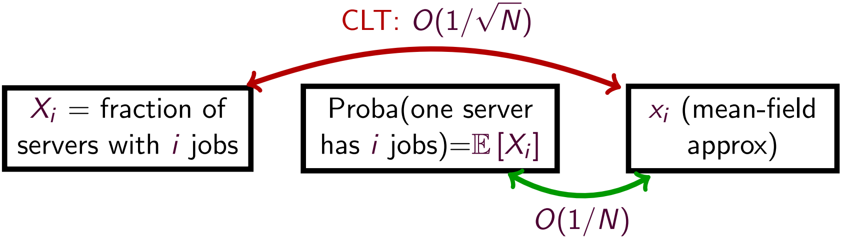
$$m^N(\rho) = \Theta_{\rho\to 1}\Big(\log\frac{1}{1-\rho}\Big) + \frac{1}{N}\underbrace{\Theta_{\rho\to 1}\Big(\frac{1}{1-\rho}\Big)}_{\text{order of magnitude larger}} + o\Big(\frac{1}{N}\Big),$$

Outline

Recap

1. Convergence of mean-field model is $O(1/N)$.
   - Works for transient and steady-state
   - Works for infinite-dimensional state space.

2. Our approach is to focus on the expected values



Extension and open questions

1. Technical question:
   - Can we compute the constant in $O(1/N)$?
   - Steady-state + only Lipschitz-continuous: is the convergence rate $O(1/\sqrt{N})$?
2. Hitting/mixing-time + fluid approximation.
3. Non-homogeneous population.
   - *e.g.*, caching

Thank you!

http://mescal.imag.fr/membres/nicolas.gast

nicolas.gast@inria.fr

Mean-field and decoupling

Benaïm, Le Boudec 08 — *A class of mean field interaction models for computer and communication systems*, M.Benaïm and J.Y. Le Boudec., Performance evaluation, 2008.

Le Boudec 10 — *The stationary behaviour of fluid limits of reversible processes is concentrated on stationary points.*, J.-Y. L. Boudec. , Arxiv:1009.5021, 2010

Darling Norris 08 — *R. W. R. Darling and J. R. Norris*, Differential equation approximations for Markov chains, Probability Surveys 2008

G. 16 — *Construction of Lyapunov functions via relative entropy with application to caching*, Gast, N., ACM MAMA 2016

Budhiraja et al. 15 — *Limits of relative entropies associated with weakly interacting particle systems.*, A. S. Budhiraja, P. Dupuis, M. Fischer, and K. Ramanan. , Electronic journal of probability, 20, 2015.

References (continued)

Optimal control and mean-field games:

G.,Gaujal Le Boudec 12 — *Mean field for Markov decision processes: from discrete to continuous optimization*, N.Gast,B.Gaujal,J.Y.Le Boudec, IEEE TAC, 2012

G. Gaujal 12 — *Markov chains with discontinuous drifts have differential inclusion limits.*, Gast N. and Gaujal B., Performance Evaluation, 2012

Puterman — *Markov decision processes: discrete stochastic dynamic programming*, M.L. Puterman, John Wiley & Sons, 2014.

Lasry Lions — *Mean field games*, J.-M. Lasry and P.-L. Lions, Japanese Journal of Mathematics, 2007.

Tembine at al 09 — *Mean field asymptotics of markov decision evolutionary games and teams*, H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman., GameNets 00

Applications: caches, bikes

Don and Towsley — *An approximate analysis of the LRU and FIFO buffer replacement schemes*, A. Dan and D. Towsley., SIGMETRICS 1990

G. Van Houdt 15 — *Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms.*, Gast, Van Houdt., ACM Sigmetrics 2015

Fricker-Gast 14 — *Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity.*, C. Fricker and N. Gast. , EJTL, 2014.

Fricket et al. 13 — *Mean field analysis for inhomogeneous bike sharing systems*, Fricker, Gast, Mohamed, Discrete Mathematics and Theoretical Computer Science DMTCS

G. et al 15 — *Probabilistic forecasts of bike-sharing systems for journey planning*, N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, CIKM 2015