A Tutorial on Mean Field and Refined Mean Field Approximation

Nicolas Gast

Inria, Grenoble, France

YEQT XI, December 2018, Toulouse

Good system design needs performance evaluation

Example : load balancing

N servers

Which allocation policy?

- Random
- Round-robin
- *JSQ*
- *JSQ*(d)
- *JIQ*

Good system design needs performance evaluation

Example : load balancing

N servers

Which allocation policy?

- Random
- Round-robin
- *JSQ*
- *JSQ*(d)
- *JIQ*

We need methods to characterize emerging behavior starting from a stochastic model of interacting objects

- We use ~~simulation~~ analytical methods and approximations.

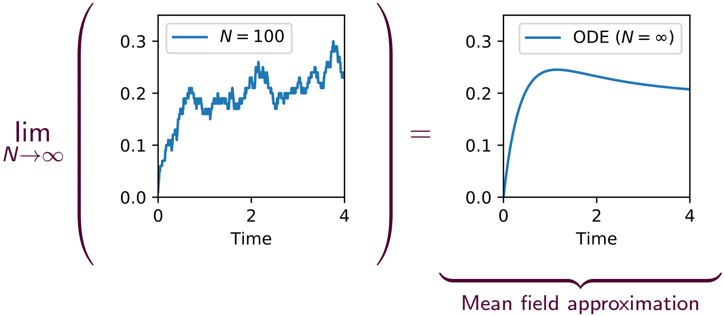
The main difficulty of probability : correlations

$$\mathbf{P}[A, B] \neq \mathbf{P}[A]\,\mathbf{P}[B]$$

Problem: state space explosion
S states per object, N objects $\Rightarrow S^N$ states

“Mean field approximation” simplifies many problems

But how to apply it?



Where has it been used?

- Performance of load balancing / caching algorithms
- Communication protocols (CSMA, MPTCP, Simgrid)
- Mean field games (evacuation, Mexican wave)
- Stochastic approximation / learning
- Theoretical biology

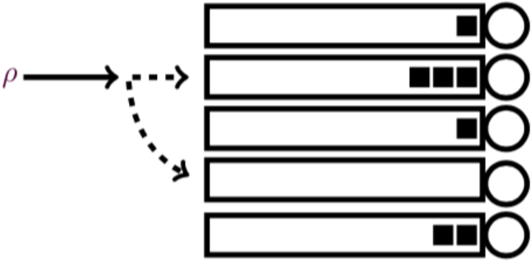
Outline: Demystifying Mean Field Approximation

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

The supermarket model (SQ(2))



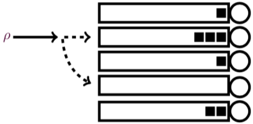
Arrival at each server ρ.

- Sample $d-1$ other queues.
- Allocate to the shortest queue

Service rate$=1$.

$SQ(d)$: state representation

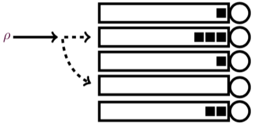
- Let $S_n(t)$ be the queue length of the nth queue at time t.



$$S = (1, 3, 1, 0, 2)$$

$SQ(d)$: state representation

- Let $S_n(t)$ be the queue length of the nth queue at time t.



$$S = (1, 3, 1, 0, 2)$$

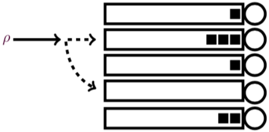
- Alternative representation:

$$X_i(t) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{\{S_n(t) \geq i\}},$$

which is the fraction of queues with queue length $\geq i$.

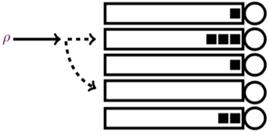
$$X = (1, 0.8, 0.4, 0.2, 0, 0, 0, \dots)$$

$SQ(d)$: state transitions



- Arrival: $x \mapsto x + \frac{1}{N}\mathbf{e_i}$.
- Departures: $x \mapsto x - \frac{1}{N}\mathbf{e_i}$.

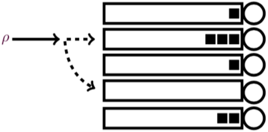
$SQ(d)$: state transitions



- Arrival: $x \mapsto x + \dfrac{1}{N}\mathbf{e_i}$.
- Departures: $x \mapsto x - \dfrac{1}{N}\mathbf{e_i}$.

Recall that x_i is the fraction of servers with i jobs or more. Pick two servers at random, what is the probability the least loaded has $i-1$ jobs?

$SQ(d)$: state transitions



- Arrival: $x \mapsto x + \frac{1}{N}\mathbf{e_i}$.
- Departures: $x \mapsto x - \frac{1}{N}\mathbf{e_i}$.

Recall that x_i is the fraction of servers with i jobs or more. Pick two servers at random, what is the probability the least loaded has $i-1$ jobs?

$$x_{i-1}^2 - x_i^2 \qquad \text{when picked with replacement}$$

$$x_{i-1}\frac{Nx_{i-1}-1}{N-1} - x_i\frac{Nx_i-1}{N-1} \qquad \text{when picked without replacement}$$

Note: this becomes asymptotically the same as N goes to infinity.

Transitions and Mean Field Approximation

State changes on x:

$$x \mapsto x + \frac{1}{N}\mathbf{e_i} \text{ at rate } N\rho(x_{i-1}^d - x_i^d)$$
$$x \mapsto x - \frac{1}{N}\mathbf{e_i} \text{ at rate } N(x_i - x_{i+1})$$

The mean field approximation is to consider the ODE associated with the drift (average variation):

$$\dot{x}_i = \underbrace{\rho(x_{i-1}^d - x_i^d)}_{\text{Arrival}} - \underbrace{(x_i - x_{i+1})}_{\text{Departure}}$$

Variants: push-pull model, centralized solution

Suppose that:

- At rate r, each server that has $i \geq 2$ or more jobs probes a server and pushes a job to it if this server has 0 jobs. Transitions are:

$$x \mapsto x + \frac{1}{N}(-e_i + e_1) \text{ at rate } Nr(x_{i-1} - x_i)(1 - x_1)$$

Variants: push-pull model, centralized solution

Suppose that:

- At rate r, each server that has $i \geq 2$ or more jobs probes a server and pushes a job to it if this server has 0 jobs. Transitions are:
$$x \mapsto x + \frac{1}{N}(-e_i + e_1) \text{ at rate } Nr(x_{i-1} - x_i)(1 - x_1)$$
- At rate $N\gamma$, a centralized server serves a job from the longests queue. Transitions is:
$$x \mapsto x - \frac{1}{N}e_i \text{ at rate } N\gamma x_i \mathbf{1}_{\{x_{i+1}=0\}}$$

Variants: push-pull model, centralized solution

Suppose that:

- At rate r, each server that has $i \geq 2$ or more jobs probes a server and pushes a job to it if this server has 0 jobs. Transitions are:

$$x \mapsto x + \frac{1}{N}(-e_i + e_1) \text{ at rate } Nr(x_{i-1} - x_i)(1 - x_1)$$

- At rate $N\gamma$, a centralized server serves a job from the longests queue. Transitions is:

$$x \mapsto x - \frac{1}{N}e_i \text{ at rate } N\gamma x_i \mathbf{1}_{\{x_{i+1}=0\}}$$

The mean field approximation becomes (for $i > 1$):

$$\dot{x}_i = \underbrace{\rho(x_{i-1}^d - x_i^d)}_{\text{Arrival}} - \underbrace{(x_i - x_{i+1})}_{\text{Departure}} - \underbrace{r(x_{i-1} - x_i)(1 - x_1)}_{\text{Push}} - \underbrace{N\gamma x_i \mathbf{1}_{\{x_{i+1}=0\}}}_{\text{Centralized}}$$

$$\dot{x}_1 = \underbrace{\rho(x_0^d - x_1^d)}_{\text{Arrival}} - \underbrace{(x_1 - x_2)}_{\text{Departure}} + \sum_{i=2}^{\infty} \underbrace{r(x_{i-1} - x_i)(1 - x_1)}_{\text{Push}} - \underbrace{N\gamma x_1 \mathbf{1}_{\{x_2=0\}}}_{\text{Centralized}}$$

Density dependent population process (Kurtz, 70s)

A population process is a sequence of CTMCs $X^N(t)$ indexed by the population size N, with state space $E^N \subset E$ and transitions (for $\ell \in \mathcal{L}$):

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

Density dependent population process (Kurtz, 70s)

A population process is a sequence of CTMCs $X^N(t)$ indexed by the population size N, with state space $E^N \subset E$ and transitions (for $\ell \in \mathcal{L}$):

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

The Mean field approximation

The drift is $f(x) = \sum_\ell \ell\beta_\ell(x)$ and the mean field approximation is the solution of the ODE $\dot{x} = f(x)$.

Density dependent population process (Kurtz, 70s)

A population process is a sequence of CTMCs $X^N(t)$ indexed by the population size N, with state space $E^N \subset E$ and transitions (for $\ell \in \mathcal{L}$):

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

The Mean field approximation

The drift is $f(x) = \sum_\ell \ell\beta_\ell(x)$ and the mean field approximation is the solution of the ODE $\dot{x} = f(x)$.

Example: SQ(d) load balancing

$$\dot{x}_i = \rho(x_{i-1}^d - x_i^d) - (x_i - x_{i+1})$$

It has a unique attractor: $\pi_i = \rho^{(d^i-1)/(d-1)}$.

Accuracy of the mean field approximation

Numerical example of SQ(d) load balancing ($d = 2$)

N	Simulation (steady-state average queue length)					Fixed Point
	10	20	30	50	100	∞ (mean field)
$\rho = 0.7$	1.2194	1.1735	1.1584	1.1471	1.1384	1.1301
$\rho = 0.9$	2.8040	2.5665	2.4907	2.4344	2.3931	2.3527
$\rho = 0.95$	4.2952	3.7160	3.5348	3.4002	3.3047	3.2139

Fairly good accuracy for $N = 100$ servers.

Accuracy of the mean field approximation

Numerical example of SQ(d) load balancing ($d = 2$)

	Simulation (steady-state average queue length)					Fixed Point
N	10	20	30	50	100	∞ (mean field)
$\rho = 0.7$	1.2194	1.1735	1.1584	1.1471	1.1384	1.1301
$\rho = 0.9$	2.8040	2.5665	2.4907	2.4344	2.3931	2.3527
$\rho = 0.95$	4.2952	3.7160	3.5348	3.4002	3.3047	3.2139

Fairly good accuracy for $N = 100$ servers.

Accuracy of the mean field approximation

Pull-push model (servers with ≥ 2 jobs push to empty)

	Simulation (steady-state ave. queue length)				Fixed point
N	10	20	50	100	∞
$\rho = 0.8$	1.5569	1.4438	1.3761	1.3545	1.3333
$\rho = 0.90$	2.3043	1.9700	1.7681	1.7023	1.6364
$\rho = 0.95$	3.4288	2.6151	2.1330	1.9720	1.8095

Fairly good accuracy for $N = 100$ servers.

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Examples: the cache-replacement policy RAND

Model: There are n objects and a cache of size m.

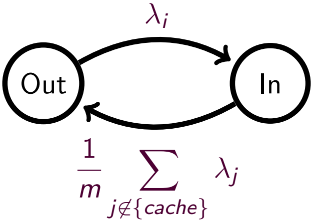
- Objects i is requested according to a Poisson process of intensity λ_i.
- An requested object that is not the cache goes into the cache and ejects a random object.

Examples: the cache-replacement policy RAND

Model: There are n objects and a cache of size m.

- Objects i is requested according to a Poisson process of intensity λ_i.
- An requested object that is not the cache goes into the cache and ejects a random object.

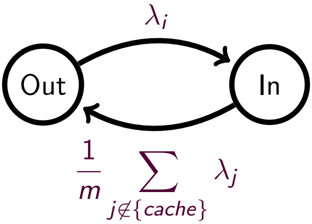
The state of object i is $\{\text{Out},\text{In}\}$.



Extension: list-based caching (G. Van Houdt, Sigmetrics 2015)

RAND: mean field approximation

Original model



RAND: mean field approximation

Original model

$$\begin{tikzcd} \text{Out} \arrow[r, bend left, "\lambda_i"] & \text{In} \arrow[l, bend left, "\frac{1}{m}\sum_{j\notin\{cache\}}\lambda_j"] \end{tikzcd}$$

MF approx: let $x_i(t) = \mathbf{P}\left[i \notin \{cache\}\right]$.
If all objects are independent:

$$\begin{tikzcd} \text{Out} \arrow[r, bend left, "\lambda_i"] & \text{In} \arrow[l, bend left, "\frac{1}{m}\sum_{j=1}^{n} x_j\lambda_j"] \end{tikzcd}$$

The "mean field" equations for the approximation model are:

$$\dot{x}_i = -\lambda_i x_i + \frac{1}{m}\sum_{j=1}^{n} x_j(t)\lambda_j(1 - x_i).$$

RAND: mean field approximation

Original model

$$\begin{tikzcd} \text{Out} \arrow[r, bend left, "\lambda_i"] & \text{In} \arrow[l, bend left, "\frac{1}{m}\sum_{j\notin\{cache\}}\lambda_j"] \end{tikzcd}$$

MF approx: let $x_i(t) = \mathbf{P}\left[i \notin \{cache\}\right]$.
If all objects are independent:

$$\begin{tikzcd} \text{Out} \arrow[r, bend left, "\lambda_i"] & \text{In} \arrow[l, bend left, "\frac{1}{m}\sum_{j=1}^{n} x_j\lambda_j"] \end{tikzcd}$$

The "mean field" equations for the approximation model are:

$$\dot{x}_i = -\lambda_i x_i + \frac{1}{m}\sum_{j=1}^{n} x_j(t)\lambda_j(1-x_i).$$

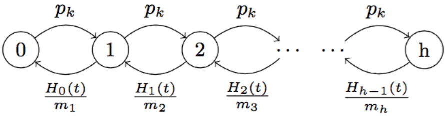
It has a unique fixed point that satisfies:

$$\pi_i = \frac{z}{z+\pi_i} \qquad \text{with } z \text{ such that } \sum_{i=1}^{n}(1-\pi_i) = m.$$

Same equations as Fagins (77).

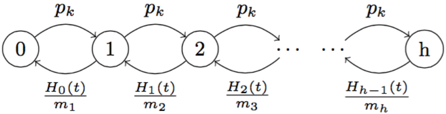
Extension to the RAND(m) model (G, Van Houdt SIGMETRICS 2015)

Let $H_i(t)$ be the popularity in list i.



Extension to the RAND(m) model (G, Van Houdt SIGMETRICS 2015)

Let $H_i(t)$ be the popularity in list i.



If $x_{k,i}(t)$ is the probability that item k is in list i at time t, we approximately have:

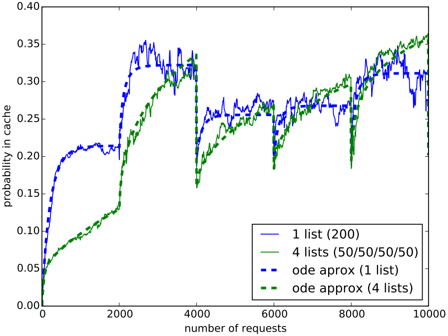
$$\dot{x}_{k,i}(t) = p_k x_{k,i-1}(t) - \overbrace{\boxed{\sum_j p_j x_{j,i-1}(t)}}^{\text{Popularity in cache i-1}} \frac{x_{k,i}(t)}{m_i}$$

$$+ \mathbf{1}_{\{i<h\}} \Bigg(\underbrace{\boxed{\sum_j p_j x_{j,i}(t)}}_{\text{Popularity in cache i}} \frac{x_{k,i+1}(t)}{m_{i+1}} - p_k x_{k,i}(t) \Bigg)$$

This approximation is of the form $\dot{x} = xQ(x)$.

The mean field approximation is very accurate

$n = 1000$ objects with Zipf popularities.



The popularities change every 2000 requests

m_1	m_2	m_3	m_4	exact	mean field
2	2	96	–	0.3166	0.3169
10	30	60	–	0.3296	0.3299
20	2	78	–	0.3273	0.3276
90	8	2	–	0.4094	0.4100
1	4	10	85	0.3039	0.3041
5	15	25	55	0.3136	0.3139
25	25	25	25	0.3345	0.3348
60	2	2	36	0.3514	0.3517

Steady-state miss probabilities

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Benaïm-Le Boudec's model (PEVA 2007)

Time is discrete.

$$X_i(k) = \text{Proportion of object in state } i \text{ at time step } k$$

$$R(k) = \text{State of the "resource" at time } k \text{ (discrete)}$$

Benaïm-Le Boudec's model (PEVA 2007)

Time is discrete.

$$X_i(k) = \text{Proportion of object in state } i \text{ at time step } k$$
$$R(k) = \text{State of the "resource" at time } k \text{ (discrete)}$$

Assumptions:

- Only $O(1)$ objects change state at each time step and

$$f(x,r) = \frac{1}{N}\mathbb{E}\left[X(k+1) - X(k) | X(k) = x, R(k) = r\right].$$

- R evolves fast in a discrete state-space and:

$$\mathbf{P}\left[R(k+1) = j | X(k) = x, R(k) = i\right] = P_{ij}(x).$$

For all x, $P(x)$ is irreducible and has a unique stationary measure $\pi(x,.)$.

Mean Field Approximation

Examples with resource: CSMA protocols, Opportunistic networks.

$$\dot{x} = \sum_r f(x,r)\pi(x,r),$$

where $\pi(x,r)$ is the stationary measure of the resource given x.

Mean Field Approximation

Examples with resource: CSMA protocols, Opportunistic networks.

$$\dot{x} = \sum_r f(x,r)\pi(x,r),$$

where $\pi(x,r)$ is the stationary measure of the resource given x.

The analysis of such models is done by considering stochastic approximation algorithms. For example, without resource one has:

$$X(k+1) = X(k) + \frac{1}{N}\left[f(X(k)) + M(k+1)\right],$$

where M is some noise process.
This is a noisy Euler discretization of an ordinary differential equation.

Take-home message on this part

Three ways to construct mean field approximation:

- Density dependent population process.
- Independence assumption $\dot{x} = xQ(x)$.
- Discrete-time model with vanishing intensity.

In what follows, I will assume that X is a density dependent population process (ex: $SQ(d)$, pull-push). Analysis of other models are similar.

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Convergence Result as N Goes to Infinity

Theorem (under some mild conditions, mostly Lipschitz continuity): If $X^N(0)$ converges to x_0, then for any finite T:

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

where $x(t)$ is the unique solution of the ODE $\dot{x} = f(x)$.

Illustration: An Infection Model

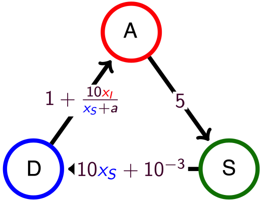
Nodes can be **D**ormant, **A**ctive or **S**usceptible.

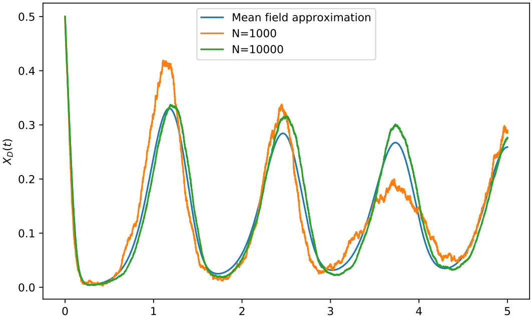
	Transition	Rate
Activation	$(D, A, S) \mapsto (D - \frac{1}{N}, A + \frac{1}{N}, S)$	$N(0.15 + 10X_A)X_D$
Immunization	$(D, A, S) \mapsto (D, A - \frac{1}{N}, S + \frac{1}{N})$	$N5X_A$
De-immunization	$(D, A, S) \mapsto (D + \frac{1}{N}, A, S - \frac{1}{N})$	$N(1 + \frac{10X_A}{X_D + \delta})X_S$

Illustration: An Infection Model

Nodes can be **D**ormant, **A**ctive or **S**usceptible.

	Transition	Rate
Activation	$(D,A,S) \mapsto (D-\frac{1}{N}, A+\frac{1}{N}, S)$	$N(0.15+10X_A)X_D$
Immunization	$(D,A,S) \mapsto (D, A-\frac{1}{N}, S+\frac{1}{N})$	$N5X_A$
De-immunization	$(D,A,S) \mapsto (D+\frac{1}{N}, A, S-\frac{1}{N})$	$N(1+\frac{10X_A}{X_D+\delta})X_S$





The fixed point method

Markov chain

Transient regime	$\dot{p} = pK$
	$\downarrow t \to \infty$
Stationary	$\pi K = 0$

The fixed point method

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

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 the fixed point method always work?

	Transition	Rate
Activation	$(D,A,S) \mapsto (D-\frac{1}{N}, A+\frac{1}{N}, S)$	$N(a+10X_A)X_D$
Immunization	$(D,A,S) \mapsto (D, A-\frac{1}{N}, S+\frac{1}{N})$	$N5X_A$
De-immunization	$(D,A,S) \mapsto (D+\frac{1}{N}, A, S-\frac{1}{N})$	$N(1+\frac{10X_A}{X_D+\delta})X_S$

- Markov chain is irreducible
- Mean field approximation has a unique fixed point $xQ(x) = 0$.

Does the fixed point method always work?

	Transition	Rate
Activation	$(D, A, S) \mapsto (D - \frac{1}{N}, A + \frac{1}{N}, S)$	$N(a + 10X_A)X_D$
Immunization	$(D, A, S) \mapsto (D, A - \frac{1}{N}, S + \frac{1}{N})$	$N5X_A$
De-immunization	$(D, A, S) \mapsto (D + \frac{1}{N}, A, S - \frac{1}{N})$	$N(1 + \frac{10X_A}{X_D + \delta})X_S$

- Markov chain is irreducible
- Mean field approximation has a unique fixed point $xQ(x) = 0$.

	Fixed point $xQ(x) = 0$		Stat. measure (simulation)		
	π_D	π_A	π_D	π_A	
$a = .3$	0.211	0.241	0.219	0.242	$(N = 10^3)$
			0.212	0.242	$(N = 10^4)$

Does the fixed point method always work?

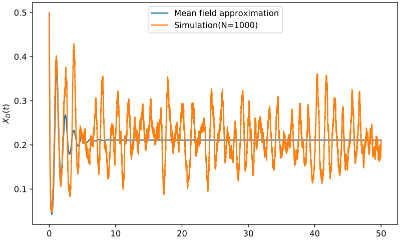
	Transition	Rate
Activation	$(D, A, S) \mapsto (D - \frac{1}{N}, A + \frac{1}{N}, S)$	$N(a + 10X_A)X_D$
Immunization	$(D, A, S) \mapsto (D, A - \frac{1}{N}, S + \frac{1}{N})$	$N5X_A$
De-immunization	$(D, A, S) \mapsto (D + \frac{1}{N}, A, S - \frac{1}{N})$	$N(1 + \frac{10X_A}{X_D + \delta})X_S$

- Markov chain is irreducible
- Mean field approximation has a unique fixed point $xQ(x) = 0$.

	Fixed point $xQ(x) = 0$		Stat. measure (simulation)		
	π_D	π_A	π_D	π_A	
$a = .3$	0.211	0.241	0.219	0.242	$(N = 10^3)$
			0.212	0.242	$(N = 10^4)$
$a = .15$	0.115	0.177	0.154	0.197	$N = 10^3$
			0.151	0.195	$N = 10^4$

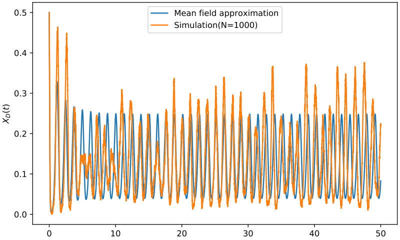
What happened?

$a = 0.30$



Fixed point = attractor
Fixed point method works!

$a = 0.15$



ODE has a cyclic behavior
Fixed point method does not work.

Convergence result (steady-state)

Theorem If the mean field approximation has a unique attractor $x(\infty)$, then

$$\left\| X^N(\infty) - x(\infty) \right\| \to 0$$

Fixed points?

Markov chain

Transient regime $\dot{p} = pK$

$t \to \infty$

Stationary $\pi K = 0$

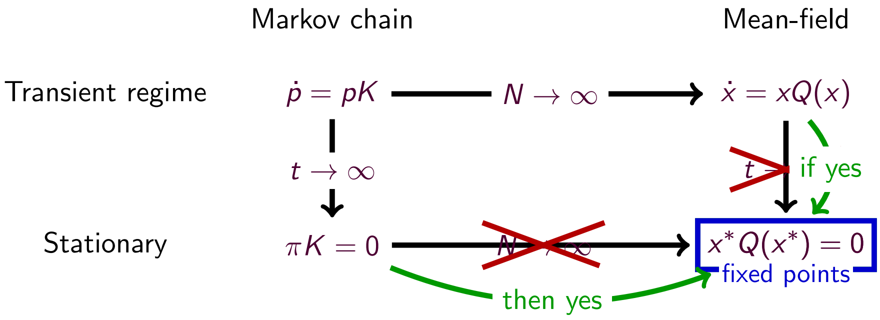
Fixed points?

	Markov chain		Mean-field
Transient regime	$\dot{p} = pK$	$\xrightarrow{N \to \infty}$	$\dot{x} = xQ(x)$
	$\downarrow t \to \infty$		
Stationary	$\pi K = 0$	$\xrightarrow{?}$	$\underset{\text{fixed points}}{\boxed{x^* Q(x^*) = 0}}$

Fixed points?

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

Fixed points?



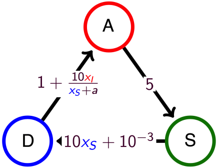
Theorem (Benaim Le Boudec 08)

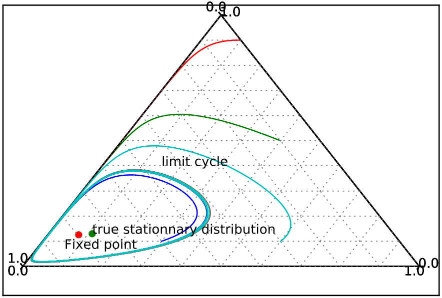
If all trajectories of the ODE converges to the fixed points, the stationary distribution π^N concentrates on the fixed points

In that case, we also have:

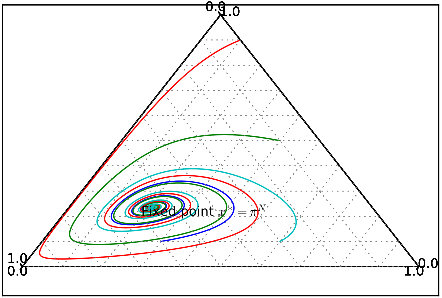
$$\lim_{N\to\infty} \mathbf{P}\left[S_1 = i_1 \dots S_k = i_k\right] = x_1^* \dots x_k^*.$$

Steady-state: illustration





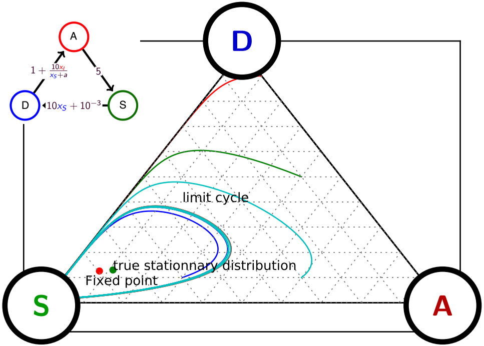
$a = .1$



$a = .3$

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As the trajectory converge to a fixed point, there is no such stationary distribution.

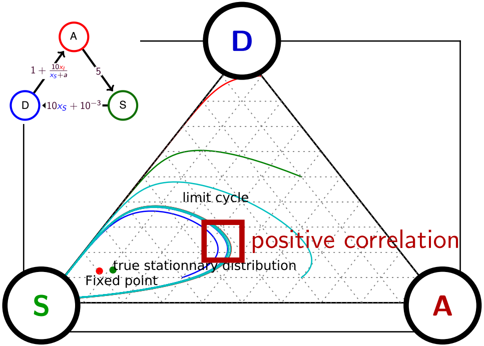
(B) $P(S_1 = S, S_2 = S) \approx P(S_1 = S)P(S_2 = S)$

(C) $P(S_1 = S, S_2 = S) > P(S_1 = S)P(S_2 = S)$

(D) $P(S_1 = S, S_2 = S) < P(S_1 = S)P(S_2 = S)$

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As the trajectory converge to a fixed point, there is no such stationary distribution.

(B) $P(S_1 = S, S_2 = S) \approx P(S_1 = S)P(S_2 = S)$

(C) $P(S_1 = S, S_2 = S) > P(S_1 = S)P(S_2 = S)$

(D) $P(S_1 = S, S_2 = S) < P(S_1 = S)P(S_2 = S)$

Answer: C

$P(S_1(t) = S, S_2(t) = S) = x_1(t)^2$. Thus: positively correlated.

How to show that trajectories converge to a fixed point?

Main solutions:

- Find a Lyapunov function
  - How to find a Lyapunov function: Energy? Entropy? Luck? (ex: G. 2016 for cache)
- Use reversibility (Le Boudec 2013)
- Monotonicity property (ex, load-balancing, see Van Houdt 2018)

Fixed point method in practice

From the examples coming from queuing theory, many models have a unique attractor.

- This holds for classical load balancing policies such as SQ(d), pull-push, JIQ,...
  - Often comes from monotonicity
- This holds in many cases in statistical physics
  - Lyapunov methods (entropy, reversibility)
- It does not always work
  - Theoretical biology / chemistry
  - Multi-stable models (ex: Kelly)
  - Counter-examples for specific CSMA models (Cho, Le Boudec, Jiang 2011)

Outline

1 Construction of the Mean Field Approximation: 3 models
- Density Dependent Population Processes
- A Second Point of View: Zoom on One Object
- Discrete-Time Models

2 On the Accuracy of Mean Field : Positive and Negative Results
- Transient Analysis
- Steady-state Regime

3 The Refined Mean Field
- Main Results
- Generator Comparison and Stein's Method
- Alternative View: System Size Expansion Approach

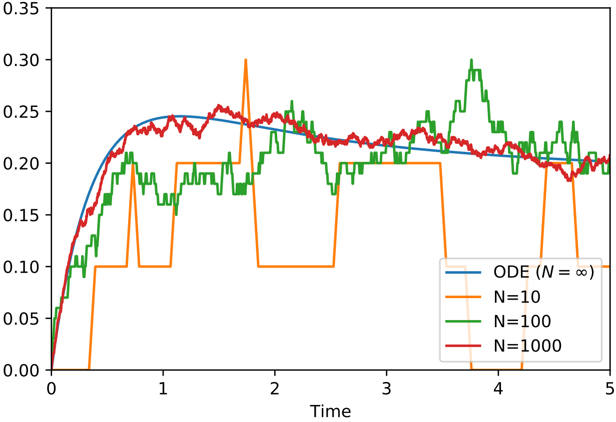
4 Demo

5 Conclusion and Open Questions

Mean Field Accuracy

Theorem (Kurtz (1970s), Ying (2016)):

If the drift f is Lipschitz-continuous:	If in addition the ODE has a unique attractor π:
$$X^N(t) \approx x(t) + \frac{1}{\sqrt{N}} G_t$$	$$\mathbb{E}\left[X^N(\infty) - \pi\right] = O(1/\sqrt{N})$$



Expected values estimated by mean field are 1/N-accurate

Some experiments (for SQ(2) with $\rho = 0.9$):

N	10	100	1000	∞
Average queue length (simulation)	2.8040	2.3931	2.3567	2.3527
Error of mean field	0.4513	0.0404	0.0040	0

Expected values estimated by mean field are 1/N-accurate

Some experiments (for SQ(2) with $\rho = 0.9$):

N	10	100	1000	∞
Average queue length (simulation)	2.8040	2.3931	2.3567	2.3527
Error of mean field	0.4513	0.0404	0.0040	0

Error seems to decrease as $1/N$

Expected values estimated by mean field are 1/N-accurate

Some experiments (for SQ(2) with $\rho = 0.9$):

N	10	100	1000	∞
Average queue length (simulation)	2.8040	2.3931	2.3567	2.3527
Error of mean field	0.4513	0.0404	0.0040	0

Error seems to decrease as $1/N$

Theorem (Kolokoltsov 2012, G. 2017& 2018). If the drift f is C^2 and has a unique exponentially stable attractor, then for any $t \in [0, \infty) \cup \{\infty\}$, there exists a constant V_t such that:

$$\mathbb{E}\left[h(X^N(t))\right] = h(x(t)) + \frac{V(t)}{N} + O(1/N^2)$$

The refined mean field approximation...

... is defined as the classic mean field plus the $1/N$ correction term:

$$\mathbb{E}\left[X^N\right] = x(t) + \frac{V(t)}{N},$$

where $V(t)$ is computed analytically.

The refined mean field approximation...

... is defined as the classic mean field plus the $1/N$ correction term:

$$\mathbb{E}\left[X^N\right] = x(t) + \frac{V(t)}{N},$$

where $V(t)$ is computed analytically.

To compute $V(t)$, we need:

- Derivative of the drifts:

$$F_j^i(t) = \frac{\partial f_i}{\partial x_j}(x(t)) \text{ and } F_{jk}^i(t) = \frac{\partial^2 f_i}{\partial x_j \partial x_k}(x(t))$$

- A variance term:

$$Q(t) = \sum_{\ell} \ell \otimes \ell \beta_{\ell}(X(t))$$

Computational methods

Theorem (G, Van Houdt 2018) Given a density dependent process with twice-differentiable drift. Let $h : E \to \mathbb{R}$ be a twice-differentiable function, then for $t > 0$:

$$\mathbb{E}\left[h(X^N(t))\right] = h(x(t)) + \frac{1}{N}\left(\sum_i \frac{\partial h(x(t))}{\partial x_i} V_i(t) + \frac{1}{2}\sum_{ij} \frac{h(x(t))}{\partial x_i \partial x_j} W_{ij}(t)\right) + O(\frac{1}{N^2}),$$

where

$$\frac{d}{dt}V^i = \sum_j F^i_j V^j + \sum_{jk} F^i_{j,k} W^{j,k}$$

$$\frac{d}{dt}W^{j,k} = Q^{jk} + \sum_m F^j_m W^{m,k} + \sum_m W^{j,m} F^k_m$$

Theorem (G, Van Houdt 2018) The previous theorem also holds for the stationary regime ($t = +\infty$) if the ODE has a unique exponentially stable attractor.

The supermarket model (SQ(2))

N	10	20	30	50	100	∞
$\rho = 0.7$						
Simulation	1.2194	1.1735	1.1584	1.1471	1.1384	–
Refined mf	1.2150	1.1726	1.1584	1.1471	1.1386	1.1301
$\rho = 0.9$						
Simulation	2.8040	2.5665	2.4907	2.4344	2.3931	–
Refined mf	2.7513	2.5520	2.4855	2.4324	2.3925	2.3527
$\rho = 0.95$						
Simulation	4.2952	3.7160	3.5348	3.4002	3.3047	–
Refined mf	4.1017	3.6578	3.5098	3.3915	3.3027	3.2139

Average queue length: Refined mean field approximation gives a significant improvement.

The supermarket model (SQ(2))

N	10	20	30	50	100	∞
$\rho = 0.7$						
Simulation	1.2194	1.1735	1.1584	1.1471	1.1384	–
Refined mf	1.2150	1.1726	1.1584	1.1471	1.1386	1.1301
$\rho = 0.9$						
Simulation	2.8040	2.5665	2.4907	2.4344	2.3931	–
Refined mf	2.7513	2.5520	2.4855	2.4324	2.3925	2.3527
$\rho = 0.95$						
Simulation	4.2952	3.7160	3.5348	3.4002	3.3047	–
Refined mf	4.1017	3.6578	3.5098	3.3915	3.3027	3.2139

Average queue length: Refined mean field approximation gives a significant improvement.

The supermarket model (SQ(2))

N	10	20	30	50	100	∞
$\rho = 0.7$						
Simulation	1.2194	1.1735	1.1584	1.1471	1.1384	–
Refined mf	1.2150	1.1726	1.1584	1.1471	1.1386	1.1301
$\rho = 0.9$						
Simulation	2.8040	2.5665	2.4907	2.4344	2.3931	–
Refined mf	2.7513	2.5520	2.4855	2.4324	2.3925	2.3527
$\rho = 0.95$						
Simulation	4.2952	3.7160	3.5348	3.4002	3.3047	–
Refined mf	4.1017	3.6578	3.5098	3.3915	3.3027	3.2139

Mean field approximation

Average queue length: Refined mean field approximation gives a significant improvement.

Pull-push model (servers with ≥ 2 jobs push to empty)

N	10	20	50	100	∞
$\rho = 0.8$					
Simulation	1.5569	1.4438	1.3761	1.3545	–
Refined mean field	1.5473	1.4403	1.3761	1.3547	1.3333
$\rho = 0.90$					
Simulation	2.3043	1.9700	1.7681	1.7023	–
Refined mean field	2.2945	1.9654	1.7680	1.7022	1.6364
$\rho = 0.95$					
Simulation	3.4288	2.6151	2.1330	1.9720	–
Refined mean field	3.4369	2.6232	2.1350	1.9723	1.8095

Mean field approximation

Average queue length: Refined mean field approximation is remarkably accurate

$SQ(2)$: the impact of choosing with/without replacement

Reminder: the least loaded of two servers has i jobs with probability:

$$x_{i-1}^2 - x_i^2 \qquad \text{when picked with replacement}$$

$$x_{i-1}\frac{Nx_{i-1}-1}{N-1} - x_i\frac{Nx_i-1}{N-1} \qquad \text{when picked without replacement}$$

Asymptotically equal but there is a $1/N$-difference!

$SQ(2)$: the impact of choosing with/without replacement

Reminder: the least loaded of two servers has i jobs with probability:

$$x_{i-1}^2 - x_i^2 \qquad \text{when picked with replacement}$$

$$x_{i-1}\frac{Nx_{i-1}-1}{N-1} - x_i\frac{Nx_i-1}{N-1} \qquad \text{when picked without replacement}$$

Asymptotically equal but there is a $1/N$-difference!

	$N = 10$ servers	Simulation	Refined mean field	Mean field
$\rho = 0.7$	with	1.215	1.215	1.1301
	without	1.173	1.169	1.1301
	with-without	0.042	0.046	–
$\rho = 0.9$	with	2.820	2.751	2.3527
	without	2.705	2.630	2.3527
	with-without	0.115	0.121	–
$\rho = 0.95$	with	4.340	4.102	3.2139
	without	4.169	3.923	3.2139
	with-without	0.171	0.179	–

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Main Elements of the Proof

1: Semi-groups and generators

For a Markov process, we define the operator Ψ_t that associates to a function h the functions $\Psi_t h$.

$$\Psi_t hx = \mathbb{E}\left[h(X(t)) \mid X(0) = x\right].$$

The generator is the derivative of Ψ_t at time 0:

$$Gh(x) = \frac{1}{dt}\mathbb{E}\left[h(X(t+dt)) - h(X(t)) \mid X(t) = x\right].$$

Main Elements of the Proof

1: Semi-groups and generators

For a Markov process, we define the operator Ψ_t that associates to a function h the functions $\Psi_t h$.

$$\Psi_t hx = \mathbb{E}\left[h(X(t)) \mid X(0) = x\right].$$

The generator is the derivative of Ψ_t at time 0:

$$Gh(x) = \frac{1}{dt}\mathbb{E}\left[h(X(t+dt)) - h(X(t)) \mid X(t) = x\right].$$

Examples:

- For a Markov process that jumps from i to j at rate Q_{ij}:
$$Gh(i) = \sum_j (h(j) - h(i))Q_{ij}$$
- For a deterministic ODE $\dot{x} = f(x)$:
$$Gh(x) = Dh(x) \cdot f(x).$$

Main Elements of the Proof

2: Comparison of Generators

The generators of the system N and the mean field approximation are:

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

$$(\Lambda h)(x) = \sum_{\ell \in \mathcal{L}} \beta_\ell(x) Dh(x) \cdot \ell = Dh(x) \cdot f(x)$$

Main Elements of the Proof

2: Comparison of Generators

The generators of the system N and the mean field approximation are:

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

$$(\Lambda h)(x) = \sum_{\ell\in\mathcal{L}} \beta_\ell(x) Dh(x)\cdot\ell = Dh(x)\cdot f(x)$$

If h is a twice-differentiable function, then:

$$\lim_{N\to\infty} N(L^{(N)} - \Lambda)h(x) = \frac{1}{2}\sum_{\ell\in\mathcal{L}} \beta_\ell(x) D^2h(x)\cdot(\ell,\ell)$$

Main Elements of the Proof

3. Stein's method

If X^N is distributed according to the stationary distribution of $L^{(N)}$, then for any function g:

$$\mathbb{E}\left[(L^{(N)}g)(X^N)\right] = 0$$

Main Elements of the Proof

3. Stein's method

If X^N is distributed according to the stationary distribution of $L^{(N)}$, then for any function g:

$$\mathbb{E}\left[(L^{(N)}g)(X^N)\right] = 0$$

Now, assume that there exists a function g such that

$$h(x) - h(\pi) = (\Lambda g)(x)$$

Main Elements of the Proof

3. Stein's method

If X^N is distributed according to the stationary distribution of $L^{(N)}$, then for any function g:

$$\mathbb{E}\left[(L^{(N)}g)(X^N)\right] = 0$$

Now, assume that there exists a function g such that

$$h(x) - h(\pi) = (\Lambda g)(x)$$

Then, we have:

$$\begin{aligned}
N\mathbb{E}\left[h(X^N) - h(\pi)\right] &= N\mathbb{E}\left[(\Lambda g)(X^N)\right] \\
&= N\mathbb{E}\left[(\Lambda - L^{(N)})(g)(X^N)\right] \\
&= \frac{1}{2}\mathbb{E}\left[\sum_{\ell} \beta_\ell(X^N) D^2 g(X^N) \cdot (\ell, \ell)\right] + O(1/N) \\
&\to \frac{1}{2}\sum_{\ell} \beta_\ell(\pi) D^2 g(\pi) \cdot (\ell, \ell).
\end{aligned}$$

Main Elements of the Proof

4. Perturbation theory

Let g be $g(x) = \int_0^\infty (h(\pi) - h(\Phi_t(x)))dt$, where $\Phi_t(x)$ is the solution of the ODE $\dot{x} = f(x)$ starting in x at time 0. Then:

$$\begin{aligned} g(x) &= \int_0^{dt} (h(\pi) - h(\Phi_t(x)))dt + \int_{dt}^\infty (h(\pi) - h(\Phi_t(x)))dt \\ &\approx (h(\pi) - h(x))dt + g(\Phi_{dt}(x)) \end{aligned}$$

This “shows” that $(\Lambda g)(x) = h(x) - h(\pi)$.

Main Elements of the Proof

4. Perturbation theory

Let g be $g(x) = \int_0^\infty (h(\pi) - h(\Phi_t(x)))dt$, where $\Phi_t(x)$ is the solution of the ODE $\dot{x} = f(x)$ starting in x at time 0. Then:

$$\begin{aligned} g(x) &= \int_0^{dt} (h(\pi) - h(\Phi_t(x)))dt + \int_{dt}^\infty (h(\pi) - h(\Phi_t(x)))dt \\ &\approx (h(\pi) - h(x))dt + g(\Phi_{dt}(x)) \end{aligned}$$

This "shows" that $(\Lambda g)(x) = h(x) - h(\pi)$.

To finish, we need to show that g is twice-differentiable. This comes from perturbation theory.

$$D^2 g(x) = -\int_0^t D^2 h(\Phi_t(x))dt$$

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Where does the $O(1/N)$-term comes from?

Going back to the $SQ(2)$ example

Transitions on X_i: $+\frac{1}{N}$ at rate $N(x_{i-1}^2 - x_i^2)$ and $-\frac{1}{N}$ at rate $N(x_i - x_{i+1})$. Hence:

$$\frac{d}{dt}\mathbb{E}[X_i(t)] = \mathbb{E}\left[X_{i-1}^2(t) - X_i^2(t) - (X_i(t) - X_{i+1}(t))\right] \quad \text{(exact)}$$

$$= \mathbb{E}\left[X_{i-1}^2(t)\right] - \mathbb{E}\left[X_i^2(t)\right] - \mathbb{E}[X_i(t)] + \mathbb{E}[X_{i+1}(t)]$$

$$\approx \mathbb{E}[X_{i-1}(t)]^2 - \mathbb{E}[X_i(t)]^2 - \mathbb{E}[X_i(t)] + \mathbb{E}[X_{i+1}(t)] \quad \text{(mean field approx.}$$

Where does the $O(1/N)$-term comes from?

Going back to the $SQ(2)$ example

Transitions on X_i: $+\frac{1}{N}$ at rate $N(x_{i-1}^2 - x_i^2)$ and $-\frac{1}{N}$ at rate $N(x_i - x_{i+1})$. Hence:

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}\left[X_i(t)\right] &= \mathbb{E}\left[X_{i-1}^2(t) - X_i^2(t) - (X_i(t) - X_{i+1}(t))\right] \quad \text{(exact)}\\
&= \mathbb{E}\left[X_{i-1}^2(t)\right] - \mathbb{E}\left[X_i^2(t)\right] - \mathbb{E}\left[X_i(t)\right] + \mathbb{E}\left[X_{i+1}(t)\right]\\
&\cancel{\approx \mathbb{E}\left[X_{i-1}(t)\right]^2 - \mathbb{E}\left[X_i(t)\right]^2 - \mathbb{E}\left[X_i(t)\right] + \mathbb{E}\left[X_{i+1}(t)\right]} \quad \text{(mean field approx.}
\end{aligned}
$$

If we now consider how $\mathbb{E}\left[X_i^2\right]$ evolves, we have:

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}\left[X_i^2\right] &= \mathbb{E}\left[(2X_i + \frac{1}{N})(X_{i-1}^2 - X_i^2) + (-2X_i + \frac{1}{N})(X_i - X_{i+1})\right]\\
&= \mathbb{E}\left[\underbrace{2X_iX_{i-1}^2}_{\mathbb{E}[X_iX_{i-1}^2 \approx ?]} + \dots\dots\dots\dots\right]
\end{aligned}
$$

where we denote X instead of $X(t)$ for simplicity.

System Size Expansion Approach

Recall that the transitions are $X \mapsto X + \frac{\ell}{N}$ at rate $N\beta_\ell(x)$.

$$\frac{d}{dt}\mathbb{E}\left[X\right] = \mathbb{E}\left[\sum_\ell \beta_\ell(X)\ell\right] = \mathbb{E}\left[f(X)\right] \qquad \text{(Exact)}$$

$$\frac{d}{dt}x = f(x) \qquad \text{(Mean Field Approx.)}$$

System Size Expansion Approach

Recall that the transitions are $X \mapsto X + \frac{\ell}{N}$ at rate $N\beta_\ell(x)$.

$$\frac{d}{dt}\mathbb{E}[X] = \mathbb{E}\left[\sum_\ell \beta_\ell(X)\ell\right] = \mathbb{E}[f(X)] \qquad \text{(Exact)}$$

$$\frac{d}{dt}x = f(x) \qquad \text{(Mean Field Approx.)}$$

We can now look at the second moment:

$$\begin{aligned}\mathbb{E}[(X-x)\otimes(X-x)] = {} & \mathbb{E}[(f(X)-f(x))\otimes(X-x)] \qquad \text{(Exact)}\\ & + \mathbb{E}[(X-x)\otimes(f(X)-f(x))]\\ & + \frac{1}{N}\mathbb{E}\left[\sum_{\ell\in\mathcal{L}} \beta_\ell(X)\ell\otimes\ell\right]\end{aligned}$$

System Size Expansion Approach

Recall that the transitions are $X \mapsto X + \frac{\ell}{N}$ at rate $N\beta_\ell(x)$.

$$\frac{d}{dt}\mathbb{E}\left[X\right] = \mathbb{E}\left[\sum_{\ell}\beta_\ell(X)\ell\right] = \mathbb{E}\left[f(X)\right] \qquad \text{(Exact)}$$

$$\frac{d}{dt}x = f(x) \qquad \text{(Mean Field Approx.)}$$

We can now look at the second moment:

$$\begin{aligned}\mathbb{E}\left[(X-x)\otimes(X-x)\right] &= \mathbb{E}\left[(f(X)-f(x))\otimes(X-x)\right] \qquad \text{(Exact)}\\ &+ \mathbb{E}\left[(X-x)\otimes(f(X)-f(x))\right]\\ &+ \frac{1}{N}\mathbb{E}\left[\sum_{\ell\in\mathcal{L}}\beta_\ell(X)\ell\otimes\ell\right]\end{aligned}$$

... We can also look at higher order moments

$$\begin{aligned}\mathbb{E}\left[(X-x)^{\otimes 3}\right] &= 3\mathrm{Sym}\mathbb{E}\left[(f(X)-f(x))\otimes(X-x)\otimes(X-x)\right]\\ &+ \frac{3}{N}\mathrm{Sym}\mathbb{E}\left[\sum_{\ell\in\mathcal{L}}\beta_\ell(X)\ell\otimes\ell\otimes(X-x)\right] + \frac{1}{N}\mathbb{E}\left[\sum_{\ell\in\mathcal{L}}\beta_\ell(X)\ell\otimes\ell\otimes\ell\right]\end{aligned}$$

System Size Expansion and Moment Closure

Let $x(t)$ be the mean field approximation and $Y(t) = X(t) - x(t)$, and $Y(t)^{(k)} = \underbrace{Y(t) \otimes \cdots \otimes Y(t)}_{k \text{ times}}$

$\frac{d}{dt}\mathbb{E}\left[Y(t)^{(k)}\right]$ can be expressed as **an exact** function of $Y(t)^{(j)}$ for $j \in \{0 \ldots, k+1\}$.

System Size Expansion and Moment Closure

Let $x(t)$ be the mean field approximation and $Y(t) = X(t) - x(t)$, and $Y(t)^{(k)} = \underbrace{Y(t) \otimes \cdots \otimes Y(t)}_{k \text{ times}}$

$\frac{d}{dt}\mathbb{E}\left[Y(t)^{(k)}\right]$ can be expressed as **an exact** function of $Y(t)^{(j)}$ for $j \in \{0 \ldots, k+1\}$.

You can close the equations by assuming that $Y^{(k)} = 0$ for $k \geq K$.

- For $K = 1$, this gives the mean field approximation ($1/N$-accurate)
- For $K = 3$, this gives the refined mean field ($1/N^2$-accurate).
- For $K = 5$, this gives a second order expansion ($1/N^3$-accurate).

Limit of the approach: For a system of dimension d, $Y(t)^{(k)}$ has d^k equations.

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Outline

1. Construction of the Mean Field Approximation: 3 models
   - Density Dependent Population Processes
   - A Second Point of View: Zoom on One Object
   - Discrete-Time Models
2. On the Accuracy of Mean Field : Positive and Negative Results
   - Transient Analysis
   - Steady-state Regime
3. The Refined Mean Field
   - Main Results
   - Generator Comparison and Stein's Method
   - Alternative View: System Size Expansion Approach
4. Demo
5. Conclusion and Open Questions

Recap and extensions

For a mean field model with twice differentiable drift, then :

1. The accuracy of the classical mean field approximation is $O(1/N)$.
2. We can use this to define a refined approximation.
3. The refined approximation is often accurate for $N = 10$.

Extensions:

- Transient regime
- Discrete-time (Synchronous)
- Next expansion term in $1/N^2$.

In many cases, the refined approximation is very accurate

"Truth"	Refined mean field approximation	Mean field approximation
$\mathbb{E}\left[X^N\right]$	$\pi + \frac{V}{N}$	π (=fixed point)

	Coupon	Supermarket	Pull/push
Simulation ($N = 10$)	1.530	2.804	2.304
Refined mean field ($N = 10$)	1.517	2.751	2.295
Mean field ($N = \infty$)	1.250	2.353	1.636

[1]

[1] Ref : G., Van Houdt, 2018

Some References

Job opening – Game theory, privacy and mean field.

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

nicolas.gast@inria.fr

- A Refined Mean Field Approximation by Gast and Van Houdt. SIGMETRICS 2018 (best paper award)
- Size Expansions of Mean Field Approximation: Transient and Steady-State Analysis Gast, Bortolussi, Tribastone
- Expected Values Estimated via Mean Field Approximation are O(1/N)-accurate by Gast. SIGMETRICS 2017.
- https://github.com/ngast/rmf_tool/