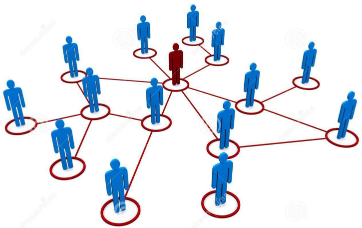
Computing the bias of mean field approximation

Nicolas Gast (Inria, Grenoble)

joint work with Sebastian Allmeier (Inria) and Benny Van Houdt (Univ. Antwerp)

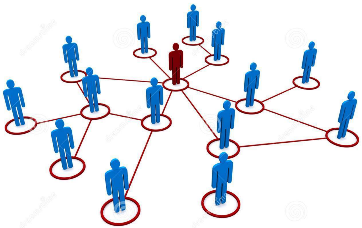
Journées COSMOS, Avignon, October 2022

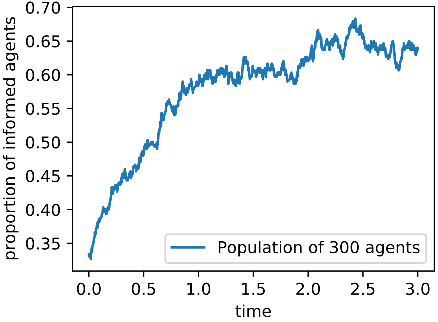
Stochastic information models on 'dense' graphs



- Propagation of an information over time
- Steady-state properties (e.g. % of informed people)

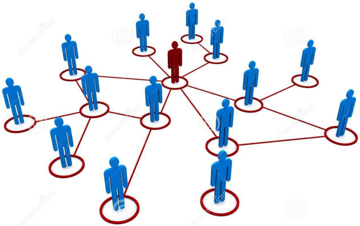
Stochastic information models on 'dense' graphs

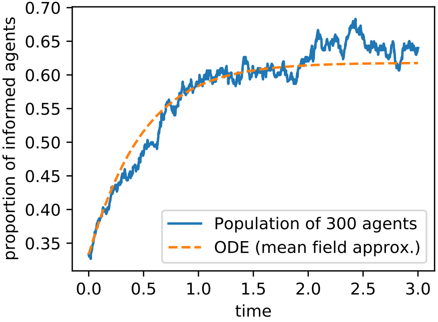




- Propagation of an information over time
- Steady-state properties (e.g. % of informed people)

Stochastic information models on ‘dense’ graphs





- Propagation of an information over time
- Steady-state properties (e.g. % of informed people)

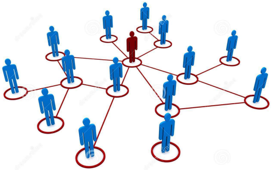
Objective of the talk (and outline)

1. What is mean field approximation?
2. How to characterize the bias of this approximation?
3. What about multiscale models?
4. Conclusion

Outline

1. What is mean field approximation?
2. How to characterize the bias of this approximation?
3. What about multiscale models?
4. Conclusion

Running example: Simple information propagation model.



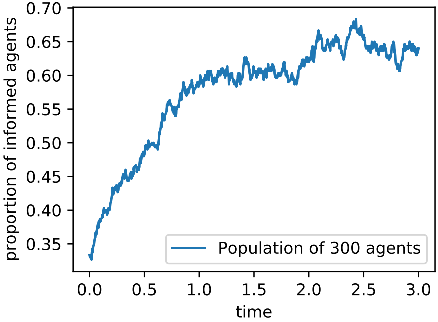
Population of n persons where each person can be "Informed" or "Outdated". x is the proportion of "informed" people.

- Informed persons loose information at rate 1.
- Outdated persons become informed at rate $1 + x$

Stochastic model

If X is the proportion of "informed" people, then:

$$X \mapsto X - \frac{1}{n} \text{ at rate } nX$$
$$X \mapsto X + \frac{1}{n} \text{ at rate } n(1-X^2) = n(1-X)(1+X)$$



Mean field approximation

$$X \mapsto X - \frac{1}{n} \text{ at rate } nX \qquad \textit{average change: } -X$$

$$X \mapsto X + \frac{1}{n} \text{ at rate } n(1 - X^2) \qquad \textit{average change: } 1 - X^2$$

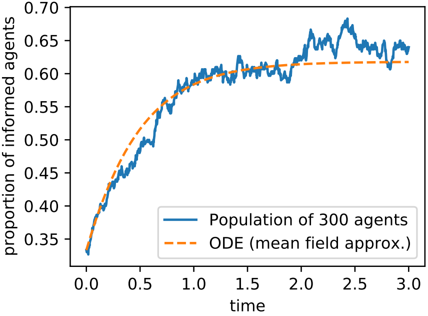
The mean field approximation is the solution of the ODE $\dot{x} = 1 - x^2 - x$.

Mean field approximation

$$X \mapsto X - \frac{1}{n} \text{ at rate } nX \qquad \textit{average change: } -X$$

$$X \mapsto X + \frac{1}{n} \text{ at rate } n(1 - X^2) \qquad \textit{average change: } 1 - X^2$$

The mean field approximation is the solution of the ODE $\dot{x} = 1 - x^2 - x$.

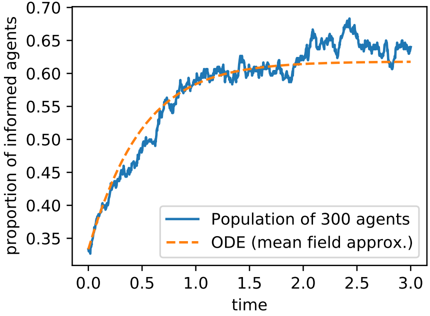


Mean field approximation

$$X \mapsto X - \frac{1}{n} \text{ at rate } nX \qquad \textit{average change: } -X$$

$$X \mapsto X + \frac{1}{n} \text{ at rate } n(1 - X^2) \qquad \textit{average change: } 1 - X^2$$

The mean field approximation is the solution of the ODE $\dot{x} = 1 - x^2 - x$.



With this approximation, we study:

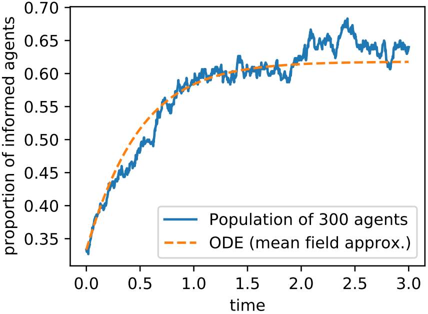
- The transient regime.
- The fixed point: $x(\infty) = (\sqrt{5} - 1)/2$.

Mean field approximation

$$X \mapsto X - \frac{1}{n} \text{ at rate } nX \qquad \textit{average change: } -X$$

$$X \mapsto X + \frac{1}{n} \text{ at rate } n(1-X^2) \qquad \textit{average change: } 1-X^2$$

The mean field approximation is the solution of the ODE $\dot{x} = 1 - x^2 - x$.



With this approximation, we study:

- The transient regime.
- The fixed point: $x(\infty) = (\sqrt{5}-1)/2$.

How accurate is this approximation?

Accuracy of the mean field approximation

n	5	10	100	∞
$\mathbb{P}$(someone informed)	0.593	0.601	0.61679642	$(\sqrt{5}-1)/2 \approx 0.618$.
Error	0.025	0.125	0.0012	0

Table: Steady-state values

Theorem (Folk). For dense graphs of interactions, mean field approximation is asymptotically exact. Accuracy is $O(1/\sqrt{n})$.

Accuracy of the mean field approximation

n	5	10	100	∞
$\mathbb{P}$(someone informed)	0.593	0.601	0.61679642	$(\sqrt{5}-1)/2 \approx 0.618.$
Error	0.025	0.125	0.0012	0

Table: Steady-state values

Theorem (Folk). For dense graphs of interactions, mean field approximation is asymptotically exact. Accuracy is $O(1/\sqrt{n})$.

Theorem: (G. Bortolussi, Tribastone) For this model, if $\mathbb{E}\left[X^{(n)}\right]$ is the probability that someone is "informed" and if the $x(\infty) = (\sqrt{5}-1)/2$ is the mean field approximation, then:

$$\mathbb{E}\left[X^{(n)}\right] = x(\infty) + \frac{1}{20n}(\sqrt{5}-1) + \frac{1}{50n^2}(\sqrt{5}-3) + O(1/n^3).$$

Outline

We study a generic interaction model

We consider a population of n objects with two types of interactions:

- Unilateral transitions:

  Object k jumps from state i to j at rate $r_{ij}^{(k)}$

- Pairwise interactions:

  Object k, k' *simultaneously* jump from states (i, i') to (j, j') at rate $r_{ij,i'j'}^{(k,k')}/n$

If the rates do not depend on k, we call the model homogeneous.

Mean field approximation for homogeneous models

$$X_s^{(n)}(t) = \frac{1}{n}\{\# \text{ objects in state } s \text{ at } t\}$$

The transitions are:

$$\mathbf{X}^{(n)} \to \mathbf{X}^{(n)} + \frac{1}{n}(e_j - e_i) \qquad \text{at rate } n r_{ij} X_i.$$

$$\mathbf{X}^{(n)} \to \mathbf{X}^{(n)} + \frac{1}{n}(e_j - e_i + e_{j'} - e_{i'}) \qquad \text{at rate } n r_{ij,i'j'} X_i X_{i'}.$$

This is a density dependent population process (Kurtz 70s).

(one example is our information propagation model)

Mean field method for non-homogeneous models

$X_s^{(n)}(t) = \frac{1}{n}\{\#\text{ objects in state } s \text{ at } t\}$ $\Rightarrow X^{(n)}$ is not Markovian.

Mean field method for non-homogeneous models

~~$X_s^{(n)}(t) = \frac{1}{n}\{\#\text{ objects in state } s \text{ at } t\}$ $\Rightarrow X^{(n)}$ is not Markovian.~~

Solution: represent model using indicators:

$$Y_{(k,s)}^{(n)}(t) = \begin{cases} 1 & \text{if object } k \text{ is in state } s \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{Y}^{(n)}$ is Markovian.

$$\mathbf{Y}^{(n)} \to \mathbf{Y}^{(n)} + e_{k,j} - e_{k,i} \qquad \text{at rate } r_{ij}^{(k)} Y_{k,i}.$$

$$\mathbf{Y}^{(n)} \to \mathbf{Y}^{(n)} + e_{k,j} - e_{k,i} + e_{k',j'} - e_{k',i'} \qquad \text{at rate } r_{ij,i'j'}^{(k,k')} Y_{k,i} Y_{k',i'}.$$

Mean field approximaiton and result

The drift is:

$$f(\mathbf{y}) = \sum_{\text{all transitions}} \text{Transition change for } \mathbf{y} \times \text{Rate of transition at } \mathbf{y}.$$

The mean field approximation is the solution of the ODE $\dot{y} = f(y)$.

Mean field approximaiton and result

The drift is:

$$f(\mathbf{y}) = \sum_{\text{all transitions}} \text{Transition change for } \mathbf{y} \times \text{Rate of transition at } \mathbf{y}.$$

The mean field approximation is the solution of the ODE $\dot{y} = f(y)$.

Theorem (Allmeier, G. 2022) There exists an (easily computable) vector $V(t)$ such that for any finite time:

$$\mathbf{P}\left[\text{ Object } k \text{ is in state } s \text{ at } t\right] = y_{k,s}(t) + \frac{1}{N} V_{k,s}(t) + O(\frac{1}{N^2}).$$

$V_{k,s}(t)$ is the bias of the mean field approximation.

Idea of proof

1. We can show that $\mathrm{cov}(Y_{k,s}Y_{k',s'})(t) = \frac{1}{N}W(t) + O(1/N^2)$, where $W(t)$ satisfies a (time inhomogeneous) linear ODE:

$$\dot{W} = A(y(t))W + WA^T(y(t)) + Q(x(t)).$$

2. We then have $\mathbb{E}\left[Y_{k,s}(t)\right] = y(t) + V(t)$, where $V(t)$ satisfies a (time inhomogeneous) linear ODE:

$$\dot{V} = A(y(t))V + B(x(t)) \cdot W(t).$$

Where does the $1/N$-term

The moment closure approach

Consider a system for which X becomes $X + 1/n$ at rate nX^2. We have:

$$\frac{d}{dt}\mathbb{E}\left[X\right] = \mathbb{E}\left[X^2\right]$$

This equation is not closed

Where does the $1/N$-term

The moment closure approach

Consider a system for which X becomes $X + 1/n$ at rate nX^2. We have:

$$\frac{d}{dt}\mathbb{E}[X] = \mathbb{E}\left[X^2\right] \qquad \approx \mathbb{E}[X]^2 \text{ (mean field approx.)}$$

This equation is ~~not~~ closed

Where does the $1/N$-term

The moment closure approach

Consider a system for which X becomes $X + 1/n$ at rate nX^2. We have:

$$\frac{d}{dt}\mathbb{E}[X] = \mathbb{E}\left[X^2\right] \qquad \not\approx \mathbb{E}[X]^2 \text{ (mean field approx.)}$$

$$\frac{d}{dt}\mathbb{E}\left[X^2\right] = 2\mathbb{E}\left[X^3\right] + \frac{1}{n}\mathbb{E}\left[X^2\right]$$

This equation is not closed

Where does the $1/N$-term

The moment closure approach

Consider a system for which X becomes $X + 1/n$ at rate nX^2. We have:

$$\frac{d}{dt}\mathbb{E}[X] = \mathbb{E}\left[X^2\right] \qquad \approx \mathbb{E}[X]^2 \text{ (mean field approx.)}$$

$$\frac{d}{dt}\mathbb{E}\left[X^2\right] = 2\mathbb{E}\left[X^3\right] + \frac{1}{n}\mathbb{E}\left[X^2\right] \qquad \approx 2(3\mathbb{E}\left[X^2\right]\mathbb{E}[X] - 2\mathbb{E}[X]^2) + \frac{1}{n}\mathbb{E}\left[X^2\right]$$

(refined approximation)

This equation is ~~not~~ closed

Where does the $1/N$-term

The moment closure approach

Consider a system for which X becomes $X + 1/n$ at rate nX^2. We have:

$$\frac{d}{dt}\mathbb{E}[X] = \mathbb{E}\left[X^2\right] \qquad \sout{\approx \mathbb{E}[X]^2 \text{ (mean field approx.)}}$$

$$\frac{d}{dt}\mathbb{E}\left[X^2\right] = 2\mathbb{E}\left[X^3\right] + \frac{1}{n}\mathbb{E}\left[X^2\right] \qquad \sout{\approx 2(3\mathbb{E}\left[X^2\right]\mathbb{E}[X] - 2\mathbb{E}[X]^2) + \frac{1}{n}\mathbb{E}\left[X^2\right]}$$

$$\frac{d}{dt}\mathbb{E}\left[X^3\right] = \mathbb{E}\left[\frac{3X^4}{n} + \frac{4X^3}{n^2} + \frac{X^2}{n^3}\right]$$

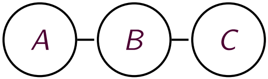
$$\vdots$$

The moment equations are never closed.

- They can be closed by assuming $\mathbb{E}\left[(X - \mathbb{E}[X])^d\right] \approx 0$
- This gives a $O(1/n^{\lfloor (d+1)/2 \rfloor})$-accurate approximation.

Outline

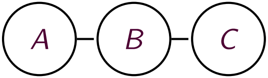
Communication with interference



Interference graph.

n nodes per class A, B or C

Communication with interference



Interference graph.

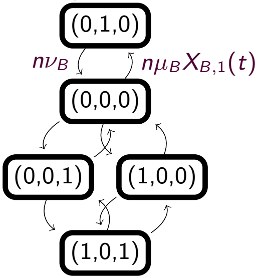
n nodes per class A, B or C

State is X, Y

- $X_{i,s} =$ proportion of nodes of class i with $\geq S$ messages.
- $Y_i = 1$ if class i talks.

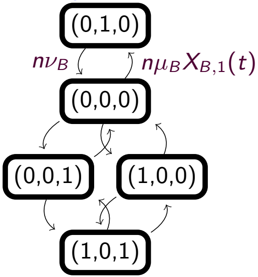
This is a two timescale model

"Fast process": Y.



This is a two timescale model

“Fast process”: Y.



“Slow process”: X.

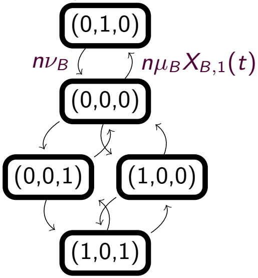
Arrival/departure:

$$X_{i,s} \mapsto X_{i,s} \pm \frac{1}{N}$$

Rate depends on y.

This is a two timescale model

"Fast process": Y.



"Slow process": X.

Arrival/departure:

$$X_{i,s} \mapsto X_{i,s} \pm \frac{1}{N}$$

Rate depends on y.

Two approximations:

$\mathbf{P}[Y(t) = y] \approx \pi_y(X(t))$ | Drift $f(X, Y)$

$$\dot{x} = \sum_y \pi_y(x) f(x, y)$$

(Averaging technique):

Accuracy results (Allmeier, G. 2022)

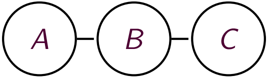
Theorem. If $X(t)$ is the two timescale process, if the rates are twice differentiable and the evolution the the fast process is "unichain", then:

$$\mathbb{E}[X(t)] = x(t) + \frac{1}{N}C(t) + O(1/N^2).$$

Holds uniformly in time if the ODE has an exponentially stable attractor.

Numerical example

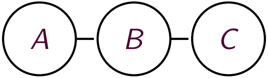
With $n = 1$ node per class!



Jobs arrive at rate 1, activation rate $= 3$. Job duration is $1/3$.

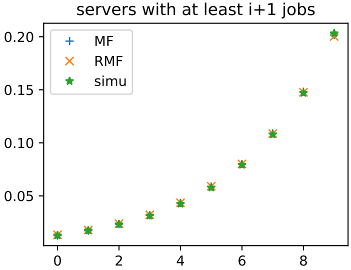
Numerical example

With $n = 1$ node per class!



Jobs arrive at rate 1, activation rate = 3. Job duration is 1/3.

Class A or C



Class B

Outline

Conclusion

Mean field approximation is a widely used heuristic.

- It consists in assuming independence.

We question its validity / accuracy.

Conclusion

Mean field approximation is a widely used heuristic.

- It consists in assuming independence.

We question its validity / accuracy.

- We characterize the bias for different models (smooth homogeneous, heterogeneous, multi-scale).
- To do so, we take correlations into account.
- Numerical library: `https://pypi.org/project/rmftool/`

Conclusion

Mean field approximation is a widely used heuristic.

- It consists in assuming independence.

We question its validity / accuracy.

- We characterize the bias for different models (smooth homogeneous, heterogeneous, multi-scale).
- To do so, we take correlations into account.
- Numerical library: `https://pypi.org/project/rmftool/`

Many open questions: (sparse) geometric models, non-Markovian, controlled systems

More slides and references: `http://polaris.imag.fr/nicolas.gast`

References

Results on which this talk is based:

- Mean Field and Refined Mean Field Approximations for Heterogeneous Systems: It Works! by Allmeier and Gast. SIGMETRICS 2022.
- 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. Performance 2018.
- Two-scale: Bias and Refinement of Multiscale Mean Field Models. Allmeier, Gast, 2022 (available soon).
  - CSMA model from CSMA networks in a many-sources regime: A mean-field approach. Cecchi, Borst, van Leeuwaarden, Whiting. Infocom 2016.

Paper cited as open problems:

- Pair-approximation: The Power of Two Choices on Graphs: the Pair-Approximation is Accurate by Gast. Mama 2015.
- Non-Markovian: Randomized Load Balancing with General Service Time Distributions by Bramson, Ly and Prabhakar. Sigmetrics 2010 and The PDE Method for the Analysis of Randomized Load Balancing Networks by Aghajani, Li, Ramanan.SIGMETRICS 2018