Bias of two timescale “replica” mean field approximation

Nicolas Gast (Inria, Grenoble)

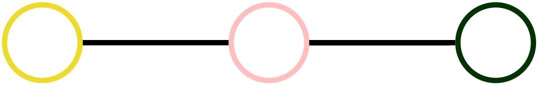
joint work with Sebastian Allmeier (Inria)

Informs APS, Nancy, June 2023

Motivation: independent sets with arrivals

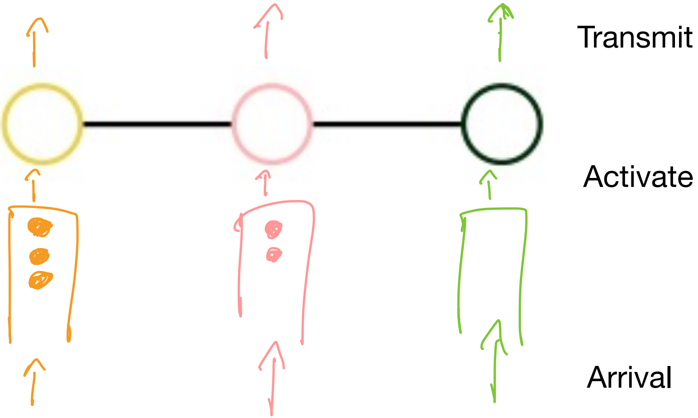
Two-timescale "replica" mean field

CSMA model from Cecchi et al. 2015



Two-timescale "replica" mean field

CSMA model from Cecchi et al. 2015

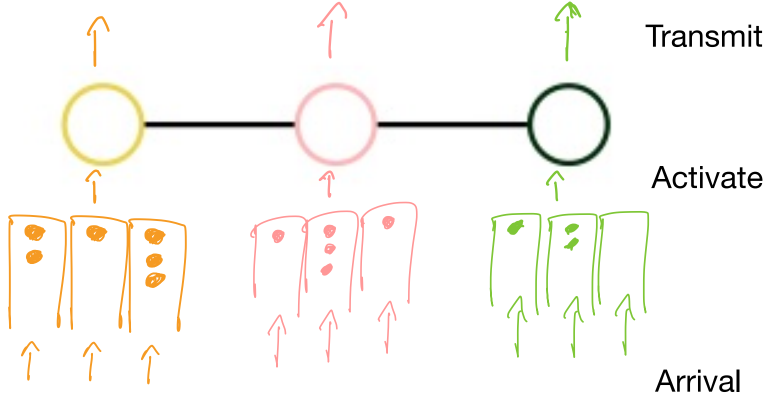


Objective: estimate steady-state.

$$\mathbf{P}[S_k = i]$$

Two-timescale "replica" mean field

CSMA model from Cecchi et al. 2015



Objective: estimate steady-state.

$$\mathbf{P}\left[S_k = i\right]$$

Scaling: *replica*

- N severs per node
- Arrival rate $\times\dfrac{1}{N}$.

Reults in a nutshell

$$X_{k,i} = \frac{1}{N}\#\{\text{Objects of type } k \text{ in state } i\}$$
$$Y = \text{“activation set”}$$

Reults in a nutshell

$$X_{k,i} = \frac{1}{N}\#\{\text{Objects of type } k \text{ in state } i\}$$
$$Y = \text{``activation set''}$$

We show that $X_{k,i}$ has a fluid limit x_i and that:

$$P(S_k = i) = x_{k,i} + \frac{1}{N} v_{k,i} + O(1/N^2).$$

This is very accurate even for small N.

Outline

Two timescale mean field model

Population of N objects.

- Object k has a state $S_k(t) \in \mathcal{S}$.
- Shared resource $Y(t) \in \mathcal{Y}$.

$$X_i = \text{ fraction of objects in state } i.$$

Two timescale mean field model

Population of N objects.

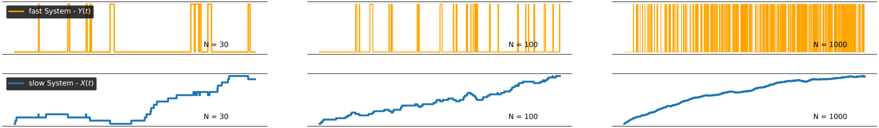
- Object k has a state $S_k(t) \in \mathcal{S}$.
- Shared resource $Y(t) \in \mathcal{Y}$.

$$X_i = \text{ fraction of objects in state } i.$$

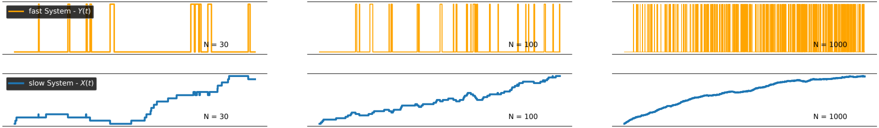
Model

- Object n jumps from i to j at rate $Q_{i,j}(\mathbf{X}, \mathbf{Y})$
- Resource $\mathbf{Y}$ jumps from y to y' at rate $K_{y,y'}(\mathbf{X})$.

We use two approximations to construct a fluid limit.



We use two approximations to construct a fluid limit.



"Average" mean field approximation

Let $\pi_y(\mathbf{x})$ be the stationary distribution of $K(\mathbf{x})$. We define:

$$\bar{Q}(\mathbf{x}) = \sum_y \pi_y(\mathbf{x}) Q(\mathbf{x}, y).$$

The mean field approximation is the solution of the ODE:

$$\dot{\mathbf{x}} = \mathbf{x}\bar{Q}(\mathbf{x}),$$

The approximation is asymptocally exact. Its bias is v/N.

Assume that

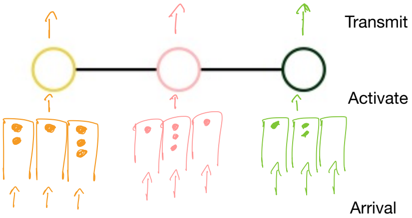
- K is unichain for all $\mathbf{x}$.
- K and Q are twice differenciable
- $\dot{\mathbf{x}} = \mathbf{x}\bar{Q}(\mathbf{x})$ has a unique attractor x^*.

Theorem

There exists a *computable* V such that, in steady-state:

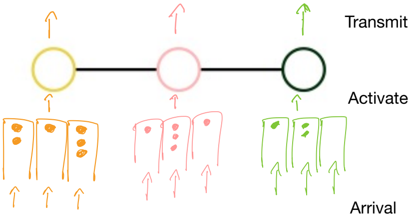
$$\mathbf{P}\left[S_k = i\right] = \underbrace{\underbrace{x_i^*}_{\text{mean field approximation}} + \frac{1}{n}V_i}_{\text{refined approximation}} + O(1/n^2).$$

Illustration of the theorem



Transmission rates	1.4	1.3	1.7
Activation rates	1.2	2	1.5
Arrival rates	0.5	0.2	0.5

Illustration of the theorem



Transmission rates	1.4	1.3	1.7
Activation rates	1.2	2	1.5
Arrival rates	0.5	0.2	0.5

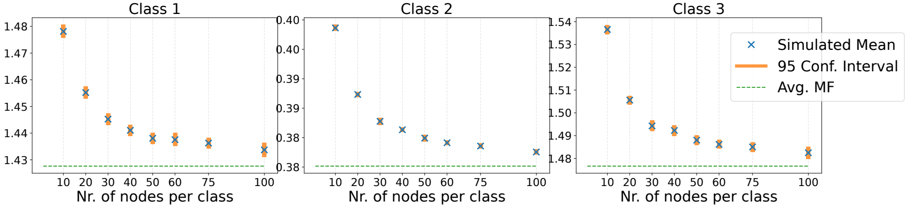
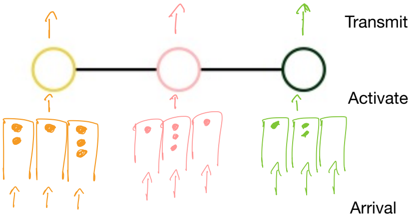
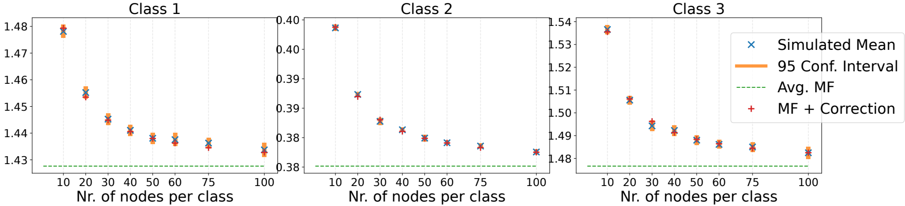


Illustration of the theorem

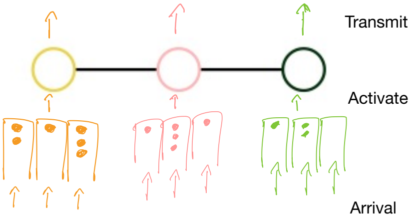


Transmission rates	1.4	1.3	1.7
Activation rates	1.2	2	1.5
Arrival rates	0.5	0.2	0.5

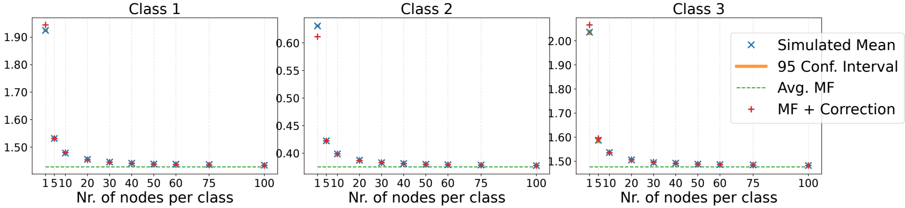


MF+correction is almost exact

Illustration of the theorem



Transmission rates	1.4	1.3	1.7
Activation rates	1.2	2	1.5
Arrival rates	0.5	0.2	0.5



MF+correction is almost exact

Outline

If Y is not here, we can directly use Stein's method

Let G^{sto} be the Generator of the stochastic system. For $h : \mathcal{X} \to \mathbb{R}$:

1. $$G^{\mathrm{sto}} h(X) = \sum_{i,j} (h(X + \frac{1}{n}(e_j - e_i)) - h(X)) n x_i Q_{ij}(x)$$

If Y is not here, we can directly use Stein's method

Let G^{sto} be the Generator of the stochastic system. For $h : \mathcal{X} \to \mathbb{R}$:

1. $$G^{\mathrm{sto}} h(X) = \sum_{i,j} (h(X + \frac{1}{n}(e_j - e_i)) - h(X)) n x_i Q_{ij}(x)$$
$$= \underbrace{\nabla h \cdot xQ(X)}_{\text{Generator of ODE } \dot{x} = xQ(x).} + \underbrace{O(1/n)}_{\text{if } h \text{ is } C^1}.$$

2. $\mathbb{E}\left[G^{\mathrm{sto}} h(X)\right] = 0$ if X is in steady-state.

If Y is not here, we can directly use Stein's method

and introduce a Poisson equation for the slow system

Let $G^{\rm sto}$ be the Generator of the stochastic system. For $h : \mathcal{X} \to \mathbb{R}$:

1. $$G^{\rm sto}h(X) = \sum_{i,j}\Big(h\big(X + \frac{1}{n}(e_j - e_i)\big) - h(X)\Big) n x_i Q_{ij}(x)$$
$$= \underbrace{\nabla h \cdot xQ(X)}_{\text{Generator of ODE } \dot{x} = xQ(x).} + \underbrace{O(1/n)}_{\text{if } h \text{ is } C^1}.$$

2. $\mathbb{E}\left[G^{\rm sto}h(X)\right] = 0$ if X is in steady-state.

Let G be such that $\nabla G \cdot xQ(x) = x - x^*$ (Poisson equation). We have:

$$\begin{aligned}
\mathbb{E}[X - x^*] &= \mathbb{E}[\nabla G \cdot XQ(X)] \\
&= \mathbb{E}\left[(\nabla G - G^{\rm sto}) \cdot XQ(X)\right] && \text{(by (2))} \\
&= o(1/n) && \text{(by (1))}.
\end{aligned}$$

When Y is here, we need to treat the fast system

Let $h : \mathcal{X} \to \mathbb{R}$ be a test function. We have:

$$G^{\mathrm{sto}} h(X, Y) = \sum_{i,j} (h(X + \frac{1}{n}(e_j - e_i)) - h(X)) n x_i Q_{ij}(X, Y)$$
$$= \underbrace{\nabla h \cdot XQ(X, Y)}_{\neq G^{\mathrm{ODE}} = \nabla h \cdot \bar{Q}(x).} + O(1/n),$$

When Y is here, we need to treat the fast system

Let $h : \mathcal{X} \to \mathbb{R}$ be a test function. We have:

$$\begin{aligned} G^{\mathrm{sto}} h(X,Y) &= \sum_{i,j} \left(h(X + \frac{1}{n}(e_j - e_i)) - h(X)\right) n x_i Q_{ij}(X,Y) \\ &= \underbrace{\nabla h \cdot X Q(X,Y)}_{\neq G^{\mathrm{ODE}} = \nabla h \cdot \bar{Q}(x).} + O(1/n), \end{aligned}$$

We are left with $Q(X,Y) - \bar{Q}(X)$.

When Y is here, we need to treat the fast system

and introduce a Poisson equation for the fast system.

Let $h : \mathcal{X} \to \mathbb{R}$ be a test function. We have:

$$\begin{aligned} G^{\mathrm{sto}} h(X, Y) &= \sum_{i,j} (h(X + \frac{1}{n}(e_j - e_i)) - h(X)) n x_i Q_{ij}(X, Y) \\ &= \underbrace{\nabla h \cdot XQ(X, Y)}_{\neq G^{\mathrm{ODE}} = \nabla h \cdot \bar{Q}(x).} + O(1/n), \end{aligned}$$

We are left with $Q(X, Y) - \bar{Q}(X)$.

Lemma: There exists a K^+ that is C^2 such that for all $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$:

$$h(X, Y) - \bar{h}(X) = K(x) K^+(x) h(X, Y).$$

Rapping up the proof

Let $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a test function.

$$G^{\mathrm{sto}} h(X,Y) = \underbrace{nK(X)h(X,Y)}_{\text{fast}} + \underbrace{\nabla_x h \cdot XQ(X,Y)}_{\text{slow}} + O(1/n)$$

Hence, $K^{fast} - \frac{1}{n} G^{\mathrm{sto}} = o(1/n)$ if h is C^1.

Rapping up the proof

Let $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a test function.

$$G^{\mathrm{sto}} h(X, Y) = \underbrace{nK(X)h(X, Y)}_{\text{fast}} + \underbrace{\nabla_x h \cdot XQ(X, Y)}_{\text{slow}} + O(1/n)$$

Hence, $K^{fast} - \frac{1}{n} G^{\mathrm{sto}} = o(1/n)$ if h is C^1.

This shows that in steady-state:

$$\begin{aligned}
\mathbb{E}\left[h(X) - \bar{h}(X)\right] &= \mathbb{E}\left[K(x)K^+(x)h(X, Y)\right] \\
&= \mathbb{E}\left[(K(x) - \frac{1}{n} G^{\mathrm{sto}})K^+(x)h(X, Y)\right] && \text{(steady-state).} \\
&= O(1/n) && \text{(expansion above).}
\end{aligned}$$

Outline

Conclusion

We study the accuracy of mean field approximation for two time-scale.

- The bias is of order $O(1/N)$. It can be computed.
- This also works for most “smooths” models (e.g., heterogeneous).

Conclusion

We study the accuracy of mean field approximation for two time-scale.

- The bias is of order $O(1/N)$. It can be computed.
- This also works for most "smooths" models (e.g., heterogeneous).

Two-timescale models:

- Shared resource or synchronization (e.g., CSMA)
- Q-learning type algorithm: Stochastic approximation algorithms with Markovian noise. Huo et al. 2023

Conclusion

We study the accuracy of mean field approximation for two time-scale.

- The bias is of order $O(1/N)$. It can be computed.
- This also works for most "smooths" models (e.g., heterogeneous).

Two-timescale models:

- Shared resource or synchronization (e.g., CSMA)
- Q-learning type algorithm: Stochastic approximation algorithms with Markovian noise. Huo et al. 2023

Many open questions: non-smooth, (sparse) geometric models, non-Markovian.

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

References

Results on which this talk is based:

- Bias and Refinement of Multiscale Mean Field Models. Allmeier, Gast, 2022. Sigmetrics 2023.
- CSMA networks in a many-sources regime: A mean-field approach. Cecchi, Borst, van Leeuwaarden, Whiting. Infocom 2016.

Q-learning and bias:

- Bias and Extrapolation in Markovian Linear Stochastic Approximation with Constant Stepsizes. Dongyan Huo, Yudong Chen, Qiaomin Xie. Sigmetrics 2023.

Related refined mean-field approximation papers:

- 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.