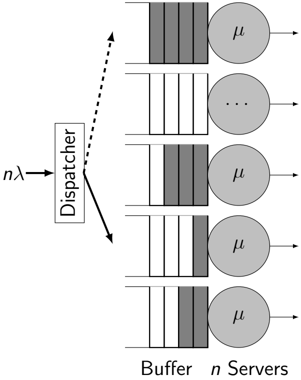
Mean Field and Refined Mean Field Approximations for Heterogeneous Systems: It Works!

Sebastian Allmeier [1]   Nicolas Gast [1]

[1]Inria Grenoble - Rhône-Alpes;
Univ. Grenoble Alpes

SIGMETRICS 2022, 6th - 10th June

Classical Mean Field Setting



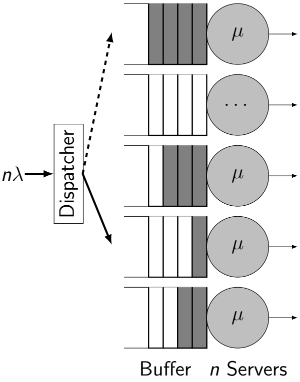
Example: Two-Choice Model

Mean Field Methodology:

- $M_s^{(n)}(t) = \frac{1}{n}\{\#$ objects in state s at time $t\}$
- $M_s^{(n)}(t) \xrightarrow{n\to\infty}$ Mean Field Approximation(ODE)

Works for: objects w/ homogeneous transitions
groups of objects w/ similar statistical behavior

Classical Mean Field Setting



Example: Two-Choice Model

Mean Field Methodology:

- $M_s^{(n)}(t) = \frac{1}{n}\{\#\text{ objects in state } s \text{ at time } t\}$
- $M_s^{(n)}(t) \xrightarrow{n\to\infty}$ Mean Field Approximation(ODE)

Works for: objects w/ homogeneous transitions
groups of objects w/ similar statistical behavior

Based on object independence assumption for large n.

Can be made more accurate by considering refinements which take dependencies into account.

Need to Model Heterogeneity

Importance of Heterogeneity

- heterogeneity has a dramatic impact, e.g. for caching, epidemic modelling, load balancing
- many homogeneous models ignore heterogeneity
- in general **no theoretical guarantees** for accuracy

Mean Field Approximation

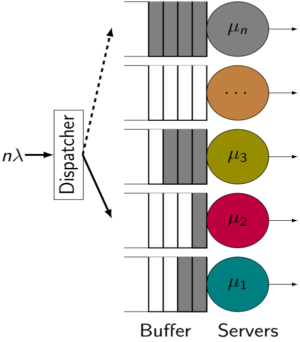
Experiments suggest similar behavior as for homogeneous case.
Advantages remain: simplify analysis + fast computation

Key Question

Can we mathematically justify that the Mean Field approximation is a valid technique?

Example of a Heterogeneous Two-Choice Model

represent model using indicators: $X^{(n)}_{(k,s)}(t) = \mathbf{1}_{\{\text{server } k \text{ has } s \text{ jobs in its queue at time } t\}}$

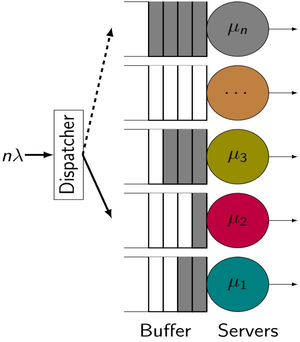


$X_{(k,s)}(t) \sim$ indicator if the buffer of server k is filled with s jobs at time t

Example: Het. Two-Choice Model

Example of a Heterogeneous Two-Choice Model

represent model using indicators: $X^{(n)}_{(k,s)}(t) = \mathbf{1}_{\{\text{server } k \text{ has } s \text{ jobs in its queue at time } t\}}$



$X_{(k,s)}(t) \sim$ indicator if the buffer of server k is filled with s jobs at time t

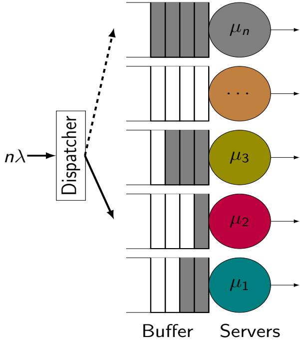
Dynamics:

- arrival - job added to shorter queue (**pairwise transitions**)
- departure - w/ **heterogeneous** rates μ_i (**unilateral transitions**)

Example: Het. Two-Choice Model

Example of a Heterogeneous Two-Choice Model

$X^{(n)}_{(k,s)}(t) = \mathbf{1}_{\{\text{server } k \text{ has } s \text{ jobs in its queue at time } t\}}$



Example: Het. Two-Choice Model

Departure

$$\mathbf{X} \mapsto \mathbf{X} - e_{(k,s)} + e_{(k,s-1)} \qquad \text{(unilateral)}$$

$$\text{at rate } \mu_k \, X_{(k,s)}$$

Arrival

$$\mathbf{X} \mapsto \mathbf{X} + e_{(k,s+1)} - e_{(k,s)} \qquad \text{(pairwise)}$$

$$\text{at rate } \lambda/n \, X_{(k,s)} X_{(k',s)}$$

$$\text{or } 2\lambda/n \, X_{(k,s)} X_{(k',j)} \quad (j > s)$$

The Generic Heterogeneous Framework

Allowed Dynamics: d-wise interactions

Interactions of d objects $\mathbf{k} = (k_1, \dots, k_d)$ jumping from states $\mathbf{s}$ to $\mathbf{s}'$ at rate $\frac{1}{n^{d-1}} r_{\mathbf{k},\mathbf{s}\to\mathbf{s}'}$:

- $\mathbf{X} \mapsto \mathbf{X} - e_{(k,s)} + e_{(k,s')}$

  at rate $r_{k,s\to s'} X_{(k,s)}$ (unilateral)

- $\mathbf{X} \mapsto \mathbf{X} - e_{(k_1,s_1)} + e_{(k_2,s_1')} - e_{(k_2,s_2)} + e_{(k_2,s_2')}$

  at rate $\frac{1}{n} r_{\mathbf{k},\mathbf{s}\to\mathbf{s}'} X_{(k_1,s_1)} X_{(k_2,s_2)}$ (pairwise)

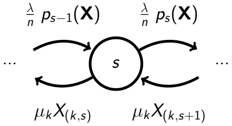
- $\dots$

Definition of the Mean Field Approximation

Obtaining the drift

$$f^{(n)}(\mathbf{X}(t)) \approx \sum_{\substack{\text{set of jumps} \\ \text{from } \mathbf{X}^{(n)}(t)}} \text{jump} \times \text{jump rate}$$

Zoom on single item for het. JSQ(2)



$$f^{(n)}_{(k,s)}(\mathbf{X}) = \frac{\lambda}{n} p_{(s-1)}(\mathbf{X}) + \mu_k X_{(k,s+1)} - \frac{\lambda}{n} p_{(s)}(\mathbf{X}) - \mu_k X_{(k,s)}$$

The Mean Field Approximation is solution to the IVP:

$$\frac{d}{dt}\mathbf{x}^{(n)}(t) = f^{(n)}(\mathbf{x}^{(n)}(t)); \quad \mathbf{x}^{(n)}(0) = \mathbf{X}^{(n)}(0)$$

Results

- By definition: Drift is Lipschitz continuous and smooth
- Assumption: uniform bound on rates independent of system size n

Theorem

For the Mean Field approximation $\mathbf{x}(t)$ and refinement term $\mathbf{v}(t)$

$$\mathbb{P}\,(\text{item } k \text{ in state } s \text{ at time } t) = x_{(k,s)}(t) + O(1/n),$$
$$\mathbb{P}\,(\text{item } k \text{ in state } s \text{ at time } t) = x_{(k,s)}(t) + v_{(k,s)}(t) + O(1/n^2).$$

$\rightarrow$ Allows to obtain approximations for linear transformations such as average values (e.g. average queue length)

Refinement & Proof

The Heterogeneous Framework allows to define a **refinement term $\boldsymbol{v(t)}$** following the ideas of [Gast, Van Houdt] [Gast, Bortolussi, Tribastone].

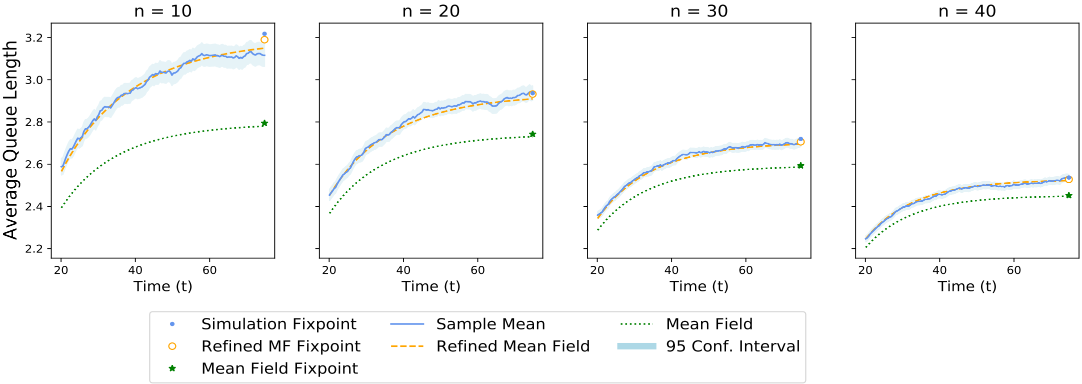
The **Refined Mean Field approximation** $\underline{\mathbf{x}(t) + \mathbf{v}(t)}$ increases accuracy by considering dependencies among objects in the approximation.

Proof Idea

- $\mathbb{P}\,(\text{item } k \text{ in state } s \text{ at time } t) = \mathbb{E}\left[X_{(k,s)}(t)\right]$
- use generator comparison of stochastic and deterministic process (related to Stein's method)
- apply Taylor approximation and bound remainder
- use transition rate structure to obtain $1/n$ bounds

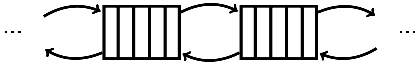
Numerical Results

Increasing Mean Field Accuracy as n grows
Almost exact Refined MF



Approximation Error for a Cache model

The RANDOM cache model in a nutshell:



	Mean Field		Refined Mean Field		Simulation	
n	Error	(time)	Error	(time)	Error	(time)
10	0.0142	(10ms)	0.00197	(10ms)	0.00026	(4.3s)
30	0.0050	(14ms)	0.00022	(17ms)	0.00047	(4.9s)
50	0.0031	(17ms)	0.00008	(30ms)	0.00055	(5.7s)

Refined Mean Field is more accurate than simulations of 10^8 requests

Computation Time for a Cache model

	Transient up to $T = 1000$	
n	mean field	refined m.f.
10	30ms	180ms
30	30ms	370ms
50	35ms	1s
100	60ms	14s
300	170ms	–*
500	300ms	–*
1000	970ms	–*

* "–" means that the ODE solver did not finish before 30 seconds.

Takeaways

The heterogeneous framework and derived Mean Field Approximations

- can be applied to **fully** and **partially heterogeneous** interaction models
- can be **refined** through the refinement term $\mathbf{v}$
- can be **efficiently implemented**
  (i.e., by using the rmftool: `https://pypi.org/project/rmftool/`)
- exhibit **small hidden constants** in practice

Takeaways

The heterogeneous framework and derived Mean Field Approximations

- can be applied to **fully** and **partially heterogeneous** interaction models
- can be **refined** through the refinement term $\mathbf{v}$
- can be **efficiently implemented**
  (i.e., by using the rmftool: `https://pypi.org/project/rmftool/`)
- exhibit **small hidden constants** in practice

Thank you!

Sebastian Allmeier
sebastian.allmeier@inria.fr

Mean Field and Refined Mean Field
Approximations for Heterogeneous Systems: It Works!

References

Nicolas Gast and Benny Van Houdt 2017

A Refined Mean Field Approximation. Proceedings of the ACM on Measurement and Analysis of Computing Systems 1, 2

Nicolas Gast, Luca Bortolussi, and Mirco Tribastone 2019

Size Expansions of Mean Field Approximation: Transient and Steady-State Analysis. Performance Evaluation 129 (Feb. 2019)

Sebastian Allmeier and Nicolas Gast 2021

Mean Field and Refined Mean Field Approximations for Heterogeneous Systems: It Works! Proc. ACM Meas. Anal. Comput. Syst., 6(1), (Feb 2022)

Sebastian Allmeier and Nicolas Gast 2021

rmf_tool – a library to compute (refined) mean field approximation(s). SIGMETRICS Perform. Eval. Rev. (Workshop).