# Mean Field Approximation of Uncertain Stochastic Models

## Luca Bortolussi<sup>1</sup>, Nicolas Gast<sup>2</sup>

DSN conference 2016, Toulouse, France

<sup>1</sup>CNR-Univ Trieste, Italy <sup>2</sup>Inria, France

Nicolas Gast - 1 / 30

# Why do we need models?



• Design, prediction, optimization, correctness, etc.

# Uncertainties in models: stochastic v.s. non-determinism

Stochastic	Non-determinism
ex: Markov chains	ex: ODE, timed-automata

# Uncertainties in models: stochastic v.s. non-determinism

Stochastic ex: Markov chains	Non-determinism ex: ODE, timed-automata	
Quantitative analysis.	Worst case / correctness	
<ul><li>+ can be simulated</li><li>- How to choose parameters?</li></ul>	<ul> <li>Symbolic computation</li> <li>+ No problem of parameters</li> </ul>	
Combination of the two Uncertain Markov chains		

# Building a

# continuous time Markov chain

• agents engage in actions at some rate.



# Building an uncertain continuous time Markov chain

• agents engage in actions at some rate.



## Main problem: the state space grows exponentially



We need to keep track

$$\mathbb{P}(X_1(t)=i_1,\ldots,X_n(t)=i_n)$$

and solve the differential inclusion:

$$\frac{d}{dt}P(t) \in \bigcup_{\theta \in [\theta_{\min}, \theta_{\max}]} P(t)Q(\theta)$$

 $3^{13} \approx 10^6$  states.

#### Is there any hope?

# Contributions (and Outline)

Some systems simplify when the population grows.

- Mean-field approach
- 2 We can add non-determinism to these models
- **③** We can build and use numerical algorithms.

# Outline

Population Processes and Classical Mean Field Methods

- 2 Uncertain and Imprecise Population Processes
- 3 Numerical Algorithms and Comparisons
  - Numerical algorithms (transient regime)
  - Steady-state
  - General processor sharing example



# Outline

## Population Processes and Classical Mean Field Methods

- 2 Uncertain and Imprecise Population Processes
- 3 Numerical Algorithms and Comparisons
  - Numerical algorithms (transient regime)
  - Steady-state
  - General processor sharing example

## 4 Conclusion

Mean field methods have been used in a multiple contexts

ex: model-checking, performance of SSD, load balancing, MAC protocol,...

- SPAA 98 Analyses of Load Stealing Models Based on Differential Equations by Mitzenmacher
- JSAC 2000 Performance Analysis of the IEEE 802.11 Distributed Coordination Function by Bianchi
- FOCS 2002 Load balancing with memory by Mitzenmacher et al.

.

- DSN 2013 A logic for model-checking mean-field models by Kolesnichenko et al
- DSN 2013 Lumpability of fluid models with heterogeneous agent types by lacobelli and Tribastone
- SIGMETRICS 2013 A mean field model for a class of garbage collection algorithms in flash-based solid state drives by Van Houdt





We view the population of objects more abstractly, assuming that individuals are indistinguishable.



We view the population of objects more abstractly, assuming that individuals are indistinguishable.

An occupancy measure records the proportion of agents that are currently exhibiting each possible state.



We view the population of objects more abstractly, assuming that individuals are indistinguishable.

An occupancy measure records the proportion of agents that are currently exhibiting each possible state.



We view the population of objects more abstractly, assuming that individuals are indistinguishable.

An occupancy measure records the proportion of agents that are currently exhibiting each possible state.

## Population CTMC

We consider a sequence  $\mathbf{X}^N$ , indexed by the population size N, with state spaces  $\mathbf{E}^N \subset E \subset \mathbb{R}^d$ . The transitions are:

$$X\mapsto X+rac{\ell}{N}$$
 at rate  $Neta_\ell(X).$ 

for a finite number of  $\ell \in \mathcal{L}$ . The drift is  $f(x) = \sum_{\ell} \ell \beta_{\ell}(x)$ .

Example :





$$\ell_1 = (-1, +1, 0) \text{ at rate } NX_S X_I \\ \ell_2 = (0, -1, +1) \text{ at rate } NX_I \\ \ell_3 = (+1, 0, -1) \text{ at rate } NX_R \\ \ell_4 = (-1, 0, +1) \text{ at rate } NX_S$$

## Kurtz' convergence theorem

**Theorem**: Let **X** be a population model. If  $X^N(0)$  converges (in probability) to a point x, then the stochastic process **X**<sup>N</sup> converges (in probability) to the solutions of the differential equation  $\dot{x} = f(x)$ , where f is the drift.



## Kurtz' convergence theorem

Theorem: Let **X** be a population model. If  $X^N(0)$  converges (in probability) to a point x, then the stochastic process **X**<sup>N</sup> converges (in probability) to the solutions of the differential equation  $\dot{x} = f(x)$ , where f is the drift.



## Kurtz' convergence theorem

**Theorem:** Let **X** be a population model. If  $X^N(0)$  converges (in probability) to a point x, then the stochastic process **X**<sup>N</sup> converges (in probability) to the solutions of the differential equation  $\dot{x} = f(x)$ , where f is the drift.



# Outline

## Population Processes and Classical Mean Field Methods

#### 2 Uncertain and Imprecise Population Processes

## 3 Numerical Algorithms and Comparisons

- Numerical algorithms (transient regime)
- Steady-state
- General processor sharing example

## 4 Conclusion

# Uncertain and imprecise models

Instead of  $\beta_{\ell}(x)$ , the rates can depend on a parameter  $\vartheta$ :  $\beta_{\ell}(x, \vartheta)$ .

We distinguish two kinds of uncertainties:

Uncertain	Imprecise
$\vartheta \in \Theta$ is constant but its value is not known precisely. • Uncertainties in the model	$artheta = artheta(t) \in \Theta$ can vary (measur- ably) as a function of time • human behavior, environment, adversary,

## Uncertain and imprecise population models

We consider a sequence  $\mathbf{X}^N$  of Imprecise or Uncertain population processes, indexed by the size N, with state spaces  $\mathbf{E}^N \subset E \subset \mathbb{R}^d$ . The transitions are (for  $\ell \in \mathcal{L}$ ):

$$X\mapsto X+rac{\ell}{N}$$
 at rate  $Neta_\ell(X,artheta)$ 

The drifts corresponding to parameter  $\vartheta$  is  $f(x, \vartheta) = \sum_{\ell \in \mathcal{L}} \ell \beta_{\ell}(x, \vartheta)$ .

#### Theorem (Bortolussi, G. 2016)

if  $X^{N}(0)$  converges (in probability) to a point x, then the uncertain (or imprecise) stochastic process  $\mathbf{X}^{N}$  converges in probability to:

. ,	-	
Uncertain		Imprecise
A solution of $\dot{x} = f(x, \vartheta)$ (	for a given $\vartheta$ )	A solution of $\dot{x} \in \bigcup_{\vartheta \in \Theta} f(x, \vartheta)$

# Example of the SIR model

The state is  $(X_S, X_I, X_R)$  and the transitions are







## Consequence on the stochastic system



#### Uncertain

Imprecise

#### N = 1000

Remark : the proportion of infected is non-monotone on the infection rate.

Nicolas Gast - 17 / 30

## Consequence on the stochastic system



#### Uncertain

Imprecise

#### N = 10000

Remark : the proportion of infected is non-monotone on the infection rate.

Nicolas Gast - 17 / 30

# Outline

1 Population Processes and Classical Mean Field Methods

#### 2 Uncertain and Imprecise Population Processes

### 3 Numerical Algorithms and Comparisons

- Numerical algorithms (transient regime)
- Steady-state
- General processor sharing example

### 4 Conclusion

# Numerical algorithms

- Uncertain (fixed parameter)
  - Exhaustive search
  - Online learning
- Imprecise (varying parameter). Difficulty = non-linear.
  - Exact: reachability (ex: solvable by Pontryagin's principle)
  - Approximation: polygons (ex: differential hull)



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\dot{\overline{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x}, \overline{x})}{\overline{f}_i(\underline{x}, \overline{x})} = \min_{\substack{x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t) \\ x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t)}} \min F_i(x)$$



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$$



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) }} \min F_i(x)$$



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

 $\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$ 



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

 $\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$ 



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\dot{\overline{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

 $\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$ 



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x}, \overline{x})}{\overline{f}_i(\underline{x}, \overline{x})} = \min_{\substack{x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t) \\ x \in [\underline{x}, \overline{x}] : x_i = \underline{x}_i(t)}} \min F_i(x)$$



For any solution of the differential equation x, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\overline{\dot{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$$



For any solution of the differential equation *x*, we have:

 $\underline{x}(t) \leq x(t) \leq \overline{x}(t)$ 

where  $\underline{x}$  and  $\overline{x}$  satisfy  $\underline{\dot{x}} = \underline{f}(\underline{x}, \overline{x})$ and  $\dot{\overline{x}} = \overline{f}(\underline{x}, \overline{x})$ , with

$$\frac{f_i(\underline{x},\overline{x})}{\overline{f}_i(\underline{x},\overline{x})} = \min_{\substack{x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t) \\ x \in [\underline{x},\overline{x}]: x_i = \underline{x}_i(t)}} \min F_i(x)$$

# An alternative is to formulate the problem as an optimization problem

$$x_i^{\max}(T) := \max_{ heta} x_i(T)$$
 such that for all  $t \in [0; T]$ :  

$$\begin{cases} x(t) = x + \int_0^t f(x(s), \theta(s)) ds \\ \theta(t) \in \Theta \end{cases}$$

- Pontryagin's principle can be used.
- Easier than MDP

# SIR model: we can compute an minimal/maximal trajectory by using Pontraygin's maximum principle



Stationary regime of imprecise models

Stationary regime of imprecise models

• Asymptotic reachable set of the differential inclusion  $A_F$ :

$$A_F = \bigcap_{T>0 \ x,t \ge T, \mathbf{x} \in S_{F,x}} \{\mathbf{x}(t)\}$$

- Theorem: Let **X** be an imprecise population process, then  $\lim_{N\to\infty} \lim_{t\to\infty} d(X^N(t), A_F) = 0 \qquad \text{in probability.}$
- Theorem: Let X be an imprecise population process such that X<sup>N</sup> is a Markov chain that has a stationary measure μ<sup>N</sup>. Let μ be a limit point of μ<sup>N</sup> (for the weak convergence). Then, the support of μ is included in the Birkhoff centre of F: μ(B<sub>F</sub>) = 1.

## Asymptotically reachable set: example for the SIR model



Note: comparison with the differential hull approach: bounds are very loose.

# SIR model: stationary regime



• No policy can make the stochastic system exit the blue zone (for large N).

# In the paper, we also study a Generalized Processor Sharing model

![](_page_52_Figure_1.jpeg)

• Parameters:  $\mu_1 = 5$ ,  $\mu_2 = 1$ ,  $\phi_1 = \phi_2 = 1$ ,  $a_1 = 1$  and  $a_2 = 2$ .  $\lambda_i, \lambda'_i$  imprecise with  $\lambda_1^{\min} = 1, \lambda_1^{\max} = 7, \lambda_2^{\min} = 2$ ,  $\lambda_2^{\max} = 3$ ,  $\lambda'_i^{\min} = 1/(1/a_i + a/\lambda_i^{\min})$ , and  $\lambda'_i^{\max} = 1/(1/a_i + a/\lambda_i^{\max})$ 

• A model of two tandem queues  $Q_1$ ,  $Q_2$  sharing a processor.  $Q_i$  gets a fraction  $\phi_i N_i Q_i / (\phi_1 N_1 Q_1 + \phi_2 N_2 Q_2)$  of the capacity C of the server. Each queue serves a job of type i, with average completion time  $\mu_i$ . Arrivals are Poisson ( $D_i$  - delay station - rate  $\lambda'_i$ ) or MAP (two delay stations in series  $E_i$ ,  $D_i$  rates  $a_i$ ,  $\lambda_i$ ).

Generalized Processor Sharing: for the imprecise model, a higher arrival rate does not imply a larger queuing delay.

![](_page_53_Figure_1.jpeg)

• Optimization (for imprecise) of  $\phi_1$  to minimize the maximum queue length at time t:  $\bar{Q}(t) = \max_{\theta} (Q_1^{\theta}(t) + Q_2^{\theta}(t)).$ 

# Outline

Population Processes and Classical Mean Field Methods

- 2 Uncertain and Imprecise Population Processes
- 3 Numerical Algorithms and Comparisons
  - Numerical algorithms (transient regime)
  - Steady-state
  - General processor sharing example

![](_page_54_Picture_7.jpeg)

# Recap and Future Work

Mean field methods are useful to study large stochastic systems.

We extended the mean field results for imprecise and uncertain PCTMCs, both at transient and at steady state.

We developed numerical method to bound the reachable sets.

Future work: scalability of numerical algorithms, integration in a toolset (EU project Quanticol), algorithmic complexity.

![](_page_55_Figure_5.jpeg)

# Thank you!

Slides are online:

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

nicolas.gast@inria.fr

This paper	$\begin{array}{l} \textit{Mean Field Approximation of Uncertain Stochastic Models, L. Bortolussi and N. }\\ \textit{Gast, DSN 2016} \end{array}$
	Other mean-field references
B-G 16	Mean-Field Limits Beyond Ordinary Differential Equations, L. Bortolussi, N. Gast., SFM Quanticol summer school
Benaïm, Le Boudec 08	A class of mean field interaction models for computer and communication
Le Boudec 08	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., JY. L. Boudec. , Arxiv:1009.5021, 2010
G. Van Houdt 15	Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms., Gast, Van Houdt., ACM Sigmetrics 2015