IMT SCHOOL FOR ADVANCED STUDIES LUCCA

Reconciling discrete and continuous modelling for the analysis of large-scale Markov chains

Mirco Tribastone

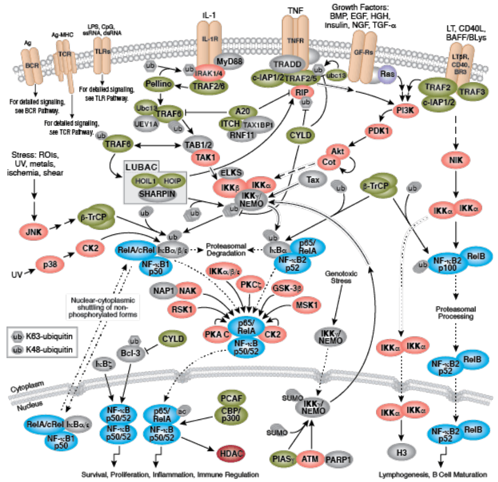
IMT Lucca

PRIN 2020 - Nirvana

Ca’ Foscari – 8 July 2022

Motivation

Markov chains are a popular tool for stochastic analysis

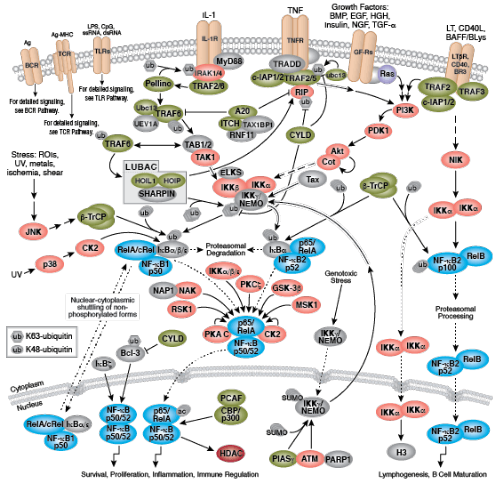


- PageRank and other centrality measures
- Epidemiological models
- Chemistry
- Systems biology
- Queuing systems
- (Probabilistic) programming languages
- (Probabilistic) population protocols
- …

> *"Indeed, the whole of the mathematical study of random processes can be regarded as a generalization in one way or another of the theory of Markov chains."*
>
> J.R. Norris, *Markov chains*, 1998

Motivation

Markov chains are a popular tool for stochastic analysis



- PageRank and other centrality measures
- **Epidemiological models**
- Chemistry
- **Systems biology**
- **Queuing systems**
- (Probabilistic) programming languages
- **(Probabilistic) population protocols**
- …

Markov Population Processes

> *"Indeed, the whole of the mathematical study of random processes can be regarded as a generalization in one way or another of the theory of Markov chains."*
>
> J.R. Norris, *Markov chains*, 1998

Markov Population Process

- A collection of *N* interacting agents evolving over *m* local states
- Typically, *N* is very large and *m* is small

Reaction Network

$$S + I \longrightarrow I + I, \quad \alpha S I$$
$$I \longrightarrow R, \quad \beta I$$

Agent states ($m = 3$)

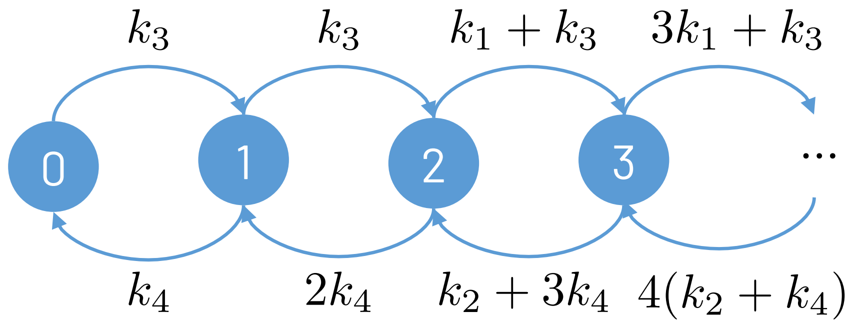
Reaction rates

Dynamics of a Markov Population Process

Schlögl's system

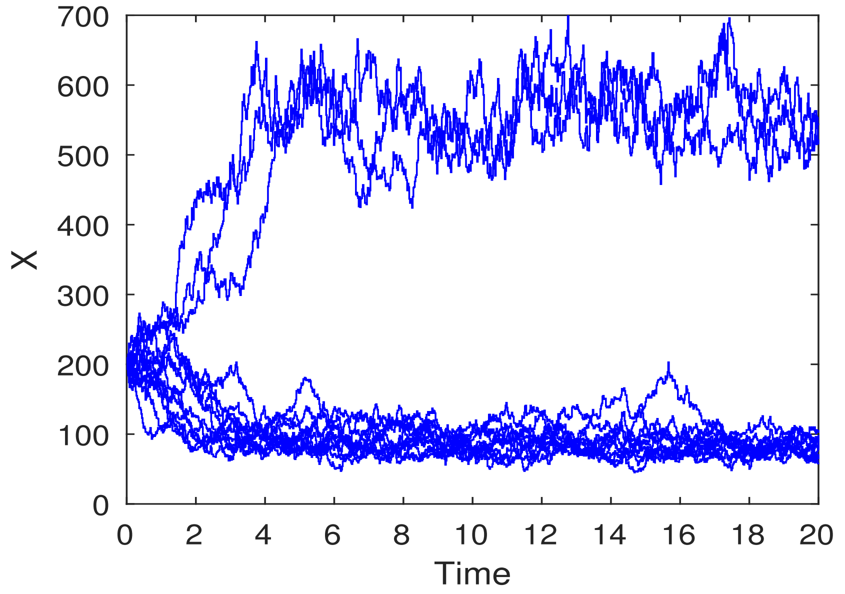
$$
\begin{aligned}
2X &\longrightarrow 3X, & k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, & k_2 X(X-1)(X-2)/6 \\
&\longrightarrow X, & k_3 \\
X &\longrightarrow \;\;, & k_4 X
\end{aligned}
$$

Markov chain



Stochastic Simulation

- Holding time at each state is **exponentially distributed** with the sum of outgoing rates
- Probability of a choosing a given transition after holding time equals its transition rate divided by the rate of the residence time

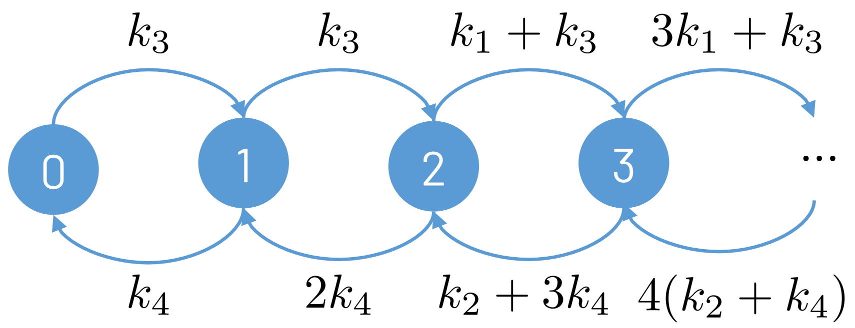


Dynamics of a Markov Population Process

Schlögl's system

$$
\begin{aligned}
2X &\longrightarrow 3X, \quad && k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad && k_2 X(X-1)(X-2)/6 \\
&\longrightarrow X, \quad && k_3 \\
X &\longrightarrow \ \ , \quad && k_4 X
\end{aligned}
$$

Markov chain



Master/Forward/Kolmogorov Equations

- Analytical computation of the transient probability distribution
- A system of coupled linear ordinary differential equations, each giving the time course of the Markov chain in any discrete state

$$
\begin{aligned}
\frac{d\pi_0}{dt} &= -k_3\pi_0 + k_4\pi_1 \\
\frac{d\pi_1}{dt} &= -(k_3 + k_4)\pi_1 + k_3\pi_0 + 2k_4\pi_2 \\
\frac{d\pi_2}{dt} &= -(k1 + k_3 + 2k_4)\pi_2 + k_3\pi_1 + (k_2 + 3k_4)\pi_3 \\
&\ \ \dots
\end{aligned}
$$

Mean-field Approximation

- Stochastic analysis is often **expensive**
  - Many simulations required for complex systems to obtain tight confidence intervals
  - Analytical solution of the Markov chain possible when the number of states is small enough (and approximations are usually needed for infinite state Markov chains)
- "Full" stochastic analysis is often **too informative**
  - In many applications the modeller is interested on average behaviour (and perhaps a few higher order moments)

Mean-field approximation (aka deterministic rate equations)

Analytical technique to approximate the **average dynamics of a Markov population process** using a (smaller) system of ordinary differential equations

Mean-field Approximation: Example

Schlögl's system

$$\begin{aligned} 2X &\longrightarrow 3X, \quad && k_1 X(X-1)/2 \\ 3X &\longrightarrow 2X, \quad && k_2 X(X-1)(X-2)/6 \\ &\longrightarrow \ X, \quad && k_3 \\ X &\longrightarrow \ \ , \quad && k_4 X \end{aligned}$$

Properties

- Self-consistent, **compact system** of equations (one per type of agent)
- Correct in the limit when the population levels go to infinity (**Kurtz's theorem**)
- Derivation can be generalized to obtain equations for higher-order moments (**moment-closure approximation**)

Derivation

Dynkin's Formula

$$\frac{d\mathbb{E}[X]}{dt} = \frac{k_1}{2}\mathbb{E}[X(X-1)] - \frac{k_2}{6}\mathbb{E}[X(X-1)(X-2)] + + k_3 - k_4\mathbb{E}[X]$$

(large-scale approximation)

$$\approx \frac{k_1}{2}\mathbb{E}[X^2] - \frac{k_2}{6}\mathbb{E}[X^3] + k_3 - k_4\mathbb{E}[X]$$

(expectation of a function vs. function of the expectations)

$$\approx \frac{k_1}{2}\mathbb{E}[X]^2 - \frac{k_2}{6}\mathbb{E}[X]^3 + k_3 - k_4\mathbb{E}[X]$$

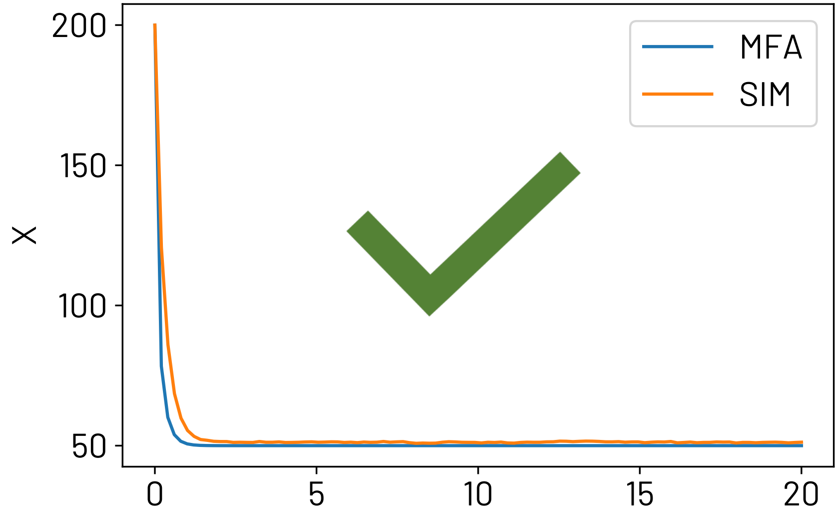
Mean-field Approximation: Results

Schlögl's system

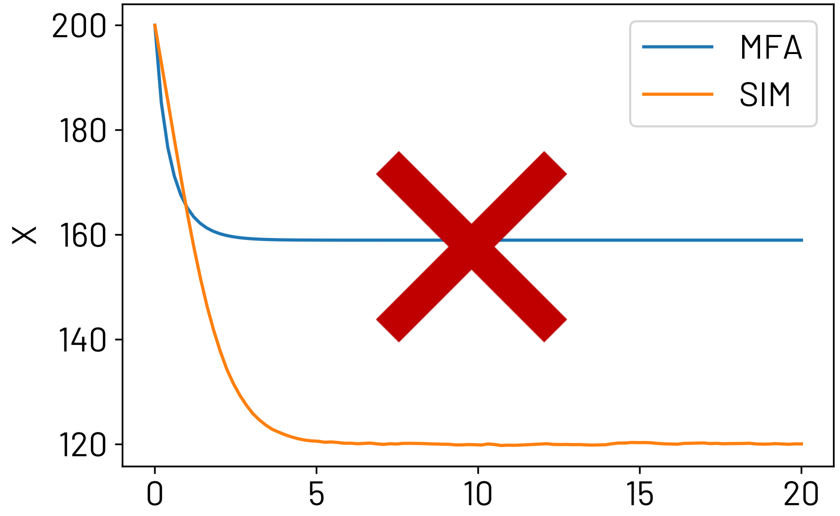
$$
\begin{aligned}
2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\
&\longrightarrow \ X, \quad k_3 \\
X &\longrightarrow \quad , \quad k_4 X
\end{aligned}
$$

Properties

- Quality of the approximation can be model- and parameter-dependent
- Always correct for linear systems and for a limited class of nonlinear systems
- Exact corrections available for special cases



$k_1 = 0.03$
$k_2 = 0.0004$
$k_3 = 200$
$k_4 = 4.5$



$k_1 = 0.03$
$k_2 = 0.0001$
$k_3 = 200$
$k_4 = 3.5$

Finite State Expansion

Intuition

- Markov population process gives a **discrete** description of the system
- Mean-field approximation gives a **continuous** one
- These can be seen are two extremes of a **lattice of approximations** where a subset of the whole states is kept discrete, and the rest is approximated continuously
- **Finite state expansion** is such a *hybrid* analytical method

Method

- Fix an **observation bound** for each agent type: it gives how many entities of that class to observe discretely
- Create a new reaction network adding **new agents types**, one for each discrete configuration
- **Rewrite** each original reaction to track discrete changes as far as possible, using the original agent types when behaviour goes beyond the chosen observation bounds

Finite State Expansion: Worked Example

$$\begin{aligned} 2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\ 3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\ &\longrightarrow X, \quad k_3 \\ X &\longrightarrow \ , \quad k_4 X \end{aligned}$$

Reaction expansion

$$[\![n]\!] \xrightarrow{f_1(n)} [\![n]\!] + X, \qquad n = \overline{O}_X$$

$$[\![n]\!] \xrightarrow{f_1(n)} [\![n+1]\!], \qquad 0 \leq n < \overline{O}_X$$

$$f_1(n) = [\![n]\!] k_1 (X+n)(X+n-1)/2$$

Finite State Expansion: Worked Example

$$\begin{aligned} 2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\ 3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\ &\longrightarrow X, \quad k_3 \\ X &\longrightarrow \;\;, \quad k_4 X \end{aligned}$$

Reaction expansion

$$[\![n]\!] \xrightarrow{f_2(n)} [\![n-1]\!], \qquad 0 < n \leq \overline{O}_X$$

$$X + [\![0]\!] \xrightarrow{f_2(n)} [\![0]\!], \qquad n = 0$$

$$f_2(n) = [\![n]\!] k_2 (X+n)(X+n-1)(X+n-2)/6$$

conditional probability

coupled mass-action propensity

Finite State Expansion: Worked Example

$$\begin{aligned} 2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\ 3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\ &\longrightarrow X, \quad k_3 \\ X &\longrightarrow \ , \quad k_4 X \end{aligned}$$

Reaction expansion

$$[\![n]\!] \xrightarrow{f_3(n)} [\![n]\!] + X, \qquad n = \overline{O}_X$$

$$[\![n]\!] \xrightarrow{f_3(n)} [\![n+1]\!], \qquad 0 \leq n < \overline{O}_X$$

$$f_3(n) = [\![n]\!] k_3$$

Finite State Expansion: Worked Example

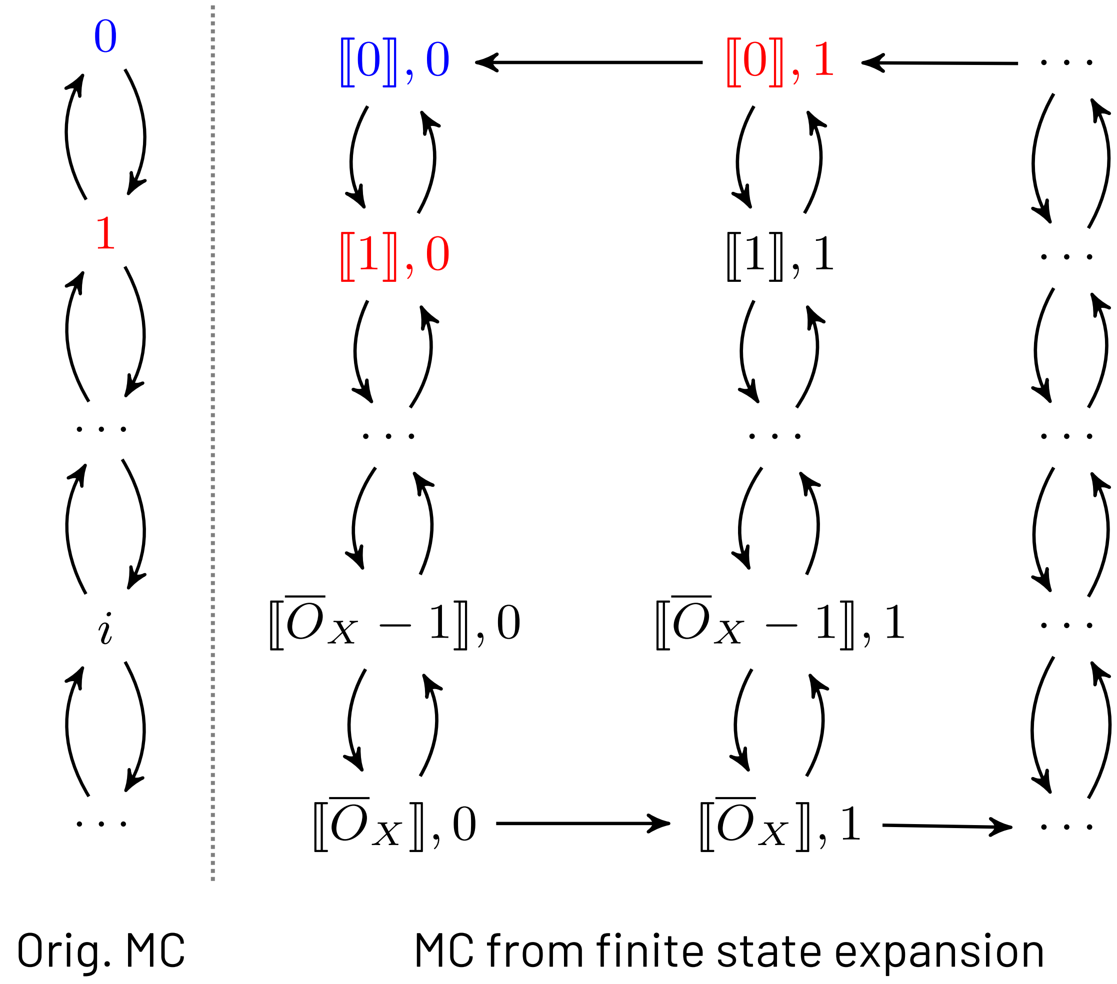
$$
\begin{aligned}
2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\
&\longrightarrow \ \ X, \quad k_3 \\
X &\longrightarrow \quad\ \ , \quad k_4 X
\end{aligned}
$$

Reaction expansion

$$
\begin{aligned}
[\![n]\!] &\xrightarrow{f_4(n)} [\![n-1]\!], & 0 < n \le \overline{O}_X \\
X + [\![n]\!] &\xrightarrow{f_4(n)} [\![n]\!], & n = 0 \\
f_4(n) &= [\![n]\!] k_4 (X + n)
\end{aligned}
$$

“Consistency” of Finite State Expansion

1. Finite state expansion always preserves the stochastic dynamics of the reaction network: the original and the expanded Markov chains are related by **ordinary lumpability** (aka probabilistic bisimulation)
2. When the observation bound is zero, the MFA corresponds to the original one
3. When the observation bound is infinity, **the MFA corresponds to the original forward equations**



Mean-field Approximation: Comparison

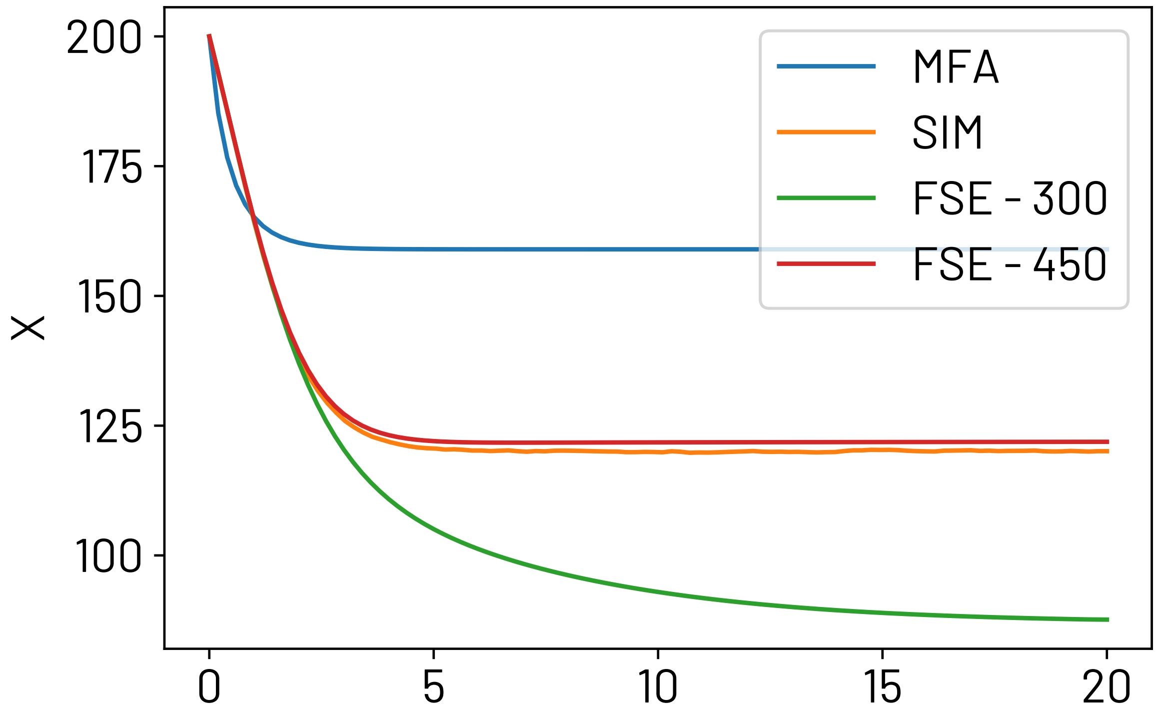
Original Schlögl

$$\frac{dX}{dt} = k_1 X^2/2 - k_2 X^3/6 + k_3 - k_4 X$$

FSE Schlögl

$$\begin{aligned}\frac{d[\![n]\!]}{dt} = &-\mathbf{I}_{n<\overline{O}_X}\{f_1(n) + f_3(n)\}\\ &- \mathbf{I}_{n>0}\{f_2(n) + f_4(n)\}\\ &+ \mathbf{I}_{1\le n\le \overline{O}_X} f_1(n-1)\\ &+ \mathbf{I}_{0\le n\le \overline{O}_X - 1} f_2(n+1)\\ &+ \mathbf{I}_{1\le n\le \overline{O}_X} f_3(n-1)\\ &+ \mathbf{I}_{0\le n\le \overline{O}_X - 1} f_4(n+1)\end{aligned}$$

$$\frac{dX}{dt} = f_1(\overline{O}_X) + f_3(\overline{O}_X) - f_2(0) - f_4(0)$$

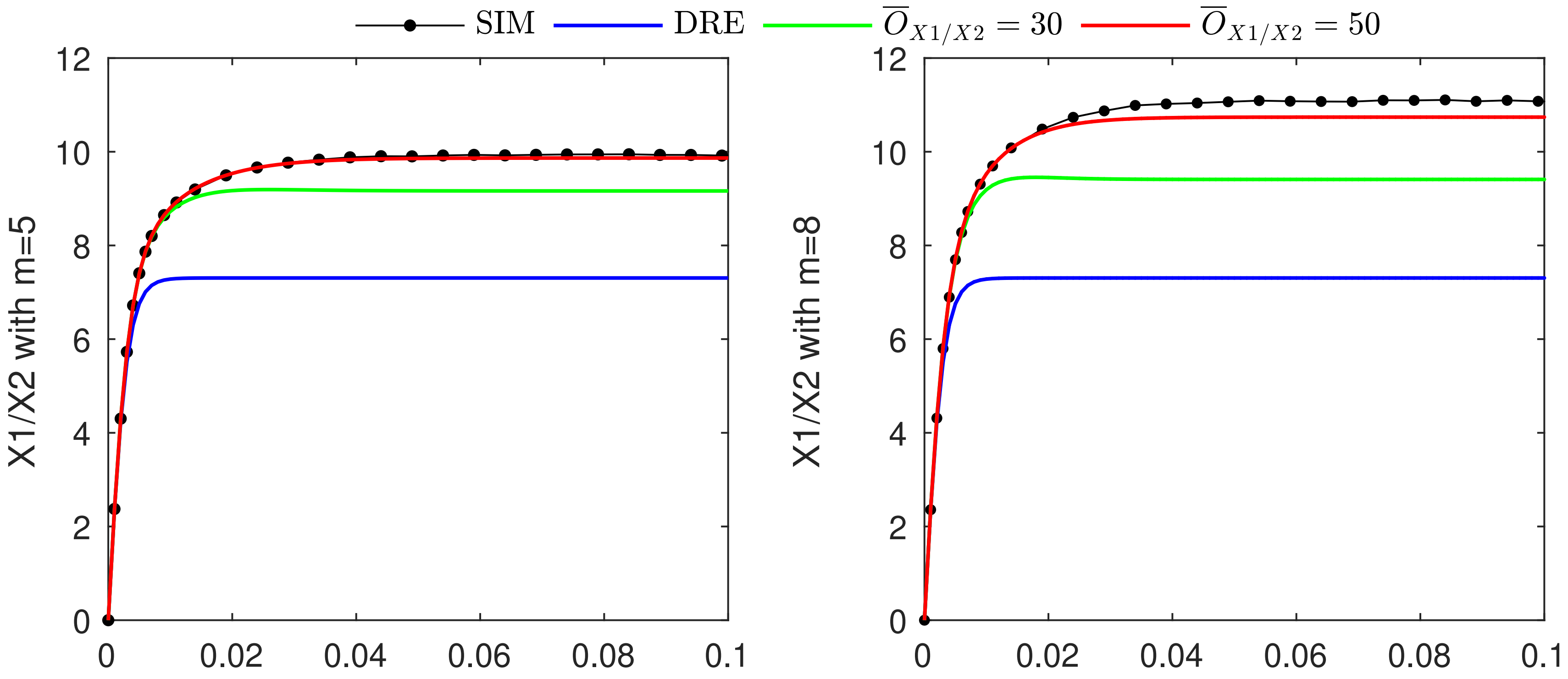


Case Studies

Heterodimerization

$$\xrightarrow{k_1} mX_1 \qquad \xrightarrow{k_1} mX_2 \qquad X_1 + X_2 \xrightarrow{k_2} X_3$$

$$X_1 \xrightarrow{k_3} \qquad X_2 \xrightarrow{k_3} \qquad X_3 \xrightarrow{k_3}$$

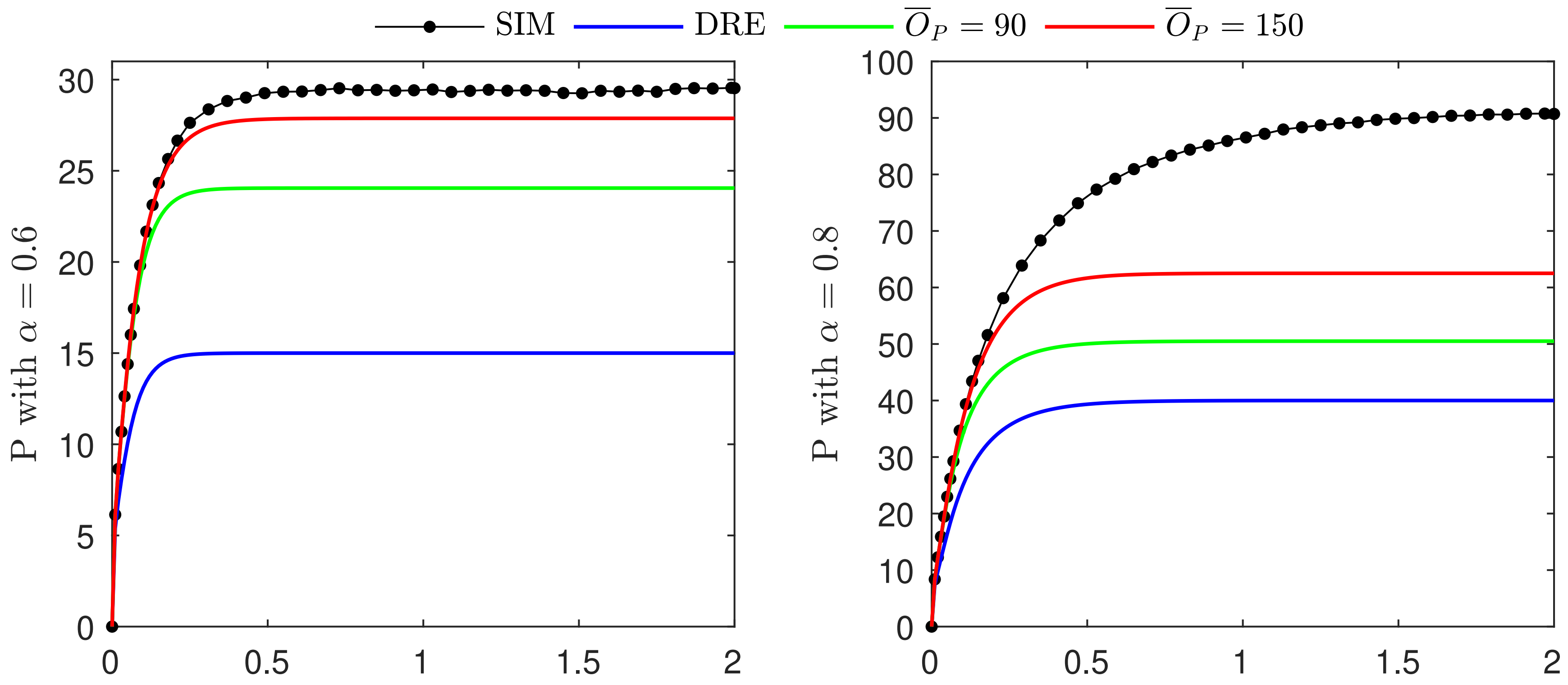


Case Studies

Protein degradation

$$\xrightarrow{k_1} mP \qquad P_d \xrightarrow{k_5}$$

$$P + E \xrightarrow{k_2} C \qquad C \xrightarrow{k_3} P + E \qquad C \xrightarrow{k_4} P_d + E$$



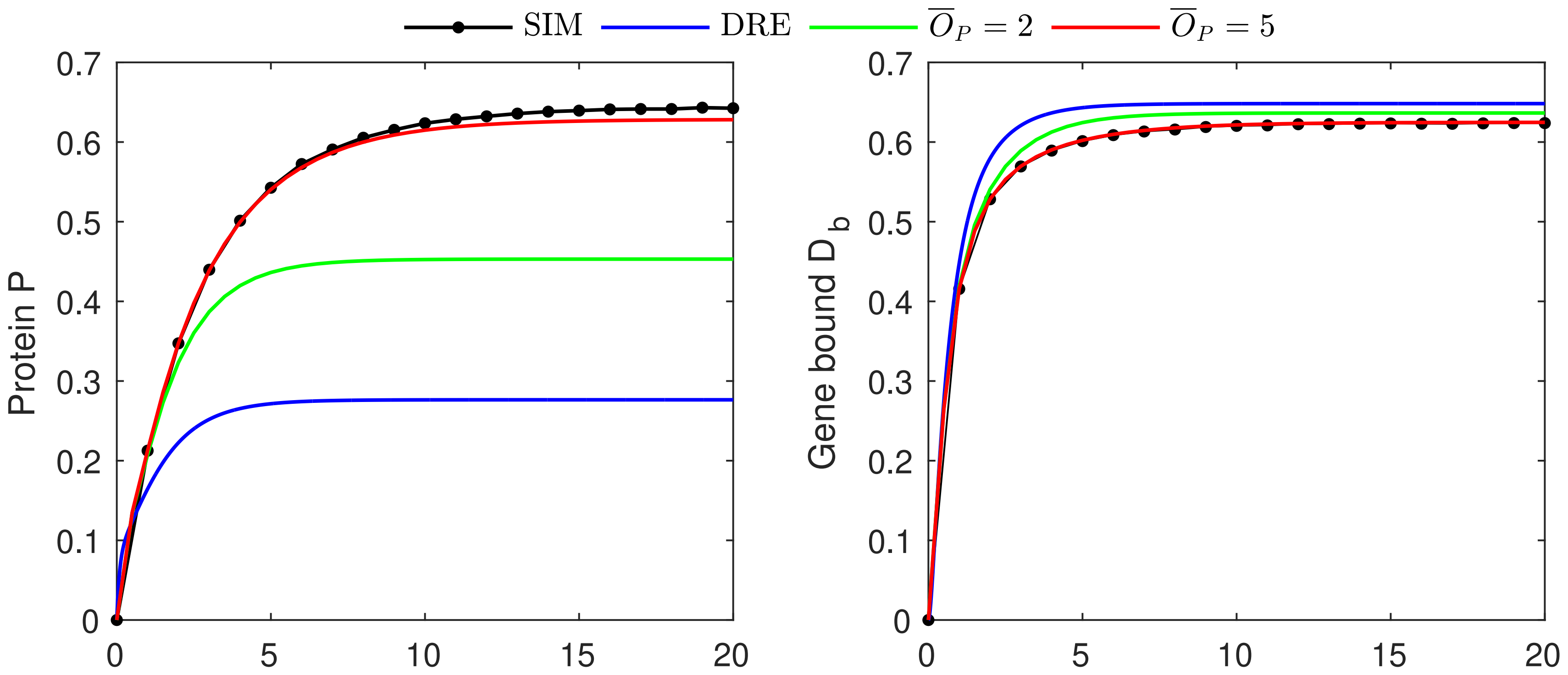
Case Studies

Gene feedback switch

$$D_u \xrightarrow{r_u} D_u + P \qquad D_b \xrightarrow{s_u} D_u + P$$

$$D_b \xrightarrow{r_b} D_b + P \qquad D_u + P \xrightarrow{s_b} D_b$$

$$D_b \xrightarrow{k_b} D_u \qquad P \xrightarrow{k_f}$$



Queuing System

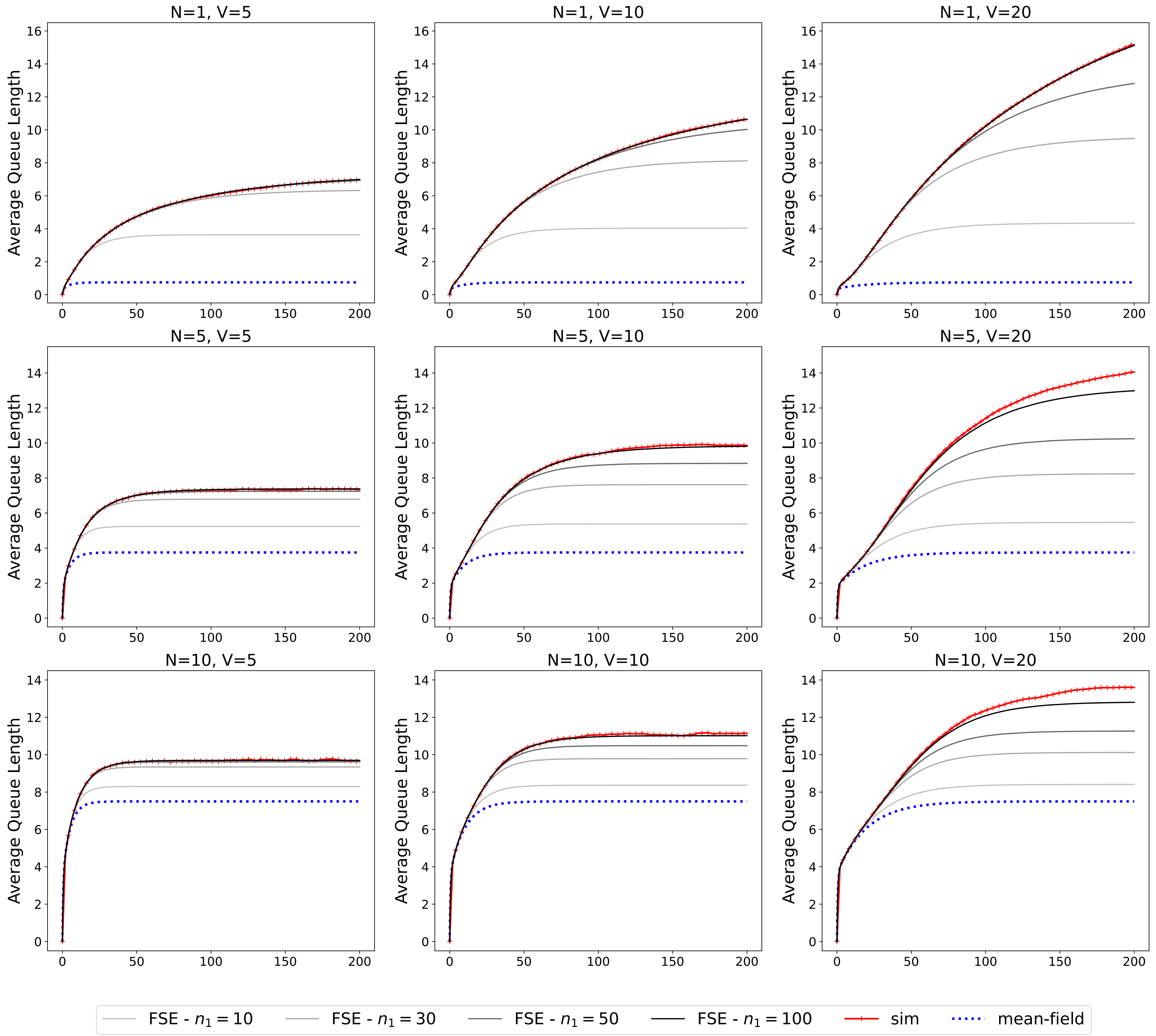
$$\xrightarrow{\lambda} Q_1$$

$$Q_1 + S_1 \xrightarrow{(1-p)\mu_1 \min(Q_1, S_1)} S_1$$

$$Q_1 + S_1 \xrightarrow{p\mu_1 \min(Q_1, S_1)} Q_2 + S_2$$

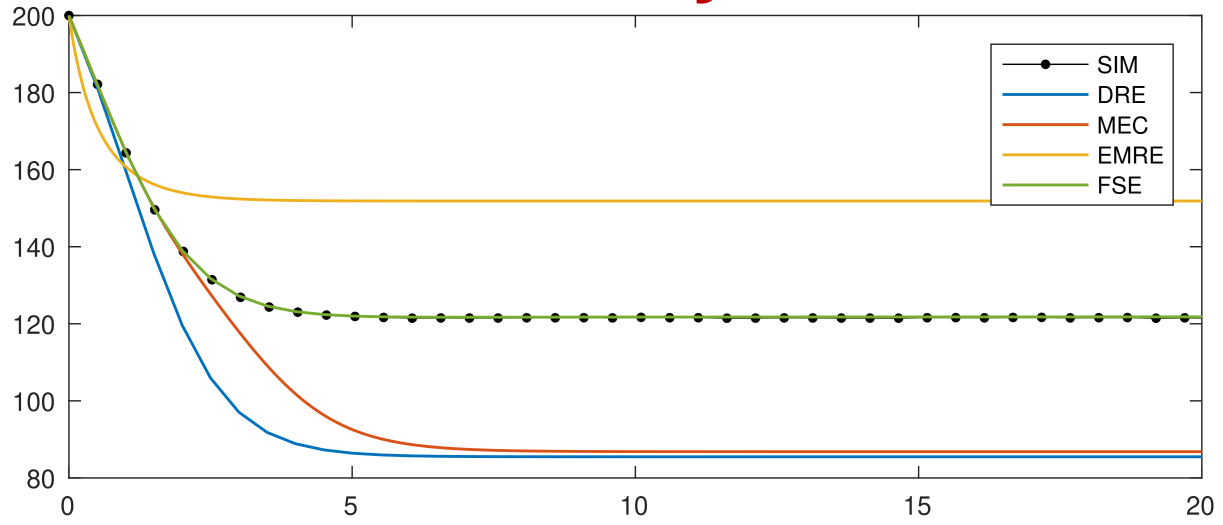
$$Q_2 + S_2 \xrightarrow{\mu_2 \min(Q_2, S_2)} S_1$$

- Textbook model with Poisson arrivals and Coxian-distributed service times with same mean and **increasing variance *V***
- *N* servers can simultaneously process client's requests
- Mean-field approximation is **insensitive to variance**
- Finite state expansion can track increasing variances

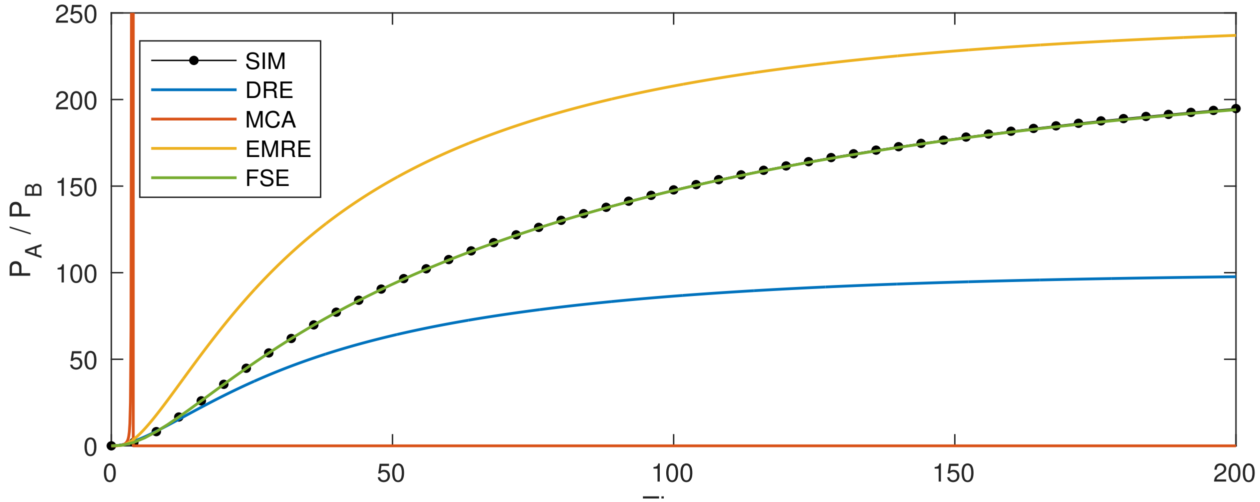


Comparison against State of the Art

Schlögl

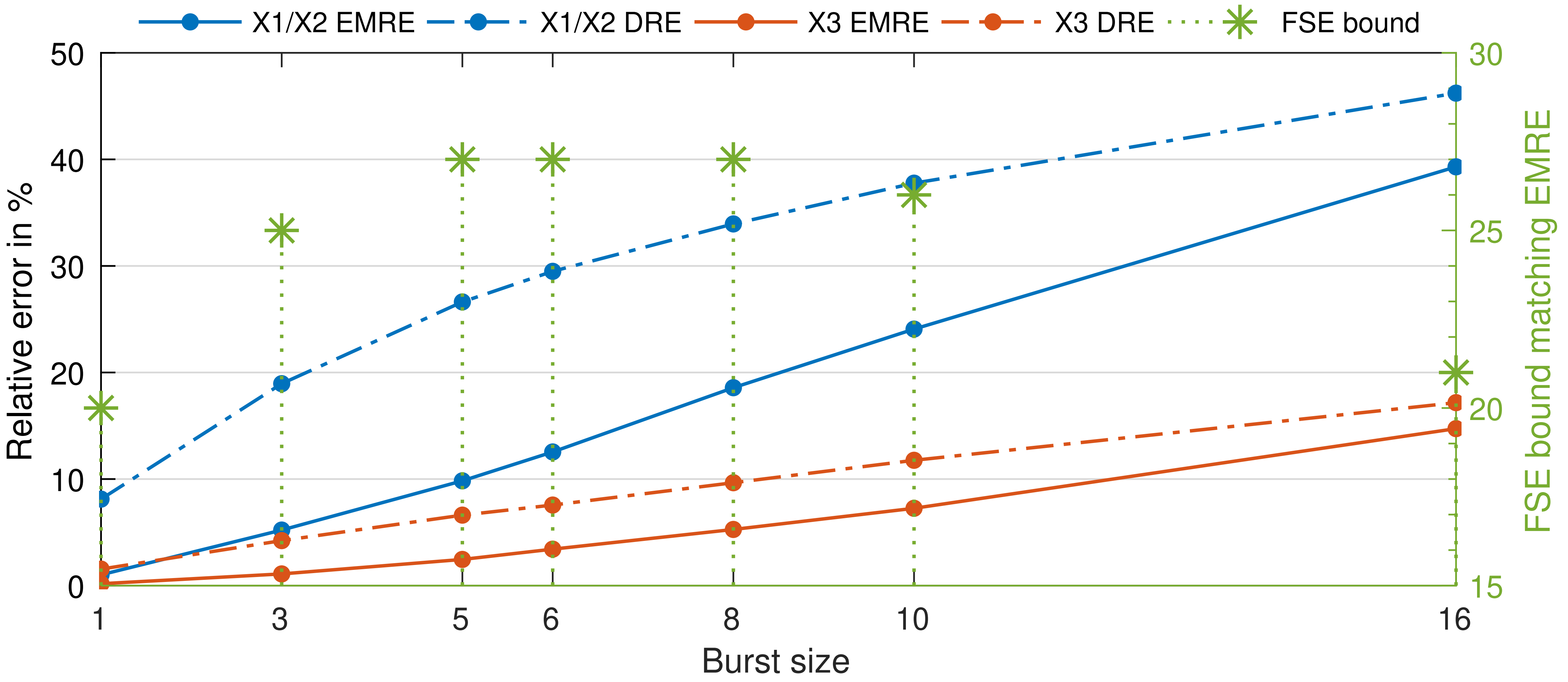


Toggle switch



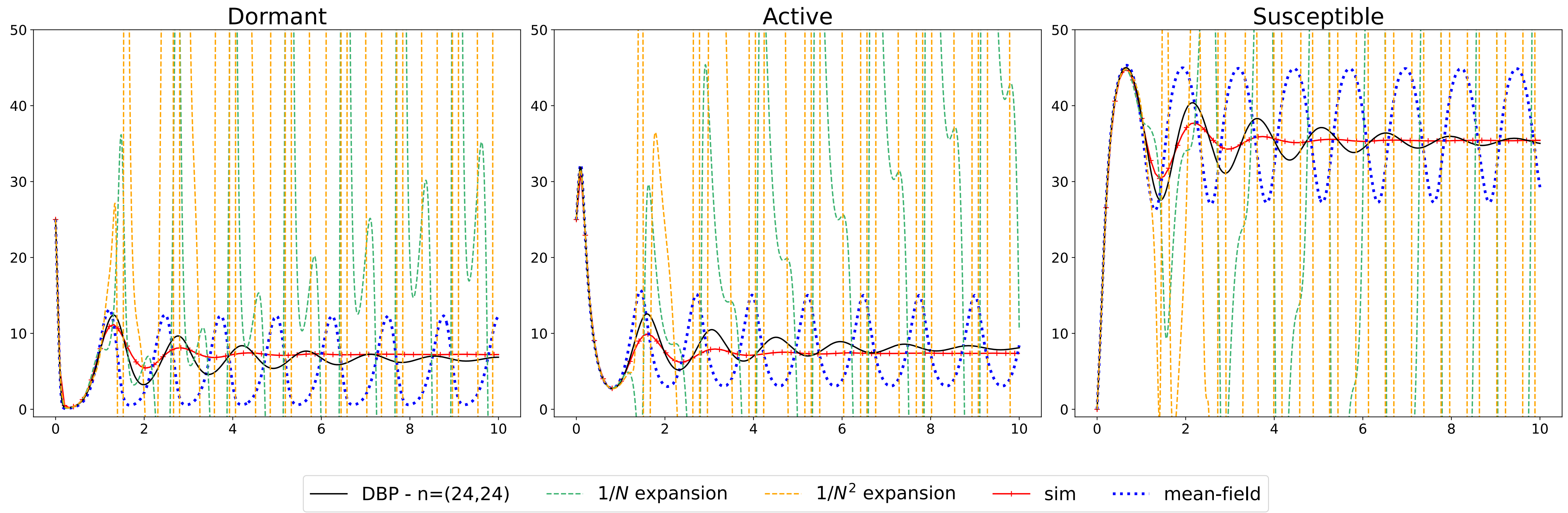
- EMRE (effective mesoscopic rate equations) and MEC (moment closure approximation) provide equations for first- and second-order moment
- In some cases (e.g. toggle switch) MCA gives **unphysical results**
- FSE improves on the accuracy of both; but it does so requiring **possibly many more equations**
- EMRE **cannot be used** if rates are not differentiable

Finite State Expansion on a Budget



Comparison Against 1/N and $1/N^2$ Expansions

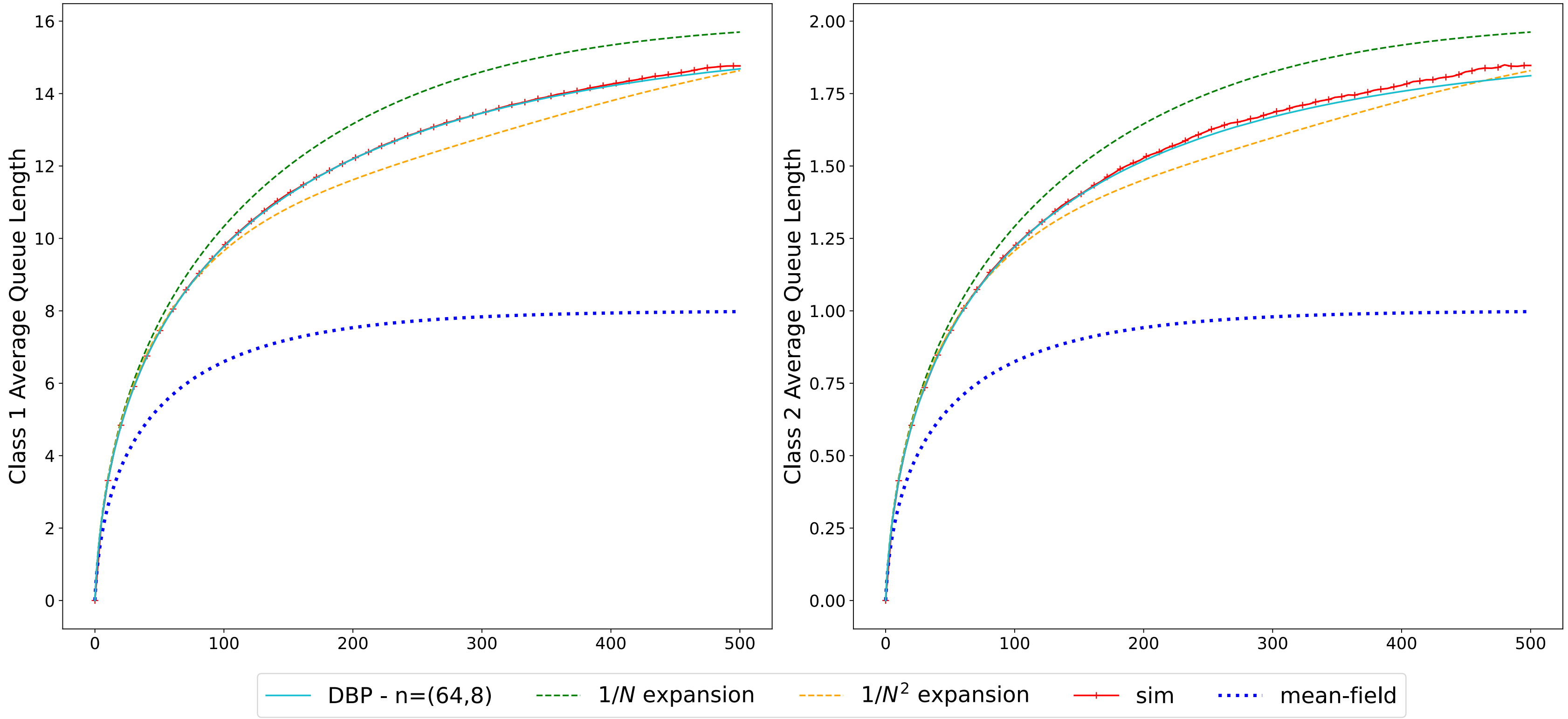
Malware progagation model



- 1/N expansion by Gast & Van Houdt, SIGMETRICS 2017
- $1/N^2$ expansion by Gast et al., PERFORMANCE 2019
- Both require differentiability of the vector field

Comparison Against 1/N and $1/N^2$ Expansions

Egalitarian Processor Sharing Queueing System

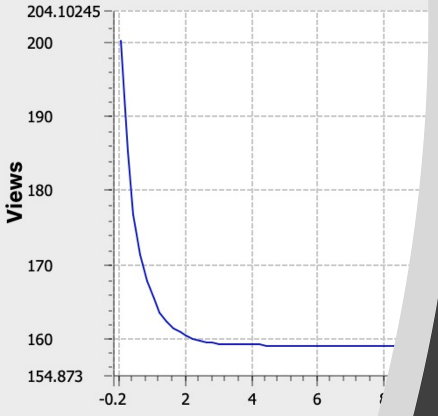


- 1/N expansion by Gast & Van Houdt, SIGMETRICS 2017
- $1/N^2$ expansion by Gast et al., PERFORMANCE 2019
- Both require differentiability of the vector field

```
begin model schloegl
begin parameters
k1 = 0.03
k2 = 0.0001
k3 = 200
k4 = 3.5
end parameters
begin init
X = 200
B = 1
SINK = 0
end init
begin reactions

2*X -> 3*X, k1
3*X -> 2*X, k2

B -> X + B, k3
X -> SINK, k4
end reactions

begin views
vX = X
end views

simulateCTMC(tEnd=20, repeats=3000,
    viewPlot=VIEWS, csvFile="schloegl_ssa")
fse(fileOut="schloegl_650.ode",limits=[X:650])
fse(fileOut="schloegl_450.ode",limits=[X:650])
simulateODE(tEnd=20, steps=100, viewPlot=VIEWS,
    library=SUNDIALS, csvFile="schloegl_ode")
end model
```

:=VIEWS,library=SUNDIALS,csvFile=/Use

schloegl - ODE solutions



Console

ERODE-schloegl_650-[13/09/2021

```
***************
 *************
*************
*************

Reading schl
        Par
        Sr

Solvi
Writ
Wr
```

www.erode.eu

Conclusion

- Starting point: **lumping of mean-field equations** – when reaction networks are huge to start with

  Cardelli et al., PNAS 2017

- How about the opposite, i.e. **expanding reaction networks**? FSE as a specific expansion algorithm...

  Waizmann et al., PRSA 2021

- ...but is actually an instance of a family of algorithms related to **finite state *projection***

  Randone et al., SIGMETRICS 2021

PNAS

Check for updates

Maximal aggregation of polynomial dynamical systems

Luca Cardelli[a,b,1], Mirco Tribastone[c,1,2], Max Tschaikowski[c,1], and Andrea Vandin[c,1]

[a]Microsoft Research, Cambridge CB1 2FB, United Kingdom; [b]Department of Computing, University of Oxford, Oxford OX1 3QD, United Kingdom; and [c]Scuola IMT Alti Studi Lucca, 55100 Lucca, Italy

Edited by Moshe Y. Vardi, Rice University, Houston, TX, and approved July 28, 2017 (received for review February 16, 2017)

Ordinary differential equations (ODEs) with polynomial derivatives are a fundamental tool for understanding the dynamics of systems across many branches of science, but our ability to gain mechanistic insight and effectively conduct numerical evaluations is critically hindered when dealing with large models. Here we propose an aggregation technique that rests on two notions of equivalence relating ODE variables whenever they have the same solution (backward criterion) or if a self-consistent system can be written for describing the evolution of sums of variables in the same equivalence class (forward criterion). A key feature of our proposal is to encode a polynomial ODE system into a finitary structure akin to a formal chemical reaction network. This enables the development of a discrete algorithm to efficiently compute the largest equivalence, building on approaches rooted in computer science to minimize basic models of computation through iterative partition refinements. The physical interpretability of the aggregation is shown on polynomial ODE systems for biochemical reaction networks, gene regulatory networks, and evolutionary game theory.

polynomial dynamical systems | aggregation | partition refinement

variables in a single block. Furthermore, the freedom in choosing an arbitrary initial partition is instrumental to producing reductions that preserve the dynamics of desired original variables, which are then not aggregated.

Mathematically, our approach is a generalization of well-known equivalence relations for Markov chains named lumpability (14). Ordinary lumpability relates states that have the same aggregate transition rate toward every equivalence class (thus, it is a forward criterion); in exact lumpability, two equivalent states have the same aggregate rate from every equivalence class (thus, it is a backward criterion). In a conceptually similar spirit, we define forward equivalence as a relation whereby each equivalence class describes the evolution of the sum of ODE variables in the original model. Backward equivalence identifies variables that have the same solutions at all time points (hence, they must start from the same initial conditions). Indeed, forward and backward equivalence collapse to ordinary and exact lumpability, respectively, when the (linear) ODE system is the equation of motion for the transient probability distribution of a continuous time Markov chain (15).

COMPUTER SCIENCES

PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

CrossMark
click for updates

Article submitted to journal

Subject Areas:
systems theory, computational biology

Improved estimations of stochastic chemical kinetics by finite state expansion

Tabea Waizmann[1], Luca Bortolussi[2], Andrea Vandin[3], and Mirco Tribastone[1]

[1]IMT School for Advanced Studies, Lucca, 55100, Italy,
[2]Department of Mathematics and Geosciences, University of Trieste, 34127, Italy, and
[3]Sant'Anna School of Advanced Studies, Pisa, 56127, Italy.

25

Refining Mean-field Approximations by Dynamic State Truncation

FRANCESCA RANDONE, IMT School For Advanced Studies Lucca, Italy
LUCA BORTOLUSSI, Università degli Studi di Trieste, Italy
MIRCO TRIBASTONE, IMT School For Advanced Studies Lucca, Italy

Perspectives

- Can we lump the discrete part while controlling overall accuracy? (Exact lumping is not possible in general) **YES, QEST 2022**
- How to effectively choose observation bounds?
- Monotonicity results? (Does not hold in general)
- Would other expansions improve the approximation with the same computational budget
- Can we approximate higher-order moments?
- ...

THANKS!

Advert

WE ARE HIRING!

Expression of interest for researchers working in:

- dynamical systems/stochastic processes
- software verification
- software engineering
- (cyber-)security

POSITIONS AVAILABLE AT POST-DOC AND ASSISTANT-PROFESSOR (TENURE- TRACK)

mirco.tribastone@imtlucca.it