Mean-field limit for a class of density-dependent stochastic processes

Illés Horváth

MTA–BME Information Systems Research Group

Nov. 4, 2014

based on joint work with Richard Hayden and Miklós Telek

Introduction

Starting model: Density-dependent Markov population process (1970s, Kurtz)

- finite population of identical components
- each component in a local state $s \in \mathcal{S}$
- each component performs Markovian (exponential) transitions simultaneously on $\mathcal{S}$
- transition rates depend on the *global state* of the sytem, e.g. the percentage of the population in each class

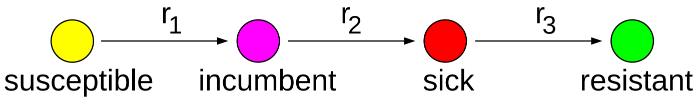
Example 1: Virus epidemic

A virus is spreading in a population of size N. Each individual can be in one of the following states:

- susceptible: healthy, but subject to infection
- incumbent: infectious, but no symptoms yet
- sick: symptoms visible, under treatment
- resistant: recovered, not susceptible to infection

Initially, a small percentage of the population is incumbent, the rest are susceptible.

Example 1: Virus epidemic



x_1, x_2, x_3, x_4: percentage of the population in each state.

r_1, r_2, r_3 are the transition rates, which may depend on $\mathbf{x}$, e.g. $r_1(\mathbf{x}) = c_1 x_2 + c_2 x_3$: the more individuals are infected in the general population, the more likely it is for a single individual to get infected. $r_2(\mathbf{x}) = c_3$ and $r_3(\mathbf{x}) = c_4$ are constants.

The mean-field limit

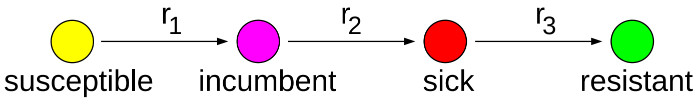
What is the behaviour of the system when N is large?

The system is random locally, but, as $N \to \infty$, the global evolution of the system converges to a deterministic limit called the mean-field limit.

The limit is defined by a system of $|\mathcal{S}|$ ordinary differential equations (ODE). The equations can be defined automatically.

Example 1: Virus epidemic

Back to the example:

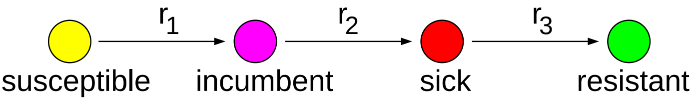


The system of ODEs corresponding to this system:

$$
\begin{aligned}
\dot{v}_1(t) &= -r_1(\mathbf{v}(t))v_1(t) \\
\dot{v}_2(t) &= r_1(\mathbf{v}(t))v_1(t) - r_2(\mathbf{v}(t))v_2(t) \\
\dot{v}_3(t) &= r_2(\mathbf{v}(t))v_2(t) - r_3(\mathbf{v}(t))v_3(t) \\
\dot{v}_4(t) &= r_3(\mathbf{v}(t))v_3(t)
\end{aligned}
$$

Example 1: Virus epidemic

Back to the example:



The system of ODEs corresponding to this system:

$$\dot{v}_1(t) = -(c_1 v_2(t) + c_2 v_3(t))v_1(t)$$
$$\dot{v}_2(t) = (c_1 v_2(t) + c_2 v_3(t))v_1(t) - c_3 v_2(t)$$
$$\dot{v}_3(t) = c_3 v_2(t) - c_4 v_3(t)$$
$$\dot{v}_4(t) = c_4 v_3(t)$$

Kurtz's theorem

Let $\mathbf{x}^N(t)$ be the random Markovian population process with population size N and let $\mathbf{v}(t)$ be the solution of the ODE. Assume also that $\|\mathbf{x^N}(\mathbf{0}) - \mathbf{v}(\mathbf{0})\| \to \mathbf{0}$ in probability and that r_c are Lipschitz-continuous (which holds in Example 1).

Then for any fixed $T > 0$ and $\varepsilon > 0$,

$$P\left(\sup_{t\in[0,T]} \|\mathbf{v}(\mathbf{t}) - \mathbf{x^N}(\mathbf{t})\| > \varepsilon\right) \to 0$$

as $N \to \infty$.

Main goal

We want to extend the setup to include non-Markovian transitions (generally-timed delays).

In many scenarios, exponential delays are not realistic. We are interested in introducing generally timed delays into the model described, and also to see how the mean field limit changes as a result.

History, related results

- Started in the 1970s (Kurtz)
- Related work in biology and chemistry; however, the results are very specific to their respective models
- Results with Petri nets (German 2000)
- Recent results with deterministic delays in a more general framework (Hayden 2012, Bortolussi–Hillston 2012)

Non-Markovian transitions

Introducing non-Markovian transitions (e.g. generally-timed delays) into density-dependent Markov population processes:

- some of the local states have an active generally-timed clock
- a cumulative distribution function F_e assigned to each active clock e
- when a component enters a state with an active clock, the clock sets to a random time according to F_e
- when the clock goes off, the component makes a (non-Markovian) transition

Non-Markovian transitions

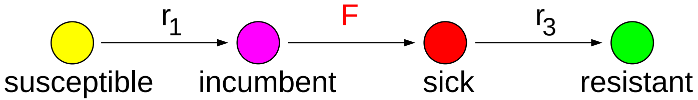
Introducing non-Markovian transitions (e.g. generally-timed delays) into density-dependent Markov population processes:

- some of the local states have an active generally-timed clock
- a cumulative distribution function F_e assigned to each active clock e
- when a component enters a state with an active clock, the clock sets to a random time according to F_e
- when the clock goes off, the component makes a (non-Markovian) transition

Main restrictions:

- at most one active clock in each local state; however, globally, any number of clocks may be active simultaneously
- two active clocks may not follow each other immediately
- an active clock can not be interrupted by Markovian transitions

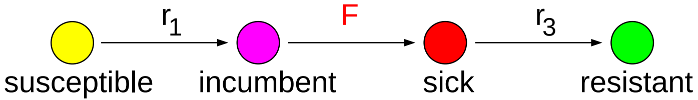
Example 1': Virus epidemic



Assume the length of the incumbent period is not exponential, but some general distribution with cumulative distribution function F instead. How to change the system of equations?

$$
\begin{aligned}
\dot{v}_1(t) &= -(c_1 v_2(t) + c_2 v_3(t)) v_1(t) \\
\dot{v}_2(t) &= (c_1 v_2(t) + c_2 v_3(t)) v_1(t) - ? \\
\dot{v}_3(t) &= ? - c_4 v_3(t) \\
\dot{v}_4(t) &= c_4 v_3(t)
\end{aligned}
$$

Example 1': Virus epidemic



Assume the length of the incumbent period is not exponential, but some general distribution with cumulative distribution function F instead. How to change the system of equations?

$$\dot{v}_1(t) = -(c_1 v_2(t) + c_2 v_3(t))v_1(t)$$

$$\dot{v}_2(t) = (c_1 v_2(t) + c_2 v_3(t))v_1(t) - \int_0^t (c_1 v_2(s) + c_2 v_3(s))v_1(s)\mathrm{d}F(s)$$

$$\dot{v}_3(t) = \int_0^t (c_1 v_2(s) + c_2 v_3(s))v_1(s)\mathrm{d}F(s) - c_4 v_3(t)$$

$$\dot{v}_4(t) = c_4 v_3(t)$$

Comparison of Markov and generalised processes

Purely Markov population process:

- memoryless
- approximated by the solution of a system of deterministic ordinary differential equations (ODE)

Comparison of Markov and generalised processes

Purely Markov population process:

- memoryless
- approximated by the solution of a system of deterministic ordinary differential equations (ODE)

Markov population process with generally-timed delays (or *generalised semi-Markov process, GSMP*):

- with memory (due to generally-timed delays)
- approximated by the solution of a system of deterministic *delayed differential equations* (DDE)

Example 2: Peer-to-peer software update

Two types of nodes: *old* and *updated*.

- old nodes search for updates in peer-to-peer fashion when turned on
- search is successful with a rate proportional to the number of updated nodes that are currently turned on
- if unsuccessful, the node gives up after a timeout and stays on for some time
- both updated and old nodes turn on and off

We assume that the off times are generally-distributed (in the following simulations, Pareto), everything else is Markovian.

Example 2: peer-to-peer software update

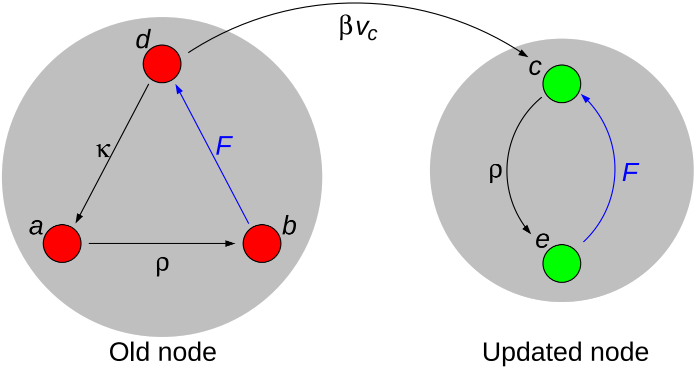
Local states.
Updated nodes:

- c: updated node turned on
- e: updated node turned off

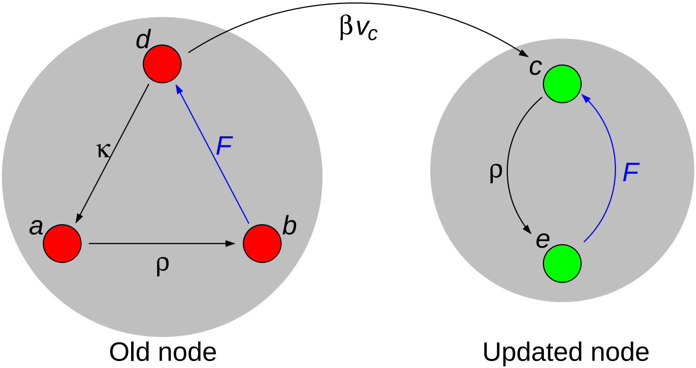
Old nodes:

- d: old node turned on, searching for updates
- a: old node turned on, no search for updates
- b: old node turned off

Example 2: peer-to-peer software update

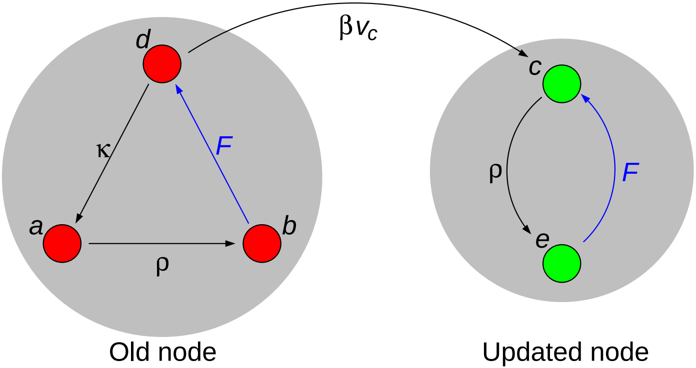


Example 2: peer-to-peer software update



$$\dot{v}_d(t) = -\kappa v_d(t) - \beta v_c(t) v_d(t) + \rho \int_0^t v_a(t-s)\,\mathrm{d}F(s)$$

Example 2: peer-to-peer software update



$$\dot{v}_b(t) = \rho v_a(t) - \rho \int_0^t v_a(t-s)\mathrm{d}F(s)$$

Example 2: peer-to-peer software update

$$\dot{v}_a(t) = \kappa v_d(t) - \rho v_a(t)$$

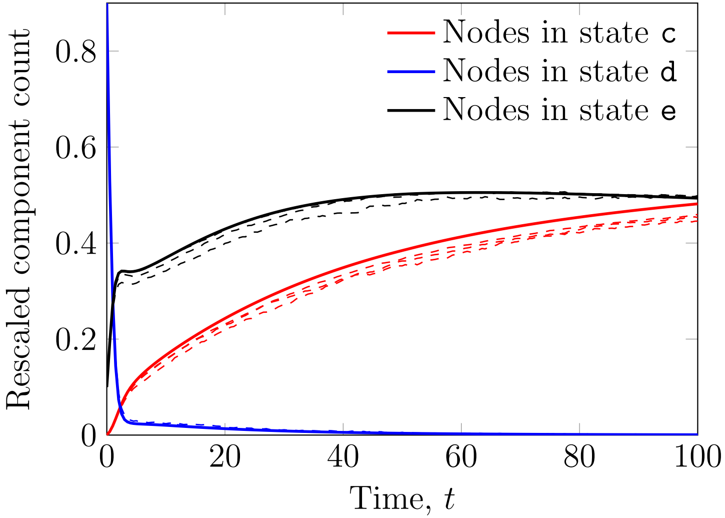
$$\dot{v}_b(t) = \rho v_a(t) - \rho \int_0^t v_a(t-s)\mathrm{d}F(s)$$

$$\dot{v}_c(t) = \beta v_d(t) v_c(t) - \rho v_c(t) + + \rho \int_0^t v_c(t-s)\mathrm{d}F(s)$$

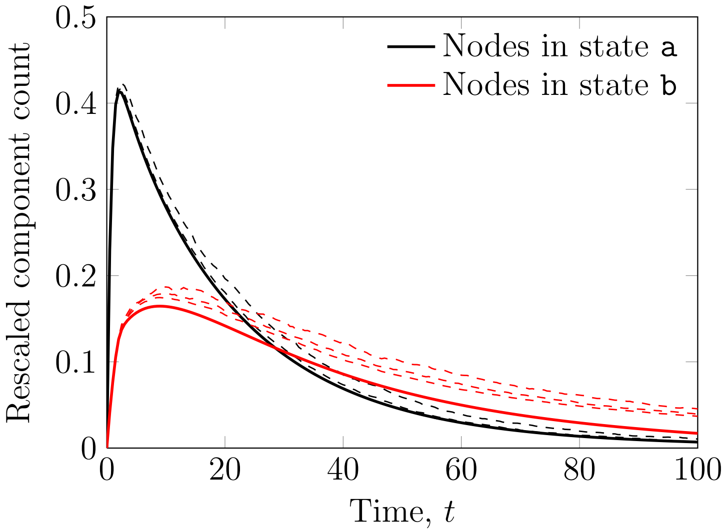
$$\dot{v}_d(t) = -\kappa v_d(t) - \beta v_c(t) v_d(t) + \rho \int_0^t v_a(t-s)\mathrm{d}F(s)$$

$$\dot{v}_e(t) = \rho v_c(t) - \rho \int_0^t v_c(t-s)\mathrm{d}F(s)$$

Example 2: peer-to-peer software update



Example 2: peer-to-peer software update



General setup

General setup:

- set of local states $\mathcal{S}$
- set of Markovian transitions $\mathcal{C} \subseteq \mathcal{S} \times \mathcal{S}$
- r_c: aggregate rate of Markovian transition $c \in \mathcal{C}$
- set of generally-timed transitions $\mathcal{E} \subseteq \mathcal{S} \times \mathcal{S}$
- F_e: CDF of general delay $e \in \mathcal{E}$

Let $\mathbf{x^N}$ be the random GSMP on N components.

General setup

The general DDE:

$$\dot{v}_s(t) = \sum_{c \in \mathcal{C}} l_s^c r_c(\mathbf{v}(t)) + \sum_{e \in \mathcal{E}} \sum_{c \in \mathcal{C}} h_s^{c,e} \int_0^t r_c(\mathbf{v}(t-u)) \, \mathrm{d}F_e(u)$$

(l_s^c and $h_s^{c,e}$ are ± 1 or 0.)

Let $\mathbf{v}$ be the solution.

General setup

Given the prior assumptions plus the following:

- $\|\mathbf{v}(\mathbf{0}) - \mathbf{x}^{\mathbf{N}}(\mathbf{0})\| \to \mathbf{0}$ in probability
- r_c are Lipschitz-continuous

Theorem (Hayden–Telek–H., 2014)

For any fixed $T > 0$ and $\varepsilon > 0$,

$$P\Big(\sup_{t \in [0,T]} \|\mathbf{v}(\mathbf{t}) - \mathbf{x}^{\mathbf{N}}(\mathbf{t})\| > \varepsilon \Big) \to \mathbf{0}$$

as $N \to \infty$.

Main elements of the proof: Gronwall's inequality, Poisson-representation, probability concentration theorems (law of large numbers, Azuma's inequality).

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

The main practical advantages of this approximation:

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

The main practical advantages of this approximation:

- generally-timed delays are often relevant in real-world computer and communication systems

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

The main practical advantages of this approximation:

- generally-timed delays are often relevant in real-world computer and communication systems
- no state space explosion – the number of equations is independent of N

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

The main practical advantages of this approximation:

- generally-timed delays are often relevant in real-world computer and communication systems
- no state space explosion – the number of equations is independent of N
- the system of DDEs can be derived automatically from the GSMP

Advantages

The main theoretical result is a rigorous mathematical proof of the mean-field limit for a large class of random processes (GSMPs).

The main practical advantages of this approximation:

- generally-timed delays are often relevant in real-world computer and communication systems
- no state space explosion – the number of equations is independent of N
- the system of DDEs can be derived automatically from the GSMP
- the system of DDEs can be solved efficiently numerically

Further questions I - the race case

What if we drop the assumption that non-Markovian transitions can not be interrupted?

In this case, Markovian and non-Markovian transitions *race* – if a Markovian transition occurs before an active clock would go off, the clock is cancelled.

The race case

Let $\mathbf{v}$ be the solution of the DDE (in integral form)

$$v_s(t) = v_s(0) + \sum_{c\in\mathcal{C}} l_s^c \int_0^t r_c(\mathbf{v}(s))\,\mathrm{d}s + \sum_{e\in\mathcal{E}} \sum_{c\in\mathcal{C}} h_s^{c,e} \times$$
$$\times \int_{z=0}^{t} \int_{x=0}^{t-z} \exp\left(- \int_{\tau=z}^{z+x} q_s(\mathbf{v}(\tau))\mathrm{d}\tau \right) \mathrm{d}F_e(x) r_c(\mathbf{v}(z))\mathrm{d}z$$

where

$$q_s(\mathbf{v}) = \frac{1}{v_s} \sum_c r_c(\mathbf{v}),$$

where the sum goes over all transitions c which go out from local state s. q_s may be interpreted as the *rate of risk* for a component in local state s to be interrupted by a Markovian transition.

Convergence not fully proven yet, work in progress.

Other methods to handle non-Markovian transitions

Phase-type distributions: the vanishing time of a vanishing Markov process. More general than exponential distributions, and can be used to approximate general distributions.

Other methods to handle non-Markovian transitions

Phase-type distributions: the vanishing time of a vanishing Markov process. More general than exponential distributions, and can be used to approximate general distributions.

By approximating generally distributed delays by phase-type distributions, the obtained model is fully Markovian, albeit on a much larger state space, making this approach impractical to examine population models.

Other methods to handle non-Markovian transitions

Phase-type distributions: the vanishing time of a vanishing Markov process. More general than exponential distributions, and can be used to approximate general distributions.

By approximating generally distributed delays by phase-type distributions, the obtained model is fully Markovian, albeit on a much larger state space, making this approach impractical to examine population models.

There is a direct analogue of this approach for differential equations: a system of DDEs can be approximated by a larger system of ODEs (Maset, 2003).

Further questions

- Second-order approximation (fluctuation around the mean-field limit)

Further questions

- Second-order approximation (fluctuation around the mean-field limit)
- Multiple generally-timed transitions racing – description by PDEs? – would require completely different methods

Further questions

- Second-order approximation (fluctuation around the mean-field limit)
- Multiple generally-timed transitions racing – description by PDEs? – would require completely different methods

Thank you for your attention!