Mean field for performance models with deterministic delays and interrupts

Illés Horváth[a,*], Miklós Telek[b,]

[a]MTA-BME Information Systems Research Group, Magyar Tudósok krt. 2., 1117 Budapest, Hungary
[b]Budapest University of Technology and Economics, Department of Networked Systems and Services, Magyar Tudósok krt. 2., 1117 Budapest, Hungary

Abstract

The mean-field limit of stochastic models with exponential and deterministic delays has been proved for the case when the deterministic delays cannot be interrupted by an exponential one.

In this paper we extend the mean-field limit for the class of stochastic models with exponential and deterministic delays where the activities with exponential delays can interrupt the ones with deterministic delays. Our main focus is the rigorous proof of the mean-field limit for this case.

Keywords: mean-field limit, deterministic delay, delayed differential equation

1. Introduction

The classic result of Kurtz [17] states that the mean-field limit of a density dependent Markov population process can be described as the solution of a system of ordinary differential equations (ODEs) corresponding to the system. In this paper, we extend this result to a class of processes called *population generalised semi-Markov processes (PGSMPs)*, where individuals can enable both Markovian and deterministic transitions. The main novelty of the paper is that it encompasses the *race case*, e.g. deterministic transitions compete with exponential transitions (as opposed to delay-only PGSMPs). The main focus is the mathematically rigorous setup and proof of the transient mean-field convergence. To the best of our knowledge, this is the first rigorous result that deals with the race case.

The main motivation is that the delay-only restriction may be severe even in the most simple models. Allowing the race between exponential and deterministic transitions offers more flexibility in modelling. This will be addressed in more detail in the running example.

Interruptible delays arise naturally in many applications in many different fields whenever multiple effects apply simultaneously. In computer systems, for example, searches for software updates are interrupted when an update is found (actually, a peer-to-peer software update model will serve as our running example). Another example is disease spreading, when the situation of a patient may change due to effects such as vaccination, diagnosis or quarantine.

Previous research for PGSMP models have been initiated in various fields, each carrying various notation and focusing on different aspects of the topic. Closest related work is due to [14] and [16]. [14] presents essentially the same framework, but with the deterministic transitions delay-only, that is, deterministic transitions are non-interruptible by Markovian transitions. [16] discusses generally-timed delay-only transitions. [6] presents a different approach, highlighting to connection to ODE approximations of DDEs [19] which is analogous to the Erlang approximation of the delay in the PGSMP. These papers approach the problem with an eye to accurate modelling of real-world computer and networking protocols.

Other related work is present in biology and chemistry literature. DDEs have been used to approximate reaction networks where deterministic delays occur after reactions [3, 7, 20]. However, these models are typically very specific to the given application, while the current paper takes a general modeling approach, making it suitable for a much larger class of population models. Some recent related work in biology may be found in [4], where mean-field methodology is applied in a slightly different manner, with a more detailed (and thus more precise) system of equations, albeit for a less general model.

The motivation for the mean-field approach is the same as in the continuous-time Markov chain (CTMC) case [17] — unsurprisingly, GSMP models with many components also become computationally intractable to explicit state techniques [8, 10] rapidly as a result of the familiar state-space explosion problem. Our approach is based on the derivation of *delayed differential equations (DDEs)* from PGSMP models and

*Corresponding author
Email addresses: horvath.illes.antal@gmail.com (Illés Horváth), telek@hit.bme.hu (Miklós Telek)

generalises the traditional mean-field approach as applied to CTMC models based on ordinary differential equations (ODEs) [2, 13, 5, 15].

The class of models to which our approach applies is quite broad — the only significant restriction we make is that at most one deterministic transition may be enabled by each individual in any given local state and that the execution of a deterministic transition cannot activate a new deterministic delay immediately. However, globally, there is no restriction on the concurrent enabling of deterministic transitions by different individuals.

Handling interruptible deterministic delays presents a new challenge in the analysis of PGSMPs and calls for technology to handle them. In the present paper, the classic concept of Poisson representation that keeps track of the number of transitions according to transition type is enhanced by indicator variables that correspond to the survival or interruption of deterministic clocks (see (1)). The novel elements of the proof, e.g. Lemma 2, also focus on these terms. In the mean-field limit, the indicators turn into exponential terms (see (2)), keeping the formulas relatively simple. Numerical solutions are still feasible as well.

The rest of the paper is structured as follows. In Section 2, we set up notation and give a precise definition of the model, the evolution of the system; the mean-field limit is also described, and an example is included. In Section 3, we state the main result and give a mathematically rigorous proof. In Section 4, we conclude and give an outlook at further related questions. The Appendix contains proofs of some of the lemmas used along the way and also a numerical example.

2. Setup

2.1. Definitions and assumptions

A PGSMP model consists of N interacting components (individuals), each of which is in a state from a finite *set of local states* $\mathcal{S}$.

Each component is subject to Markovian transitions in continuous time. When it is in state $i \in \mathcal{S}$, it transitions to state $j \in \mathcal{S}$ with rate r_{ij}^N ($r_{ij}^N : \mathbb{Z}^N \to [0, \infty]$). The rates may depend on the *global state of* the system; the global state of the system is defined as the total number of individuals in each state, that is, a vector $\mathbf{x}^N \in (\{0, 1, \ldots, N\})^{\mathcal{S}}$ with $x_1^N + \cdots + x_{|\mathcal{S}|}^N = N$. The normalized global state of the system is defined as $\bar{\mathbf{x}}^N = \frac{\mathbf{x}^N}{N}$, so $\bar{\mathbf{x}}^N \in [0,1]^{\mathcal{S}}$ with $\bar{x}_1^N + \cdots + \bar{x}_{|\mathcal{S}|}^N = 1$. $\bar{\mathbf{x}}^N(t)$ tracks the *ratio* of the components in each state at time t, and will be used as the underlying stochastic process of the model.

We assume r_{ij}^N to be *density-dependent*, that is, for each pair $i, j \in \mathcal{S}$

$$r_{ij}^N(N\mathbf{x}) = r_{ij}(\mathbf{x})$$

for a common function $r_{ij} : [0,1]^{\mathcal{S}} \to [0, \infty]$ *independent of* N (in fact, r_{ij} needs to be defined only on the simplex $x_1 + \cdots + x_N = 1$). To simplify notation, we assume $r_{ii} \equiv r_{ii}^N \equiv 0$.

We assume that r_{ij} are Lipschitz-continuous with common Lipschitz-constant R^L. R^M is an upper bound on r_{ij}. Let $R = \max\{R^L, R^M\}$.

We also include deterministic transitions. Some of the states have a so-called *active clock*. We will call these *active states* and denote the set of active states by $\mathcal{S}_A \subseteq \mathcal{S}$ and the set of inactive states by $\mathcal{S}_I = \mathcal{S} \setminus \mathcal{S}_A$. When an individual enters a state $i \in \mathcal{S}_A$, a deterministic clock with a deterministic clock time T_i is initialized.

If the component performs a Markovian transition before the deterministic clock goes off, we say that the deterministic clock is interrupted; in other words, the Markovian transitions race with the deterministic transitions. If no Markovian transitions occur in the component for time T_i after the deterministic clock started, the component makes a deterministic transition. We assume that the deterministic transition from state i always targets the same state j; we formulate this by introducing a transition vector associated with type i deterministic transitions. If a component makes a type i deterministic transition from state i to state j then p_{ij} is equal to 1 and $p_{ih} = 0$ for $\forall h \in \mathcal{S} \setminus j$. This technical restriction on the same destination of deterministic transitions is easy to relax on the expense of handling more cases and notations. The important modelling restriction is that deterministic transitions do not follow each other directly. That is if for some $i \in \mathcal{S}_A$ $p_{ij} = 1$ then $j \in \mathcal{S}_I$. For notational convenience we also define p_{ij} for $i \in \mathcal{S}_I$ such that $p_{ij} = 0$ for $\forall i \in \mathcal{S}_I, j \in \mathcal{S}$. Note that a Markovian transition from an active state to another active state is allowed.

We also assume that the initial state of the system is concentrated on $\mathcal{S}_I$; in other words, no generally-timed clocks are active initially. For a model where general initial condition is examined, see [6].

We are interested in the evolution of the process $\bar{\mathbf{x}}^N(t)$. $\bar{\mathbf{x}}^N(t)$ in itself is not a continuous-time Markov-process since it does not include the times when the deterministic clocks are set to go off. However, this information is available from the past of the process, as we shall see in the next subsection.

2.2. Evolution of the process

The formal definition of the evolution of the process $\bar{x}^N$ is done by Poisson representation:

$$\begin{aligned}\bar{x}_i^N(t) = \bar{x}_i^N(0) - \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ij}\left(N\int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) + \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ji}\left(N\int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) \\ + \sum_{j\in\mathcal{S}_A}\int_{z=0}^{t-T_j} p_{ji} B^{j,N}_{\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right) \\ - \sum_{j\in\mathcal{S}}\int_{z=0}^{t-T_i} p_{ij} B^{i,N}_{\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right) \quad (1)\end{aligned}$$

for $i \in \mathcal{S}$. For each i, j, let P_{ij} be an independent Poisson-process with rate 1. $B_k^{i,N}(z)$ is an indicator random variable that is equal to 1 if clock number k of type i (which refers to active clocks in state i) starts at time z, it is *not interrupted by Markovian transitions*, and, after time T_i, makes a transition to state j. (The clocks are ordered by starting time.)

Remarks for the formula (1):

1. The first term corresponds to the initial condition.
2. The second and third term correspond to Markovian transitions to and from state i. Note that

$$N\bar{x}_i^N(t) r_{ij}(\bar{\mathbf{x}}^N(t)) = x_i^N(t) r_{ij}^N(\mathbf{x}^N(t))$$

   is the *aggregate rate* of an $i \to j$ Markovian transition at time t, that is, the rate that any of the $x_i^N(t)$ components in state i makes a transition to state j.
3. The fourth term counts the deterministic transitions to i up to time t. Per our assumption that an active state is always followed by an inactive state the fourth term may be nonzero only if $i \in \mathcal{S}_I$, otherwise $p_{ji} = 0, \forall j \in \mathcal{S}_A$. Consider an active state j with $p_{ji} = 1$. Clocks of type j are activated whenever a Markov transition to j occurs from some state h. z denotes the points in time of such activations: the counting process $\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)$ jumps whenever a transition to state j from some state h occurs. So the notations for a single clock are the following: a Markovian transition from h to j at time z activates a clock of type j. If $z > t - T_j$, this clock will have no effect before time t. If $z \le t - T_j$, the clock will result in a deterministic transition before time t iff it has not been interrupted by Markovian transitions; this event is encoded in $B_k^{j,N}(z)$. The variables $B_k^{j,N}(z)$ are further discussed later.
4. Note that in case clock j can not be interrupted because $r_{ji} = 0$ for all $i \in \mathcal{S}$, then $B_k^{j,N}(z) = 1$.
5. The fifth term counts the deterministic transitions from i. Similarly, because an active state is always followed by an inactive state, the fifth term may be nonzero only for $i \in \mathcal{S}_A$, otherwise $p_{ij} = 0, \forall j \in \mathcal{S}$. So at most one of the fourth and fifth terms may be nonzero for each state i.
6. The Poisson representation offers a coupling of the system for various different values of N, placing all models (for different values of N) in the same probability space. This allows to state some results which hold almost surely. Nevertheless, for the main result, we will stick to convergence in probability.

The Poisson representation also offers a pseudo-algorithm for discrete event simulation; nevertheless, we describe how to generate such a process in more detail.

2.3. Simulation

The simulator represents the system state by the population vector $\bar{\mathbf{x}}^N$ and by the set of markers of each type indicating the scheduled transition time of active deterministic transitions.

Initially no element of the population is in an active state and consequently there is no active marker. Thus the first transition will be Markovian. Since an individual transition from i to j occurs with total rate $N\bar{x}_i(0) r_{ij}(\bar{\mathbf{x}}^N(0))$, the first transition may be generated by considering the first arrival from among all processes $P_{ij}(N\bar{x}_i(0) r_{ij}(\bar{\mathbf{x}}^N(0))t)$, e.g. select i and j according to

$$\mathrm{argmin}_{(i,j)} \inf_{t>0} \left\{t : P_{ij}(N\bar{x}_i(0) r_{ij}(\bar{\mathbf{x}}^N(0))t) = 1\right\}$$

and let s_1 be the time of this first arrival. In the $(0, s_1)$ interval the population vector is $\bar{\mathbf{x}}^N(0)$. After the transition the new population vector is $\bar{\mathbf{x}}^N - e_i/N + e_j/N$ (where e_i is the ith unit vector of size $|\mathcal{S}|$) and the rates of all Poisson-processes change according to the change of the population vector.

We say that after this initial step the simulation time advances to time s_1 and the simulation goes on based on the system state at the current simulation time. That is, the next interval is simulated similarly based on the population vector (and the active markers, if any) at the beginning of the next interval.

In general, when deterministic clocks might also be initialized, the elementary step of the simulation modifies as follows. Assuming the simulation time is t_n we find the next Markovian transition by

$$\operatorname{argmin}_{(i,j)} \inf_{t>0} \left\{ t : P_{ij}(N\bar{x}_i(t_n) r_{ij}(\bar{\mathbf{x}}^N(t_n))t) = 1 \right\}$$

and the associated time s_n (to avoid cumbersome notation, here we assume that a new Poisson process is used in each step) and check if the minimum of all markers is less than $t_n + s_n$ or not.

- In case the first Markovian transition at $t_n + s_n$ precedes all the deterministic markers, a Markovian transition from state i to state j will occur. Depending on whether i and j are active or not we have the following cases.
  - $i \in \mathcal{S}_I$, $j \in \mathcal{S}_I$: the occupation vector changes to $\bar{\mathbf{x}}^N(t_n) - e_i/N + e_j/N$ and the simulation time advances to $t_n + s_n$.
  - $i \in \mathcal{S}_A$, $j \in \mathcal{S}_I$: the occupation vector changes to $\bar{\mathbf{x}}^N(t_n) - e_i/N + e_j/N$, the simulation time advances to $t_n + s_n$ and the transition interrupts the active clock in one of the components in state i. The interrupted clock is uniformly selected from the $N\bar{x}_i^N(t_n)$ active clocks. In the simulation it is implemented by deleting a uniformly sampled marker of type i. (If the selected clock was the kth activated type i clock then it means that $B_k^{i,N}$ is 0.)
  - $i \in \mathcal{S}_I$, $j \in \mathcal{S}_A$: the occupation vector changes to $\bar{\mathbf{x}}^N(t_n) - e_i/N + e_j/N$, the simulation time advances to $t_n + s_n$, and a new deterministic clock of type j is activated. That is a new marker of type j is set to $t_n + T_j$.
  - $i \in \mathcal{S}_A$, $j \in \mathcal{S}_A$: the occupation vector changes to $\bar{\mathbf{x}}^N(t_n) - e_i/N + e_j/N$, the simulation time advances to $t_n + s_n$, a new deterministic clock of type j is activated, and an active clock of type i is interrupted (as detailed above).
- In case a deterministic marker of type m is minimal among the markers and is less than the first Markovian transition at $t_n + s_n$, the corresponding deterministic transition from state m to state ℓ will occur (where ℓ is such that $p_{m\ell} = 1$) and the occupation vector changes accordingly; the simulation time advances to the time of this minimal marker; and the marker with minimal time gets deleted. (If the clock with the minimal marker was the kth activated type m clock then it means that $B_k^{m,N}$ is 1. The activation of a new deterministic clock is not possible in this case because a deterministic transition can not activate deterministic clock.)

We refer to [1] for a simulation code of this procedure in Wolfram Mathematica.

2.4. The mean-field equation

We define the mean-field equation corresponding to a PGSMP defined in Section 2.1 as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} v_i(t) = &- \sum_{j\in\mathcal{S}} v_i(t) r_{ij}(\mathbf{v}(t)) + \sum_{j\in\mathcal{S}} v_j(t) r_{ji}(\mathbf{v}(t)) \\
&+ \sum_{h\in\mathcal{S}_I} \sum_{j\in\mathcal{S}_A} p_{ji} \exp\left(- \int_{\tau=0}^{T_j} q_j(\mathbf{v}(t - T_j + \tau)) \mathrm{d}\tau \right) v_h(t - T_j) r_{hj}(\mathbf{v}(t - T_j)) \\
&- \sum_{h\in\mathcal{S}_I} \sum_{j\in\mathcal{S}} p_{ij} \exp\left(- \int_{\tau=0}^{T_i} q_i(\mathbf{v}(t - T_i + \tau)) \mathrm{d}\tau \right) v_h(t - T_i) r_{hi}(\mathbf{v}(t - T_i)),
\end{aligned} \tag{2}$$

where

$$q_i(\mathbf{v}) = \sum_{j\in\mathcal{S}} r_{ij}(\mathbf{v}).$$

q_i may be interpreted as the *rate of risk* for a component in local state i to be interrupted by a Markovian transition. q_i are Lipschitz-continuous with Lipschitz constant $Q = |\mathcal{S}|R$.

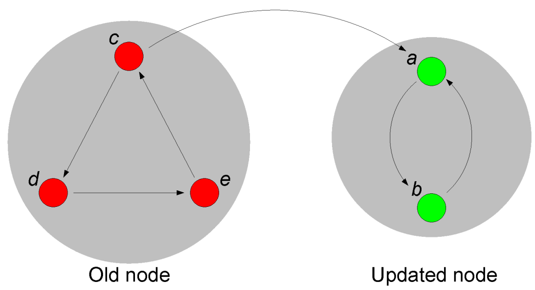

Figure 1: Example: peer-to-peer software update

Assume that for the initial conditions,

$$\|\bar{\mathbf{x}}^N(0) - \mathbf{v}(0)\| \to 0 \qquad \text{in probability}$$

(where $\|.\| = \|.\|_\infty$) and also that both $\bar{\mathbf{x}}^N(0)$ and $\mathbf{v}(0)$ are concentrated on $\mathcal{S}_I$.

We assume that the solution of (2) uniquely exists.

Rewriting the equations in integral form gives

$$\begin{aligned} v_i(t) =& v_i(0) - \sum_{j\in\mathcal{S}} \int_0^t v_i(u) r_{ij}(\mathbf{v}(u))\mathrm{d}u + \sum_{j\in\mathcal{S}} \int_0^t v_j(u) r_{ji}(\mathbf{v}(u))\mathrm{d}u \\ &+ \sum_{h\in\mathcal{S}_I} \sum_{j\in\mathcal{S}_A} \int_{z=0}^{t-T_j} p_{ji} \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\mathbf{v}(\tau))\mathrm{d}\tau\right) v_h(z) r_{hj}(\mathbf{v}(z))\mathrm{d}z \\ &- \sum_{h\in\mathcal{S}_I} \sum_{j\in\mathcal{S}} \int_{z=0}^{t-T_i} p_{ij} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right) v_h(z) r_{hi}(\mathbf{v}(z))\mathrm{d}z\ . \end{aligned} \tag{3}$$

The main result of the paper, stated and proved in Section 3, is that the mean-field equations are indeed the mean-field limit of the corresponding PGSMP. Before that, we include an example.

2.5. Example: peer-to-peer software update model

We recall an example of a peer-to-peer software update model from [16] and demonstrate through this example the more general possibilities opened up by allowing the race between deterministic and Markovian transitions.

Each component corresponds to a computer in a network. A certain software is updated throughout the network in a peer-to-peer manner; components try to obtain the update from other components which are already udpated. Also, each computer may be turned on or off.

Our model is the following. Components that are updated may be turned on or off. When a component that is not yet updated ("old") is turned on, it is searching for an update for some time; the probability (rate) of finding an update is proportional to the number of updated components turned on. If it finds an update, it applies it, otherwise, it gives up after some time, but stays turned on. After some more time, it is turned off. Accordingly, $\mathcal{S} = \{a, b, c, d, e\}$ where the local states are the following:

(a) updated component turned on
(b) updated component turned off
(c) old component searching for an update
(d) old component turned on, but not searching for an update
(e) old component turned off

Altogether, the structure of the local states and transitions is depicted in Figure 1.

In a purely Markovian version of this model, all the transitions in the above model are Markovian, that is, the waiting times are exponential. We intend to make the above model more general by replacing some of the Markovian transitions with deterministic transitions. The question is: which transition may be replaced?

In the version of this example in [16], the transitions (e→c) and (b→a) were replaced by generally distributed delays. These correspond to non-interruptible delays since there are no other transitions originating from either state e or state b.

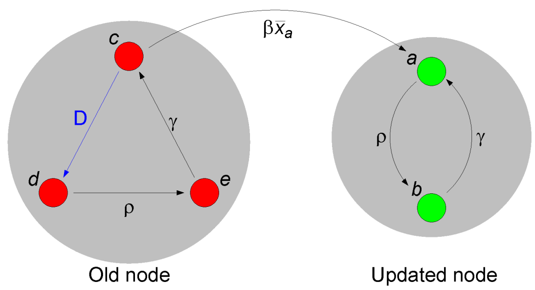

Figure 2: Peer-to-peer software update with transition types and rates

This is not the case for state c though, where transitions (c→a) and (c→d) are competing. Replacing either one by a deterministic transition results in a racing situation which is not covered by [16] (or any other result in the literature that we are aware of).

So, to finish the proper setup of the example, we set the following Markovian transition rates:

- a component in state (a) transitions to state (b) with rate $r_{ab} = \rho$ (rate of turning off);
- a component in state (b) transitions to state (a) with rate $r_{ba} = \gamma$ (rate of turning on);
- a component in state (c) transitions to state (a) with rate $r_{ca} = \beta \bar{x}_a$, where $\bar{x}_a$ denotes the ratio of components in state (a) (rate of update);
- a component in state (d) transitions to state (e) with rate $r_{de} = \rho$ (rate of turning off);
- a component in state (e) transitions to state (c) with rate $r_{ec} = \gamma$ (rate of turning on).

Additionally,

- a component in state (c) transitions to state (d) after a deterministic delay D (unless it "loses the race" and makes a Markovian transition to (a) first).

The transition types and rates are denoted in Figure 2.

For the peer-to-peer software update model, the Poisson representation is the following:

$$
\begin{aligned}
\bar{x}_a(t) =& \bar{x}_a(0) - \frac{1}{N} P_{ab}\left(N\int_0^t \rho\bar{x}_a(u)\mathrm{d}u\right) + \frac{1}{N}P_{ca}\left(N\int_0^t \beta\bar{x}_c(u)\bar{x}_a(u)\mathrm{d}u\right) + \frac{1}{N}P_{ba}\left(N\int_0^t \gamma\bar{x}_b(u)\mathrm{d}u\right)\\
\bar{x}_b(t) =& \bar{x}_b(0) - \frac{1}{N} P_{ba}\left(N\int_0^t \gamma\bar{x}_b(u)\mathrm{d}u\right) + \frac{1}{N}P_{ab}\left(N\int_0^t \rho\bar{x}_a(u)\mathrm{d}u\right)\\
\bar{x}_c(t) =& \bar{x}_c(0) - \frac{1}{N} P_{ca}\left(N\int_0^t \beta\bar{x}_c(u)\bar{x}_a(u)\mathrm{d}u\right) + \frac{1}{N}P_{ec}\left(N\int_0^t \gamma\bar{x}_e(u)\mathrm{d}u\right)\\
&- \int_{z=0}^{t-D} B_{P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)}(z)\frac{1}{N}\mathrm{d}\left(P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)\right) \qquad (4)\\
\bar{x}_d(t) =& \bar{x}_d(0) - \frac{1}{N} P_{de}\left(N\int_0^t \rho\bar{x}_d(u)\mathrm{d}u\right)\\
&+ \int_{z=0}^{t-D} B_{P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)}(z)\frac{1}{N}\mathrm{d}\left(P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)\right)\\
\bar{x}_e(t) =& \bar{x}_e(0) - \frac{1}{N} P_{ec}\left(N\int_0^t \gamma\bar{x}_e(u)\mathrm{d}u\right) + \frac{1}{N}P_{de}\left(N\int_0^t \rho\bar{x}_d(u)\mathrm{d}u\right).
\end{aligned}
$$

The corresponding set of DDEs is the following:

Figure 3: Peer-to-peer model for $N = 200$

Figure 4: Peer-to-peer model for $N = 1000$

Figure 5: Peer-to-peer model for $N = 5000$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}v_a(t) =& -\rho v_a(t) + \beta v_c(t)v_a(t) + \gamma v_b(t) \\
\frac{\mathrm{d}}{\mathrm{d}t}v_b(t) =& -\gamma v_b(t) + \rho v_a(t) \\
\frac{\mathrm{d}}{\mathrm{d}t}v_c(t) =& -\beta v_a(t)v_c(t) - \exp\left(-\int_{\tau=0}^{D}\beta v_a(t-D+\tau)\mathrm{d}\tau\right)\gamma v_e(t-D) + \gamma v_e(t) \\
\frac{\mathrm{d}}{\mathrm{d}t}v_d(t) =& -\rho v_d(t) + \exp\left(-\int_{\tau=0}^{D}\beta v_a(t-D+\tau)\mathrm{d}\tau\right)\gamma v_e(t-D) \\
\frac{\mathrm{d}}{\mathrm{d}t}v_e(t) =& -\gamma v_e(t) + \rho v_d(t).
\end{aligned}
\tag{5}
$$

The novel term that appears in (4) is

$$
\int_{z=0}^{t-D} B_{P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)}(z)\frac{1}{N}\mathrm{d}\left(P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)\right). \tag{6}
$$

(6) counts the total number of $c \to d$ deterministic transitions up to time t. Only active clocks started by time $t-D$ may fire before t. An active clock starts whenever an $e \to c$ Markovian transition occurs, which is at times z when the Poisson process $P_{ec}\left(N\int_0^z \gamma\bar{x}_e(u)\mathrm{d}u\right)$ jumps. For example, if the Poisson process jumps at $w_1, w_2, w_3, \dots, w_K$, then the integral is equal to $\frac{1}{N}\sum_{k=1}^{K} B_k^{i,N}(w_k)$. For the total number of $c \to d$ deterministic transitions, we only need to count the ones that survived (were not interrupted).

(5) is the differential form of the mean-field equation for the example. The novel term in (5) is

$$
\exp\left(-\int_{\tau=0}^{D}\beta v_a(t-D+\tau)\mathrm{d}\tau\right)\gamma v_e(t-D). \tag{7}
$$

The integral form of (7) can be written as

$$
\int_{z=0}^{t}\exp\left(-\int_{\tau=0}^{D}\beta v_a(z-D+\tau)\mathrm{d}\tau\right)\gamma v_e(z-D)\mathrm{d}z. \tag{8}
$$

In the mean-field limit, (6) (or, equivalently, the sum $\frac{1}{N}\sum_{k=1}^{K} B_k^{i,N}(w_k)$) turns into (8). (7) represents the occurrence of deterministic transition at time t. For that, a transition at time $t-D$ with $\gamma v_e(t-D)$ is required, and this deterministic transition should not be interrupted by a competing Markovian transition whose probability is $\exp\left(-\int_{\tau=0}^{D}\beta v_a(t-D+\tau)\mathrm{d}\tau\right)$.

2.5.1. Numerical simulation

In order to demonstrate how the Poisson representation turns into a simulation, a step-by-step simulation of the peer-to-peer software update model for a small population size is carried out in Appendix A.

We also carried out larger simulations to compare $\bar{\mathbf{x}}^N$ with the solution $\mathbf{v}$ of the system of equations (5).

Figures 3-5 display simulation results for $N = 200, 1000$ and 5000 along with the (numerical) solution of the system of equations (5). The graphs a–e show the ratio of components in each class as time progresses; the smooth lines are the solution of (5) and the thick jagged lines are the results of the simulation. The parameters used are $\gamma = 0.2, \beta = 10, \rho = 0.1, D = 1$. The initial condition for both $\bar{\mathbf{x}}^N$ and $\mathbf{v}^N$ is $\mathbf{v}(0) = \bar{\mathbf{x}}^N(0) = (0.1, 0, 0, 0.9, 0)$, so 1/10 of the components start from state (a) (updated, on) and 9/10 of the components start from state (d) (old, not searching for updates). The time horizon is $T = 30$. Wolfram Mathematica code for the simulations can be accessed at [1].

Note that as N increases, the results of the simulation converge to the mean-field limit. That said, in the present paper we do not pursue an explicit bound on the speed of convergence.

3. Main result

This section is devoted to the main result and its proof.

Theorem 1. *Under the assumptions of Section 2.1-2.4, notably:*

- *the Markovian transition rates r_{ij} are Lipschitz-continuous,*
- *there is at most one deterministic clock in each active state,*
- *from an active state i, the deterministic transition takes the component to the same inactive state deterministically,*
- *the initial conditions $\bar{\mathbf{x}}^N(0)$ and $\mathbf{v}(0)$ are concentrated on $\mathcal{S}_I$,*
- *for the initial conditions, $\|\bar{\mathbf{x}}^N(0)-\mathbf{v}(0)\| \to 0$ in probability, and*
- *the solution of (2) uniquely exists.*

We have, for any $T>0$:

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t)-\mathbf{v}(t)\| = 0$$

in probability.

Sketch of the proof.

The formulas (1) and (3) are very similar in form and spirit. Both define the corresponding process from its past behaviour. They both contain terms corresponding to the same types of transition, (1) for the Poisson representation of a random population model with a finite population, while (3) for the deterministic mean-field limit. The fact that the two processes have a similar evolution formula makes it natural to apply Grönwall's lemma to prove that the two processes are indeed close over a finite interval.

We are going to introduce two auxiliary processes $\mathbf{y}^N$ and $\mathbf{w}^N$. The idea behind these processes is that they are obtained by gradually "averageing out" the randomness in the process $\bar{\mathbf{x}}^N$. Replacing any process (e.g. $\bar{\mathbf{x}}^N$) with a version of it with some of the randomness averaged out (e.g. $\mathbf{y}^N$) leads naturally to martingale concentration results (for $\bar{\mathbf{x}}^N-\mathbf{y}^N$). We provide more specific explanation after the definition of $\mathbf{y}^N$ and $\mathbf{w}^N$ (formulas (9) and (10)).

On to the actual proof now.

Define the auxiliary processes

$$\begin{aligned}
y_i^N(t) :=\ & \bar{x}_i^N(0) - \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ij}\left(N\int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) + \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ji}\left(N\int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) \\
&+ \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}_A} \int_{z=0}^{t-T_j} p_{ji} \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right) \\
&- \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}} \int_{z=0}^{t-T_i} p_{ij} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\ , \qquad (9)
\end{aligned}$$

and

$$\begin{aligned}
w_i^N(t) :=\ & \bar{x}_i^N(0) - \sum_{j\in\mathcal{S}} \int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u + \sum_{j\in\mathcal{S}} \int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u \\
&+ \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}_A} \int_{z=0}^{t-T_j} p_{ji} \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hj}(\bar{\mathbf{x}}^N(z))\mathrm{d}z \\
&- \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}} \int_{z=0}^{t-T_i} p_{ij} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z))\mathrm{d}z. \qquad (10)
\end{aligned}$$

The definition of $\mathbf{y}^N$ is obtained from the definition of $\bar{\mathbf{x}}^N$ in (1) by replacing the indicator term $B^{j,N}_{\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)$ inside the formula (1) corresponding to the interruption by its conditional expectation along the trajectory of $\bar{\mathbf{x}}^N$.

The definition of $\mathbf{w}^N$ is obtained from the definition of $\mathbf{y}^N$ by replacing the Poisson process terms corresponding to the Markovian transitions by their conditional expectation along the trajectory of $\bar{\mathbf{x}}^N$ (which correspond to the actual transition rates).

The above interpretations provide an intuitive explanation as to why $\bar{\mathbf{x}}^N - \mathbf{y}^N$ and $\mathbf{y}^N - \mathbf{w}^N$ are martingales; nevertheless, the calculations need to be carried out properly.

The rest of the proof of Theorem 1 is structured as follows: we state Lemmas 2, 3 and 4 that respectively show that $\bar{\mathbf{x}}^N - \mathbf{y}^N$, $\mathbf{y}^N - \mathbf{w}^N$ and $\mathbf{w}^N - \mathbf{v}$ are small. We finish the proof of Theorem 1 using the lemmas, then the lemmas are proved in separate subsections.

Lemma 2.

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\| = 0$$

almost surely (and in probability).

Lemma 3.

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left\|\mathbf{y}^N(t) - \mathbf{w}^N(t)\right\| = 0$$

almost surely (and in probability).

Lemma 4.

$$\|\mathbf{w}^N(t) - \mathbf{v}(t)\| \le \|\bar{\mathbf{x}}^N(0) - \mathbf{v}(0)| + C\int_{u=0}^{t} \|\bar{\mathbf{x}}^N(u) - \mathbf{v}(u)\|\,\mathrm{d}u$$

where C is a constant independent of N.

Using Lemma 4 we get that

$$\begin{aligned}&\|\bar{\mathbf{x}}^N(t) - \mathbf{v}(t)\| \le \|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\| + \|\mathbf{y}^N(t) - \mathbf{w}^N(t)\| + \|\mathbf{w}^N(t) - \mathbf{v}(t)\| \le \\ &D_N + C\int_{u=0}^{t} \|\bar{\mathbf{x}}^N(u) - \mathbf{v}(u)\|\,\mathrm{d}u\ ,\end{aligned}$$

where

$$D_N = \sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\| + \sup_{t\in[0,T]} \left\|\mathbf{y}^N(t) - \mathbf{w}^N(t)\right\| + \|\bar{\mathbf{x}}^N(0) - \mathbf{v}(0)|.$$

From Lemmas 2 and 3 (and the assumption on the initial conditions) we get that $D_N \to 0$ in probability as $N \to \infty$.

An application of Grönwall's lemma ([12], Appendix 5) yields that

$$\sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t) - \mathbf{v}(t)\| \le D_N \exp(CT),$$

which thus goes to 0 in probability as $N \to \infty$, proving Theorem 1.

3.1. Proof of Lemma 2

Sketch of the proof: $\bar{\mathbf{x}}^N - \mathbf{y}^N$ is estimated by a martingale concentration argument in two steps. In Subsection 3.1.1, the proof goes for a fixed t first; the difference is extended to a continuous time process which is a sum of a random number of martingales of 0 expectation. Then it is rewritten as a discrete-time process, sampled at random times. Azuma's inequality gives an exponential bound on the discrete-time martingales, and the random number of terms and random time of sampling weaken the estimate only by a linear factor each. In Subsection 3.1.2, the whole argument is upgraded from a fixed t to $\sup_t(.)$ by checking the value *at* the points of a partition and *between* the points separately. The estimate at the points introduces yet another linear factor, while between the points we need routine large deviation estimates. Then Borel–Cantelli lemma is applied.

The proof is rather long and technical; we highlight Lemma 5 and, more specifically, (B.2) as a central point in the proof, setting up the connection between the interrupts in the population model and the function $q(.)$, that is, the rate of interrupt for the mean-field model.

Since

$$\begin{aligned}
&\mathbf{x}_i^N(t)-y_i^N(t)= \\
&\sum_{j\in\mathcal{S}_A}\int_{z=0}^{t-T_j} p_{ji}\left[B^{j,N}_{\sum_{h\in\mathcal{S}_I}P_{hj}\left(N\int_0^z \bar{x}_h^N(u)r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)-\right. \\
&\left.\exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right]\times\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I}P_{hj}\left(N\int_0^z \bar{x}_h^N(u)r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right) \\
&-\sum_{j\in\mathcal{S}}\int_{z=0}^{t-T_i} p_{ij}\left[B^{i,N}_{\sum_{h\in\mathcal{S}_I}P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)-\right. \\
&\left.\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right]\times\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I}P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right),
\end{aligned}\tag{11}$$

to obtain Lemma 2, it is enough to prove

$$\begin{aligned}
&\lim_{N\to\infty}\sup_{t\in[0,T]}\left|\int_{z=0}^{t-T_i}\left[B^{i,N}_{\sum_{h\in\mathcal{S}_I}P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)-\right.\right. \\
&\left.\left.\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right]\times\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I}P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right|=0
\end{aligned}\tag{12}$$

almost surely for each i.

3.1.1. Estimate of $\|\bar{\mathbf{x}}^N(t)-\mathbf{y}^N(t)\|$ for fixed t

We rewrite (12) as

$$\begin{aligned}
&\int_{z=0}^{t-T_i}\left[B^{i,N}_{\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z)-\right. \\
&\left.\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right]\times\frac{1}{N}\mathrm{d}\left(\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)= \\
&\frac{1}{N}\sum_{k=1}^K B_k^{i,N}(w_k)-\exp\left(-\int_{\tau=w_k}^{w_k+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right),
\end{aligned}\tag{13}$$

where $w_k=w_{k,i}$ (to simplify notation we neglect explicit reference to state i) denotes the starting time of clock number k from deterministic clocks of type i, and $K=K_i=\max\{k: w_k<t-T_i\}$ denotes the total number of such clocks up to time t. $B_k^{i,N}(w_k)$ denotes the indicator variable that clock number k of type i is not interrupted. K and w_k are random. First we will show that a single term in the above sum is a martingale; then we will deal with the fact that K is random.

Denote

$$Z_k=Z_{k,i}:=B_k^{i,N}(w_k)-\exp\left(-\int_{\tau=w_k}^{w_k+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right),$$

and

$$M_u^k:=\begin{cases}0, & \text{if } u<w_k, \\ \mathbf{E}(Z_k|\mathcal{F}_u), & \text{if } u\geq w_k,\end{cases}\tag{14}$$

where the filtration $\{\mathcal{F}_u\}$ is defined as follows: $\mathcal{F}_u$ contains all the information known up to time u, notably the following:

- the type and time of all Markovian transitions happened before time u;
- all deterministic delays started before time u;
- for each deterministic delay, whether it was interrupted or not up to time u.

Technically, in an actual simulation it is enough to keep track of the following information:

- the global state of the system at time u;
- the list of active deterministic delays at time u (that is, deterministic delays that are set for a transition after time u which have not yet been interrupted).

A natural way to define martingales is as the conditional expectation of a random variable with respect to a filtration. (14) defines M_u^k almost in this manner (as $M_u^k = \mathbf{E}(Z_k|\mathcal{F}_u)$), but not quite; in order to prove it is indeed a martingale with respect to $\{\mathcal{F}_u\}$, we need the following lemma.

Lemma 5.

$$\mathbf{E}(Z_k|\mathcal{F}_{w_k}) = 0.$$

The proof of Lemma 5 can be found in Appendix B.

It follows from Lemma 5 that for any fixed s, $\sum_{k=1}^{s} M_u^k$ is also a continuous-time martingale with 0 expectation. To apply Azuma's inequality, we set it up as a discrete-time martingale.

Let us consider the following three types of points in time (between 0 and T):

- whenever a Markovian jump occurs,
- all points in time when a deterministic clock is set to go off (regardless of whether the clock actually goes off or is interrupted before that time),
- all points of the form $\frac{a}{N}$, $a = 0, 1, \ldots, \lfloor NT \rfloor$.

Let z_m denote the m-th such point in increasing order. These points are an increasing sequence of *stopping times* with respect to the filtration $\mathcal{F}_u$, so it makes sense to define the discrete-time filtration

$$\mathcal{F}_m := \mathcal{F}_{z_m}.$$

We argue that for any fixed s, the corresponding discrete-time process

$$N_m^s := \sum_{k=1}^{s} M_{z_m}^k \ , \qquad m = 1, 2, \ldots$$

is indeed a discrete-time martingale due to the Optional Sampling Theorem (see for example Theorem 10.10 in [23]); the stopping times are all bounded by T, and N_m^s has bounded increments: it may increase only by 1 (since any Z_k may increase only by 1 and the process is stopped whenever this happens) and decrease only by sQ/N (since Z_k may decrease only via the exponential term, which, in a time interval of length at most $1/N$, may decrease at most Q/N due to being Lipschitz-continuous). For $s \leq 2NRT$,

$$sQ/N \leq 2RTQ.$$

For simplicity, we assume $2RTQ \geq 1$; otherwise, just increase Q (or R) so that it holds. In this case, the martingale N_m^s has increments bounded by $2RTQ$.

Now (13) is equal to

$$\sum_{k=1}^{K} Z_k = \sum_{k=1}^{K} M_{z_P}^k = N_P^K,$$

where K is the total number of deterministic clocks of type i which are initialized to go off before time t and z_P is the time when clock number K is set to go off (note that z_P is inherently included among the points $\{z_k\}$). Note that K and P are both random, so we can not apply Azuma's inequality directly at this point; instead, we first prove a concentration estimate (Azuma) for $\sum_{k=1}^{s} M_{z_p}^k$ for fixed s and p.

Let $\varepsilon > 0$ be fixed, and let $k \leq 2NRT, p \leq N(3R+1)T$. Azuma's inequality ([9], page 94, Theorem 5.2) guarantees that

$$\mathbf{P}\left(\left|N_p^k\right| > \varepsilon N\right) \leq 2\exp\left(-\frac{(\varepsilon N)^2}{(kQ/N)^2 p}\right) \leq 2\exp(-C_1 N) \ ,$$

where

$$C_1 = \frac{\varepsilon^2}{4R^2T^3Q^2(3R+1)}.$$

Then we upgrade from a fixed p and k to random P and K. The main point here is that K and P are not likely to be larger than linear in N.

$$\begin{aligned}
&\mathbf{P}\left(\left|N_P^K\right| > \varepsilon N\right) \leq \\
&\sum_{k=1}^{2NRT} \sum_{p=NT}^{N(3R+1)T} \mathbf{P}\left(\left|N_p^k\right| > \varepsilon N\right) + \mathbf{P}(K > 2NRT) + \mathbf{P}(P > N(3R+1)T) \leq \\
&(2NRT)(3NRT)2\exp(-C_1N) + \mathbf{P}(K > 2NRT) + \mathbf{P}(P > N(3R+1)T).
\end{aligned}$$

K is stochastically dominated by Poisson(NRT); that is,

$$\mathbf{P}(K > x) \leq \mathbf{P}(X > x) \quad \forall x \in \mathbb{R}$$

where X is Poisson(NRT) distributed. The stochastic dominance holds since K is the total number of Markovian transitions and the rate of the Markovian transitions may never go above NR (and the time horizon is T). P counts three types of points, two of which are similarly stochastically dominated by Poisson(NRT) and the third is deterministic, NT. Cramér's large deviation theorem (see e.g. Theorem II.4.1 in [11]) guarantees that

$$\mathbf{P}(K > 2NRT) \leq \exp(-C_2N), \qquad \mathbf{P}(P > N(3R+1)T) \leq \exp(-C_3N),$$

where

$$C_2 = (2\ln 2 - 1)RT, \qquad C_3 = (3\ln(3/2) - 1)RT.$$

(The Cramér rate function of the Poisson-distribution with parameter λ is $I(x) = x\ln(x/\lambda) - x + \lambda$.)

In any case, for a fixed ε,

$$\mathbf{P}\left(\frac{1}{N}\left|\sum_{k=1}^{K} \mathbf{1}\left\{B_k^{i,N}(w_k)\right\} - \exp\left(-\int_{\tau=w_k}^{w_k+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right| > \varepsilon\right)$$

is exponentially small in N and thus

$$\mathbf{P}\left(\left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| > \varepsilon\right) \leq \exp(-C_4N) \tag{15}$$

for some constant C_4 independent of t.

3.1.2. Estimate of $\sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\|$

Next we upgrade the estimate (15) from a fixed t to $\sup_{t\leq T}(.)$. Let

$$t_l := \frac{lT}{N},\ l = 0, 1, \ldots N,$$

then

$$\mathbf{P}\left(\max_{0\leq l\leq N} \frac{1}{N}\left|\bar{\mathbf{x}}^N(t_l) - \mathbf{y}^N(t_l)\right| > \varepsilon\right) \leq N\exp(-C_4N).$$

Let $t \in (t_l, t_{l+1})$. From (11),

$$
\begin{aligned}
&|\bar{x}_i^N(t) - y_i^N(t)| \le |\bar{x}_i^N(t_l) - y_i^N(t_l)| + \\
&\sum_{j\in\mathcal{S}_A} \int_{z=t_l-T_j}^{t-T_j} p_{ji} \Bigg| \mathbf{1}\left\{ B^{i,N}_{\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z) \right\} - \\
&\exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \Bigg| \times \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right) \\
&+ \sum_{j\in\mathcal{S}} \int_{z=t_l-T_i}^{t-T_i} p_{ij} \Bigg| \mathbf{1}\left\{ B^{i,N}_{\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z) \right\} - \\
&\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \Bigg| \times \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right).
\end{aligned}
$$

We only deal with the second of the last two terms, the other one is similar. Since

$$
\int_{z=t_l-T_i}^{t-T_i} \mathbf{1}\left\{ B^{i,N}_{\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z) \right\} \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)
$$

and

$$
\int_{z=t_l-T_i}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)
$$

are both increasing in t, we only need to check that neither of them increases by more than some ε'. Both functions inside the integral are bounded from above by $1/N$, so in fact both are dominated stochastically by

$$
\frac{1}{N}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^{t_l-T_i} r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right) - \sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^{t-T_i} r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right),
$$

and the probability that either integral is larger than ε' is less than

$$
2\mathbf{P}(X > \varepsilon' N), \qquad X \sim \text{Poisson}(RT)\ .
$$

Once again, this probability is exponentially small in N, and setting

$$
\varepsilon' = \varepsilon \cdot |\mathcal{S}|^{-2}
$$

we obtain

$$
\begin{aligned}
&\mathbf{P}\left(\sup_{t\in[0,T]} \left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| > 2\varepsilon\right) \le \\
&\mathbf{P}\left(\max_{0\le l\le N} \left|\bar{\mathbf{x}}^N(t_l) - \mathbf{y}^N(t_l)\right| > \varepsilon\right) + \\
&\quad \sum_{l=0}^{N-1}\sum_{i\in\mathcal{S}} \mathbf{P}\left(\sup_{t\in[t_l,t_{l+1}]} |(\bar{x}_i^N(t) - \bar{x}_i^N(t_l)) - (y_i^N(t) - y_i^N(t_l))| > \varepsilon\right) \le \exp(-C_5 N)
\end{aligned}
$$

for a suitable $C_5 > 0$. Since

$$
\sum_{N=1}^{\infty} \mathbf{P}\left(\sup_{t\in[0,T]} \left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| > 2\varepsilon\right) \le \sum_{N=1}^{\infty} \exp(-C_5 N) < \infty,
$$

according to the first Borel–Cantelli lemma, only finitely many of the events

$$\left\{ \sup_{t\in[0,T]} \left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| > 2\varepsilon \right\}$$

may occur almost surely, so for N large enough,

$$\sup_{t\in[0,T]} \left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| \leq 2\varepsilon.$$

This holds for any $\varepsilon > 0$, thus

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left|\bar{\mathbf{x}}^N(t) - \mathbf{y}^N(t)\right| = 0$$

almost surely.

3.2. Proof of Lemma 3

Techniques of the proof: $\mathbf{y}^N - \mathbf{w}^N$ is estimated via the Functional Strong Law of Large Numbers for the Poisson-process (which itself may be viewed as a basic martingale concentration theorem) and bounded functions.

$$\begin{aligned}
&\left|y_i^N(t) - w_i^N(t)\right| \leq \\
&\sum_{j\in\mathcal{S}} \left| \frac{1}{N} P_{ij}\left(N\int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) - \int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u \right| \\
&+\sum_{j\in\mathcal{S}} \left| \frac{1}{N} P_{ji}\left(N\int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right) - \int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u \right| \\
&+\sum_{j\in\mathcal{S}_A} p_{ji} \left| \int_{z=0}^{t-T_j} \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right. \\
&\qquad \left. - \sum_{h\in\mathcal{S}_I} \int_{z=0}^{t-T_j} \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hj}(\bar{\mathbf{x}}^N(z))\mathrm{d}z \right| \\
&+\sum_{j\in\mathcal{S}} p_{ij} \left| \int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right. \\
&\qquad \left. - \sum_{h\in\mathcal{S}_I} \int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z))\mathrm{d}z \right|
\end{aligned} \tag{16}$$

The first two terms go to 0 almost surely by the functional strong law of large numbers (FSLLN) for the Poisson process ([21], Section 3.2, [22]). That is,

$$\left| \frac{1}{N} P_{ij}\left(N\int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right) - \int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u \right| \leq \sup_{t\in[0,T]} \left| \frac{1}{N} P_{ij}\left(NRt\right) - Rt \right| \to 0$$

almost surely, and the same holds for the second term too.

The last two terms are handled in the following Lemma. The technique used is essentially the same as in Lemma 2 in [16]. We state the lemma for fixed i and j, dropping the sums in h and j and with notation corresponding to the last term in (16), but it applies to the third term in (16) just the same.

Lemma 6.

$$\begin{aligned}
&\left| \int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right. \\
&\left. - \sum_{h\in\mathcal{S}_I} \int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z \right| \to 0
\end{aligned}$$

almost surely as $N \to \infty$.

The proof of Lemma 6 is in Appendix C. Lemma 6 finishes Lemma 3.

3.3. Proof of Lemma 4

For $w_i^N(t) - v_i(t)$ we have

$$\begin{aligned}
w_i^N(t) - v_i(t) = {}& \bar{x}_i^N(0) - v_i(0) \\
& - \sum_{j\in\mathcal{S}} \int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u)) - v_i(u) r_{ij}(\mathbf{v}(u))\mathrm{d}u \\
& + \sum_{j\in\mathcal{S}} \int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u)) - v_j(u) r_{ji}(\mathbf{v}(u))\mathrm{d}u \\
& + \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}_A} p_{ji} \int_{z=0}^{t-T_j} \left[\exp\left(-\int_{\tau=z}^{z+T_j} q_j(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hj}(\bar{\mathbf{x}}^N(z)) - \right. \\
& \qquad \left. \exp\left(-\int_{\tau=z}^{z+T_j} q_j(\mathbf{v}(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hj}(\mathbf{v}(z)) \right] \mathrm{d}z \\
& - \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}} p_{ij} \int_{z=0}^{t-T_i} \left[\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z)) - \right. \\
& \qquad \left. \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\mathbf{v}(z)) \right] \mathrm{d}z\ .
\end{aligned} \tag{17}$$

The first term (the initial conditions) is left unchanged. To estimate the second and third terms in (17), we use the straightforward bound

$$|xr(x) - vr(v)| \le |x(r(x) - r(v))| + |(x-v)r(v)| \le R \cdot |x| \cdot |x-v| + |x-v| \cdot R \le 2R|x-v|,$$

which yields

$$\left| \sum_{j\in\mathcal{S}} \int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u)) - v_i(u) r_{ij}(\mathbf{v}(u))\mathrm{d}u \right| \le 2R|\mathcal{S}| \int_{u=0}^t \left\| \bar{\mathbf{x}}^N(u) - \mathbf{v}(u) \right\| \mathrm{d}u.$$

The third term is the same.

For the second of the last two terms in (17), we have

$$\begin{aligned}
& \left| \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_i^N(z) r_{ij}(\bar{\mathbf{x}}^N(z)) - \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right) \bar{x}_i^N(z) r_{ij}(\mathbf{v}(z)) \right| \le \\
& \left| \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \cdot \left(\bar{x}_i^N(z) r_{ij}(\bar{\mathbf{x}}^N(z)) - \bar{x}_h^N(i) r_{ij}(\mathbf{v}(z)) \right) \right| + \\
& \qquad \left| \left(\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) - \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right) \right) \cdot \bar{x}_i^N(z) r_{ij}(\mathbf{v}(z)) \right| \le \\
& 1 \cdot \left| r_{ij}(\bar{\mathbf{x}}^N(z)) - r_{ij}(\mathbf{v}(z)) \right| + R \left| \int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau - \int_{\tau=z}^{z+T_i} q_i(\mathbf{v}(\tau))\mathrm{d}\tau \right| \le \\
& R \left| \bar{\mathbf{x}}^N(z) - \mathbf{v}(z) \right| + RQ \int_{\tau=z}^{z+T_i} \left\| \bar{\mathbf{x}}^N(\tau) - \mathbf{v}(\tau) \right\| \mathrm{d}\tau\ ,
\end{aligned}$$

and thus

$$\left|\sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}}p_{ij}\int_{z=0}^{t-T_i}\left[\exp\left(-\int_{\tau=z}^{z+T_i}q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\bar{x}_h^N(z)r_{hi}(\bar{\mathbf{x}}^N(z))-\right.\right.$$
$$\left.\left.\exp\left(-\int_{\tau=z}^{z+T_i}q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right)\bar{x}_h^N(z)r_{hi}(\mathbf{v}(z))\right]\mathrm{d}z\right|\leq$$
$$\sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}}p_{ij}\int_{z=0}^{t-T_i}RQ\int_{\tau=z}^{z+T_i}\|\bar{\mathbf{x}}^N(\tau)-\mathbf{v}(\tau)\|\mathrm{d}\tau\mathrm{d}z\leq$$
$$\sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}}p_{ij}RQ\int_{z=0}^{t}\int_{\tau=0}^{t}\|\bar{\mathbf{x}}^N(\tau)-\mathbf{v}(\tau)\|\mathrm{d}\tau\mathrm{d}z=$$
$$\sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}}p_{ij}RQT\int_{\tau=0}^{t}\|\bar{\mathbf{x}}^N(\tau)-\mathbf{v}(\tau)\|\mathrm{d}\tau.$$

The remaining term in (17) is similar.

Summing all of these terms up we obtain

$$|w_i^N(t)-v_i(t)|\leq|\bar{x}_i^N(0)-v_i(0)|+C\int_{u=0}^{t}\|\bar{\mathbf{x}}^N(u)-\mathbf{v}(u)\|\,\mathrm{d}u$$

for $C=4R|\mathcal{S}|+2QT|\mathcal{S}|^2$, proving Lemma 4.

4. Conclusion and outlook

We have presented a model of PGSMPs, where individuals can enable both Markovian and deterministic transitions that compete with each other (as opposed to delay-only PGSMPs), and we have given a rigorous proof for the transient mean-field convergence.

4.1. Comparison with previous work

The extra level of randomness originating from the interrupts (and represented by the variables $B_k^{i,N}(z)$) is not present when the clocks are delay-only (non-interruptible); this is the main technical novelty compared to [14], and is also the reason why the proof in the current paper involves more concentration results.

We remark that the concentration tools used for the proof are somewhat reminiscent to the tools used in [16]. However, in [16], the "source of extra randomness" is different: in [16], the non-Markovian clocks are generally-distributed (hence the extra randomness), but they may not be interrupted. In the present paper, the extra randomness is due to the possibility of interruption. Accordingly, the calculations in Lemma 5 are original to this paper.

As far as we know, this is the first rigorous result concerning the race case.

4.2. Outlook

That said, there are a number of questions left open. In [16], the non-racing case was examined, albeit with the non-Markovian clocks having a general cumulative distribution function F_i instead of being deterministic. In fact, it is natural to ask the mean-field limit for the common generalization of the two cases: generally distributed clocks with interrupts. We will not go into too much detail, but, since most of the notation is already in place, we give the Poisson representation for $\mathbf{x}^N$:

$$
\begin{aligned}
&\bar{x}_i^N(t) = \bar{x}_i^N(0) - \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ij}\left(N\int_0^t \bar{x}_i^N(u) r_{ij}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right)\\
&+ \sum_{j\in\mathcal{S}} \frac{1}{N} P_{ji}\left(N\int_0^t \bar{x}_j^N(u) r_{ji}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right)\\
&+ \sum_{j\in\mathcal{S}_A} \int_{z=0}^t p_{ji} \int_{x=0}^{t-z} B^{j,N}_{\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z,x)\times\\
&\qquad \mathrm{d}\mathbf{1}\left\{T^j_{\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)} \le x\right\}\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hj}\left(N\int_0^z \bar{x}_h^N(u) r_{hj}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\\
&- \sum_{j\in\mathcal{S}} \int_{z=0}^{t-T_i} p_{ij} \int_{x=0}^{t-z} B^{i,N}_{\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)}(z,x)\times\\
&\qquad \mathrm{d}\mathbf{1}\left\{T^i_{\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)} \le x\right\}\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right),
\end{aligned}
$$

where (T_k^i), $k = 1, 2, \dots$ are independent random variables distributed according to F_i (notably the clock times). There are a few changes in the notation: the variables $B_k^{i,n}(z, x)$ now include x, the random clock time to which the clock is initialized (so the clock activates at time z and, unless interrupted, goes off at time $x + z$).

The integral form of the corresponding DDE is

$$
\begin{aligned}
v_i(t) =& v_i(0) - \sum_{j\in\mathcal{S}} \int_0^t v_i(u) r_{ij}(\mathbf{v}(u))\mathrm{d}u + \sum_{j\in\mathcal{S}} \int_0^t v_j(u) r_{ji}(\mathbf{v}(u))\mathrm{d}u\\
&+ \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}_A} p_{ji} \int_{z=0}^t \int_{x=0}^{t-z} \exp\left(-\int_{\tau=z}^{z+x} q_j(\mathbf{v}(\tau))\mathrm{d}\tau\right)\mathrm{d}F_j(x)\bar{x}_h^N(z) r_{hj}(\mathbf{v}(z))\mathrm{d}z\\
&- \sum_{h\in\mathcal{S}_I}\sum_{j\in\mathcal{S}} p_{ij} \int_{z=0}^t \int_{x=0}^{t-z} \exp\left(-\int_{\tau=z}^{z+x} q_i(\mathbf{v}(\tau))\mathrm{d}\tau\right)\mathrm{d}F_i(x)\bar{x}_h^N(z) r_{hi}(\mathbf{v}(z))\mathrm{d}z,
\end{aligned}
$$

where

$$
q_i(\mathbf{v}) = \sum_{j\in\mathcal{S}} r_{ij}(\mathbf{v}).
$$

These formulas are the general versions of (1) and (3).

Conjecture 1. *Under the assumptions of Section 2, we have, for any $T > 0$:*

$$
\lim_{N\to\infty} \sup_{t\in[0,T]} \|\bar{\mathbf{x}}^N(t) - \mathbf{v}(t)\| = 0
$$

in probability.

We believe a framework similar to the present paper should work; the auxiliary processes involved are more complicated and we have not been able to check the martingale property yet (most specifically, we have difficulty with the corresponding version of Lemma 5 in the more general setting). Nevertheless, we have no doubt about the validity of the conjecture.

A question in a different direction is second order approximation, that is, fluctuations around the mean-field limit. For the original Markov population model, the fluctuations are Gaussian, with the covariance satisfying a system of ordinary differential equations [18] [12]. For PGSMP's, it is reasonable to expect Gaussian fluctuations, with the covariance satisfying a system of delayed differential equations instead, but no results are available yet. This is also related to the speed of convergence, which was not examined in the present paper.

Acknowledgments

This research is partially supported by the OTKA K100473 and OTKA K101150 projects.

[1] Code for simulation and dde solution. `http://webspn.hit.bme.hu/~illes/race.nb`. Accessed: 2016-01-20.

[2] J. Anselmi and M. Verloop. Energy-aware capacity scaling in virtualized environments with performance guarantees. *Perf. Eval.*, 68(11):1207–1221, 2011.

[3] Roberto Barbuti, Giulio Caravagna, Andrea Maggiolo-Schettini, and Paolo Milazzo. Delay stochastic simulation of biological systems: a purely delayed approach. *Transactions on Computational Systems Biology XIII*, 6575:61–84, January 2011.

[4] A. Bátkai, Á. Havasi, R. Horváth, D. Kunszenti-Kovács, and P.L. Simon. PDE approximation of large systems of differential equations. *Operators and Matrices*, pages 147–163, 2015.

[5] Michel Benaïm and Jean-Yves Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11-12):823–838, November 2008.

[6] Luca Bortolussi and Jane Hillston. Fluid approximation of CTMC with deterministic delays. *Int. Conf. on Quantitative Evaluation of Systems*, pages 53–62, 2012.

[7] Giulio Caravagna and Jane Hillston. Bio-PEPAd: A non-Markovian extension of Bio-PEPA. *Theoretical Computer Science*, 419:26–49, 2012.

[8] Hoon Choi, Vidyadhar Kulkarni, Kishor Trivedi, and Marco Ajmone Marsan. Transient analysis of deterministic and stochastic Petri nets. In *Application and Theory of Petri Nets*, volume 691, pages 166–185. Springer Berlin / Heidelberg, 1993.

[9] Fan Chung and Linyuan Lu. Concentration inequalities and martingale inequalities: A survey. *Internet Mathematics*, 3(1):79–127, 2006.

[10] David R. Cox. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(03):433–441, July 1955.

[11] R. Ellis. *Entropy, Large Deviations, and Statistical Mechanics.* Classics in Mathematics. Springer, 2005.

[12] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, 2005.

[13] Nicolas Gast and Gaujal Bruno. A mean field model of work stealing in large-scale systems. *SIGMETRICS Perform. Eval. Rev.*, 38(1):13–24, 2010.

[14] R.A. Hayden. Mean field for performance models with deterministically-timed transitions. In *Quantitative Evaluation of Systems (QEST), 2012 Ninth International Conference on*, pages 63–73, Sept 2012.

[15] Richard A. Hayden and Jeremy T. Bradley. A fluid analysis framework for a Markovian process algebra. *Theoretical Computer Science*, 411:2260–2297, 2010.

[16] Richard A. Hayden, Illés Horváth, and Miklós Telek. Mean field for performance models with generally-distributed timed transitions. In Gethin Norman and William Sanders, editors, *Quantitative Evaluation of Systems*, volume 8657 of *Lecture Notes in Computer Science*, pages 90–105. Springer International Publishing, 2014.

[17] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

[18] Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications*, 6(3):223–240, 1978.

[19] S. Maset. Numerical solution of retarded functional differential equations as abstract Cauchy problems. *J. Comput. Appl. Math.*, 161(2):259–282, 2003.

[20] Robert Schlicht and Gerhard Winkler. A delay stochastic process with applications in molecular biology. *Journal of Mathematical Biology*, 57(5):613–48, 2008.

[21] Ward Whitt. Internet supplement to Stochastic-Process Limits, 2002.

[22] Ward Whitt. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer, 2002.

[23] D. Williams. *Probability with Martingales*. Cambridge Mathematical Textbooks, 1991.

Appendix A. Step-by-step numerical simulation

In order to demonstrate how the Poisson representation turns into a simulation, we play out the first few steps of an actual realization of the peer-to-peer software update model.

Let $N = 5$ and $\bar{\mathbf{x}}(0) = (0.2, 0, 0, 0, 0.8)$ (or, equivalently, $\mathbf{x}(0) = (1, 0, 0, 0, 4)$). The parameter values are set to $\rho = \gamma = 1, \beta = 2$ and $D = 1$. We start by listing the first few jumps of the Poisson processes involved in (4):

$$P_{ab} : 0.3, 1.7, \ldots, \quad P_{ba} : 0.2, 2.4, \ldots, \quad P_{ca} : 1.4, 4.2, \ldots, \quad P_{de} : 0.5, 2.3, \ldots, \quad P_{ec} : 0.4, 0.7, 2.5, \ldots$$

Initially, only two exponential transitions are possible: $e \to c$ and $a \to b$. For the Poisson representation (4), this means that all the Poisson terms except the $P_{ec}(.)$ and $P_{ab}(.)$ ones have 0 rate inside initially. A transition $e \to c$ would occur at the solution of

$$\int_0^t N\gamma\bar{x}_e(u)\mathrm{d}u = \int_0^t 1 \cdot 4\mathrm{d}u = 0.4$$

which gives $t = 0.1$, while a transition $a \to b$ would occur at the solution of

$$\int_0^t N\rho\bar{x}_a(u)\mathrm{d}u = \int_0^t 1 \cdot 1\mathrm{d}u = 0.3$$

which gives $t = 0.3$. The solution is smaller for $e \to c$, so the first transition in the system is of type $e \to c$ at time $t = 0.1$.

At time $t = 0.1$, the system looks like the following: the global state is $(1, 0, 1, 0, 3)$, and a deterministic clock (of type c, the only type) is set to go off at time 1.1. For the next transition, we need to consider three exponential transitions and one deterministic transition. The potential exponential transitions are $e \to c$, $a \to b$ and $c \to a$; their corresponding jump times are the solutions of

$$\begin{aligned}
\int_0^t N\gamma\bar{x}_e(u)\mathrm{d}u = 0.4 + \int_{0.1}^t 1 \cdot 3\mathrm{d}u = 0.7 &\implies t = 0.2 \\
\int_0^t N\rho\bar{x}_a(u)\mathrm{d}u = \int_0^t 1 \cdot 1\mathrm{d}u = 0.3, \quad \text{and} &\implies t = 0.3 \\
\int_0^t N\beta\bar{x}_a(u)\bar{x}_c(u)\mathrm{d}u = \int_{0.1}^t 2 \cdot 1 \cdot 1\mathrm{d}u = 1.4 &\implies t = 0.65
\end{aligned}$$

.

so the next transition is $e \to c$ again, at $t = 0.2$.

The global state $\mathbf{x}(t)$ of the system and the list of the active clocks changes according to the following:

- $(1, 0, 0, 0, 4)$ at time 0, no active clock
- $(1, 0, 1, 0, 3)$ at time 0.1, one active clock of type c set to go off at time 1.1
- $(1, 0, 2, 0, 2)$ at time 0.2, two active clocks of type c set to go off at times 1.1 and 1.2
- $(0, 1, 2, 0, 2)$ at time 0.3, two active clocks of type c set to go off at times 1.1 and 1.2
- $(1, 0, 2, 0, 2)$ at time 0.5, two active clocks of type c set to go off at times 1.1 and 1.2

At time 0.7, a transition $c \to a$ occurs. At this point, there are two active clocks of type c. Either one will be interrupted with probability 1/2; we need to choose at random. Assume that the result of the random choice is to interrupt the clock set to go off at time 1.2. Then $B_2^{c,5}(1.2) = 0$. Note that at this point, we are not yet sure if $B_1^{c,5}(1.1) = 1$ as the first active clock may still be interrupted. However, if we continue the simulation, it turns out that the next transition that occurs is the $c \to d$ deterministic transition at time 1.1. Accordingly, the next two transitions of the global state are:

- $(2, 0, 1, 0, 2)$ at time 0.7, one active clock of type c set to go off at time 1.1
- $(2, 0, 0, 1, 2)$ at time 1.1, no active clock.

At this point we know that $B_1^{c,5}(1.1) = 1$. We stop with the step-by-step simulation here.

Appendix B. Proof of Lemma 5

k and i are fixed throughout the proof. Let

$$u_l := w_k + l\frac{T_i}{n}, \qquad l = 0, 1, \ldots, n\ ,$$

and

$$q_l := q_i(\bar{\mathbf{x}}^N(u_l)), \qquad l = 0, 1, \ldots, n\ .$$

Then

$$\begin{aligned}
&\mathbf{E}\left(B_k^{i,N}(w_k) \middle| \mathcal{F}_{w_k} \right) = \mathbf{E}\left(B_k^{i,N}(w_k) \middle| \mathcal{F}_{t_0} \right) = \\
&\mathbf{E}\left(\prod_{l=1}^{n} \mathbf{1}\left\{\text{clock } k \text{ is not interrupted during the interval } (u_{l-1}, u_l]\right\} \middle| \mathcal{F}_{u_0} \right). \qquad \text{(B.1)}
\end{aligned}$$

Note that

$$\mathbf{E}\left(\mathbf{1}\left\{\text{clock } k \text{ is not interrupted during } (u_l, u_{l+1})\right\} \middle| \mathcal{F}_{u_l} \right) = 1 - \frac{T_i}{n} q_l + o\left(\frac{1}{n}\right)\ , \qquad \text{(B.2)}$$

since q_l is the exact rate of interruption at time u_l. To see the first order term, note that the number of components at time u_l in state i is $\mathbf{x}_i^N(u_l)$, while the total rate of Markovian transitions that interrupt *one of the components* in state i is $x_i^N(u_l) \sum_j r_{ij}^N(\mathbf{x}^N(u_l)) = x_i^N(u_l) \sum_j r_{ij}(\bar{\mathbf{x}}^N(u_l))$, thus the rate of risk for a single component is

$$\sum_j r_{ij}(\bar{\mathbf{x}}^N(u_l)) = q_i(\bar{\mathbf{x}}^N(u_l)) = q_l.$$

The error term in (B.2) requires some care since $\bar{\mathbf{x}}^N(u)$ and thus $q_i(\bar{\mathbf{x}}^N(u))$ are not continuous in u. First note that the total number of transitions over a time period of length $\frac{T_i}{n}$ is stochastically dominated by a Poisson-variable of parameter $2NRT_i/n$; the total rate of Markovian transitions is bounded by NR at all times, and the factor 2 is due to the deterministic transitions, which were enabled by Markovian transitions earlier. Thus the contribution of the change to $\bar{\mathbf{x}}^N$ is no more than $2RT_i/n$ and the contribution to the change in the rate $q_s(\bar{\mathbf{x}}^N)$ is no more than $2QRT_i/n$, which, over a time interval of length T_i/n remains $O\left(\frac{1}{n^2}\right) = o\left(\frac{1}{n}\right)$.

This yields

$$
\begin{aligned}
&\mathbf{E}\left(\mathbf{E}\left(\prod_{i=1}^{n}\mathbf{1}\left\{\text{clock } k \text{ is not interrupted during } (u_{i-1},u_i)\right\}\middle|\,\mathcal{F}_{u_1}\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\Bigg(\mathbf{1}\left\{\text{clock } k \text{ is not interrupted during } (u_0,u_1)\right\}\times\\
&\quad\mathbf{E}\left(\prod_{l=2}^{n}\mathbf{1}\left\{\text{clock } k \text{ is not interrupted during } (u_{l-1},u_l)\right\}\middle|\,\mathcal{F}_{u_1}\right)\middle|\,\mathcal{F}_{u_0}\Bigg)=\cdots=\\
&=\mathbf{E}\left(\prod_{l=1}^{n}\mathbf{E}\left(\mathbf{1}\left\{\text{clock } k \text{ is not interrupted during } (u_{l-1},u_l)\right\}\middle|\,\mathcal{F}_{u_{l-1}}\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left(\prod_{l=1}^{n}\left(1-\frac{T_i}{n}q_{l-1}+o\left(\frac{1}{n}\right)\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left(\prod_{l=1}^{n}\exp\left(-\frac{T_i}{n}q_{l-1}+o\left(\frac{1}{n}\right)\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left((1+o(1))\exp\left(-\frac{T_i}{n}\sum_{l=1}^{n}q_{l-1}\right)\middle|\,\mathcal{F}_{u_0}\right).
\end{aligned}
\tag{B.3}
$$

On the other hand,

$$
\begin{aligned}
&\mathbf{E}\left(\exp\left(-\int_{\tau=w_k}^{w_k+T_i}q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left(\exp\left(-\sum_{l=1}^{n}\int_{\tau=t_{l-1}}^{t_l}q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left(\exp\left(-\sum_{i=1}^{n}\frac{T_i}{n}q_{l-1}+o\left(\frac{1}{n}\right)\right)\middle|\,\mathcal{F}_{u_0}\right)=\\
&=\mathbf{E}\left((1+o(1))\exp\left(-\frac{T_i}{n}\sum_{l=1}^{n}q_{l-1}\right)\middle|\,\mathcal{F}_{u_0}\right).
\end{aligned}
$$

Letting $n\to\infty$ shows

$$
\mathbf{E}\left(B_k^{i,N}(w_k)\middle|\,\mathcal{F}_{u_0}\right)=\mathbf{E}\left(\exp\left(-\int_{\tau=w_k}^{w_k+T_i}q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\middle|\,\mathcal{F}_{u_0}\right)
$$

and finishes the proof of Lemma 5.

Appendix C. Proof of Lemma 6

We first assert that the function

$$
f(z)=f_{i,N}(z):=\exp\left(-\int_{\tau=z}^{z+T_i}q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)
$$

as a function of z has bounded total variation over the interval $[0, T - T_i]$; moreover, its total variation has a bound independent from N. Consider

$$
\begin{aligned}
&|f(z+\delta(z)) - f(z)| = \\
&\left|\exp\left(-\int_{\tau=z+\delta(z)}^{z+\delta(z)+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) - \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\right| \leq \\
&\left|\int_{\tau=z+\delta(z)}^{z+\delta(z)+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau - \int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right| \leq \\
&\left|\int_{\tau=z}^{z+\delta(z)} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right| + \left|\int_{\tau=z+T_i}^{z+T_i+\delta(z)} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right| \leq 2\delta(z)Q,
\end{aligned}
$$

meaning that the total variation over the interval $[0, T - T_i]$ is certainly no more than $2QT$. (Lipschitz-continuity of the function $\exp(-x)$ for $x \geq 0$ was used.)

For a fixed $\varepsilon > 0$, we intend to write $f(z)$ in the form

$$
f(z) = g_\varepsilon(z) + h_\varepsilon(z) \qquad z \in [0, T - T_i],
$$

where $g = g_{i,N,\varepsilon}$ is a piecewise constant function with $0 \leq g(z) \leq 1$ and $\|h\|_\infty \leq \varepsilon$. Their exact definition is as follows. Take the ε–quantiles of the variance function of $f(z)$; that is, let $z_0 = 0$ and

$$
z_l = \inf\left\{z : z > z_{l-1},\, |f(z) - f(z_{l-1})| \geq \varepsilon\right\} .
$$

Let K denote the number of quantiles up to T ($z_K = T$); due to the bounded total variation, K is certainly no more than $\lceil 2QT/\varepsilon \rceil$, independent of N.

Let g be the piecewise constant function

$$
g(z) = f(z_l) \quad \text{if} \quad z \in (z_{l-1}, z_l];
$$

the choice of z_l's guarantees that for $h(z) = f(z) - g(z)$, $|h(z)| \leq \varepsilon$. Also, $0 \leq g(z) \leq 1$.

$$
\begin{aligned}
&\left|\int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right. \\
&\left. - \sum_{h\in\mathcal{S}_I}\int_{z=0}^{t-T_i} \exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right) \bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right| = \\
&\left|\int_{z=0}^{t-T_i} (g(z) + h(z))\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u) r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right. \\
&\left. - \sum_{h\in\mathcal{S}_I}\int_{z=0}^{t-T_i} (g(z) + h(z))\bar{x}_h^N(z) r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right|.
\end{aligned}
$$

The $g(z)$ and $h(z)$ parts are estimated slightly differently.

$$
\begin{aligned}
&\left|\int_{z=0}^{t-T_i} g(z)\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)-\sum_{h\in\mathcal{S}_I}\int_{z=0}^{t-T_i} g(z)\bar{x}_h^N(z)r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right|\leq\\
&\sum_{l=1}^K g(z_l)\left|\int_{z=z_{l-1}}^{z_l}\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)-\sum_{h\in\mathcal{S}_I}\int_{z=z_{l-1}}^{z_l}\bar{x}_h^N(z)r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right|\leq\\
&\sum_{l=1}^K 1\cdot\left|\int_{z=z_{l-1}}^{z_l}\frac{1}{N}\mathrm{d}\left(\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)-\sum_{h\in\mathcal{S}_I}\int_{z=z_{l-1}}^{z_l}\bar{x}_h^N(z)r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right|\leq\\
&K\cdot\sup_{s,t\in[0,T]}\left|\frac{1}{N}\sum_{h\in\mathcal{S}_I} P_{hi}\left(N\int_s^t \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)-\sum_{h\in\mathcal{S}_I}\int_s^t \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right|\leq\\
&\lceil 2QT/\varepsilon\rceil\cdot\sum_h\sup_{s,t\in[0,RT]}\left|\frac{1}{N}\left(P_{hi}(Nt)-P_{hi}(Ns)\right)-(t-s)\right|. \qquad \text{(C.1)}
\end{aligned}
$$

The last expression is very similar to the functional strong law of large numbers for the Poisson process, and it is indeed a straightforward consequence:

$$
\sup_{s,t\in[0,T]}\left|\frac{1}{N}P_{hi}\left(N\int_s^t r_{hi}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right)-\int_s^t r_{hi}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right|\to 0
$$

almost surely since

$$
\sup_{s,t\in[0,T]}\left|\int_s^t\cdot\right|=\sup_{s,t\in[0,T]}\left|\int_0^t\cdot-\int_0^s\cdot\right|\leq 2\sup_{t\in[0,T]}\left|\int_0^t\cdot\right|.
$$

For the $h(z)$ part,

$$
\begin{aligned}
&\left|\int_{z=0}^{t-T_i} h(z)\frac{1}{N}\mathrm{d}\left(\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)-\sum_h\int_{z=0}^{t-T_i} h(z)\bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right|\leq\\
&\int_{z=0}^{t-T_{hi}}|h(z)|\frac{1}{N}\mathrm{d}\left(\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)+\sum_h\int_{z=0}^{t-T_i}|h(z)|\bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\leq\\
&\int_{z=0}^{t-T_i}\varepsilon\frac{1}{N}\mathrm{d}\left(\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)+\sum_h\int_{z=0}^{t-T_i}\varepsilon\bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\leq\\
&\varepsilon\sum_h\sup_{s\in[0,T]}\left(\left|\frac{1}{N}P_{hi}\left(\int_0^s \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right)\right|+\left|\int_0^s \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\mathrm{d}u\right|\right)\leq\\
&\varepsilon\sum_h\left(2RT+\sup_{s\in[0,RT]}\left|\frac{1}{N}P_{hi}\left(Ns\right)-s\right|\right). \qquad \text{(C.2)}
\end{aligned}
$$

Given that $\lceil 2QT/\varepsilon\rceil$ is independent of N in (C.1), we get that

$$
\begin{aligned}
&\left|\int_{z=0}^{t-T_i}\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\frac{1}{N}\mathrm{d}\left(\sum_h P_{hi}\left(N\int_0^z \bar{x}_h^N(u)r_{hi}(\bar{\mathbf{x}}^N(u))\,\mathrm{d}u\right)\right)\right.\\
&\left.-\sum_h\int_{z=0}^{t-T_i}\exp\left(-\int_{\tau=z}^{z+T_i} q_i(\bar{\mathbf{x}}^N(\tau))\mathrm{d}\tau\right)\bar{x}_h^N(z)r_{hi}(\bar{\mathbf{x}}^N(z))\,\mathrm{d}z\right|\leq\\
&\leq 2\varepsilon RT|\mathcal{S}|+J_N,
\end{aligned}
$$

where $J_N\to 0$ almost surely as $N\to\infty$ for any fixed ε. Letting $\varepsilon\to 0$ finishes the proof of Lemma 6.