Research Article Open Access

Quan-Lin Li

Nonlinear Markov processes in big networks

Abstract: Big networks express multiple classes of large-scale networks in many practical areas such as computer networks, internet of things, cloud computation, manufacturing systems, transportation networks, and healthcare systems. This paper analyzes such big networks, and applies the mean-field theory and the nonlinear Markov processes to constructing a broad class of nonlinear continuous-time block-structured Markov processes, which can be used to deal with many practical stochastic systems. Firstly, a nonlinear Markov process is derived from a large number of big networks with weak interactions, where each big network is described as a continuous-time block-structured Markov process. Secondly, some effective algorithms are given for computing the fixed points of the nonlinear Markov process by means of the UL-type RG-factorization. Finally, the Birkhoff center, the locally stable fixed points, the Lyapunov functions and the relative entropy are developed to analyze stability or metastability of the system of weakly interacting big networks, and several interesting open problems are proposed with detailed interpretation. We believe that the methodology and results given in this paper can be useful and effective in the study of big networks.

Keywords: Nonlinear Markov process; Big network; Mean-field theory; Fixed point; Metastability

MSC: 60J10, 15B21

1 Introduction

In this paper, we consider a large number of big networks with weak interactions, where each big network is described as a continuous-time block-structured Markov process, which can be applied to deal with many practical stochastic systems. As the number of big networks goes to infinite, the interactions between any two subsets of the big networks become negligible or are asymptotically independent, and the overall effect of the interactions can be replaced by an empirical measure under a mean-field setting. Based on this, the evolution of such a big network is expressed as a time-inhomogeneous block-structured Markov process, which leads to that transient performance of the big network can be discussed by means of a system of ordinary differential equations, while its stationary performance measures can be computed in terms of a fixed point, which satisfies a system of nonlinear equations.

The purpose of this paper is to develop the mean-field computational theory both for performance evaluation and for performance optimization of big networks. During the last three decades considerable attention has been paid to studying the mean-field theory of large-scale stochastic systems, which has been well documented, for example, an excellent survey paper by Sznitman [78]; interacting Markov processes by, such as, Spitzer [76], Kurtz [52–54], Dawson [23], Shiga and Tanaka [75], Dawson and Zheng [24], Duffield and Werner [29], Duffield [28], Le Boudecet at al. [55], Darling and Norris [22], Bordenave at al. [12], Benaim and Le Boudec [9, 10] and Li [57]; interacting Markov decision processes by Gast and Gaujal [38] and Gast at al. [39]; and more generally, interacting particle systems introduced by three books of Kipnis and Landim [48], Chen [21] and

Quan-Lin Li: School of Economics and Management Sciences, Yanshan University, Qinhuangdao 066004, P.R. China, E-mail: liquanlin@tsinghua.edu.cn

Liggett [62]. For limit theory of stochastic process sequences, readers may refer to three books of Ethier and Kurtz [31], Whitt [85], and Jacod and Shiryaev [45].

During the last two decades the mean-field theory has widely been applied to studying some practical networks (or systems) including queueing systems, communication networks, manufacturing systems, transportation networks and so forth. (a) For queueing systems, Baccelli et al. [4] and Kelly [47] first applied the mean-field theory to the study of queueing networks. Subsequent papers have been published on this theme, among which see Borovkov [13], Bobbio et al. [11]; Delcoigne and Fayolle [25] for polling systems; Karpelevich and Rybko [46] for symmetric closed queuing networks; Baccelli et al. [5] for varying topology networks; Hayden et al. [43] for stochastic process algebra; and Hunt and Kurtz [44], and Zachary and Ziedins [86] for large loss networks. (b) To discuss randomized load balancing, some work has been done on two different research directions: Supermarket models, and work stealing models. For the supermarket models, readers may refer to, such as, Vvedenskaya et al. [83] for operator semigroup, Mitzenmacher [68] for density-dependent jump Markov processes, Turner [80] for martingale limits. Subsequent papers have been published on this theme, important examples include Vvedenskaya and Suhov [84], Graham [41, 42], Luczak and McDiarmid [63, 64], Li et al. [58, 59] and Li [60]. As a key generalization of the supermarket models, the fast Jackson networks were investigated by Martin and Suhov [66], Martin [65], and Suhov and Vvedenskaya [77]. In contrast, the available results of the work stealing models based on the mean-field theory, Markov processes and queueing theory are still few in the literature, e.g., see Gast and Gaujal [37] and Li and Yang [61]. In addition, readers may refer to the computer and communication systems by Benaim and Le Boudec [9, 10], and Antunes et al. [2, 3]; the bike sharing systems by Fricker et al. [35] and Fricker and Gast [34]; and the transportation networks by Oseledets and Khmelev [71].

Nonlinear Markov processes play an important role in the study of big networks. Important examples include Rybko and Shlosman [74], Peng [72], Turner [81], Benaim and Le Boudec [9], Frank [32], Kolokoltsov [49], Gast and Gaujal [38], Kolokoltsov [49, 50], Kolokoltsov at al. [51], Muzychka and Vaninsky [69], Dupuis and Fischer [30], Gast at al. [39], Vaninsky et al. [82], Budhiraja et al. [18, 19], Budhiraja and Majumder [20] and Benaim [8].

Metastability is an ubiquitous and important phenomenon of the dynamical behavior of communication networks, e.g., see Marbukh [67], Gibbens at al. [40], Kelly [47] (see Page 349), Dobrushin [27], Antunes et al. [3] and Tibi [79]. For metastability in Markov processes, readers may refer to Galves at al. [36], Bovier et al. [16, 17], Olivieri and Vares [70], Freidlin and Wentzell [33], Bovier [14, 15], den Hollander [26], and Beltran and Landim [6, 7].

The main contributions of this paper are threefold. The first one is to set up a broad class of nonlinear continuous-time block-structured Markov processes when applying the mean-field theory to analysis of big networks with weak interactions. The second one is to propose some effective algorithms for computing the fixed points of the nonlinear Markov processes by means of the UL-type *RG*-factorization, and show for some big networks that there possibly exist multiple fixed points, which lead to the metastability. The third one is to use the Birkhoff center, the locally stable fixed points, the Lyapunov functions and the relative entropy to analyze either stability or metastability of the big networks, and to give several interesting open problems with detailed interpretation. Furthermore, this paper provides a new method for computing the locally stable fixed points in the study of big networks. We believe that the methodology and results given in this paper can be useful and effective in performance evaluation and performance optimization of big networks.

The remainder of this paper is organized as follows. In Section 2, we derive a class of nonlinear Markov processes through an asymptotic analysis of the weakly interacting big networks. In Section 3, we provide some effective algorithms for computing the fixed points of the dynamic system of mean-field equations. In Section 4, we discuss the Birkhoff center and the locally stable fixed points of the dynamic system of mean-field equations, and apply the Lyapunov functions and the relative entropy to study the stability or metastability of the big network. Also, we provide several interesting open problems with detailed interpretation. Some concluding remarks are given in the final section.

2 Nonlinear Markov Processes

In this section, we derive a class of nonlinear Markov processes through an asymptotic analysis for a collection of weakly interacting big networks, in which each big network evolves as a continuous-time block-structured Markov process, which can be applied to deal with many practical stochastic systems.

To discuss a system of weakly interacting big networks, we assume that any individual of the big networks evolves as a continuous-time block-structured Markov process $\mathcal{X}$ whose infinitesimal generator is given by

$$
Q = \left(\begin{array}{ccccc}
Q_{0,0} & Q_{0,1} & Q_{0,2} & Q_{0,3} & \cdots \\
Q_{1,0} & Q_{1,1} & Q_{1,2} & Q_{1,3} & \cdots \\
Q_{2,0} & Q_{2,1} & Q_{2,2} & Q_{2,3} & \cdots \\
Q_{3,0} & Q_{3,1} & Q_{3,2} & Q_{3,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{array} \right),
\tag{1}
$$

where the size of the matrix $Q_{j,j}$ is m_j for $j \geq 0$, and the sizes of other matrices can be determined accordingly. It is easy to see that the matrix $Q_{j,j}$ is also the infinitesimal generator of a continuous-time Markov process with m_j states for $j \geq 0$. We assume that the continuous-time Markov process Q is irreducible, aperiodic and positive recurrent, and its state space may be expressed as a two-dimensional structure: $\Omega = \{(k, j) : k \geq 0, 1 \leq j \leq m_k\}$. See Li [56] for more details.

From the continuous-time block-structured Markov chain $\mathcal{X}$, the system of N weakly interacting big networks is described as an $\mathcal{X}^N$-valued Markov process, and the states of the N big networks are denoted as $X^{1,N}(t)$, $X^{2,N}(t)$, ..., $X^{N,N}(t)$, respectively.

Let $\mathbf{X}^N(t) = \left(X^{1,N}(t), X^{2,N}(t), \dots, X^{N,N}(t) \right)$. Then the empirical measure of the system of N weakly interacting big networks is given by

$$
\mu^N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{X^{i,N}(t)},
\tag{2}
$$

where δ_x is the Dirac measure at x.

We denote by $\mathfrak{P}(\Omega)$ the space of probability vectors on the state space Ω, which is equipped with the usual topology of weak convergence. If $p \in \mathfrak{P}(\Omega)$, we write $p = (p_0, p_1, p_2, \dots)$, where the size of the vector p_j is m_j for $j \geq 0$. At the same time, it is clear that $\mu^N(t) \in \mathfrak{P}(\Omega)$ is a random variable for $t \geq 0$, and $\left\{ \mu^N(t) : t \geq 0 \right\}$ is a continuous-time Markov process.

For the $\mathcal{X}^N$-valued continuous-time block-structured Markov process, we define that the probability distribution of $\mathbf{X}^N(t)$ is exchangeable, if for any level permutation $(k_{i_1}, k_{i_2}, \dots, k_{i_N})$ of $(k_1, k_2, \dots, k_N)$ and any phase permutation $(j_{i_1}, j_{i_2}, \dots, j_{i_N})$ of $(j_1, j_2, \dots, j_N)$,

$$
P \left\{ X^{1,N}(t) = (k_1, j_1), X^{2,N}(t) = (k_2, j_2), \dots, X^{N,N}(t) = (k_N, j_N) \right\}
\tag{3}
$$

$$
= P \left\{ X^{i_1,N}(t) = \left(k_{i_1}, j_{i_1} \right), X^{i_2,N}(t) = \left(k_{i_2}, j_{i_2} \right), \dots, X^{i_N,N}(t) = \left(k_{i_N}, j_{i_N} \right) \right\}.
$$

In the system of N weakly interacting big networks, the effect of a tagged big network on the dynamics of the system of N weakly interacting big networks is of order $1/N$, and the jump intensity of any given big network depends on the configuration of other big networks only through the empirical measure $\mu^N(t)$. To study the system of N weakly interacting big networks, it is seen from probability one that at most one big network will jump, i.e., change state, at a given time, and the jump intensities of any given big network depend only on its own state and the state of the empirical measure at that time. In addition, the jump intensities of the N weakly interacting big networks have the same functional form. Based on this, for the $\mathcal{X}^N$-valued Markov process, if the initial probability distribution of $\mathbf{X}^N(0)$ is exchangeable, then at any time $t \geq 0$, the probability distribution of $\mathbf{X}^N(t)$ is also exchangeable.

For the system of N weakly interacting big networks, if the probability distribution of $\mathbf{X}^N(t)$ is exchangeable, then the N big networks are indistinguishable, thus we apply the mean-field theory to discussion of this system through only considering the Markov process of a tagged big network (such as, the first big network);

while analysis of the total system can be completed by the propagation of chaos (as $N \to \infty$). Based on this, the infinitesimal generator of the Markov process corresponding to the tagged big network is defined as

$$\Gamma^{(N)}\left(\mu^N(t)\right) = \begin{pmatrix} \Gamma^{(N)}_{0,0}\left(\mu^N(t)\right) & \Gamma^{(N)}_{0,1}\left(\mu^N(t)\right) & \Gamma^{(N)}_{0,2}\left(\mu^N(t)\right) & \Gamma^{(N)}_{0,3}\left(\mu^N(t)\right) & \cdots \\ \Gamma^{(N)}_{1,0}\left(\mu^N(t)\right) & \Gamma^{(N)}_{1,1}\left(\mu^N(t)\right) & \Gamma^{(N)}_{1,2}\left(\mu^N(t)\right) & \Gamma^{(N)}_{1,3}\left(\mu^N(t)\right) & \cdots \\ \Gamma^{(N)}_{2,0}\left(\mu^N(t)\right) & \Gamma^{(N)}_{2,1}\left(\mu^N(t)\right) & \Gamma^{(N)}_{2,2}\left(\mu^N(t)\right) & \Gamma^{(N)}_{2,3}\left(\mu^N(t)\right) & \cdots \\ \Gamma^{(N)}_{3,0}\left(\mu^N(t)\right) & \Gamma^{(N)}_{3,1}\left(\mu^N(t)\right) & \Gamma^{(N)}_{3,2}\left(\mu^N(t)\right) & \Gamma^{(N)}_{3,3}\left(\mu^N(t)\right) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}, \tag{4}$$

where the size of the matrix $\Gamma^{(N)}_{j,j}\left(\mu^N(t)\right)$ is m_j for $j \geq 0$, and the sizes of other matrices can be determined accordingly. Since $\mu^N(t)$ is a random variable, it is clear that $\Gamma^{(N)}\left(\mu^N(t)\right)$ is a random matrix of infinite order. On the other hand, it is seen from the law of large number that the limit of the empirical measure $\mu^N(t)$ is deterministic under suitable conditions.

Now, we analyze some convergence of the sequence $\left\{\mu^N(t) : t \geq 0\right\}$ of Markov processes for $N = 1, 2, 3, \ldots$, and our aim is to provide a basic support for our later study of various convergences involved. To this end, we consider the empirical measure: $\mu^N(t) = \frac{1}{N}\sum_{i=1}^{N}\delta_{X^{i,N}(t)}$ with samples in $\mathcal{P}(\mathbb{D}(R_+, \mathbf{N}))$, where $R_+ = [0, +\infty)$, $\mathbf{N} = \{(k, j) : k \geq 0, 1 \leq j \leq m_k\}$, and $\mathbb{D}(R_+, \mathbf{N})$ is the Skorohod space, i.e., the set of mappings $x : R_+ \to \mathbf{N}$ which are right continuous with left-hand limits (in short, càdlàg). Readers may refer to Chapter 3 of Ethier and Kurtz [31] for more details. Notice that the convergence in the Skorohod topology means the convergence in distribution (or weak convergence) for the Skorohod topology on the space of trajectories, we assume that the sequence $\left\{\mu^N(t) : t \geq 0\right\}$ of Markov processes converges in probability (or converges weakly), for the Skorohod topology, to a given probability vector $p(t)$. At the same time, for this weak convergence, we write $\mu^N(t) \Longrightarrow p(t)$ for $t \geq 0$, as $N \to \infty$.

Let $\mu^N(t) \Longrightarrow p(t)$ and $\Gamma^{(N)}\left(\mu^N(t)\right) \Longrightarrow \Gamma(p(t))$ for $t \geq 0$, as $N \to \infty$. Then $p(t)$ is a given probability vector. Furthermore, using some probability analysis, we may obtain an infinite-dimensional dynamic system of mean-field equations as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}p(t) = p(t)\,\Gamma(p(t)) \tag{5}$$

with the initial condition

$$p(0) = q. \tag{6}$$

Obviously, the dynamic system of mean-field equations, given in (5) and (6), is related to a nonlinear Markov process whose infinitesimal generator is given by

$$\Gamma(p(t)) = \begin{pmatrix} \Gamma_{0,0}(p(t)) & \Gamma_{0,1}(p(t)) & \Gamma_{0,2}(p(t)) & \Gamma_{0,3}(p(t)) & \cdots \\ \Gamma_{1,0}(p(t)) & \Gamma_{1,1}(p(t)) & \Gamma_{1,2}(p(t)) & \Gamma_{1,3}(p(t)) & \cdots \\ \Gamma_{2,0}(p(t)) & \Gamma_{2,1}(p(t)) & \Gamma_{2,2}(p(t)) & \Gamma_{2,3}(p(t)) & \cdots \\ \Gamma_{3,0}(p(t)) & \Gamma_{3,1}(p(t)) & \Gamma_{3,2}(p(t)) & \Gamma_{3,3}(p(t)) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}. \tag{7}$$

Remark 1. *To establish the infinitesimal generator $\Gamma(p(t))$ of a nonlinear Markov process, readers may also refer to some recent publications, for example, the discrete-time Markov chains by Benaim and Le Boudec [9] and Budhiraja and Majumder [20], the Markov decision processes by Gast and Gaujal [38] and Gast at al. [39], the continuous-time Markov chains by Dupuis and Fischer [30] and Budhiraja et al. [18, 19], and some practical examples include Mitzenmacher [68], Bobbio et al. [11], Li et al. [58, 59], and Li and Lui [60].*

In what follows, it is necessary to provide some useful interpretation or proofs for how to establish the dynamic system of mean-field equations (5) and (6).

(a) Existence and Uniqueness

Consider the infinite-dimensional ordinary differential equation: $\frac{d}{dt}p(t) = p(t)\Gamma(p(t))$ with $p(0) = q$. A solution in the classical sense is a (continuously) differential function $p(t)$ such that $\frac{d}{dt}p(t) = p(t)\Gamma(p(t))$ with $p(0) = q$. A classical method is the Picard approximation as follows. If $\Gamma(x)$ is (locally) Lipschitz on a set $E \subseteq \mathfrak{P}(\Omega)$, that is, there exists a positive constant C such that

$$\|\Gamma(x) - \Gamma(y)\| \le C\|x - y\|, \quad x, y \in \mathfrak{P}(\Omega),$$

and $p(0) = q$ is in the interior of E, then there exists a unique global solution to the ordinary differential equation: $\frac{d}{dt}p(t) = p(t)\Gamma(p(t))$ with $p(0) = q$, within E.

To deduce whether the $\Gamma(x)$ is (locally) Lipschitz on a set $E \subseteq \mathfrak{P}(\Omega)$, Li et al. [58] and Li and Lui [60] gave an algorithmic method through dealing with some matrices of infinite orders.

(b) The limiting processes

To discuss the limit: For $t \ge 0$, $\mu^N(t) \Longrightarrow p(t)$ as $N \to \infty$, we need to set up some suitable conditions in order to guarantee the existence of such a limit.

Let $e_{k,j}$ be the unit vector of infinite dimension in which the (k, j)th entry is one and all the others are zero. Note that the empirical measure process $\mu^{(N)} = \left\{\mu^N(t) : t \ge 0\right\}$ is a Markov process on the state space $\mathfrak{P}_N(\Omega)$ where $\mathfrak{P}_N(\Omega) = \mathfrak{P}(\Omega) \cap \left(\frac{1}{N}\Omega\right)$, the possible jumps of $\mu^{(N)}$ are of the form $(e_{l,i} - e_{k,j})/N$ for $(l, i) \ne (k, j)$, and $(k, j), (l, i) \in \Omega$. If $\mu^N(t) = x \in \mathfrak{P}_N(\Omega)$ at time $t \ge 0$, then $Nx_{k,j}$ denotes State (k, j) of the big network. Hence the transition rate of the Markov process from State (k, j) to State (l, i), corresponding to the tagged big network, is given by $Nx_{k,j}\Gamma^{(N)}_{k,j;l,i}(x)$. Based on this, the generator $\mathbf{A}^{(N)}$ of the Markov process $\mu^{(N)}$ is given by

$$\mathbf{A}^{(N)}f(x) = \sum_{(k,j)\in\Omega} \sum_{\substack{(l,i)\in\Omega \\ (l,i)\ne(k,j)}} Nx_{k,j}\Gamma^{(N)}_{k,j;l,i}(x) \left[f\left(x + \frac{1}{N}\left(e_{l,i} - e_{k,j}\right)\right) - f(x)\right],$$

where $f(x)$ is a real function on $\mathfrak{P}_N(\Omega)$, and there are two types of boundary states: That a task enters the big network corresponds to an arriving boundary state: $(k, j) = (0, j)$; while a task is completed and immediately leaves the big network corresponds to a departed boundary state: $(l, i) = (0, i)$. It is easy to see that as $N \to \infty$

$$\mathbf{A}^{(N)}f(x) \to \sum_{(k,j)\in\Omega} \sum_{\substack{(l,i)\in\Omega \\ (l,i)\ne(k,j)}} x_{k,j}\Gamma_{k,j;l,i}(x) \left[\frac{\partial}{\partial x_{l,i}}f(x) - \frac{\partial}{\partial x_{k,j}}f(x)\right] \overset{\text{def}}{=} \mathbf{A}f(x).$$

Theorem 1. *Suppose that for $(k, j), (l, i) \in \Omega$ with $(k, j) \ne (l, i)$, there exists a Lipschitz continuous function $\Gamma^{(N)}_{k,j;l,i}(p) : \mathfrak{P}(\Omega) \to [0, +\infty)$ such that $\Gamma^{(N)}_{k,j;l,i}(p) \to \Gamma_{k,j;l,i}(p)$ uniformly on $\mathfrak{P}(\Omega)$. If $\left\{\mu^{(N)}(0)\right\}$ converges in probability to $q \in \mathfrak{P}(\Omega)$, then $\left\{\mu^{(N)}(t)\right\}$ converges uniformly on compact time intervals in probability to $p(t) \in \mathfrak{P}(\Omega)$ for $t \ge 0$, where the probability vector $p(t)$ is the unique global solution to the ordinary differential equation: $\frac{d}{dt}p(t) = p(t)\Gamma(p(t))$ with $p(0) = q$.*

Proof. The proof may directly follow from Theorem 2.11 in Kurtz [52]. Here, we only give a simple outline as follows. Firstly, notice that

$$F^{(N)}(p) = \sum_{(k,j),(l,i)\in\Omega} Np_{k,j}\left(\frac{1}{N}e_{l,i} - \frac{1}{N}e_{k,j}\right)\Gamma^{(N)}_{k,j;l,i}(p)$$

and

$$F(p) = \sum_{(k,j),(l,i)\in\Omega} p_{k,j}\left(e_{l,i} - e_{k,j}\right)\Gamma_{k,j;l,i}(p),$$

where as $N \to \infty$

$$\Gamma^{(N)}_{k,j;l,i}(p) \to \Gamma_{k,j;l,i}(p),$$

and $\Gamma(x)$ is (locally) Lipschitz on a set $E \subseteq \mathfrak{P}(\Omega)$, thus for the sequence $\left\{\mu^{(N)}(t), t \ge 0\right\}$ of Markov processes, it follows from Equation (III.10.13) in Rogers and Williams [73] or Page 162 in Ethier and Kurtz [31] that

$$\mathbf{M}^{(N)}(t) = \mu^{(N)}(t) - \mu^{(N)}(0) - \int_0^t \mu^{(N)}(x)\Gamma\left(\mu^{(N)}(x)\right)dx$$

is a martingale with respect to each $N \geq 1$. Therefore, if $\left\{ \mu^{(N)}(0) \right\}$ converges weakly to $q \in \mathfrak{P}(\Omega)$ as $N \to \infty$, then $\left\{ \mu^{(N)}(t), N \geq 1 \right\}$ converges weakly in $D_{\mathcal{F}}[0, +\infty)$ endowed with the Skorohod topology to the solution $p(t)$ to the ordinary differential equation: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ with $p(0) = q$, within $\mathfrak{P}(\Omega)$. This completes the proof. □

Remark 2. *Benaim and Le Boudec [9] applied the mean-field theory to studying the discrete-time system of N weakly interacting objects, and for t = 0, 1, 2, ..., the states of this entire system are expressed as $\mathbf{Y}^{(N)}(t) = \left(X_1^{(N)}(t), X_2^{(N)}(t), \ldots, X_N^{(N)}(t); R^{(N)}(t) \right)$ where $X_k^{(N)}(t)$ is the state of the kth object for $1 \leq k \leq N$, and $R^{(N)}(t)$ is the state of the common resource of the N objects. Differently from that of Benaim and Le Boudec [9], this paper uses the mean-field theory to studying the continuous-time system of N weakly interacting objects, where the states of each object is described as a block-structured Markov process (i.e., a Markov process under a stochastic environment $J^{(N)}(t)$), and it is easy to see that the stochastic environment $J^{(N)}(t)$ may be regarded as the resource $R^{(N)}(t)$. On the other hand, we provide a simple method to describe the weak interaction of the N objects through introduction to the matrix $\Gamma^{(N)}\left(\mu^N(t) \right)$ where $\mu^N(t)$ is the empirical measure of the system of the N weakly interacting objects. Although our method is simple to deal with the weak convergence $\Gamma^{(N)}\left(\mu^N(t) \right) \Longrightarrow \Gamma(p(t))$ as $N \to \infty$, it is useful and effective in the mean-field study of many continuous-time big networks, readers may refer to Li et al. [58, 59] and Li [60] for more details.*

3 The Fixed Points

In this section, we use the UL-type RG-factorization to provide some effective algorithms for computing the fixed points of the ordinary differential equation: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ with $p(0) = q$. Further, we set up a nonlinear characteristic equation of the censoring matrix to level 0, which is satisfied by the fixed points.

A point $\pi \in \mathfrak{P}(\Omega)$ is said to be a fixed point of the ordinary differential equation: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ with $p(0) = q$, if $p(t) \to \pi$ as $t \to +\infty$, and

$$\lim_{t \to +\infty} \left[\frac{\mathrm{d}}{\mathrm{d}t} p(t) \right] = 0.$$

In this case, it is clear that

$$\pi \Gamma(\pi) = 0, \tag{8}$$

which is an infinite-dimensional system of nonlinear equations. In general, there are more difficulties and challenging due to both the infinite order of and the nonlinear structure of the matrix $\Gamma(\pi)$ when solving the fixed point equation (8) together with $\pi e = 1$, where e is a column vector of ones with a suitable size.

It is easy to check that for every $\pi \in \mathfrak{P}(\Omega)$, $\Gamma(\pi)$ is the infinitesimal generator of an irreducible continuous-time Markov process. Based on Li [56], we can develop the UL-type RG-factorization of the matrix $\Gamma(\pi)$. To that end, we partition the matrix $\Gamma(\pi)$ as

$$\Gamma(\pi) = \left(\begin{array}{cc} T(\pi) & U(\pi) \\ V(\pi) & W(\pi) \end{array} \right)$$

according to the level sets $L_{\leq n}$ and $L_{\geq n+1}$ for $n \geq 0$. Since the Markov chain $\Gamma(\pi)$ is irreducible, it is clear that the two truncated chains with infinitesimal generators $T(\pi)$ and $W(\pi)$ are all transient, and also the matrices $T(\pi)$ and $W(\pi)$ are all invertible from a different understanding that the inverse of the matrix $T(\pi)$ is ordinary, but the invertibility of the matrix $W(\pi)$ is different under an infinite-dimensional meaning. Although the matrix $W(\pi)$ of infinite size may have multiple inverses, we in general are interested in the maximal non-positive inverse $W_{\max}^{-1}(\pi)$ of $W(\pi)$, i.e., $W^{-1}(\pi) \leq W_{\max}^{-1}(\pi) \leq 0$ for any non-positive inverse $W^{-1}(\pi)$. Of course, $0 \leq [-W(\pi)]_{\min}^{-1} \leq [-W(\pi)]^{-1}$ for any nonnegative inverse $[-W(\pi)]^{-1}$ of $-W(\pi)$, that is, $[-W(\pi)]_{\min}^{-1}$ is the

minimal nonnegative inverse of $-W(\pi)$. Based on this, for $n \geq 0$ we write

$$\Gamma^{[\leq n]}(\pi) = T(\pi) + U(\pi)\left[-W(\pi)\right]_{\min}^{-1} V(\pi) = \begin{pmatrix} \phi_{0,0}^{(n)}(\pi) & \phi_{0,1}^{(n)}(\pi) & \cdots & \phi_{0,n}^{(n)}(\pi) \\ \phi_{1,0}^{(n)}(\pi) & \phi_{1,1}^{(n)}(\pi) & \cdots & \phi_{1,n}^{(n)}(\pi) \\ \vdots & \vdots & & \vdots \\ \phi_{n,0}^{(n)}(\pi) & \phi_{n,1}^{(n)}(\pi) & \cdots & \phi_{n,n}^{(n)}(\pi) \end{pmatrix},$$

where the size of the matrix $\phi_{j,j}^{(n)}(\pi)$ is m_j for $0 \leq j \leq n$, and the sizes of other matrices can be determined accordingly. It is clear from Section 7 of Chapter 2 in Li [56] that for $n \geq 0$, $0 \leq i, j \leq n$,

$$\phi_{i,j}^{(n)}(\pi) = \Gamma_{i,j}(\pi) + \sum_{k=n+1}^{\infty} \phi_{i,k}^{(k)}(\pi)\left[-\phi_{k,k}^{(k)}(\pi)\right]^{-1}\phi_{k,j}^{(k)}(\pi).$$

Let

$$\Psi_n(\pi) = \phi_{n,n}^{(n)}(\pi), \quad n \geq 0;$$

$$R_{i,j}(\pi) = \phi_{i,j}^{(j)}(\pi)\left[-\phi_{j,j}^{(k)}(\pi)\right]^{-1}, \quad 0 \leq i < j;$$

and

$$G_{i,j}(\pi) = \left[-\phi_{i,i}^{(k)}(\pi)\right]^{-1}\phi_{i,j}^{(i)}(\pi), \quad 0 \leq j < i.$$

Then the UL-type *RG*-factorization of the matrix $\Gamma(\pi)$ is given by

$$\Gamma(\pi) = [I - R_U(\pi)]\,\Psi_D(\pi)\,[I - G_L(\pi)], \tag{9}$$

where

$$R_U(\pi) = \begin{pmatrix} 0 & R_{0,1}(\pi) & R_{0,2}(\pi) & R_{0,3}(\pi) & \cdots \\ & 0 & R_{1,2}(\pi) & R_{1,3}(\pi) & \cdots \\ & & 0 & R_{2,3}(\pi) & \cdots \\ & & & 0 & \cdots \\ & & & & \ddots \end{pmatrix},$$

$$\Psi_D(\pi) = \mathrm{diag}\left(\Psi_0(\pi), \Psi_1(\pi), \Psi_2(\pi), \Psi_3(\pi), \ldots\right)$$

and

$$G_L(\pi) = \begin{pmatrix} 0 & & & & \\ G_{1,0}(\pi) & 0 & & & \\ G_{2,0}(\pi) & G_{2,1}(\pi) & 0 & & \\ G_{3,0}(\pi) & G_{3,1}(\pi) & G_{3,2}(\pi) & 0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Based on the UL-type *RG*-factorization (9), it follows from Subsection 2.7.3 in Li [56] that the fixed point π is given by

$$\begin{cases} \pi_0 = \tau x_0(\pi), \\ \pi_k = \sum_{i=0}^{k-1} \pi_i R_{i,k}(\pi), \quad k \geq 1, \end{cases} \tag{10}$$

where $x_0(\pi)$ is the stationary probability vector of the censored Markov chain $\Psi_0(\pi)$ to level 0, and the scalar τ is determined by $\sum_{k=0}^{\infty} \pi_k e = 1$ uniquely.

Using the expression (10) of the fixed point π, we set up an important relation as follows:

$$\pi = \left(\tau x_0(\pi), \pi_0 R_{0,1}(\pi), \sum_{i=0}^{1} \pi_i R_{i,k}(\pi), \sum_{i=0}^{2} \pi_i R_{i,k}(\pi), \ldots\right), \tag{11}$$

which is called a fixed point equation with *R*-measure.

In what follows we consider two special cases in order to further explain the fixed point equation (11) with R-measure.

Case one: Nonlinear Markov processes of GI/M/1 type

In this case, the infinitesimal generator $\Gamma(\pi)$ is given by

$$\Gamma(\pi) = \begin{pmatrix} B_1(\pi) & B_0(\pi) & & & \\ B_2(\pi) & A_1(\pi) & A_0(\pi) & & \\ B_3(\pi) & A_2(\pi) & A_1(\pi) & A_0(\pi) & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Let $R(\pi)$ be the minimal nonnegative solution to the nonlinear matrix equation

$$\sum_{k=0}^{\infty} R^k(\pi) A_k(\pi) = 0.$$

Then

$$\pi_k = \pi_1 R^{k-1}(\pi), \quad k \geq 1,$$

where the two vectors π_0 and π_1 satisfy the following system of nonlinear matrix equations

$$(\pi_0, \pi_1) \begin{pmatrix} B_1(\pi) & B_0(\pi) \\ \sum_{k=0}^{\infty} R^k(\pi) B_{k+2}(\pi) & \sum_{k=0}^{\infty} R^k(\pi) A_{k+1}(\pi) \end{pmatrix} = 0$$

and

$$\pi_0 e + \pi_1 [I - R(\pi)]^{-1} e = 1.$$

Thus, the fixed point equation (11) with R-measure is simplified as

$$\pi = \left(\pi_0, \pi_1, \pi_1 R(\pi), \pi_1 R^2(\pi), \ldots \right).$$

Case two: Nonlinear Markov processes of M/G/1 type

In this case, the infinitesimal generator $\Gamma(\pi)$ is given by

$$\Gamma(\pi) = \begin{pmatrix} B_1(\pi) & B_2(\pi) & B_3(\pi) & B_4(\pi) & \cdots \\ B_0(\pi) & A_1(\pi) & A_2(\pi) & A_3(\pi) & \cdots \\ & A_0(\pi) & A_1(\pi) & A_2(\pi) & \cdots \\ & & A_0(\pi) & A_1(\pi) & \cdots \\ & & & \ddots & \ddots \end{pmatrix}.$$

Let $G(\pi)$ be the minimal nonnegative solution to the nonlinear matrix equation

$$\sum_{k=0}^{\infty} A_k(\pi) G^k(\pi) = 0.$$

Then

$$\Psi_0(\pi) = B_1(\pi) + \sum_{k=2}^{\infty} B_k(\pi) G^{k-2}(\pi) G_1(\pi),$$

$$\Psi(\pi) = A_1(\pi) + \sum_{k=2}^{\infty} A_k(\pi) G^{k-1}(\pi);$$

and the R-measure

$$R_{0,j}(\pi) = \left[\sum_{k=j+1}^{\infty} B_k(\pi) G^{k-1}(\pi) \right] [-\Psi(\pi)]^{-1}, \quad j \geq 1,$$

for $i \geq 1$

$$R_j(\pi) = \left[\sum_{k=j+1}^{\infty} A_k(\pi) G^{k-1}(\pi) \right] [-\Psi(\pi)]^{-1}, \quad j \geq 1.$$

The fixed point π is given by

$$\begin{cases} \pi_0 = \tau x_0(\pi), \\ \pi_k = \pi_0 R_{0,k}(\pi) + \sum_{i=1}^{k-1} \pi_i R_{k-i}(\pi), \quad k \geq 1, \end{cases}$$

where $x_0(\pi)$ is the stationary probability vector of the censored Markov chain $\Psi_0(\pi)$ to level 0, and the scalar τ is determined by $\sum_{k=0}^{\infty} \pi_k e = 1$ uniquely. Thus, the fixed point equation (11) with R-measure is simplified as

$$\pi = \left(\tau x_0(\pi), \pi_0 R_{0,1}(\pi), \pi_0 R_{0,2}(\pi) + \pi_1 R_1(\pi), \pi_0 R_{0,3}(\pi) + \sum_{i=1}^{2} \pi_i R_{k-i}(\pi), \dots \right).$$

Now, we write the fixed point equation (11) with R-measure as a functional form: $\pi = \mathbf{F}(\mathfrak{R}(\pi))$ where $\mathfrak{R}(\pi)$ is related to the R-measure, as shown in the above two special cases. Based on this, we can provide an approximative algorithm as follows:

Algorithm I: Computation of the fixed points

Step one: Taking any initial probability vector: $\pi^{(0)} \in \mathfrak{P}(\Omega)$.

Step two: Computing the infinitesimal generator: $\Gamma\left(\pi^{(0)}\right)$; and then compute the R-measure, which gives $\pi^{(1)} = \mathbf{F}\left(\mathfrak{R}\left(\pi^{(0)}\right)\right)$.

Step three: For $N \geq 2$, compute $\pi^{(N+1)} = \mathbf{F}\left(\mathfrak{R}\left(\pi^{(N)}\right)\right)$.

Step four: For a sufficiently small $\varepsilon > 0$, if $\left\| \pi^{(N+1)} - \pi^{(N)} \right\| < \varepsilon$, then the computation is over; otherwise we go to Step three.

Note that it is possible for some practical big networks that there exist multiple fixed points because the infinitesimal generator $\Gamma(\pi)$ is more general, e.g., see such examples given in Marbukh [67], Gibbens at al. [40], Kelly [47] (see Page 349), Dobrushin [27], Antunes et al. [3] and Tibi [79]. In this case, it is a key to design a suitable initial probability vector: $\pi^{(0)} \in \mathfrak{P}(\Omega)$, for example:

(1) For any given integer $m \geq 1$, we take

$$\pi^{(0)} = \left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}, 0, \dots \right).$$

(2) For any given parameter $\rho \in (0, 1)$, we take

$$\pi^{(0)} = \left(1 - \rho, \rho(1 - \rho), \rho^2(1 - \rho), \rho^3(1 - \rho), \dots \right).$$

(3) For any given parameter $\lambda > 0$, we take

$$\pi^{(0)} = \left(e^{-\lambda}, e^{-\lambda}\frac{\lambda}{1!}, e^{-\lambda}\frac{\lambda^2}{2!}, e^{-\lambda}\frac{\lambda^3}{3!}, \dots \right).$$

(4) It is obvious that we can construct a suitable vector $\pi^{(0)}$ through some practical observation and experience, for instance, a discrete-time PH distribution of order 2 is established by means of only the first three moments of a random variable.

Now, we provide another algorithm for computing the fixed points. To do end, we set up a characteristic equation of the censoring matrix $\Psi_0(\pi)$ to level 0, while the characteristic equation is satisfied by the fixed points.

Note that for the censored Markov chain $\Psi_0(\pi)$ to level 0, we have

$$x_0(\pi) \Psi_0(\pi) = 0, \quad x_0(\pi) e = 1;$$

$$\pi_0 = \tau x_0 (\pi), \quad \pi_0 e = \tau \in (0, 1).$$

Thus it is easy to see from the irreducibility of the matrix $\Gamma (\pi)$ that the matrix $\Psi_0 (\pi)$ of size m_0 is also irreducible, hence this gives $\text{rank}(\Psi_0 (\pi)) = m_0 - 1$. For the **M** matrix $\Psi_0 (\pi)$, its eigenvalue with the maximal real part is equal to zero due to $\Psi_0 (\pi) e = 0$. Let the characteristic equation be $f_x (\pi) = \det (xI - \Psi_0 (\pi)) = 0$. Then the fixed points satisfy the characteristic equation $f_0 (\pi) = \det (-\Psi_0 (\pi)) = 0$, and hence $\det (\Psi_0 (\pi)) = 0$. Hence the fixed points satisfy the system of nonlinear equations as follows:

$$\begin{cases} \det (\Psi_0 (\pi)) = 0, \\ \text{rank} (\Psi_0 (\pi)) = m_0 - 1. \end{cases} \tag{12}$$

Note that (12) provide another algorithm for computing the fixed points as follows:

Algorithm II: Computation of the fixed points

Step one: Providing a numerical solution $\widehat{\pi}$ to the nonlinear characteristic equation: $\det (\Psi_0 (\pi)) = 0$.

Step two: Check whether $\text{rank}(\Psi_0 (\widehat{\pi})) = m_0 - 1$. If Yes, then $\widehat{\pi}$ is a fixed point. If No, then going to Step one.

4 Stability and Metastability

In this section, we first discuss the Birkhoff center and the locally stable fixed points of the dynamic system of mean-field equations: $\frac{d}{dt} p (t) = p (t) \Gamma (p (t))$ with $p (0) = q$. Then we apply the Lyapunov functions and the relative entropy to studying the stability or metastability of the big networks. Furthermore, we provide several interesting open problems with detailed interpretation.

We write

$$\mathbb{S} = \{\pi : \pi \Gamma (\pi) = 0, \ \pi e = 1\}.$$

Then it is clear that

$$\mathbb{S} = \left\{ \pi : \pi = \left(\tau x_0 (\pi), \pi_0 R_{0,1} (\pi), \sum_{i=0}^{1} \pi_i R_{i,k} (\pi), \sum_{i=0}^{2} \pi_i R_{i,k} (\pi), \dots \right), \ \pi e = 1 \right\}$$

or

$$\mathbb{S} = \{\pi : \det (\Psi_0 (\pi)) = 0, \text{rank} (\Psi_0 (\pi)) = m_0 - 1, \ \pi e = 1\}$$

with $\pi_0 = \tau x_0 (\pi)$ and $\pi_k = \sum_{i=0}^{k-1} \pi_i R_{i,k} (\pi)$ for $k \geq 1$.

Since the vector equation $\pi \Gamma (\pi) = 0$, together with $\pi e = 1$, is nonlinear, it is possible for some practical big networks that there exist multiple elements in the set $\mathbb{S}$. Such practical examples include Marbukh [67], Gibbens at al. [40], Kelly [47] (see Page 349), Dobrushin [27], Antunes et al. [3] and Tibi [79]. At the same time, an argument by analytic functions can indicate an important result: The elements of the set $\mathbb{S}$ are isolated.

To describe the isolated element structure of the set $\mathbb{S}$, we often need to use the Birkhoff center of the dynamic system of mean-field equations. Notice that the Birkhoff center is used to check whether the fixed point is unique or not. Based on this, our discussion includes the following two cases:

Case one: $N \to \infty$. In this case, we denote by $\Phi (t)$ a solution to the system of differential equations $\frac{d}{dt} p (t) = p (t) \Gamma (p (t))$ with $p (0) = q$. Thus, the Birkhoff center of the solution $\Phi (t)$ is defined as

$$\Theta = \left\{ \overline{P} \in \mathfrak{P} (\Omega) : \overline{P} = \lim_{k \to \infty} \Phi (t_k) \text{ for any scale sequence} \right.$$

$$\left. \{t_k\} \text{ with } t_l \geq 0 \text{ for } l \geq 1 \text{ and } \lim_{k \to \infty} t_k = +\infty \right\}.$$

Notice that perhaps Θ contains the limit cycles or the equilibrium points (the local minimal points, or the local maximal points, or the saddle points). Thus it is clear that $\mathbb{S} \subset \Theta$. Obviously, the limiting empirical

Markov process $\{\mathbf{Y}(t) : t \geq 0\}$ spends most of its time in the Birkhoff center Θ, where $\mathbf{Y}(t) = \lim_{N \to \infty} \mu^N(t)$ weakly.

Case two: $t \to +\infty$. In this case, we write

$$\pi^{(N)} = \lim_{t \to +\infty} \mu^N(t), \quad \text{weakly,}$$

if for each $N = 1, 2, 3, \ldots$, the system of N weakly interacting big networks is stable.

Let

$$\Xi = \left\{ \overline{\pi} \in \mathfrak{P}(\Omega) : \overline{\pi} = \lim_{k \to \infty} \pi^{(N_k)} \text{ for any positive integer sequence} \right.$$

$$\left. \{N_k\} \text{ with } 1 \leq N_1 \leq N_2 \leq N_3 \leq \cdots \text{ and } \lim_{k \to \infty} N_k = \infty \right\}.$$

It is easy to see that

$$\mathbb{S} \subset \Xi \subset \Theta.$$

At the same time, it is clear that

$$\mathbb{S} = \{\text{the local minimal points in } \Theta\} \cup \{\text{the local maximal points in } \Theta\}$$
$$\cup \{\text{the saddle points in } \Theta\}.$$

and

$$\Theta - \mathbb{S} = \{\text{the limit cycles in } \Theta\}.$$

In what follows, we discuss stability or metastability of the big networks.

To analyze the stability or metastability, a key is to determine a Lyapunov function for the dynamic system of mean-field equations: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ with $p(0) = q$. The Lyapunov function g defined on $\mathfrak{P}(\Omega)$ is constructed such that

$$y\Gamma(y) \cdot \nabla g(y) \leq 0, \quad y \in \mathfrak{P}(\Omega). \tag{13}$$

It is easy to see that if $\pi \in \mathbb{S}$, then $\pi\Gamma(\pi) \cdot \nabla g(\pi) = 0$ due to the fact that $\pi\Gamma(\pi) = 0$. On the other hand, if $\pi\Gamma(\pi) \cdot \nabla g(\pi) = 0$, then $\pi \in \mathbb{S}$.

Let $|\mathbb{S}|$ be the number of elements in the set $\mathbb{S}$. If $|\mathbb{S}| = 1$, then

$$\lim_{N \to \infty} \lim_{t \to +\infty} \mu^N(t) = \lim_{t \to +\infty} \lim_{N \to \infty} \mu^N(t) = \pi, \quad \text{weakly.}$$

If $|\mathbb{S}| \geq 2$, then the system of big networks exhibits a metastability property, that is, the evolutionary process of the tagged big network with a mean-field modification switches from one stable point to the other one after a long residence time. In the study of metastability, it is a key to estimate the expected value of such a residence time. See Bovier [15], and Olivieri and Vares [70] for more details.

An interesting issue in the study of big networks is to analyze stability or metastability of the corresponding nonlinear Markov processes. On this line, it is a key to construct a Lyapunov function or a local Lyapunov function. Note that the relative entropy function in some sense can define a globally attracting Lyapunov function.

For $p, q \in \mathfrak{P}(\Omega)$, we define the relative entropy of p with respect to q as

$$R(p||q) = \sum_{x \in \Omega} p_x \log \left(\frac{p_x}{q_x} \right).$$

(a) The linear case

Let $p(t)$ and $q(t)$ be two different solutions to the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Lambda, \quad p(0) = q,$$

where Λ is the infinitesimal generator of an irreducible continuous-time Markov process. In this case, Dupuis and Fischer [30] indicated that

$$\frac{\mathrm{d}}{\mathrm{d}t} R\left(p\left(t\right) \| q\left(t\right)\right) = -\sum_{\substack{x,y\in\Omega \\ x\neq y}} \Psi\left(\frac{p_y\left(t\right)q_x\left(t\right)}{p_x\left(t\right)q_y\left(t\right)}\right) p_x\left(t\right)\frac{q_y\left(t\right)}{q_x\left(t\right)}\Lambda_{y,x} \leq 0, \tag{14}$$

where $\Psi\left(z\right) = z\log z - z + 1$. At the same time, they indicated that $\frac{\mathrm{d}}{\mathrm{d}t}R\left(p\left(t\right)\|q\left(t\right)\right) = 0$ if and only if $p\left(t\right) = q\left(t\right)$ for $t \geq 0$; and also $\frac{\mathrm{d}}{\mathrm{d}t}R\left(p\left(t\right)\|\pi\right) = 0$ if and only if $p\left(t\right) = \pi$ for $t \geq 0$.

(b) The nonlinear case

For the ordinary differential equation: $\frac{\mathrm{d}}{\mathrm{d}t}p\left(t\right) = p\left(t\right)\Gamma\left(p\left(t\right)\right)$ with $p\left(0\right) = q$, Dupuis and Fischer [30] demonstrated that the the relative entropy relation (14) can not be applied directly. In this case, they first defined $\mathbf{P}^{(N)}\left(t\right)$ as the state probability of the system of N big networks at time $t \geq 0$, and let $\mathbf{P}^{(N)}\left(0\right) = \otimes^N q$. Based on this, they gave an approximate method to construct the Lyapunov function as follows:

$$F\left(q\right) = \lim_{N\to\infty}\lim_{T\to+\infty}\frac{1}{N}R\left(\mathbf{P}^{(N)}\left(0\right)\|\mathbf{P}^{(N)}\left(T\right)\right) = \lim_{N\to\infty}\frac{1}{N}R\left(\otimes^N q\|\otimes^N\pi\right).$$

For applying the relative entropy to construct a Lyapunov function, readers may also refer to Budhiraja et al. [18, 19] for more details.

Now, we introduce the locally stable fixed points in the set $\mathbb{S}$, and explain convergence of the sequence $\left\{\mu^{(N)}(t)\right\}$ of Markov processes.

If there exists only an element π in the set $\mathbb{S}$, then

$$\mu^{(N)}(t) \Rightarrow p\left(t\right), \text{ as } N\to\infty, \tag{15}$$

and

$$p\left(t\right) \to \pi, \text{ as } t\to+\infty. \tag{16}$$

When there exist multiple elements in the set $\mathbb{S}$, we hope to find a similar property to that in (15) and (16). To this end, it is necessary to introduce the locally stable fixed points, which are first defined in Budhiraja et al. [18]. For convenience of readers, we restate the definition of locally stable fixed points here.

We define the relative interior of $\Omega^\diamond$ of Ω as

$$\Omega^\diamond = \left\{\mathbf{g}\in\Omega : \mathbf{g} > 0\right\}.$$

Notice that the Markov chain $\Gamma\left(\pi\right)$ is irreducible, it is easy to see from the system of fixed point equations $\pi\Gamma\left(\pi\right) = 0$ and $\pi e = 1$ that any fixed point $\pi > 0$, hence we obtain $\mathbb{S}\subset\Omega^\diamond$.

Definition 1. *A fixed point $\pi\in\Omega^\diamond$ is said to be locally stable if there exists a relative open set $\mathfrak{F}$ of Ω that contains π and has the property that whenever $p\left(0\right) = q\in\mathfrak{F}$, the solution $p\left(t\right)$ to the dynamic system of mean-field equations (5) and (6) converges to π as $t\to+\infty$.*

Based on Definition 1, every element in the set $\mathbb{S}$ is a locally stable fixed point. In this case, let $\mathfrak{F}^{\langle k\rangle}$ be a relative open set of Ω for $k \geq 1$, and we write

$$\mathbb{S} = \left\{\pi^{\langle 1\rangle}, \pi^{\langle 2\rangle}, \pi^{\langle 3\rangle}, \ldots\right\}, \ \ \pi^{\langle k\rangle}\in\mathfrak{F}^{\langle k\rangle} \text{ for } k \geq 1,$$

and

$$\mathbb{Q} = \left\{q^{\langle 1\rangle}, q^{\langle 2\rangle}, q^{\langle 3\rangle}, \ldots\right\}, \ \ q^{\langle k\rangle}\in\mathfrak{F}^{\langle k\rangle} \text{ for } k \geq 1.$$

If for each $k \geq 1$,

$$\mu^{(N)}(0) \Rightarrow q^{\langle k\rangle}\in\mathfrak{F}^{\langle k\rangle}, \text{ as } N\to\infty,$$

then

$$\mu^{(N)}(t) \Rightarrow p\left(t, q^{\langle k\rangle}\right), \text{ as } N\to\infty, \tag{17}$$

and

$$p\left(t, q^{\langle k \rangle}\right) \to \pi^{\langle k \rangle}, \text{ as } t \to +\infty, \tag{18}$$

where $p\left(t, q^{\langle k \rangle}\right)$ denotes that this solution $p(t)$ depends on the initial vector $q^{\langle k \rangle} \in \mathfrak{F}^{\langle k \rangle}$, where $p\left(0, q^{\langle k \rangle}\right) = q^{\langle k \rangle}$.

From the above analysis, for each locally stable fixed point in $\mathbb{S}$, we find the similar property (17) and (18) to that in (15) and (16). Based on this, computation of the locally stable fixed points is on a unified line, but the the initial vector $q^{\langle k \rangle} \in \mathfrak{F}^{\langle k \rangle}$ with $\mu^{(N)}(0) \Rightarrow q^{\langle k \rangle}$ as $N \to \infty$ has a large impact on the limit $\pi^{\langle k \rangle}$ of the solution $p(t)$ as $t \to +\infty$, thus a basic task for computing the locally stable fixed points is to find a suitable collection of the relative open sets of Ω as follows:

$$\mathcal{F} = \left\{ \mathfrak{F}^{\langle 1 \rangle}, \mathfrak{F}^{\langle 2 \rangle}, \mathfrak{F}^{\langle 3 \rangle}, \ldots \right\},$$

which sufficiently corresponds to the set $\mathbb{S} = \left\{ \pi^{\langle 1 \rangle}, \pi^{\langle 2 \rangle}, \pi^{\langle 3 \rangle}, \ldots \right\}$.

In the remainder of this section, we provide several interesting open problems with detailed interpretation.

Open problem one: The mean drift condition.

We consider an irreducible QBD process whose infinitesimal generator is given by

$$\Gamma(p) = \begin{pmatrix} B_1(p) & B_0(p) & & \\ B_2(p) & A_1(p) & A_0(p) & \\ & A_2(p) & A_1(p) & A_0(p) \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $\Gamma(p) e = 0$, the sizes of the matrices $B_1(p)$ and $A_1(p)$ are m_0 and m, respectively, and the sizes of other matrices can be determined accordingly. We assume that for any $p \in \mathfrak{P}(\Omega)$, the Markov process: $A(p) = A_0(p) + A_1(p) + A_2(p)$, is irreducible, aperiodic and positive recurrent. Let θ_p be the stationary probability vector of the the Markov process $A(p)$. Then it is clear that for for any $p \in \mathfrak{P}(\Omega)$, the Markov process $\Gamma(p)$ is positive recurrent if and only if $\theta_p A_2(p) e > \theta_p A_0(p) e$.

It is interesting to study how the mean drift condition: $\theta_p A_2(p) e > \theta_p A_0(p) e$ for any $p \in \mathfrak{P}(\Omega)$, can influence stability or metastability of the ordinary differential equation: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ with $p(0) = q$.

Open problem two: The censoring Markov processes.

For the infinitesimal generator $\Gamma(p)$ given in (7), it is easy to give the infinitesimal generator $\Psi_0(p)$ of the censoring Markov processes to level 0. It is very interesting (but difficult) to set up some useful relations of stability or metastability between two ordinary differential equations: $\frac{\mathrm{d}}{\mathrm{d}t} p(t) = p(t) \Gamma(p(t))$ and $\frac{\mathrm{d}}{\mathrm{d}t} p_0(t) = p_0(t) \Psi_0(p(t))$, where $p(t) = (p_0(t), p_1(t), p_2(t), \ldots)$.

5 Concluding Remarks

This paper sets up a broad class of nonlinear continuous-time block-structured Markov processes by means of applying the mean-field theory to the study of big networks, and proposes some effective algorithms for computing the fixed points of the nonlinear Markov process by means of the UL-type *RG*-factorization. Furthermore, this paper considers stability or metastability of the big networks, and gives several interesting open problems with detailed interpretation. Along such a line, there are a number of interesting directions for potential future research, for example:

–   providing algorithms for computing the fixed points of big networks with multiple stable points;
–   studying the influence of the censoring Markov processes on the metastability;
–   discussing how to apply the *RG*-factorizations given in Li [56] to compute the expected residence times in the study of metastability; and
–   analyzing some big networks with a heterogeneous geographical environment, and set up their simultaneous systems of nonlinear Markov processes.

References

[1] N. Antunes, C. Fricker, P. Robert, D. Tibi, Metastability of CDMA cellular systems. In: *Proceedings of the 12th Annual International ACM Conference on Mobile Computing and Networking*, (2006), pp 206–214.

[2] N. Antunes, C. Fricker, P. Robert, D. Tibi, Analysis of Loss Networks with Routing, *The Annals of Applied Probability*, **16**(4) (2006) 2007–2026.

[3] N. Antunes, C. Fricker, P. Robert, D. Tibi, Stochastic networks with multiple stable points, *The Annals of Probability*, **36**(1) (2008) 255–278.

[4] F. Baccelli, F.I. Karpelevich, M.Y. Kelbert, A.A. Puhalskii, A.N. Rybko, Y.M. Suhov, A mean-field limit for a class of queueing networks, *Journal of Statistical Physics*, **66**(3–4) (1992) 803–825.

[5] F. Baccelli, A.N. Rybko, S. Shlosman, Queuing networks with varying topology–a mean-field approach, arXiv preprint: arXiv:1311.3898, (2013).

[6] J. Beltran, C. Landim, Tunneling and metastability of continuous time Markov chains, *Journal of Statistical Physics*, **140**(6) (2010) 1065–1114.

[7] J. Beltran, C. Landim, Tunneling and metastability of continuous time Markov chains II, the nonreversible case, *Journal of Statistical Physics*, **149**(4) (2012) 598–618.

[8] M. Benaim, On gradient like properties of population games, learning models and self reinforced processes, arXiv preprint: arXiv:1409.4091, (2014).

[9] M. Benaim, J.Y. Le Boudec, A class of mean-field interaction models for computer and communication systems, *Performance Evalution*, **65**(11–12) (2008) 823–838.

[10] M. Benaim, J.Y. Le Boudec, On mean field convergence and stationary regime, arXiv preprint: arXiv:1111.5710, (2011).

[11] A. Bobbio, M. Gribaudo, M. Telek, Analysis of large scale interacting systems by mean field method. In: *The Fifth International IEEE Conference on Quantitative Evaluation of Systems*, (2008), pp 215–224.

[12] C. Bordenave, D. McDonald, A. Proutiere, A particle system in interaction with a rapidly varying environment: Mean-field limits and applications, *Networks and Heterogeneous Media*, **5**(1) (2010) 31–62.

[13] K.A. Borovkov, Propagation of chaos for queueing networks, *Theory of Probability & Its Applications*, **42**(3) (1998) 385–394.

[14] A. Bovier, Markov processes and metastability, Lecture Notes TUB, (2003), pp 1–75.

[15] A. Bovier, Metastability: A potential theoretic approach. In: *Proceedings oh the International Congress of Mathematicians: Invited Lectures*, (2006), pp 499–518.

[16] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability in stochastic dynamics of disordered mean-field models, *Probability Theory and Related Fields*, **119**(1) (2001) 99–161.

[17] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability and low lying spectral in reversible Markov chains, *Communications in Mathematical Physics*, **228**(2) (2002) 219–255.

[18] A. Budhiraja, P. Dupuis, M. Fischer, K. Ramanan, Local stability of Kolmogorov forward equations for finite state nonlinear Markov processes, arXiv preprint: arXiv:1412.5555, (2014).

[19] A. Budhiraja, P. Dupuis, M. Fischer, K. Ramanan, Limits of relative entropies associated with weakly interacting particle systems, arXiv preprint: arXiv:1412.5553, (2014).

[20] A. Budhiraja, A.P. Majumder, Long time results for a weakly interacting particle system in discrete time, arXiv preprint: arXiv:1401.3423, (2014).

[21] M.F. Chen, *From Markov Chains to Non-equilibrium Particle Systems*. World Scientific, (2004).

[22] R.W.R. Darling, J.R. Norris, Differential equation approximations for Markov chains, *Probability surveys*, **5**(1) (2008) 37–79.

[23] D.A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior, *Journal of Statistical Physics*, **31**(1) (1983) 29–85.

[24] D.A. Dawson, X. Zheng, Law of large numbers and central limit theorem for unbounded jump mean-field models, *Advances in Applied Mathematics*, **12**(3) (1991) 293–326.

[25] F. Delcoigne, G. Fayolle, Thermodynamical limit and propagation of chaos in polling systems, *Markov Processes and Related Fields*, **5**(1) (1999) 89–124.

[26] F. den Hollander, Metastability under stochastic dynamics, *Stochastic Processes and their Applications*, **114**(1) (2004) 1–26.

[27] R.L. Dobrushin, Queuing networks - without explicit solutions and without computer simulation. In: *Conference on Applied Probability in Engineering, Computer and Communication Sciences: Keynote Lecture*, Paris, (1993).

[28] N.G. Duffield, Local mean-field Markov processes: An application to message-switching networks, *Probability Theory and Related Fields*, **93**(4) (1992) 485–505.

[29] N.G. Duffield, R.F. Werner, Local dynamics of mean-field quantum systems, *Helvetica Physica Acta*, **65**(8) (1991) 1016–1054.

[30] P. Dupuis, M. Fischer, On the construction of Lyapunov functions for nonlinear Markov processes via relative entropy. Submitting for publication, (2011).

[31] S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence*. John Wiley & Sons, (1986).

[32] T.D. Frank, Nonlinear Markov processes, *Physics Letters A*, **372**(25) (2008) 4553–4555.

[33] M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamic Systems*. Springer, (1984).

[34] C. Fricker, N. Gast, Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity, *EURO Journal on Transportation and Logistics*, Published Online: June 7, (2014), pp 1–31.

[35] C. Fricker, N. Gast, H. Mohamed. Mean field analysis for inhomogeneous bike sharing systems. In: *DMTCS Proceedings*, **1** (2012), pp 365–376.

[36] A. Galves, E. Olivieri, M.E. Vares, Metastability for a class of dynamical systems subject to small random perturbations, *The Annals of Probability*, **15**(4) (1987) 1288–1305.

[37] N. Gast, B. Gaujal, A mean field model of work stealing in large-scale systems, *ACM SIGMETRICS Performance Evaluation Review*, **38**(1) (2010) 13–24.

[38] N. Gast, B. Gaujal, A mean field approach for optimization in discrete time, *Discrete Event Dynamic Systems*, **21**(1) (2011) 63–101.

[39] N. Gast, B. Gaujal, J.Y. Le Boudec, Mean field for Markov decision processes: from discrete to continuous optimization, *IEEE Transactions on Automatic Control*, **57**(9) (2012) 2266–2280.

[40] R.J. Gibbens, P.J. Hunt, F.P. Kelly, Bistability in communication networks. In: *Disorder in Physical Systems: A Volume in Honour of John M. Hammersley*, (Eds G. Grimmett and D. Welsh), Oxford University Press, (1990), pp 113–127.

[41] C. Graham, Chaoticity on path space for a queueing network with selection of the shortest queue among several, *Journal of Applied Probability*, **37**(1) (2000) 198–201.

[42] C. Graham, Functional central limit theorems for a large network in which customers join the shortest of several queues, *Probability Theory and Related Fields*, **131**(1) (2004) 97–120.

[43] R.A. Hayden, A. Stefanek, J.T. Bradley, Fluid computation of passage-time distributions in large Markov models, *Theoretical Computer Science*, **413**(1) (2012) 106–141.

[44] P.J. Hunt, T.G. Kurtz, Large loss networks, *Stochastic Processes and their Applications*, **53**(2) (1994) 363–378.

[45] J. Jacod, A. Shiryaev, *Limit Theorems for Stochastic Processes*. Springer, (2003).

[46] F.I. Karpelevich, A.N. Rybko, Thermodynamical limit for symmetric closed queuing networks, *Translations of the American Mathematical Society*, **198**(2) (2000) 133–156.

[47] F.P. Kelly, Loss networks, *The Annals of Applied Probability*, **1**(3) (1991) 319–378.

[48] C. Kipnis, C. Landim, *Scaling Limits of Interacting Particle Systems*. Springer, (1999).

[49] V.N. Kolokoltsov, *Nonlinear Markov Processes and Kinetic Equations*. Cambridge University Press, (2010).

[50] V.N. Kolokoltsov, Nonlinear Lévy and nonlinear Feller processes: An analytic introduction, arXiv preprint: arXiv:1103.5591, (2011).

[51] V.N. Kolokoltsov, J.J. Li, W. Yang, Mean field games and nonlinear Markov processes, arXiv preprint: arXiv: 1112.3744, (2011).

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

[53] T.G. Kurtz, Limit theorems for sequences of jump Markov processes approximating ordinary differential processes, *Journal of Applied Probability*, **8**(2) (1971) 344–356.

[54] T.G. Kurtz, Strong approximation theorems for density dependent Markov chains, *Stochastic Processes and Their Applications*, **6**(3) (1978) 223–240.

[55] J.Y. Le Boudec, D. McDonald, J. Mundinger, A generic mean-field convergence result for systems of interacting objects. In: *Proc. Conf. IEEE on the Quantitative Evaluation of Systems*, (2007), pp 3–18.

[56] Q.L. Li, *Constructive Computation in Stochastic Models with Applications: The RG-Factorizations*. Springer, (2010).

[57] Q.L. Li, Tail probabilities in queueing processes, *Asia-Pacific Journal of Operational Research*, **31**(2) (2014) 1–31.

[58] Q.L. Li, G.R. Dai, J.C.S. Lui, Y. Wang, The mean-field computation in a supermarket model with server multiple vacations, *Discrete Event Dynamic Systems*, **24**(4) (2014) 473–522.

[59] Q.L. Li, Y. Du, G.R. Dai, M. Wang, On a doubly dynamically controlled supermarket model with impatient customers, *Computers & Operations Research*, **55** (2014) 76–87.

[60] Q.L. Li, J.C.S. Lui, Block-structured supermarket models, *Discrete Event Dynamic Systems*, Published Online: June 29, (2014), pp 1–36.

[61] Q.L. Li, F.F. Yang, Mean-field analysis for heterogeneous work stealing models. In: *Information Technologies and Mathematical Modelling: Queueing Theory and Applications*, Springer, (2015), pp 28–40.

[62] T.M. Liggett, *Interacting Particle Systems*. Springer, (2005).

[63] M.J. Luczak, C. McDiarmid, On the maximum queue length in the supermarket model, *The Annals of Probability*, **34**(2) (2006) 493–527.

[64] M.J. Luczak, C. McDiarmid, Asymptotic distributions and chaos for the supermarket model, *Electronic Journal of Probability*, **12**(1) (2007) 75–99.

[65] J.B. Martin, Point processes in fast Jackson networks, *The Annals of Applied Probability*, **11**(3) (2001) 650–663

[66] J.B. Martin, Y.M. Suhov, Fast Jackson networks, *The Annals of Applied Probability*, **9**(3) (1999) 854–870.

[67] V.V. Marbukh, Investigation of a fully connected channel-switching network with many nodes and alternative routes, *Automation & Remote Control*, **44**(12) (1984) 1601–1608.

[68] M.D. Mitzenmacher, The Power of Two Choices in Randomized Load Balancing, Ph.D. thesis, Department of Computer Science, University of California at Berkeley, USA, (1996).

[69] S.A. Muzychka, K.L. Vaninsky, A class of nonlinear random walks related to the Ornstein-Uhlenbeck process, *Markov Processes and Related Fields*, **17**(2) (2011) 277–304.

[70] E. Olivieri, M.E. Vares, *Large Deviations and Metastability*. Cambridge University Press, (2005).

[71] V.I. Oseledets, D.V. Khmelev, Stochastic transportation networks and stability of dynamical systems, *Theory of Probability & Its Applications*, **46**(1) (2002) 154–161.

[72] S.G. Peng, Nonlinear expectations and nonlinear Markov chains, *Chinese Annals of Mathematics*, **26**(2) (2005) 159–184.

[73] L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales, Vol. 1: Foundations*. John Wiley & Sons, (1994).

[74] A.N. Rybko, S. Shlosman, Poisson Hypothesis for information networks (a study in non-linear Markov processes), arXiv preprint: arXiv:0303010, (2003).

[75] T. Shiga, H. Tanaka, Central limit theorem for a system of Markovian particles with mean field interactions, *Probability Theory and Related Fields*, **69**(3) (1985) 439–459.

[76] F. Spitzer, Interaction of Markov processes, *Advances in Mathematics*, **5**(2) (1970) 246–290.

[77] Y.M. Suhov, N.D. Vvedenskaya, Fast Jackson networks with dynamic routing, *Problems of Information Transmission*, **38**(2) (2002) 136–153.

[78] A. Sznitman, *Topics in Propagation of Chaos*. Springer-Verlag lecture notes in mathematics 1464, (1989), pp 165–251.

[79] D. Tibi, Metastability in communication networks, arXiv preprint: arXiv:1002.0796, (2010).

[80] S.R.E. Turner, Resource Pooling in Stochastic Networks, Ph.D. Thesis, Statistical Laboratory, Christ's College, University of Cambridge, (1996).

[81] A.G. Turner, Convergence of Markov processes near saddle fixed points, *The Annals of Probability*, **35**(3) (2007) 1141–1171.

[82] K. Vaninsky, S. Myzuchka, A. Lukov, A multi-agent nonlinear Markov model of the order book, arXiv preprint: arXiv:1208.3083, (2012).

[83] N.D. Vvedenskaya, R.L. Dobrushin, F.I. Karpelevich, Queueing system with selection of the shortest of two queues: An asymptotic approach, *Problems of Information Transmission*, **32**(1) (1996) 20–34.

[84] N.D. Vvedenskaya, Y.M. Suhov, Dobrushin's mean-field limit for a queue with dynamic routing, *Markov Processes and Related Fields*, **3**(3) (1997) 493–526.

[85] W. Whitt, *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer, (2002).

[86] S. Zachary, I. Ziedins, A refinement of the Hunt-Kurtz theory of large loss networks, with an application to virtual partitioning, *The Annals of Applied Probability*, **12**(1) (2002) 1–22.