Nonlinear Markov Processes in Big Networks

Quan-Lin Li
School of Economics and Management Sciences
Yanshan University, Qinhuangdao 066004, P.R. China

August 6, 2015

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 network, 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; RG-factorization; Fixed point; Stability; Metastability; Lyapunov function; Relative entropy

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 a system of ordinary differential equations, while its stationary performance measures can be computed by 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 large-scale stochastic systems. During the last three decades considerable attention has been paid to studying the mean-field theory of big networks. Readers may refer to recent publications, among which are Kurtz [47], Dawson [22], Shiga and Tanaka [65], Sznitman [66], Dawson and Zheng [23], Duffield and Werner [27], Duffield [26], Kipnis and Landim [43], Chen [20], Liggett [54], Le Boudecet at al. [48], Darling and Norris [21], Bordenave at al. [11], Gast and Gaujal [36], Kolokoltsov at al. [46], Gast at al. [37], and Li [50].

During the last two decades the mean-field theory has been applied to studying some practical networks, such as, queueing systems, computer networks, manufacturing systems and transportation networks. Readers may refer to, for example, Baccelli et al. [3], Vvedenskaya et al. [71], Vvedenskaya and Suhov [72], Mitzenmacher [58], Turner [68], Borovkov [12], Martin and Suhov [57], Delcoigne and Fayolle [24], Graham [39, 40], Karpelevich and Rybko [42], Oseledets and Khmelev [61], Luczak and McDiarmid [53, 56], Bobbio et al. [10], Benaim and Le Boudec [8, 9], Antunes et al. [2], Gast and Gaujal [35], Hayden et al. [41], Fricker et al. [33], Baccelli et al. [4], Li et al. [51, 52], Li [53], and Fricker and Gast [32].

Nonlinear Markov processes play an important role in the study of big networks. Important examples include Rybko and Shlosman [64], Peng [62], Turner [69], Benaim and Le Boudec [8], Frank [30], Kolokoltsov [44], Gast and Gaujal [36], Kolokoltsov [44, 45], Kolokoltsov at al. [46], Muzychka and Vaninsky [59], Dupuis and Fischer [28], Gast at al. [37], Vaninsky et al. [70], Budhiraja et al. [17], Budhiraja and Majumder [19] and Benaim [7].

Metastability is an ubiquitous and important phenomenon of the dynamical behavior of communication networks, e.g., see Gibbens et al. [38], Antunes et al. [12] and Tibi [67].
For metastability in Markov processes, readers may refer to Galves et al. [34], Bovier et al. [15, 16], Olivieri and Vares [60], Freidlin and Wentzell [31], Bovier [13, 14], den Hollander [25], and Beltran and Landim [5, 6].

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 for a collection of 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 $X$ whose
infinitesimal generator is given by
\[
Q = \begin{pmatrix}
Q_{0,0} & Q_{0,1} & Q_{0,2} & Q_{0,3} & \cdots \\
Q_{1,0} & Q_{1,1} & Q_{1,2} & Q_{1,3} & \cdots \\
Q_{2,0} & Q_{2,1} & Q_{2,2} & Q_{2,3} & \cdots \\
Q_{3,0} & Q_{3,1} & Q_{3,2} & Q_{3,3} & \cdots \\
& & & & \\
& & & & \\
& & & & \\
\end{pmatrix},
\]
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 [49] 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, where the states of the \(N\) big networks are denoted as \(X^{1,N}(t), X^{2,N}(t), \ldots, X^{N,N}(t)\), respectively.

Let \(X^N(t) = (X^{1,N}(t), X^{2,N}(t), \ldots, X^{N,N}(t))\). Then the empirical measure of the system of \(N\) weakly interacting big networks is given by
\[
\mu^N(t) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^{i,N}(t)},
\]
where \(\delta_x\) is the Dirac measure at \(x\).

We denote by \(\mathfrak{P}(\Omega)\) the space of probability vectors on the state space \(\Omega\), which is equipped with the usual topology of weak convergence. If \(p \in \mathfrak{P}(\Omega)\), we write \(p = (p_0, p_1, p_2, \ldots)\), 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 \(\{\mu^N(t) : t \geq 0\}\) 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 \(X^N(t)\) is exchangeable, if for any level permutation \((k_{i_1}, k_{i_2}, \ldots, k_{i_N})\) of \((k_1, k_2, \ldots, k_N)\) and any phase permutation \((j_{i_1}, j_{i_2}, \ldots, j_{i_N})\) of \((j_1, j_2, \ldots, j_N)\),
\[
P \{ X^{1,N}(t) = (k_1, j_1), X^{2,N}(t) = (k_2, j_2), \ldots, X^{N,N}(t) = (k_N, j_N) \}
= P \{ X^{i_1,N}(t) = (k_{i_1}, j_{i_1}), X^{i_2,N}(t) = (k_{i_2}, j_{i_2}), \ldots, X^{i_N,N}(t) = (k_{i_N}, j_{i_N}) \}.
\]
In the system of $N$ weakly interacting big networks, the effect of a typical big network on the dynamics of the system of 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 $X^N(0)$ is exchangeable, then at any time $t \geq 0$, the probability distribution of $X^N(t)$ is also exchangeable.

For the system of $N$ weakly interacting big networks, if the probability distribution of $X^N(t)$ 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 any given big network (such as, the first big network); while analysis of the total system will be completed by the propagation of chaos (as $N \to \infty$). Based on this, the infinitesimal generator of the Markov process corresponding to the first big network can be defined as

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

(4)

where the size of the matrix $\Gamma^{(N)}_{j,j}(\mu^N(t))$ 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 $\Gamma^{(N)}(\mu^N(t))$ 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.

Let $\mu^N(t) \to p(t)$ and $\Gamma^{(N)}(\mu^N(t)) \to \Gamma(p(t))$ (a.s.) for $t \geq 0$, as $N \to \infty$. Then $p(t)$ is a probability vector. Furthermore, using some probability analysis, we may obtain
an infinite-dimensional dynamic system of mean-field equations as follows:
\[
\frac{d}{dt} p(t) = p(t) \Gamma (p(t)) \tag{5}
\]
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
\[
\Gamma (p(t)) = \begin{pmatrix}
\Gamma_{0,0}(p(t)) & \Gamma_{0,1}(p(t)) & \Gamma_{0,2}(p(t)) & \Gamma_{0,3}(p(t)) & \cdots \\
\Gamma_{1,0}(p(t)) & \Gamma_{1,1}(p(t)) & \Gamma_{1,2}(p(t)) & \Gamma_{1,3}(p(t)) & \cdots \\
\Gamma_{2,0}(p(t)) & \Gamma_{2,1}(p(t)) & \Gamma_{2,2}(p(t)) & \Gamma_{2,3}(p(t)) & \cdots \\
\Gamma_{3,0}(p(t)) & \Gamma_{3,1}(p(t)) & \Gamma_{3,2}(p(t)) & \Gamma_{3,3}(p(t)) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix} \tag{7}
\]

**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 [8] and Budhiraja and Majumder [19], the Markov decision processes by Gast and Gaujal [36] and Gast et al. [37], the continuous-time Markov chains by Dupuis and Fischer [28] and Budhiraja et al. [17, 18], and some nice practical examples include Mitzenmacher [58], Bobbio et al. [10], Li et al. [51, 52], and Li and Lui [53].

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 result is the Picard approximation as follows. If \( \Gamma (x) \) is (locally) Lipschitz on a set \( E \subseteq \mathcal{P}(\Omega) \), that is, there exists a positive constant \( C \) such that
\[
\|\Gamma (x) - \Gamma (y)\| \leq C \|x - y\|, \quad x, y \in \mathcal{P}(\Omega),
\]
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 $E \subseteq \mathcal{P}(\Omega)$, Li et al. \cite{51} and Li and Lui \cite{53} gave an algorithmic method through dealing with some matrices of infinite orders.

(b) The limiting processes

To discuss the limit: $\mu^N(t) \to p(t)$ (a.s.) for $t \geq 0$, 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 $\mu^{(N)} = \{\mu^N(t) : t \geq 0\}$ is a Markov process on the state space $\mathcal{P}_N(\Omega)$ where $\mathcal{P}_N(\Omega) = \mathcal{P}(\Omega) \cap (\frac{1}{N}\Omega)$, the possible jumps of $\mu^{(N)}$ are of the form $(e_{l,i} - e_{k,j})/N$ for $(l,i) \neq (k,j)$, and $(k,j), (l,i) \in \Omega$. If $\mu^N(t) = x \in \mathcal{P}_N(\Omega)$ at time $t \geq 0$, 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 any big network, is given by $Nx_{k,j}\Gamma^{(N)}_{k,j,l,i}(x)$. Based on this, the generator $A^{(N)}$ of the Markov process $\mu^{(N)}$ is given by

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

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

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

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

$$\frac{\partial}{\partial t} p(t) = p(t) \Gamma(p(t)) \text{ with } p(0) = q.$$

**Proof:** The proof may directly follow from Theorem 2.11 in Kurtz \cite{47}. Here, we only give a simple interpretation 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
\[ F(p) = \sum_{(k,j),(l,i) \in \Omega} p_{k,j} (e_{l,i} - e_{k,j}) \Gamma_{k,j,l,i}(p), \]
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 \( E \subseteq \mathcal{P}(\Omega) \), thus for the sequence \( \{\mu^{(N)}(t), t \geq 0\} \) of Markov processes, it follows from Equation (III.10.13) in Rogers and Williams [63] or Page 162 in Ethier and Kurtz [29] that
\[ M^{(N)}(t) = \mu^{(N)}(t) - \mu^{(N)}(0) - \int_0^t \left\{ \mu^{(N)}(x)\Gamma\left(\mu^{(N)}(x)\right) \right\} \, dx \]
is a martingale with respect to \( N \geq 1 \). Therefore, if \( \{\mu^{(N)}(0)\} \) converges weakly to \( q \in \mathcal{P}(\Omega) \) as \( N \to \infty \), then \( \{\mu^{(N)}(t), N \geq 1\} \) converges weakly in \( D_F[0, +\infty) \) endowed with the Skorohod topology to the solution \( p(t) \) to the ordinary differential equation:
\[ \frac{d}{dt} p(t) = p(t) \Gamma(p(t)) \] with \( p(0) = q \), within \( \mathcal{P}(\Omega) \). This completes the proof.■

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:
\[ \frac{d}{dt} p(t) = p(t) \Gamma(p(t)) \] 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 \( \pi \in \mathcal{P}(\Omega) \) is said to be a fixed point of the ordinary differential equation:
\[ \frac{d}{dt} p(t) = p(t) \Gamma(p(t)) \] with \( p(0) = q \), if \( p(t) \to \pi \) as \( t \to +\infty \), and
\[ \lim_{t \to +\infty} \left[ \frac{d}{dt} p(t) \right] = 0. \]
In this case, it is clear that
\[ \pi \Gamma(\pi) = 0, \quad (8) \]
which is an infinite-dimensional system of nonlinear equations. In general, there are more difficulties and challenging due to the infinite order 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 \( \pi \in \Phi (\Omega) \), \( \Gamma (\pi) \) is the infinitesimal generator of an irreducible continuous-time Markov process. Based on Li [49], we can develop the UL-type \( RG \)-factorization of the matrix \( \Gamma (\pi) \). To that end, we partition the matrix \( \Gamma (\pi) \) as

\[
\Gamma (\pi) = \begin{pmatrix}
T (\pi) & U (\pi) \\
V (\pi) & W (\pi)
\end{pmatrix}
\]

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 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 \( W^{-1}_{\max} (\pi) \) of \( W (\pi) \), i.e., \( W^{-1} (\pi) \leq W^{-1}_{\max} (\pi) \leq 0 \) for any non-positive inverse \( W^{-1} (\pi) \). Of course, \( 0 \leq [-W (\pi)]^{-1}_{\min} \leq [-W (\pi)]^{-1} \) for any non-negative inverse \( [-W (\pi)]^{-1} \) of \( -W (\pi) \), that is, \( [-W (\pi)]^{-1}_{\min} \) is the minimal nonnegative inverse of \( -W (\pi) \). Based on this, for \( n \geq 0 \) we write

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

where the size of the matrix \( \phi^{(n)}_{j,j} (\pi) \) 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 [49] that for \( n \geq 0, 0 \leq i, j \leq n \),

\[
\phi^{(n)}_{i,j} (\pi) = \Gamma_{i,j} (\pi) + \sum_{k=n+1}^{\infty} \phi^{(k)}_{i,k} (\pi) [-\phi^{(k)}_{k,k} (\pi)]^{-1} \phi^{(k)}_{k,j} (\pi).
\]

Let

\[
\Psi_n (\pi) = \phi^{(n)}_{0,n} (\pi), \quad n \geq 0;
\]

\[
R_{i,j} (\pi) = \phi^{(j)}_{i,j} (\pi) [-\phi^{(k)}_{j,j} (\pi)]^{-1}, \quad 0 \leq i < j;
\]

and

\[
G_{i,j} (\pi) = [-\phi^{(k)}_{i,i} (\pi)]^{-1} \phi^{(i)}_{i,j} (\pi), \quad 0 \leq j < i.
\]
Then the UL-type $RG$-factorization of the matrix $\Gamma (\pi)$ is given by
\[
\Gamma (\pi) = [I - R_U (\pi)] \Psi_D (\pi) [I - G_L (\pi)],
\]  
where
\[
R_U (\pi) = \begin{pmatrix}
0 & R_{0,1} (\pi) & R_{0,2} (\pi) & R_{0,3} (\pi) & \cdots \\
0 & R_{1,2} (\pi) & R_{1,3} (\pi) & \cdots \\
0 & R_{2,3} (\pi) & \cdots \\
0 & \cdots \\
\end{pmatrix},
\]  
\[
\Psi_D (\pi) = \text{diag} \left( \Psi_0 (\pi), \Psi_1 (\pi), \Psi_2 (\pi), \Psi_3 (\pi), \ldots \right)
\]
and
\[
G_L (\pi) = \begin{pmatrix}
0 & G_{1,0} (\pi) & 0 \\
G_{2,0} (\pi) & G_{2,1} (\pi) & 0 \\
G_{3,0} (\pi) & G_{3,1} (\pi) & G_{3,2} (\pi) & 0 \\
\vdots & \vdots & \vdots & \vdots \\
\end{pmatrix}.
\]

Based on the UL-type $RG$-factorization (9), it follows from Subsection 2.7.3 in Li [49] that the fixed point $\pi$ is given by
\[
\begin{cases}
\pi_0 = \tau x_0 (\pi), \\
\pi_k = \sum_{i=0}^{k-1} \pi_i R_{i,k} (\pi), \quad k \geq 1,
\end{cases}
\]
where $x_0 (\pi)$ is the fixed point of the censored Markov chain $\Psi_0 (\pi)$ to level 0, and the scalar $\tau$ is determined by $\sum_{k=0}^{\infty} \pi_k e = 1$ uniquely.

Using the expression (10) of the fixed point $\pi$, we set up an important relation as follows:
\[
\pi = \left( \tau x_0 (\pi), \pi_0 R_{0,1} (\pi), \sum_{i=0}^{1} \pi_i R_{i,k} (\pi), \sum_{i=0}^{2} \pi_i R_{i,k} (\pi), \ldots \right).
\]

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

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

where the two vectors $\pi_0$ and $\pi_1$ satisfy the following system of nonlinear matrix equations

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

and

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

Thus, the fixed point equation \[11\] with $R$-measure is simplified as

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

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

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

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

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
\[ \Psi_0 (\pi) = B_1 (\pi) + \sum_{k=2}^{\infty} B_k (\pi) G^{k-2} (\pi) G_1 (\pi) \]
for \( k \geq 1 \)
\[ \Psi (\pi) = A_1 (\pi) + \sum_{k=2}^{\infty} A_k (\pi) G^{k-1} (\pi) ; \]
and the \( R \)-measure
\[ R_{0,j} (\pi) = \left[ \sum_{k=j+1}^{\infty} B_k (\pi) G^{k-j} (\pi) \right] [\Psi (\pi)]^{-1} , \quad j \geq 1, \]
for \( i \geq 1 \)
\[ R_j (\pi) = \left[ \sum_{k=j+1}^{\infty} A_k (\pi) G^{k-j} (\pi) \right] [\Psi (\pi)]^{-1} , \quad j \geq 1. \]
The fixed point \( \pi \) 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 fixed point of the censored Markov chain \( \Psi_0 (\pi) \) to level 0, and the scalar \( \tau \) is determined by \( \sum_{k=0}^{\infty} \pi_k e = 1 \) uniquely. Thus, the fixed point equation (11) with \( R \)-measure is simplified as
\[ \pi = \left( \tau x_0 (\pi), \pi_0 R_{0,1} (\pi), \pi_0 R_{0,2} (\pi) + \pi_1 R_1 (\pi), \pi_0 R_{0,3} (\pi) + \sum_{i=1}^{2} \pi_i R_{k-i} (\pi), \ldots \right). \]

Now, we write the fixed point equation (11) with \( R \)-measure as a functional form:
\[ \pi = F \left( R (\pi) \right), \]
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: \( \pi^{(0)} \in \mathcal{P} (\Omega) \).

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

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

**Step four:** For a sufficiently small \( \varepsilon > 0 \), if \( \| \pi^{(N+1)} - \pi^{(N)} \| < \varepsilon \), 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. In this case, it is a key...
to design a suitable initial probability vector: \( \pi^{(0)} \in \mathcal{P}(\Omega) \), for example, for any integer \( m \geq 1 \) we take

\[
\pi^{(0)} = \left( \frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}, 0, \ldots \right).
\]

Now, we provide another algorithm for computing the fixed points. To do so, 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

\[
\pi_0 \Psi_0(\pi) = 0, \quad \pi_0 e = \tau \in (0, 1).
\]

Thus it is easy to see from the irreducibility of the matrix \( \Gamma(\pi) \) that for the matrix \( \Psi_0(\pi) \) of size \( m_0 \), \( \text{rank}(\Psi_0(\pi)) = m_0 - 1 \) according to the irreducibility of the matrix \( \Psi_0(\pi) \), and its eigenvalue with the maximal real part is equal to zero. 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:

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

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 \( \hat{\pi} \) to the nonlinear characteristic equation:

\[
\det(\Psi_0(\pi)) = 0.
\]

**Step two:** Check whether \( \text{rank}(\Psi_0(\hat{\pi})) = m_0 - 1 \). If Yes, then \( \hat{\pi} \) 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: \( \frac{dp(t)}{dt} = p(t) \Gamma(p(t)) \) 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
\[ S = \{ \pi : \pi \Gamma (\pi) = 0, \pi e = 1 \} . \]

Then it is clear that
\[ S = \left\{ \pi : \pi = \left( \tau x_0 (\pi), \pi_0 R_{0,1} (\pi), \sum_{i=0}^{1} \pi_i R_{i,k} (\pi), \sum_{i=0}^{2} \pi_i R_{i,k} (\pi), \ldots \right), \pi e = 1 \right\} \]
or
\[ S = \{ \pi : \det (\Psi_0 (\pi)) = 0, \operatorname{rank} (\Psi_0 (\pi)) = m_0 - 1, \pi e = 1 \} \]
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 \( S \). At the same time, an argument by analytic function can indicate that the elements of the set \( S \) are isolated.

To describe the isolated element structure of the set \( S \), 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
\[ \frac{d}{dt} \mu (t) = p (t) \Gamma (p (t)) \]
with \( p (0) = q \). Thus, the Birkhoff center of the solution \( \Phi (t) \) is defined as
\[ \Theta = \left\{ \bar{P} \in \mathcal{P} (\Omega) : \bar{P} = \lim_{k \to \infty} \Phi (t_k) \right\} \]
for any scale sequence \( \{t_k\} \) with \( t_l \geq 0 \) for \( l \geq 1 \) and \( \lim_{k \to \infty} t_k = +\infty \).

Notice that perhaps \( \Theta \) 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 \( S \subset \Theta \). Obviously, the limiting empirical Markov process \( \{ \mathbf{Y} (t) : t \geq 0 \} \) spends most of its time in the Birkhoff center \( \Theta \), where \( \mathbf{Y} (t) = \lim_{N \to \infty} \mu_N (t) \).

**Case two:** \( t \to +\infty \). In this case, we write
\[ \pi^{(N)} = \lim_{t \to +\infty} \mu_N (t), \text{ a.s.,} \]
since for each \( N = 1, 2, 3, \ldots \), if the system of \( N \) weakly interacting big networks is stable.
Let
\[ \Xi = \left\{ \pi \in \mathcal{P}(\Omega) : \pi = \lim_{k \to \infty} \pi^{(N_k)} \text{ for any positive integer sequence } \{N_k\} \text{ with } 1 \leq N_1 \leq N_2 \leq N_3 \leq \cdots \text{ and } \lim_{k \to \infty} N_k = \infty \right\}. \]

It is easy to see that
\[ \mathcal{S} \subset \Xi \subset \Theta. \]

At the same time, it is clear that
\[ \mathcal{S} = \{ \text{the local minimal points in } \Theta \} \cup \{ \text{the local maximal points in } \Theta \} \cup \{ \text{the saddle points in } \Theta \}. \]

and
\[ \Theta - \mathcal{S} = \{ \text{the limit cycles in } \Theta \}. \]

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:
\[ \frac{d}{dt} p(t) = p(t) \Gamma(p(t)) \text{ with } p(0) = q. \]

The Lyapunov function \( g \) defined on \( \mathcal{P}(\Omega) \) is constructed such that
\[ y \Gamma(y) \cdot \nabla g(y) \leq 0, \quad y \in \mathcal{P}(\Omega). \] (13)

It is easy to see that if \( \pi \in \mathcal{S} \), 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 \( \pi \in \mathcal{S} \).

Let \( |\mathcal{S}| \) be the number of elements in the set \( \mathcal{S} \). If \( |\mathcal{S}| = 1 \), then
\[ \lim_{N \to \infty} \lim_{t \to +\infty} \mu^N(t) = \lim_{t \to +\infty} \lim_{N \to \infty} \mu^N(t) = \pi, \text{ a.s..} \]

If \( |\mathcal{S}| \geq 2 \), then the system of big networks exhibits a metastability property, that is, the state of the given big network switches from one stable point to the other 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 [14] and Olivieri and Vares [60] 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 \( p, q \in \mathcal{P}(\Omega) \), we define the relative entropy of \( p \) with respect to \( q \) as

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

Let \( \Psi(z) = z \log z - z + 1 \). Then if \( p(t) \) and \( q(t) \) are two different solutions to the ordinary differential equation

\[
\frac{d}{dt} p(t) = p(t) \Lambda,
\]

where \( \Lambda \) is the infinitesimal generator of an irreducible continuous-time Markov process. In this case, Dupuis and Fischer [28] indicated that

\[
\frac{d}{dt} R(p(t)||q(t)) = -\sum_{x,y \in \Omega} \Psi \left( \frac{p_y(t) q_x(t)}{p_x(t) q_y(t)} \right) \frac{q_y(t)}{q_x(t)} \Lambda_{y,x} \leq 0,
\]

and \( \frac{d}{dt} R(p(t)||q(t)) = 0 \) if and only if \( p(t) = q(t) \) for \( t \geq 0 \). Obviously, \( \frac{d}{dt} R(p(t)||\pi) = 0 \) if and only if \( p(t) = \pi \) for \( t \geq 0 \).

Dupuis and Fischer [28] further demonstrated that for the ordinary differential equation: \( \frac{d}{dt} p(t) = p(t) \Gamma(p(t)) \), the relative entropy relation (14) cannot be applied directly. In this case, they first defined \( P^{(N)}(t) \) as the state probability of the system of \( N \) big networks at time \( t \geq 0 \), and let \( P^{(N)}(0) = \otimes^N q \). Then they gave an approximate method to construct the Lyapunov function as follows:

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

For applying the relative entropy to construct a Lyapunov function, readers may also refer to Budhiraja et al. [17, 18] for more details.

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

If there exists only an element \( \pi \) in the set \( S \), then

\[
\mu^{(N)}(t) \Rightarrow p(t), \quad \text{as } N \to \infty,
\]

and

\[
p(t) \to \pi, \quad \text{as } t \to +\infty.
\]

When there exist multiple elements in the set \( S \), 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. [17]. For convenience of readers, we restate the definition of locally stable fixed points here.

We define the relative interior of $\Omega^{\circ}$ of $\Omega$ as

$$
\Omega^{\circ} = \{ g \in \Omega : g > 0 \}.
$$

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

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

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

$$
S = \left\{ \pi^{(1)}, \pi^{(2)}, \pi^{(3)}, \ldots \right\}, \quad \pi^{(k)} \in \mathcal{F}^{(k)} \text{ for } k \geq 1,
$$

and

$$
Q = \left\{ q^{(1)}, q^{(2)}, q^{(3)}, \ldots \right\}, \quad q^{(k)} \in \mathcal{F}^{(k)} \text{ for } k \geq 1.
$$

If for each $k \geq 1$,

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

then

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

and

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

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

Based on the above analysis, for each locally stable fixed point in $S$, we find the similar property (17) and (18) to that in (15) and (16). Thus computation of the locally stable fixed points is on a unified line, but the the initial vector $q^{(k)} \in \mathcal{F}^{(k)}$ with $\mu^{(N)}(0) \Rightarrow q^{(k)}$ as $N \to \infty$ has a large impact on the limit $\pi^{(k)}$ 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:

\[ \mathcal{F} = \left\{ \mathcal{F}^{(1)}, \mathcal{F}^{(2)}, \mathcal{F}^{(3)}, \ldots \right\}, \]

which sufficiently corresponds to the set \( \mathcal{S} = \left\{ \pi^{(1)}, \pi^{(2)}, \pi^{(3)}, \ldots \right\}. \)

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 \( \theta_p \) 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:

\[
\frac{d}{dt} p(t) = p(t) \Gamma (p(t)).
\]

**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: \( \frac{d}{dt} p(t) = p(t) \Gamma (p(t)) \) and \( \frac{d}{dt} p_0 (t) = p_0 (t) \Psi_0 (p(t)). \)

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 network, 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 \cite{49} 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.

Acknowledgement

We would like to thank Professor Jeffrey J. Hunter for his useful helps and suggestions, and acknowledge his pioneering research on Markov processes, Markov renewal processes and generalized inverses in which those interesting results play a foundational role in the area of applied probability, such as, the Poisson equations, the mixing times, and many stationary computation. At the same time, this work is partly supported by the National Natural Science Foundation of China under grant (#71271187, #71471160), and the Fostering Plan of Innovation Team and Leading Talent in Hebei Universities under grant (# LJRC027).

References

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

[2] N. Antunes, C. Fricker, P. Robert and D. Tibi (2008). Stochastic networks with multiple stable points. The Annals of Probability, Vol. 36, 255–278.
[3] F. Baccelli, F.I. Karpelevich, M.Y. Kelbert, A.A. Puhalskii, A.N. Rybko and Y.M. Suhov (1992). A mean-field limit for a class of queueing networks. *Journal of Statistical Physics*, Vol. 66, 803–825.

[4] F. Baccelli, A.N. Rybko and S. Shlosman (2013). Queuing networks with varying topology – a mean-field approach. arXiv preprint: [arXiv:1311.3898](http://arxiv.org/abs/1311.3898).

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

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

[7] M. Benaim (2014). On gradient like properties of population games, learning models and self reinforced processes. arXiv preprint: [arXiv:1409.4091](http://arxiv.org/abs/1409.4091).

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

[9] M. Benaim and J.Y. Le Boudec (2011). On mean field convergence and stationary regime. arXiv preprint: [arXiv:1111.5710](http://arxiv.org/abs/1111.5710).

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

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

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

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

[14] A. Bovier (2006). Metastability: A potential theoretic approach. In: *Proceedings oh the International Congress of Mathematicians: Invited Lectures*, Pages 499–518.
[15] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein (2001). Metastability in stochastic
dynamics of disordered mean-field models. Probability Theory and Related Fields, Vol.
119, 99–161.

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

[17] A. Budhiraja, P. Dupuis, M. Fischer and K. Ramanan (2014). Local stability of Kol-
mogorov forward equations for finite state nonlinear Markov processes. arXiv preprint:

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

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

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

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

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

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

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

[25] F. den Hollander (2004). Metastability under stochastic dynamics. Stochastic Pro-
cesses and their Applications, Vol. 114, 1–26.
[26] N.G. Duffield (1992). Local mean-field Markov processes: An application to message-switching networks. *Probability Theory and Related Fields*, Vol. 93, 485–505.

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

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

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

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

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

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

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

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

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

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

[37] N. Gast, B. Gaujal and J.Y. Le Boudec (2012). Mean field for Markov decision processes: from discrete to continuous optimization. *IEEE Transactions on Automatic Control*, Vol. 57, 2266–2280.
[38] R.J. Gibbens, P.J. Hunt and F.P. Kelly (1990). Bistability in communication networks. *Disorder in Physical Systems*, 113–128.

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

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

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

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

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

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

[45] V.N. Kolokoltsov (2011). Nonlinear Lévy and nonlinear Feller processes: An analytic introduction. arXiv preprint: [arXiv:1103.5591](https://arxiv.org/abs/1103.5591).

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

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

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

[49] Q.L. Li (2010). *Constructive Computation in Stochastic Models with Applications: The RG-Factorizations*. Springer, Tsinghua Press.
[50] Q.L. Li (2014). Tail probabilities in queueing processes. Asia-Pacific Journal of Operational Research, Vol. 31, No. 2, 1–31.

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

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

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

[54] T.M. Liggett (2006). Interacting Particle Systems. Springer.

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

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

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

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

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

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

[61] V.I. Oseledets and D.V. Khmelev (2002). Stochastic transportation networks and stability of dynamical systems. Theory of Probability & Its Applications, Vol. 46, 154–161.
[62] S.G. Peng (2005). Nonlinear expectations and nonlinear Markov chains. *Chinese Annals of Mathematics*, Vol. 26, No. 2, 159–184.

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

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

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

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

[67] D. Tibi (2010). Metastability in communication networks. arXiv preprint: [arXiv:1002.0796](https://arxiv.org/abs/1002.0796).

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

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

[70] K. Vaninsky, S. Myzuchka and A. Lukov (2012). A multi-agent nonlinear Markov model of the order book. arXiv preprint: [arXiv:1208.3083](https://arxiv.org/abs/1208.3083).

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

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