Second order moment approximations for asymptotically density dependent Markov chains

Yilun Shang[1,2]

Abstract

We consider differential equation approximations for continuous time Markov chains with asymptotically density dependent transition rates. Based on some operator semigroup techniques, we show that the second order moment of the Markov process can be approximated uniformly by the solution of an appropriately chosen mean-field equation. The convergence rate is shown to be given by $O(n^{-1})$ with n being the size of the state space.

Key words: mean-field approximation, one-parameter operator semigroup, Markov chain.

MSC 2010: 47D06, 60J28.

1 Introduction

Differential equation (mean-field) approximations for stochastic processes are receiving a growing research attention in a variety of fields, e.g., disease dissemination, statistical physics and natural evolution, see [1]. These approximations often provide fine insight on the behavior of the stochastic system since deterministic processes can usually be addressed more easily. From a mathematical point of view many stochastic systems can be translated to Markovian random processes, like continuous time Markov chains [2]. A continuous time Markov chain with finite state space $\{0, 1, \cdots, n\}$ ($n \in \mathbb{N}$), for example, is described by its Kolmogorov equations

$$\dot{p}_k(t) = b_{k-1}p_{k-1}(t) - (b_k + d_k)p_k(t) + d_{k+1}p_{k+1}(t), \quad k = 0, 1, \cdots, n, \tag{1}$$

[1]Department of Mathematics, Tongji University, Shanghai 200092, China; Einstein Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem 91904, Israel

[2]Email: shyl@tongji.edu.cn

where $p_k(t)$ represents the probability that the process is in state k at time t. Alternatively, with the aid of the infinitesimal generator A_n^T with

$$A_n = \begin{pmatrix} -b_0 - d_0 & b_0 & 0 & \cdots & 0 & 0 \\ d_1 & -b_1 - d_1 & b_1 & \cdots & 0 & 0 \\ 0 & d_2 & -b_2 - d_2 & \ddots & \vdots & \vdots \\ \vdots & \vdots & d_3 & \ddots & b_{n-2} & 0 \\ \vdots & \vdots & \vdots & \ddots & -b_{n-1} - d_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & d_n & -b_n - d_n \end{pmatrix},$$

and $p(t) = (p_0(t), p_1(t), \cdots, p_n(t))^T$, the Kolmogorov equations (1) are equivalent to $\dot{p}(t) = A_n^T p(t)$. The particular tri-diagonal structure of A_n corresponds to a birth-death process with transition rates [2]

$$\begin{aligned} k &\longrightarrow k+1 \quad \text{at rate} \quad b_k, \\ k &\longrightarrow k-1 \quad \text{at rate} \quad d_k, \end{aligned}$$

with $b_n = d_0 = 0$.

It is the goal of this paper to prove the convergence of second order moment of the Markov chain (1) to the solution of a mean-field differential equation as n tends to infinity. From the differential equation point of view, the Markovian stochastic system is governed by its Kolmogorov equations, which are linear ordinary differential equations describing the time evolution of the system. By using Trotter type approximation theorems, Kurtz [3, 4] showed that the pure jump density dependent Markov chain converges in probability to the solution of a mean-field model. McVinish and Pollett [5] showed the weak convergence of a density dependent Markov chain with individual variation to a deterministic process by constructing measure-valued Markov chains. Based on primary operator semigroup approaches [12], the expected value of an asymptotically density dependent Markov chain is shown to converge to the solution of a mean-field equation by Bátkai et al. in [6]. For more approximation techniques and applications, we refer the reader to [7, 8, 9, 10] and references therein.

Let $b_k = B_n(k)$ and $d_k = D_n(k)$, the Markov chain (1) is said to be density dependent [4] if there exist $b(x), d(x) \in C[0,1]$ for all $x \in [0,1]$ and $n \in \mathbb{N}$, $b(x) = B_n(nx)/n$ and $d(x) = D_n(nx)/n$ hold. The asymptotical density dependence [6, 11] means that

$$b(x) = \lim_{n\to\infty} \frac{B_n(nx)}{n} \quad \text{and} \quad d(x) = \lim_{n\to\infty} \frac{D_n(nx)}{n}$$

hold for all $x \in [0,1]$. Given $n \in \mathbb{N}$, let $X_n(t)$ be a Markov chain corresponding to (1) taking values in $S = \{0, 1/n, 2/n, \cdots, 1\}$. Then we have $p_k(t) = P(X_n(t) = k/n)$. Our main result goes as follows.

Theorem 1. *Let*

$$y(t) = \sum_{k=0}^{n} \frac{k^2}{n^2} p_k(t) \tag{2}$$

be the second order moment of the Markov process $X_n(t)$. Suppose that $x(t)$ is the solution of the initial value problem

$$\begin{cases} \dot{x}(t) = 2(b(x) - d(x)) \quad for\ t \geq 0, \\ x(0) = y(0), \end{cases} \tag{3}$$

where $b(x), d(x) \in C^2[0,1]$ satisfying

$$\left| b(x) - \frac{B_n(\alpha_n x)}{\alpha_n} \right| = O\left(\frac{1}{n}\right) \quad and \quad \left| d(x) - \frac{D_n(\alpha_n x)}{\alpha_n} \right| = O\left(\frac{1}{n}\right) \tag{4}$$

for any $\alpha_n \leq n^2$ and $x \in [0,1]$ with $\alpha_n x \in S$. Then for any $t_0 > 0$, there exists a constant C such that

$$|x(t) - y(t)| \leq \frac{C}{n}$$

for all $t \in [0, t_0]$.

The expected value of the Markov chain $X_n(t)$ is given by $z(t) = \sum_{k=0}^{n} (k/n) p_k(t)$. Bátkai et al. [6, Theorem 1] showed that $z(t)$ can be approximated, under the same condition of Theorem 1, by the solution of an initial value problem similar (differ by a factor of two) to (3). Our methodology is to first introduce a new Markov chain X'_n which has different state space from that of X_n but share essentially the same infinitesimal generator with it. Same operator semigroup techniques used in [6] are then applied to X'_n, which in turn gives the second order moment approximation of the original chain X_n.

Our investigation of convergence of the second order moment instead of weak/stochastic convergence is motivated by the following two considerations. Firstly, no sophisticated probabilistic techniques are needed, which yields a less technical proof and a more transparent demonstration of the result. Secondly, the convergence of second order moment can be readily combined with that of the first order moment ([6, Theorem 1]) to test the goodness (via deviation) of the approximation effectively. As is known, the weak/stochastic convergence is rarely examined in practice.

The rest of the paper is organized as follows. In Section 2, we derive the mean-field equation in (3). In Section 3, we present the proof of Theorem 1.

2 Derivation of mean-field equation

In this section, we derive the approximating differential equation in (3) for the second order moment of Markov chain $X_n(t)$. Following Bátkai et al. [6], by using (2) and the Kolmogorov equations (1), we obtain

$$\begin{aligned}\dot{y}(t) &= \sum_{k=0}^{n} \frac{k^2}{n^2}\dot{p}_k(t) \\ &= \sum_{k=0}^{n}\left(\frac{(k+1)^2}{n^2}b_k - \frac{k^2}{n^2}(b_k+d_k) + \frac{(k-1)^2}{n^2}d_k\right)p_k(t) \\ &= 2\sum_{k=0}^{n}\frac{k(b_k-d_k)}{n^2}p_k(t) + \sum_{k=0}^{n}\left(\frac{b_k+d_k}{n^2}\right)p_k(t), \qquad (5)\end{aligned}$$

where we used the convention that $d_0 = b_n = 0$.

It follows from (4) and the continuity of $b(x), d(x)$ that b_k/n and d_k/n are bounded. Hence, the last term in (5) approaches 0 as $n \to \infty$. We have

$$\begin{aligned}\dot{y}(t) &\approx 2\sum_{k=0}^{n}\frac{k(b_k-d_k)}{n^2}p_k(t) \\ &= 2\sum_{k=0}^{n}\left(\frac{kB_n(n^2\frac{k}{n^2})}{n^2} - \frac{kD_n(n^2\frac{k}{n^2})}{n^2}\right)p_k(t). \qquad (6)\end{aligned}$$

Again applying the asymptotic density dependence (4), the right-hand side of (6) can be approximated by

$$2\sum_{k=0}^{n}\left(kb\left(\frac{k}{n^2}\right) - kd\left(\frac{k}{n^2}\right)\right)p_k(t).$$

In view of the Taylor formula for the functions $b(\cdot)$ and $d(\cdot)$, higher-order moments for the Markov process X_n will be needed. To obtain a self-contained equation at the level of the second order moment, we resort to the following closure

$$\sum_{k=0}^{n}\left(kb\left(\frac{k}{n^2}\right) - kd\left(\frac{k}{n^2}\right)\right)p_k(t) \approx b\left(\sum_{k=0}^{n}\frac{k^2}{n^2}p_k(t)\right) - d\left(\sum_{k=0}^{n}\frac{k^2}{n^2}p_k(t)\right).$$

Recall the definition (2) and we finally obtain the following mean-field approximation

$$\dot{y}(t) \approx 2(b(y) - d(y)).$$

For other moment closure techniques used in differential equation approximations, we refer the interested reader to [1, Sec. IV.D] and references therein.

3 Proof of Theorem 1

In this section, we will prove Theorem 1 employing operator semigroup approaches [6, 12].

To start with, we define a new continuous time Markov chain $X'_n(t)$ taking values in $S' = \{0, (1/n)^2, (2/n)^2, \cdots, 1\}$. Its infinitesimal generator is the same as A_n^T defined above. We have $p_k(t) = P(X'_n(t) = (k/n)^2)$ and the transition probabilities for the Markov chain X'_n are represented by

$$p_{j,k}(t) = P\left(X'_n(t) = \frac{k^2}{n^2}\middle| X'_n(0) = \frac{j^2}{n^2}\right).$$

Therefore, we obtain $p_k(t) = \sum_{j=0}^{n} p_{j,k}(t)p_j(0)$ and the matrix semigroup $(T_n(t))_{t\geq 0}$ defined by

$$T_n(t) := (p_{j,k}(t)) = e^{tA_n}$$

is a uniformly continuous operator semigroup on $\mathbb{R}^{n+1}$ with the following identification $\mathbb{R}^{n+1} = \big\{f : f \text{ maps } \{0, (1/n)^2, (2/n)^2, \cdots, 1\} \text{ to } \mathbb{R}\big\}$. As is known, A_n is the generator of semigroup $(T_n(t))_{t\geq 0}$, and

$$(T_n(t)f)\left(\frac{j^2}{n^2}\right) = \sum_{k=0}^{n} f\left(\frac{k^2}{n^2}\right) p_{j,k}(t). \tag{7}$$

Let $\varphi(t, x(0))$ be the solution of the Cauchy problem (3). The operator semigroup $(T(t))_{t\geq 0}$ defined by

$$(T(t)f)(x(0)) := f(\varphi(t, x(0))) \tag{8}$$

for $f \in C[0, 1]$ is strongly continuous on $C[0, 1]$ [12, pp. 91-92]. Its generator $(A, D(A))$ with domain $D(A) = C^1[0, 1]$ satisfies

$$(Af)(x(0)) = 2(b(x(0)) - d(x(0)))f'(x(0)). \tag{9}$$

To approximate the semigroup $(T_n(t))_{t\geq 0}$ using the semigroup $(T(t))_{t\geq 0}$, the following projection-like linear operators between the approximation spaces $\mathbb{R}^{n+1}$ and the space $C^1[0, 1]$ are necessitated. We choose operators [13]

$$J_n : \mathbb{R}^{n+1} \to C^1[0, 1], \quad J_n(f) := g,$$

and

$$P_n : C^1[0, 1] \to \mathbb{R}^{n+1}, \quad P_n(g) := f,$$

with $f((k/n)^2) = g((k/n)^2)$ for $k = 0, 1, \cdots, n$, such that $\|J_n\| \le 1$, $\|P_n\| \le 1$, $P_n J_n = \mathrm{id}_{\mathbb{R}^{n+1}}$ for $n \in \mathbb{N}$, and $\lim_{n\to\infty} J_n P_n f = f$ for $f \in C^1[0,1]$.

The following lemma is crucial to the proof of Theorem 1.

Lemma 1. *Suppose that the conditions in Theorem 1 hold. For $f \in C^2[0,1]$ and $t_0 > 0$, there exists a constant $C = C(f, t_0) > 0$ such that*

$$\|(P_n T(t) - T_n(t) P_n) f\| \le \frac{C}{n}$$

for all $t \in [0, t_0]$.

Proof. The families $\tilde{T}_n(t) := J_n T_n(t) P_n$, $t \ge 0$, define a strongly continuous semigroup on $C^1[0,1]$ with generator $\tilde{A}_n$ given by $\tilde{A}_n = J_n A_n P_n$. Using the variation of parameters formula [12, p. 161], we obtain

$$\begin{aligned} (P_n T(t) - T_n(t) P_n) f &= P_n (T(t) f - \tilde{T}_n(t) f) \\ &= \int_0^t P_n \tilde{T}_n(t-s)(A - \tilde{A}_n) T(s) f \mathrm{d}s \\ &= \int_0^t T_n(t-s)(P_n A - A_n P_n) T(s) f \mathrm{d}s, \end{aligned} \tag{10}$$

for any $f \in C^2[0,1]$.

In the following, we estimate the discrepancy $P_n A - A_n P_n$. It follows from (9) and the definition of A_n that

$$(P_n A f)\left(\frac{k^2}{n^2}\right) = 2\left(b\left(\frac{k^2}{n^2}\right) - d\left(\frac{k^2}{n^2}\right)\right) f'\left(\frac{k^2}{n^2}\right)$$

and

$$(A_n P_n f)\left(\frac{k^2}{n^2}\right) = \frac{2k b_k}{n^2} \cdot \frac{f(\frac{(k+1)^2}{n^2}) - f(\frac{k^2}{n^2})}{\frac{2k}{n^2}} - \frac{2k d_k}{n^2} \cdot \frac{f(\frac{k^2}{n^2}) - f(\frac{(k-1)^2}{n^2})}{\frac{2k}{n^2}}$$

for $f \in C^2[0,1]$ and $k = 0, 1, \cdots, n$. Hence,

$$\begin{aligned} (P_n A f)\left(\frac{k^2}{n^2}\right) - (A_n P_n f)\left(\frac{k^2}{n^2}\right) &= 2\left(b\left(\frac{k^2}{n^2}\right) - \frac{k b_k}{n^2} - d\left(\frac{k^2}{n^2}\right) + \frac{k d_k}{n^2}\right) f'\left(\frac{k^2}{n^2}\right) \\ &\quad + \frac{2k b_k}{n^2}\left(f'\left(\frac{k^2}{n^2}\right) - \frac{f(\frac{(k+1)^2}{n^2}) - f(\frac{k^2}{n^2})}{\frac{2k}{n^2}}\right) \\ &\quad + \frac{2k d_k}{n^2}\left(\frac{f(\frac{k^2}{n^2}) - f(\frac{(k-1)^2}{n^2})}{\frac{2k}{n^2}} - f'\left(\frac{k^2}{n^2}\right)\right). \end{aligned} \tag{11}$$

From (4) we have for all $k = 0, 1, \cdots, n$,

$$\left| b\left(\frac{k^2}{n^2}\right) - \frac{k b_k}{n^2} \right| = \left| b\left(\frac{k^2}{n^2}\right) - \frac{B_n(\frac{n^2}{k}\frac{k^2}{n^2})}{\frac{n^2}{k}} \right| \le \frac{L}{n}$$

and similarly

$$\left|d\left(\frac{k^2}{n^2}\right)-\frac{kd_k}{n^2}\right|=\left|d\left(\frac{k^2}{n^2}\right)-\frac{D_n(\frac{n^2}{k}\frac{k^2}{n^2})}{\frac{n^2}{k}}\right|\le\frac{L}{n},$$

for some $L>0$. Moreover, $kb_k/n^2\le C$ and $kd_k/n^2\le C$ for some constant $C>0$ since the functions $b(\cdot),d(\cdot)\in C^2[0,1]$. Applying Taylor's formula for $f\in C^2[0,1]$, we have for every $k=0,1,\cdots,n$, there exists $\eta_k\in((k/n)^2,((k+1)/n)^2)$ such that

$$\left|f'\left(\frac{k^2}{n^2}\right)-\frac{f(\frac{(k+1)^2}{n^2})-f(\frac{k^2}{n^2})}{\frac{2k}{n^2}}\right|\le\left|\frac{(2k+1)f''(\eta_k)}{2n^2}\right|\le\frac{\|f''\|}{n-1}$$

holds. Since $f'((k/n)^2)\le\|f'\|$, we have from (11) that

$$\|(P_nA-A_nP_n)f\|\le\frac{K\|f''\|}{n}+\frac{L\|f'\|}{n} \tag{12}$$

for some constant $K>0$ and all n large enough.

Inserting (12) into (10), and recall that $T(t)$ maps $C^2[0,1]$ into itself (by the continuous dependence theorem), we can derive that for $f\in C^2[0,1]$ there exist two constant $\hat{K},\hat{L}>0$ such that

$$\|(P_nT(t)-T_n(t)P_n)f\|\le\int_0^t\frac{\hat{K}\|(T(s)f)''\|}{n}+\frac{\hat{L}\|(T(s)f)'\|}{n}\mathrm{d}s.$$

Since $b(\cdot),d(\cdot)\in C^2[0,1]$, we obtain $\varphi(t,\cdot)\in C^2[0,1]$ by continuous dependence. Consequently, for each $s\in[0,t]$, there exist $M_1,M_2>0$ such that

$$\|(T(s)f)'\|=\|(f(\varphi(s,\cdot)))'\|=\|f'(\varphi(s,\cdot))\varphi'(s,\cdot)\|\le M_1\|f\|_{H^2}$$

and

$$\|(T(s)f)''\|=\|f''(\varphi(s,\cdot))(\varphi'(s,\cdot))^2+f'(\varphi(s,\cdot))\varphi''(s,\cdot)\|\le M_2\|f\|_{H^2},$$

where $\|f\|_{H^2}<\infty$ is the Sobolev space norm. This readily concludes the proof of Lemma 1. □

Now, take $f=\mathrm{id}:[0,1]\to[0,1]$ in Lemma 1. We obtain that there exists a constant $C>0$ such that

$$\|(P_nT(t)-T_n(t)P_n)\mathrm{id}\|\le\frac{C}{n} \tag{13}$$

for all $t \in [0, t_0]$. From (2) and (7) it follows that

$$\begin{aligned} y(t) &= \sum_{k=0}^{n} \frac{k^2}{n^2} p_k(t) \\ &= \sum_{k=0}^{n} \frac{k^2}{n^2} \sum_{j=0}^{n} p_{j,k}(t) p_j(0) \\ &= \sum_{j=0}^{n} p_j(0) \sum_{k=0}^{n} \mathrm{id}\left(\frac{k^2}{n^2}\right) p_{j,k}(t) \\ &= \sum_{j=0}^{n} p_j(0) (T_n(t) P_n \mathrm{id}) \left(\frac{j^2}{n^2}\right). \end{aligned}$$

It suffices to examine the initial condition $p(0)$ with the mth component being 1 and the other components being 0. Accordingly, $y(0) = x(0) = (m/n)^2$ and $x(t) = \varphi(t, (m/n)^2) = (P_n T(t)\mathrm{id})((m/n)^2)$. Thus, by using (13) we obtain

$$|x(t) - y(t)| = \left| (P_n T(t)\mathrm{id}) \left(\frac{m^2}{n^2}\right) - (T_n(t) P_n \mathrm{id}) \left(\frac{m^2}{n^2}\right) \right| \leq \frac{C}{n}.$$

This proves Theorem 1.

References

[1] J. Goutsias, G. Jenkinson, Markovian dynamics on complex reaction networks. *Phys. Rep.*, 529(2013) 199–264

[2] J. R. Norris, *Markov Chains.* Cambridge University Press, 1997

[3] T. G. Kurtz, Extenstions of Trotter's operator semigroup approximation theorems. *J. Funct. Anal.*, 3(1969) 354–375

[4] T. G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.*, 7(1970) 49–58

[5] R. McVinish, P. K. Pollett, The deterministic limit of a stochastic logistic model with individual variation. *Math. Biosci.*, 241(2013) 109–114

[6] A. Bátkai, I. Z. Kiss, E. Sikolya, P. L. Simon, Differential equation approximations of stochastic network processes: an operator semigroup approach. *Netw. Heterog. Media.*, 7(2012) 43–58

[7] L. Bortolussi, Hybrid limits of continuous time Markov chains. *8th International Conference on Quantitative Evaluation of Systems*, Aachen, 2011, 3–12

[8] R. W. R. Darling, J. R. Norris, Differential equation approximations for Markov chains. *Probab. Surv.*, 5(2008) 37–79

[9] G. Kesidis, T. Konstantopoulos, P. Sousi, A stochastic epidemiological model and a deterministic limit for BitTorrent-like peer-to-peer file-sharing networks. In: *Network Control and Optimization*, LNCS 5425(2009) 26–36

[10] T. House, M. J. Keeling, Insights from unifying modern approximations to infections on networks. *J. R. Soc. Interface*, 8(2011) 67–73

[11] P.K. Pollett, Diffusion approximations for ecological models. *Proc. International Congress on Modelling and Simulation*, Modelling and Simulation Society of Australia and New Zealand, 2000, 843–848

[12] K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations.* Springer-Verlag, New York, 2000

[13] A. Bátkai, P. Csomós, G. Nickel, Operator splittings and spatial approximations for evolution equations. *J. Evol. Equ.*, 9(2009) 613–636