Edgeworth expansions in operator form

Zbigniew S. Szewczak*

21 November 2018

Abstract

An operator form of asymptotic expansions for Markov chains is established. Coefficients are given explicitly. Such expansions require a certain modification of the classical spectral method. They prove to be extremely useful within the context of large deviations.

Key words: Asymptotic expansions, large deviations, Perron-Frobenius theorem, transition probability function.

Mathematics Subject Classification (2000): 60F05, 60F10, 60J10, 47N30.

1 Introduction

Let $\{\xi_k\}_{k\in\mathbb{Z}_+}$ be a homogeneous Markov chain defined on a probability space $(\Omega,\mathcal{F},\mathrm{P})$. Denote by $\mathbb{S}$ and $\mathcal{S}$, respectively, the phase-space and its $\sigma-$algebra of measurable subsets. Further, denote by $P(x,A),\ x\in\mathbb{S},\ A\in\mathcal{S}$ the transition probability kernel of the chain. It means that for each $A\in\mathcal{S},\ P(x,A)$ is a non-negative measurable function on $\mathbb{S}$ while for each $x\in\mathbb{S},\ P(x,A)$ is a probability measure on $\mathcal{S}$. In what follows we assume that the chain is uniformly ergodic. So, there exists a stationary distribution denoted by π.

Consider the sequence of random variables $X_0=f(\xi_0),\dots,X_n=f(\xi_n)$ determined by a measurable function $f\colon\mathbb{S}\to\mathbb{R}$. In what follows we assume that

$$\sigma^2=\mathrm{E}_\pi[X_0^2]+2\sum_{n=1}^{\infty}\mathrm{E}_\pi[X_0X_n]>0. \tag{1.1}$$

There exists a huge literature concerning the limit theorems for successive sums $S_n=\sum_{i=1}^n X_i,\quad n=1,2,\dots$. For our purposes, it is enough to keep in mind only the works of S. Nagaev (1957) and (1961) and the monograph by Sirazhdinov and Formanov (1979). Despite the theory of limit theorems is well developed, some settings seem to be set aside. For example, in Szewczak (2005) it was shown that the Cramér method of conjugate distributions assumes a special form of the local limit theorem that was not considered before. The case studied in Szewczak (2005) concerns Markov chains with a finite number of states. It worth noting that the large deviation theorems, established there,

*Nicolaus Copernicus University, Faculty of Mathematics and Computer Science, ul. Chopina 12/18, 87-100 Toruń, Poland, e-mail: zssz@mat.uni.torun.pl

proved to be very useful in statistics of Markov chains (see A. Nagaev (2001) and (2002)). The mentioned form of the local limit theorem means the weak convergence of the measures

$$Q_x^{(n)}(A \times B) = \sigma\sqrt{2\pi n}\mathrm{P}_x^{(n)}\,[\,X_1 + \ldots + X_n \in A,\, \xi_n \in B\,], \qquad (1.2)$$

where

$$\mathrm{P}_x^{(n)}(\xi_1 \in A_1, \ldots, \xi_n \in A_n) = \int_{A_1} P(x, \mathrm{d}x_1) \int_{A_2} P(x_1, \mathrm{d}x_2) \ldots \int_{A_n} P(x_{n-1}, \mathrm{d}x_n),$$

$A_k \in \mathcal{S},\ k = 1, \ldots, n,\ B \in \mathcal{S},\ A \in \mathcal{B}(\mathbb{R})$ and $x \in \mathbb{S}$.

Define the linear operators

$$(\mathbf{K}_n g)(x) = \int P(x, \mathrm{d}x_1) \cdots \int P(x_{n-1}, \mathrm{d}x_n) g(x_n) K_n(x_1, \ldots, x_n); \quad g \in L^\infty(\mu), \qquad (1.3)$$

where K_n, $n = 1, 2, \ldots$, are measurable kernels, and $L^\infty(\mu)$ is the Banach space of measurable functions equipped with the essential supremum norm

$$\| g \| = \operatorname*{ess\,sup}_x |g(x)| = \inf\{a;\ \mu\{x;\ |g(x)| > a\} = 0\},$$

μ is the initial distribution, i.e. $\mu(A) = \mathrm{P}[\,\xi_0 \in A\,],\ A \in \mathcal{S}$.

Various probability measures of interest can be represented as a set indexed family of the operators (1.3). If one puts

$$K_{n,A}(x_1, \ldots, x_n) = \sigma\sqrt{2\pi n} I_A(f(x_1) + \cdots + f(x_n)), \quad g(x_n) = I_B(x_n),\ A \in \mathcal{B}(\mathbb{R}),$$

then (1.2) takes the form

$$Q_x^{(n)}(A \times B) = (\mathbf{K}_{n,A} I_B)(x). \qquad (1.4)$$

Similarly,

$$\mathrm{P}_x[\frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}} \in A, \xi_n \in B\,] = (\mathbf{K}_{n,A} g)(x) \qquad (1.5)$$

provided

$$K_{n,A}(x_1, \ldots, x_n) = I_A\left(\frac{f(x_1) + \cdots + f(x_n)}{\sigma\sqrt{n}}\right), \quad g(x_n) = I_B(x_n).$$

When x is fixed the weak convergence of the measures (1.4) (or (1.5)) means a form of the classical local limit (or central limit theorem). Naturally, we expect that the measures (1.4) weakly converge to $\lambda \times \pi$ while (1.5) converge to $\nu \times \pi$ where λ is the Lebesgue measure on $\mathcal{B}(\mathbb{R})$ and

$$\nu(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{u^2}{2}} \mathrm{d}u.$$

Such statement can be embedded into the following scheme of convergence.

Define

$$\|\mathbf{K}\|_+ = \sup_{\{g\geq 0;\, g\in L^\infty(\mu),\, \left| g \right| \leq 1\}} \left| \mathbf{K}g \right| . \tag{1.6}$$

Consider a family of sequences $\{\mathbf{K}_{n,A}\}$, $A \in \mathcal{A} \subset \mathcal{B}(\mathbb{R})$. We say that a sequence $\mathbf{K}_{n,A}$ is $L^\infty_{\mathcal{A}}(\mu)$-strongly convergent to $\mathbf{K}_A$ if

$$\sup_{A\in\mathcal{A}} \|\mathbf{K}_{n,A} - \mathbf{K}_A\|_+ \to 0 \qquad \text{as} \qquad n \to \infty. \tag{1.7}$$

Let $\mathcal{A} = \{z \in \mathbb{R} \,|\, (-\infty, z\sigma)\}$. If the sequence of operators is defined as in (1.5) then the limit operator in (1.7) has the form

$$(\mathbf{K}_A g)(x) = \psi(x)\nu(A) \int_{\mathbb{S}} g(s)\mu(\mathrm{d}s)$$

where $\psi(x) \equiv 1$. This fact is formally more general than e.g. Th. 2.2 in Nagaev (1957) though its proof does not require serious efforts. It is of much greater interest to establish the *operator form* of the asymptotic expansions for the sequence $\{\mathbf{K}_{n,z}\}$ determined by the kernels

$$K_{n,z}(x_1, \ldots, x_n) = I_{(-\infty,z)} \left(\frac{f(x_1) + \cdots + f(x_n)}{\sigma\sqrt{n}} \right), \ z \in \mathbb{R}.$$

Such asymptotic expansions is basic goal of the present paper. The paper is organized as follows. In Section 2 the main results are stated. In Section 3 a new estimate for the so-called characteristic operator in the neighborhood of zero is established (Cf. Lemma 1.6 in Nagaev (1961)). The proofs are given in Section 4.

2 The main results

In order to state the main results of the paper we have to introduce the indispensable notation. We are going to establish an asymptotic expansion of the form

$$\|\mathbf{K}_{n,z} - \sum_{m=0}^{k-2} n^{-\frac{m}{2}} \mathbf{A}_{m,z}\|_+ = o(n^{-\frac{k-2}{2}}), \tag{2.8}$$

where $\mathbf{A}_{m,z}$ are linear operators defined on $L^\infty(\mu)$, $m = 0, 1, \ldots,$ $z \in \mathbb{R}$. The operators $\mathbf{A}_{m,z}$ are expressed through the Hermite polynomials H_k and certain derivatives of the so-called characteristic operator

$$\hat{\mathbf{P}}(\theta)(g)(x) = \int e^{i\theta f(y)} g(y) P(x, \mathrm{d}y),$$

where $g \in L^\infty(\mu)$. More precisely, let $\lambda(\theta)$ be the principal eigenvalue of $\hat{\mathbf{P}}(\theta)$ and $\hat{\mathbf{P}}_1(\theta)$ be the projection on the eigenspace corresponding to $\lambda(\theta)$. Assume that $\hat{\mathbf{P}}(\theta)$ is k-times strongly differentiable at $\theta = 0$ and $\mathbf{P} = \hat{\mathbf{P}}(0)$ is L^∞-regular (or primitive), i.e. there exist $C > 0$ and $\gamma,\ 0 \leq |\gamma| < 1,$ such that

$$\left| \mathbf{P}^n g - \mathbf{\Pi} g \right| \leq C|\gamma|^n \left| g \right| , \qquad g \in L^\infty(\mu), \tag{2.9}$$

where

$$(\mathbf{\Pi}g)(x) = \psi(x)\int_{\mathbb{S}} g(s)\mu(\mathrm{d}s) = \psi(x)\mathrm{E}_\pi[g]$$

(Cf. Gudynas (2000)). Then $\hat{\mathbf{P}}_1(\theta)$ and $\ln\lambda(\theta)$ admit the following MacLaurin expansions:

$$\hat{\mathbf{P}}_1(\theta) = \sum_{m=0}^{k}\frac{(i\theta)^m}{m!}\hat{\mathbf{P}}_1^{(m)} + o(|\theta|^k), \qquad \text{and} \qquad \ln\lambda(\theta) = \sum_{m=0}^{k}\frac{(i\theta)^m}{m!}\gamma_m + o(|\theta|^k).$$

Here, the operators $\hat{\mathbf{P}}_1^{(m)}$ can be explicitly expressed in terms of $\mathbf{P}$ and $\mathbf{\Pi}$ (see Lemma 3). The coefficients γ_m, $m = 0, 1, \ldots$ are called *cumulants.* In what follows we assume $\gamma_1 = \mathrm{E}_\pi[f] = 0$ thus $\gamma_2 = \sigma^2$, where σ^2 is defined by (1.1) and $\gamma_3 = \mu_3$ is defined in Lemma 1.2 in Nagaev (1961). Let $\mathfrak{N}$ and $\mathfrak{n}$ denote the distribution function and the density function of the standard normal law. Introduce the operators defined on $L^\infty(\mu)$: $\mathbf{A}_{0,z} = \mathfrak{N}(z)\mathbf{\Pi}$, $\mathbf{A}_{\nu,z} = \sum_{j=0}^{\nu} a_j(z)\hat{\mathbf{P}}_1^{(j)}$, where

$$a_j(z) = -\mathfrak{n}(z)\sum_{(k_1,k_2,\ldots,k_{\nu-j})\in\mathcal{K}_{\nu-j}} a_{j,\nu-j}H_{\nu-1+2\sum\limits_{i=1}^{\nu-j}k_i}(z), \quad a_\nu = -\mathfrak{n}(z)H_{\nu-1},$$

$$a_{j,\nu-j} = \frac{1}{j!\sigma^j}\prod_{m=1}^{\nu-j}\frac{1}{k_m!}\Big(\frac{\gamma_{m+2}}{(m+2)!\sigma^{m+2}}\Big)^{k_m},$$

and $\mathcal{K}_m = \{(k_1, \ldots, k_m)\,;\, \sum_{i=1}^m ik_i = m,\, k_i \geq 0, i = 1, \ldots, m\}$. Thus, the operators $\mathbf{A}_{\nu,z}$ are well-defined provided $\hat{\mathbf{P}}(\theta)$ is k-times strongly differentiable at $\theta = 0$ and $\sigma > 0$.

Let $r(\theta)$ be the spectral radius of $\hat{\mathbf{P}}(\theta)$. It is well known that $r(\theta)$ inherits many principal properties of the characteristic functions. In order to establish asymptotic expansions (2.8) we have to assume that

$$r(\theta) < 1,\ \theta \neq 0, \qquad \text{and} \qquad \limsup_{|\theta|\to\infty} r(\theta) < 1. \tag{2.10}$$

The second inequality in (2.10) is analogous to the well-known Cramér condition (C). As to the first one it guarantees that the distributions of $\sum_{i=1}^n X_i$ for all sufficiently large n is non-lattice.

The operator form of asymptotic expansions implies such properties of the considered Markov chain as strong differentiability of $\hat{\mathbf{P}}(\theta)$, primitiveness and (2.10). Of course, one could simply assume that these properties take place. Another way is to give a simply verified condition that guarantees these properties. As such we take the following

Condition (Ψ):
there exist $\alpha > 0$ and $\beta < \infty$ such that for every Borel set A of a positive measure μ we have $\alpha\mu(A) \leq P(x, A) \leq \beta\mu(A)$ for μ-a.a. $x \in \mathbb{S}$.

This condition enables us to verify the required properties by the initial distribution μ. For example if Condition (Ψ) is fulfilled then (2.10) takes place provided $\mu_f = \mu\circ f^{-1}$ is non-lattice and

$$\limsup_{|\theta|\to\infty} |\widehat{\mu}_f(\theta)| < 1, \tag{2.11}$$

where $\widehat{\mu}_f = \int e^{i\theta f(y)}\mu(\mathrm{d}y)$. Moreover, if $\int |f(y)|^k \mu(\mathrm{d}y) < \infty$ then $\hat{\mathbf{P}}(\theta)$ is k-times strongly differentiable. It should be noted (see the proof of Lemma 3.1 in Jensen (1991)) that Condition (Ψ) implies $\sigma^2 = \gamma_2 > 0$.

Now we, are able to state the main results.

Theorem 1 *Let Condition (Ψ) is fulfilled. If $\int |f(x)|^k \mu(\mathrm{d}x) < \infty$, $k > 3$, and μ_f satisfies (2.11) then (2.8) holds.*

As in the case of asymptotic expansions for i.i.d. variables (see Gnedenko and Kolmogorov, 1954, §42, Th. 2) the following statement does not require the condition (2.11).

Theorem 2 *Let Condition (Ψ) is fulfilled. If $\int |f(x)|^3 \mu(\mathrm{d}x) < \infty$ and μ_f is non-lattice then (2.8) holds with $k = 3$.*

In order to clarify the specificity of the limit theorems given in the operator form consider two examples. First, let $\{\xi_k\}$ be a finite state Markov chain, i.e. $\mathbb{S} = \{1, \ldots, d\}$, $d \geq 3$. Denote by $\mathbf{P}$ the transition matrix. The entries of $\mathbf{P}^\nu, \nu \geq 0$, we denote by $p_{ij}^{(\nu)}, i, j \in \mathbb{S}$, $p_{ij}^{(0)} = \delta_{ij}$. For a real function f on $\mathbb{S}$ define the matrix $\mathbf{P}^{(1)}$ with the elements $f(j)p_{ij}$, $i, j \in \mathbb{S}$. The following statement is of independent interest.

Corollary 1 *Suppose that transition matrix $\mathbf{P}$ is strictly positive. If $f(\xi_0)$ is non-lattice and $\sum_{k=1}^d \pi_k f(k) = 0$ then uniformly in $z \in \mathbb{R}$ the matrix*

$$(\mathrm{P}[S_n < z\sigma\sqrt{n}\,;\, \xi_n = j \,|\, \xi_0 = i])_{i,j\in\mathbb{S}}$$

is approximated by the matrix

$$\mathfrak{N}(z)\mathbf{\Pi} + n^{-1/2}\mathfrak{n}(z)(\frac{\mu_3}{6\sigma^3}(1 - z^2)\mathbf{\Pi} - \frac{1}{\sigma}\sum_{\nu\geq 0} \mathbf{\Pi}\mathbf{P}^{(1)}(\mathbf{P}^\nu - \mathbf{\Pi}) + (\mathbf{P}^\nu - \mathbf{\Pi})\mathbf{P}^{(1)}\mathbf{\Pi}) \tag{2.12}$$

with an error $o(n^{-1/2})$. Here,

$$\mathbf{\Pi} = \begin{pmatrix} \pi_1 & \pi_2 & \ldots & \pi_d \\ \ldots & \ldots & \ldots & \ldots \\ \pi_1 & \pi_2 & \ldots & \pi_d \end{pmatrix}.$$

Another particular case of independent interest is covered by the following statement.

Corollary 2 *Let $\mathbb{S} = [0,\ 1]$. Suppose that the transition density $p(x, y)$ is such that $0 < p_- \leq p(x, y) \leq p_+ < \infty$. If $f(\xi_0)$ is non-lattice and $\int f(u)\pi(\mathrm{d}u) =$*

0 *then the linear operator (2.12)* $L^\infty_{\mathcal{A}}$*-strongly approximates the operator* $g \mapsto \mathrm{E}_x[I_{[S_n<z\sigma\sqrt{n}]}g(\xi_n)]$ *with an error* $o(n^{-1/2})$*. Here,*

$$(\mathbf{\Pi}g)(x) = \int_{\mathbb{S}} g(s)\mu(\mathrm{d}s)\psi(x).$$

Note that the classical scalar form of the presented statement is:

$$\mathrm{P}_\pi[S_n < z\sigma\sqrt{n}] - \mathfrak{N}(z) = n^{-1/2}\mathfrak{n}(z)\frac{\mu_3}{6\sigma^3}(1-z^2) + o(n^{-1/2}) \tag{2.13}$$

(see e.g. Th. 2 in Nagaev (1961)). The corollaries show that the operator form of asymptotic expansions is much more sensitive to the initial conditions than the scalar one. It should be emphasized that the spectral method suggested by S. Nagaev (see e.g. Nagaev, 1957) remains efficient under this new setting though requires some modification. Furthermore, the cumbersome calculations, that are typical for asymptotic expansions, can be implemented using the package *Maple.* This power software proved to be very efficient for such purposes.

3 Characteristic operator

Given $m \in \mathbb{N}$ let us define operator $\mathbf{P}^{(m)}g = \mathbf{P}f^m g$. The following lemma is an extension of the well-known result due to S. Nagaev (Cf. Nagaev, 1961, pp. 71–75).

Lemma 1 *Suppose that (2.9) holds and* $\mathbf{P}^{(1)}$ *is a bounded endomorphism. Then there exists* $\xi = \xi(C, |\gamma|, \|\mathbf{P}^{(1)}\|)$ *such that for* $|\theta| < \xi$,

$$\hat{\mathbf{P}}^n(\theta) = \lambda^n(\theta)\hat{\mathbf{P}}_1(\theta) + \hat{\mathbf{Q}}_n(\theta) + (\mathbf{P}^n - \mathbf{\Pi}) \tag{3.14}$$

and $|\lambda(\theta) - 1| < \delta$*, where* $\|\hat{\mathbf{P}}_1(\theta) - \mathbf{\Pi}\| = O(|\ \theta\ |)$, $\|\hat{\mathbf{Q}}_n(\theta)\| = O(\kappa^n\ |\ \theta\ |)$, $\kappa = \frac{1}{3} + \frac{2}{3}|\gamma|$, $\delta = \frac{1}{3} - \frac{1}{3}|\gamma|$.

Proof of Lemma 1

Write $\Gamma_0 = \{|\zeta| = \kappa\}$, $\Gamma_1 = \{|\zeta - 1| = \delta\}$ and $D = \{|\zeta| \geq \kappa\} \cap \{|\zeta - 1| \geq \delta\}$, where $\zeta \in \mathbb{C}$. Denote by $\hat{\mathbf{R}}(\zeta,\theta)$ the resolvent of $\hat{\mathbf{P}}(\theta)$ and set $\mathbf{R}(\zeta) = \hat{\mathbf{R}}(\zeta, 0)$. Let $\xi = \frac{1}{2\|\mathbf{P}^{(1)}\|}\left(\frac{1-|\gamma|}{3(3+C)}\right)^2$. Consequently for $|\ \theta\ | < \xi$, $\zeta \in D$ (see §1 in Nagaev (1961)) we may define the projections

$$\hat{\mathbf{P}}_1(\theta) = \frac{1}{2\pi i}\oint_{\Gamma_1}\hat{\mathbf{R}}(\zeta,\theta)\mathrm{d}\zeta, \qquad \hat{\mathbf{P}}_2(\theta) = \frac{1}{2\pi i}\oint_{\Gamma_0}\hat{\mathbf{R}}(\zeta,\theta)\mathrm{d}\zeta. \tag{3.15}$$

Thus (3.14) holds with $\hat{\mathbf{Q}}_n(\theta) = \hat{\mathbf{P}}^n(\theta)\hat{\mathbf{P}}_2(\theta) - (\mathbf{P}^n - \mathbf{\Pi})$. We see at once that

$$\hat{\mathbf{P}}(\theta)\hat{\mathbf{R}}(\zeta,\theta) = -\mathbf{I} + \zeta\hat{\mathbf{R}}(\zeta,\theta)$$

therefore,

$$\hat{\mathbf{P}}^n(\theta)\hat{\mathbf{R}}(\zeta,\theta) = -\sum_{k=1}^{n}(\hat{\mathbf{P}}(\theta))^{n-k}\zeta^{k-1} + \zeta^n\hat{\mathbf{R}}(\zeta,\theta).$$

So it easily seen that

$$\begin{aligned}
\|\hat{\mathbf{Q}}_n(\theta)\| &= \|\frac{1}{2\pi i}\oint_{\Gamma_0} \zeta^n(\hat{\mathbf{R}}(\zeta,\theta)-\mathbf{R}(\zeta))\mathrm{d}\zeta\| \\
&\leq \frac{1}{2\pi}\int_0^{2\pi} \kappa^n \frac{2(3(3+C))^3\|\hat{\mathbf{P}}(\theta)-\mathbf{P}\|}{(1-|\gamma|)^2(6(3+C)-1+|\gamma|)}\kappa\mathrm{d}\phi = O(\kappa^n \mid \theta \mid).
\end{aligned}$$

Similarly we have $\|\hat{\mathbf{P}}_1(\theta)-\mathbf{\Pi}\| = O(\mid \theta \mid)$. The proof is completed. □

The following lemma deals with the existence of the "operator" moments.

Lemma 2 *If*

$$\lim_{L\to\infty} \Big\| \int_{|f(y)|>L} |f(y)|^k P(x,\mathrm{d}y) \Big\| = 0 \tag{3.16}$$

then

$$\hat{\mathbf{P}}(\theta) = \sum_{m=0}^{k} \frac{(i\theta)^m}{m!}\mathbf{P}^{(m)} + o(|\theta|^k), \tag{3.17}$$

where $\mathbf{P}^{(m)}$ *are bounded for* $0 \leq m \leq k$.

Proof of Lemma 2

Indeed, we have

$$\begin{aligned}
h^{-1}(i^{(k-1)}\hat{\mathbf{P}}^{(k-1)}(\theta+h)g - i^{(k-1)}\hat{\mathbf{P}}^{(k-1)}(\theta)g) - i^k \int e^{i\theta y} g(y) f^k(y) P(\,\cdot\,,\mathrm{d}y) \\
= i^k \int e^{i\theta f(y)} g(y) f^k(y) \int_0^1 (e^{ihsf(y)}-1)\mathrm{d}s P(\,\cdot\,,\mathrm{d}y).
\end{aligned}$$

Now, choose L be sufficiently large positive number. Since,

$$\begin{aligned}
&\inf\{K\,;\,\mu\{x\,;\,|\int (if(y))^k \int_0^1 (e^{ihsf(y)}-1)\mathrm{d}s P(x,\mathrm{d}y)| > K\} = 0\} \\
&\quad \leq \inf\{K\,;\,\mu\{x\,;\,|\int_{|f(y)|\leq L} (if(y))^k \int_0^1 (e^{ihsf(y)}-1)\mathrm{d}s P(x,\mathrm{d}y)| > K\} = 0\} \\
&\quad + \inf\{K\,;\,\mu\{x\,;\,|\int_{|f(y)|> L} (if(y))^k \int_0^1 (e^{ihsf(y)}-1)\mathrm{d}s P(x,\mathrm{d}y)| > K\} = 0\} \\
&\quad \leq \frac{1}{2}L^{k+1}|h| + 2\Big\|\int_{|f(y)|>L} |f(y)|^k P(x,\mathrm{d}y)\Big\|
\end{aligned}$$

the lemma follows by the Taylor formula. □

The next lemma presents a series expansion for characteristic projector.

Lemma 3 *If a primitive operator* $\mathbf{P}$ *satisfies (3.17) then*

$$\hat{\mathbf{P}}_1(\theta) = \sum_{m=0}^{k} \frac{(i\theta)^m}{m!}\hat{\mathbf{P}}_1^{(m)} + o(|\theta|^k).$$

Proof of Lemma 3

Put, for short $\mathbf{E}(\zeta) = \sum_{n\geq 0}(\mathbf{P}^n - \mathbf{\Pi})\zeta^{-n-1}$, and $\mathbf{E} = \mathbf{E}(1)$. In view of (1.10) in Nagaev (1957) and (3.17) we obtain for $|\theta| < \xi$

$$\hat{\mathbf{R}}(\zeta,\theta) = \mathbf{R}(\zeta) + \sum_{n\geq 1}\mathbf{R}(\zeta)(\sum_{m=1}^{k}\mathbf{P}^{(m)}\mathbf{R}(\zeta)\frac{(i\theta)^m}{m!})^n + o(|\theta|^k).$$

Hence taking in the above coefficient at $i\theta$ and using (3.15) we get for $k = 1$

$$\begin{aligned}\hat{\mathbf{P}}_1^{(1)} &= \frac{1}{2\pi i}\oint_{\Gamma_1}(\frac{\mathbf{\Pi}}{\zeta-1} + \mathbf{E}(\zeta))\mathbf{P}^{(1)}(\frac{\mathbf{\Pi}}{\zeta-1} + \mathbf{E}(\zeta))\mathrm{d}\zeta \\ &= \mathbf{\Pi}\mathbf{P}^{(1)}\mathbf{\Pi}\frac{1}{2\pi i}\oint_{\Gamma_1}\frac{1}{(\zeta-1)^2}\mathrm{d}\zeta + \frac{1}{2\pi i}\oint_{\Gamma_1}\mathbf{\Pi}\mathbf{P}^{(1)}\frac{\mathbf{E}(\zeta)}{\zeta-1}\mathrm{d}\zeta \\ &\quad + \frac{1}{2\pi i}\oint_{\Gamma_1}\frac{\mathbf{E}(\zeta)}{\zeta-1}\mathbf{P}^{(1)}\mathbf{\Pi}\mathrm{d}\zeta = \mathbf{\Pi}\mathbf{P}^{(1)}\mathbf{E} + \mathbf{E}\mathbf{P}^{(1)}\mathbf{\Pi}\end{aligned}$$

by Cauchy's integral formula. For $1 < m \leq k$ arguments are similar. We have to replace every $\mathbf{R}(\zeta)$ by $\frac{\mathbf{\Pi}}{\zeta-1} + \mathbf{E}(\zeta)$ in

$$\hat{\mathbf{P}}_1^{(m)} = \frac{m!}{2\pi i}\oint_{\Gamma_1}\sum_{\nu_1+\nu_2+\ldots+\nu_l=m}\mathbf{R}(\zeta)\frac{\mathbf{P}^{(\nu_1)}}{\nu_1!}\mathbf{R}(\zeta)\frac{\mathbf{P}^{(\nu_2)}}{\nu_2!}\cdots\mathbf{R}(\zeta)\frac{\mathbf{P}^{(\nu_l)}}{\nu_l!}\mathbf{R}(\zeta)\mathrm{d}\zeta,\ \nu_k \geq 1.$$

□

Now, we are in a position to represent the principal eigenvalue of the characteristic operator in a power series.

Lemma 4 *If a primitive operator* $\mathbf{P}$ *satisfies (3.17) then*

$$\lambda(\theta) = 1 + \frac{(i\theta)}{1!}\mu_1 + \frac{(i\theta)^2}{2!}\mu_2 + \frac{(i\theta)^3}{3!}\mu_3 + \cdots + \frac{(i\theta)^k}{k!}\mu_k + o(|\theta|^k).$$

Proof of Lemma 4

It follows from (3.14) that

$$\pi\hat{\mathbf{P}}(\theta)\hat{\mathbf{P}}_1(\theta)\psi = \lambda(\theta)\pi\hat{\mathbf{P}}_1(\theta)\psi. \tag{3.18}$$

Denote $\hat{\lambda}^{(k)} = \pi\hat{\mathbf{P}}_1^{(k)}\psi$. By virtue of (3.17) and Lemma 3

$$\lambda(\theta)\sum_{\nu=0}^{k}\hat{\lambda}^{(\nu)}\frac{(i\theta)^\nu}{\nu!} = \sum_{m=0}^{k}\Big(\sum_{\nu=0}^{m}\binom{m}{\nu}\pi\mathbf{P}^{(\nu)}\hat{\mathbf{P}}_1^{(m-\nu)}\psi\frac{(i\theta)^m}{m!}\Big) + o(|\theta|^k).$$

Since $\lambda^{(0)} = \hat{\lambda}^{(0)} = 1$, $\lambda^{(1)} = \hat{\lambda}^{(1)} = 0$, and $\lambda^{(k)}$ exists so by the Leibniz formula

$$\lambda^{(m)} = \mu_m = \sum_{\nu=1}^{m}\binom{m}{\nu}\pi\mathbf{P}^{(\nu)}\hat{\mathbf{P}}_1^{(m-\nu)}\psi - \sum_{\nu=2}^{m-2}\binom{m}{\nu}\lambda^{(\nu)}\hat{\lambda}^{(m-\nu)}. \tag{3.19}$$

By (3.19) $\gamma_2 = \lambda^{(2)}$, for $m > 2$ also use the equation (1.13) in Petrov (1996). □

The following theorem is the main result of the present Section.

Theorem 3 *If a primitive operator* $\mathbf{P}$ *satisfies (3.17),* $k \geq 3$ *and* $\sigma^2 > 0$ *then there exists* $\eta_k > 0$ *such that for* $T_n = \eta_k \sigma \sqrt{n}$ *and* $|\theta| \leq T_n$ *we have*

$$\|\hat{\mathbf{P}}^n(\frac{\theta}{\sigma\sqrt{n}}) - e^{-\frac{\theta^2}{2}}\Big(\sum_{m=0}^{k-2}\sum_{j=0}^{m}\frac{(i\theta)^j}{n^{\frac{m}{2}}j!\sigma^j}\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_1^{(j)}\Big) - (\mathbf{P}^n - \mathbf{\Pi})\| \tag{3.20}$$

$$\leq \frac{o(1)}{n^{\frac{k-2}{2}}}(|\theta|^{k-2} + |\theta|^{k-1} + |\theta|^k + |\theta|^{3(k-2)})e^{-\frac{\theta^2}{4}} + O(\frac{|\theta|}{\sqrt{n}}\kappa^n),$$

where

$$\mathfrak{P}_\nu(i\theta) = \sum_{(k_1,k_2,\ldots,k_\nu)\in\mathcal{K}_\nu}\prod_{m=1}^{\nu}\frac{1}{k_m!}\left(\frac{\gamma_{m+2}(i\theta)^{m+2}}{(m+2)!\sigma^{m+2}}\right)^{k_m}.$$

Proof of Theorem 3

Let $0 < \eta_3 \leq \xi$ be such that $\sup_{|\theta|\leq\eta_3}|\lambda^{(3)}(\theta) - \mu_3| \leq \sigma^3$. Put, for short $T_n = \min\{\frac{\sigma^2}{5(\frac{3}{2}|\mu_3|+\sigma^3)}, \eta_3\}\sigma\sqrt{n}$. By Taylor's formula for $|\theta| \leq T_n$ we have

$$\begin{aligned}|\lambda(\frac{\theta}{\sigma\sqrt{n}})| &\geq 1 - \frac{\theta^2}{2n} - \frac{|\theta|^3(|\mu_3| + \sigma^3 + \frac{1}{2}|\mu_3|)}{6n^{\frac{3}{2}}\sigma^3} \geq 1 - \frac{T_n^2}{2n} - \frac{T_n^3(\frac{3}{2}|\mu_3| + \sigma^3)}{6n^{\frac{3}{2}}\sigma^3}\\ &\geq 1 - \frac{\sigma^6}{50(\frac{3}{2}|\mu_3| + \sigma^3)^2} - \frac{\sigma^6}{6\cdot 125(\frac{3}{2}|\mu_3| + \sigma^3)^2} > 1 - \frac{2}{50} = \frac{24}{25}\cdot\end{aligned}$$

Hence for $|\theta| \leq T_n$ by Taylor's formula and Lemma 4

$$\begin{aligned}n\ln\lambda(\frac{\theta}{\sigma\sqrt{n}}) &= -\frac{\theta^2}{2} + \frac{(i\theta)^3\mu_3}{6\sqrt{n}\sigma^3} + \frac{(i\theta)^4\mu_4}{24n\sigma^4} - \frac{(i\theta)^4}{8n} + \cdots\\ &\quad + \frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}\gamma_k + \frac{(i\theta)^k}{(k-1)!n^{\frac{k-2}{2}}\sigma^k}\int_0^1(1-x)^{k-1}W_k(\frac{x\theta}{\sigma\sqrt{n}})\mathrm{d}x,\end{aligned} \tag{3.21}$$

where $W_k(x) = \frac{\partial^k}{\partial y^k}\ln\lambda(y)|_{y=x} - \gamma_k$. Further, it is evident that we can insert $\eta_k \leq \eta_3$ in $T_n = \min\{\frac{\sigma^2}{5(\frac{3}{2}|\mu_3|+\sigma^3)}, \eta_k\}\sigma\sqrt{n}$, such that for $|\theta| \leq T_n$ we have

$$6\sigma^2(\frac{|\theta|^2}{24n\sigma^4}|\gamma_4| + \cdots + \frac{|\theta|^{k-2}}{k!n^{\frac{k-2}{2}}\sigma^k}(|\gamma_k| + c_k)) < 7,$$

where $c_k = \sup_{x\in[0,1]}|W_k(x)|$. Since $W_k(\frac{x\theta}{\sigma\sqrt{n}}) \to_n 0$, so by the Lebesgue dominated convergence theorem we get $\int_0^1(1-x)^{k-1}W_k(\frac{x\theta}{\sigma\sqrt{n}})\mathrm{d}x = o(1)$. By virtue of (3.14), Lemma 1 and (3.21) we obtain

$$\|\hat{\mathbf{P}}^n(\frac{\theta}{\sigma\sqrt{n}}) - e^{-\frac{\theta^2}{2} + \frac{(i\theta)^3\mu_3}{6\sqrt{n}\sigma^3} + \frac{(i\theta)^4\mu_4}{24n\sigma^4} - \frac{(i\theta)^4}{8n} + \cdots + \frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}\gamma_k}\hat{\mathbf{P}}_1(\frac{\theta}{\sigma\sqrt{n}}) - (\mathbf{P}^n - \mathbf{\Pi})\|$$

$$\leq e^{-\frac{\theta^2}{2} + \cdots + \frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}\gamma_k}\left|\exp\{\frac{(i\theta)^k\int_0^1(1-x)^{k-1}W_k(\frac{x\theta}{\sigma\sqrt{n}})\mathrm{d}x}{(k-1)!n^{\frac{k-2}{2}}\sigma^k}\} - 1\right|O(1)$$

$$+ O(\frac{|\theta|}{\sigma\sqrt{n}}\kappa^n).$$

By (3.21) and the inequality $|e^x - 1| \leq |x|e^{|x|}$, we find that for $|\theta| \leq T_n$ we have

$$|\exp\{\frac{(i\theta)^k}{(k-1)!n^{\frac{k-2}{2}}\sigma^k}\int_0^1 (1-x)^{k-1}W_k(\frac{x\theta}{\sigma\sqrt{n}})\mathrm{d}x\}-1| \leq \frac{o(1)|\theta|^k}{n^{\frac{k-2}{2}}}\exp\{\frac{c_k|\theta|^k}{k!n^{\frac{k-2}{2}}\sigma^k}\}.$$

Hence,

$$\begin{aligned}
&\|\hat{\mathbf{P}}^n(\frac{\theta}{\sigma\sqrt{n}}) - e^{-\frac{\theta^2}{2}+\frac{(i\theta)^3\mu_3}{6\sqrt{n}\sigma^3}+\frac{(i\theta)^4\mu_4}{24n\sigma^4}-\frac{(i\theta)^4}{8n}+\cdots+\frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}\gamma_k}\hat{\mathbf{P}}_1(\frac{\theta}{\sigma\sqrt{n}}) - (\mathbf{P}^n - \mathbf{\Pi})\| \\
&\quad \leq \exp\{-\frac{\theta^2}{2} + \cdots + \frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}(|\gamma_k| + c_k)\}\frac{|\theta|^k}{n^{\frac{k-2}{2}}}o(1) + O(\frac{|\theta|}{\sigma\sqrt{n}}\kappa^n) \\
&\quad \leq \exp\{-\frac{\theta^2}{2} + \frac{\theta^2}{2}(\frac{1}{15}\frac{\frac{3}{2}|\mu_3|}{\frac{3}{2}|\mu_3| + \sigma^3} + \frac{7}{15}\frac{\sigma^3}{\frac{3}{2}|\mu_3| + \sigma^3})\}\frac{|\theta|^k}{n^{\frac{k-2}{2}}}o(1) + O(\frac{|\theta|}{\sigma\sqrt{n}}\kappa^n) \\
&\quad \leq o(1)\frac{|\theta|^k}{n^{\frac{k-2}{2}}}\exp\{-\frac{\theta^2}{4}\} + O(\frac{|\theta|}{\sigma\sqrt{n}}\kappa^n).
\end{aligned}$$

Thus expanding $\exp\{\frac{(i\theta)^3\mu_3}{6\sqrt{n}\sigma^3}+\cdots+\frac{(i\theta)^k}{k!n^{\frac{k-2}{2}}\sigma^k}\gamma_k\}$ and using Lemma 3 and Taylor's formula for $\hat{\mathbf{P}}_1(\frac{\theta}{\sigma\sqrt{n}})$ we obtain (3.20). $\square$

The following lemma provides an estimate for the iterates of characteristic operator (for the proof see Lemma 1.5 in Nagaev (1961)).

Lemma 5 *Let Condition* (Ψ) *is fulfilled. Then for* $n \geq 1$ *and* $\| g \| \leq 1$

$$\| \hat{\mathbf{P}}^n(\theta)g \| \leq \left(\sqrt{1 - \frac{\alpha^4}{2\beta}(1 - |\widehat{\mu}_f(\theta)|^2)}\right)^{n-1}.$$

4 Proofs

Proof of Theorems 1 and 2

By virtue of Condition (Ψ) and §1 in Nagaev (1957) $\mathbf{P}$ is primitive in $L^\infty(\mu)$ (alternatively one can use Proposition 3.13 in Wu (2000)). Moreover, it follows also that

$$\lim_{L\to\infty} \| \int_{|f(y)|>L} |f(y)|^k P(x, \mathrm{d}y) \| \leq \beta \lim_{L\to\infty} \int_{|f(y)|>L} |f(y)|^k \mu(\mathrm{d}y) = 0$$

so that (3.16) holds. Write $F_{gn}(z) = F_{g(\cdot),n}(z) = (\mathbf{K}_{n,z}g)(\cdot)$, and $G_{gn}(z) = \sum_{m=0}^{k-2} n^{-\frac{m}{2}}\mathbf{A}_{m,z}g$. Let $K_{gn}(z)$ be the distribution function that assigns the mass $(\mathbf{P}^n - \mathbf{\Pi})g(\cdot)$ at 0. Put,

$$H_{gn}(z) = G_{gn}(z) + K_{gn}(z),\ \hat{H}_{gn}(\theta) = \int e^{i\theta x}\mathrm{d}H_{gn}(x),\ \hat{F}_{gn}(\theta) = \int e^{i\theta x}\mathrm{d}F_{gn}(x).$$

Note that $\hat{F}_{gn}(\theta) = \hat{\mathbf{P}}^n(\frac{\theta}{\sigma\sqrt{n}})(g)$ and

$$\hat{G}_{gn}(\theta) = \int e^{i\theta x}\mathrm{d}G_{gn}(x) = e^{-\frac{\theta^2}{2}}\Big(\sum_{m=0}^{k-2}\frac{1}{(\sqrt{n})^m}\sum_{j=0}^{m}\frac{1}{j!}(\frac{i\theta}{\sigma})^j\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_1^{(j)}g(\cdot)\Big).$$

Because of

$$\Vert H_{gn}(z+y)-H_{gn}(z)\Vert \le \Vert G_{gn}(z+y)-G_{gn}(z)\Vert + \Vert (\mathbf{P}^n-\mathbf{\Pi})g\Vert,$$

$$G_{gn}(z+y)-G_{gn}(z) = y\frac{\partial}{\partial z}G_{gn}(z)+\mathrm{sgn}(y)\int\limits_{\frac{y-|y|}{2}}^{\frac{y+|y|}{2}}(\frac{\partial}{\partial u}G_{gn}(z+u)-\frac{\partial}{\partial z}G_{gn}(z))\mathrm{d}u$$

thus in view of Th. 5.3 on pp. 146–147 in Petrov (1996) and (2.9) we have

$$\begin{aligned}\Vert F_{gn}(z)-G_{gn}(z)\Vert &\le \Vert F_{gn}(z)-G_{gn}(z)-K_{gn}(z)\Vert + C|\gamma|^n\Vert g\Vert \\ &\le \frac{1}{\pi}\int_{|\theta|\le T}\Vert \hat{F}_{gn}(\theta)-\hat{H}_{gn}(\theta)\Vert\frac{\mathrm{d}\theta}{|\theta|} \qquad (4.22)\\ &\quad +\frac{3c^2(\frac{1}{\pi})}{\pi T}\sup_z\Vert\frac{\partial}{\partial z}G_{gn}(z)\Vert + C|\gamma|^n(1+\frac{2c(\frac{1}{\pi})}{\pi})\Vert g\Vert.\end{aligned}$$

Now, since $\sup_z\Vert\frac{\partial}{\partial z}G_{gn}(z)\Vert$ is bounded and $|\gamma|<1$ whence by (4.22) for $T=n^k, k\ge 4$, we get

$$\Vert F_{gn}(z)-G_{gn}(z)\Vert \le \frac{1}{\pi}\int_{|\theta|\le n^k}\Vert \hat{F}_{gn}(\theta)-\hat{H}_{gn}(\theta)\Vert\frac{\mathrm{d}\theta}{|\theta|}+o(\frac{\Vert g\Vert}{n^{\frac{k-2}{2}}})\cdot \qquad (4.23)$$

By virtue of Th. 3

$$\begin{aligned}&\int_{|\theta|\le T_n}\Vert \hat{F}_{gn}(\theta)-\hat{H}_{gn}(\theta)\Vert\frac{\mathrm{d}\theta}{|\theta|} \qquad (4.24)\\ &\quad\le \frac{o(\Vert g\Vert)}{n^{\frac{k-2}{2}}}\int_{|\theta|\le T_n}(|\theta|^{k-3}+|\theta|^{k-1}+|\theta|^k+|\theta|^{3k-7})e^{-\frac{\theta^2}{4}}\mathrm{d}\theta+\frac{T_n^2}{\sqrt{n}}O(\kappa^n)\cdot\end{aligned}$$

This established, we have to show that

$$\int_{T_n<|\theta|\le n^k}\frac{\Vert \hat{F}_{gn}(\theta)-\hat{H}_{gn}(\theta)\Vert}{|\theta|}\mathrm{d}\theta \le o(\frac{\Vert g\Vert}{n^{\frac{k-2}{2}}})\cdot \qquad (4.25)$$

For this observe that

$$\int_{T_n<|\theta|\le n^k}\frac{\Vert \hat{G}_{gn}(\theta)\Vert}{|\theta|}\mathrm{d}\theta \le 2\int_{T_n}^{\infty}e^{-\frac{\theta^2}{2}}\left\Vert\sum_{m=0}^{k-2}\sum_{j=0}^{m}\frac{(i\theta)^j}{n^{\frac{m}{2}}j!\sigma^j}\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_1^{(j)}\right\Vert\Vert g\Vert\frac{\mathrm{d}\theta}{|\theta|}$$

and that by (2.9)

$$\int_{T_n<|\theta|\le n^k}\Vert \hat{H}_{gn}(\theta)-\hat{G}_{gn}(\theta)\Vert\frac{\mathrm{d}\theta}{|\theta|}\le 2C_k|\gamma|^n\Vert g\Vert\ln n = o(\frac{\Vert g\Vert}{n^{\frac{k-2}{2}}})\cdot$$

Further, by Lemma 5 and (2.11) there exists θ_0 such that for any $\tau>\theta_0$

$$n^{\frac{k-2}{2}}\int\limits_{T_n<|\theta|\le n^k}\Vert \hat{F}_{gn}(\theta)\Vert\frac{\mathrm{d}\theta}{|\theta|}\le n^{\frac{k-2}{2}}\int\limits_{\tau\le|\theta|\le n^k}\Vert(\hat{\mathbf{P}}(\theta))^n g\Vert\frac{\mathrm{d}\theta}{|\theta|}\le C_k n^{\frac{k-2}{2}}e^{-cn}\Vert g\Vert\ln n$$

which with the latter inequalities proves (4.25). Consequently, the substitution of (4.24) and (4.25) into (4.23) yields (2.8). For the case $k = 3$ set $T = T_n r_n$ and choose a sequence $r_n \to \infty$ such that we have

$$\int_{T_n<|\theta|\le T_n r_n} \left| \hat{F}_{gn}(\theta) \right| \frac{\mathrm{d}\theta}{|\theta|} = \int_{T_n<|\theta|\sigma\sqrt{n}\le T_n r_n} \left| (\hat{\mathbf{P}}(\theta))^n g \right| \frac{\mathrm{d}\theta}{|\theta|} = \left| g \right| o(n^{-1/2}).$$

This completes the proof. □

Acknowledgment. The author thanks A. Nagaev for his comments concerning this exposition.

References

Gnedenko, B.V., Kolmogorov, A.N., 1954. Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading, Mass.

Gudynas, P., 2000. Refinements of the Central Limit Theorem for Homogeneous Markov Chains, in: Yu. V. Prokhorov and V. Statulevičius, eds. Limit Theorems of Probability Theory. Springer, Berlin, pp. 165–183.

Jensen, J.L., 1991. Saddlepoint expansions for sums of Markov dependent variables on a continuous state space. Probab. Theory Related Fields 89, 181–199.

Nagaev, A.V., 2001. An asymptotic formula for the Bayes risk in discriminating between two Markov chains. J. Appl. Probab. 38A, 131–141.

Nagaev, A.V., 2002. An asymptotic formula for the Neyman–Pearson risk in dicriminating between two Markov chains. J. Math. Sci. 111, 3592–3600.

Nagaev, S.V., 1957. Some limit theorems for stationary Markov chains. Teor. Veroyatnost. i Primenen. 2, 389–416.

Nagaev, S.V., 1961. More exact statements of limit theorems for homogeneous Markov chains. Teor. Veroyatnost. i Primenen. 6, 67–86.

Petrov, V.V., 1996. Limit Theorems of Probability Theory. Sequences of Indenpendent Random Variables. Oxford Studies in Probability 4, Oxford.

Sirazhdinov, S.H., Formanov, S.K., 1979. Limit Theorems for Sums of Random Vectors Connected in a Markov Chain. FAN, Tashkent.

Szewczak, Z.S. 2005. A remark on large deviation theorem for Markov chain with finite number of states. Teor. Veroyatnost. i Primenen. 50 3, 612–622.

Wu, L.M., 2000. Uniformly Integrable Operators and Large Deviations for Markov Processes. J. Funct. Anal. 172, 301–376.