SIAM J. APPL. MATH.
Vol. 58, No. 6, pp. 1862–1877, December 1998

010

ASYMPTOTIC THEORY OF LARGE DEVIATIONS FOR MARKOV CHAINS*

G. LERMAN† AND Z. SCHUSS‡

Abstract. A formal asymptotic expansion is constructed for the joint probability density function (pdf) of a stationary ergodic Markov chain, X_n, and its averages, Y_n. Since the pair (X_n, Y_n) is Markovian, the joint pdf satisfies a forward Kolmogorov equation whose solution is expanded asymptotically for large n. An algorithm is proposed for the calculation of the full asymptotic series, but only the three leading terms are found explicitly. It is found that for small values of the average, the asymptotic expansion coincides with the appropriate version of the central limit theorem. The ideas and methods are generalized to a large class of averages and to vector valued Markov chains.

Key words. large deviations, Markov chains, asymptotics

AMS subject classifications. 60J10, 60F10, 45L05

PII. S003613999630051

1. Introduction.

We consider an ergodic stationary Markov process, $\{X_n\}$, with zero mean, and its sample averages

$$Y_n \equiv \frac{1}{n}\sum_{j=1}^{n} X_j\,. \tag{1.1}$$

The problem of large deviations is to determine the uniform asymptotic behavior of the probability distribution function (PDF) of Y_n and the joint PDF of (X_n, Y_n) for large n.

The problem of large deviations arises in many problems in statistical decision theory, communications, information theory, and statistical physics [1, 2]. Specific applications of the present theory to the Neyman–Pearson maximum-likelihood test and to block coding will be presented in a forthcoming paper.

The large deviations principle [2]–[4] characterizes the exponential decay of the tails of these PDFs by a rate function. The rate function controls the logarithmic asymptotics of the PDFs but does not provide a full asymptotic expansion.

In the case of independent identically distributed (i.i.d.) $\{X_n\}$, a full asymptotic expansion was formulated in [5] under the assumption that the moment generating function exists near zero. The expansion for the complementary PDF of Y_n was given in the form

$$1 - F_n(y) \sim \frac{1}{\sqrt{2\pi n}} e^{-n\psi(y)}\left[K_0(y) + \frac{1}{n}K_1(y) + \frac{1}{n^2}K_2(y) + \cdots\right]. \tag{1.2}$$

The exponential rate function, $\psi(y)$, was determined as the solution of a differential equation. The explicit expressions for the higher functions K_j were later calculated

* Received by the editors March 15, 1996; accepted for publication (in revised form) April 23, 1997; published electronically August 18, 1998.
http://www.siam.org/journals/siap/58-6/30055.html

† Department of Applied Mathematics, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978, Israel (gilad.lerman@yale.edu). This work is based in part on this author's M.Sc. dissertation, Department of Applied Mathematics, Tel-Aviv University, 1995.

‡ Department of Applied Mathematics, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978, Israel (schuss@math.tau.ac.il).

in [6], [7], [8], and [9] by different methods. In this case, the joint probability density function (pdf) of $\{X_n, Y_n\}$ can be obtained from the pdfs $\rho(x)$ and $p_n(y)$ of X_n and Y_n, respectively, as

$$p_n(x,y) = \frac{\partial}{\partial y}\Pr(Y_n \le y | X_n = x)\rho(x) = \left(\frac{n}{n-1}\right)\rho(x) p_{n-1}\left(\frac{ny-x}{n-1}\right). \tag{1.3}$$

For Markovian $\{X_n\}$, the joint pdf is not such a simple function of the marginals.

The Edgeworth expansion [10] is essentially a small deviations approximation in the sense that its asymptotic validity is limited to the region $y = o(n^{-\frac{1}{3}})$ [10] and thus can lead to a significant relative error in the tail region. For example, if X_n is supported in the interval $[-1, 1]$, the Edgeworth expansion results in an infinite relative error for $|y| > 1$.

The construction of a full asymptotic expansion in the case where X_n is a Markov process is a considerably harder problem than that for the i.i.d. case. The case of a finite state Markov chain was solved in [11] and [12] (also see [1]), and [12] also considers the general case. In [11] a restricted case of continuous state space is considered and a full asymptotic expansion of the pdf of Y_n is given, but not of the joint pdf of the pair $\{X_n, Y_n\}$. The expansions depend on the initial distribution of the chain.

The purpose of this paper is to propose a straightforward method for the explicit calculation of the full formal asymptotic expansion of the pdfs $p_n(y)$ and $p_n(x, y)$ for a discrete or continuous state stationary Markov chain. The main results of this paper are (2.30) and (2.31) for the leading terms in the asymptotic expansion of the pdfs. The calculations can be extended to higher order terms in the expansion (see the Appendix). These results reduce to those of [11], [12], and [13] for a finite state stationary Markov chain. Formulas (2.30) and (2.31) are convenient for calculations and give an explicit expression for the pre-exponential factor in the case of an asymmetric telegraph process. The expansion is formal and no proof of asymptotic convergence is presented.

Our method is based on the observation [12, 14] that the pair $\{X_n, Y_n\}$ is Markovian so that its joint pdf (whenever it exists) satisfies the forward Kolmogorov equation. We construct an asymptotic solution to this equation in the Wentzel-Kramers-Brillouin (WKB) form [15]. The routine application of the WKB technique leads to a breakdown of the expansion scheme given in [9] so that a modification is introduced to remove the degeneration.

In section 2, we present the main results of the paper. We also consider averages of the form

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i),$$

where f is a bounded measurable function. In section 3, we generalize the results to include vector valued random variables. A discrete Markov chain example is presented in section 4. Finally, in section 5 an application is given to "small deviations" theory.

2. Asymptotic expansion of the joint pdf. Let $\{X_n\}$ be an ergodic stationary Markov process with zero mean and stationary pdf $\phi_0(x)$. The first two moments of the stationary pdf are denoted

$$m_1 \equiv \int t\phi_0(t)\,dt = 0, \qquad m_2 \equiv \int t^2\phi_0(t)\,dt. \tag{2.1}$$

The Kolmogorov forward equation (or the master equation) for the transition probability density function (tpdf), $p_n(y|x)$, is given by [10]

$$p_{n+1}(y|x) = \int \rho(y,t) p_n(t|x)\, dt \equiv \mathbf{L} p_n(y|x), \tag{2.2}$$

where

$$\rho(y,t) \equiv \frac{\partial}{\partial y} \Pr\{X_{n+1} \le y | X_n = t\}. \tag{2.3}$$

We use the term pdf in the sense of distributions, and the partial derivative in (1.3) and (2.3) is understood in this sense as well. Thus, the pdf of a discrete variable is a sum of delta functions.

The initial condition for (2.2) is

$$p_0(y|x) = \delta(x-y). \tag{2.4}$$

The operator $\mathbf{L}$ is bounded in $L^1(R)$ with $||\mathbf{L}||_1 \le 1$. The stationary pdf, $\phi_0(x)$, is an eigenfunction of $\mathbf{L}$ corresponding to the eigenvalue $\mu_0 = 1$.

The joint pdf of $\{X_n, Y_n\}$ satisfies the forward Kolmogorov equation

$$p_{n+1}(x,y) = \frac{n+1}{n} \int \rho(x,\xi) p_n\left(\xi, y + \frac{y-x}{n}\right) d\xi. \tag{2.5}$$

Following the method of [9], an asymptotic solution of (2.5) is constructed in the form

$$p_n(x,y) = K_n(x,y) e^{-n\psi(y)}, \tag{2.6}$$

where

$$K_n(x,y) = \sqrt{n}\left[q^0(x,y) + \frac{1}{n}\, q^1(x,y) + \frac{1}{n^2}\, q^2(x,y) + \cdots\right]. \tag{2.7}$$

Substituting (2.6) and (2.7) into (2.5) and expanding in negative powers of n, it is found at the leading order that

$$q^0(x,y)\, e^{-\psi(y)+y\psi'(y)} = \int \rho(x,t) q^0(t,y)\, e^{x\psi'(y)}\, dt. \tag{2.8}$$

The function $\psi(y)$ and some information on $q^0(x,y)$ are determined from (2.8) as follows. A one parameter family of operators, $\mathbf{M}(y)$, acting on functions of x in $L^1(R)$, for any fixed y, is defined by

$$\mathbf{M}(y) f(x) \equiv e^{x\psi'(y)} \int \rho(x,t) f(t)\, dt \equiv e^{x\psi'(y)} \mathbf{L} f(x).$$

We now write (2.8) in the form

$$\mu(y) q^0(x,y) = \mathbf{M}(y) q^0(x,y), \quad \mu(y) = e^{-\psi(y)+y\psi'(y)},$$

that is, $q^0(x,y)$ is an eigenfunction of $\mathbf{M}(y)$ with the eigenvalue $\mu(y)$. The operator $\mathbf{M}(y)$ has a nonnegative kernel. We assume that the operator $\mathbf{M}(y)$ has the greatest positive eigenvalue $\mu(y)$ with geometric multiplicity one and that the corresponding eigenfunction, $q^0(x,y)$, is the only positive eigenfunction up to normalization ([16, p. 287], [10, p. 271], [14]).

It is well known that the rate function is convex; therefore $\psi'(y)$, wherever it exists, is an increasing function and thus has an inverse in its range. It follows that $\mu(y)$ can be considered as a function of $\psi'(y)$. We use interchangeably, with some abuse of notation, both $\mu(\psi'(y))$ and $\mu(y)$. Equation (2.8) is a first order differential equation for $\psi(y)$,

$$\mu(\psi'(y)) = e^{-\psi(y)+y\psi'(y)}\,, \tag{2.9}$$

or equivalently,

$$-\psi(y) + y\psi'(y) = \log \mu(\psi'(y))\ . \tag{2.10}$$

Equation (2.10) can be reduced to an implicit equation for $\psi'(y)$ by differentiating (2.10) with respect to y,

$$y = \frac{\mu'(\psi'(y))}{\mu(\psi'(y))}\ . \tag{2.11}$$

The case of independent random variables can be recovered from (2.11) as follows. If X_n are i.i.d. random variables, then $\rho(x,t)$ is independent of t. Therefore, integrating (2.8) with respect to x, it is found that $\mu(\psi'(y))$ is the moment generating function of $\rho(x)$, the pdf of X_n, that is,

$$\mu(\psi'(y)) = \int e^{\psi'(y)x}\rho(x)\,dx\,.$$

Thus (2.10) reduces to the well-known result of large deviations theory [5].

The large deviations principle (2.6) with (2.9)–(2.11) is equivalent to that mentioned in [4, Chap. 4, section 1] in the sense that $\psi(y)$ is the Legendre transform of the spectral radius of the operator $\mathbf{M}(y)$.

Equation (2.8) determines $q^0(x,y)$ as an eigenfunction of $\mathbf{M}(y)$, up to a normalization factor which is a function of y. This function is usually found from the next order equation, obtained by comparing the coefficients of $n^{-1/2}$ in the expansion (2.6), (2.7) on both sides of the Kolmogorov equation (2.5). The resulting equation degenerates to 0=0, as can be seen from (2.24). To remove the degeneration, we define an approximate sample average [9],

$$Y_{n+1}^{\epsilon} = \frac{n+\epsilon}{n+1}\,Y_n^{\epsilon} + \frac{1}{n+1}X_{n+1}, \tag{2.12}$$

where ϵ is a small parameter, ultimately to be set equal to zero. Kolmogorov's equation (2.5) now takes the form

$$p_{n+1}^{\epsilon}(x,y) = \frac{n+1}{n+\epsilon}\int \rho(x,\xi)\,p_n^{\epsilon}\left(\xi, y + \frac{(1-\epsilon)y - x}{n+\epsilon}\right)\,d\xi\,.$$

Consequently, the eikonal equation (2.8) takes the form

$$\{\mu(y) - \mathbf{M}(y)\}\,q_{\epsilon}^0(x,y) = 0, \tag{2.13}$$

where

$$\mu(y) = e^{-\psi_\epsilon(y)+(1-\epsilon)y\psi_\epsilon'(y)},$$

and (2.11) becomes

$$(1-\epsilon)y - \frac{\mu'(\psi'_\epsilon(y))}{\mu(\psi'_\epsilon(y))} = \frac{\epsilon\,\psi'_\epsilon(y)}{\psi''_\epsilon(y)}.$$

To simplify notation, we drop the ϵ in the notation for functions of the perturbed process (2.12).

The next order equation can be written as

$$\left\{\left(\frac{1}{2}-\epsilon\right) + \epsilon[(1-\epsilon)y-x]\psi'(y) - \frac{1}{2}\psi''(y)[(1-\epsilon)y-x]^2\right\}\mu(y)q^0(x,y) \tag{2.14}$$
$$+\,[(1-\epsilon)y-x]\mathbf{M}(y)q^0_y(x,y) = [\mu(y)-\mathbf{M}(y)]q^1(x,y)\,.$$

Equation (2.15) is handled by considering the adjoint operator to $\mathbf{M}(y)$, denoted $\mathbf{M}^*(y)$. The operator $\mathbf{M}^*(y)$ acts on functions in $L\infty(R)$ and is defined by the pairing

$$\langle f,g\rangle = \int f(x)g(x)\,dx, \qquad f\in L^\infty(R)\, g\in L^1(R),$$

as

$$\mathbf{M}^*(y)f(t) \equiv \int \rho(x,t)\,e^{x\psi'(y)}f(x)\,dx.$$

The eigenvalue $\mu(y)$ is also the greatest positive eigenvalue of $\mathbf{M}^*(y)$ with a corresponding positive eigenfunction of multiplicity one, denoted $\tilde{p}(x,y)$. We conclude from (2.8) that

$$\mathbf{M}^*(y)\tilde{p}(t,y) \equiv \int \rho(x,t)\,e^{x\psi'(y)}\tilde{p}(x,y)\,dx = \,e^{-\psi(y)+y\psi'(y)}\,\tilde{p}(t,y). \tag{2.15}$$

By orthogonality considerations,

$$\int \tilde{p}(x,y)\,[\mathbf{M}(y)-\mu(y)]\,f(x)\,dx \,=\, 0 \tag{2.16}$$

for any function $f(x)$ in the domain of $\mathbf{M}(y)$. We write $q^0(x,y)$ as

$$q^0(x,y) = k_0(y)\tilde{q}(x,y)\,, \tag{2.17}$$

where $\tilde{q}(x,y)$ is normalized by

$$\int \tilde{q}(x,y)\,dx = 1, \tag{2.18}$$

and $k_0(y)$ is determined from (2.15). We normalize $\tilde{p}(x,y)$ by

$$\int \tilde{q}(x,y)\tilde{p}(x,y)\,dx = 1. \tag{2.19}$$

Multiplying (2.15) by $\tilde{p}(x,y)$ and integrating with respect to x, we obtain the differential equation

$$\left(\frac{1}{2}-\epsilon\right)\mu(y)k_0(y) + \epsilon\psi'(y)\int[(1-\epsilon)y-x]\mu(y)q^0(x,y)\tilde{p}(x,y)\,dx$$
$$-\frac{1}{2}\psi''(y)\int[(1-\epsilon)y-x]^2\mu(y)q^0(x,y)\tilde{p}(x,y)\,dx$$
$$+\int[(1-\epsilon)y-x]\mathbf{M}(y)q^0_y(x,y)\tilde{p}(x,y)dx = 0. \tag{2.20}$$

Equation (2.20) is used to determine the factor $k_0(y)$.

We need the following two identities. By differentiating (2.13) with respect to y, multiplying the result by $\tilde{p}(x,y)$, and integrating with respect to x, we get

$$\int [(1-\epsilon)y - x]\mu(y)q^0(x,y)\tilde{p}(x,y)dx = \epsilon\frac{\psi'(y)}{\psi''(y)}\mu(y)k_0(y)\,. \tag{2.21}$$

For $\epsilon = 0$, this identity reduces to

$$\int (y-x)\mu(y)q^0(x,y)\tilde{p}(x,y)\,dx = 0\,. \tag{2.22}$$

Next, differentiating eq. (2.13) with respect to y twice, multiplying the result by $\tilde{p}(x,y)$, integrating with respect to x, and applying eq. (2.21), we get the identity

$$\begin{aligned}
&\int [(1-\epsilon)y - x]\tilde{p}(x,y)\mathbf{M}(y)q^0_y(x,y)\,dx \\
&\quad - \frac{1}{2}\psi''(y)\int [(1-\epsilon)y - x]^2\mu(y)q^0(x,y)\tilde{p}(x,y)\,dx \\
&= \epsilon\frac{\psi'(y)}{\psi''(y)}\int \tilde{p}(x,y)\mathbf{M}(y)q^0_y(x,y)dx + \frac{\epsilon}{2}\int x\psi'(y)\mu(y)q^0(x,y)\tilde{p}(x,y)\,dx \\
&\quad + \left\{\frac{\epsilon(\epsilon-1)}{2}y\psi'(y) - \frac{\epsilon\psi'''(y)\psi'(y)}{2(\psi'')^2} - \frac{(1-2\epsilon)}{2}\right\}\mu(y)k_0(y)\,.
\end{aligned} \tag{2.23}$$

Now, substituting (2.21) and (2.23) into (2.20), we get

$$\begin{aligned}
&\int \frac{\epsilon\psi'(y)}{\psi''(y)}\tilde{p}(x,y)\mathbf{M}(y)q^0_y(x,y)dx + \int \frac{\epsilon}{2}\psi'(y)\,x\mu(y)q^0_y(x,y)\tilde{p}(x,y)\,dx \\
&= \left\{\frac{\epsilon}{2}\,y\psi'(y) + \frac{\epsilon}{2}\frac{\psi'''(y)\psi'(y)}{[\psi''(y)]^2} - \frac{\epsilon^2}{2}\psi'(y)y - \frac{\epsilon^2\psi'^2(y)}{\psi''(y)}\right\}\mu(y)k_0(y).
\end{aligned} \tag{2.24}$$

Dividing (2.24) by ϵ, taking the limit $\epsilon \to 0$, and applying (2.22), we obtain

$$\frac{1}{2}\frac{\psi'''(y)}{\psi''(y)}\mu(y)k_0(y) - \int \tilde{p}(x,y)\mathbf{M}(y)q^0_y(x,y)\,dx = 0\,, \tag{2.25}$$

for which the notation without ϵ is correct. We observe that by (2.16)

$$\int \tilde{p}(x,y)\mathbf{M}(y)q^0_y(x,y)\,dx = \int \mu(y)q^0_y(x,y)\tilde{p}(x,y)\,dx\,. \tag{2.26}$$

Finally, using (2.17) and (2.26) in (2.25), we obtain a simplified equation for $k_0(y)$,

$$\frac{k_0'(y)}{k_0(y)} = \frac{1}{2}\frac{\psi'''(y)}{\psi''(y)} - \int \tilde{q}_y(x,y)\tilde{p}(x,y)\,dx. \tag{2.27}$$

We denote

$$f(y) \equiv \int \tilde{q}_y(x,y)\tilde{p}(x,y)\,dx, \qquad F(y) \equiv \int_0^y f(z)\,dz. \tag{2.28}$$

Then the solution of (2.27) is

$$k_0(y) = C\sqrt{\psi''(y)}e^{-F(y)}\,,$$

where C is a constant. By the normalization requirement of $\sqrt{n}k_0(y)e^{-n\psi(y)}$ in the limit of large n, using Laplace's method for the asymptotic evaluation of integrals [15], one obtains that

$$k_0(y) = \sqrt{\frac{\psi''(y)}{2\pi}}e^{-F(y)}. \tag{2.29}$$

In the special case where X_n are i.i.d. random variables, the expression (2.29) is simplified. By (2.15) $\tilde{p}(x, y) = 1$, resulting in

$$f(y) = \int \tilde{q}_y(x, y)\tilde{p}(x, y)dx = 0,$$

in agreement with [9]. It is also found by (2.8) that, for the i.i.d. case,

$$\tilde{q}(x, y) = \frac{\rho(x)e^{\psi'(y)x}}{\mu(\psi'(y))}.$$

Finally, we can present the large deviations result for the joint pdf of $\{X_n, Y_n\}$,

$$p_n(x, y) \sim \left\{\sqrt{\frac{n\psi''(y)}{2\pi}}\tilde{q}(x, y)e^{-F(y)}\right\} e^{-n\psi(y)}, \tag{2.30}$$

where $\tilde{q}(x, y)$ is the eigenfunction of the operator $\mathbf{M}(y)$, corresponding to the largest eigenvalue $\mu(y)$, normalized by (2.18). The asymptotic representation of the pdf of Y_n is given by

$$p_n(y) \sim \left\{\sqrt{\frac{n\psi''(y)}{2\pi}}e^{-F(y)}\right\} e^{-n\psi(y)}. \tag{2.31}$$

Both results, (2.30) and (2.31), seem to be new.

Next, we consider the more general case of the sum

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i),$$

where $f(x)$ is a measurable bounded function. Kolmogorov's equation is given by

$$p_{n+1}(x, y) = \frac{n+1}{n}\int \rho(x, \xi)p_n\left(\xi, y + \frac{y - f(x)}{n}\right) d\xi. \tag{2.32}$$

We assume that $\{f(X_n)\}$ is an ergodic stationary process for which the first two moments exist.

The asymptotic expansion for (2.32) is constructed as above. The eikonal equation is

$$q^0(x, y)\, e^{-\psi(y)+y\psi'(y)} = \int \rho(x, t)q^0(t, y)\, e^{f(x)\psi'(y)}\, dt. \tag{2.33}$$

We find that $\mathbf{M}(y) \equiv e^{f(x)\psi'(y)}\mathbf{L}(y)$, its greatest eigenvalue $\mu(y)$, the normalized eigenfunctions $\tilde{q}(x, y)$ and $\tilde{p}(x, y)$, and the rate function $\psi(y)$ are all dependent on the function $f(x)$. We obtain

$$p_n(x, y) \sim \left\{\sqrt{\frac{n\psi''(y)}{2\pi}}\tilde{q}(x, y)e^{-F(y)}\right\} e^{-n\psi(y)}, \tag{2.34}$$

and

$$p_n(y) \sim \left\{ \sqrt{\frac{n\psi''(y)}{2\pi}} e^{-F(y)} \right\} e^{-n\psi(y)}, \tag{2.35}$$

where

$$F(y) \equiv \int_0^y \int_{-\infty}^{\infty} \tilde{q}_z(x,z)\tilde{p}(x,z)\,dx\,dz.$$

The explicit form of the pre-exponential factor in the expansion (2.35) can be used to improve the large deviation theory (LDT) estimate of the sample size needed to achieve a given significance level in statistical tests or of the word length required to ensure a given bound on the decoding error in block coding. These applications will be discussed in a separate paper.

3. *N*-dimensional Markov processes. The results obtained in section 2 can be generalized by considering $\{\mathbf{X}_n\}$ to be an N-dimensional ergodic stationary Markov process with stationary probability density function (pdf) $\phi_0(\mathbf{x}) \equiv \phi_0(x_1,\ldots,x_N)$. The moments of the stationary pdf, up to second order, are denoted

$$\mathbf{m}_{k_1\cdots k_N} \equiv \int \cdots \int z_1^{k_1}\cdots z_N^{k_N}\,\phi_0(z_1\ldots z_N)\,dz_1\cdots dz_N, \qquad k_1+\cdots+k_N \le 2.$$

The Kolmogorov forward equation for the tpdf, $p_n(\mathbf{y}|\mathbf{x})$, is given by

$$p_{n+1}(\mathbf{y}|\mathbf{x}) = \int \rho(\mathbf{y},\mathbf{z})p_n(\mathbf{z}|\mathbf{x})\,d\mathbf{z} \equiv \mathbf{L}p_n(\mathbf{z}|\mathbf{x})\,, \qquad p_0(\mathbf{y}|\mathbf{x}) = \delta(\mathbf{x}-\mathbf{y}),$$

where

$$\rho(\mathbf{y},\mathbf{z}) \equiv \frac{\partial^N}{\partial y^1\cdots\partial y^N}\,\Pr\{X_{n+1}\le \mathbf{y}|X_n=\mathbf{z}\}.$$

We consider the average

$$\mathbf{Y}_n \equiv \frac{1}{n}\sum_{j=1}^{n}\mathbf{f}(\mathbf{X}_j),$$

where $\mathbf{f}$ is a bounded measurable function on R^M, for any natural M, and we assume that $\{\mathbf{f}(\mathbf{X_n})\}$ is an ergodic and stationary process for which the stationary moments up to second order exist. As in the scalar case, for a small parameter ϵ an average $\mathbf{Y}_n^\epsilon$ is constructed as above, and we obtain the forward Kolmogorov equation for the joint pdf of the pair $\{\mathbf{X}_n,\mathbf{Y}_n^\epsilon\}$,

$$p_{n+1}^\epsilon(\mathbf{x},\mathbf{y}) = \left(\frac{n+1}{n+\epsilon}\right)^N \int\cdots\int \rho(\mathbf{x},\boldsymbol{\xi})p_n^\epsilon\left(\boldsymbol{\xi},\mathbf{y}+\frac{(1-\epsilon)\mathbf{y}-\mathbf{f}(\mathbf{x})}{n+\epsilon}\right)d\xi_1\cdots d\xi_N\,.$$

The asymptotic solution now takes the form

$$p_n(\mathbf{x},\mathbf{y}) = K_n(\mathbf{x},\mathbf{y})\,e^{-n\psi(\mathbf{y})},$$

where

$$K_n(\mathbf{x},\mathbf{y}) = (\sqrt{n})^N \left[q^0(\mathbf{x},\mathbf{y}) + \frac{1}{n}q^1(\mathbf{x},\mathbf{y}) + \frac{1}{n^2}q^2(\mathbf{x},\mathbf{y}) + \cdots\right].$$

To simplify notation, we denote

$$\frac{\partial}{\partial y_j}\psi(\mathbf{y}) \equiv \psi_j(\mathbf{y}) \equiv \psi_j,$$

and summation over repeated indices is assumed. Approximation at leading order leads to the equation

$$q^0(\mathbf{x},\mathbf{y})e^{-\psi(\mathbf{y})+(1-\epsilon)\mathbf{y}\cdot\nabla\psi(\mathbf{y})} = \int\cdots\int \rho(\mathbf{x},\boldsymbol{\xi})q^0(\boldsymbol{\xi},\mathbf{y})e^{\mathbf{f}(\mathbf{x})\cdot\nabla\psi(\mathbf{y})}\,d\xi_1\cdots d\xi_N. \tag{3.1}$$

For $\epsilon = 0$ it reduces to

$$q^0(\mathbf{x},\mathbf{y})e^{-\psi(\mathbf{y})+\mathbf{y}\cdot\nabla\psi(\mathbf{y})} = \int\cdots\int \rho(\mathbf{x},\boldsymbol{\xi})q^0(\boldsymbol{\xi},\mathbf{y})e^{\mathbf{f}(\mathbf{x})\cdot\nabla\psi(\mathbf{y})}\,d\xi_1\cdots d\xi_N.$$

We set

$$\boldsymbol{\theta} \equiv \nabla\psi(\mathbf{y})$$

and consider a one parameter family of operators defined by

$$\mathbf{M}(\boldsymbol{\theta})\phi(\mathbf{x},\mathbf{y}) \equiv e^{\mathbf{f}(\mathbf{x})\cdot\boldsymbol{\theta}}\int\cdots\int \rho(\mathbf{x},\mathbf{t})\phi(\mathbf{t},\mathbf{y})\,dt_1\cdots dt_N \equiv e^{\mathbf{f}(\mathbf{x})\cdot\boldsymbol{\theta}}\mathbf{L}\phi(\mathbf{x},\mathbf{y}).$$

By the same considerations as in the scalar case, the operator $\mathbf{M}(\boldsymbol{\theta})$ possesses a positive largest eigenvalue, denoted $\mu(\boldsymbol{\theta})$. Thus, $\psi(\mathbf{y})$ is a solution of the first order partial differential equation

$$\mu(\boldsymbol{\theta}) = e^{-\psi(\mathbf{y})+(1-\epsilon)\mathbf{y}\cdot\nabla\psi(\mathbf{y})}.$$

The function $q^0(\mathbf{x},\mathbf{y})$ is a positive eigenfunction of $\mathbf{M}(\theta)$ determined up to a factor $k_0(\mathbf{y})$, which is found from the next order equation

$$\begin{aligned}&\left\{N\left(\frac{1}{2}-\epsilon\right) + \epsilon[(1-\epsilon)y_i - f_i(\mathbf{x})]\psi_i(\mathbf{y})\right.\\
&\left.-\frac{1}{2}[(1-\epsilon)y_i - f_i(\mathbf{x})]\,\psi_{ij}(\mathbf{y})\,[(1-\epsilon)y_j - f_j(\mathbf{x})]\right\}\mu(\boldsymbol{\theta})q^0(\mathbf{x},\mathbf{y})\\
&+[(1-\epsilon)y_i - f_i(\mathbf{x})]\mathbf{M}(\theta)q_i^0(\mathbf{x},\mathbf{y}) = [\mu(\boldsymbol{\theta}) - \mathbf{M}(\boldsymbol{\theta})]q^1(\mathbf{x},\mathbf{y}).\end{aligned}$$

To determine $k_0(\mathbf{y})$, we follow the considerations used in the scalar case. The eigenfunction of the adjoint operator $\mathbf{M}^*(\boldsymbol{\theta})$ is denoted $p_0(\mathbf{x},\mathbf{y})$ and is normalized, together with the eigenfunction $q^0(\mathbf{x},\mathbf{y})$ to $\tilde{q}(\mathbf{x},\mathbf{y})$ and $\tilde{p}(\mathbf{x},\mathbf{y})$ as in (2.18) and (2.19). Equation (3.1) is differentiated, yielding two identities, as above. All these lead to the equation

$$(\psi_{lk})^{-1}\psi_l\left[\int \tilde{q}_k(\mathbf{x},\mathbf{y})\tilde{p}(\mathbf{x},\mathbf{y})\,d\mathbf{x} + \frac{\partial}{\partial y_k}\log k_0(\mathbf{y}) - \frac{1}{2}\psi_{ikj}(\psi_{ij})^{-1}\right] = 0\,. \tag{3.2}$$

The following property of the Hessian matrix can be easily verified:

$$\left(\frac{\partial}{\partial y_k}\psi_{ij}\right)(\psi_{ij})^{-1} \equiv \frac{\partial}{\partial y_k}\left[\ln\left(\det\left[\psi_{ij}\right]\right)\right]. \tag{3.3}$$

Using (3.3) in (3.2) yields

$$(\psi_{lk})^{-1}\psi_l \frac{\partial}{\partial y_k}\left(\ln\left[\frac{k_0(\mathbf{y})}{\sqrt{\det[\psi_{ij}]}}\right]\right) = -(\psi_{lk})^{-1}\psi_l \int \frac{\partial}{\partial y_k}\tilde{q}(\mathbf{x},\mathbf{y})\tilde{p}(\mathbf{x},\mathbf{y})d\mathbf{x}. \tag{3.4}$$

A normalized solution for (3.4) is

$$k_0(\mathbf{y}) = \left(\frac{1}{2\pi}\right)^{N/2}\sqrt{\det[\psi_{ij}(\mathbf{y})]}\ \exp\left\{-\sum_{k=1}^{N}\int_0^{y_k}\left(\int\cdots\int \tilde{p}(\mathbf{x},\mathbf{y})\frac{\partial}{\partial y_k}\tilde{q}(\mathbf{x},\mathbf{y})\,d\mathbf{x}\right)dy_k\right\}.$$

The large deviations result for the joint pdf of $\{\mathbf{X}_n, \mathbf{Y}_n\}$ is

$$p_n(\mathbf{x},\mathbf{y}) \sim n^{N/2}k_0(\mathbf{y})\tilde{q}(\mathbf{x},\mathbf{y})e^{-n\psi(\mathbf{y})},$$

where $\tilde{q}(\mathbf{x},\mathbf{y})$ is the eigenfunction of the operator $\mathbf{M}(\boldsymbol{\theta})$, corresponding to the largest eigenvalue $\mu(\boldsymbol{\theta})$, normalized by (2.18). The asymptotic representation of the pdf of $\mathbf{Y}_n$ is given by

$$p_n(\mathbf{y}) \sim n^{N/2}k_0(\mathbf{y})e^{-n\psi(\mathbf{y})}. \tag{3.5}$$

4. An example. We consider a Markov process X_n that stays at -1 with probability p_1 and jumps to 1 with probability $1-p_1$, stays at 1 with probability p_2 and jumps to -1 with probability $1-p_2$. The one step transition density is given by

$$\begin{aligned}\rho(x,t) = &\{\,p_1\delta(x+1) + (1-p_1)\delta(x-1)\}\,\delta(t+1)\\ &+\{\,(1-p_2)\delta(x+1) + p_2\delta(x-1)\}\,\delta(t-1).\end{aligned}$$

Hence, $\mathbf{L}$ can be represented as the matrix

$$\begin{pmatrix} p_1 & 1-p_2 \\ 1-p_1 & p_2 \end{pmatrix}$$

operating on column vectors. The stationary pdf of X_n is given by

$$p(x) = \frac{1-p_2}{2-p_1-p_2}\delta(x+1) + \frac{1-p_1}{2-p_1-p_2}\delta(x-1).$$

Assumption (2.1) is

$$p_2 = p_1 \equiv p.$$

Hence, the stationary pdf of X_n is

$$p(x) = \frac{1}{2}\delta(x+1) + \frac{1}{2}\delta(x-1).$$

We denote $\theta = \psi'(y)$ and write the eikonal equation (2.8) in the form

$$\mu(\theta)\begin{bmatrix} q^0(-1,y) \\ q^0(1,y)\end{bmatrix} = \begin{bmatrix} e^{-\theta}p & e^{-\theta}(1-p) \\ e^{\theta}(1-p) & e^{\theta}p\end{bmatrix}\begin{bmatrix} q^0(-1,y) \\ q^0(1,y)\end{bmatrix}$$

and obtain that

$$\mu(\theta) = p\cosh(\theta) + \sqrt{(p-1)^2 + p^2\sinh^2\theta}\,. \tag{4.1}$$

The normalized eigenfunctions, with some abuse of notations, are given by

$$\begin{aligned}\tilde{q}(x,y) &\equiv \tilde{q}(-1,y)\delta(x+1)+\tilde{q}(1,y)\delta(x-1),\\ \tilde{p}(x,y) &\equiv \tilde{p}(-1,y)\delta(x+1)+\tilde{p}(1,y)\delta(x-1),\end{aligned} \tag{4.2}$$

where

$$\begin{bmatrix} \tilde{q}(-1,y) \\ \tilde{q}(1,y) \end{bmatrix} = \begin{bmatrix} e^{-\theta}(1-p) \\ \mu(\theta)-e^{-\theta}p \end{bmatrix} \frac{1}{e^{-\theta}(1-2p)+\mu(\theta)}$$

and

$$\begin{bmatrix} \tilde{p}(-1,y) \\ \tilde{p}(1,y) \end{bmatrix} = \begin{bmatrix} e^{\theta}(1-p) \\ \mu(\theta)-e^{-\theta}p \end{bmatrix} \frac{e^{-\theta}(1-2p)+\mu(\theta)}{(1-p)^2+(\mu(\theta)-e^{-\theta}p)^2}\,.$$

Next, we determine $\psi'(y)$. We write (2.11) in the form

$$\begin{aligned}&\sqrt{(p-1)^2+(2p-1)\tanh^2\theta}\,(p\tanh\theta-yp)\\ &=y\left((p-1)^2+(2p-1)\tanh^2\theta\right)-p^2\tanh\theta,\end{aligned}$$

and solve it with respect to the variable $\tanh\theta$ to obtain

$$\tanh\theta=\frac{y(1-p)}{\sqrt{y^2(1-2p)+p^2}}\,,$$

so we get that

$$\psi'(y)\equiv\theta=\tanh^{-1}\left[\frac{y(1-p)}{\sqrt{y^2(1-2p)+p^2}}\right]. \tag{4.3}$$

The rate function $\psi(y)$ is

$$\begin{aligned}\psi(y)=y\psi'(y)&-\frac{1}{2}\ln[(1-y^2)p^2(1-2p)^2]\\ &+\frac{1}{2}\ln\left|p^2+y(1-2p)+(p-1)\sqrt{p^2+y^2(1-2p)}\right|\\ &+\frac{1}{2}\ln\left|p^2-y(1-2p)+(p-1)\sqrt{p^2+y^2(1-2p)}\right|\,.\end{aligned}$$

To determine the factor $k_0(y)$ from (2.29), we find the function $\sqrt{\psi''(y)/2\pi}$ by differentiating (4.3) with respect to y, which yields

$$\sqrt{\frac{\psi''(y)}{2\pi}}=\sqrt{\frac{1-p}{2\pi(1-y^2)}}\,\frac{1}{[y^2(1-2p)+p^2]^{\frac{1}{4}}}.$$

Differentiating $\tilde{q}(x,y)$ with respect to y, we get that

$$\begin{bmatrix} \tilde{q}_y(-1,y) \\ \tilde{q}_y(1,y) \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \frac{\psi''(y)e^{-\theta}\,(1-p)\,\mu(\theta)\,(1+y)}{[e^{-\theta}\,(1-2p)+\mu(\theta)]^2},$$

and thus,

$$f(y) = \int_{-\infty}^{\infty} \tilde{q_y}(x,y)\tilde{p}(x,y)dx \equiv \langle \tilde{q_y}(x,y), \tilde{p}(x,y)\rangle$$
$$= \frac{[e^{-\theta}(1-p)\psi''(y)\mu(\theta)(1+y)][\mu(\theta) - e^{\theta}(1-p) - e^{-\theta}p]}{[e^{-\theta}(1-2p) + \mu(\theta)][\theta(1-p)^2 + (\mu(\theta) - e^{-\theta}p)^2]}.$$

The function $f(y)$ vanishes only for $p = \frac{1}{2}$ and then our approximation is the same as that for i.i.d. Bernoulli random variables.

5. Small deviations. We evaluate the probability densities $p_n(x,y)$ and $p_n(y)$ for small values of y, that is, for small deviations from the mean. We can change y to a new variable $\theta = \psi'(y)$, which is zero when y is zero. Then (2.8) can be solved for small y by expanding the eigenvalue $\mu(\theta)$, the corresponding eigenfunction $f(x,\theta)$, and the exponential function in powers of θ. We apply only quadratic order Taylor expansion. Thus, writing

$$\mu(\theta) = \mu_0 + \theta\mu_1 + \theta^2\mu_2 + o\left(\theta^2\right),$$
$$f(x,\theta) = f_0(x) + \theta f_1(x) + \theta^2 f_2(x) + o\left(\theta^2\right),$$
$$e^{x\theta} = 1 + \theta x + \frac{\theta^2 x^2}{2} + o\left(\theta^2\right),$$

and equating coefficients of like powers of θ in (2.8), we obtain

$$(\mu_0\mathbf{I} - \mathbf{L})f_0(x) = 0, \tag{5.1}$$
$$(\mu_0\mathbf{I} - \mathbf{L})f_1(x) = -(\mu_1\mathbf{I} - x\mathbf{L})f_0(x), \tag{5.2}$$
$$(\mu_0\mathbf{I} - \mathbf{L})f_2(x) = -(\mu_1\mathbf{I} - x\mathbf{L})f_1(x) - \left(\mu_2\mathbf{I} - \frac{x^2}{2!}\mathbf{L}\right) f_0(x). \tag{5.3}$$

From (5.1) it follows that $\mu_0 = 1$ and $f_0(x) = \phi_0(x)$. Integrating (5.2) with respect to x, it is found from (2.1) that $\mu_1 = 0$ and that

$$(\mathbf{I} - \mathbf{L})f_1(x) = x\phi_0(x).$$

Then, if $x\phi_0(x)$ is in the range of $\mathbf{I} - \mathbf{L}$, we can consider the operator $(\mathbf{I} - \mathbf{L})^{-1}$ on a subspace containing $x\phi_0(x)$ to obtain

$$f_1(x) = (\mathbf{I} - \mathbf{L})^{-1}x\phi_0(x) \equiv \sum_{i=0}^{\infty} \mathbf{L}^i x\phi_0(x) + cf_0(x), \tag{5.4}$$

where c is an arbitrary constant. Integrating (5.3) with respect to x, and using (5.4), it is found that

$$\mu_2 = \frac{m_2}{2} + \int x\mathbf{L}f_1(x)\,dx = \frac{m_2}{2} + \int \sum_{i=1}^{\infty} x\mathbf{L}^i x\phi_0(x)\,dx. \tag{5.5}$$

Next, the exponent in (2.9) is expanded in Taylor series in y,

$$e^{-\psi(y)+y\psi'(y)} = 1 + \frac{1}{2}\psi''(0)\,y^2 + o\left(y^2\right). \tag{5.6}$$

Setting

$$\theta = \psi'(y) = \psi''(0)\, y + o(y)$$

in (5.6), we obtain

$$\mu(\theta) = 1 + \frac{\theta^2}{2\psi''(0)} + o\left(\theta^2\right),$$

and thus,

$$2\mu_2 = \frac{1}{\psi''(0)},$$

and by (5.5), we conclude that

$$\psi''(0) = \frac{1}{2\sum_{i=1}^{\infty} \int \, x\mathbf{L}^i x\phi_0(x)\, dx + m_2}\,. \tag{5.7}$$

It can easily be found that the denominator in (5.7) is the spectral density of the process $\{X_n\}$ at frequency 0, denoted $\mathcal{S}(0)$. Indeed, the autocorrelation function of $\{X_n\}$ is given by

$$R(n) = \int \int xy\phi_0(x)p_n(y|x)\, dx\, dy.$$

From (2.2) and (2.4), it follows that for $n \geq 1$,

$$p_n(y|x) = \rho_n(y,x),$$

where

$$\rho_1(y,x) = \rho(y,x),$$

and

$$\rho_{n+1}(y,x) = \int \rho_n(y,t)\rho(t,x)\, dt.$$

With this notation the integrals in the denominator of (5.7) can be written as

$$\int x\mathbf{L}^i x\phi_0(x)\, dx = \int\int xy\phi_0(x)\rho_i(y,x)\, dx\, dy. \tag{5.8}$$

For $n \geq 1$, the autocorrelation function can be written as

$$R(n) = \int\int xy\phi_0(x)\rho_n(x,y)\, dx\, dy, \tag{5.9}$$

while for $n < 0$, we have

$$R(n) = R(-n), \tag{5.10}$$

and for $n = 0$

$$R(0) = m_2. \tag{5.11}$$

Applying (5.8)–(5.11), we write (5.7) as

$$\psi''(0) = \frac{1}{\sum_{n=-\infty}^{\infty} R(n)} \equiv \frac{1}{S(0)},$$

where $S(0)$ is the spectral density at frequency 0. The asymptotic pdf of Y_n for small deviations is found by integrating (2.6) with respect to x and by expanding for small y. Thus Y_n is asymptotically a zero mean normal variable with variance $\mathcal{S}(0)/n$ (see [17]) , that is,

$$p_n(y) \sim \sqrt{\frac{n}{2\pi S(0)}}\, e^{\frac{-ny^2}{2S(0)}},$$

$$p_n(x,y) \sim \sqrt{\frac{n}{2\pi S(0)}}\, \phi_0(x)\, e^{\frac{-ny^2}{2S(0)}}\,.$$

Appendix. Analysis of the next order terms. The full asymptotic expansion for $p_n(x,y)$, (2.6), is in the form of an infinite series. If the relations $q^{i+1}(x,y)/q^i(x,y)$ are bounded uniformly in x and y, expansion (2.6) is indeed an asymptotic series and the leading order approximation for $p_n(x,y)$ can be limited only to the first terms $\psi(y)$ and $q^0(x,y)$. In certain cases, e.g., when X_n have bounded support, the WKB expansion may fail at the boundary of the support and must be fixed there by a boundary layer. This problem was extensively studied in [7] for the case of i.i.d. random variables.

The calculations for evaluating all terms are very long. We evaluate only the next term $q^1(x,y)$. The quotients $q^{i+1}(x,y)/q^i(x,y)$ seem to be asymptotically similar to $q^1(x,y)/q^0(x,y)$.

We substitute the approximation (2.6) into (2.5) and expand it in negative powers of n. The zeroth order equation is (2.8). The first order equation is an equation for $q^1(x,y)$,

$$\begin{aligned} &[\mu(y) - \mathbf{M}(y)]q^1(x,y) \\ &= \{1 - \psi''(y)(y-x)^2\}\frac{1}{2}\mu(y)q^0(x,y) + (y-x)\mathbf{M}(y)q^0_y(x,y), \end{aligned} \tag{A.1}$$

which is (2.15) for $\epsilon = 0$. We solve $q^1(x,y)$ as follows. By differentiating (2.8), we obtain the two identities

$$\begin{aligned} &[\mathbf{M}(\mathbf{y})(y) - \mu(y)]q^0_y(x,y) = (y-x)\psi''(y)\mu(y)q^0(x,y), \\ &[\mathbf{M}(y) - \mu(y)]q^0_{yy}(x,y) \\ &= \left\{\psi''(y) + (y-x)\psi'''(y) + (y-x)^2\psi''^2(y)\right\}\mu(y)q^0(x,y) \\ &+2(y-x)\psi''(y)\mu(y)q^0_y(x,y). \end{aligned} \tag{A.2}$$

Hence,

$$q^1_p(x,y) = \frac{\psi'''(y)}{2\psi''^2(y)}q^0_y(x,y) - \frac{1}{2\psi''(y)}q^0_{yy}(x,y)$$

is a particular solution for (A.1). Thus,

$$q^1(x,y) = A(y)q^0(x,y) + q_p^1(x,y), \tag{A.3}$$

where $A(y)$ is determined from the next order equation (A.5) below.

The second order equation is

$$\begin{aligned}
&[\mu(y) - \mathbf{M}(y)]q^2(x,y) \\
&= \mu(y)\left[\frac{1}{8}q^0(x,y) + \frac{1}{2}q^1(x,y)\right] \\
&+ \left[\frac{1}{8}(y-x)^4\psi''^2(y) - \frac{1}{6}(y-x)^3\psi'''(y) - \frac{1}{2}(y-x)^2\psi''(y)\right]\mu(y)q^0(x,y) \\
&+ \left[(y-x) - \frac{1}{2}(y-x)^3\psi''(y)\right]\mathbf{M}(y)q_y^0(x,y) + \frac{1}{2}(y-x)^2\mathbf{M}(y)q_{yy}^0(x,y) \\
&+(y-x)\mathbf{M}(y)q_y^1(x,y) + \left[1 - \frac{1}{2}(y-x)^2\psi''(y)\right]\mathbf{M}(y)q^1(x,y).
\end{aligned} \tag{A.4}$$

In order to determine $A(y)$, we first multiply eq. (A.5) by $\tilde{p}(x,y)$ and integrate with respect to x. Then, we simplify it by differentiating eq.(A.1) up to second order and (2.8) up to fourth order. We also substitute (A.3) for $q^1(x,y)$. This was done using *Mathematica*, and we have found that

$$A(y) = \frac{1}{8}\frac{\psi^{(iv)}(y)}{\psi''^2(y)} - \frac{5}{24}\frac{\psi'''^2(y)}{\psi''^3(y)},$$

and thus

$$\begin{aligned}
q^1(x,y) = &\left\{\frac{1}{8}\frac{\psi^{(iv)}(y)}{\psi''^2(y)} - \frac{5}{24}\frac{\psi'''^2(y)}{\psi''^3(y)}\right\}q^0(x,y) \\
&+\frac{\psi'''(y)}{2\psi''^2(y)}q_y^0(x,y) - \frac{1}{2\psi''(y)}q_{yy}^0(x,y).
\end{aligned} \tag{A.5}$$

We substitute (2.17) into (A.5) and apply (2.18). Note that (2.18) also means that

$$\int \tilde{q}_y(x,y)\,dx = \int \tilde{q}_{yy}(x,y)\,dx = 0.$$

We thus obtain that

$$\begin{aligned}
k_1(y) &\equiv \int q^1(x,y)\,dx \\
&= \frac{1}{8}k_0(y)\frac{\psi^{iv}(y)}{\psi''^2(y)} - \frac{5}{24}k_0(y)\frac{\psi'''^2(y)}{\psi''^3(y)} \\
&+k_0'(y)\frac{\psi'''(y)}{2\psi''^2(y)} - \frac{1}{2\psi''(y)}k_0''(y).
\end{aligned} \tag{A.6}$$

Finally, we apply (2.29) into (A.6) and get that

$$\begin{aligned}
\frac{k_1(y)}{k_0(y)} = &\frac{1}{6}\frac{\psi'''^2(y)}{\psi''^3(y)} - \frac{1}{8}\frac{\psi^{iv}(y)}{\psi''^2(y)} + \frac{1}{2\psi''(y)}\int \tilde{q}_y(x,y)\tilde{p}_y(x,y)\,dx \\
&+\frac{1}{2\psi''(y)}\left[\int \tilde{q}_{yy}(x,y)\tilde{p}(x,y)\,dx - \left(\int \tilde{q}_y(x,y)\tilde{p}(x,y)\,dx\right)^2\right].
\end{aligned} \tag{A.7}$$

The expansion for $p_n(y)$ remains asymptotic as long as this quotient is bounded.

REFERENCES

[1] J. A. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley Series in Probability and Mathematical Statistics, John Wiley, New York, 1990.

[2] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, MA, 1992.

[3] R. S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Springer-Verlag, New York, 1985.

[4] J. D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, New York, 1989.

[5] P. H. Cramer, *Sur un nouveaux theorèm-limite de la theorie des probabilitiés"*, in Actualités Scientifiques et Industrielles 736, Colloque consacré à la théorie des probablités October 1937, Vol. 3. Hermann, Paris, 1938, pp. 5–23.

[6] R. P. Bahadur and R. Ranga Rao, *On deviations of the sample mean*, Ann. Math. Statist., 31 (1960), pp. 1015–1027.

[7] H. E. Daniels, *Saddlepoint approximations in statistics*, Ann. Math. Stat., 25 (1954), pp. 631–650.

[8] O. Barndorf-Nielsen and D.R. Cox, *Edgeworth and saddle-point approximations with statistical applications*, J. Roy. Statist. Soc. B, 41 (1979), pp. 279–312.

[9] C. Knessl, B. J. Matkowsky, Z. Schuss, and C. Tier, *An asymptotic theory of large deviations for Markov jump processes*, SIAM J. Appl. Math., 46 (1985), pp. 1006–1028.

[10] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd ed., John Wiley, New York, 1970.

[11] H. D. Miller, *A convexity property in the theory of random variables on a finite Markov chain*, Ann. Math. Statist., 32 (1961), pp. 1260–1270.

[12] J. L. Jensen, *Saddlepoint expansions for sums of Markov dependent variables on a continuous state space*, Probab. Theory Related Fields, 89 (1991), pp. 181–199.

[13] J. L. Jensen, *Saddlepoint Approximations*, Oxford Statistical Science Series, Oxford, 1995.

[14] I. Iscoe, P. Ney, and E. Nummelin, *Large deviations of uniformly recurrent Markov additive processes*, Adv. Appl. Math., 6 (1985), pp. 373–412.

[15] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw–Hill, New York, 1987.

[16] J. A. Cochran, *The Analysis of Linear Integral Equations*, McGraw–Hill, New York, 1972.

[17] P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1968.