J. Appl. Prob. **52,** 1133–1145 (2015)
Printed in England

ESTIMATION OF INTEGRALS WITH RESPECT TO INFINITE MEASURES USING REGENERATIVE SEQUENCES

KRISHNA B. ATHREYA * AND
VIVEKANANDA ROY,** *Iowa State University*

Abstract

Let f be an integrable function on an infinite measure space $(S, \mathcal{S}, \pi)$. We show that if a *regenerative sequence* $\{X_n\}_{n\geq 0}$ with canonical measure π could be generated then a consistent estimator of $\lambda \equiv \int_S f \,\mathrm{d}\pi$ can be produced. We further show that under appropriate second moment conditions, a confidence interval for λ can also be derived. This is illustrated with estimating countable sums and integrals with respect to absolutely continuous measures on $\mathbb{R}^d$ using a simple symmetric random walk on $\mathbb{Z}$.

Keywords: Markov chain; Monte Carlo; improper target; random walk; regenerative sequence

2010 Mathematics Subject Classification: Primary 65C05
Secondary 60F05

1. Introduction

Let $(S, \mathcal{S}, \pi)$ be a measure space. Let $f\colon S \to \mathbb{R}$ be $\mathcal{S}$ measurable, and $\int_S |f| \,\mathrm{d}\pi < \infty$. The goal is to estimate $\lambda \equiv \int_S f \,\mathrm{d}\pi$. If π is a probability measure, that is, $\pi(S) = 1$, a well-known statistical tool is to estimate λ by sample averages $\bar{f}_n \equiv \sum_{j=1}^n f(\xi_j)/n$ based on independent and identically distributed (i.i.d.) observations, $\{\xi_j\}_{j=1}^n$, from π. This i.i.d. Monte Carlo (i.i.d. MC) method is a fundamental notion in statistics and has made the subject very useful in many areas of science. A refinement of this result is via the central limit theorem (CLT) from which it follows that, if $\int_S f^2 \,\mathrm{d}\pi < \infty$, then an asymptotic $(1 - \alpha)$-level confidence interval for λ can be obtained as $I_n \equiv (\bar{f}_n - z_\alpha \sigma_n/\sqrt{n}, \bar{f}_n + z_\alpha \sigma_n/\sqrt{n})$, where $\sigma_n^2 = \sum_{j=1}^n f^2(\xi_j)/n - \bar{f}_n^2$ and z_α is such that $\mathbb{P}(|Z| > z_\alpha) = \alpha$, where Z is an $N(0, 1)$ random variable. Here, $\mathbb{P}(\lambda \in I_n) \to 1 - \alpha$ as $n \to \infty$.

On the other hand, if it is difficult to sample directly from π then the above classical i.i.d./ MC method cannot be used to estimate λ. In the pioneering work of [13], the target distribution π was the so-called Gibbs measure on the configuration space (a finite but large set) in statistical mechanics, but it was difficult to generate an i.i.d. sample from this. In [13] a Markov chain $\{X_n\}_{n\geq 0}$ was constructed that was appropriately *irreducible* and had π as its stationary distribution. The authors used a law of large numbers for such chains, that asserts that if $\{X_n\}_{n\geq 0}$ is a suitably irreducible Markov chain and has a probability measure π as its invariant distribution, then for any initial distribution of X_0, the 'time average' $\sum_{j=1}^n f(X_j)/n$ converges almost surely (a.s.) to the 'space average' $\lambda = \int_S f \,\mathrm{d}\pi$ as $n \to \infty$; see [14, Theorem 17.0.1]. So $\sum_{j=1}^n f(X_j)/n$ provides a consistent estimator of λ. In the late 1980s and early 1990s a

Received 23 May 2014; revision received 19 September 2014.
* Postal address: Department of Mathematics, Iowa State University, Ames, IA 50011, USA.
** Postal address: Department of Statistics, Iowa State University, Ames, IA 50011, USA.
Email address: vroy@iastate.edu

https://doi.org/10.1239/jap/1450802757 Published online by Cambridge University Press

number of statisticians became aware of the work of [13] and adapted it to solve some statistical problems. Thus, a new statistical method (for estimating integrals with respect to probability distributions) known as the Markov chain Monte Carlo (MCMC) method was born. Since then the subject has exploded in both theory and applications; see, e.g. [16]. Here also, if $\int_S f^2 \,\mathrm{d}\pi < \infty$ then under certain conditions on mixing rates of the chain $\{X_n\}_{n\geq 0}$, a CLT is available for the time average estimator $\sum_{j=1}^{n} f(X_j)/n$, from which a confidence interval estimate for λ can be produced.

Recently, [2] have shown that the standard Monte Carlo (both i.i.d. MC and MCMC) methods are not applicable for estimating λ in the case of improper targets, that is, when $\pi(S) = \infty$. In particular, the authors showed that the usual time average estimator, $\sum_{i=1}^{n} f(X_i)/n$, based on a recurrent Markov chain $\{X_n\}_{n\geq 0}$ with invariant measure π (with $\pi(S) = \infty$) converges to 0 with probability 1 and, hence, is inappropriate. The authors provided consistent estimators of λ based on *regenerative sequences* of random variables whose canonical measure is π.

A sequence of random variables is *regenerative* if it probabilistically restarts itself at random times and can thus be broken up into i.i.d. pieces. Below is the formal definition of regenerative sequences.

Definition 1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(S, \mathcal{S})$ be a measurable space. A sequence of random variables $\{X_n\}_{n\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(S, \mathcal{S})$ is called *regenerative* if there exists a sequence of (random) times $0 < T_1 < T_2 < \cdots$ such that the excursions $\{X_n\colon T_j \leq n < T_{j+1}, \tau_j\}_{j\geq 1}$ are i.i.d., where $\tau_j = T_{j+1} - T_j$ for $j = 1, 2, \ldots$, that is,

$$\begin{aligned}\mathbb{P}(\tau_j = k_j, X_{T_j+q} \in A_{q,j}, 0 \leq q < k_j, j = 1, \ldots, r)\\ = \prod_{j=1}^{r} \mathbb{P}(\tau_1 = k_j, X_{T_1+q} \in A_{q,j}, 0 \leq q < k_j)\end{aligned}$$

for all $k_1, \ldots, k_r \in \mathbb{N}$, the set of positive integers, $A_{q,j} \in \mathcal{S}, 0 \leq q < k_j, j = 1, \ldots, r$, and $r \geq 1$ and these are independent of the initial excursion $\{X_j\colon 0 \leq j < T_1\}$. The random times $\{T_n\}_{n\geq 1}$ are called *regeneration times*.

The standard example of a regenerative sequence is a Markov chain that is suitably irreducible and recurrent. A regenerative sequence need not have the Markov property. In particular, it need not be a Markov chain (see [2] for examples). Let

$$\pi(A) \equiv \mathbb{E}\bigg(\sum_{j=T_1}^{T_2-1} \mathbf{1}_A(X_j)\bigg) \quad \text{for } A \in \mathcal{S}, \tag{1}$$

where $\mathbf{1}_A$ is the indicator function for some set A. The measure π is called the *canonical* (or, *occupation*) measure for regenerative sequence $\{X_n\}_{n\geq 0}$ with regeneration times $\{T_n\}_{n\geq 0}$. Let $N_n = k$ if $T_k \leq n < T_{k+1}, k, n = 1, 2, \ldots$. That is, N_n denotes the number of regenerations by time n. Athreya and Roy [2] showed that the following estimator $\hat{\lambda}_n$, called the *regeneration estimator* for estimating $\lambda \equiv \int_S f \,\mathrm{d}\pi$ (assuming $\int_S |f| \,\mathrm{d}\pi < \infty$) is indeed consistent. That is,

$$\hat{\lambda}_n \equiv \frac{\sum_{j=0}^{n} f(X_j)}{N_n} \to \lambda \quad \text{a.s.} \tag{2}$$

Thus, given a (proper or improper) measure π, if we can find a regenerative sequence with π as its canonical measure, then $\lambda \equiv \int_S f \,\mathrm{d}\pi$ can be estimated by (2). It may be noted that this regenerative sequence Monte Carlo (RSMC) works whether $\pi(S)$ is infinite or finite. If $\pi(S) < \infty$ then the strong law of large numbers implies that N_n, the number of regenerations by time n, grows at the rate $n/\pi(S)$ (since $\mathbb{E}(T_2 - T_1) = \pi(S)$) as $n \to \infty$. Thus, when $\pi(S)$ is finite we have at least three choices of Monte Carlo methods of estimation, namely the i.i.d. MC, MCMC, and RSMC. This last Monte Carlo method, that is, RSMC has a natural universality property, namely, it works whether one knows the target π is a finite or infinite measure.

The regenerative property of positive recurrent Markov chains has been used in the MCMC literature for calculating standard errors of MCMC-based estimates for integrals with respect to a probability distribution; see, for example, [1, Section IV.4], [10], [15], and [16, Chapter 12]. Regenerative methods for analyzing simulation-based output also have a long history in the operations research literature; see, for example, [6] and [9]. But in these methods, the excursion time τ_1 is assumed to have finite second moment, which does not hold when the target distribution is improper. In fact, $\mathbb{E}(\tau_1) = \mathbb{E}(T_2 - T_1) = \pi(S) = \infty$ when π is improper. On the other hand, the RSMC method does not require the existence of even the first moment of τ_1.

Athreya and Roy [2] developed algorithms based mainly on random walks for estimating λ when S is countable as well as $S = \mathbb{R}^d$ for *any* $d \geq 1$. This leads to the very important question of how to construct a confidence interval for λ based on $\hat{\lambda}_n$. An approximate distribution of $(\hat{\lambda}_n - \lambda)$ can be used for estimating the Monte Carlo error of the regeneration estimator $\hat{\lambda}_n$. In this paper we obtain an asymptotic confidence interval for λ under the assumption of finite second moments of $\sum_{i=T_1}^{T_2-1} f(X_i)$ (this is not the same as requiring $\mathbb{E}(T_2 - T_1)^2 < \infty$) and a regularly varying tail of the distribution of the regeneration time τ_1. We make use of a deep result due to [12] in order to obtain the limiting distribution of (suitably normalized) $(\hat{\lambda}_n - \lambda)$. We then apply our general results to the algorithms based on the simple symmetric random walk (SSRW) on $\mathbb{Z}$ presented in [2] for the case when S is countable as well as $S = \mathbb{R}^d$ for some $d \geq 1$. We provide simple conditions on f under which a confidence interval based on $\hat{\lambda}_n$ is available in both cases, that is, when S is countable or $S = \mathbb{R}^d$ and π is absolutely continuous. The algorithms based on the SSRW [2] are used for estimating λ.

2. Main results

Theorem 1. *Let $\{X_n\}_{n\geq 0}$ be a regenerative sequence as in Definition 1. Let π be its canonical measure as defined in (1). Let $f\colon S \to \mathbb{R}$ be $\mathcal{S}$ measurable.*

(i) *Assume that*

$$\mathbb{E}\bigg(\sum_{j=T_1}^{T_2-1} |f(X_j)|\bigg) = \int_S |f| \,\mathrm{d}\pi < \infty, \qquad \mathbb{E}\bigg(\sum_{j=T_1}^{T_2-1} f(X_j)\bigg)^2 < \infty.$$

Let $U_i = \sum_{j=T_i}^{T_{i+1}-1} f(X_j)$, $i = 1, 2, 3, \ldots$, and $0 < \sigma^2 \equiv \mathbb{E}U_1^2 - \lambda^2 < \infty$. Let

$$Y_k = \frac{\sum_{j=1}^{k}(U_j - \lambda)}{\sigma\sqrt{k}} \quad \textit{for } k = 1, 2, \ldots.$$

Then

(a) *as $k \to \infty$,*

$$Y_k \xrightarrow{\mathrm{D}} N(0,1).$$

(b) *Let* $Y_k(t), t \geq 0$ *be the linear interpolation of* Y_k *on* $[0, \infty)$*, that is,*

$$Y_k(t) \equiv Y_{[kt]} + (kt - [kt])\frac{(U_{[kt]+1} - \lambda)}{\sqrt{k}\sigma}.$$

Then as $k \to \infty$, $\{Y_k(t) : t \geq 0\} \xrightarrow{\text{D}} \{B(t) : t \geq 0\}$, *in* $C[0, \infty)$, *where* $\{B(t) : t \geq 0\}$ *is the standard Brownian motion.*

(ii) *Assume that*

$$\mathbb{P}(\tau_1 > x) \sim x^{-\alpha} L(x) \quad \text{as } x \to \infty, \tag{3}$$

where $0 < \alpha < 1$ *and* $L(\cdot)$ *is slowly varying, i.e. for all* $0 < c < \infty$, $L(cx)/L(x) \to 1$ *as* $x \to \infty$. *Then*

(a) $\pi(S) = \mathbb{E}(T_2 - T_1) = \infty$.

(b) *Let* $N_n = k$ *if* $T_k \leq n < T_{k+1}, k \geq 1, n \geq 1$. *Then*

$$\frac{N_n}{n^\alpha / L(n)} \xrightarrow{\text{D}} V_\alpha \quad \text{as } n \to \infty, \tag{4}$$

where $\mathbb{P}(V_\alpha > 0) = 1$, *and for* $0 < x < \infty$, $\mathbb{P}(V_\alpha \leq x) = \mathbb{P}(\tilde{V}_\alpha \geq x^{-1/\alpha})$ *with* $\tilde{V}_\alpha$ *being a positive random variable with a stable distribution with index* α *such that*

$$\mathbb{E}(\exp(-s\tilde{V}_\alpha)) = \exp(-s^\alpha \Gamma(1 - \alpha)), \qquad 0 \leq s < \infty$$

where $\Gamma(p) = \int_0^\infty x^{p-1} \exp(-x) \, \mathrm{d}x$, $0 < p < \infty$ *is the gamma function.*

(iii) *Assume that*

$$\mathbb{E}\biggl(\sum_{j=T_1}^{T_2-1} |f(X_j)|\biggr)^2 < \infty \tag{5}$$

and (3) holds. Then

$$\frac{(\hat{\lambda}_n - \lambda)\sqrt{N_n}}{\sigma} \xrightarrow{\text{D}} N(0, 1) \quad \text{as } n \to \infty, \tag{6}$$

$$\frac{(\hat{\lambda}_n - \lambda)\sqrt{n^\alpha / L(n)}}{\sigma} \xrightarrow{\text{D}} Q \quad \text{as } n \to \infty, \tag{7}$$

where $Q \equiv B(V_\alpha)/V_\alpha$, $\{B(t) : t \geq 0\}$ *is the standard Brownian motion and* V_α *is as in (4) and independent of* $\{B(t) : t \geq 0\}$.

(a) *Let*

$$\sigma_n^2 = \sum_{i=1}^{N_n} \frac{U_i^2}{N_n} - \hat{\lambda}_n^2 \tag{8}$$

Then

$$\frac{(\hat{\lambda}_n - \lambda)\sqrt{N_n}}{\sigma_n} \xrightarrow{\text{D}} N(0, 1) \quad \text{as } n \to \infty, \tag{9}$$

$$\frac{(\hat{\lambda}_n - \lambda)\sqrt{n^\alpha / L(n)}}{\sigma_n} \xrightarrow{\text{D}} Q \quad \text{as } n \to \infty, \tag{10}$$

where Q *is as in (7).*

https://doi.org/10.1239/jap/1450802757 Published online by Cambridge University Press

Using the results in Theorem 1 we can construct asymptotic confidence intervals for λ based on the regenerative sequence $\{X_i\}_{i=0}^n$ as discussed below in Corollary 1.

Corollary 1. *Fix* $0 < p < 1$, *and let* z_p, q_p *be such that* $\mathbb{P}(|Z| > z_p) = p$ *and* $\mathbb{P}(Q > q_p) = p$, *where* $Z \sim N(0, 1)$ *and* Q *is as in (7). Let* $I_{n1} \equiv (\hat{\lambda}_n - z_p\sigma_n/\sqrt{N_n}, \hat{\lambda}_n + z_p\sigma_n/\sqrt{N_n})$, *and* $I_{n2} \equiv (\hat{\lambda}_n - q_{p/2}\sigma_n/\sqrt{n^\alpha/L(n)}, \hat{\lambda}_n - q_{1-p/2}\sigma_n/\sqrt{n^\alpha/L(n)})$. *Let* $l(I_{n1}) = 2z_p\sigma_n/\sqrt{N_n}$, *and* $l(I_{n2}) = (q_{p/2} - q_{1-p/2})\sigma_n/\sqrt{n^\alpha/L(n)}$ *be the lengths of the intervals* I_{n1} *and* I_{n2}, *respectively. Then we have the following:*

(i) $\mathbb{P}(\lambda \in I_{ni}) \to 1 - p$ *as* $n \to \infty$ *for* $i = 1, 2$;

(ii) $\sqrt{n^\alpha/L(n)}\, l(I_{n1}) \xrightarrow{\text{D}} 2z_p\sigma/\sqrt{V_\alpha}$, *where* V_α *is as in (7);*

(iii) $\sqrt{n^\alpha/L(n)}\, l(I_{n2}) \to (q_{p/2} - q_{1-p/2})\sigma$ *a.s.*

Below we consider a special case of Theorem 1 in the case when S is countable. We use an algorithm [2, Algorithm I] based on the SSRW on $\mathbb{Z}$ for consistently estimating countable sums. We provide a simple sufficient condition for the second moment hypothesis (5) in this case so that we can obtain confidence interval as well. Since S is countable, without loss of generality, we can take $S = \mathbb{Z}$ in this case.

Theorem 2. *Let* $\{X_n\}_{n\geq 0}$ *be an SSRW on* $\mathbb{Z}$ *starting at* $X_0 = 0$. *That is,*

$$X_{n+1} = X_n + \delta_{n+1}, \qquad n \geq 0,$$

where $\{\delta_n\}_{n\geq 1}$ *are i.i.d. with distribution* $\mathbb{P}(\delta_1 = +1) = \frac{1}{2} = \mathbb{P}(\delta_1 = -1)$ *and independent of* X_0. *Let* $N_n = \sum_{j=0}^n \mathbf{1}_{\{X_j=0\}}$ *be the number of visits to* 0 *by* $\{X_j\}_{j=0}^n$. *Assume that* $\pi_i \equiv \pi(i) \geq 0$ *for all* i. *Let* $f\colon \mathbb{Z} \to \mathbb{R}$ *be such that* $\sum_{j\in\mathbb{Z}} |f(j)|\pi(j) < \infty$. *Then*

$$\hat{\lambda}_n \equiv \frac{\sum_{j=0}^n f(X_j)\pi(X_j)}{N_n} \to \lambda \equiv \sum_{i\in\mathbb{Z}} f(i)\pi_i \quad \text{as } n \to \infty \text{ a.s.} \tag{11}$$

Assume that, in addition, $\sum_{j\in\mathbb{Z}} |f(j)|\pi(j)\sqrt{|j|} < \infty$. *Then*

(i) $\mathbb{E}(\sum_{j=0}^{T_1-1} |f(X_j)|\pi(X_j))^2 < \infty$, *where* $T_1 = \min\{n\colon n \geq 1, X_n = 0\}$,

(ii) *and*

$$\frac{\sqrt{N_n}(\hat{\lambda}_n - \lambda)}{\sigma} \xrightarrow{\text{D}} N(0, 1) \quad \text{as } n \to \infty, \tag{12}$$

where $\sigma^2 \equiv \mathbb{E}(\sum_{j=0}^{T_1-1} f(X_j)\pi(X_j))^2 - \lambda^2$,

(iii) *then*

$$\frac{(\hat{\lambda}_n - \lambda)n^{1/4}}{\sigma} \xrightarrow{\text{D}} \frac{B(V_{1/2})}{V_{1/2}} \quad \text{as } n \to \infty, \tag{13}$$

where $\{B(t)\colon t \geq 0\}$ *is the standard Brownian motion,* $V_{1/2}$ *is a random variable independent of* $\{B(t)\colon t \geq 0\}$, *and* $V_{1/2}$ *has the same distribution as* $\sqrt{\pi/2}|Z|$, $Z \sim N(0, 1)$.

(iv) *Also the analogues of (9), (10), and Corollary 1 hold.*

Lastly, we consider the algorithm presented in [2, Algorithm III] that is based on the SSRW on $\mathbb{Z}$ and a randomization tool to estimate integrals with respect to an absolutely continuous measure π on *any* $\mathbb{R}^d$, $d < \infty$. Let $f: \mathbb{R}^d \to \mathbb{R}$ and π be an absolutely continuous measure on $\mathbb{R}^d$ with Radon–Nikodym derivative $p(\cdot)$. Assume that $\int_{\mathbb{R}^d} |f(x)| p(x)\,\mathrm{d}x < \infty$. The following theorem provides a consistent estimator as well as an interval estimator of $\lambda \equiv \int_{\mathbb{R}^d} f(x)p(x)\,\mathrm{d}x$.

Theorem 3. *Let $\{X_n\}_{n\ge 0}$ be a SSRW on $\mathbb{Z}$ with $X_0 = 0$. Let $\{U_{ij}: i = 0, 1, \ldots; j = 1, 2, \ldots, d\}$ be a sequence of i.i.d. uniform $(-\frac{1}{2}, \frac{1}{2})$ random variables and independent of $\{X_n\}_{n\ge 0}$. Assume that $\kappa: \mathbb{Z} \to \mathbb{Z}^d$ be $1-1$, onto. Define $W_n := \kappa(X_n) + U^n$, $n = 0, 1, \ldots,$ where $U^n = (U_{n1}, U_{n2}, \ldots, U_{nd})$. Note that the sequence $\{W_n\}_{n\ge 0}$ is regenerative with regeneration times $\{T_n\}_{n\ge 0}$ being the returns of SSRW $\{X_n\}_{n\ge 0}$ to 0. Then*

$$\hat{\lambda}_n \equiv \frac{\sum_{j=0}^n f(W_j)p(W_j)}{N_n} \to \lambda \equiv \int_{\mathbb{R}^d} f(x)p(x)\,\mathrm{d}x \quad a.s., \tag{14}$$

where $N_n = \sum_{j=0}^n \mathbf{1}_{\{X_j=0\}}$ is the number of visits to 0 by $\{X_j\}_{j=0}^n$.

Let $g: \mathbb{Z} \to \mathbb{R}_+$ be defined as

$$\begin{aligned} g(r) &\equiv \sqrt{\mathbb{E}(|f(\kappa(r)+U)|p(\kappa(r)+U))^2} \\ &= \left(\int_{[-1/2,1/2]^d} (|f(\kappa(r)+u)|p(\kappa(r)+u))^2\,\mathrm{d}u\right)^{1/2}, \end{aligned} \tag{15}$$

where $U = (U_1, U_2, \ldots, U_d)$ with the U_is, $i = 1, 2, \ldots, d$, are i.i.d. uniform $(-\frac{1}{2}, \frac{1}{2})$ random variables, and $[-\frac{1}{2}, \frac{1}{2}]^d$ is the d-dimensional rectangle with each side being $[-\frac{1}{2}, \frac{1}{2}]$. Assume that $\sum_{r\in\mathbb{Z}} g(r)\sqrt{|r|} < \infty$. Then

(i) $\mathbb{E}(\sum_{j=0}^{T_1-1} |f(W_j)|\pi(W_j))^2 < \infty$, *where $T_1 = \min\{n: n \ge 1, X_n = 0\}$,*

(ii) *and*

$$\frac{\sqrt{N_n}(\hat{\lambda}_n - \lambda)}{\sigma} \to N(0,1) \quad as\ n \to \infty,$$

where $\sigma^2 \equiv \mathbb{E}(\sum_{j=0}^{T_1-1} f(W_j)\pi(W_j))^2 - \lambda^2$,

(iii) *then*

$$\frac{(\hat{\lambda}_n - \lambda)n^{1/4}}{\sigma} \to \frac{B(V_{1/2})}{V_{1/2}} \quad as\ n \to \infty,$$

where $\{B(t): t \ge 0\}$, and $V_{1/2}$ are as in (13).

(iv) *Also the analogues of (9), (10), and Corollary 1 hold.*

Remark 1. A sufficient condition for $\sum_{r\in\mathbb{Z}} g(r)\sqrt{|r|} < \infty$ in Theorem 3 is as follows. Let

$$h(r) = \sup_{u\in[-1/2,1/2]^d} |f(\kappa(r)+u)|p(\kappa(r)+u).$$

From (15), it follows that $g(r) \le h(r)$ for all $r \in \mathbb{Z}$ and so a sufficient condition for $\mathbb{E}(\sum_{j=0}^{T_1-1} |f(W_j)|\pi(W_j))^2 < \infty$ is $\sum_{r\in\mathbb{Z}} h(r)\sqrt{|r|} < \infty$.

The proofs of Theorems 1–3 are given in Section 4. The proof of Corollary 1 follows from the proof of Theorem 1 and Slutsky's theorem and, hence, is omitted.

3. Examples

In this section we demonstrate the use of the results in Section 2 with some examples. We first consider estimating $\lambda = \sum_{m=1}^{\infty} 1/m^2$. The SSRW chain mentioned in Theorem 2 was used in [2] to estimate λ, that is, the authors used $\hat{\lambda}_n$ defined in (11) to consistently estimate λ. Note that, in this case $f(j) = 1/j^2$ if $j \geq 1$, $f(j) = 0$ otherwise, and $\pi(j) = 1$ for all $j \in \mathbb{Z}$. Since $\sum_{j\in\mathbb{Z}} f(j)\pi(j)\sqrt{|j|} = \sum_{j\geq 1} j^{-(1+1/2)} < \infty$, we can use Theorem 2 to provide a confidence interval for λ based on $\hat{\lambda}_n$. In particular, an asymptotic 95% confidence interval for λ is given by $(\hat{\lambda}_n \pm 1.96\sigma_n/\sqrt{N_n})$, where σ_n^2 is defined in (8). In Figure 1a we show the point as well as the 95% interval estimates for six values ($\log_{10} n = 3, 4, \ldots, 8$, where $\log_{10}$ denotes logarithm base 10). The point and 95% interval estimates for $n = 10^8$ are 1.636 and (1.580, 1.693), respectively. Note that, the true value λ is $\pi^2/6 = 1.645$. The time taken to run the SSRW chain for 10^8 steps using R ([17]) on an Intel® Core™2 Q9550 2.83GHz machine running Windows® 7 is about three seconds.

The next example was originally considered in [5]. Let

$$f(x, y) = \exp(-xy), \qquad 0 < x, y < \infty. \tag{16}$$

Let

$$f_{X\mid Y}(x \mid y) := \frac{f(x, y)}{\int_{\mathbb{R}_+} f(x', y)\,\mathrm{d}x'} = y\exp(-xy), \qquad 0 < x, y < \infty.$$

Thus, for each y, the conditional density of X given $Y = y$ is an exponential density. Consider the Gibbs sampler $\{(X_n, Y_n)\}_{n\geq 0}$ that uses the two conditional densities $f_{X\mid Y}(\cdot \mid y)$ and $f_{Y\mid X}(\cdot \mid x)$, alternately. Casella and George [5] found that the usual estimator

$$\sum_{j=0}^{n} \frac{f_{X\mid Y}(\check{x} \mid Y_j)}{n}$$

for the marginal density $f_X(\check{x}) = \int_{\mathbb{R}_+} f(\check{x}, y)\,\mathrm{d}y = \int_{\mathbb{R}_+} f_{X\mid Y}(\check{x} \mid y) f_Y(y)\,\mathrm{d}y = 1/\check{x}$ breaks down. It was shown in [2] that the Gibbs sampler $\{(X_n, Y_n)\}_{n\geq 0}$ is regenerative with improper invariant measure whose density with respect to the Lebesgue measure is $f(x, y)$ as defined in (16). Thus, [2, Theorem 3] implies that $\sum_{j=0}^{n} f_{X\mid Y}(\check{x} \mid Y_j)/n$ converges to 0 with probability 1. In [2], the authors used $\hat{\lambda}_n$ defined in (14) for consistently estimating

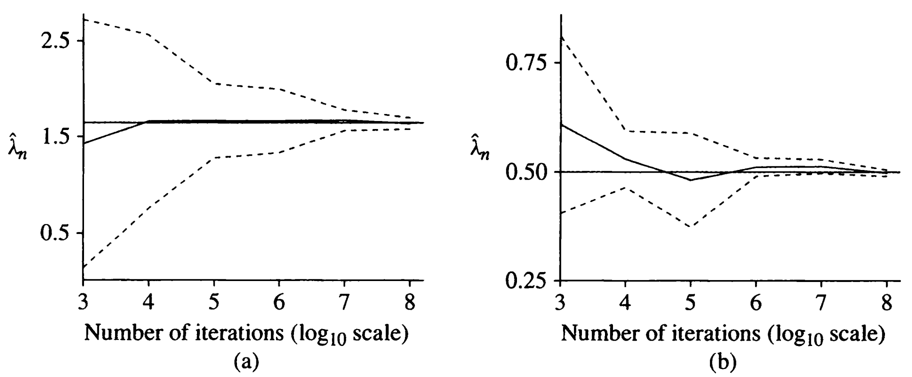

FIGURE 1: Point and 95% interval estimates of (a) $\sum_{m=1}^{\infty} 1/m^2$ and (b) $f_X(0.5)$.

$f_X(\check{x})$. In this example, using Remark 1, we have $h(r) = 0$ for all $r \leq -1$, $h(0) = 1$, and $h(r) = \exp(-\check{x}(r - \frac{1}{2}))$ for all $r \geq 1$. Since

$$\sum_{r \in \mathbb{Z}} g(r)\sqrt{|r|} \leq \sum_{r \geq 1} \exp(-\check{x}(r - \tfrac{1}{2}))\sqrt{r} < \infty,$$

from Theorem 3, we obtain a confidence interval for $f_X(\check{x})$ based on $\hat{\lambda}_n$. In Figure 1(b) we show the (point and 95% interval) estimates of $f_X(2) = \frac{1}{2}$ for the same six n values mentioned in the previous example. The estimates for $n = 10^8$ are 0.497 and (0.490, 0.505), respectively.

4. Proofs of results

We begin with a short lemma that is used in the proof of Theorem 1.

Lemma 1. *Let $\{\xi_i\}_{i \geq 1}$ be i.i.d. random variables with $\mathbb{E}|\xi_1| < \infty$. Then $\xi_n / n \to 0$ a.s.*

Proof. Since $\mathbb{E}|\xi_1| < \infty$, for all $\varepsilon > 0$, $\sum_{n=1}^{\infty} \mathbb{P}(|\xi_1| > \varepsilon n) < \infty$. By the Borel–Cantelli lemma, $\sum_{n=1}^{\infty} \mathbb{P}(|\xi_n| > \varepsilon n) < \infty$ implies that $\mathbb{P}(|\xi_n|/n > \varepsilon \text{ i.o.}) = 0$, where i.o. stands for infinitely often. This implies that $\limsup |\xi_n|/n \leq \varepsilon$ with probability 1 for all $\varepsilon > 0$. This in turn implies that $\xi_n/n \to 0$ a.s.

Proof of Theorem 1. From (1) it follows that $\lambda = \mathbb{E}(U_1)$. Since the U_is are i.i.d. random variables with $\operatorname{var}(U_1) = \sigma^2$, Theorem 1(i)(a) and Theorem 1(i)(b) follow from the classical CLT and the functional CLT for i.i.d. random variables; see [4].

From (1), we have $\pi(S) = \mathbb{E}(T_2 - T_1)$. Since (3) holds and $0 < \alpha < 1$, $\mathbb{E}(T_2 - T_1) = \mathbb{E}(\tau_1) = \infty$ implying Theorem 1(ii)(a).

The proof of (4) is given in [8] (see also [2] and [12]).

Now we establish (6). Note that

$$\frac{(\hat{\lambda}_n - \lambda)\sqrt{N_n}}{\sigma} = \frac{\sum_{j=0}^{T_1 - 1} f(X_j)}{\sqrt{N_n}\sigma} + \frac{\sum_{i=1}^{N_n}(U_i - \lambda)}{\sqrt{N_n}\sigma} + \frac{\sum_{j=T_{N_n}}^{n} f(X_j)}{\sqrt{N_n}\sigma}. \tag{17}$$

Now since $\mathbb{P}(T_1 < \infty) = 1$, $\mathbb{P}(|\sum_{j=0}^{T_1-1} f(X_j)| < \infty) = 1$. Also, $N_n \to \infty$ with probability 1 as $n \to \infty$ and $0 < \sigma < \infty$. This implies that $\sum_{j=0}^{T_1-1} f(X_j)/\sqrt{N_n}\sigma \to 0$ a.s. Next,

$$\frac{|\sum_{j=T_{N_n}}^{n} f(X_j)|}{\sqrt{N_n}\sigma} \leq \frac{\sum_{j=T_{N_n}}^{T_{N_n+1}-1} |f(X_j)|}{\sqrt{N_n}\sigma} \equiv \frac{\eta_{N_n}}{\sqrt{N_n}\sigma},$$

where $\eta_i \equiv \sum_{j=T_i}^{T_{i+1}-1} |f(X_j)|$, $i = 1, 2, \ldots$. Since the condition (5) is in force, we have $\mathbb{E}\eta_1^2 < \infty$. By Lemma 1, $\eta_n^2/n \to 0$ a.s. This implies that $\eta_n/\sqrt{n} \to 0$ a.s. Since $N_n \to \infty$ a.s., as $n \to \infty$, we have $\eta_{N_n}/\sqrt{N_n} \to 0$ a.s.

So to establish (6), it suffices to show that

$$\frac{\sum_{i=1}^{N_n}(U_i - \lambda)}{\sqrt{N_n}\sigma} \xrightarrow{\text{D}} N(0, 1).$$

Let, for $0 \leq t < \infty$,

$$B_n(t) \equiv \frac{\sum_{i=1}^{[nt]}(U_i - \lambda)}{\sqrt{n}\sigma} + (nt - [nt])\frac{(U_{[nt]+1} - \lambda)}{\sqrt{n}\sigma} \tag{18}$$

https://doi.org/10.1239/jap/1450802757 Published online by Cambridge University Press

and

$$A_n(t) \equiv \frac{T_{[nt]}}{a_n} + (nt - [nt])\frac{\tau_{[nt]}}{a_n},$$

where $\{a_n\}$ is such that $na_n^{-\alpha}L(a_n) \to 1$. Then, it is known (see [4]) by Donsker's invariance principle that $\{B_n(\cdot)\colon 0 \le t < \infty\}$ converges weakly in $C[0,\infty)$ as $n \to \infty$ to a standard Brownian motion $B(\cdot)$. Also, it is known (see [8, p. 448]) that for any $0 < t_1 < t_2 < \cdots < t_k < \infty$, $(A_n(t_1), A_n(t_2), \dots, A_n(t_k))$ convergence in distribution as $n \to \infty$ to $(A(t_1), A(t_2), \dots, A(t_k))$, where $\{A(t)\colon t \ge 0\}$ is a nonnegative stable process of order α with $A(0) = 0$, a.s. and $\mathbb{E}(\exp(-sA(1))) = \exp(-s^\alpha\Gamma(1-\alpha)), 0 \le s < \infty$. It has been pointed out in [12, p. 525] that [18] has shown that this finite-dimensional convergence of $A_n(\cdot)$ to $A(\cdot)$ implies the convergence in law in the Skorokhod space $D[0,\infty)$. Next, it can be shown that $(A_n(\cdot), B_n(\cdot))$ converges in the sense of finite-dimensional distributions as $n \to \infty$. Since both $\{A_n(\cdot)\}_{n\ge1}$ and $\{B_n(\cdot)\}_{n\ge1}$ converge weakly in $D[0,\infty)$ (as pointed out above) both are tight. This implies that the bivariate sequence $\{A_n(\cdot), B_n(\cdot)\}_{n\ge1}$ is also tight as processes in $D^2[0,\infty) \equiv D[0,\infty) \times D[0,\infty)$. Since the finite-dimensional distributions of $(A_n(\cdot), B_n(\cdot))$ converge as $n \to \infty$, this yields the weak convergence of $\{A_n(\cdot), B_n(\cdot)\}_{n\ge1}$ as $n \to \infty$ in $D^2[0,\infty)$. For the limit process $(A(\cdot), B(\cdot))$, we conclude that the process $C(\cdot) = A(\cdot) + B(\cdot)$ is a Lévy process on $[0,\infty)$. Now, since $B(\cdot)$ has continuous trajectory and $A(\cdot)$ has strictly increasing nonnegative sample paths, it follows by the uniqueness of the Lévy–Itô decomposition of $C(\cdot)$ that the processes $A(\cdot)$ and $B(\cdot)$ have to be independent. This argument is due to [12].

As noted by [18] (see also [4]) it is possible to produce a sequence of processes $(\tilde{A}_n(\cdot), \tilde{B}_n(\cdot))$ and a process $(\tilde{A}(\cdot), \tilde{B}(\cdot))$ all defined in the same probability space such that for each n, $(\tilde{A}_n(\cdot), \tilde{B}_n(\cdot))$ has the same distribution as $(A_n(\cdot), B_n(\cdot))$ on $D^2[0,\infty)$, and $(\tilde{A}(\cdot), \tilde{B}(\cdot))$ has the same distribution as $(A(\cdot), B(\cdot))$, and $(\tilde{A}_n(\cdot), \tilde{B}_n(\cdot))$ converges to $(\tilde{A}(\cdot), \tilde{B}(\cdot))$ with probability 1 in $D^2[0,\infty)$. More specifically, we can generate on the same probability space sequences $\{\tilde{U}_{n,i}\}_{i\ge1,n\ge1}$ and $\{\tilde{T}_{n,i}\}_{i\ge1,n\ge1}$ such that for each n, the sequence $\{\tilde{U}_{n,i}, \tilde{T}_{n,i}\}_{i\ge1}$ has the same distribution as $\{U_i, T_i\}_{i\ge1}$ and for each n, the processes $\tilde{A}_n(\cdot)$ and $\tilde{B}_n(\cdot)$ are defined in terms of the sequence $\{\tilde{U}_{n,i}, \tilde{T}_{n,i}\}_{i\ge1}$ and another sequence $\{\tilde{U}_i, \tilde{T}_i\}_{i\ge1}$ also having the same distribution as $\{U_i, T_i\}_{i\ge1}$ such that $(\tilde{A}(\cdot), \tilde{B}(\cdot))$ is defined using $\{\tilde{U}_i, \tilde{T}_i\}_{i\ge1}$.

Next, let $A_n^{-1}(\cdot)$ and $A^{-1}(\cdot)$ be the inverses of the nondecreasing nonnegative functions $A_n(\cdot), A(\cdot)$ from $[0,\infty)$ to $[0,\infty)$. (For a nondecreasing nonnegative function H on $[0,\infty)$, we define the inverse $H^{-1}(\cdot)$ by $H^{-1}(y) \equiv \inf\{x\colon H(x) \ge y\}, 0 \le y < \infty$.) It can be shown (see also [11, Theorem A.1]) that $(\tilde{A}_n, \tilde{A}_n^{-1}, \tilde{B}_n)$ converges to $(\tilde{A}, \tilde{A}^{-1}, \tilde{B})$ with probability 1 in $D^3[0,\infty)$. This, in turn, yields by the continuous mapping theorem and the fact that $\mathbb{P}(\tilde{A}^{-1}(1) > 0) = 1$, as $n \to \infty$,

$$\frac{\tilde{B}_n(\tilde{A}_n^{-1}(1))}{\sqrt{\tilde{A}_n^{-1}(1)}} \to \frac{\tilde{B}(\tilde{A}^{-1}(1))}{\sqrt{\tilde{A}^{-1}(1)}} \quad \text{a.s.} \tag{19}$$

Now by the independence of $\tilde{B}$ and $\tilde{A}$ and the fact that $\mathbb{P}(\tilde{A}^{-1}(1) > 0) = 1$, the limiting random variable on the right-hand side of (19) is distributed as $N(0, 1)$.

Let $b_n \uparrow \infty$ be a sequence such that $a_{b_n}/n \to 1$ as $n \to \infty$, where $\{a_n\}$ is as defined earlier, satisfies $na_n^{-\alpha}L(a_n) \to 1$. Such a sequence $\{b_n\}$ exists as $a_n \uparrow \infty$ as $n \to \infty$. By definition

$$\tilde{A}_n^{-1}(1) = \inf\{x\colon \tilde{A}_n(x) \ge 1\}.$$

Let $y < \tilde{A}_n^{-1}(1)$, then $\tilde{A}_n(y) < 1$. This implies that $\tilde{T}_{n,[ny]}/a_n < 1$. Let $\{\tilde{N}_{n,m}\}_{m\geq 1}$ be the sequence of regeneration times associated with $\{\tilde{U}_{n,i}, \tilde{T}_{n,i}\}_{i\geq 1}$. Then, $\tilde{T}_{n,[ny]}/a_n < 1$ implies that $\tilde{N}_{n,a_n} \geq [ny] \geq ny - 1$. So, $\tilde{N}_{n,a_n}/n \geq y - 1/n$. Since this holds for all $y < \tilde{A}_n^{-1}(1)$, we have

$$\frac{\tilde{N}_{n,a_n}}{n} \geq \tilde{A}_n^{-1}(1) - \frac{1}{n}. \tag{20}$$

Similarly, letting $y > \tilde{A}_n^{-1}(1)$, we conclude that $\tilde{N}_{n,a_n}/n \leq y + 1/n$. As this holds for all $y > \tilde{A}_n^{-1}(1)$, we have

$$\frac{\tilde{N}_{n,a_n}}{n} \leq \tilde{A}_n^{-1}(1) + \frac{1}{n}. \tag{21}$$

From (20) and (21), we have

$$\tilde{A}_n^{-1}(1) - \frac{1}{n} \leq \frac{\tilde{N}_{n,a_n}}{n} \leq \tilde{A}_n^{-1}(1) + \frac{1}{n},$$

and, more specifically,

$$\tilde{A}_{b_n}^{-1}(1) - \frac{1}{b_n} \leq \frac{\tilde{N}_{b_n,a_{b_n}}}{b_n} \leq \tilde{A}_{b_n}^{-1}(1) + \frac{1}{b_n}. \tag{22}$$

Since $a_{b_n}/n \to 1$ as $n \to \infty$, for all $\varepsilon > 0$, $n(1-\varepsilon) \leq a_{b_n} \leq n(1+\varepsilon)$ for all large n. This implies that for all large n, $\tilde{N}_{b_n,n(1-\varepsilon)} \leq \tilde{N}_{b_n,a_{b_n}} \leq \tilde{N}_{b_n,n(1+\varepsilon)}$. This yields, by (22),

$$\tilde{A}_{b_n}^{-1}(1) - \frac{1}{b_n} \leq \frac{\tilde{N}_{b_n,n(1+\varepsilon)}}{b_n} = \frac{\tilde{N}_{b_n,n(1+\varepsilon)}}{b_{n(1+\varepsilon)}} \frac{b_{n(1+\varepsilon)}}{b_n}. \tag{23}$$

As $\{b_n\}$ is such that $a_{b_n}/n \to 1$ and $na_n^{-\alpha}L(a_n) \to 1$, which implies that $b_n(a_{b_n})^{-\alpha}L(a_{b_n}) \to 1$, that is, $b_n \sim a_{b_n}^{\alpha}/L(a_{b_n}) \sim n^{\alpha}/L(n)$. So $b_{n(1+\varepsilon)}/b_n \sim (1+\varepsilon)^{\alpha}L(n)/L(n(1+\varepsilon)) \to (1+\varepsilon)^{\alpha}$ as $n \to \infty$ for all $\varepsilon > -1$. Since $\tilde{A}_{b_n}^{-1}(1) \to \tilde{A}^{-1}(1)$ a.s. and (23) holds for all $\varepsilon > 0$, we conclude that

$$\tilde{A}^{-1}(1) \leq \liminf \frac{\tilde{N}_{b_n,n}}{b_n} \quad \text{with probability 1.}$$

Similarly, $\tilde{A}^{-1}(1) \geq \limsup \tilde{N}_{b_n,n}/b_n$ with probability 1 and, hence, $\lim \tilde{N}_{b_n,n}/b_n = \tilde{A}^{-1}(1)$ with probability 1.

By the definition of $(\tilde{A}_n, \tilde{B}_n)$,

$$\frac{B_{b_n}(N_n/b_n)}{\sqrt{N_n/b_n}}$$

has the same distribution as

$$\frac{\tilde{B}_{b_n}(\tilde{N}_{b_n,n}/b_n)}{\sqrt{\tilde{N}_{b_n,n}/b_n}}.$$

Since $\tilde{N}_{b_n,n}/b_n \to \tilde{A}^{-1}(1)$ a.s., $\tilde{B}_{b_n}(\cdot) \to \tilde{B}(\cdot)$ a.s. in $C[0,\infty)$, and $\tilde{B}(\cdot)$ has continuous trajectory

$$\frac{\tilde{B}_{b_n}(\tilde{N}_{b_n,n}/b_n)}{\sqrt{\tilde{N}_{b_n,n}/b_n}} \to \frac{\tilde{B}(\tilde{A}^{-1}(1))}{\sqrt{\tilde{A}^{-1}(1)}} \quad \text{a.s.}$$

https://doi.org/10.1239/jap/1450802757 Published online by Cambridge University Press

From (18), we see that

$$\frac{\sum_{i=1}^{N_n}(U_i - \lambda)}{\sqrt{N_n}\sigma} = \frac{B_{b_n}(N_n/b_n)}{\sqrt{N_n/b_n}};$$

hence, (6) is proved.

Next, to prove (7), we see from (17) it suffices to show that

$$\frac{\sum_{i=1}^{N_n}(U_i - \lambda)}{N_n\sigma}\sqrt{\frac{n^\alpha}{L(n)}} \xrightarrow{\text{D}} Q \equiv \frac{B(V_\alpha)}{V_\alpha}.$$

Applying the above embedding used to prove (6), it is enough to show that

$$\frac{\tilde{B}_{b_n}(\tilde{N}_{b_n,n}/b_n)}{\tilde{N}_{b_n,n}/b_n}\sqrt{\frac{n^\alpha}{L(n)b_n}} \to Q \quad \text{a.s.}$$

This follows from the argument used in the proof of (6) and the fact that $n^\alpha/\{L(n)b_n\} \to 1$ as $n \to \infty$.

Since by the strong law of large numbers, $\sigma_n^2 \to \sigma^2$ a.s. as $n \to \infty$, (9) and (10) follow from Slutsky's theorem, (6) and (7).

Proof of Theorem 2. The SSRW Markov chain $\{X_n\}_{n\geq 0}$ is null recurrent (see, e.g. [14, Section 8.4.3]) with the counting measure on $\mathbb{Z}$ being the unique (up to multiplicative constant) invariant measure for $\{X_n\}_{n\geq 0}$. Hence, the SSRW Markov chain $\{X_n\}_{n\geq 0}$ is regenerative with regeneration times $T_0 = 0$, $T_{r+1} = \inf\{n\colon n \geq T_r + 1, X_n = 0\}$, $r = 0, 1, 2, \dots$ and the proof of (11) follows from the strong law of large numbers; see also [2].

Let $N(j) \equiv \sum_{i=0}^{T_1-1}\mathbf{1}_{\{X_i=j\}}$ be the number of visits to the state j during the first excursion $\{X_i\}_{i=0}^{T_1-1}$ for $j \in \mathbb{Z}$. Note that $X_0 = 0$ and $N(0) = 1$. Without loss of generality, for the rest of the proof we assume that $\pi_j = 1$ for all $j \in \mathbb{Z}$. Since

$$\sum_{j=0}^{T_1-1}|f(X_j)| = \sum_{j\in\mathbb{Z}}|f(j)|N(j),$$

by Minkowski's inequality, we have

$$\mathbb{E}\bigg(\sum_{j=0}^{T_1-1}|f(X_j)|\bigg)^2 = \mathbb{E}\bigg(\sum_{j\in\mathbb{Z}}|f(j)|N(j)\bigg)^2 \leq \bigg(\sum_{j\in\mathbb{Z}}|f(j)|\sqrt{E(N(j))^2}\bigg)^2.$$

For the SSRW on $\mathbb{Z}$, it has been shown by [3] that for $r \neq 0$, $\mathbb{E}(N(r)) = 1$ and $\operatorname{var}(N(r)) = 4|r| - 2$. So

$$\mathbb{E}(N(r))^2 = \operatorname{var}(N(r)) + 1 = \begin{cases} 4|r| - 1 & \text{if } r \neq 0, \\ 1 & \text{if } r = 0. \end{cases}$$

Thus, $\sum_{j\in\mathbb{Z}}|f(j)|\sqrt{|j|} < \infty$ implies that $\mathbb{E}(\sum_{j=0}^{T_1-1}|f(X_j)|)^2 < \infty$.

Since $\mathbb{P}(T_1 > n) \sim \sqrt{2/\pi}\, n^{-1/2}$ as $n \to \infty$ (see, e.g. [7, p. 203]), from (4), we have

$$\frac{N_n}{\sqrt{n}} \xrightarrow{\text{D}} \sqrt{\frac{\pi}{2}}|Z|,$$

where $Z \sim \text{N}(0, 1)$; see, e.g. [8, p. 173].

Then (12), (13), and Theorem 2(iv) follow from (6), (7), and Theorem 1(iii)(c).

Proof of Theorem 3. The proof of (14) is given in [2]. We now show that $\sum_{r\in\mathbb{Z}} g(r)\sqrt{|r|} < \infty$ implies that $\mathbb{E}(\sum_{j=0}^{T_1-1}|f(W_j)|\pi(W_j))^2 < \infty$. Without loss of generality, we assume that $p(x) \equiv 1$ for all $x \in \mathbb{R}^d$. Since the $\{U_{ij} : i = 0, 1, \ldots; j = 1, 2, \ldots, d\}$ are i.i.d. uniform $(-\frac{1}{2}, \frac{1}{2})$ and are independent of $\{X_n\}_{n\geq 0}$, we have

$$\mathbb{E}\bigg(\sum_{j=0}^{T_1-1} |f(W_j)|\bigg)^2 = \mathbb{E}\bigg(\sum_{r\in\mathbb{Z}}\sum_{i=1}^{N(r)} |f(\kappa(r) + U^i)|\bigg)^2,$$

where $N(r)$ is as defined in the proof of Theorem 2, the number of visits to the state r during the first excursion $\{X_i\}_{i=0}^{T_1-1}$ and $U^i \equiv (U_{i1}, U_{i2}, \ldots, U_{id})$, with the U_{ij} being i.i.d. uniform$(-\frac{1}{2}, \frac{1}{2})$. By Minkowski's inequality, we have

$$\bigg\{\mathbb{E}\bigg(\sum_{r\in\mathbb{Z}}\sum_{i=1}^{N(r)} f(\kappa(r) + U^i)\bigg)^2\bigg\}^{1/2} \leq \sum_{r\in\mathbb{Z}}\bigg\{\mathbb{E}\bigg[\sum_{i=1}^{N(r)} f(\kappa(r) + U^i)\bigg]^2\bigg\}^{1/2}. \tag{24}$$

For any fixed $r \in \mathbb{Z}$, another application of Minkowski's inequality yields

$$\mathbb{E}\bigg[\sum_{i=1}^{N(r)} f(\kappa(r) + U^i)\bigg]^2 = \mathbb{E}\bigg(\mathbb{E}\bigg\{\bigg[\sum_{i=1}^{N(r)} f(\kappa(r) + U^i)\bigg]^2 \,\bigg|\, N(r)\bigg\}\bigg) \leq g(r)^2\mathbb{E}(N(r)^2),$$

where $g(r)$ is defined in (15). Hence, the rest of the proof follows from (24) and using similar arguments as in the proof of Theorem 2.

Acknowledgements

The authors thank the anonymous referee and the editor for helpful comments and valuable suggestions.

References

[1] Asmussen, S. and Glynn, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.

[2] Athreya, K. B. and Roy, V. (2014). Monte Carlo methods for improper target distributions. *Electron. J. Statist.* **8,** 2664–2692.

[3] Baron, M. and Rukhin, A. L. (1999). Distribution of the number of visits of a random walk. *Commun. Statist. Stoch. Models* **15,** 593–597.

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

[5] Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *Amer. Statistician* **46,** 167–174.

[6] Crane, M. A. and Iglehart, D. L. (1975). Simulating stable stochastic systems. III. Regenerative processes and discrete-event simulations. *Operat. Res.* **23,** 33–45.

[7] Durrett, R. (2010). *Probability: Theory and Examples*, 4th edn. Cambridge University Press.

[8] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.

[9] Glynn, P. W. and Iglehart, D. L. (1987). A joint central limit theorem for the sample mean and regenerative variance estimator. *Ann. Operat. Res.* **8,** 41–55.

[10] Hobert, J. P., Jones, G. L., Presnell, B. and Rosenthal, J. S. (2002). On the applicability of regenerative simulation in Markov chain Monte Carlo. *Biometrika* **89,** 731–743.

[11] Karlsen, H. A. and Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *Ann. Statist.* **29,** 372–416.

[12] Kasahara, Y. (1984). Limit theorems for Lévy processes and Poisson point processes and their applications to Brownian excursions. *J. Math. Kyoto Univ.* **24,** 521–538.

[13] Metropolis, N. *et al.* (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21,** 1087–1092.

[14] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
[15] MYKLAND, P., TIERNEY, L. AND YU, B. (1995). Regeneration in Markov chain samplers. *J. Amer. Statist. Assoc.* **90,** 233–241.
[16] ROBERT, C. P. AND CASELLA, G. (2004). *Monte Carlo Statistical Methods*, 2nd edn. Springer, New York.
[17] R DEVELOPMENT CORE TEAM (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. Available at http://www.r-project.org.
[18] SKOROKHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theory Prob. Appl.* **2,** 138–171.