APPLICATIONES MATHEMATICAE
40,2 (2013), pp. 129–165

Witold Bednorz (Warszawa)

THE KENDALL THEOREM AND ITS APPLICATION TO THE GEOMETRIC ERGODICITY OF MARKOV CHAINS

Abstract. We give an improved quantitative version of the Kendall theorem. The Kendall theorem states that under mild conditions imposed on a probability distribution on the positive integers (i.e. a probability sequence) one can prove convergence of its renewal sequence. Due to the well-known property (the first entrance last exit decomposition) such results are of interest in the stability theory of time-homogeneous Markov chains. In particular this approach may be used to measure rates of convergence of geometrically ergodic Markov chains and consequently implies estimates on convergence of MCMC estimators.

1. Introduction. Let $(X_n)_{n\geq 0}$ be a time-homogeneous Markov chain on a measurable space $(\mathcal{S},\mathcal{B})$, with transition probabilities $\mathbf{P}^n(x,\cdot)$, $n\geq 0$, and a unique stationary measure π. Let P be the transition operator given on the Banach space of bounded measurable functions on $(\mathcal{S},\mathcal{B})$ by $Pf(x) = \int f(y)\,\mathbf{P}(x,dy)$. Under mild conditions imposed on $(X_n)_{n\geq 0}$ the chain is ergodic, i.e.

$$\|\mathbf{P}^n(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}}\to 0 \quad \text{as } n\to\infty \tag{1.1}$$

for all starting points $x\in\mathcal{S}$ in the usual total variation norm

$$\|\mu\|_{\mathrm{TV}} = \sup_{|f|\leq 1}\Big|\int f\,d\mu\Big|,$$

where μ is a real measure on $(\mathcal{S},\mathcal{B})$. It is known that the aperiodicity, the Harris recurrence property and the finiteness of π are equivalent to (1.1) (see [10, Theorem 13.0.1]). Consequently, the recurrence property is necessary to prove the convergence of X_n distributions to the invariant measure in the total variation norm regardless of the starting point $X_0=x$. Whenever one

---

2010 *Mathematics Subject Classification*: Primary 60J20; Secondary 60K05, 65C05.
Key words and phrases: geometric ergodicity, renewal theory, Markov chain Monte Carlo.

DOI: 10.4064/am40-2-1

needs to apply the ergodicity for MCMC estimators, a stronger form of the result is required, namely one expects the exponential rate of convergence and a reasonable method to estimate this rate (cf. [13]).

One of the possible generalizations of the total variation convergence is to consider functions controlled from above by $V : \mathcal{S} \to \mathbb{R}$ with $V \geq 1$, $\pi(V) < \infty$. Therefore we define B_V to be the Banach space of all measurable functions on $(\mathcal{S}, \mathcal{B})$ such that $\sup_{x\in\mathcal{S}} |f(x)|/V(x) < \infty$ with the norm

$$\|f\|_V := \sup_{x\in\mathcal{S}} \frac{|f(x)|}{V(x)}.$$

Then instead of the total variation distance one applies

$$\|\mu\|_V := \sup_{|f|\leq V} \left|\int f\, d\mu\right|.$$

Geometric convergence of $\mathbf{P}^n(x,\cdot)$ to a unique stationary measure π means there exists $\rho_V < r \leq 1$ such that

$$\left\|(P^n g)(x) - \int g\, d\pi\right\|_V \leq M_V(r) r^n \|g\|_V, \qquad g \in B_V, \tag{1.2}$$

where ρ_V is the spectral radius of $P - 1\otimes\pi$ acting on $(B_V, \|\cdot\|_B)$, and $M_V(r)$ is the optimal constant. In applications one often works with test functions g from a smaller space B_W, where $W : \mathcal{S} \to \mathbb{R}$ and $1 \leq W \leq V$. In this case we expect

$$\left\|(P^n g)(x) - \int g\, d\pi\right\|_V \leq M_W(r) r^n \|g\|_W, \qquad g \in B_W,$$

which is valid at least for $\rho_V \leq r \leq 1$, and $M_W(r)$ is the optimal constant. The most important case is when $W \equiv 1$, i.e. non-uniform (with respect to $x \in \mathcal{S}$) geometric convergence in the total variation norm. More precisely,

$$\|\mathbf{P}^n(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq M_1(r) V(x) r^n$$

for all $x \in \mathcal{S}$, $r > \rho_V$.

Whenever it exists, we call ρ_V the *convergence rate of geometric ergodicity* for the chain $(X_n)_{n\geq 0}$. For a class of examples one can prove geometric convergence (see [10, Chapter 15]) and it is closely related to the existence of the exponential moment of the return time for a set $C \in \mathcal{B}$ of positive π-measure.

The main tool to measure the convergence rate of geometric ergodicity is the *drift condition*, i.e. the existence of Lyapunov function $V : \mathcal{S} \to \mathbb{R}$, $V \geq 1$, which is contracted outside a small set C. The standard formulation of the required properties is the following:

(1) *Minorization condition.* There exist $C \in \mathcal{B}$, $\bar{b} > 0$ and a probability measure ν on $(\mathcal{S}, \mathcal{B})$ such that

$$\mathbf{P}(x, A) \geq \bar{b}\nu(A) \quad \text{ for all } x \in C \text{ and } A \in \mathcal{B}.$$

(2) *Drift condition.* There exist a measurable function $V : S \to [1, \infty)$ and constants $\lambda < 1$ and $K < \infty$ satisfying

$$PV(x) \le \begin{cases} \lambda V(x) & \text{if } x \notin C, \\ K & \text{if } x \in C. \end{cases}$$

(3) *Strong aperiodicity.* There exists $b > 0$ such that $\bar{b}\nu(C) \ge b > 0$.

The first property means there exists a small set C on which the regeneration of $(X_n)_{n\ge0}$ takes place (see [10, Chapter 5]). The assumption is relatively weak since each Harris recurrent chain admits the existence of such a small set at least for some of its m-skeletons (i.e. processes $(X_{nm})_{n\ge0}$, $m \ge 1$)—see [10, Theorem 5.3.2]. The existence of the small set is used in the split chain construction (see Section 3 and [12] for details) to extend $(X_n)_{n\ge0}$ to a new Markov chain on a larger probability space $\mathcal{S} \times \{0, 1\}$, so that $(C, 1)$ is a true atom of the new chain and its marginal distribution on $\mathcal{S}$ equals the distribution of $(X_n)_{n\ge0}$.

The second condition is the existence of a Lyapunov function V which is contracted by the semigroup related operator P with rate $\lambda < 1$, for all points outside the small set. Finally, strong aperiodicity means that the regeneration set C is of positive measure for the basic transition probability for all starting points in C. Therefore the regeneration can occur in one step assuming the chain is in the set C.

Our main result concerns convergence rates of ergodic Markov chains. Since our approach is based on reduction to renewal sequences, we first prove an abstract theorem that treats renewal sequences and which strengthens previous forms of the result (known as the Kendall theorem). Only then do we analyze the atomic case and show how to apply the idea to the case when a true atom exists and what is the natural setting for our approach. However, the idea is valid for general Harris chains. It requires additional work, the split chain construction. Results of this type are used whenever exact estimates on the ergodicity matter (cf. [1], [8] and [7]).

The organization of the paper is as follows: the history of the abstract Kendall theorem as well as our main improvement are contained in Section 2; in Section 3 we compare our extensions with what was previously known; then in Section 4 we discuss how the abstract Kendall theorem affects estimation of convergence rates for atomic Markov chains; using the method of chain split, in Section 5 we extend the results to general Harris chains; we leave the tedious estimates of constants (improving previous results of this type) to Appendix A; finally, in Appendix B we analyze the result for basic toy examples.

2. The abstract Kendall theorem. Let $(\tau_k)_{k\ge0}$ be a random walk on $\mathbb{N}$ starting from zero, i.e. $\tau_0 = 0$, $\tau_k - \tau_{k-1}$, $k \ge 1$, are independent distributed like τ, that is,

$$\mathbf{P}(\tau_k - \tau_{k-1} = n) = \mathbf{P}(\tau = n) = b_n, \quad n \geq 1.$$

By the definition, the sequence $(b_n)_{n\geq 1}$ is stochastic, which means $b_n \geq 0$ and $\sum_{n=1}^{\infty} b_n = 1$. From the applications' point of view such a random walk is generated by consecutive visits of an atomic Markov chain to a given true atom. The renewal process for the sequence $(\tau_{k\geq 0})$ is defined by $V_m = \inf\{\tau_n - m : \tau_n \geq m\}$, $m \geq 0$. In the language of Markov chains the process measures how long it is before the next visit to the true atom. Let $u_n = \mathbf{P}(V_n = 0)$, $n \geq 0$, i.e. the probability that the process $(V_m)_{m\geq 0}$ renews (goes to zero) in the nth time step. The sequence $(u_n)_{n\geq 0}$ is of importance for the study of ergodic properties of Markov chains, which will be the main issue of the next sections. In particular, u_n equals the probability that the suitable atomic Markov chain stays in the given atom in the nth time step. Observe that $u_0 = 1$ and $u_n = \sum_{k=1}^{n} u_{n-k} b_k$, hence denoting $b(z) = \sum_{n=1}^{\infty} b_n z^n$ and $u(z) = \sum_{n=0}^{\infty} u_n z^n$ for $z \in \mathbb{C}$, one can state the renewal equation as follows:

$$u(z) = 1/(1 - b(z)) \quad \text{for } |z| < 1. \tag{2.1}$$

The equation means that to study the properties of $(u_n)_{n\geq 0}$ it suffices to concentrate on the properties of $(b_n)_{n\geq 1}$. In particular one can ask when the sequence $(u_n)_{n\geq 0}$ is ergodic, that is, when $\lim_{n\to\infty} u_n$ exists. Historically, the first result that matches these properties with geometric ergodicity was due to Kendall [6], who proved

THEOREM 2.1. *Assume that* $b_1 > 0$ *and* $\sum_{n=1}^{\infty} b_n r^n < \infty$ *for some* $r > 1$. *Then the limit* $u_\infty = \lim_{n\to\infty} u_n$ *exists and equals* $u_\infty = (\sum_{n=1}^{\infty} n b_n)^{-1}$; *moreover, the radius of convergence of* $\sum_{n=0}^{\infty} (u_n - u_\infty) z^n$ *is strictly greater than* 1.

The Kendall theorem states that the sequence $(u_n)_{n\geq 0}$ is ergodic whenever $b(z)$ is bounded on the disc of radius strictly greater than 1 and we have slight control over b_1. However, the question is: does Theorem 2.1 imply any rates of convergence? This obviously requires basic information about the upper bound on $b(z)$, i.e. $b(R) \leq L$ for a given $R > 1$, and the lower bound $b_1 \geq b > 0$. The data b, R, L come from conditions 1–3 of the introduction, and are easy to compute in the atomic case. Consequently, the main question we treat in this section is what one can say about the rate of convergence of u_n, $n \geq 0$, to u_∞ having information on b, R, L. This is an abstract Kendall-type question on renewal processes, where we search for r_0, a lower bound on the radius of convergence for $\sum_{n=0}^{\infty} (u_n - u_\infty) z^n$, and $K_0(r)$, a computable upper bound on $\sup_{|z|=r} |\sum_{n=0}^{\infty} (u_n - u_\infty) z^n|$ for $1 \leq r < r_0$.

The Kendall theorem was improved first in [11] and then in [2, Theorem 3.2]. There are also several results where some additional assumptions on the distribution of τ are made. For example, [3] elaborates on how to provide an optimal bound on the rate of convergence, but under additional conditions on the τ distribution. Whenever the general Kendall question is considered, the bounds obtained up to now are still far from being optimal or easy to use.

The goal of this paper is to give a more accurate estimate on the rate of convergence which significantly improves upon the previous results. Our approach is based on introducing u_∞ as a parameter, namely we prove that the following result holds:

THEOREM 2.2. *Suppose that* $(b_n)_{n\geq 1}^\infty$ *satisfies* $b_1 \geq b > 0$ *and* $b(r) = \sum_{n=1}^\infty b_n r^n < \infty$ *for some* $r > 1$. *Then* $u_\infty = (\sum_{n=1}^\infty n b_n)^{-1}$ *and*

$$\sup_{|z|=r} \Big| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \Big| \leq \frac{c(r) - c(1)}{c(1)(r-1)\big([(1-b)D(\alpha) - c(r) + c(1)]_+\big)},$$

where $c(r) = \frac{b(r)-1}{r-1}$, $c(1) = u_\infty^{-1}$ *and*

$$D(\alpha) = \frac{\big|1 + \frac{b}{1-b}(1 - e^{\frac{i\pi}{1+\alpha}})\big| - 1}{|1 - e^{\frac{i\pi}{1+\alpha}}|}, \quad \textit{where} \quad \alpha = \frac{c(1)-1}{1-b},$$

Proof. Let $b(z)$ and $u(z)$ be the complex generating functions for b_i, $i \geq 1$, and u_i, $i \geq 0$, respectively. The main tool we use is the renewal equation (2.1), i.e.

$$1 - b(z) = \frac{1}{u(z)}, \qquad |z| < 1.$$

Note that the equation remains valid on the disc $|z| \leq R$ in the sense of analytic functions. By Theorem 2.1 we know that $u_\infty < \infty$ and the renewal generating function $\sum_{n=0}^\infty (u_n - u_\infty) z^n$ is convergent on some disc of radius greater than 1. Denote $c(z) = (b(z) - 1)/(z - 1)$ (cf. [2, proof of Theorem 3.2]) and observe that $c(z)$ is well defined on $|z| \leq R$, because

$$c(R) = \frac{b(R) - 1}{R - 1} = \frac{L - 1}{R - 1} < \infty.$$

Since $u_\infty = c(1)^{-1}$ we have

$$(2.2) \quad \sum_{n=0}^{\infty} (u_n - u_\infty) z^n = u(z) - \frac{1}{c(1)(1-z)} = \frac{1}{1 - b(z)} - \frac{1}{c(1)(z-1)}$$
$$= \frac{1}{1-z}\left(\frac{1}{c(z)} - \frac{1}{c(1)}\right) = \frac{c(z) - c(1)}{z - 1} \frac{1}{c(1)c(z)}.$$

The main problem is to estimate $|c(z)|$ from below, for which we use the simple trick

$$(2.3) \qquad |c(re^{i\theta})| = |c(e^{i\theta})| - |c(re^{i\theta}) - c(e^{i\theta})| = |c(e^{i\theta})| - c(r) + c(1).$$

Consequently, the problem is reduced to finding a lower bound on $|c(e^{i\theta})|$. We recall that by definition $c_i = \sum_{j>i} b_j$ and $c(1) = \sum_{i=0}^{\infty} c_i$. To provide a sharp estimate of (2.3) we use the fact that for $\pi/(l+1) < |\theta| \le \pi/l$, $l \ge 1$, there is a better control on the first l summands in $c(e^{i\theta}) = \sum_{j=1}^{\infty} c_j e^{ij\theta}$. First we note that

$$|c(e^{i\theta})| = \frac{|1 - \sum_{j=1}^{\infty} b_j e^{ij\theta}|}{|1-e^{i\theta}|} \ge \frac{|1 - \sum_{j=1}^{l} b_j e^{ij\theta}| - \sum_{j>l} b_j}{|1-e^{i\theta}|},$$

which is equivalent to

$$|c(e^{i\theta})| \ge \frac{|c_l + \sum_{j=1}^{l} b_j(1-e^{ij\theta})| - c_l}{|1-e^{i\theta}|}.$$

A geometrical observation gives, for $\pi/(l+1) < |\theta| \le \pi/l$,

$$\Big|c_l + \sum_{j=1}^{l} b_j(1-e^{ij\theta})\Big| \ge \Big|c_l + \Big(\sum_{j=1}^{l} b_j\Big)(1-e^{i\theta})\Big| = |c_l + (1-c_l)(1-e^{i\theta})|,$$

hence we conclude that

$$|c(e^{i\theta})| \ge c_l|1-e^{i\theta}|^{-1}\big(|1+(1-c_l)c_l^{-1}(1-e^{i\theta})| - 1\big).$$

Since $1 - c_l \ge b$, for $l \ge 1$ we see that

$$|1+(1-c_l)c_l^{-1}(1-e^{i\theta})| \ge \sqrt{1 + bc_l^{-2}|1-e^{i\theta}|^2}.$$

It remains to verify that $f(x) = x^{-1}[\sqrt{1+bx^2} - 1]$ is increasing, which is ensured by

$$f'(x) = -x^{-2}(\sqrt{1+bx^2} - 1) + x^{-2}\frac{bx^2}{\sqrt{1+4bx^2}} \ge 0. \tag{2.4}$$

Therefore we finally obtain, for $\pi/(l+1) < |\theta| \le \pi/l$,

$$|c(e^{i\theta})| \ge c_l|1-e^{\frac{i\pi}{l+1}}|\Big(\sqrt{1 + bc_l^{-2}|1-e^{\frac{i\pi}{l+1}}|^2} - 1\Big). \tag{2.5}$$

Due to (2.4) and (2.5), when estimating the global minimum of $|c(e^{i\theta})|$ it suffices to find a bound from above on $c_l|1-e^{\frac{i\pi}{l+1}}|^{-1}$. We will show that

$$c_l|1-e^{\frac{i\pi}{l+1}}|^{-1} \le (1-b)|1-e^{\frac{i\pi}{1+\alpha}}|^{-1}, \tag{2.6}$$

where we recall that $\alpha = (c(1)-1)/(1-b)$. First observe that (2.6) is trivial for $l \le \alpha$, since $c_l \le 1-b$ and $|1-e^{\frac{i\pi}{l+1}}| \ge |1-e^{\frac{i\pi}{1+\alpha}}|$. On the other hand, for $l > \alpha$,

$$c_l|1-e^{\frac{i\pi}{l+1}}|^{-1} \ge (c_l l)(l|1-e^{\frac{i\pi}{l+1}}|)^{-1} \ge (c_l l)(\alpha|1-e^{\frac{i\pi}{1+\alpha}}|)^{-1}. \tag{2.7}$$

Using that $c(1) = \sum_{j=0}^{\infty} c_j$ we deduce

$$c_l l \le \sum_{j=1}^{l} c_j \le c(1) - 1 = \alpha(1-b), \tag{2.8}$$

and combining (2.7) and (2.8) we obtain

$$c_l |1 - e^{\frac{i\pi}{l+1}}|^{-1} \le (1-b)|1 - e^{\frac{i\pi}{1+\alpha}}|^{-1},$$

which is (2.6).

As already noted, (2.4) implies that

$$|c(e^{i\theta})| \ge (1-b)|1 - e^{\frac{\pi i}{1+\alpha}}|^{-1}\Big(\sqrt{1 + b(1-b)^{-2}|1 - e^{\frac{\pi i}{1+\alpha}}|^2} - 1\Big),$$

which is equivalent to

$$|c(e^{i\theta})| \ge |1 - e^{\frac{\pi i}{1+\alpha}}|^{-1}\big(|(1-b) + b(1 - e^{\frac{\pi i}{1+\alpha}})| - (1-b)\big). \tag{2.9}$$

Plugging (2.9) into (2.3) we derive

$$|c(re^{i\theta})| \ge \frac{|(1-b) + b(1 - e^{\frac{\pi i}{1+\alpha}})| - (1-b)}{|1 - e^{\frac{\pi i}{1+\alpha}}|} - c(r) + c(1).$$

Finally, using (2.2) we conclude that

$$\sup_{|z|=r} \Big|\sum_{n=0}^{\infty} (u_n - u_\infty) z^n\Big| \le \frac{c(r) - c(1)}{c(1)(r-1)\big((1-b)D(\alpha) - c(r) + c(1)\big)},$$

where $D(\alpha) = |1 - e^{\frac{\pi i}{1+\alpha}}|^{-1}\big(\big|1 + \frac{b}{1-b}(1 - e^{\frac{\pi i}{1+\alpha}})\big| - 1\big)$, which completes the proof of Theorem 2.2. ∎

Consequently, whenever one can control $c(r) = (b(r) - 1)/(r - 1)$ from above, there is a bound on the rate of convergence for the renewal process. The simplest case is when $c(1) = u_\infty^{-1}$ is known and we can control $c(r)$ at a certain point, i.e. $c(R) \le N < \infty$ for some $R > 1$. Observe that if $b(R) \le L$, then due to $c(R) = (b(R) - 1)/(R - 1)$ one derives that $c(R) \le N = (L-1)/(R-1)$, which will be our basic setting. Note that by the Hölder inequality, for all $1 \le r \le R$,

$$c(r) - c(1) = (c(1) - 1)\left(\frac{c(r) - 1}{c(1) - 1} - 1\right) \le (1-b)\alpha(r^{\kappa(\alpha)} - 1),$$

where

$$\kappa(\alpha) = \frac{\log\big(\frac{N-1}{c(1)-1}\big)}{\log R} = \frac{\log\big(\frac{N-1}{(1-b)\alpha}\big)}{\log R}, \qquad \alpha = \frac{c(1) - 1}{1 - b}.$$

We summarize this in the following assertion:

COROLLARY 2.3. *Suppose that $c(1) = u_\infty^{-1}$ is known, $b_1 \ge b$ and $b(R) \le L$. Then $\sum_{n=0}^{\infty}(u_n - u_\infty)z^n$ is convergent for $|z| < r_0$, where*

$$r_0 = \min\{R, (1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}\}. \tag{2.10}$$

Moreover, for $r < r_0$,

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le K_0(r) = \frac{u_\infty(r^{\kappa(\alpha)} - 1)}{(r-1)(\alpha^{-1}D(\alpha) - r^{\kappa(\alpha)} + 1)}.$$

REMARK 2.4. Observe that the bound $(1+D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ increases with b assuming that $L, R, c(1)$ are fixed.

In applications we have to treat $c(1) = u_\infty^{-1}$ as a parameter. The advantage of this approach is that there is a sharp upper bound on $c(1)$ or rather $\alpha = (c(1) - 1)/(1 - b)$. Using the inequality

$$\begin{aligned} R^\alpha = R^{\sum_{n=1}^{\infty}(n-1)b_n/(1-b)} &\le \frac{\sum_{n=2}^{\infty} b_n R^{n-1}}{1-b} \le \frac{b(R) - bR}{(1-b)R} \\ &\le \frac{L - bR}{(1-b)R}, \end{aligned} \tag{2.11}$$

we deduce that $\alpha \le \alpha_0$, where $\alpha_0 = \log\big(\frac{L-bR}{(1-b)R}\big)/\log R$. On the other hand, if $b = b_1$, then $c(1) - 1 \ge 1 - b$ and therefore by Remark 2.4 we can always require that $c(1) - 1 \ge 1 - b$ or equivalently $\alpha \ge 1$. Therefore to find an estimate on the rate of convergence we search for the possible minimum of $(1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}$, $\alpha \in [1, \alpha_0]$.

COROLLARY 2.5. *Suppose that $b_1 \ge b$ and $b(R) \le L$. Then the series $\sum_{n=0}^{\infty}(u_n - u_\infty)z^n$ is convergent for $|z| < r_0$, where*

$$r_0 = \min\Big\{R, \min_{1\le\alpha\le\alpha_0}(1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}\Big\}. \tag{2.12}$$

Moreover, for $r < r_0$,

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le K_0(r) = \max_{1\le\alpha\le\alpha_0}\frac{r^{\kappa(\alpha)} - 1}{(r-1)(\alpha^{-1}D(\alpha) - r^{\kappa(\alpha)} + 1)}. \tag{2.13}$$

The above corollary should be compared with [2, Theorem 3.2]; we defer the discussion to the following section.

3. Comparing with the previous bounds. Recall that our bound on the radius of convergence is of the form

$$r_0 = \min\{R, \hat{r}_0\}, \qquad \hat{r}_0 = \min_{1\le\alpha\le\alpha_0}(1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}.$$

This estimate will be shown to be always better than the main bound in [2, Theorem 3.2]. Then we will investigate the reason for this improvement.

Using the limit case with b, L fixed and $R \to 1$, we check that the minimum of $\alpha \mapsto (1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ can be attained in the interval $[1, \alpha_0]$ and that it is a data depending problem one cannot avoid. On the other hand, we stress that in the usual setting the minimum of $\alpha \mapsto (1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ should be attained at α_0. The intuition behind this phenomenon is that the smaller $c(1) = u_\infty^{-1}$, the worse rate of convergence one should expect. The intuition fails only when L is chosen to be close to 1 with respect to the other data: b, R.

Observe that the minimum of the function $(1+D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ is attained at the unique point α that satisfies

$$\log\left(\frac{N-1}{1-b}\right) = \log\alpha + \log\left(1 + \frac{D(\alpha)}{\alpha}\right)\frac{D(\alpha)+\alpha}{D(\alpha) - \alpha D'(\alpha)}. \tag{3.1}$$

Obviously, to find the minimum on the interval $[1, \alpha_0]$, the solution α of (3.1) must be compared with 1 and α_0. Consequently, $\hat{r}_0 = (1 + D(1))^{1/\kappa(1)}$ when such α is smaller than 1, and $\hat{r}_0 = (1 + D(\alpha_0)/\alpha_0)^{1/\kappa(\alpha_0)}$ when it is greater than α_0, otherwise the solution of (3.1) is the worst possible α that minimizes our bound on the radius of convergence. The same discussion concerns maximization of $K_0(r)$. Clearly the problem reduces to finding the maximum of the function $\alpha(D(\alpha))^{-1}(r^{\kappa(\alpha)} - 1)$ which is attained at the unique point α that satisfies the equation

$$\left(1 + \frac{D'(\alpha)\alpha}{D(\alpha)}\right)(r^{\kappa(\alpha)} - 1) = \frac{\log r}{\log R} r^{\kappa(\alpha)}. \tag{3.2}$$

To find the maximum of $\alpha(D(\alpha))^{-1}(r^{\kappa(\alpha)} - 1)$ on $[1, \alpha_0]$ we compare the solution α of (3.2) with 1 and α_0. If $\alpha > \alpha_0$ then

$$\alpha_0(D(\alpha_0))^{-1}(r^{\kappa(\alpha)_0} - 1)$$

is the optimal bound on $\max_{1\le\alpha\le\alpha_0} \alpha(D(\alpha))^{-1}(r^{\kappa(\alpha)} - 1)$. Similarly if $\alpha \le 1$ then $(D(1))^{-1}(r^{\kappa(1)} - 1)$ is the bound and otherwise the solution of (3.2) is the maximum point for $\max_{1\le\alpha\le\alpha_0} \alpha(D(\alpha))^{-1}(r^{\kappa(\alpha)} - 1)$.

Remark 3.1. It is possible that the bound L is so good that R is the optimal lower bound on the radius of convergence of $\sum_{n=0}^{\infty}(u_n - u_\infty)z^n$, i.e. $r_0 = R$. This is the case when the solution of (3.1) is smaller than 1, i.e. when

$$\left(1 + \frac{D'(1)}{D(1)}\right)(R^{\kappa(1)} - 1) \ge R^{\kappa(1)}.$$

We now look for computable bounds on $K_0(r)$ in the case when u_∞ is unknown. Note that the function $D(\alpha)$ is decreasing and therefore $D(\alpha) \ge D(\alpha_0)$. Consequently, one can rewrite Corollary 2.5 with $D(\alpha)$ replaced by $D(\alpha_0)$ and in this way obtain new bounds: $K_1(r) \ge K_0(r)$ and $r_1 \le r_0$, where $r_1 = \min\{R, \hat{r}_1\}$, $\hat{r}_1 = \min_{1\le\alpha\le\alpha_0}(1 + D(\alpha_0)/\alpha)^{1/\kappa(\alpha)}$ and

$$K_1(r) = \max_{1\le\alpha\le\alpha_0} \frac{r^{\kappa(\alpha)} - 1}{(r-1)(D(\alpha_0)\alpha^{-1} - r^{\kappa(\alpha)+1})}.$$

Consequently, to find $K_1(r)$ it suffices to compute the maximum of $\alpha(r^{\kappa(\alpha)}-1)$ on $[1, \alpha_0]$. The maximum of $\alpha(r^{\kappa(\alpha)} - 1)$ is attained at α that satisfies

$$r^{\kappa(\alpha)} - 1 = \frac{\log r}{\log R} r^{\kappa(\alpha)}. \tag{3.3}$$

The solution of (3.3) is

$$\alpha = \frac{N-1}{1-b}\Big(1 - \frac{\log r}{\log R}\Big)^{\frac{\log R}{\log r}}. \tag{3.4}$$

Again the solution must be compared with 1 and α_0, which finally provides the direct form of $K_1(r)$. We have proved the following result:

Corollary 3.2. *Suppose that $b_1 \ge b$ and $b(R) \le L$.*

(i) *If $1 \ge \frac{N-1}{1-b}\big(1 - \frac{\log r}{\log R}\big)^{\frac{\log R}{\log r}}$, then*

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le K_1(r) = (r-1)^{-1}\bigg(\bigg[\frac{D(\alpha_0)}{(r^{\kappa(1)-1})} - 1\bigg]_+\bigg)^{-1}.$$

(ii) *If $1 \le \frac{N-1}{1-b}\big(1 - \frac{\log r}{\log R}\big)^{\frac{\log R}{\log r}} \le \alpha_0$, then*

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le K_1(r)$$

$$= (r-1)^{-1}\bigg(\bigg[\frac{(1-b)D(\alpha_0)}{N-1}\,\frac{\log R}{\log r}\Big(1 - \frac{\log r}{\log R}\Big)^{-\frac{\log R}{\log r}+1} - 1\bigg]_+\bigg)^{-1}.$$

(iii) *If $\alpha_0 \le \frac{N-1}{1-b}\big(1 - \frac{\log r}{\log R}\big)^{\frac{\log R}{\log r}}$, then*

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le K_1(r) = (r-1)^{-1}\bigg(\bigg[\frac{D(\alpha_0)}{\alpha_0(r^{\kappa(\alpha_0)-1})} - 1\bigg]_+\bigg)^{-1}.$$

Corollary 3.2 implies some interpretation of $\hat{r}_1$ as a solution of an equation which we need to compare our bound with the previous results. Let $x_\alpha = r$, $\alpha \ge 1$, be the unique solution of

$$\alpha = \frac{N-1}{1-b}\Big(1 - \frac{\log r}{\log R}\Big)^{\frac{\log R}{\log r}} \tag{3.5}$$

if $\frac{N-1}{(1-b)\alpha} \ge e$, and $x_\alpha = 1$ otherwise. From Corollary 3.2 we deduce

COROLLARY 3.3. *Suppose that $b_1 \ge b$ and $b(R) \le L$. Let $\bar r$ be the unique r satisfying*

$$\frac{(1-b)D(\alpha_0)}{N-1} = \frac{\log r}{\log R}\left(1 - \frac{\log r}{\log R}\right)^{\frac{\log R}{\log r}-1}.$$

If $\bar r \le x_1$ then $\hat r_1 = (1 + D(\alpha_0))^{1/\kappa(1)}$; if $x_1 \le \bar r \le x_{\alpha_0}$ then $\hat r_1 = \bar r$; and if $\bar r \ge x_{\alpha_0}$ then $\hat r_1 = (1 + D(\alpha_0)/\alpha_0)^{1/\kappa(\alpha_0)}$.

Clearly $r_1 \le r_0$; we now turn to showing that r_1 is better than the main bound in [2, Theorem 3.2], which we denote by r_2. Again $r_2 = \min\{R, \hat r_2\}$ and $\hat r_2$ is the unique r satisfying

$$\frac{r-1}{r}\frac{1}{\log^2(R/r)} = \frac{b}{2N}. \tag{3.6}$$

Our aim is to show that $r_2 \le r_1$. First observe that by definition

$$\hat r_1^{\kappa(\alpha)} - 1 = \frac{D(\alpha_0)}{\alpha}$$

for some $\alpha \in [1, \alpha_0]$. Again by definition $R^{\kappa(\alpha)} = (N-1)/((1-b)\alpha)$, which yields

$$\kappa(\alpha)\hat r_1^{\kappa(\alpha)}\frac{\hat r_1 - 1}{\hat r_1} \ge \hat r_1^{\kappa(\alpha)} - 1 \ge \frac{(1-b)R^{\kappa(\alpha)}D(\alpha_0)}{N-1}. \tag{3.7}$$

By (3.7) and the inequality

$$D(\alpha_0) = \frac{\sqrt{(1-b)^2 + 4b\sin^2\left(\frac{\pi}{2(1+\alpha_0)}\right)} - (1-b)}{2(1-b)\sin\left(\frac{\pi}{2(1+\alpha_0)}\right)} \ge \frac{b}{(1-b)(1+\alpha_0)},$$

we obtain

$$\kappa(\alpha)\frac{\hat r_1^{\kappa(\alpha)}}{R^{\kappa(\alpha)}}\frac{\hat r_1 - 1}{\hat r_1} \ge \frac{b}{(1+\alpha_0)(N-1)}. \tag{3.8}$$

It suffices to note that $1 + \alpha_0 \le 2\kappa(\alpha_0) \le 2\kappa(\alpha)$, which is a consequence of $\kappa(\alpha_0) \le \kappa(\alpha)$ and the fact that

$$R^{\kappa(\alpha_0)} = R\frac{R^{\alpha_0} - 1}{R - 1},$$

which can be used to show that for a given R, the function $\kappa(\alpha_0)/(1+\alpha_0)$ is increasing with α_0. Thus since $\kappa(\alpha_0)/(1+\alpha_0) = 1/2$ for $\alpha_0 = 1$ we deduce that $1 + \alpha_0 \le 2\kappa(\alpha_0)$. Plugging the estimate $2\kappa(\alpha) \ge 1 + \alpha_0$ into (3.8) we derive

$$\kappa(\alpha)^2\frac{\hat r_1^{\kappa(\alpha)}}{R^{\kappa(\alpha)}}\frac{\hat r_1 - 1}{\hat r_1} \ge \frac{b}{2(N-1)}.$$

It remains to check that $\kappa(\alpha) = 2/\log(R/\hat{r}_1)$ is the maximum point of $\kappa(\alpha)^2(\hat{r}_1/R)^{\kappa(\alpha)}$, which implies that

$$\frac{\hat{r}_1}{\hat{r}_1 - 1}\frac{1}{\log^2(R/\hat{r}_1)} \geq \frac{be^2}{8(N-1)}.$$

This shows that $\hat{r}_1 \geq \hat{r}_2$ and in fact $\hat{r}_2$ can be treated as the lower bound in the worst possible case of our result. We stress that using α_0 instead of the minimization over all α_0 usually gives a major numerical improvement.

To provide a convincing numerical argument for exploiting the parameter α_0 let us consider the simplest renewal model where there are only two possible states 1 and α_0 (for simplicity assume that $\alpha_0 \in \mathbb{N}$). Then the optimal rate of convergence is closely related to a specific solution of $\frac{bz+(1-b)z^{\alpha_0}-1}{z-1} = 0$, namely it is the inverse of the smallest absolute value of solutions of this equation. Denoting the root by z_0 one can show that

$$|z_{\alpha_0}| = 1 + \frac{2b\pi^2}{(1-b)^2\alpha_0^3} + o(\alpha_0^{-3}) \tag{3.9}$$

(see discussion after [2, Theorem 3.2]) and α_0 is exactly our parameter. Therefore whenever the estimate $(1+D(\alpha_0)/\alpha_0)^{1/\kappa(\alpha_0)}$ is applied, one cannot improve it up to a numerical constant.

We turn to studying this phenomenon in the limit case where b, L are fixed and $R \to 1$.

Corollary 3.4. *Suppose that $R \to 1$ and $b_1 \geq b$, $b(R) \leq L$.*

(i) *If $\big(\frac{L-1}{1-b}\big)/\log\big(\frac{L-b}{1-b}\big) \geq e^{1/2}$, then*

$$r_0(R) = 1 + \frac{b\pi(R-1)^3}{2(1-b)^2}\log^{-2}\left(\frac{L-b}{1-b}\right)\log^{-1}\left((L-1)/\log\frac{L-b}{1-b}\right) + o((R-1)^3).$$

(ii) *If $\big(\frac{L-1}{1-b}\big)/\log\big(\frac{L-b}{1-b}\big) \leq e^{1/2}$, then*

$$r_0(R) = 1 + \frac{be\pi(R-1)^3}{(L-1)^2} + o((R-1)^3).$$

Proof. Observe that

$$\lim_{\alpha\to\infty} \alpha D(\alpha) = \frac{b\pi}{2(1-b)^2}, \tag{3.10}$$

thus we can treat $\pi b(2(1-b)^2\alpha)^{-1}$ as the right approximation of $D(\alpha)$ when α tends to infinity. As stated in Corollary 2.5, to find

$$\hat{r}_0(R) = \inf_{1\leq\alpha\leq\alpha_0(R)} (1 + D(\alpha)/\alpha)^{1/\kappa(\alpha)} \tag{3.11}$$

one should solve the equation (3.1), i.e. find $\alpha(R)$ that satisfies

$$\log\left(\frac{N(R)-1}{1-b}\right) = \log\alpha + \log\left(1+\frac{D(\alpha)}{\alpha}\right)\frac{D(\alpha)+\alpha}{D(\alpha)-\alpha D'(\alpha)}, \tag{3.12}$$

where $N(R) = (L-1)/(R-1)$, and compare the outcome with 1 and $\alpha_0(R)$. In particular we deduce from (3.12) that $\alpha(R)$ necessarily tends to infinity as $R \to 1$, hence using

$$\lim_{\alpha\to\infty}\left(1+\frac{\alpha}{D(\alpha)}\right)\log\left(1+\frac{D(\alpha)}{\alpha}\right) = 1 \quad \text{and} \quad \lim_{\alpha\to\infty}\left(1-\frac{\alpha D'(\alpha)}{D(\alpha)}\right) = 2,$$

we obtain

$$\log\alpha(R) = -\frac{1}{2} + \log\left(\frac{N(R)-1}{1-b}\right) + o(1).$$

The solution must be compared with $\alpha_0(R)$, so if

$$\lim_{R\to\infty}\frac{N(R)-1}{\alpha_0(R)R-1} = \frac{L-1}{1-b}\log^{-1}\left(\frac{L-b}{1-b}\right) < e^{1/2}$$

we have to use $\alpha(R)$ (at least for small R) to minimize $(1+D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ over $[1, \alpha_0(R)]$, otherwise $\alpha_0(R)$ is the minimum point. In the first case we have

$$\alpha(R) = e^{-1/2}\frac{L-1}{(1-b)(R-1)} + o(1) \quad \text{and} \quad \kappa(\alpha(R)) = \frac{1}{2(R-1)} + o(1),$$

thus using (3.10) and (3.11) we obtain

$$\begin{aligned}
\hat{r}_0(R) &= \left(1+\frac{D(\alpha(R))}{\alpha(R)}\right)^{1/\kappa(\alpha(R))} = 1 + \frac{D(\alpha(R))}{\alpha(R)\kappa(\alpha(R))} + o((R-1)^3)\\
&= 1 + \frac{\pi b}{2(1-b)^2\alpha^2(R)\kappa(\alpha(R))} + o((R-1)^3)\\
&= 1 + \frac{\pi e b(R-1)^3}{(L-1)^2} + o((R-1)^3).
\end{aligned}$$

In the same way if $\frac{L-1}{1-b}\log^{-1}\left(\frac{L-b}{1-b}\right) \geq e^{1/2}$, then

$$\alpha_0(R) = \frac{\log\left(\frac{L-b}{1-b}\right)}{R-1} + o(1), \qquad \kappa(\alpha_0(R)) = \frac{\frac{L-1}{1-b}}{(R-1)\log\left(\frac{L-b}{1-b}\right)} + o(1),$$

and hence

$$\begin{aligned}
\hat{r}_0(R) &= \left(1+\frac{D(\alpha_0(R))}{\alpha_0(R)}\right)^{1/\kappa(\alpha_0(R))} = 1 + \frac{D(\alpha_0(R))}{\alpha_0(R)\kappa(\alpha_0(R))} + o((R-1)^3)\\
&= 1 + \frac{\pi b}{2(1-b)^2\alpha_0^2(R)\kappa(\alpha_0(R))} + o((R-1)^3)\\
&= 1 + \frac{\pi b(R-1)^3}{2(1-b)^2}\log^{-2}\left(\frac{L-b}{1-b}\right)\log^{-1}\left((L-1)/\log\left(\frac{L-b}{1-b}\right)\right)\\
&\quad + o((R-1)^3).
\end{aligned}$$

It is clear that $\hat{r}_0(R) \le R$ for R small enough, so the asymptotic for $\hat{r}_0(R)$ is the same as for $r_0(R)$. This completes the proof of the corollary. ∎

In particular Corollary 3.4 shows that whenever $\frac{L-1}{1-b}\log^{-1}\big(\frac{L-b}{1-b}\big) \ge e^{1/2}$ then

$$r_0(R) = 1 + \frac{\pi b}{2(1-b)^2\alpha_0^2(R)\kappa(\alpha_0(R))} + o(\alpha_0(R)^{-3}),$$

which when compared with (3.9) proves that our result cannot be improved up to a numerical constant (recall that $(1+\alpha_0)/2 \le \kappa(\alpha_0) \le \alpha_0$). On the other hand, Corollary 3.4 makes it possible to compare our result with [2, Theorem 3.2]. The following estimate holds for $r_2(R)$ in the same setting (see [2, Section 3]):

$$r_2(R) = 1 + \frac{e^2 b(R-1)^3}{8(L-1)} + o((R-1)^3).$$

Therefore if $L-1$ is much larger than $1-b$ our answer is better by a factor of $(L-1)/(1-b)^2$ and if $L-1$ is close to $1-b$ then by a factor of $L-1$.

We stress that there are indeed two data-depending cases: either L is far from 1 with respect to b, L, and then the minimum of $(1+D(\alpha)/\alpha)^{1/\kappa(\alpha)}$ is attained at $\alpha_0(R)$; or L is close to 1 (again with respect to b and L) and then we have to use minimization inside $[1, \alpha_0(R)]$ even for $R \to 1$. This explains that one cannot avoid minimization over $\alpha \in [1, \alpha_0]$ from the discussion of r_0 estimates.

4. The atomic case. In this section we follow the classic idea of the first entrance last exit decomposition to obtain rates of convergence for ergodic Markov chains under the assumption that a true atom exists.

For this section we assume that $\bar{b} = 1$. Note that in this setting one can rewrite the minorization condition (1) (from the introduction) as

$$\mathbf{P}(x, A) = \nu(A) \quad \text{for all } x \in C,$$

which implies that C is an atom and $\nu = \mathbf{P}(a, \cdot)$ for any $a \in C$. It remains to translate conditions (2)–(3) (from the introduction) into a simpler form which can be used later to prove geometric ergodicity. Let $\tau = \tau(C) = \inf\{n \ge 1 : \ X_n \in C\}$ and then define τ_k, $k \ge 1$, as the successive visits to C. For simplicity let also $\tau_0 = \sigma(C) = \inf\{n \ge 0 : X_n \in C\}$, which means $\tau_0 = 0$ whenever we start the chain from $a \in C$. In this way we construct a random walk of the form stated in the previous section such that $b_n = \mathbf{P}_a(\tau = n)$. Moreover denoting $u_n = \mathbf{P}_a(X_n \in C)$ for $n \ge 0$, we obtain the renewal sequence for $(\tau_k)_{k\ge 0}$.

As already mentioned, the behavior of $(u_n)_{n\ge 0}$ is closely related to the ergodicity of the Markov chain. In particular, assuming ergodicity, $\lim_{n\to\infty} u_n$ exists and equals $u_\infty = \pi(C)$. Following [2] we define $G(r, x) = \mathbf{E}_x r^\tau$ for all

$x \in S$ and $0 < r \le \lambda^{-1}$. The main property of $G(r, x)$ is that it is a lower bound for $V(x)$ on $\mathcal{S} \setminus C$, namely we have

PROPOSITION 4.1 (see [2, Proposition 4.1]). *Assume only the drift condition* (2).

(i) *For all* $x \in S$, $\mathbf{P}_x(\tau < \infty) = 1$.
(ii) *For* $1 \le r \le \lambda^{-1}$,

$$G(r,x) \le \begin{cases} V(x) & \text{if } x \notin C, \\ rK & \text{if } x \in C. \end{cases}$$

The renewal approach is based on the first entrance last exit property. To state it we need the notation $H_W(r, x) = \mathbf{E}_x(\sum_{n=1}^{\tau} r^n W(X_n))$, for all $r > 0$ for which the definition makes sense. The following result holds:

PROPOSITION 4.2 (see [2, Proposition 4.2]). *Assume only that the Markov chain is geometrically ergodic with a* (*unique*) *invariant probability measure* π, *that* C *is an atom, and that* $W : \mathcal{S} \to \mathbb{R}$ *is such that* $W \ge 1$. *Suppose* $g : S \to \mathbb{R}$ *satisfies* $\|g\|_W \le 1$. *Then for all* $r \ge 1$ *for which the right-hand sides below are finite we have:*

$$\sup_{|z|=r} \Big| \sum_{n=1}^{\infty} \Big(P^n g(a) - \int g\, d\pi \Big) z^n \Big|$$
$$\le H_W(r,a) \sup_{|z|\le r} \Big| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \Big| + \pi(C) \frac{H_W(r,a) - rH_W(1,a)}{r-1}$$

for all $a \in C$, *and*

$$\sup_{|z|=r} \Big| \sum_{n=1}^{\infty} \Big(P^n g(x) - \int g\, d\pi \Big) z^n \Big|$$
$$\le H_W(r,x) + G(r,x) H_W(r,a) \Big| \sup_{|z|\le r} \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \Big|$$
$$+ \pi(C) \frac{H_W(r,a) - rH_W(1,a)}{r-1} G(r,x) + \pi(C) H_W(1,a) \frac{r(G(r,x)-1)}{r-1}$$

for all $x \notin C$.

Now the problem of proving geometric convergence splits into two parts: in the first one we have to provide some estimate on $H_W(r,x)$, $x \in \mathcal{S}$, on the interval $1 \le r \le \lambda^{-1}$, and this is of importance when we want to obtain reasonable bounds on $M_W(r)$, whereas in the second part we search for r_0, a lower bound for the inverse of the radius of convergence of $\sum_{n=0}^{\infty}(u_n - u_\infty) z^n$, and then for some upper bound $K_0(r)$ on $\sup_{|z|=r} |\sum_{n=0}^{\infty}(u_n - u_\infty) z^n|$ for $r < r_0$. The second question is exactly the Kendall theorem in the setting when $R = \lambda^{-1}$, $L = \lambda^{-1} K$ (note that $b(r) = G(r, a)$ and thus $b(\lambda^{-1}) \le$

$\lambda^{-1}K$ by Proposition 4.1) and b is the bound for strong aperiodicity (i.e. $b_1 \geq b$). The additional parameter is $u_\infty = \pi(C)$; the better our knowledge of $\pi(C)$, the better the bound that comes from Theorem 2.2. If one knows the exact value of $\pi(C)$ one can use Corollary 2.3; in general, in the absence of information on $\pi(C)$, one can apply Corollary 2.5.

As for the first issue, we consider two cases. The simplest setting is when $W \equiv 1$, which implies that $H_1(r,x) = r(G(r,x)-1)/(r-1)$, $H_1(1,a) = \mathbf{E}_a\tau = \pi(C)^{-1}$. The following estimate slightly improves upon what is known for general V (cf. [2, Proposition 4.1]):

PROPOSITION 4.3. *Assume only the drift condition* (2).

(i) *For* $1 \leq r \leq \lambda^{-1}$,

$$H_1(r,x) \leq \begin{cases} \dfrac{r\lambda(V(x)-1)}{1-\lambda} & \text{if } x \notin C, \\ \dfrac{r(K-\lambda)}{1-\lambda} & \text{if } x \in C. \end{cases}$$

(ii) *For* $1 \leq r \leq \lambda^{-1}$,

$$\frac{H_1(r,a) - rH_1(1,a)}{r-1} \leq \frac{r\lambda(K-1)}{(1-\lambda)^2}.$$

Proof. To show (i) it suffices to observe that $r^{-1}H_1(r,x)$ attains its maximum on $[1,\lambda^{-1}]$ at λ^{-1}. Using Proposition 4.1 we obtain

$$r^{-1}H_1(r,x) \leq \lambda H_1(\lambda^{-1},x) = \frac{G(\lambda^{-1},x)-1}{\lambda^{-1}-1} \leq \frac{V(x)-1}{\lambda^{-1}-1}.$$

Consequently, $H_1(r,x) \leq \frac{r\lambda(V(x)-1)}{1-\lambda}$ for $x \notin C$ and in the same way we show that $H_1(r,x) \leq \frac{r(K-\lambda)}{1-\lambda}$ if $x \in C$. (ii) can be derived in a similar way: first we note that $r^{-1}(r-1)^{-1}(H_1(r,a) - rH_1(1,a))$ is increasing and then we use the bound

$$\lambda\frac{H_1(\lambda^{-1},a) - \lambda^{-1}H_1(1,a)}{\lambda^{-1}-1} \leq \frac{\frac{K-\lambda}{1-\lambda} - 1}{1-\lambda} = \frac{K-1}{(1-\lambda)^2}. \ \blacksquare$$

Combining the estimates from Propositions 4.1 and 4.3 with Proposition 4.2 and Corollaries 2.3, 2.5 we obtain our first result on atomic chains.

THEOREM 4.4. *Suppose* $(X_n)_{n\geq 0}$ *satisfies conditions* (1)–(3) *with* $\bar{b} = 1$. *Then* $(X_n)_{n\geq 0}$ *is geometrically ergodic—it satisfies* (1.2) *and we have the bounds*

$$\rho_V \leq r_0^{-1}, \qquad M_1(r) \leq \frac{2r\lambda}{1-\lambda} + \frac{r\lambda(K-1)}{(1-\lambda)^2} + \frac{r(K-\lambda)}{1-\lambda}K_0(r),$$

where $r_0 = r_0(b,\lambda^{-1},\lambda^{-1}K)$ *and* $K_0(r) = K_0(r,b,\lambda^{-1},\lambda^{-1}K)$ *are defined in Corollaries* 2.3 *and* 2.5.

On the other hand, when $W \equiv V$ there are weaker bounds on $H_V(r)$, which are stated in [2, Proposition 4.2]:

PROPOSITION 4.5. *Assume only the drift condition* (2).

(i) *For* $1 \le r \le \lambda^{-1}$,

$$H_V(r,x) \le \begin{cases} \dfrac{r\lambda(V(x)-1)}{1-r\lambda} & \text{if } x \notin C,\\ \dfrac{r(K-r\lambda)}{1-r\lambda} & \text{if } x \in C. \end{cases}$$

in particular $H_V(1,x) \le \frac{K-\lambda}{1-\lambda}$ *for all* $x \in C$.

(ii) *For* $1 \le r \le \lambda^{-1}$,

$$\frac{H_V(r,a) - rH_V(1,a)}{r-1} \le \frac{r\lambda(K-1)}{(1-\lambda)(1-r\lambda)}.$$

Using Proposition 4.5 instead of 4.3 in the proof of Theorem 4.4 we obtain a similar result, yet with a worse control on $M_W(r)$ (that necessarily goes to infinity near $r = \lambda^{-1}$).

THEOREM 4.6. *Suppose that* $(X_n)_{n\ge0}$ *satisfies conditions* (1)–(3) *with* $\bar{b} = 1$. *Then* $(X_n)_{n\ge0}$ *is geometrically ergodic—it satisfies* (1.2) *and we have the bounds*

$$\rho_V \le r_0^{-1},$$

$$M_V(r) \le \frac{r\lambda}{1-r\lambda} + \frac{r\lambda(K-\lambda)}{(1-\lambda)^2} + \frac{r\lambda(K-1)}{(1-\lambda)(1-r\lambda)} + \frac{r(K-r\lambda)}{1-r\lambda}K_0(r),$$

where $r_0 = r_0(b, \lambda^{-1}, \lambda^{-1}K)$ *and* $K_0(r) = K_0(r, b, \lambda^{-1}, \lambda^{-1}K)$ *are defined in Corollaries* 2.3 *and* 2.5.

5. Non-atomic case. For general Markov chains we have to assume that $\bar{b} \le 1$, which means that a true atom may not exist. However, there is a simple trick (cf. Meyn–Tweedie [10], Numellin [12]) which reduces this case to the atomic one. Consider the split chain $(X_n, Y_n)_{n\ge0}$ defined on the state space $\bar{\mathcal{S}} = \mathcal{S} \times \{0,1\}$ with the σ-field $\bar{\mathcal{B}}$ generated by $\mathcal{B} \times \{0\}$ and $\mathcal{B} \times \{1\}$. We define transition probabilities as follows:

$$\mathbf{P}(Y_n = 1 \mid \mathcal{F}_n^X, \mathcal{F}_{n-1}^Y) = \bar{b}1_C(X_n),$$

$$\mathbf{P}(X_{n+1} \in A \mid \mathcal{F}_n^X, \mathcal{F}_n^Y) = \begin{cases} \nu(A) & \text{if } Y_n = 1,\\ \dfrac{\mathbf{P}(X_n, A) - \bar{b}1_C(X_n)\nu(A)}{1-\bar{b}1_C(X_n)} & \text{if } Y_n = 0, \end{cases}$$

where $\mathcal{F}_n^X = \sigma(X_k : 0 \le k \le n)$ and $\mathcal{F}_n^Y = \sigma(Y_k : 0 \le k \le n)$. Thus the chain evolves in such a way that whenever X_n is in C we pick $Y_n = 1$ with probability $\bar{b}$. Then if $Y_n = 1$ we choose X_{n+1} from the ν distribution whereas if $Y_n = 0$ then we just apply the normalized probability measure version of $\mathbf{P}(X_n, \cdot) - \bar{b}1_C\nu$. The split chain is designed so that it has an atom $\mathcal{S} \times \{1\}$ and so that its first component $(X_n)_{n\geq 0}$ is a copy of the original Markov chain. Therefore we can apply the approach from the previous section to the split chain (X_n, Y_n) and the stopping time

$$T = \min\{n \geq 1 : Y_n = 1\}.$$

Let $\mathbf{P}_{x,i}$, $\mathbf{E}_{x,i}$ denote the probability and the expectation for the split chain started with $X_0 = x$ and $Y_0 = i$. Observe that for a fixed point $a \in C$ we have $\mathbf{P}_{x,1} = \mathbf{P}_{a,1}$ and $\mathbf{E}_{x,1} = \mathbf{E}_{a,1}$ for all $x \in C$. Following the method used in the atomic case we define the renewal sequence $\bar{u}_n = \mathbf{P}_{a,1}(Y_n = 1)$ and the corresponding increment sequence $\bar{b}_n = \mathbf{P}_{a,1}(T = n)$ for $n \geq 1$. Clearly $\bar{u}_n = \mathbf{P}_{a,1}(X_n \in C,\, Y_n = 1) = \bar{b}\mathbf{P}_\nu(X_{n-1} \in C)$ for $n \geq 1$, so

$$\bar{b}_1 = \bar{b}\nu(C) \geq b \quad \text{and} \quad \bar{u}_\infty = \bar{b}\pi(C). \tag{5.1}$$

We define

$$\bar{G}(r, x, i) := \mathbf{E}_{x,i}(r^T), \qquad \bar{H}_W(r, x, i) := \mathbf{E}_{x,i}\Big(\sum_{n=1}^{T} r^n W(X_n)\Big),$$

for all $x \in \mathcal{S}$, $i = 0, 1$ and all $r > 0$ for which the right hand sides are well defined. We also need the following expectation:

$$\mathbf{E}_x := (1 - \bar{b}1_C(x))\mathbf{E}_{x,0} + \bar{b}1_C(x)\mathbf{E}_{x,1},$$

which agrees with the usual $\mathbf{E}_x$ on $\mathcal{F}^X$. There exists a unique stationary measure $\bar{\pi}$ on $(\bar{\mathcal{S}}, \bar{\mathcal{B}})$ so that $\int \bar{g}\, d\bar{\pi} = \int g\, d\pi$ (where $g(x) = \bar{g}(x, 0) = \bar{g}(x, 1)$ for all $x \in \mathcal{S}$). In particular, $\bar{\pi}(\mathcal{S} \times \{1\}) = \bar{b}\pi(C)$. The first entrance last exit decomposition leads to the following result:

Proposition 5.1 ([2, Proposition 4.2]). *For all $a \in C \times \{1\}$,*

$$\sup_{|z|=r}\Big|\sum_{n=1}^{\infty}\Big(P^n\bar{g}(a) - \int g\, d\pi\Big)z^n\Big| \leq \bar{H}_W(r, a, 1)\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(\bar{u}_n - \bar{u}_\infty)z^n\Big| + \bar{b}\pi(C)\frac{\bar{H}_W(r, a, 1) - r\bar{H}_W(1, a, 1)}{r - 1}, \tag{5.2}$$

and for all $x \in \mathcal{S} \times \{0\}$,

$$\sup_{|z|=r}\Big|\sum_{n=1}^{\infty}\Big(P^n\bar{g}(x) - \int g\, d\pi\Big)z^n\Big| \leq \bar{H}_W(r, x, 0) + \bar{G}(r, x, 0)\bar{H}_W(r, a, 1)\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(\bar{u}_n - \bar{u}_\infty)z^n\Big| \tag{5.3}$$

$$+\bar{b}\pi(C)\frac{\bar{H}_W(r,a,1)-r\bar{H}_W(1,a,1)}{r-1}\bar{G}(r,x,0)$$
$$+\bar{b}\pi(C)\bar{H}_W(1,a,1)\frac{r(\bar{G}(r,x,0)-1)}{r-1}.$$

Proof. The proof mimics the proof of Proposition 4.3 in [2], the only difference being that one has to control which part of the space the point x comes from: $\mathcal{S}\times\{0\}$ or the atom $C\times\{1\}$. ■

As in the atomic case, now the problem splits into two parts. The first is to derive bounds on all the quantities $\bar{H}$ and $\bar{G}$ in Proposition 5.1. Generally it is a very tedious task, yet we detail the bounds in Appendix A improving what was known especially in the case of $W\equiv 1$. The second problem is to find bounds on r_0, the radius of convergence of $\sum_{n=0}^{\infty}(\bar{u}_n-\bar{u}_\infty)z^n$, as well as on $\bar{K}_0(r)$, the bounding constant for $\sup_{|z|=r}|\sum_{n=0}^{\infty}(\bar{u}_n-\bar{u}_\infty)z^n|$. Here the problem is that we have some information on the basic sequence yet we need control on the sequence $(\bar{b}_n)_{n\geq 0}$.

We sketch briefly what can be done about the ergodicity of $(\bar{u}_n)_{n\geq 0}$. Recall that $(\bar{u}_n)_{n\geq 0}$ is the renewal sequence for $(\bar{b}_n)_{n\geq 1}$. As in the atomic case, let $\bar{b}(z)$, $\bar{u}(z)$, $z\in C$, be the corresponding generating functions and $\bar{c}(z)=(\bar{b}(z)-1)/(z-1)$. Clearly $\bar{b}_1=\bar{b}\nu(C)\geq b$ and $\bar{c}(1)=\bar{b}^{-1}\pi(C)$ so as in the atomic case we have control on the limiting behavior of $\bar{c}(z)-\bar{c}(1)$, namely Theorem 2.2 implies that whenever $\bar{c}(r)<\infty$, then

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(\bar{u}_n-\bar{u}_\infty)z^n\Big|\leq\frac{\bar{c}(r)-\bar{c}(1)}{\bar{c}(1)(r-1)\big([(1-b)D(\bar{\alpha})-\bar{c}(r)+\bar{c}(1)]_+\big)},\tag{5.4}$$

where $\bar{c}(r)=\frac{\bar{b}(r)-1}{r-1}$, $\bar{c}(1)=\bar{u}_\infty^{-1}=\bar{b}^{-1}\pi(C)^{-1}$ and

$$D(\bar{\alpha})=\frac{\big|1+\frac{b}{1-b}(1-e^{\frac{i\pi}{1+\bar{\alpha}}})\big|-1}{|1-e^{\frac{i\pi}{1+\bar{\alpha}}}|},\quad\text{where}\quad\bar{\alpha}=\frac{\bar{c}(1)-1}{1-b}.$$

In this way the problem reduces to estimating $\bar{b}(r)$. The main difficulty is that in the non-atomic case condition (2) from the introduction together with Proposition 4.1 provides only that for $R=\lambda^{-1}>1$,

$$b_x(R)=\mathbf{E}_xR^{\tau}\leq L=KR\quad\text{for all } x\in C,\tag{5.5}$$

whereas one needs a bound on the generating function of $(\bar{b}_n)_{n\geq 1}$. We discuss this question in the Appendix, showing in Proposition A.2 that for all $1\leq r\leq\min\{R,(1-b)^{-1/(1+\alpha_1)}\}$,

$$\bar{b}(r)\leq L(r)=\max\bigg\{\frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}},\frac{br+(\bar{b}-b)r^{1+\alpha_2}}{1-(1-\bar{b})r}\bigg\},\tag{5.6}$$

where $\alpha_1=\log\big(\frac{L-\bar{b}R}{(1-\bar{b})R}\big)/\log R$ and $\alpha_2=\log\big(\frac{L-(1-\bar{b}+b)R}{(\bar{b}-b)R}\big)/\log R$. Moreover if

$1+b \geq 2\bar{b}$ then simply

$$L(r) = \frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}}.$$

Using (5.6) is the best what the renewal approach can offer to bound $\bar{b}(r)$. The meaning of the result is that there are only two generating functions that are important to bound $\bar{b}(r)$. If $\bar{b}$ is close to 1 then we are in a similar setting to the atomic case and surely one can expect a bound on $\bar{b}(r)$ of the form $\frac{br+(\bar{b}-b)r^{1+\alpha_2}}{1-(1-\bar{b})r}$, whereas if $\bar{b}$ is far from 1 only the split chain construction matters and the bound on $\bar{b}(r)$ should be like $\frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}}$.

As in the atomic case, we will need a bound on $\bar{\alpha} = \frac{\bar{c}(1)-1}{1-b}$. We show in Corollary A.3 that

$$\bar{\alpha} \leq \bar{b}^{-1} \max\left\{\frac{1-\bar{b}}{1-b}(1+\alpha_1), \frac{1-\bar{b}}{1-b} + \frac{\bar{b}-b}{1-b}\alpha_2\right\}. \tag{5.7}$$

In fact the maximum equals $\bar{b}^{-1}\frac{1-\bar{b}}{1-b}(1+\alpha_1)$ if $1+b \geq 2\bar{b}$, and $\bar{b}^{-1}\frac{1-\bar{b}}{1-b} + \frac{\bar{b}-b}{1-b}\alpha_2$ otherwise.

Now we turn to the basic idea for all the approach presented in the paper, i.e. a certain convexity property of the function $r \mapsto c(r)$. Observe that $\frac{\bar{c}(r)-1}{\bar{c}(1)-1}$ satisfies the Hölder inequality, i.e. for $p+q=1$, $p,q>0$,

$$\left(\frac{\bar{c}(r_1)-1}{\bar{c}(1)-1}\right)^p \left(\frac{\bar{c}(r_2)-1}{\bar{c}(1)-1}\right)^q \geq \frac{\bar{c}(r_1^p r_2^q)-1}{\bar{c}(1)-1},$$

which means that $F_0(x) = \log\left(\frac{\bar{c}(e^x)-1}{\bar{c}(1)-1}\right)$ is convex and $F_0(0)=0$. By (5.6) we have $\bar{c}(e^x) \leq L(e^x)$ and hence

$$F_0(x) \leq F_1(x) = \log\left(\frac{L(e^x)-e^x}{(1-b)\bar{\alpha}(e^x-1)}\right). \tag{5.8}$$

Therefore we can easily compute the largest possible function $\bar{F}(x)$ that satisfies the conditions:

1. $\bar{F}(x) \leq F_1(x)$ for $0 \leq x \leq \min\{\log R, -\frac{1}{1+\alpha_1}\log(1-\bar{b})\}$;
2. $\bar{F}(0)=0$ and $\bar{F}$ is convex;
3. $\bar{F}$ is maximal of all the functions with properties 1–2, namely if there exists F that satisfies the above conditions then $F(x) \leq \bar{F}(x)$ for all $0 \leq x \leq \min\{\log R, -\frac{1}{1+\alpha_1}\log(1-\bar{b})\}$.

The role of $\bar{F}$ is to answer the question: how to find a suitable value of $e^x \in [1,R]$ and a suitable bound on $\bar{b}(e^x)$ to apply our main Kendall theorem. Under the basic data contained in conditions 1–3 of the introduction one can propose an upper bound F_1 on F_0. On the other hand we may benefit from the fact that F_0 is convex and starts from zero. Consequently, we consider all the possible functions that have these properties. It occurs that there is a maximizer $\bar{F}$ in this class and this function should be considered as a generator of the optimal bound on $\bar{b}(e^x)$ one should apply in Theorem 2.2. Namely

$$\bar{b}(e^x) \leq (1-b)\bar{\alpha}(e^x-1)\exp(\bar{F}(x)) + e^x$$

for all $0 \leq x \leq \min\{\log R, -\frac{1}{1+\alpha_1}\log(1-\bar{b})\}$.

Let x_0 be the unique solution of the equation

$$F_1'(x)x = F_1(x). \tag{5.9}$$

Note that $x_0 \leq -\frac{1}{1+\alpha_1}\log(1-\bar{b})$. If additionally $x_0 \leq \log R$ then the optimal $\bar{F}(x)$ is of the form

$$\bar{F}(x) = \begin{cases} F_1'(x_0)x & \text{for all } 0 \leq x \leq x_0, \\ F_1(x) & \text{for all } x_0 \leq x \leq \min\{\log R, -\frac{1}{1+\alpha_1}\log(1-\bar{b})\}; \end{cases} \tag{5.10}$$

otherwise if $x_0 > \log R$ then

$$\bar{F}(x) = \frac{F_1(\log R)}{\log R}x \quad \text{for all } 0 \leq x \leq \log R. \tag{5.11}$$

To make the notation similar to the atomic case let $\bar{\kappa}(\bar{\alpha}, r) = \bar{F}(\log r)/\log r$. In particular if $\log r < x_0 \leq \log R$ then $\bar{\kappa}(\bar{\alpha}, r) = F_1'(x_0)$ and similarly $\bar{\kappa}(\alpha, r) = \bar{F}(\log R)/\log R$ if $\log R < x_0$. The above discussion leads to the following conclusion:

$$\bar{c}(r)-\bar{c}(1) \leq (1-b)\bar{\alpha}r^{\kappa(\bar{\alpha},r)} \quad \text{for all } 1 \leq r \leq \min\{R, (1-\bar{b})^{-1/(1+\alpha_1)}\}, \tag{5.12}$$

furthermore $\bar{\kappa}(\bar{\alpha}, r)$ as a function of r is constant at least on part of the interval $[1, \min\{R, (1-\bar{b})^{-1/(1+\alpha_1)}\}]$. Consequently, applying (5.4) for the case where $\pi(C)$ is known, we obtain our main estimate in the non-atomic case.

THEOREM 5.2. *Suppose that $\bar{b}_1 \geq b$ and $\bar{b}(r)$ satisfies (5.6), and $\bar{u}_\infty = \bar{b}\pi(C)$ is known. Then $\sum_{n=0}^{\infty}(\bar{u}_n - \bar{u}_\infty)z^n$ is convergent for $|z| < r_0$, where*

$$\bar{r}_0 = \min\{R, (1-\bar{b})^{-1/(1+\alpha_1)}, \bar{r}_0(\bar{\alpha})\},$$

where $\bar{r}_0(\bar{\alpha})$ is the unique solution of the equation

$$r = (1 + D(\bar{\alpha})/\bar{\alpha})^{1/\bar{\kappa}(\bar{\alpha},r)}.$$

Moreover, for $r < \bar{r}_0$,

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \leq \bar{K}_0(r) = \frac{\bar{u}_\infty(r^{\bar{\kappa}(\bar{\alpha},r)}-1)}{(r-1)(\bar{\alpha}D(\bar{\alpha}) - r^{\bar{\kappa}(\bar{\alpha},r)}+1)}.$$

REMARK 5.3. Observe that if

$$\log(1 + D(\bar{\alpha})/\bar{\alpha})/F_1'(x_0) \leq x_0 \leq \log R, \tag{5.13}$$

then $\bar{r}_0 = (1 + D(\bar{\alpha})/\bar{\alpha})^{1/F_1'(x_0)}$. Due to (5.9), the condition (5.13) is equivalent to $x_0 \le \log R$ and

$$1 + \frac{D(\bar{\alpha})}{\alpha} \le \frac{L(e^{x_0}) - e^{x_0}}{(1-b)\bar{\alpha}(e^{x_0}-1)}.$$

Therefore for a large class of examples we have a computable direct bound on the rate of convergence even for general ergodic Markov chains.

If $\bar{u}_\infty = \bar{b}\pi(C)$ is unknown then we have to treat it as a parameter and use a bound on $\bar{\alpha}$. As for the upper bounds, we can use (5.7); on the other hand we show in Corollary A.3 that if $\bar{b}\nu(C) = b$ then $\bar{\alpha} \ge \bar{b}^{-1}$. Since in the same way as in the atomic case $(1 + D(\bar{\alpha})/\bar{\alpha})^{1/\bar{\kappa}(\bar{\alpha},r)}$ increases with b assuming that $\bar{b}, L, R$ are fixed, we can always assume $\bar{\alpha} \ge \bar{b}^{-1}$. Let $\bar{\alpha}_0 = \max\{\frac{1-\bar{b}}{1-b}(1+\alpha_1), \frac{1-\bar{b}}{1-b} + \frac{\bar{b}-b}{1-b}\alpha_2\}$.

Theorem 5.4. *Let $\bar{b}_1 \ge b$, and suppose that $\bar{b}(r)$ satisfies* (5.6). *Then $\sum_{n=0}^{\infty}(\bar{u}_n - \bar{u}_\infty)z^n$ is convergent for $|z| < \bar{r}_0$, where*

$$\bar{r}_0 = \min\Big\{R, (1-\bar{b})^{-1/(1+\alpha_1)}, \min_{\bar{b}^{-1} \le \bar{\alpha} \le \bar{b}^{-1}\bar{\alpha}_0} \bar{r}_0(\bar{\alpha})\Big\},$$

where $\bar{r}_0(\bar{\alpha})$ is the unique solution of the equation

$$r = (1 + D(\bar{\alpha})/\bar{\alpha})^{1/\bar{\kappa}(\bar{\alpha},r)}.$$

Moreover, for $r < \bar{r}_0$,

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\Big| \le \bar{K}_0(r) = \max_{\bar{b}^{-1} \le \bar{\alpha} \le \bar{b}^{-1}\bar{\alpha}_0} \frac{\bar{b}(r^{\bar{\kappa}(\bar{\alpha},r)} - 1)}{(r-1)(\bar{\alpha}^{-1}D(\bar{\alpha}) - r^{\bar{\kappa}(\bar{\alpha},r)} + 1)}.$$

We show by examples that the approach presented in Theorems 5.2 and 5.4 is comparable with the coupling method (see [2, Section 7] for a short introduction). Therefore we obtain a computable tool for the general question of rates of convergence of ergodic Markov chains under the geometric drift condition.

We postpone the detailed computation of all the bounds required in Proposition 5.1 to Appendix A. This knowledge enabled us to formulate the main results for general Markov chains. The first one concerns the case of $W \equiv 1$. By Proposition 5.1 and Propositions A.2, A.5 from Appendix A we obtain the first result for general Markov chains.

Theorem 5.5. *Suppose $(X_n)_{n\ge 0}$ satisfies conditions* (1)–(3) *from the introduction. Then $(X_n)_{n\ge 0}$ is geometrically ergodic—it satisfies* (1.2) *and*

$$\rho_V \le \bar{r}_0^{-1},$$

$$M_1(r) \le \frac{2\lambda r}{1-\lambda} + \frac{2(1-\bar{b})(r^{1+\alpha_1}-1)r}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})} + \frac{\bar{b}}{1-(1-\bar{b})r^{1+\alpha_1}}\frac{r\lambda(K-1)}{(1-\lambda)^2} + \frac{(r-1)\bar{K}_0(r) + \bar{b}(1-\bar{b})(r^{1+\alpha_1}-1)}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})^2}\,\frac{r(K-\lambda)}{1-\lambda},$$

where $\bar{K}_0(r) = \bar{K}_0(r, b, \bar{b}, \lambda^{-1}, K\lambda^{-1})$ *and* $\bar{r}_0 = \bar{r}_0(b, \bar{b}, \lambda^{-1}, K\lambda^{-1})$ *are given in Theorems* 5.2 *and* 5.4.

Proof. Note that $\bar{b}\pi(C)\bar{H}_1(1, a, 1) = 1$. We apply Proposition 5.1 so that we sum (5.2) with weight $1 - \bar{b}1_C(x)$ and (5.3) with weight $\bar{b}1_C(x)$. Then we use (A.5) to bound $\bar{b}1_C(x) + (1-\bar{b}1_C(x))\bar{G}(r, x, 0)$, and (A.19) to bound $(1-\bar{b}1_C(x))\bar{H}_1(r, x, 0) = (1-\bar{b}1_C(x))\frac{r\bar{G}(r,x,0)-1}{r-1}$. Finally (A.20) and (A.21) are estimates for $\bar{H}(r, a, 1)$ and $(\bar{H}_1(r, a, 1) - r\bar{H}_1(1, a, 1))/(r-1)$. ∎

The second case is when $W \equiv V$. Propositions 5.1, A.2 and A.6 imply our result in the most general form.

THEOREM 5.6. *Suppose* $(X_n)_{n\ge0}$ *satisfies conditions* (1)–(3) *from the introduction. Then* $(X_n)_{n\ge0}$ *is geometrically ergodic—it satisfies* (1.2) *and*

$$\rho_V \le \bar{r}_0^{-1}$$

and

$$M_V(r) \le \frac{\lambda r}{1-r\lambda} + \left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{r}{1-(1-\bar{b})r^{1+\alpha_1}} + \frac{K-\lambda}{1-\lambda}\left(\frac{r\lambda}{1-\lambda} + \frac{(1-\bar{b})(r^{1+\alpha_1}-1)r}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})}\right) + \frac{\bar{b}}{1-(1-\bar{b})r^{1+\alpha_1}}\left(\frac{r(K-1)}{(1-\lambda)(1-r\lambda)} + \left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{1}{1-(1-\bar{b})r^{1+\alpha_1}}\frac{r(K-\lambda)}{1-\lambda}\right) + \frac{\bar{K}_0(r)}{1-(1-\bar{b})r^{1+\alpha_1}}\left(\frac{r(K-r\lambda)}{1-r\lambda} + \left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{r-1}{1-(1-\bar{b})r^{1+\alpha_1}}\,\frac{r(K-\lambda)}{1-\lambda}\right),$$

where $\bar{K}_0(r) = \bar{K}_0(r, b, \bar{b}, \lambda^{-1}, K\lambda^{-1})$ *and* $\bar{r}_0 = \bar{r}_0(b, \bar{b}, \lambda^{-1}, K\lambda^{-1})$ *are given in Theorems* 5.2 *and* 5.4.

Proof. Observe that $\pi(C) \le 1$. As in the proof of Theorem 5.5, we use Proposition 5.1, summing (5.2) with weight $1 - \bar{b}1_C(x)$ and (5.3) with weight $\bar{b}1_C(x)$. Again we use (A.5) to bound $\bar{b}1_C(x) + (1-\bar{b}1_C(x))\bar{G}(r, x, 0)$, then (A.25), (A.26), (A.27) to bound respectively $(1-\bar{b}1_C(x))\bar{H}_V(r, x, 0)$, $\bar{H}_V(r, a, 1)$ and $(\bar{H}_V(r, a, 1) - r\bar{H}_V(1, a, 1))/(r-1)$. We also use the bound

$\bar{H}_V(1,a,1) \le \bar{b}^{-1}\frac{K-\lambda}{1-\lambda}$ and (A.19) to bound $(1-\bar{b}1_C(x))\bar{H}_1(r,x,0) = (1-\bar{b}1_C(x))\frac{r\bar{G}(r,x,0)-1}{r-1}$. ■

Appendix A

A.1. Global bounds. Our method described in Corollary 5.2 implies that

$$\sup_{|z|=r}\Big|\sum_{n=0}^{\infty}(\bar{u}_n-\bar{u}_\infty)z^n\Big| \le K_0(r) \quad \text{for } 1\le r\le r_0.$$

The first step is to replace the stopping time T by $\tau=\tau_C$. For this, we define

$$G(r,x,i)=\mathbf{E}_{x,i}r^\tau, \qquad H_W(r,x,i)=\mathbf{E}_{x,i}\Big(\sum_{n=1}^{\tau}r^nW(X_n)\Big).$$

Let also

$$G(r)=\sup_{x\in C}\mathbf{E}_{x,0}r^\tau \quad\text{and}\quad H_W(r)=\sup_{x\in C}\mathbf{E}_{x,0}\sum_{n=1}^{\tau}r^nW(X_n).$$

In [2, Lemma A.1] the following inequalities are proved:

Proposition A.1. *For $r\le\lambda^{-1}$ and $(1-\bar{b})G(r)<1$,*

$$\bar{G}(r,x,i)\le\frac{\bar{b}G(r,x,i)}{1-(1-\bar{b})G(r)}, \tag{A.1}$$

$$\bar{H}_W(r,x,i)\le H_W(r,x,i)+\frac{(1-\bar{b})H_W(r)G(r,x,i)}{1-(1-\bar{b})G(r)}. \tag{A.2}$$

In the introduction we have explained that it is crucial for our approach to establish (5.6). Now we have all the necessary tools to get that result.

Proposition A.2. *For all $a\in C$ and $1\le r\le\min\{\lambda^{-1},(1-\bar{b})^{-1/(1+\alpha_1)}\}$,*

$$\bar{G}(r,a,1)\le\max\Big\{\frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}},\frac{br+(\bar{b}-b)r^{\alpha_2}}{1-(1-\bar{b})r}\Big\}, \tag{A.3}$$

where $\alpha_1=\log\big(\frac{K-\bar{b}}{1-\bar{b}}\big)/\log\lambda^{-1}$ and $\alpha_2=\log\big(\frac{K-1+\bar{b}-b}{\bar{b}-b}\big)/\log\lambda^{-1}$. Moreover, if $1+b\ge2\bar{b}$, then

$$\bar{G}(r,a,1)\le\frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}}. \tag{A.4}$$

For all $x\in\mathcal{S}$ and $1\le r\le\min\{\lambda^{-1},(1-\bar{b})^{-1/(1+\alpha_1)}\}$,

$$\bar{b}1_C(x)+(1-\bar{b}1_C(x))\bar{G}(r,x,0)\le\frac{\bar{b}V(x)}{1-(1-\bar{b})r^{1+\alpha_1}}. \tag{A.5}$$

Proof. The split chain construction implies that for any $a\in C$,

$$(1-\bar{b})\sup_{x\in C}G(r,x,0)+\bar{b}G(r,a,1)=\sup_{x\in C}G(r,x)=G(r). \tag{A.6}$$

Moreover, due to $\bar{b}\nu(C) \geq b$ we have $\bar{b}G(r,a,1) = \bar{b}\sum_{k=1}^{\infty} \mathbf{P}_\nu(\sigma = k-1)r^k$, where $\sigma = \inf\{n \geq 0 : X_n \in C\}$, has its first coefficient greater than or equal to b. Therefore by our usual argument with the Hölder inequality we deduce that

$$(1-\bar{b}) \sup_{x\in C} G(r,x,0) \leq (1-\bar{b})rv^{\frac{\log r}{\log \lambda^{-1}}}, \qquad \bar{b}G(r,a,1) \leq br + (\bar{b}-b)ru^{\frac{\log r}{\log \lambda^{-1}}},$$

where $u = G(\lambda^{-1}, a, 1)$, $v = \sup_{x\in C} G(\lambda^{-1}, x, 0)$ satisfy

$$\text{(A.7)} \quad \lambda^{-1}(b + (\bar{b}-b)u + (1-\bar{b})v) = \sup_{x\in C} G(\lambda^{-1}, x) \leq K\lambda^{-1}, \qquad u, v \geq 1.$$

Observe that by (A.1),

$$\text{(A.8)} \qquad \bar{G}(r,a,1) \leq \frac{\bar{b}G(r,a,1)}{1-(1-\bar{b})G(r)} \leq F(u,v) = \frac{br + (\bar{b}-b)ru^{\frac{\log r}{\log \lambda^{-1}}}}{1-(1-\bar{b})rv^{\frac{\log r}{\log \lambda^{-1}}}}.$$

One can check that the bounding function $F(u,v)$ is convex for all (u,v) that satisfy (A.7) and hence it takes its maximum on the boundaries of the set given by (A.7). Consequently, due to (A.8) we obtain

$$\text{(A.9)} \qquad \bar{G}(r,a,1) \leq \max\left\{\frac{\bar{b}r}{1-(1-\bar{b})r^{1+\alpha_1}}, \frac{br + (\bar{b}-b)r^{1+\alpha_2}}{1-(1-\bar{b})r}\right\}.$$

It is easy to check that whenever $1 + b \geq 2\bar{b}$ then $(1-\bar{b})\alpha_1 \geq (\bar{b}-b)\alpha_2$ and the maximum in (A.9) can be replaced by the first quantity for any $r \geq 1$. Otherwise if $1 + b < 2\bar{b}$ then $(1-\bar{b})\alpha_1 < (\bar{b}-b)\alpha_2$ and therefore for small enough r the maximum in (A.9) is attained at the second expression.

We turn to showing the second assertion. Observe that by Proposition 4.1 we have $G(r,x,0) = G(r,x) \leq V(x)$ for all $x \notin C$. Consequently, (A.1) yields

$$\text{(A.10)} \qquad \bar{G}(r,x,0) \leq \frac{\bar{b}V(x)}{1-(1-\bar{b})G(r)}$$

for all $x \notin C$. Since obviously $G(r) \leq r^{1+\alpha_1}$ we deduce that

$$\bar{G}(r,x,0) \leq \frac{\bar{b}V(x)}{1-(1-\bar{b})r^{\alpha_1}}.$$

On the other hand, by (A.1),

$$\bar{G}(r,x,0) \leq \frac{\bar{b}r^{\alpha_1}}{1-(1-\bar{b})r^{1+\alpha_1}}$$

for all $x \in C$ and therefore

$$\text{(A.11)} \qquad \bar{b} + (1-\bar{b})\bar{G}(r,x,0) \leq \frac{\bar{b}}{1-(1-\bar{b})r^{1+\alpha_1}}$$

for all $x \in C$. Since $V \geq 1$, inequalities (A.10) and (A.11) imply (A.5). ■

The next step is to obtain the estimate (5.13).

COROLLARY A.3. *The following inequality holds:*

$$\bar{b}^{-1} \le \frac{\bar{H}_1(1,a,1)-1}{1-b} \le \bar{b}^{-1}\max\left\{\frac{1-\bar{b}}{1-b}(1+\alpha_1), \frac{1-\bar{b}}{1-b}+\frac{\bar{b}-b}{1-b}\alpha_2\right\} = \bar{b}^{-1}\bar{\alpha}_0. \tag{A.12}$$

Proof. For the first inequality, we simply apply (A.3) to bound $\bar{H}_1(r,a,1) = \frac{r\bar{G}(r,a,1)-1}{r-1}$ and then let $r \to 1$. To prove the second inequality let $S = \max\{k \ge 1 : \tau_k \le T\}$, where τ_k, $k \ge 0$, are the successive visits to C by $(X_n)_{n\ge 0}$, in particular $\tau_0 = 0$. Observe that

$$\bar{H}_1(1,a,1) = \mathbf{E}_{a,1}\Big(\sum_{k=0}^{\infty} 1_{S\ge k}(\tau_k - \tau_{k-1})\Big).$$

Therefore by construction

$$\bar{H}_1(1,a,1) \ge \mathbf{E}_\nu(1+\sigma) + \mathbf{E}_{a,1}(S-1),$$

where we recall that $\sigma = \min\{n \ge 0 : X_n \in C\}$. Since S has geometric distribution with the probability of success $\bar{b}$, we obtain

$$\bar{H}_1(1,a,1) \ge \bar{b}^{-1} + \mathbf{E}_\nu\sigma.$$

It remains to notice that $\mathbf{E}_\nu\sigma \ge 1-\nu(C)$, therefore if $\bar{b}\nu(C) = b$ then

$$\bar{H}_1(1,a,1) \ge \bar{b}^{-1} + 1 - \frac{b}{\bar{b}},$$

which completes the proof. ■

Now we state an improvement of the result mentioned in the proof of Proposition 4.4 in [2].

PROPOSITION A.4. *For $r \le \lambda^{-1}$ and $(1-\bar{b})G(r) < 1$ we have*

$$\bar{H}_W(r,a,1) \le \frac{1}{\bar{b}}\sup_{x\in C} H_W(r,x) + \frac{1-\bar{b}}{\bar{b}}\,\frac{H_W(r)\sup_{x\in C}(G(r,x)-1)}{1-(1-\bar{b})G(r)} \tag{A.13}$$

and

$$\bar{H}_W(r,a,1) - r\bar{H}_W(1,a,1) \le \frac{1}{\bar{b}}\sup_{x\in C}(H_W(r,x,0) - rH_W(1,x,0)) + \frac{1-\bar{b}}{\bar{b}}H_W(r)(\bar{G}(r,a,1)-1). \tag{A.14}$$

Proof. To prove the first assertion note that (A.6) can be rewritten as

$$\frac{\bar{b}G(r,a,1)}{1-(1-\bar{b})G(r)} \le 1 + \frac{\sup_{x\in C}(G(r,x)-1)}{1-(1-\bar{b})G(r)}.$$

Combining the above inequality with (A.2) we derive

$$\bar{H}_W(r,a,1) \le H_W(r,a,1) + \frac{1-\bar{b}}{\bar{b}} H_W(r) + \frac{(1-\bar{b})H_W(r)\sup_{x\in C}(G(r,x)-1)}{\bar{b}(1-(1-\bar{b})G(r))}.$$

Since the definition of $H_W(r,x,1)$ implies that

$$\bar{b}H_W(r,a,1) + (1-\bar{b})H_W(r) \le \sup_{x\in C} H_W(r,x)$$

we obtain (A.13).

To show the second assertion we use $S = \max\{k \ge 1 : \tau_k \le T\}$ defined in the proof of Corollary A.3. Then

$$\text{(A.15)} \quad \bar{H}_W(r,a,1) - r\bar{H}_W(1,a,1) \le H_W(r,a,1) - rH_W(1,a,1) + \sum_{k=2}^{\infty} \mathbf{E}_{a,1}\Big[1_{k\le S} \sup_{x\in C}(r^{\tau_{k-1}}H_W(r,x,0) - rH_W(1,x,0))\Big].$$

As shown in Corollary A.3, $\mathbf{E}_{a,1}(S-1) = (1-\bar{b})/\bar{b}$, and we deduce that

$$\sum_{k=2}^{\infty}(\mathbf{E}_{a,1}1_{k\le N}) \sup_{x\in C}(H_W(r,x,0) - rH_W(1,x,0)) = \frac{1-\bar{b}}{\bar{b}} \sup_{x\in C}(H_W(r,x,0) - rH_W(1,x,0)),$$

which together with (A.15) provides

$$\text{(A.16)} \quad \bar{H}_W(r,a,1) - r\bar{H}_W(1,a,1) \le H_W(r,a,1) - rH_W(1,a,1) + \frac{1-\bar{b}}{\bar{b}} \sup_{x\in C}(H_W(r,x,0) - rH_W(1,x,0)) + \sum_{k=2}^{\infty}[\mathbf{E}_{a,1}1_{k\le N}(r^{\tau_{k-1}}-1)] \sup_{x\in C} H_W(r,x,0).$$

As usual we observe that

$$\text{(A.17)} \quad H_W(r,a,1) - rH_W(1,a,1) + \frac{1-\bar{b}}{\bar{b}} \sup_{x\in C}(H_W(r,x,0) - rH_W(1,x,0)) \le \frac{1}{\bar{b}} \sup_{x\in C}(H_W(r,x) - rH_W(1,x)).$$

Moreover, since Y_{τ_k} is independent of τ_{k-1} we have

$$\mathbf{E}_{a,1}r^{\tau_{k-1}}1_{k\le S} = (1-\bar{b})\mathbf{E}_{a,1}r^{\tau_{k-1}}1_{k-1\le S},$$

which implies that

$$\sum_{k=2}^{\infty}\mathbf{E}_{a,1}r^{\tau_{k-1}}1_{S=k-1}=\bar{b}\sum_{k=2}^{\infty}\mathbf{E}_{a,1}r^{\tau_{k-1}}1_{k-1\le S}=\frac{b}{1-\bar{b}}\sum_{k=2}^{\infty}\mathbf{E}r^{\tau_{k-1}}1_{k\le S},$$

and thus

$$\sum_{k=2}^{\infty}\mathbf{E}_{a,1}1_{k\le S}(r^{\tau_{k-1}}-1)=\frac{1-\bar{b}}{\bar{b}}\mathbf{E}_{a,1}(r^{T}-1)=\frac{1-\bar{b}}{\bar{b}}(\bar{G}(r,a,1)-1).$$

Consequently,

(A.18)
$$\sum_{k=2}^{\infty}[\mathbf{E}_{a,1}1_{k\le N}(r^{\tau_{k-1}}-1)]\sup_{x\in C}H_W(r,x,0)=\frac{1-\bar{b}}{\bar{b}}H_W(r)(\bar{G}(r,a,1)-1).$$

Combining (A.16)–(A.18) we conclude that

$$\begin{aligned}\bar{H}_W(r,a,1)-r\bar{H}_W(1,a,1)\le{}&\frac{1}{\bar{b}}\sup_{x\in C}(H_W(r,x,0)-rH_W(1,x,0))\\&+\frac{1-\bar{b}}{\bar{b}}H_W(r)(\bar{G}(r,a,1)-1).\end{aligned}$$

This completes the proof of (A.14). ∎

A.2. Case of $W\equiv 1$. In the case of $W\equiv 1$, the above result improves the estimation of $\bar{H}_1(r,a,1)-r\bar{H}_1(1,a,1)$, which, as mentioned in the introduction, can be used in the part of the proof where $\sup_{|z|=r}|\sum_{n=0}^{\infty}(\bar{u}_n-\bar{u}_{\infty})|$ is considered.

PROPOSITION A.5. *The following inequalities hold:*

$$(1-\bar{b}1_C(x))\bar{H}_1(r,x,0)\le\frac{r\lambda(V(x)-1)}{1-\lambda}+\frac{(1-\bar{b})(r^{1+\alpha_1}-1)rV(x)}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})}\tag{A.19}$$

for all $x\in\mathcal{S}$, $1\le r\le\min\{\lambda^{-1},(1-\bar{b})^{-1/(1+\alpha_1)}\}$;

$$\bar{H}_1(r,a,1)\le\frac{1}{1-(1-\bar{b})r^{1+\alpha_1}}\,\frac{r(K-\lambda)}{1-\lambda}\tag{A.20}$$

for all $a\in C$ and $1\le r\le\min\{\lambda^{-1},(1-\bar{b})^{-1/(1+\alpha_1)}\}$; and

$$\begin{aligned}\frac{\bar{H}_1(r,a,1)-r\bar{H}_1(r,a,1)}{r-1}\le{}&\frac{1}{\bar{b}}\,\frac{r\lambda(K-1)}{(1-\lambda)^2}\\&+\frac{1}{\bar{b}}\,\frac{(1-\bar{b})(r^{1+\alpha_1}-1)}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})}\,\frac{r(K-\lambda)}{1-\lambda}\end{aligned}\tag{A.21}$$

for all $a\in C$ and $1\le r\le\min\{\lambda^{-1},(1-\bar{b})^{-1/(1+\alpha_1)}\}$.

Proof. By (A.2) we have

$$\bar{H}_1(r,x,0) \le H_1(r,x,0) + \frac{(1-\bar{b})H_1(r)G(r,x,0)}{1-(1-\bar{b})G(r)}.$$

Together with $H_1(r,x,0) = H_1(r,x)$ and $G(r,x,0) = G(r,x) \le V(x)$ for $x \notin C$ it follows that

$$\bar{H}_1(r,x,0) \le H_1(r,x) + \frac{(1-\bar{b})H_1(r)V(x)}{1-(1-\bar{b})G(r)}.$$

Consequently, by Proposition 4.3,

$$\bar{H}_1(r,x,0) \le \frac{r\lambda(V(x)-1)}{1-\lambda} + \frac{(1-\bar{b})(G(r)-1)rV(x)}{(r-1)(1-(1-\bar{b})G(r))}.$$

Using $G(r) \le r^{1+\bar{\alpha}}$ we deduce that

$$\bar{H}_1(r,x,0) \le \frac{r\lambda(V(x)-1)}{1-\lambda} + \frac{(1-\bar{b})(r^{1+\alpha_1}-1)rV(x)}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})} \tag{A.22}$$

for all $x \notin C$. On the other hand, (A.2) implies that

$$\bar{H}_1(r,x,0) \le H_1(r)\left(1 + \frac{(1-\bar{b})G(r)}{1-(1-\bar{b})G(r)}\right) = \frac{H_1(r,x,0)}{1-(1-\bar{b})G(r)} \quad \text{for } x \in C.$$

Hence again by $G(r) \le r^{1+\alpha_1}$,

$$(1-\bar{b})\bar{H}_1(r,x,0) \le \frac{(1-\bar{b})(r^{1+\alpha_1}-1)r}{(r-1)(1-(1-\bar{b})r^{1+\alpha_1})} \quad \text{for } x \in C. \tag{A.23}$$

From (A.22) and (A.23) we deduce (A.19).

Observe that (A.1) and (A.6) imply that

$$\begin{aligned} \bar{H}_1(r,a,1) = \frac{r(\bar{G}(r,a,1)-1}{r-1} &\le \frac{\sup_{x\in C} r(G(r,x)-1)}{(r-1)(1-(1-\bar{b})G(r))} \\ &= \frac{\sup_{x\in C} H_1(r,x)}{1-(1-\bar{b})G(r)} \end{aligned} \tag{A.24}$$

and therefore

$$\bar{H}_1(r,a,1) \le \frac{\sup_{x\in C}(H_1(r,x))}{1-(1-\bar{b})r^{1+\alpha_1}} \le \frac{r(K-\lambda)}{(1-\lambda)(1-(1-\bar{b})r^{1+\alpha_1})},$$

which is (A.20).

To prove the last assertion we use (A.14) and (A.24), which imply that

$$\begin{aligned} \bar{H}_1(r,a,1) - r\bar{H}_1(1,a,1) \le{}& \frac{1}{\bar{b}} \sup_{x\in C}(H_1(r,x) - rH_1(r,x)) \\ &+ \frac{(1-\bar{b})H_1(r)\sup_{x\in C}(G(r,x)-1)}{1-(1-\bar{b})G(r)}. \end{aligned}$$

The above inequality is equivalent to

$$\bar{H}_1(r,a,1) - r\bar{H}_1(1,a,1) \le \frac{1}{\bar{b}} \sup_{x\in C}(H_1(r,x) - rH_1(r,x)) + \frac{(1-\bar{b})(G(r)-1)\sup_{x\in C} H_1(r,x)}{1-(1-\bar{b})G(r)}.$$

Clearly $H_1(r,x) = \frac{r(G(r,x)-1)}{r-1}$, thus by Proposition 4.3 we obtain

$$\frac{\bar{H}_1(r,a,1) - r\bar{H}_1(1,a,1)}{r-1} \le \frac{1}{\bar{b}} \frac{r\lambda(K-1)}{(1-\lambda)^2} + \frac{(1-\bar{b})(G(r)-1)}{\bar{b}(r-1)(1-(1-\bar{b})G(r))} \frac{r(K-\lambda)}{1-\lambda}.$$

Due to $G(r) \le r^{1+\alpha_1}$ we deduce (A.21), completing the proof. ∎

A.3. Case of $W \equiv V$. The second case we consider is when $W = V$.

PROPOSITION A.6. *The following inequalities hold:*

(A.25)
$$(1-\bar{b}1_C(x))\bar{H}_V(r,x,0) \le \frac{\lambda r(V(x)-1)}{1-r\lambda} + \left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{rV(x)}{1-(1-\bar{b})r^{1+\alpha_1}}$$

for all $x \in \mathcal{S}$ and $1 \le r \le \min\{\lambda^{-1}, (1-\bar{b})^{-1/(1+\alpha_1)}\}$;

(A.26)
$$\bar{H}_V(r,a,1) \le \bar{b}^{-1}\frac{r(K-r\lambda)}{1-r\lambda} + \bar{b}^{-1}\left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{r-1}{1-(1-\bar{b})r^{1+\alpha_1}} \frac{r(K-\lambda)}{1-\lambda}$$

for all $a\in C$ and $1\le r\le\min\{\lambda^{-1}, (1-\bar{b})^{-1/(1+\alpha_1)}\}$, in particular $\bar{H}_V(1,a,1) \le \bar{b}^{-1}\frac{K-\lambda}{1-\lambda}$; and

$$\frac{\bar{H}_V(r,a,1) - r\bar{H}(1,a,1)}{r-1} \le \bar{b}^{-1}\frac{r(K-1)}{(1-\lambda)(1-r\lambda)} + \bar{b}^{-1}\left(\frac{K-r\lambda}{1-r\lambda} - \bar{b}\right)\frac{1}{1-(1-\bar{b})r^{1+\alpha_1}} \frac{r(K-\lambda)}{1-\lambda} \tag{A.27}$$

for all $a \in C$ and $1 \le r \le \min\{\lambda^{-1}, (1-\bar{b})^{-1/(1+\alpha_1)}\}$.

Proof. We recall that (A.2) implies that

$$\bar{H}_V(r,x,0) \le H_V(r,x,0) + \frac{(1-\bar{b})H_V(r)G(r,x,0)}{1-(1-\bar{b})G(r)}.$$

Therefore since $H_V(r,x,0) = H_V(r,x)$ and $G(r,x,0) = G(r,x)$ for all $x \notin C$, we can use Propositions 4.1 and 4.5 to get

$$\bar{H}_V(r,x,0) \le \frac{\lambda r(V(x)-1)}{1-r\lambda} + \frac{(1-\bar{b})H_V(r)V(x)}{1-(1-\bar{b})G(r)}$$

for all $x \notin C$. Similarly for $x \in C$,

$$(1-\bar{b})\bar{H}_V(r,x,0) \le (1-\bar{b})H_V(r)\Big(1+\frac{(1-\bar{b})G(r)}{1-(1-\bar{b})G(r)}\Big) = \frac{(1-\bar{b})H_V(r)}{1-(1-\bar{b})G(r)}.$$

Hence using $G(r) \le r^{1+\alpha_1}$ we obtain

$$(1-\bar{b}1_C(x))\bar{H}_V(r,x,0) \le \frac{\lambda r(V(x)-1)}{1-r\lambda} + \frac{(1-\bar{b})H_V(r)V(x)}{1-(1-\bar{b})r^{1+\alpha_1}}.$$

Therefore it suffices to bound $(1-\bar{b})H_V(r)$. Note that

$$\bar{b}H_V(r,x,1) + (1-\bar{b})H_V(r) \le \sup_{x\in C} H_V(r,x) \quad \text{for all } x \in C.$$

Clearly $H_V(r,x,1) \ge r$, so by Proposition 4.5 we deduce that

$$(1-\bar{b})H_V(r) \le \frac{r(K-r\lambda)}{1-r\lambda} - \bar{b}r, \tag{A.28}$$

which establishes (A.25).

To show the remaining assertions we use (A.13), (A.14) and (A.6), obtaining

$$\bar{H}_V(r,a,1) \le \bar{b}^{-1}\sup_{x\in C} H_V(r,x) + \bar{b}^{-1}\frac{(1-\bar{b})H_V(r)\sup_{x\in C}(G(r,x)-1)}{1-(1-\bar{b})G(r)}$$

and

$$\begin{aligned}\bar{H}_V(r,a,1) - r\bar{H}_V(r,a,1) \le \bar{b}^{-1}&\sup_{x\in C}(H_V(r,x) - rH_V(1,x)) \\ &+ \bar{b}^{-1}\frac{(1-\bar{b})H_V(r)\sup_{x\in C}(G(r,x)-1)}{1-(1-\bar{b})G(r)}.\end{aligned}$$

Recall that $G(r) \le r^{1+\alpha_1}$ and by Propositions 4.1 and 4.5,

$$\frac{G(r,x)-1}{r-1} \le \frac{K-\lambda}{1-\lambda}, \qquad H_V(r,x) \le \frac{r(K-r\lambda)}{1-r\lambda};$$

consequently,

$$\bar{H}_V(r,a,1) \le \bar{b}^{-1}\frac{r(K-r\lambda)}{1-r\lambda} + \bar{b}^{-1}\frac{(r-1)(1-\bar{b})H_V(r)}{1-(1-\bar{b})r^{1+\alpha_1}}\,\frac{K-\lambda}{1-\lambda}.$$

Together with (A.28) this completes the proof of (A.26).

Finally, the same argument shows

$$\begin{aligned}\frac{\bar{H}_V(r,a,1) - r\bar{H}_V(r,a,1)}{r-1} \le \bar{b}^{-1}&\frac{r(K-1)}{(1-\lambda)(1-r\lambda)} \\ &+ \bar{b}^{-1}\frac{(1-\bar{b})H_V(r)}{1-(1-\bar{b})r^{1+\alpha_1}}\,\frac{K-\lambda}{1-\lambda}.\end{aligned}$$

Again by (A.28) we obtain (A.27), which completes the proof. ■

Appendix B. We compare our result with what was shown in [2] as a numerical test for the presented approach.

B.1. The reflecting random walk. We consider the Bernoulli random walk on $\mathbb{Z}_+$ with transition probabilities $P(i, i-1) = p > 1/2$, $P(i, i+1) = q = 1-p$ for $i \geq 1$ and boundary conditions $P(0,0) = p$, $P(0,1) = q$. We set $C = \{0\}$ and $V(i) = (p/q)^{i/2}$, and compute $\lambda = 2\sqrt{pq}$, $K = p + \sqrt{pq}$, $b = p$ and $u_\infty = \pi(C) = 1 - q/p$. The optimal radius of convergence for the reflecting random walk is λ.

Consider two cases:

(1) $p = 2/3$, so $b = 2/3$, $\lambda = 2\sqrt{2}/3$, $K = (2+\sqrt{2})/3$, $u_\infty = 1/2$.

(2) $p = 0.9$, and hence $\lambda = 0.6$, $K = 1.2$, $b = 0.9$, $u_\infty = 8/9$.

We compare our result with others in Table 1 below, where ρ and ρ_C denote estimates on the radius of convergence when u_∞ is known and when it is not. We use *Optimal* for the true value of the spectral radius, and *Bednorz*, *Baxendale*, *Meyn–Tweedie1* and *Meyn–Tweedie2* respectively for our Corollaries 2.3 and 2.5, Baxendale's Theorem 3.2 of [2], Meyn–Tweedie's result of [11] and its improved version (see [2, Section 8] for details).

Table 1

$p = 2/3$	ρ	ρ_C
Optimal	0.9428	0.9428
Bednorz	0.9737	0.9737
Baxendale	0.9994	?
Meyn–Tweedie1	0.9999	0.9988
Meyn–Tweedie2	0.9991	0.9927
$p = 0.9$	ρ	ρ_C
Optimal	0.6	0.6
Bednorz	0.6	0.6
Baxendale	0.9060	?
Meyn–Tweedie1	0.9967	0.9888
Meyn–Tweedie2	0.9470	0.9467

B.2. Metropolis–Hastings algorithm for the normal distribution. In this example we consider the convergence of a Metropolis–Hastings algorithm in the case when we want to simulate $\pi = \mathcal{N}(0,1)$ with candidate transition probability $q(x, \cdot) = \mathcal{N}(x, 1)$. The example was studied in [11] and also in [14, 15]. By the algorithm definition $\mathbf{P}(x,\cdot)$ is distributed with density

$$p(x,y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}\exp\Big(-\dfrac{(y-x)^2}{2}\Big) & \text{if } |x|\ge|y|,\\ \dfrac{1}{\sqrt{2\pi}}\exp\Big(-\dfrac{(y-x)^2+y^2-x^2}{2}\Big) & \text{if } |x|\le|y|.\end{cases}$$

The natural setting of the problem is to consider Lyapunov functions of the type $V(x)=e^{s|x|}$ and $C=[-d,d]$. Consequently (see [2] for details),

$$\lambda=\frac{PV(d)}{V(d)},\qquad K=PV(d)=e^{sd}\lambda.$$

The computed value for ρ depends on d and s, and hence we need to find the optimal ones. Moreover, to compare our result with the previous contributions to the problem, let ν be given by

$$\nu(dx)=c\exp(-x^2)1_C(x)dx$$

for a suitable normalizing constant c. In this case, $\nu(C)=1$ and we have

$$b=\bar{b}=\sqrt{2}\exp(-d^2)[\Phi(\sqrt{2}\,d)-1/2].$$

In this case we work with the additional complication of the splitting construction. The results are compared in Table 2, where again *Bednorz1* and *Bednorz2* denote our Theorems 5.4 and 5.2 (depending on whether or not we use the additional information on $\pi(C)$), *Baxendale* denotes what can be obtained by Baxendale's [2, Theorem 3.2], *Coupling* denotes the estimate obtained by the coupling approach (see in [2, Section 7] and [17]), and *Meyn–Tweedie* the result obtained in the original paper [11]. Note that we compare methods where no additional assumptions on the transition probabilities are made.

Table 2

	d	s	$1-\rho$
Bednorz1	0.96	0.065	0.00000529
Bednorz2	0.92	0.169	0.00005496
Baxendale	1	0.13	0.00000063
Coupling	1.8	1.1	0.00068
Meyn–Tweedie	1.4	0.00004	0.000000016

Another possible choice of ν is

$$\bar{b}\nu(dx)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}\exp\Big(-\dfrac{(|x|+d)^2}{2}\Big)dx & \text{if } |x|\le d,\\ \dfrac{1}{\sqrt{2\pi}}\exp(-d|x|-|x|^2)dx & \text{if } |x|\ge d.\end{cases}$$

In this case

$$b = 2(\Phi(2d) - \Phi(d)) \quad \text{and} \quad \bar{b} = b + \sqrt{2}\exp(d^2/4)(1 - \Phi(3d/\sqrt{2})).$$

Using the same notation as in Table 2 we compare the results below.

Table 3

	d	s	$1-\rho$
Bednorz1	1.03	0.0733	0.00001061
Bednorz2	0.97	0.1740	0.00013637
Baxendale	1	0.16	0.0000017
Coupling	1.9	1.1	0.00187

Observe that our method is bit worse than coupling yet it is relatively simple (does not require further examination of the Lyapunov function V).

B.3. Contracting normals. Here we consider the family of Markov chains with transition probability $\mathbf{P}(x,\cdot) = \mathcal{N}(\theta x, 1-\theta^2)$ for some parameter $\theta \in (-1,1)$. This family occurs in [16] as a component of a two-component Gibbs sampler. The example was discussed in [2], [14] and [15]. Here we take $V(x) = 1 + x^2$ and $C = [-c, c]$. Then (2) is satisfied with

$$\lambda = \theta^2 + 2\frac{1-\theta^2}{1+c^2}, \qquad K = 2 + \theta^2(c^2 - 1).$$

We choose ν concentrated on C so that

$$\bar{b}\nu(dy) = \min_{x\in C} \frac{1}{\sqrt{2\pi(1-\theta^2)}} \exp\biggl(-\frac{(\theta x - y)^2}{2(1-\theta^2)}\biggr) dy$$

for $y \in C$. Integrating with respect to y gives

$$\bar{b} = 2\biggl(\Phi\biggl(\frac{(1+|\theta|)c}{\sqrt{1-\theta^2}}\biggr) - \Phi\biggl(\frac{|\theta|c}{\sqrt{1-\theta^2}}\biggr)\biggr).$$

We compare our answer *Bednorz1*, *Bednorz2* (Theorems 5.4, 5.2 resp.) with the coupling method *Coupling* and *Baxendale2*, an approach based on a Kendall-type result (Theorem 3.3 in [2]) that requires invertibility of the transition function.

Table 4

	θ	c	$-\rho$
Bednorz1	0.5	1.5	0.000872023152
Bednorz1	0.75	1.2	0.000000964524
Bednorz1	0.9	1.1	0.000000000004
Bednorz2	0.5	1.5	0.002754672439
Bednorz2	0.75	1.2	0.000017954821
Bednorz2	0.9	1.1	0.000000000881

Table 4 (cont.)

	θ	c	$-\rho$
Baxendale2	0.5	1.5	0.050
Baxendale2	0.75	1.2	0.0042
Baxendale2	0.9	1.1	0.00002
Coupling	0.5	2,1	0.054
Coupling	0.75	1.7	0.0027
Coupling	0.9	1.5	0.00002

B.4. Reflecting random walk, continued. Here we slightly redefine our first example. Let $\mathbf{P}(0,\{0\}) = 1$ and $\mathbf{P}(0,\{1\}) = 1-\varepsilon$ for some $\varepsilon > 0$. We concentrate on the difficult case, when $\varepsilon < p$, studied in [15] and [5]. Note that when $\varepsilon \geq p$, the chain is stochastically monotone and then the result of Tweedie [9] applies. Let $V(i) = (p/q)^{i/2}$ and $C = \{0\}$ as earlier. Then $\lambda = 2\sqrt{pq}$, $K = \varepsilon + (1-\varepsilon)\sqrt{p/q}$ and $b = \varepsilon$. In this example we can calculate the formula for $b(z)$:

$$\begin{aligned} b(z) &= G(z,0) = \varepsilon z + (1-\varepsilon)zG(z,1) \\ &= \varepsilon z + \frac{1-\varepsilon}{2q}(1-(1-4pqz^2)^{1/2}) \end{aligned} \tag{B.1}$$

for $|z| < 1/\sqrt{4pq}$, where the formula for $G(z,1)$ is in [4]. Consequently,

$$\pi(\{0\})^{-1} = b'(1) = \varepsilon + \frac{2p(1-\varepsilon)}{p-q}.$$

On the other hand (B.1) leads to the optimal bound of the radius on convergence:

$$\rho = \begin{cases} \dfrac{pq+(p-\varepsilon)^2}{p-\varepsilon} & \text{if } \varepsilon < \dfrac{p-q}{1+\sqrt{q/p}}, \\ 2\sqrt{pq} & \text{otherwise.} \end{cases}$$

We compare *Bednorz1*, *Bednorz2* (our Corollaries 2.3, 2.5) with *Fort* and *Baxendale* that denotes respectively the result of Fort [5] and Baxendale's [2, Theorem 1.2]. Note that both methods use further properties of transition probability in this particular example.

Table 5

$p = 0.6$	$\varepsilon = 0.05$	$\varepsilon = 0.25$	$\varepsilon = 0.5$
Optimal	0.9864	0.9798	0.9798
Bednorz1	0.99993	0.9994	0.99783
Bednorz2	0.99993	0.9994	0.9977
Fort	0.9997	0.9995	0.9994
Bax	0.9909	0.9798	0.9798

Table 5 (cont.)

$p = 0.7$	$\varepsilon = 0.05$	$\varepsilon = 0.25$	$\varepsilon = 0.5$
Optimal	0.9165	0.9165	0.9165
Bednorz1	0.9992	0.9940	0.9783
Bednorz2	0.9991	0.9935	0.9779
Fort	0.9964	0.9830	0.9757
Bax	0.9731	0.9165	0.9165
$p = 0.8$	$\varepsilon = 0.05$	$\varepsilon = 0.25$	$\varepsilon = 0.5$
Optimal	0.9633	0.8409	0.8000
Bednorz1	0.9970	0.9780	0.9266
Bednorz2	0.9964	0.9751	0.9253
Fort	0.9793	0.9333	0.9333
Bax	0.9759	0.8796	0.8000
$p = 0.9$	$\varepsilon = 0.05$	$\varepsilon = 0.25$	$\varepsilon = 0.5$
Optimal	0.9559	0.7885	0.6250
Bednorz1	0.9927	0.9489	0.8408
Bednorz2	0.9899	0.9358	0.8280
Fort	0.9696	0.8539	0.7500
Bax	0.9687	0.8470	0.6817
$p = 0.95$	$\varepsilon = 0.5$	$\varepsilon = 0.25$	$\varepsilon = 0.5$
Optimal	0.9528	0.7679	0.5556
Bednorz1	0.9888	0.9249	0.7827
Bednorz2	0.9841	0.9024	0.7537
Fort	0.9564	0.7853	0.5814
Bax	0.9645	0.7853	0.5814

Acknowledgments. This research was partially supported by MNiSW Grant N N201 387234.

References

[1] R. Adamczak and W. Bednorz, *Exponential concentration inequalities for additive functionals of Markov chains*, arXiv:1201.3569.
[2] P. Baxendale, *Renewal theory and computable convergence rates for geometrically ergodic Markov chains*, Ann. Appl. Probab. 15 (2005), 700–738.
[3] K. S. Bernhaut and R. Lund, *Geometric renewal convergence rates from hazard rates*, J. Appl. Probab. 38 (2001), 180–194.
[4] W. Feller, *An Introduction to Probability Theory and its Applications 1*, 3rd ed., Wiley, New York, 1968.
[5] G. Fort, *Computable bounds for V-ergodicity of Markov transition kernels*, preprint, 2003.

[6] D. G. Kendall, *Unitary dilations of Markov transition operators and the corresponding integral representation of transition probability matrices*, in: Probability and Statistics, U. Grenander (ed.), Almqvist and Wiksell, Stockholm, 1959, 138–161.

[7] K. Łatuszyński, B. Miasojedow and W. Niemiro, *Nonasymptotic bounds on the estimation error for regenerative MCMC algorithms*, Bernoulli, to appear; arXiv:0307.4915.

[8] K. Łatuszyński and W. Niemiro, *Rigorous confidence bounds for MCMC under a geometric drift condition*, J. Complexity 27 (2011), 23–38.

[9] R. B. Lund and R. L. Tweedie, *Geometric convergence rates for stochastically ordered Markov chains*, Math. Oper. Res. 21 (1996), 182–194.

[10] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer, Berlin, 1993.

[11] S. Meyn and R. Tweedie, *Computable bounds for geometric convergence rates of Markov chains*, Ann. Appl. Probab. 4 (1994), 981–1011.

[12] E. Nummelin, *General Irreducible Markov Chains and Non-Negative Operators*, Cambridge Univ. Press, 1984.

[13] E. Nummelin and P. Tuominen, *Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory*, Stoch. Process. Appl. 12 (1982), 187–202.

[14] G. O. Roberts and J. S. Rosenthal, *Shift-coupling and convergence rates of ergodic averages*, Comm. Statist. Stoch. Models 13 (1999), 147–165.

[15] G. O. Roberts and R. L. Tweedie, *Bounds on regeneration times and convergence rates for Markov chains*, Stoch. Process. Appl. 80 (1999), 211–229.

[16] J. S. Rosenthal, *Minorization conditions and convergence rates for Markov chain Monte Carlo*, J. Amer. Statist. Assoc. 90 (1995), 558–566.

[17] J. S. Rosenthal, *Quantitative convergence rates of Markov chains. A simple account*, Electron. Comm. Probab. 7 (2002), 123–128.

Witold Bednorz
Institute of Mathematics
Warsaw University
02-097 Warszawa, Poland
E-mail: wbednorz@mimuw.edu.pl

Received on 31.8.2012;
revised version on 12.3.2013 (2149)