Abstract

In this paper we prove an improved quantitative version of the Kendall’s Theorem. The Kendall Theorem states that under mild conditions imposed on a probability distribution on positive integers (i.e. probabilistic 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 the 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 \(P^n(x, \cdot), n \geq 0\) and a unique stationary measure \(\pi\). Let \(P\) be the transition operator given on the Banach space of bounded measurable functions on \((\mathcal{S}, \mathcal{B})\) by \(Pf(x) = \int f(y)P(x, dy)\). Under mild conditions imposed on \((X_n)_{n \geq 0}\) the chain is ergodic, i.e.

\[
\|P^n(x, \cdot) - \pi(\cdot)\|_{TV} \to 0, \quad \text{as} \quad n \to \infty,
\]

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

\[
\|\mu\|_{TV} = \sup_{|f| \leq 1} \left| \int f \, d\mu \right|,
\]
where $\mu$ is a real measure on $(\mathcal{S}, \mathcal{B})$. It is known that the aperiodicity, the Harris recurrence property and the finiteness of $\pi$ are equivalent to (1.1), (see Theorem 13.0.1 in [12]). 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 needs to apply the ergodicity for the MCMC estimators there is required a stronger form of the result, namely one expects the exponential rate of the convergence and a reasonable method to estimate this rate (cf. [11]).

One of the possible generalizations of the total variation convergence is to consider functions controlled from above by $V : \mathcal{S} \to \mathbb{R}$, $V \geq 1$, $\pi(V) < \infty$, therefore we refer to $B_V$ as 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

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

The geometric convergence of $P^ng(x, \cdot)$ to a unique stationary measure $\pi$, means there exists $\rho_V < r \leq 1$ such that

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

where $\rho_V$ 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

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

which is valid at least on $\rho_V \leq r \leq 1$, 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

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

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

Whenever it exists we call $\rho_V$ the convergence rate of geometric ergodicity for the chain $(X_n)_{n \geq 0}$. For a class of examples one can prove the geometric
convergence (see Chapter 15 in [12]) 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 \( \pi \)-measure.

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

1. **Minorization condition.** There exist \( C \in \mathcal{B}, \tilde{b} > 0 \) and a probability measure \( \nu \) on \((S, \mathcal{B})\) such that
   \[
   P(x, A) \geq \tilde{b} \nu(A)
   \]
   for all \( x \in C \) and \( A \in \mathcal{B} \).

2. **Drift condition.** There exist a measurable function \( V : S \rightarrow [1, \infty) \) and constants \( \lambda < 1 \) and \( K < \infty \) satisfying
   \[
   PV(x) \leq \begin{cases} 
   \lambda V(x) & \text{if } x \notin C \\
   K & \text{if } x \in C.
   \end{cases}
   \]

3. **Strong aperiodicity** There exists \( b > 0 \) such that \( \tilde{b} \nu(C) \geq b > 0 \).

The first property means there exists a small set \( C \) on which the regeneration of \( (X_n)_{n \geq 0} \) takes place (see Chapter 5 in [12]). The assumption is relatively weak since each Harris recurrent chain admits the existence of a small set at least for some of its \( m \)-skeletons (i.e. processes \( (X_{nm})_{n \geq 0}, m \geq 1 \) - see Theorem 5.3.2 in [12]. The small set existence is used in the split chain construction (see Section 3 and cf. [10] for details) to extend \( (X_n)_{n \geq 0} \) to a new Markov Chain on a larger probability space \( S \times \{0, 1\} \), so that \( (C, 1) \) is a true atom of the new chain and its marginal distribution on \( S \) equals the distribution of \( (X_n)_{n \geq 0} \). The second condition reads as the existence of a Lyapunov function \( V \) which is contracted by the semigroup related operator \( P \) with the rate \( \lambda < 1 \), for all points outside the small set. Finally the 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 turn assuming the chain is in the set \( C \).

Our main result concerns convergence rates of ergodic Markov chains. Since the approach is based on the reduction to the study of renewal sequences, we first prove an abstract theorem that treats renewal sequences and which strengthen previous forms of the result (known as the Kendall’s theorem).
Only then we turn to analyze the atomic case and show how to apply the idea to the case when a true atom exists and which is the natural setting for the approach. However, the concept is valid for the general Harris chains. It requires additional work - the split chain construction. Results of this type are used whenever exact estimates on the ergodicity matters cf. [3], [8] and [9].

The organization of the paper is as follows: the history of the abstract Kendall’s theorem as well as our main improvement of the result 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’s theorem affects estimation of convergence rates for atomic Markov chains; using the method of the chain split we extend the results in Section 5 on general Harris chains; we leave the tedious computation of required estimates on constants (which improves the previous results of this type) to the Appendix A; finally in Appendix B we analyze the result for basic toy examples.

2 The abstract Kendall’s theorem

Let \((\tau_k)_{k \geq 0}\) be a random walk on \(\mathbb{N}\) starting from zero, i.e. \(\tau_0 = 0, \tau_k - \tau_{k-1}, k \geq 1\) are independent distributed like \(\tau\), namely

\[
P(\tau_k - \tau_{k-1} = n) = P(\tau = n) = b_n, \ n \geq 1.
\]

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 application’s point of view such a random walk is generated by subsequent 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 \(u_n = P(V_n = 0), n \geq 0\) i.e. the probability that the process \((V_m)_{m \geq 0}\) renews (goes to zero) in the \(n\)-th time step. The sequence \((u_n)_{n \geq 0}\) is of meaning for the study of ergodic properties for Markov chains which will be the main issue for next chapters. In particular note that \(u_n\) equals the probability that the suitable atomic Markov chain stays in the given atom in the \(n\)-th time step. Observe that \(u_0 = 1\) and \(u_n = \sum_{k=1}^{n} u_{n-k} b_k\), hence denoting \(b(z) = \sum_{n=1}^{\infty} b_n z^n, u(z) = \sum_{n=0}^{\infty} u_n z^n\), for \(z \in \mathbb{C}\), one can state the renewal equation as follows

\[
(2.1) \quad u(z) = 1/(1 - b(z)), \ \text{for} \ |z| < 1.
\]
The equation means that to study the properties of \((u_n)_{n \geq 0}\) it suffices to concentrate on properties of \((b_n)_{n \geq 1}\). In particular one can ask when the sequence \((u_n)_{n \geq 0}\) is ergodic which means \(\lim_{n \to \infty} u_n\) exists. Historically, the first result that matches these properties with the geometric ergodicity was due to Kendall [6] who proved that:

**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 is equal \(u_\infty = (\sum_{n=1}^{\infty} nb_n)^{-1}\), moreover the radius of convergence of \(\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\) is strictly greater than 1.

The Kendall’s theorem states that the sequence \((u_n)_{n \geq 0}\) is ergodic whenever \(b(z)\) is bounded on the disc of radius strictly greater then 1 and we have slight control on \(b_1\). However, the question is: does Theorem 2.1 implies any rates of the convergence? It obviously requires basic information what is the upper bound on \(b(z)\), i.e. \(b(R) \leq L\) for a given \(R > 1\) and what is the lower bound on \(b_1 \geq b > 0\). The data \(b, R, L\) stems from the conditions 1-3 formulated in the introduction, especially they 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 \(u_\infty\) having the information on \(b, R, L\). This is the isolated abstract Kendall’s-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’s theorem was improved first in [13] and then in [1] (see Theorem 3.2). There are also several results where some additional assumptions on the distribution of \(\tau\) are made. For example there is elaborated in [2] how to provide an optimal bound on the rate of convergence, yet under additional conditions on the \(\tau\) distribution. Whenever the general Kendall’s question is considered the bounds obtained up to now are still far from being optimal or easy to use. The goal of the 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 \(u_\infty\) as a parameter, namely we prove that the following result holds:

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

\[
\sup_{|z|=r} \left| \sum_{n=0}^{\infty}(u_n - u_\infty)z^n \right| \leq \frac{c(r) - c(1)}{c(1)(r-1)((1-b)D(\alpha) - c(r) + c(1))_+},
\]
where \( c(r) = \frac{b(r)-1}{r-1} \), \( c(1) = u_\infty^{-1} \) and
\[
D(\alpha) = \frac{|1 + \frac{b}{1-b}(1 - e^{\frac{\pi}{1+b}})| - 1}{|1 - e^{\frac{\pi}{1+b}}|}, \quad \text{where} \quad \alpha = \frac{c(1) - 1}{1 - b}.
\]

**Proof.** Let \( b(z) \) and \( u(z) \) be the complex generic functions for \( b_i, i \geq 1 \) and \( u_i, i \geq 0 \) sequences respectively. The main tool we use is the renewal equation (2.1), i.e.
\[
1 - b(z) = \frac{1}{u(z)}, \quad |z| < 1.
\]

Note that the equation remains valid on the disc \(|z| \leq R\) in the sense of analytic functions. By Theorem 2.1 we learn that \( u_\infty < \infty \) and the renewal generic function \( \sum_{n=0}^\infty (u_n - u_\infty)z^n \) is convergent on some disc with radius greater than 1. Denote \( c(z) = \frac{b(z)-1}{z-1} \) (cf. proof of Theorem 3.2 in [1]) and observe that \( c(z) \) is well defined on \(|z| \leq R\), because \( c(R) = \frac{b(R)-1}{R-1} = \frac{R-1}{R-1} < \infty \). Since \( u_\infty = c(1)^{-1} \) we have that
\[
\sum_{n=0}^\infty (u_n - u_\infty)z^n = u(z) - \frac{1}{c(1)(1 - z)} = \frac{1}{1 - b(z)} - \frac{1}{c(1)(z - 1)} = \frac{1}{1 - z} \left( \frac{c(z)}{c(1)} - \frac{1}{c(1)} \right) = \frac{c(z) - c(1)}{z - 1} \frac{1}{c(1)c(z)}.
\]
(2.2)

The main problem is to estimate \(|c(z)|\) from below, to which goal we use the simple technique
\[
|c(re^{i\theta})| = |c(e^{i\theta})| - |c(re^{i\theta}) - c(e^{i\theta})| = |c(e^{i\theta})| - c(r) + c(1).
\]
(2.3)

Consequently the problem is reduced to the study of the lower bound on \(|c(e^{i\theta})|\). We recall that by the definition \( c_l = \sum_{j=1}^l b_j \) and \( c(1) = \sum_{j=0}^\infty c_j \). To provide a sharp estimate in (2.3) we benefit from the fact that for \( \frac{\pi}{l+1} < |\theta| < \frac{\pi}{l} \), \( l \geq 1 \), there is a better control on the first \( l \) summands in \( c(e^{i\theta}) = \sum_{j=1}^\infty c_je^{ij\theta} \). First we note that
\[
|c(e^{i\theta})| = \left| \frac{1 - \sum_{j=1}^\infty b_je^{ij\theta}}{1 - e^{i\theta}} \right| \geq \left| \frac{1 - \sum_{j=1}^l b_je^{ij\theta} - \sum_{j=l+1}^\infty b_j}{1 - e^{i\theta}} \right|,
\]
which is equivalent to
\[
|c(e^{i\theta})| \geq \frac{|c_l + \sum_{j=1}^l b_j(1 - e^{ij\theta})|}{|1 - e^{i\theta}|}.
\]

The geometrical observation gives that for \( \frac{\pi}{l+1} < |\theta| < \frac{\pi}{l} \)
\[
|c_l + \sum_{j=1}^l b_j(1 - e^{ij\theta})| \geq |c_l + \sum_{j=1}^l b_j(1 - e^{ij\theta})| = |c_l + (1 - c_l)(1 - e^{i\theta})|,
\]
where \( c_l = \frac{b_l-1}{l-1} \), \( c_l(1) = u_\infty^{-1} \) and
\[
D(\alpha) = \frac{|1 + \frac{b}{1-b}(1 - e^{\frac{\pi}{1+b}})| - 1}{|1 - e^{\frac{\pi}{1+b}}|}, \quad \text{where} \quad \alpha = \frac{c(1) - 1}{1 - b}.
\]

\[\Box\]

W. Bednorz
Rates of convergence

hence we conclude that

\[ |c(e^{i\theta})| \geq c_l|1 - e^{i\theta}|^{-1}(|1 + (1 - c_l)c_l^{-1}(1 - e^{i\theta})| - 1). \]

Since \(1 - c_l \geq b\), for \(l \geq 1\) we see that

\[ |1 + (1 - c_l)c_l^{-1}(1 - e^{i\theta})| \geq \sqrt{1 + bc_l^{-2}|1 - e^{i\theta}|^2}. \]

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

(2.4) \[ f'(x) = -x^{-2}(\sqrt{1 + bx^2} - 1) + x^{-2}\frac{bx^2}{\sqrt{1 + 4bx^2}} \geq 0. \]

Therefore we finally obtain that for \(\frac{\pi}{l+1} < |\theta| \leq \frac{\pi}{l}\)

(2.5) \[ |c(e^{i\theta})| \geq c_l|1 - e^{i\frac{\pi}{l+1}}|(\sqrt{1 + bc_l^{-2}|1 - e^{i\frac{\pi}{l+1}}|^2} - 1). \]

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

(2.6) \[ c_l|1 - e^{i\frac{\pi}{l+1}}|^{-1} \leq (1 - b)|1 - e^{i\frac{\pi}{l+\alpha}}|^{-1}, \]

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

(2.7) \[ c_l|1 - e^{i\frac{\pi}{l+1}}|^{-1} \geq (c_l)(l|1 - e^{i\frac{\pi}{l+1}}|)^{-1} \geq (c_l)(\alpha|1 - e^{i\frac{\pi}{l+\alpha}}|)^{-1}. \]

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

(2.8) \[ c_l \leq \sum_{j=1}^{l} c_j \leq c(1) - 1 = \alpha(1 - b) \]

and thus combining (2.7) and (2.8) we obtain that

\[ c_l|1 - e^{i\frac{\pi}{l+1}}|^{-1} \leq (1 - b)|1 - e^{i\frac{\pi}{l+\alpha}}|^{-1}, \]

which is (2.6). As we have noted the bound used in (2.4) implies that

\[ |c(e^{i\theta})| \geq (1 - b)|1 - e^{i\frac{\pi}{l+1}}|^{-1}(\sqrt{1 + b(1 - b)^{-2}|1 - e^{i\frac{\pi}{l+1}}|^2} - 1), \]

which is equivalent to

(2.9) \[ |c(e^{i\theta})| \geq |1 - e^{i\frac{\pi}{l+\alpha}}|^{-1}((1 - b) + b(1 - e^{i\frac{\pi}{l+\alpha}}) - (1 - b)). \]
Moreover, for
\[ \kappa \leq 1 \]
which will be our basic setting. Note that by the Hölder inequality, for all
\[ \sum_{n=1}^{\infty} |c_n u| \leq \sum_{n=1}^{\infty} |c_n|^{\alpha} u^{\alpha} \leq c(1) \]
where \( D(\alpha) = |1 - e^{\frac{\pi i}{\alpha}}|^{-1} |1 + \frac{b}{1 - b} (1 - e^{\frac{\pi i}{\alpha}})| - 1 \) 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 exposition is when \( c(1) = u^{-1}_\infty \) is known and we can control \( c(r) \) in a certain point, i.e. \( c(R) \leq N < \infty \), for some \( R > 1 \). Observe that if \( b(R) \leq L \), then due to \( c(R) = \frac{b(R)}{R-1} - 1 \) one derives that \( c(R) \leq N = \frac{L-1}{R-1} \), which will be our basic setting. Note that by the Hölder inequality, for all \( 1 \leq r \leq R \)
\[ c(r) - c(1) = (c(1) - 1)(\frac{c(r)}{c(1)} - 1) \leq (1 - b)\alpha r^{\kappa(\alpha)} - 1, \]
where \( \kappa(\alpha) = \log(\frac{N-1}{c(1)-1})/\log R = \log(\frac{N-1}{(1-b)\alpha})/\log R, \alpha = (c(1)-1)/(1-b) \).

We summarize this concept in the following assertion:

**Corollary 2.3.** Suppose that \( c(1) = (u^{-1}_\infty) \) is known, \( b_1 \geq b \) and \( b(R) \leq L \), then \( \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \) is convergent for \( |z| < r_0 \), where
\[ r_0 = \min\{R, (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\kappa(\alpha)}} \}. \]
Moreover for \( r < r_0 \)
\[ \sup_{|z|=r} \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \leq K_0(r) = \frac{u_\infty (r^{\kappa(\alpha)} - 1)}{(r-1)(\alpha^{-1}D(\alpha) - r^{\kappa(\alpha)} + 1)}. \]

**Remark 2.4.** Observe that the bound \( (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\kappa(\alpha)}} \) is increases with \( b \) assuming that \( L, R, c(1) \) are fixed.

In applications we have to treat \( c(1) = u^{-1}_\infty \) as a parameter. The advantage of the approach is that there is a sharp upper bound on \( c(1) \) or rather \( \alpha = (c(1) - 1)/(1-b) \). Using the inequality
\[ (2.11) \]
\[ R^\alpha = R^{\Sigma_{n=1}^{(n-1)h_n}/(1-b)} \leq \frac{\sum_{n=2}^{\infty} b_n R^{n-1}}{1-b} \leq \frac{b(R) - bR}{(1-b)R} \leq \frac{L - bR}{(1-b)R}, \]
we deduce that $\alpha \leq \alpha_0$, where $
olimits \alpha_0 = \log \frac{L - bR}{(1 - b)R} / \log R$. On the other hand if $b = b_1$, then $c(1) - 1 \geq 1 - b$ and therefore by Remark 2.4 we can always require that $c(1) - 1 \geq 1 - b$ or equivalently $\alpha \geq 1$. Therefore to find an estimate on the rate of convergence we search $(1 + \alpha \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha}}$, $\alpha \in [1, \alpha_0]$ for the possible minimum.

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

$$r_0 = \min\{R, \min_{1 \leq \alpha \leq \alpha_0} (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha}}\}.$$ 

Moreover for $r < r_0$

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

The above Corollary should be compared with Theorem 3.2. in [1] 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

$$r_0 = \min\{R, \hat{r}_0\}, \quad \hat{r}_0 = \min_{1 \leq \alpha \leq \alpha_0} (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha}}.$$ 

As it will be shown, this estimate is always better than the main bound in [1] (Theorem 3.2). Then we turn to study the reason for this improvement. Using the limit case with $b, L$ fixed and $R \to 1$, we check that the minimum of $\alpha \to (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha}}$ can be attained in the interval $[1, \alpha_0]$ and that it is data depending problem one cannot avoid. On the other hand, we stress that in the usual setting the minimum of $\alpha \to (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha}}$ should be attained at $\alpha_0$. The intuition for this phenomenon is that the smaller is $c(1) = u_\infty^{-1}$ 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 rest of the data: $b, R$.

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

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

Obviously, to find the minimum on the interval $[1, \alpha_0]$, the solution of (3.1) must be compared with 1 and $\alpha_0$. Consequently $\hat{r}_0 = (1 + D(1))^{\frac{1}{\alpha_0}}$ when
such $\alpha$ is smaller than 1 and $\hat{r}_0 = (1 + \frac{D(\alpha_0)}{\alpha_0})^{-\frac{1}{\alpha(\alpha_0)}}$ when it is bigger than $\alpha_0$, otherwise the solution of (3.1) is the worst possible $\alpha$ 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 $\alpha(D(\alpha))^{-1}(r^{\alpha(\alpha)} - 1)$ which is attained at the unique point $\alpha$ that satisfies the equation

$$\alpha = \frac{\log r}{\log R} r^{\alpha(\alpha)}.$$  

(3.2)

To find the maximum of $\alpha(D(\alpha))^{-1}(r^{\alpha(\alpha)} - 1)$ on $[1, \alpha_0]$ we compare the solution of (3.2) with 1 and $\alpha_0$. If such $\alpha$ is greater than $\alpha_0$ then

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

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

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

$$1 + \frac{D'(1)}{D(1)}(R^{\alpha(1)} - 1) \geq R^{\alpha(1)}.$$  

We turn to show computable bounds on $K_0(r)$ in the case when $u_\infty$ is unknown. Note that function $D(\alpha)$ is decreasing and therefore $D(\alpha) \geq D(\alpha_0)$. Consequently one can rewrite Corollary 2.5 with $D(\alpha)$ replaced by $D(\alpha_0)$ and in this way obtain new bounds: $K_1(r) \geq K_0(r)$ and $r_1 \leq r_0$, where $r_1 = \min\{R, \hat{r}_1\}$, $\hat{r}_1 = \min_{1 \leq \alpha \leq \alpha_0} (1 + \frac{D(\alpha)}{\alpha})^{-\frac{1}{\alpha(\alpha)}}$ and

$$K_1(r) = \max_{1 \leq \alpha \leq \alpha_0} \left( \frac{r^{\alpha(\alpha)} - 1}{r_1 - 1} \right)^{\frac{1}{\alpha(\alpha)}}.$$  

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

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

(3.3)

There is explicit solution of (3.3) of the form

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

(3.4)

Again the solution must be compared with 1 and $\alpha_0$ which finally provides the direct form of $K_1(r)$. We have proved the following result:
Corollary 3.2. Suppose that $b_1 \geq b$ and $b(R) \leq L$.

1. If $1 \geq \frac{N-1}{1-b} \left(1 - \frac{\log r}{\log R}\right) \frac{\log R}{\log r}$, then

$$\sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right| \leq K_1(r) = (r - 1)^{-1} \left(\frac{D(\alpha_0)}{r^{\alpha_0(1)} - 1}\right)^{-1}.$$ 

2. If $1 \leq \frac{N-1}{1-b} \left(1 - \frac{\log r}{\log R}\right) \frac{\log R}{\log r} \leq \alpha_0$, then

$$\sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right| \leq K_1(r) = (r - 1)^{-1} \left(\frac{D(\alpha_0)}{r^{\alpha_0(1)} - 1}\right)^{-1}.$$ 

3. If $\alpha_0 \leq \frac{N-1}{1-b} \left(1 - \frac{\log r}{\log R}\right) \frac{\log R}{\log r}$, then

$$\sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right| \leq K_1(r) = (r - 1)^{-1} \left(\frac{D(\alpha_0)}{\alpha_0(\beta_0 - 1)} - 1\right)^{-1}.$$ 

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 $x_\alpha = r$, $\alpha \geq 1$ be the unique solution of

$$\alpha = \frac{N - 1}{1-b} \left(1 - \frac{\log r}{\log R}\right) \frac{\log R}{\log r}$$

if $\frac{N-1}{1-b} \geq e$ and $x_\alpha = 1$ otherwise. We deduce from Corollary 3.2 that:

Corollary 3.3. Suppose that $b_1 \geq b$ and $b(R) \leq L$. Let $\bar{r}$ be the unique solution of

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

if $\bar{r} \leq x_1$ then $\hat{r}_1 = \left(1 + D(\alpha_0)\right)^{\frac{1}{\alpha_0}}$, if $x_1 \leq \bar{r} \leq x_\alpha$ then $\hat{r}_1 = \bar{r}$ and if $\bar{r} \geq x_\alpha$ then $\hat{r}_1 = \left(1 + \frac{D(\alpha_0)}{\alpha_0}\right)^{\frac{1}{\alpha_0}}$.

Clearly $r_1 \leq r_0$, we turn to show that $r_1$ is better than the main bound in Theorem 3.2 in [1] which we denote by $r_2$. Again $r_2 = \min\{R, \hat{r}_2\}$ and $\hat{r}_2$ is the unique solution of

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

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

$$\hat{r}_1^{\alpha_0} - 1 = \frac{D(\alpha_0)}{\alpha_0},$$
for some $\alpha \in [1, \alpha_0]$. Again by the definition $R^{\kappa(\alpha)} = (N - 1)/(1 - b)\alpha$, which yields

\begin{equation}
\kappa(\alpha)\hat{r}_1 - 1 \geq \hat{r}_1^{\kappa(\alpha)} - 1 \geq \frac{(1 - b)R^{\kappa(\alpha)}D(\alpha_0)}{N - 1}.
\end{equation}

By (3.7) and the following inequality

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

we obtain

\begin{equation}
\kappa(\alpha)\hat{r}_1 - 1 \geq \frac{b}{(1 + \alpha_0)(N - 1)}.
\end{equation}

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

\[ R^{\kappa(\alpha_0)} = R^{\kappa(\alpha)} - 1, \]

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

\[ \kappa(\alpha)^2 \hat{r}_1^{\kappa(\alpha)} \hat{r}_1 - 1 \geq \frac{b}{2(N - 1)}. \]

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

\[ \frac{\hat{r}_1}{\hat{r}_1 - 1} \geq \frac{b\epsilon^2}{8(N - 1)}. \]

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 $\alpha_0$ instead of the minimization over all $\alpha_0$ usually gives a major numerical improvement.

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

\begin{equation}
|z_{\alpha_0}| = 1 + \frac{2b\pi^2}{(1 - b)^2\alpha_0^3} + o(\alpha_0^{-3}),
\end{equation}
(see discussion after Theorem 3.2 in [1]) and $\alpha_0$ is exactly our parameter. Therefore whenever the estimate $(1 + \frac{D(\alpha_0)}{\alpha_0})^{\frac{1}{\alpha_0}}$ is applied one cannot improve it up to numerical constant.

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

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

1. If $\frac{L-1}{1-b}/\log \frac{L-b}{1-b} \geq e^{1/2}$, then

\[
r_0(R) = 1 + \frac{\pi b (R - 1)^3}{2(1-b)^2} \log^{-2}(\frac{L-b}{1-b}) \log^{-1}(\frac{L-1}{\log \frac{L-b}{1-b}} + o((R-1)^3)),
\]

2. If $\frac{L-1}{1-b}/\log \frac{L-b}{1-b} \leq e^{1/2}$, then

\[
r_0(R) = 1 + \frac{b \pi (R - 1)^3}{(L-1)^2} + o((R-1)^3),
\]

**Proof.** Observe that

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

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

\[(3.11) \quad \hat{r}_0(R) = \inf_{1 \leq \alpha \leq \alpha_0(R)} (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\alpha_0}}\]

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

\[(3.12) \quad \log \frac{N(R) - 1}{1-b} = \log \alpha + \log(1 + \frac{D(\alpha)}{\alpha}) \frac{D(\alpha) + \alpha}{D(\alpha) - \alpha D'(\alpha)},\]

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 when $R \to 1$, hence using

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

we obtain that

\[
\log \alpha(R) = -\frac{1}{2} + \log \frac{N(R) - 1}{1-b} + o(1).
\]

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

\[
\lim_{R \to \infty} \frac{N(R) - 1}{\alpha_0(R) R - 1} = \frac{L - 1}{1-b} \log^{-1}(\frac{L-b}{1-b}) < e^{1/2}
\]
we have to use \( \alpha(R) \) (at least for small enough \( R \)) when minimize \( (1 + \frac{D(\alpha)}{\alpha}) \frac{1}{\pi b} \) over \([1, \alpha_0(R)]\), otherwise \( \alpha_0(R) \) is the point minimum. In the first setting we have
\[
\alpha(R) = e^{-\frac{1}{2}} \frac{L - 1}{(1 - b)(R - 1)} + o(1), \quad \text{and} \quad \kappa(\alpha(R)) = \frac{1}{2(R - 1)} + o(1),
\]
thus using (3.10) and (3.11) we obtain that
\[
\hat{r}_0(R) = (1 + \frac{D(\alpha(R))}{\alpha(R)}) \frac{1}{\pi(\alpha(R))} = 1 + \frac{D(\alpha(R))}{\alpha(\alpha) \kappa(\alpha(R))} + o((R - 1)^3) = 1 + \frac{\pi b}{2(1 - b)^2 \alpha^2(\alpha(R))} + o((R - 1)^3). = 1 + \frac{\pi b(R - 1)^3}{2(1 - b)^2 \log^2 \frac{L - b}{1 - b} \log((L - 1)/ \log((L - b)/(1 - b))) + o((R - 1)^3).}
\]
In the same way if \( \frac{L - b}{1 - b} \log^{-1} \frac{L - b}{1 - b} \geq e^1/2 \), then
\[
\alpha_0(R) = \frac{\log \frac{L - b}{1 - b}}{R - 1} + o(1), \quad \kappa(\alpha_0(R)) = \frac{\log \frac{L - 1}{L - b}}{1 - b} + o(1),
\]
and hence
\[
\hat{r}_0(R) = (1 + \frac{D(\alpha_0(R))}{\alpha_0(R)}) \frac{1}{\pi(\alpha_0(R))} = 1 + \frac{D(\alpha_0(R))}{\alpha_0(\alpha_0) \kappa(\alpha_0(R))} + o((R - 1)^3) = 1 + \frac{\pi b}{2(1 - b)^2 \alpha^2_0(\alpha(R))} + o((R - 1)^3) = 1 + \frac{\pi b(R - 1)^3}{2(1 - b)^2 \log^2 \frac{L - b}{1 - b} \log((L - 1)/ \log((L - b)/(1 - b))) + o((R - 1)^3).}
\]
It is clear that \( \hat{r}_0(R) \leq R \) for \( R \) small enough, thus the asymptotic for \( \hat{r}_0(R) \) is the same as for \( r_0(R) \). It completes the proof of the corollary.

In particular Corollary 3.4 shows that whenever \( \frac{L - b}{1 - b} \log^{-1} \frac{L - b}{1 - b} \geq e^1/2 \) the following inequality holds
\[
r_0(R) = 1 + \frac{\pi b}{2(1 - b)^2 \alpha^2(R)} \kappa(\alpha(R)) + o(\alpha(R)^{-3}),
\]
which when compared with (3.9) proves that our result cannot be improved up to a numerical constant (recall that \( (1 + \alpha_0)/2 \leq \kappa(\alpha_0) \leq \alpha_0 \)). On the other hand Corollary 3.4 makes it possible to compare our result with Theorem 3.2 in [1]. The following estimate holds for \( r_2(R) \) in the same setting (see Section 3 in [1])
\[
r_2(R) = 1 + \frac{e^2 b(R - 1)^3}{8(L - 1)} + o((R - 1)^3).
\]
Therefore if \( L - 1 \) much larger than \( 1 - b \) our answer is better upon \( (L - 1)/(1 - b)^2 \) and if \( L - 1 \) is close to \( 1 - b \) then upon \( L - 1 \).

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 \( (1 + \frac{D(\alpha)}{\alpha})^{\frac{1}{\pi^2}} \) is attained at \( \alpha_0(R) \). At the opposite if \( L \) is close to 1 (again with respect to \( b \) and \( L \)) then we have to use the minimization inside \([1, \alpha_0(R)]\) even for \( R \to 1 \). It explains that one cannot avoid the 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 obtaining rates of convergence for ergodic Markov chains under the assumption of a true atom existence.

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

\[
P(x, A) = \nu(A), \quad \text{for all } x \in C.
\]

which implies that \( C \) is an atom and \( \nu = P(a, \cdot) \), for any \( a \in C \). It remains to translate conditions 2-3 (from introduction) into a simpler form which can be used later to prove the geometric ergodicity. Let \( \tau = \tau(C) = \inf\{n \geq 1 : X_n \in C\} \) and then define \( \tau_k, k \geq 1 \) as the subsequent visits to \( C \). For simplicity let also \( \tau_0 = \sigma(C) = \inf\{n \geq 0 : X_n \in C\} \), 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 \( b_n = P_a(\tau = n) \).

Moreover denoting \( u_n = P_a(X_n \in C) \), for \( n \geq 0 \) we construct the renewal sequence for \((\tau_k)_{k \geq 0}\).

As we have mentioned the behavior of \((u_n)_{n \geq 0}\) is closely related to the ergodicity of the Markov chain. In particular assuming ergodicity \( \lim_{n \to \infty} u_n \) exists and is equal \( u_\infty = \pi(C) \). Following [1] we define function \( G(r, x) = E_x r^\tau \), for all \( x \in S \) and \( 0 < r \leq \lambda^{-1} \). The main property of \( G(r, x) \) is that it is the lower bound for \( V(x) \) on the set \( S \setminus C \), namely we have that (for the proof see Proposition 4.1 in [1])

**Proposition 4.1.** Assume only drift condition (2).

1. For all \( x \in S, P_x(\tau < \infty) = 1. \)
2. For $1 \leq r \leq \lambda^{-1}$

$$G(r, x) \leq \begin{cases} V(x) & \text{if } x \not\in C, \\ rK & \text{if } x \in C. \end{cases}$$

The renewal approach is based on the first entrance last exit property. To state the result we need additional notation $H_W(r, x) = \mathbb{E}_x \left( \sum_{n=1}^{\tau} r^n W(X_n) \right)$, for $r > 0$ for which the definition makes sense. The following result holds (for the proof see Proposition 4.2 in [1]):

**Proposition 4.2.** Assume only that the Markov chain is geometrically ergodic with (unique) invariant probability measure $\pi$, that $C$ is an atom and that $W : S \to \mathbb{R}$ is such that $W \geq 1$. Suppose $g : S \to \mathbb{R}$ satisfies $\|g\|_W \leq 1$, then for all $r \geq 1$ for which right-hand sides are finite:

$$\sup_{|z|=r} |\sum_{n=1}^{\infty} (P_n g(a) - \int g \, d\pi) z^n| \leq H_W(r, a) \sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right| + \pi(C) \frac{H_W(r, a) - r H_W(1, a)}{r - 1},$$

for all $a \in C$ and

$$\sup_{|z|=r} |\sum_{n=1}^{\infty} (P_n g(x) - \int g \, d\pi) z^n| \leq H_W(r, x) + G(r, x) H_W(r, a) \sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right| +$$

$$+ \pi(C) \frac{H_W(r, a) - r H_W(1, a)}{r - 1} G(r, x) + \pi(C) H_W(1, a) \frac{r (G(r, x) - 1)}{r - 1},$$

for all $x \not\in C$.

Now the problem of proving the geometric convergence splits into two parts: in the first one we have to provide some estimate on $H_W(r, x)$, $x \in S$ on the interval $1 \leq r \leq \lambda^{-1}$, and it is of meaning 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 $\sup_{|z|=r} \left| \sum_{n=0}^{\infty} (u_n - u_\infty) z^n \right|$, for $r < r_0$. The second question is exactly the Kendall’s theorem in the setting when $R = \lambda^{-1}$, $L = \lambda^{-1} K$ (note that $b(r) = G(r, a)$ and thus $b(\lambda^{-1}) \leq \lambda^{-1} K$ by Proposition 4.1) and $b$ is the bound assumed for the strong aperiodicity (i.e. $b_1 \geq b$). The discussed additional parameter is $u_\infty = \pi(C)$, the better is our knowledge on $\pi(C)$ the better is the bound that stems from Theorem
2.2. If one knows the exact value of $\pi(C)$ one can use Corollary 2.3, in general - in the lack of information on $\pi(C)$, one can apply Corollary 2.5.

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

**Proposition 4.3.** Assume only drift condition (2).

1. For $1 \leq r \leq \lambda^{-1}$

\[
H_1(r, x) \leq \begin{cases} \frac{r\lambda V(x)}{r - 1} & \text{if } x \not\in C, \\ \frac{r}{r - 1} & \text{if } x \in C. \end{cases}
\]

2. and 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 the first inequality it suffices to observe that $r^{-1}H_1(r, x)$ attains its maximum on the interval $[1, \lambda^{-1}]$ at $\lambda^{-1}$. Using Proposition 4.1 we obtain that

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

Consequently $H_1(r, x) \leq \frac{r\lambda(V(x) - 1)}{r - 1}$, $x \not\in C$ and in the same way we show that $H_1(r, x) \leq \frac{r(K - 1)}{r - 1}$ if $x \in C$. The second inequality can be derived in the 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

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

Combining estimates from Propositions 4.1 4.3 with Proposition 4.2 and Corollaries 2.3.2.5 we obtain our first result on the 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 verifies (1.2) and we have the following bounds on $\rho_V$, $M_1$:

\[
\rho_V \leq r_0^{-1},
\]

\[
M_1(r) \leq \frac{2r\lambda}{1 - \lambda} + \frac{r\lambda(K - 1)}{(1 - \lambda)^2} + \frac{r(K - \lambda)}{1 - \lambda}K_0(r),
\]

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 2.3.2.5.
On the other hand when $W \equiv V$ there are weaker bounds on $H_V(r)$, which are stated in Proposition 4.2 in [1]:

**Proposition 4.5.** Assume only drift condition (2).

1. For $1 \leq r \leq \lambda^{-1}$

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

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

2. and for $1 \leq r \leq \lambda^{-1}$

$$\frac{H_V(r, a) - rH_V(1, a)}{r-1} \leq \frac{r\lambda(K-1)}{(1-\lambda)(1-r\lambda)}.$$ 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 $(X_n)_{n \geq 0}$ satisfies conditions 1-3 with $\bar{b} = 1$. Then $(X_n)_{n \geq 0}$ is geometrically ergodic - it verifies (1.2) and we have the following bounds on $\rho_V, M_V$:

$$\rho_V \leq r_0^{-1}$$

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

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 Corollaries 2.3,2.5.

## 5 Non atomic case

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

$$P(Y_n = 1| \mathcal{F}_n^X, \mathcal{F}_{n-1}^Y) = \bar{b}_1c(X_n),$$

$$P(X_{n+1} \in A| \mathcal{F}_n^X, \mathcal{F}_n^Y) = \begin{cases} \nu(A), & \text{if } Y_n = 1, \\ \frac{\nu(A) - \bar{b}_1c(X_n)\nu(A)}{1-\bar{b}_1c(X_n)}, & \text{if } Y_n = 0. \end{cases}$$
where $\mathcal{F}_n^X = \sigma(X_k : 0 \leq k \leq n)$, $\mathcal{F}_n^Y = \sigma(Y_k : 0 \leq k \leq n)$. Thus the chain evolves in a way that whenever $X_n$ is in $C$ we pick $Y_n = 1$ with probability $\bar{b}$. Then if $Y_n = 1$ we chose $X_{n+1}$ from $\nu$ distribution whereas if $Y_n = 0$ then we just apply normalized probability measure version of $P(X_n, \cdot) - \bar{b}1_C \nu$. The split chain is designed so that it has an atom $S \times \{1\}$ 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 $P_{x,i}$, $E_{x,i}$ 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 $P_{x,1} = P_{a,1}$ and $E_{x,1} = E_{a,1}$ for all $x \in C$. Following the method used in the atomic case we define the renewal sequence $\bar{u}_n = P_{a,1}(Y_n = 1)$ and the corresponding increment sequence $\bar{b}_n = P_{a,1}(T = n)$ for $n \geq 1$. Clearly $\bar{u}_n = P_{a,1}(X_n \in C, Y_n = 1) = \bar{b}P_{\nu}(X_{n-1} \in C)$ for $n \geq 1$, so

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

We define

$$\tilde{G}(r, x, i) := E_{x,i}(r^T), \quad \tilde{H}_W(r, x, i) := E_{x,i}(\sum_{n=1}^{T} r^n W(X_n)),$$

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

$$E_x := (1 - \bar{b}_1 C(x))E_{x,0} + \bar{b}_1 C(x)E_{x,1},$$

which agrees with the usual $E_x$ on $\mathcal{F}^X$. There exists a unique stationary measure $\bar{\pi}$ say, on $(S, \mathcal{B})$, so that $\int g \bar{d}\bar{\pi} = \int g d\pi$ (where $g(x) = \bar{g}(x, 0) = \bar{g}(x, 1)$ for all $x \in S$). In particular we have that $\bar{\pi}(S \times \{1\}) = \bar{b}\pi(C)$.

The first entrance last exist decomposition leads to the following result (cf. Proposition 4.2 in [1]):

**Proposition 5.1.** For all $a \in C \times \{1\}$

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

$$(5.2)$$
and for all \( x \in S \times \{0\} \)
\[
\sup_{|z|=r} \left| \sum_{n=1}^{\infty} (P^n \bar{g}(x) - \int g d\pi) z^n \right| \leq \\
\leq \bar{H}_W(r, x, 0) + \bar{G}(r, x, 0) \bar{H}_W(1, a, 1) \sup_{|z|=r} \left| \sum_{n=0}^{\infty} (\bar{u}_n - \bar{u}_\infty) z^n \right| + \\
+ \bar{b}\pi(C) \frac{\bar{H}_W(r, a, 1) - r \bar{H}_W(1, a, 1)}{r-1} \bar{G}(r, x, 0) + \\
+ \bar{b}\pi(C) \frac{r(\bar{G}(r, x, 0) - 1)}{r-1}.
\]

(5.3)

**Proof.** The proof of the result mimics the proof of Proposition 4.3 in [1], the only difference is that one has to control from which part of the space point \( x \) comes from: \( S \times \{0\} \) or the atom \( C \times \{1\} \).

As in the atomic case now the question splits into two parts. The first part 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 question 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 \( \sup_{|z|=r} \left| \sum_{n=0}^{\infty} (\bar{u}_n - \bar{u}_\infty) z^n \right| \). Here the problem is that we have some information on the basic sequence yet we have to derive control on the sequence \( (\bar{b}_n)_{n \geq 0} \).

We sketch shortly what can be done about the ergodicity of \( (\bar{u}_n)_{n \geq 0} \). Recall that \( (\bar{u}_n)_{n \geq 0} \) is the corresponding renewal sequence for \( (\bar{b}_n)_{n \geq 1} \). In the same way as in the atomic case let \( \bar{b}(z), \bar{u}(z), z \in C \) be 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 a control on the limiting behavior of \( \bar{c}(z) - \bar{c}(1) \), namely applying Theorem 2.2 we obtain that whenever \( \bar{c}(r) < \infty \), then

(5.4) \[
\sup_{|z|=r} \left| \sum_{n=0}^{\infty} (\bar{u}_n - \bar{u}_\infty) z^n \right| \leq \frac{\bar{c}(r) - \bar{c}(1)}{\bar{c}(1)(r-1)((1-b)D(\bar{\alpha}) - \bar{c}(r) + \bar{c}(1))}.
\]

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

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

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

$$b_x(R) = \mathbf{E}_x R' \leq L = KR, \text{ for all } x \in C,$$

whereas one needs a bound on the generic function of $(b_n)_{n \geq 1}$. We discuss the question in Appendix, showing in Proposition A.2 that for all $1 \leq r \leq \min\{R, (1 - b)^{-\frac{1}{\alpha_0}}\}$ the following inequality holds

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

where $\alpha_1 = \log(\frac{L - \bar{b}R}{(1 - b)R})/\log R$ and $\alpha_2 = \log(\frac{L - (1 - \bar{b} + b)R}{(b - \bar{b})R})/\log R$. Moreover if $1 + b \geq 2\bar{b}$ then simply

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

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

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

$$\bar{a} \leq \tilde{b}^{-1} \max\{\frac{1 - \bar{b}}{1 - \tilde{b}}(1 + \alpha_1), \frac{1 - \bar{b}}{1 - \tilde{b}} + \frac{\bar{b} - b}{1 - b} \alpha_2\}.$$

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

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

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

which means that $F_{0}(x) = \log(\frac{\tilde{c}(\epsilon^x) - 1}{\tilde{c}(1) - 1})$ is convex and $F_{0}(0) = 0$. By (5.6) we have that $\tilde{c}(\epsilon^x) \leq L(\epsilon^x)$ and hence

$$F_{0}(x) \leq F_{1}(x) = \log(\frac{L(\epsilon^x) - \epsilon^x}{(1 - b)\bar{a}(\epsilon^x - 1)}).$$

Therefore we can easily compute the largest possible function $\tilde{F}(x)$ 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 over the functions with the properties 1-2, namely if there exists $F$ that satisfies the above condition 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 the function $\bar{F}$ 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’s theorem. Under the basic data contained in conditions 1-3 in the introduction one can propose the upper bound $F_1$ on $F_0$. On the other we may benefit from the fact that $F_0$ is convex and starts from zero. Consequently we consider all the possible functions that posses the properties. It occurs that there is a maximizer $\bar{F}$ among the 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

$$(5.9) \quad F_1'(x)x = F_1(x).$$

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

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

otherwise if $x_0 > \log R$ then

$$(5.11) \quad \bar{F}(x) = \frac{F_1(\log R)}{\log R}x \text{ for all } 0 \leq x \leq \log R.$$
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
\[
\bar{r}_0 = \min\{R, (1 - \bar{b})^{-1}, \bar{r}_0(\bar{\alpha})\},
\]
where $\bar{r}_0(\bar{\alpha})$ is the unique solution of the equation
\[
r = (1 + \frac{D(\bar{\alpha})}{\bar{\alpha}})^{\frac{1}{F(\bar{\alpha},r)}}.
\]
Moreover for $r < \bar{r}_0$
\[
\sup_{|z|=r} \left|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\right| \leq \bar{K}_0(r) = \frac{\bar{u}_\infty(r^{\bar{\alpha}(\bar{\alpha},r)} - 1)}{\bar{u}_\infty(r^{\bar{\alpha}(\bar{\alpha},r)} - 1) + 1}.
\]

Remark 5.3. Observe that if
\[
\log(1 + \frac{D(\bar{\alpha})}{\bar{\alpha}})/F_1(x_0) \leq x_0 \leq \log R,
\]
then $\bar{r}_0 = (1 + \frac{D(\bar{\alpha})}{\bar{\alpha}})^{\frac{1}{F_1(x_0)}}$. Due to (5.9), the condition (5.13) is equivalent to $x_0 \leq \log R$ and
\[
1 + \frac{D(\bar{\alpha})}{\bar{\alpha}} \leq \frac{L(e^{x_0}) - e^{x_0}}{(1 - b)\bar{\alpha}(e^{x_0} - 1)}.
\]

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 $\bar{\alpha} \geq \bar{b}^{-1}$. Since in the same way as in the atomic case $(1 + \frac{D(\bar{\alpha})}{\bar{\alpha}})^{\frac{1}{F_1(x_0)}}$ increases with $b$ assuming that $\bar{b}, L, R$ are fixed, thus we can always assume $\bar{\alpha} \geq \bar{b}^{-1}$. Let $\bar{\alpha}_0 = \max\{\frac{1-b}{1-b}, \frac{1-b}{1-b}, \frac{b-b}{1-b}\}$.  

Theorem 5.4. Let $\bar{b}_1 \geq b$, and $\bar{b}(r)$ satisfies (5.6). Then $\sum_{n=0}^{\infty}(\bar{u}_n - \bar{u}_\infty)z^n$ is convergent for $|z| < \bar{r}_0$, where
\[
\bar{r}_0 = \min\{R, (1 - \bar{b})^{-1}, \min_{\bar{\alpha}, \bar{\alpha} \in \bar{\alpha}_0} \bar{r}_0(\bar{\alpha})\},
\]
where $\bar{r}_0(\bar{\alpha})$ is the unique solution of the equation
\[
r = (1 + \frac{D(\bar{\alpha})}{\bar{\alpha}})^{\frac{1}{F(\bar{\alpha},r)}}.
\]
Moreover for $r < \bar{r}_0$
\[
\sup_{|z|=r} \left|\sum_{n=0}^{\infty}(u_n - u_\infty)z^n\right| \leq \bar{K}_0(r) = \max_{\bar{\alpha}, \bar{\alpha} \in \bar{\alpha}_0} \frac{\bar{b}(r^{\bar{\alpha}(\bar{\alpha},r)} - 1)}{(r - 1)(\bar{\alpha}^{-1}D(\bar{\alpha}) - r^{\bar{\alpha}(\bar{\alpha},r)} + 1)}.
\]
We show in examples that the approach presented in Theorems 5.2, 5.4 is comparable with the coupling method (see Section 7 in [1] for short introduction). Therefore we obtain the computable tool for the general question of rates of convergence of ergodic Markov chains under the geometric drift condition.

We move the detail computation of all the required bounds in Proposition 5.1 to Appendix A. This knowledge enable 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 \((X_n)_{n \geq 0}\) satisfies conditions 1-3 from the introduction. Then \((X_n)_{n \geq 0}\) is geometrically ergodic - it verifies (1.2) and the following bounds on \( \rho_V, M_1 \) hold:

\[
\rho_V \leq \bar{r}_0^{-1}
\]

and

\[
M_1(r) \leq \frac{2\lambda r}{1 - \lambda} + \frac{2(1 - \tilde{b})(r^{1+\alpha_1} - 1)r}{r - 1}(1 - (1 - \tilde{b})r^{1+\alpha_1}) + \frac{\tilde{b}r \lambda(K - 1)}{1 - (1 - b)r^{1+\alpha_1}(1 - \lambda)^2} + \frac{(r - 1)\bar{K}_0(r) + b(1 - \tilde{b})(r^{1+\alpha_1} - 1)r(K - \lambda)}{(r - 1)(1 - (1 - b)r^{1+\alpha_1})^2} \frac{1 - \lambda}{1 - \lambda},
\]

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

**Proof.** Note that \( \tilde{b}\pi(C)\tilde{H}_1(1, a, 1) = 1 \). We apply Proposition 5.1 in the way that we sum (5.2) with weight \((1 - \tilde{b}1_{C}(x))\) and (5.3) with weight \( \tilde{b}1_{C}(x) \). Then we use (A.5) to bound \( \tilde{b}1_{C}(x) + (1 - \tilde{b}1_{C}(x))G(r, x, 0) \) and (A.20) to bound \((1 - \tilde{b}1_{C}(x))\tilde{H}_1(r, x, 0) = (1 - \tilde{b}1_{C}(x))\frac{G(r, x, 0) - 1}{r - 1} \). Finally (A.21) and (A.22) are estimates for \( \tilde{H}(r, a, 1) \) and \( \tilde{H}_1(r, a, 1) - r\tilde{H}_1(1, a, 1)/r - 1 \).

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

**Theorem 5.6.** Suppose \((X_n)_{n \geq 0}\) satisfies conditions 1-3 from the introduction. Then \((X_n)_{n \geq 0}\) is geometrically ergodic - it verifies (1.2) and the following bounds on \( \rho_V, M_V \) hold:

\[
\rho_V \leq \bar{r}_0^{-1}
\]
and

\[
M_V(r) \leq \frac{\lambda r}{1 - r\lambda} + \frac{K - r\lambda - \bar{b}}{1 - r\lambda} \frac{r}{1 - (1 - \bar{b})r^{1+\alpha_1}} + \\
\frac{K - \lambda}{1 - \lambda} \frac{r\lambda}{1 - \lambda} + \frac{(1 - \bar{b})(r^{1+\alpha_1} - 1)r}{(r - 1)(1 - (1 - \bar{b})r^{1+\alpha_1})} + \\
\frac{\bar{b}}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{r(K - 1)}{(1 - \lambda)(1 - r\lambda)} + \\
\frac{K - r\lambda - \bar{b}}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{1}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{r(K - \lambda)}{1 - \lambda} + \\
\frac{\bar{K}_0(r)}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{r(K - \lambda)}{1 - r\lambda} + \frac{K - r\lambda - \bar{b}}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{r - 1}{1 - (1 - \bar{b})r^{1+\alpha_1}} \frac{r(K - \lambda)}{1 - \lambda},
\]

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

Proof. Observe that \(\pi(C) \leq 1\). As in the proof of Theorem 5.5 we use Proposition 5.1 summing (5.2) with weight \((1 - \bar{b}C(x))\) and (5.3) with weight \(\bar{b}C(x)\). Again we use (A.5) to bound \(\bar{b}C(x) + (1 - \bar{b}C(x))\bar{G}(r, x, 0)\), then (A.26), (A.27), (A.28) to bound respectively \((1 - \bar{b}C(x))\bar{H}_V(r, x, 0), \bar{H}_V(r, a, 1)\) and \((\bar{H}_V(r, a, 1) - r\bar{H}_V(1, a, 1))/(r - 1)\). We use also the bound \(\bar{H}_V(1, a, 1) \leq \bar{b}^{-1}K^{-\lambda}/1 - \lambda\) and (A.20) to bound \((1 - \bar{b}C(x))\bar{H}_1(r, x, 0) = (1 - \bar{b}C(x))\bar{H}_1(r, x, 0)\).

\(\blacksquare\)

Appendix A

A.1 Global bounds

Our method described in Corollary 5.2 implies that

\[
\sup_{|z| = r} \left| \sum_{n=0}^{\infty} (\bar{u}_n - \bar{u}_\infty)z^n \right| \leq K_0(r) \text{ for } 1 \leq r \leq r_0.
\]

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

\[
G(r, x, i) = E_{x,i}r^\tau, \quad H_W(r, x, i) = E_{x,i}(\sum_{n=1}^{\tau} r^nW(X_n)).
\]

Let also \(G(r) = \sup_{x \in C} E_{x,0}r^\tau, \quad H_W(r) = \sup_{x \in C} E_{x,0} \sum_{n=1}^{\tau} r^nW(X_n)\). In the Lemma A.1 in [1] there are proved following inequalities:
Proposition A.1. For $r \leq \lambda^{-1}$ and $(1 - \tilde{b}) G(r) < 1$ the inequalities hold:

(A.1) \[ \bar{G}(r, x, i) \leq \frac{\tilde{b} G(r, x, i)}{1 - (1 - \tilde{b}) G(r)} \]

and

(A.2) \[ H_W(r, x, i) \leq H_W(r, x, i) + \frac{(1 - \tilde{b}) H_W(r) G(r, x, i)}{1 - (1 - \tilde{b}) G(r)} \]

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

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

(A.3) \[ \bar{G}(r, a, 1) \leq \max\left\{ \frac{\tilde{b} r}{1 - (1 - \tilde{b}) r^{1 + \alpha_1}}, \frac{b r + (\tilde{b} - b) \alpha_2}{1 - (1 - \tilde{b}) r^{-\alpha_2}} \right\}, \]

where $\alpha_1 = \log(\frac{K - \tilde{b}}{1 - \tilde{b}})/\log \lambda^{-1}$, $\alpha_2 = \log(\frac{(K - 1 + \tilde{b} - b)}{b - \tilde{b}})/\log \lambda^{-1}$. Moreover if $1 + b \geq 2\tilde{b}$, then

(A.4) \[ \bar{G}(r, a, 1) \leq \frac{\tilde{b} r}{1 - (1 - \tilde{b}) r^{1 + \alpha_1}}. \]

For all $x \in S$, $1 \leq r \leq \min\{\lambda^{-1}, (1 - \tilde{b})^{-1 + \alpha_1}\}$

(A.5) \[ \tilde{b} 1_C(x) + (1 - \tilde{b} 1_C(x)) \bar{G}(r, x, 0) \leq \frac{\tilde{b} V(x)}{1 - (1 - \tilde{b}) r^{1 + \alpha_1}}. \]

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

(A.6) \[ (1 - \tilde{b}) \sup_{x \in C} G(r, x, 0) + \tilde{b} G(r, a, 1) = \sup_{x \in C} G(r, x) = G(r). \]

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

\[ (1 - \tilde{b}) \sup_{x \in C} G(r, x, 0) \leq (1 - \tilde{b}) r^v \log r \frac{\log r}{\log \lambda^{-1}}, \] and, $\bar{G}(r, a, 1) \leq \frac{b r + (\tilde{b} - b) \alpha_2}{1 - (1 - \tilde{b}) r^{-\alpha_2}} u^{\log r \frac{\log r}{\log \lambda^{-1}}}$,

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

(A.7) \[ \lambda^{-1}(b + (\tilde{b} - b) u + (1 - \tilde{b}) v) = \sup_{x \in C} G(\lambda^{-1}, x) \leq K \lambda^{-1}, \] $u, v \geq 1$.

Observe that by (A.1) we have the following bound

(A.8) \[ \bar{G}(r, a, 1) \leq \frac{\tilde{b} G(r, a, 1)}{1 - (1 - \tilde{b}) G(r)} \leq F(u, v) = \frac{b r + (\tilde{b} - b) \alpha_2}{1 - (1 - \tilde{b}) r^{\alpha_2}} u^{\log r \frac{\log r}{\log \lambda^{-1}}}. \]
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 that

\begin{equation}
\bar{G}(r, a, 1) \leq \max\left\{ \frac{b r}{1 - (1 - b) r^{1 + \alpha_1}}, \frac{b r + (\bar{b} - b) r^{1 + \alpha_2}}{1 - (1 - b) r} \right\}.
\end{equation}

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) must be attained at the second expression.

We turn to show 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

\begin{equation}
\bar{G}(r, x, 0) \leq \frac{\bar{b} V(x)}{1 - (1 - b) G(r)},
\end{equation}

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

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

On the other hand by (A.1)

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

for all $x \in C$ and therefore

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

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

\[ \blacksquare \]

**Corollary A.3.** The following inequality holds

\begin{equation}
\bar{b}^{-1} \leq \frac{H_1(1, a, 1) - 1}{1 - b} \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\} = \bar{b}^{-1} \alpha_0.
\end{equation}

**Proof.** As for the first assertion we simply apply (A.3) to bound $H_1(r, a, 1) = r \bar{G}(r, a, 1) - 1$ and then tend with $r$ to 1. To prove the second assertion let
\[ S = \max\{k \geq 1 : \tau_k \leq T\} \], where \( \tau_k, k \geq 0 \) are subsequent visits to \( C \) by \((X_n)_{n \geq 0}\), in particular \( \tau_0 = 0 \). Observe that

\[ \bar{H}_1(1, a, 1) = E_{a,1}(\sum_{k=0}^{\infty} 1_{S \geq k}(\tau_k - \tau_{k-1})). \]

Therefore by the construction

\[ \bar{H}_1(1, a, 1) \geq E_{\nu}(1 + \sigma) + E_{a,1}(S - 1), \]

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

\[ \bar{H}_1(1, a, 1) \geq \bar{b}^{-1} + E_{\nu}\sigma. \]

It remains to notice that \( E_{\nu}\sigma \geq 1 - \nu(C) \), therefore if \( \bar{b}\nu(C) = b \) then

\[ \bar{H}_1(1, a, 1) \geq \bar{b}^{-1} + 1 - \frac{b}{\bar{b}}, \]

which completes the proof.

Now we state an improvement of the result mentioned in the proof of Proposition 4.4 in [1].

**Proposition A.4.** For \( r \leq \lambda^{-1} \) and \( (1 - \bar{b})G(r) < 1 \) we have that

\[ (A.13) \quad \bar{H}_W(r, a, 1) \leq \frac{1}{b} \sup_{x \in C} H_W(r, x) + \frac{1 - \bar{b}}{b} \frac{H_W(r) \sup_{x \in C}(G(r, x) - 1)}{1 - (1 - b)G(r)} \]

and

\[ \bar{H}_W(r, a, 1) - r\bar{H}_W(1, a, 1) \leq \frac{1}{b} \sup_{x \in C}(H_W(r, x, 0) - rH_W(1, x, 0)) + \]

\[ (A.14) + \frac{1 - \bar{b}}{b} H_W(r)(G(r, a, 1) - 1). \]

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

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

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

\[ \bar{H}_W(r, a, 1) \leq H_W(r, a, 1) + \frac{1 - \bar{b}}{b} H_W(r) + \frac{(1 - \bar{b})H_W(r) \sup_{x \in C}(G(r, x) - 1)}{b(1 - (1 - b)G(r))}. \]
Since the definition of $H_W(r, x, 1)$ implies that
\[
\bar{b}H_W(r, a, 1) + (1 - \bar{b})H_W(r) \leq \sup_{x \in C} H_W(r, x)
\]
we obtain (A.13). To show the second assertion we use $S = \max\{k \geq 1 : \tau_k \leq T\}$ defined in the proof of Corollary A.3. The following inequality holds
\[
H_W(r, a, 1) - rH_W(1, a, 1) \leq H_W(r, a, 1) - rH_W(1, a, 1) + \sum_{k=2}^{\infty} \mathbb{E}_{a,1}[1_{k \leq S} \sup_{x \in C} (r^{\tau_{k-1}} H_W(r, x, 0) - rH_W(1, x, 0))].
\] (A.15)

As we have shown in Corollary A.3, $\mathbb{E}_{a,1}(S - 1) = (1 - \bar{b})/\bar{b}$ we deduce that
\[
\sum_{k=2}^{\infty} (\mathbb{E}_{a,1}1_{k \leq N}) \sup_{x \in C} (H_W(r, x, 0) - rH_W(1, x, 0)) = \frac{1 - \bar{b}}{\bar{b}} \sup_{x \in C} (H_W(r, x, 0) - rH_W(1, x, 0)),
\]
which together with (A.15) provides
\[
H_W(r, a, 1) - rH_W(1, a, 1) \leq H_W(r, a, 1) - rH_W(1, a, 1) + \frac{1 - \bar{b}}{\bar{b}} \sup_{x \in C} (H_W(r, x, 0) - rH_W(1, x, 0)) + \sum_{k=2}^{\infty} [\mathbb{E}_{a,1}1_{k \leq N}(r^{\tau_{k-1}} - 1)] \sup_{x \in C} H_W(r, x, 0).
\] (A.16)

As usual we observe that
\[
H_W(r, a, 1) - rH_W(1, a, 1) + \frac{1 - \bar{b}}{\bar{b}} \sup_{x \in C} (H_W(r, x, 0) - rH_W(1, x, 0)) \leq \frac{1}{\bar{b}} \sup_{x \in C} (H_W(r, x) - rH_W(1, x)).
\] (A.18)

Moreover since $Y_{\tau_k}$ is independent of $\tau_{k-1}$ we have $\mathbb{E}_{a,1}r^{\tau_{k-1}}1_{k \leq S} = (1 - \bar{b})\mathbb{E}_{a,1}r^{\tau_{k-1}}1_{k-1 \leq S}$ which implies that
\[
\sum_{k=2}^{\infty} \mathbb{E}_{a,1}r^{\tau_{k-1}}1_{S=k-1} = \bar{b} \sum_{k=2}^{\infty} \mathbb{E}_{a,1}r^{\tau_{k-1}}1_{k-1 \leq S} = \frac{b}{1 - \bar{b}} \sum_{k=2}^{\infty} \mathbb{E}_{a,1}r^{\tau_{k-1}}1_{k \leq S},
\]
and thus
\[
\sum_{k=2}^{\infty} [\mathbb{E}_{a,1}1_{k \leq S}(r^{\tau_{k-1}} - 1)] = \frac{1 - \bar{b}}{b} (\mathbb{E}_{a,1}(r^T - 1)) = \frac{1 - \bar{b}}{b} (\bar{G}(r, a, 1) - 1).
\]
Consequently
\[ \sum_{k=2}^{\infty} [E_{a,1} 1_{k \leq N}(r^{\tau_k} - 1)] \sup_{x \in C} H_W(r, x, 0) = \frac{1 - \bar{b}}{b} H_W(r)(\bar{G}(r, a, 1) - 1). \]

Combining (A.17), (A.18) and (A.19) we conclude that
\[
\bar{H}_W(r, a, 1) - r \bar{H}_W(1, a, 1) \leq \frac{1}{b} \sup_{x \in C} (H_W(r, x, 0) - r H_W(1, x, 0)) + \frac{1 - \bar{b}}{b} H_W(r)(\bar{G}(r, a, 1) - 1).
\]

It completes the proof of (A.14).

\[ \blacksquare \]

### A.2 Case of $W \equiv 1$

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

**Proposition A.5.** The following inequalities hold
\[ (A.20) \quad (1 - \bar{b}1_{C}(x)) \bar{H}_1(r, x, 0) \leq \frac{r \lambda(V(x) - 1)}{1 - \lambda} + \frac{(1 - \bar{b})(1 + \alpha_1) - 1)}{(r - 1)(1 - (1 - \bar{b})r^{1 + \alpha_1})}, \]

for all $x \in S$, $1 \leq r \leq \min\{\lambda^{-1}, (1 - \bar{b})^{-\frac{1}{1 + \alpha_1}}\}$,

\[ (A.21) \quad \bar{H}_1(r, a, 1) \leq \frac{1}{1 - (1 - \bar{b})r^{1 + \alpha_1}} \frac{r(K - \lambda)}{1 - \lambda}, \]

for all $a \in C$, $1 \leq r \leq \min\{\lambda^{-1}, (1 - \bar{b})^{-\frac{1}{1 + \alpha_1}}\}$ and
\[ (A.22) \quad \frac{\bar{H}_1(r, a, 1) - r \bar{H}_1(1, a, 1)}{r - 1} \leq \frac{r \lambda(K - 1)}{b (1 - \lambda)^2} + \frac{1}{b (r - 1)(1 - (1 - \bar{b})r^{1 + \alpha_1})} \frac{r(K - \lambda)}{1 - \lambda}, \]

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

**Proof.** By (A.2) we have that
\[ \bar{H}_1(r, x, 0) \leq H_1(r, x, 0) + \frac{(1 - \bar{b})H_1(r)G(r, x, 0)}{1 - (1 - \bar{b})G(r)}. \]
Together with $H_1(r, x, 0) = H_1(r, x)$ and $G(r, x, 0) = G(r, x) \leq V(x)$ for $x \notin C$ it follows that

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

Consequently by Proposition 4.3

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

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

(A.23) \hspace{1cm} \tilde{H}_1(r, x, 0) \leq \frac{r\lambda(V(x) - 1)}{1 - \lambda} + \frac{(1 - \tilde{b})(r^{1+\alpha} - 1)rV(x)}{(r - 1)(1 - (1 - b)r^{1+\alpha})},

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

$$\tilde{H}_1(r, x, 0) \leq H_1(r)(1 + \frac{(1 - \tilde{b})G(r)}{1 - (1 - b)G(r)}) = \frac{H_1(r, x, 0)}{1 - (1 - b)G(r)}, \text{ for } x \in C,$$

Hence again by $G(r) \leq r^{1+\alpha}$

(A.24) \hspace{1cm} (1 - \tilde{b})\tilde{H}_1(r, x, 0) \leq \frac{(1 - \tilde{b})(r^{1+\alpha} - 1)r}{(r - 1)(1 - (1 - b)r^{1+\alpha})} \text{ for } x \in C.

From (A.23) and (A.24) we conclude (A.20).

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

(A.25) \hspace{1cm} \tilde{H}_1(r, a, 1) = \frac{r\tilde{G}(r, a, 1) - 1}{r - 1} \leq \sup_{x \in C} r(G(r, x) - 1) = \frac{\sup_{x \in C} H_1(r, x)}{1 - (1 - b)G(r)}

and therefore

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

which is (A.21). To prove the last assertion we use (A.14) and (A.25), which imply that

$$\tilde{H}_1(r, a, 1) - r\tilde{H}_1(1, a, 1) \leq \frac{1}{b} \sup_{x \in C} H_1(r, x) - rH_1(r, x) +$$

+ \frac{(1 - \tilde{b})H_1(r)\sup_{x \in C}(G(r, x) - 1)}{1 - (1 - b)G(r)}.

The above inequality is equivalent to

$$\tilde{H}_1(r, a, 1) - r\tilde{H}_1(1, a, 1) \leq \frac{1}{b} \sup_{x \in C} H_1(r, x) - rH_1(r, x) +$$

+ \frac{(1 - \tilde{b})(G(r) - 1)\sup_{x \in C} H_1(r, x)}{1 - (1 - b)G(r)}.$$
Clearly $H_1(r, x) = \frac{r(G(r, x) - 1)}{r-1}$, thus by Proposition 4.3 we obtain that

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

Due to $G(r) \leq r^{1+\alpha}$ we deduce (A.22) and complete the proof of the result.

\[\blacksquare\]

### A.3 Case of $W \equiv V$

The second case we consider is when $W = V$.

**Proposition A.6.** The following inequalities hold

\[\text{(A.26)}\]

\[
(1-\bar{b}1_C(x))\bar{H}_V(r, x, 0) \leq \frac{\lambda r(V(x) - 1)}{1-r\lambda} + \frac{(K-r\lambda - \bar{b})}{1-(1-\bar{b})r^{1+\alpha_1}} \frac{rV(x)}{1-\lambda},
\]

for all $x \in S$, $1 \leq r \leq \min\{\lambda^{-1}, (1-\bar{b})^{-\frac{1}{1+\alpha_1}}\}$,

\[\text{(A.27)}\]

\[
\bar{H}_V(r, a, 1) \leq \bar{b}^{-1} \frac{r(K-r\lambda)}{1-r\lambda} + \frac{1}{1-(1-\bar{b})r^{1+\alpha_1}} \frac{r(K-\lambda)}{1-\lambda},
\]

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

\[\text{(A.28)}\]

\[
\frac{\bar{H}_V(r, a, 1) - r\bar{H}(1, a, 1)}{r-1} \leq \bar{b}^{-1} \frac{r(K-1)}{(1-\lambda)(1-r\lambda)} + \frac{1}{1-(1-\bar{b})r^{1+\alpha_1}} \frac{r(K-\lambda)}{1-\lambda},
\]

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

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

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

Therefore since $H_V(r, x, 0) = H_V(r, x)$ and $G(r, x, 0) = G(r, x)$ for all $x \not\in C$ we can use Propositions 4.1 and 4.5 to get

\[
\bar{H}_V(r, x, 0) \leq \frac{\lambda r(V(x) - 1)}{1-r\lambda} + \frac{(1-\bar{b})H_V(r)V(x)}{1-(1-\bar{b})G(r)}
\]
for all \( x \notin C \). Similarly for \( x \in C \)

\[
(1 - \bar{b})\bar{H}_V(r, x, 0) \leq (1 - \bar{b})H_V(r)(1 + \frac{(1 - \bar{b})G(r)}{1 - (1 - \bar{b})G(r)}) = \frac{(1 - \bar{b})H_V(r)}{1 - (1 - \bar{b})G(r)}.
\]

Hence using \( G(r) \leq r^{1+\alpha_1} \) we obtain that

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

Therefore it suffices to bound \((1 - \bar{b})H_V(r)\), note that

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

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

\[
(A.29) \quad (1 - \bar{b})H_V(r) \leq \frac{r(K - r\lambda)}{1 - r\lambda} - \bar{b}r,
\]

which establishes \((A.26)\). To show the remaining assertions we use bounds \((A.13)\), \((A.14)\) and \((A.6)\) obtaining that

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

and

\[
\bar{H}_V(r, a, 1) - r\bar{H}_V(r, a, 1) \leq \bar{b}^{-1}(H_V(r, a) - rH_V(1, x)) + \bar{b}^{-1}(1 - \bar{b})H_V(r)\sup_{x \in C}(G(r, x) - 1) \frac{1 - (1 - \bar{b})G(r)}{1 - (1 - \bar{b})G(r)}
\]

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

\[
\frac{G(r, x) - 1}{r - 1} \leq \frac{K - \lambda}{1 - \lambda}, \quad H_V(r, x) \leq \frac{r(K - r\lambda)}{1 - r\lambda},
\]

consequently

\[
\bar{H}_V(r, a, 1) \leq \bar{b}^{-1}r(K - r\lambda) \frac{1}{1 - r\lambda} + \bar{b}^{-1}(r - 1)(1 - \bar{b})H_V(r) K - \lambda \frac{1}{1 - \lambda}.
\]

Together with \((A.29)\) it completes the proof of \((A.27)\). Finally the same argument shows

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

Again by \((A.29)\) we obtain \((A.28)\), which completes the proof.
Appendix B

We compare our result with what was shown in [1] 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 $V(i) = (p/q)^i$, 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 $\lambda$.

Consider two cases:

1. In the first one we consider $p = 2/3$, so $b = 2/3$, $\lambda = 2\sqrt{2}/3$, $K = (2 + \sqrt{2})/3$, $u_\infty = 1/2$.

2. In the second one we set $p = 0.9$, and hence $\lambda = 0.6$, $K = 1.2$, $b = 0.9$, $u_\infty = 8/9$.

We compare our result in Table 1 below, where $\rho$, $\rho_C$ denotes estimates on the radius of convergence in the case where respectively $u_\infty$ is known or 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 [1], Meyn-Tweedie’s result given in [13] and its improved version (see Section 8 in [1] for details).

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\rho$</th>
<th>$\rho_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/3</td>
<td>0.9428</td>
<td>0.9428</td>
</tr>
<tr>
<td>Optimal</td>
<td>0.9428</td>
<td>0.9428</td>
</tr>
<tr>
<td>Bednorz</td>
<td>0.9737</td>
<td>0.9737</td>
</tr>
<tr>
<td>Baxendale</td>
<td>0.9994</td>
<td>?</td>
</tr>
<tr>
<td>Meyn-Tweedie1</td>
<td>0.9999</td>
<td>0.9988</td>
</tr>
<tr>
<td>Meyn-Tweedie2</td>
<td>0.9991</td>
<td>0.9927</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\rho$</th>
<th>$\rho_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Optimal</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Bednorz</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Baxendale</td>
<td>0.9060</td>
<td>?</td>
</tr>
<tr>
<td>Meyn-Tweedie1</td>
<td>0.9967</td>
<td>0.9888</td>
</tr>
<tr>
<td>Meyn-Tweedie2</td>
<td>0.9470</td>
<td>0.9467</td>
</tr>
</tbody>
</table>
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 [13] and also in [14] and [15]. By the algorithm definition \( P(x, \cdot) \) is distributed with a density

\[
p(x, y) = \begin{cases} 
\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-x)^2}{2}\right), & \text{if } |x| \geq |y| \\
\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-x)^2+y^2-x^2}{2}\right), & \text{if } |x| \leq |y|.
\end{cases}
\]

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 [1] for details)

\[
\lambda = \frac{PV(d)}{V(d)}, \quad K = PV(d) = e^{sd}\lambda.
\]

The computed value for \( \rho \) depends on \( d \) and \( s \), and hence to we need to find the optimal ones. Moreover to compare our result with the previous contributions to the problem, let \( \nu \) 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

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

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 whether or not we use the additional information on \( \pi(C) \)), Baxendale denotes what can be obtained by Baxendale’s Theorem 3.2, Coupling denotes the estimate obtained by the coupling approach (see Section 7 in [1] and [17]) and Meyn-Tweedie the result obtained in the original paper [13]. Note that we compare methods that no additional assumptions on the transition probabilities are made.

<table>
<thead>
<tr>
<th></th>
<th>( d )</th>
<th>( s )</th>
<th>( (1 - \rho) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bednorz1</td>
<td>0.96</td>
<td>0.065</td>
<td>0.00000529</td>
</tr>
<tr>
<td>Bednorz2</td>
<td>0.92</td>
<td>0.169</td>
<td>0.00005496</td>
</tr>
<tr>
<td>Baxendale</td>
<td>1</td>
<td>0.13</td>
<td>0.00000063</td>
</tr>
<tr>
<td>Coupling</td>
<td>1.8</td>
<td>1.1</td>
<td>0.00068</td>
</tr>
<tr>
<td>Meyn-Tweedie</td>
<td>1.4</td>
<td>0.0004</td>
<td>0.00000016</td>
</tr>
</tbody>
</table>
Another possible choice of $\nu$ is
\[
\bar{b}\nu(dx) = \begin{cases} 
\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|x|+d}{2}\right)dx, & \text{if } |x| \leq d \\
\frac{1}{\sqrt{2\pi}} \exp(-d|x|-|x|^2)dx, & \text{if } |x| \geq d
\end{cases}
\]
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>
<thead>
<tr>
<th></th>
<th>$d$</th>
<th>$s$</th>
<th>$(1-\rho)$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bednorz1</strong></td>
<td>1.03</td>
<td>0.0733</td>
<td>0.00001061</td>
</tr>
<tr>
<td><strong>Bednorz2</strong></td>
<td>0.97</td>
<td>0.1740</td>
<td>0.00013637</td>
</tr>
<tr>
<td><strong>Baxendale</strong></td>
<td>1</td>
<td>0.16</td>
<td>0.0000017</td>
</tr>
<tr>
<td><strong>Coupling</strong></td>
<td>1.9</td>
<td>1,1</td>
<td>0.00187</td>
</tr>
</tbody>
</table>

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 $P(x,\cdot) = \mathcal{N}(\theta x, 1-\theta^2)$ for some parameter $\theta \in (-1, 1)$. The family of examples occurs in [16] as one component of a two component Gibbs sampler. The example was discussed in [1], [14] and [15]. Here we take $V(x) = 1+x^2$ and $C = [-c, c]$. Then (2) is satisfied with
\[
\lambda = \theta^2 + 2\frac{1-\theta^2}{1+c^2}, \quad K = 2 + \theta^2(c^2 - 1).
\]
The we choose $\nu$ concentrated on $C$ so that
\[
\bar{b}\nu(dy) = \min_{x \in C} \frac{1}{\sqrt{2\pi(1-\theta^2)}} \exp\left(-\frac{(\theta x - y)^2}{2(1-\theta^2)}\right)dy
\]
for $y \in C$. Integrating with respect to $y$ gives
\[
\bar{b} = 2(\Phi\left(\frac{(1+|\theta|)c}{\sqrt{1-\theta^2}}\right) - \Phi\left(\frac{|\theta|c}{\sqrt{1-\theta^2}}\right)).
\]
We compare our answer **Bednorz1**, **Bednorz2** (Theorems 5.4, 5.2 resp.) with the coupling method **Coupling** and **Baxendale2** an approach based
on Kendal type result (Theorem 3.3 in [1]) that requires invertibility of the transition function.

Table 4

<table>
<thead>
<tr>
<th>θ</th>
<th>c</th>
<th>1 − ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bednorz1</td>
<td>0.5</td>
<td>1.5 0.000872023152</td>
</tr>
<tr>
<td>Bednorz1</td>
<td>0.75</td>
<td>1.2 0.000000964524</td>
</tr>
<tr>
<td>Bednorz1</td>
<td>0.9</td>
<td>1.1 0.000000000004</td>
</tr>
<tr>
<td>Bednorz2</td>
<td>0.5</td>
<td>1.5 0.002754672439</td>
</tr>
<tr>
<td>Bednorz2</td>
<td>0.75</td>
<td>1.2 0.000017954821</td>
</tr>
<tr>
<td>Bednorz2</td>
<td>0.9</td>
<td>1.1 0.000000000881</td>
</tr>
<tr>
<td>Baxendale2</td>
<td>0.5</td>
<td>1.5 0.050</td>
</tr>
<tr>
<td>Baxendale2</td>
<td>0.75</td>
<td>1.2 0.0042</td>
</tr>
<tr>
<td>Baxendale2</td>
<td>0.9</td>
<td>1.1 0.00002</td>
</tr>
<tr>
<td>Coupling</td>
<td>0.5</td>
<td>2.1 0.054</td>
</tr>
<tr>
<td>Coupling</td>
<td>0.75</td>
<td>1.7 0.0027</td>
</tr>
<tr>
<td>Coupling</td>
<td>0.9</td>
<td>1.5 0.00002</td>
</tr>
</tbody>
</table>

B.4 Reflecting random walk, continued.

Here we slightly redefine our first example. Let \( P(0, \{0\}) = 1 \) and \( P(0, \{1\}) = 1 − \varepsilon \) for some \( \varepsilon > 0 \). We concentrate on the difficult case, when \( \varepsilon < p \), that was studied in [15] and [5]. Note that when \( \varepsilon \geq p \) then the chain is stochastically monotone and then the result of Tweedie [7] apply. Let \( V(i) = (p/q)^{i/2}, \ C = \{0\} \) 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 on \( b(z) \) which is

\[
B.1 \quad b(z) = G(z, 0) = \varepsilon z + (1 - \varepsilon)G(z, 1) = \varepsilon z + \frac{1 - \varepsilon}{2q} \left( 1 - (1 - 4pqz^2)^{1/2} \right),
\]

for \( |z| < 1/\sqrt{4pq} \), where the formula for \( G(z, 1) \) is from [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 on the radius on convergence which is

\[
\rho = \begin{cases} 
\frac{pq + (p - \varepsilon)^2}{p - \varepsilon}, & \text{if } \varepsilon < \frac{p - q}{1 + \sqrt{q/p}}, \\
2\sqrt{pq}, & \text{otherwise}.
\end{cases}
\]
We compare Bednorz1,Bednorz - our Corollaries 2.3.2.5 with results Fort and Baxendale that denotes respectively the result of Fort [5] and Baxendale’s Theorem 1.2 [1]. Note that both two 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           |
| $p = 0.7$        | $\varepsilon = 0.05$ | $\varepsilon = 0.25$ | $\varepsilon = 0.5$ |
| **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           |

References


