On some possible combinations of mixing rates for strictly stationary, reversible Markov chains

Richard C. Bradley

Department of Mathematics, Indiana University, Bloomington, IN 47405, U.S.A.
bradleyr@indiana.edu

January 29, 2023

Abstract

A class of examples is constructed to show that for strictly stationary Markov chains that are reversible, the simultaneous mixing rates for the ρ-mixing and strong mixing (α-mixing) conditions can be fairly arbitrary, within certain unavoidable tight restrictions. The examples constructed here have the added property that the mixing rate for the absolute regularity (β-mixing) condition is within a constant factor of that for strong mixing.

Keywords: strictly stationary, reversible Markov chain;
strong mixing; absolute regularity; ρ-mixing

AMS 2020 Mathematics Subject Classifications: 60J10, 60G10

1 Introduction

Theory developed in the papers by Roberts and Rosenthal [19] and Roberts and Tweedie [20] together produced, among other things, a "key result" giving, for strictly stationary, reversible Markov chains, a connection between the "geometric ergodicity" condition and a certain "spectral gap" condition from functional analysis. It is well known that that "key result" can be formulated in the terminology of the dependence coefficients associated with the "absolute regularity" ("β-mixing") condition and the "ρ-mixing" condition. All that will be discussed in a little more detail after Theorem 1.2 below. In an effort to provide a relatively gentle exposition of that "key result" and some related theory from the two aforementioned papers, the author [4] employed the dependence coefficients associated with the "strong mixing" ("α-mixing") condition as a means to try to organize that exposition into a somewhat more transparent form.

That expository effort in [4] in turn leads to the question of what combinations of mixing rates, within some unavoidable narrow constraints, are possible for the ρ-mixing and strong mixing conditions together — and perhaps including the absolute regularity condition as well — for strictly stationary, reversible Markov chains. This paper here is intended to give at least a partial answer to that question. The construction given in this paper — a class of strictly stationary, reversible Markov chains — will illustrate, within the unavoidable narrow constraints alluded to above, a pretty broad spectrum of possible combinations of such mixing rates. As the examples will also involve a state space that is countable (rather than, say, the whole real number line), the arguments involved in checking the various properties of the examples here will be pretty elementary.

To set up further discussion, Setting 1.1 below will lay out some key terminology and background facts.

Setting 1.1. First a few preliminary formalities: Throughout this paper,
$\mathbb{R}$ denotes the set of all real numbers,
$\mathbb{Z}$ denotes the set of all integers, and
$\mathbb{N}$ denotes the set of all positive integers.
In this paper, a set S is said to be "countable" if S is either finite or countably infinite.

1

The usual notations (such as $\mathbb{R}^{\mathbb{Z}}$) will be used for Cartesian products of sets.

The notation $\log x$ refers to the natural logarithm.

If $(a_n, n \in \mathbb{N})$ and $(b_n, n \in \mathbb{N})$ are each a sequence of positive numbers, then the notation $a_n \asymp b_n$ as $n \to \infty$ will mean that $a_n = O(b_n)$ and $b_n = O(a_n)$ as $n \to \infty$.

If $(a_n, n \in \mathbb{N})$ is a sequence of nonnegative numbers, then the notation "$a_n \to 0$ at least exponentially fast as $n \to \infty$" means that there exists a positive number c such that $a_n = O(e^{-cn})$ as $n \to \infty$.

Part 1 (*The probability space*). The setting for the probability theory in this paper is a probability space $(\Omega, \mathcal{F}, P)$, rich enough to accommodate all random variables declared. All random variables in this paper are defined on that probability space.

For simplicity, *all random variables in this paper are real-valued*. Some may be discrete, including those in the examples constructed in this paper. (It is to be tacitly understood that some of the sources cited in this paper dealt with random variables taking their values in more general spaces than the real numbers; however, for the material in this paper, there is no need to deal with that greater generality.)

For a given family $(Y_i, i \in I)$ of (real-valued) random variables (where I is a nonempty index set), the σ-field ($\subset \mathcal{F}$, on Ω) generated by this family will be denoted $\sigma(Y_i, i \in I)$.

Part 2 (*Three measures of dependence*). Suppose $\mathcal{A}$ and $\mathcal{B}$ are any two σ-fields $\subset \mathcal{F}$. Define the following three measures of dependence: First,

$$\alpha(\mathcal{A}, \mathcal{B}) := \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A)P(B)|. \tag{1.1}$$

Next,

$$\beta(\mathcal{A}, \mathcal{B}) := \sup \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)| \tag{1.2}$$

where the supremum is taken over all pairs of finite partitions $\{A_1, A_2, \ldots, A_I\}$ and $\{B_1, B_2, \ldots, B_J\}$ of Ω such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j. [The factor of $1/2$ in (1.2) is not of special significance, but has become customary in order to make certain inequalities "a little nicer".] Finally, define the "maximal correlation coefficient"

$$\rho(\mathcal{A}, \mathcal{B}) := \sup |\mathrm{Corr}(Y, Z)| \tag{1.3}$$

where the supremum is taken over all pairs of square-integrable random variables Y and Z such that Y is $\mathcal{A}$-measurable and Z is $\mathcal{B}$-measurable.

The following inequalities are elementary and well known:

$$0 \le 2\alpha(\mathcal{A}, \mathcal{B}) \le \beta(\mathcal{A}, \mathcal{B}) \le 1; \quad \text{and} \tag{1.4}$$

$$0 \le 4\alpha(\mathcal{A}, \mathcal{B}) \le \rho(\mathcal{A}, \mathcal{B}) \le 1. \tag{1.5}$$

(See e.g. [[2], v.1, Proposition 3.11].) The quantities $\alpha(\mathcal{A}, \mathcal{B})$, $\beta(\mathcal{A}, \mathcal{B})$, and $\rho(\mathcal{A}, \mathcal{B})$ are all equal to 0 if the σ-fields $\mathcal{A}$ and $\mathcal{B}$ are independent, and are all positive otherwise.

Part 3 (*Three mixing conditions*). Suppose $X := (X_k, k \in \mathbb{Z})$ is a (not necessarily Markovian) strictly stationary sequence of (real-valued) random variables. For each integer j, define the notations $\mathcal{F}_{-\infty}^{j} := \sigma(X_k, k \le j)$ and $\mathcal{F}_j^{\infty} := \sigma(X_k, k \ge j)$.

For each positive integer n, define the following three "dependence coefficients":

$$\alpha(n) = \alpha_X(n) := \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{\infty}); \tag{1.6}$$

$$\beta(n) = \beta_X(n) := \beta(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{\infty}); \text{ and} \tag{1.7}$$

$$\rho(n) = \rho_X(n) := \rho(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{\infty}). \tag{1.8}$$

For each positive integer n, one has by strict stationarity that $\alpha(n) = \alpha(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^{\infty})$ for every integer j, and the analogous comment holds for $\beta(n)$ and for $\rho(n)$ as well.

Also, one (trivially) has that $\alpha(1) \ge \alpha(2) \ge \alpha(3) \ge \ldots$; and the analogous comment holds for the numbers $\beta(n)$ and for the numbers $\rho(n)$.

2

The strictly stationary sequence X is said to satisfy
"strong mixing" (or "α-mixing") if $\alpha_X(n) \to 0$ as $n \to \infty$;
"absolute regularity" (or "β-mixing") if $\beta_X(n) \to 0$ as $n \to \infty$;
"ρ-mixing" if $\rho_X(n) \to 0$ as $n \to \infty$.

The strong mixing (α-mixing) condition is due to Rosenblatt [21]. The absolute regularity (β-mixing) condition was first studied by Volkonskii and Rozanov [24], and was attributed there to Kolmogorov. The ρ-mixing condition is due to Kolmogorov and Rozanov [10]. (The "maximal correlation coefficient" $\rho(\mathcal{A}, \mathcal{B})$ itself, for σ-fields $\mathcal{A}$ and $\mathcal{B}$, was first studied earlier by Hirschfeld [8] in a statistical context that had no particular connection with stochastic processes.)

By (1.4) and (1.5), one has that for each positive integer n,

$$0 \leq 2\alpha(n) \leq \beta(n) \leq 1, \text{ and} \tag{1.9}$$

$$0 \leq 4\alpha(n) \leq \rho(n) \leq 1. \tag{1.10}$$

By (1.9), absolute regularity (β-mixing) implies strong mixing (α-mixing); and by (1.10), ρ-mixing implies strong mixing (α-mixing). [Part 5(E) below, and Example 1 of Remark 1.4 later on, will together illustrate the fact that of the absolute regularity and ρ-mixing conditions, neither implies the other.]

Part 4 (*Strictly stationary Markov chains*). Now suppose that $X := (X_k, k \in \mathbb{Z})$ is a strictly stationary *Markov chain* (with the random variables X_k, $k \in \mathbb{Z}$ being real-valued, possibly discrete). (No assumption of "reversibility" yet.)

As a well known consequence of the Markov property, for each positive integer n, eqs. (1.6)-(1.8) hold in the following augmented forms for the given (strictly stationary) Markov chain X:

$$\alpha(n) = \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty) = \alpha\big(\sigma(X_0), \sigma(X_n)\big); \tag{1.11}$$

$$\beta(n) = \beta(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty) = \beta\big(\sigma(X_0), \sigma(X_n)\big); \tag{1.12}$$

$$\rho(n) = \rho(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty) = \rho\big(\sigma(X_0), \sigma(X_n)\big). \tag{1.13}$$

(See e.g. [[2], v.1, Theorem 7.3].)

By strict stationarity and (1.11)-(1.13), one has that for any integer j and any positive integer n, the (strictly stationary) Markov chain X satisfies
(i) $\alpha(n) = \alpha(\sigma(X_j), \sigma(X_{j+n}))$,
(ii) $\beta(n) = \beta(\sigma(X_j), \sigma(X_{j+n}))$, and
(iii) $\rho(n) = \rho(\sigma(X_j), \sigma(X_{j+n}))$.

Here are some special facts involving the dependence coefficients $\rho(n)$, $n \in \mathbb{N}$. As an elementary consequence of (1.3), for any two σ-fields $\mathcal{A}$ and $\mathcal{B}$, $\rho(\mathcal{A}, \mathcal{B}) = \sup \|E(Y|\mathcal{B})\|_2 / \|Y\|_2$ where the supremum is taken over all square-integrable, $\mathcal{A}$-measurable random variables Y with mean 0. (When necessary, interpret $0/0 := 0$.) As a well known application of that fact (together with the equality in (iii) in the preceding paragraph), for the given strictly stationary Markov chain $X := (X_k, k \in \mathbb{Z})$, one has that for any pair of positive integers m and n,

$$\rho(m + n) \leq \rho(m) \cdot \rho(n). \tag{1.14}$$

By (1.14), for any two positive integers m and n, $\rho(m(n+1)) \leq \rho(mn) \cdot \rho(m)$. Hence by induction, for every positive integer m, one has that the given (strictly stationary) Markov chain X satisfies

$$\rho(mn) \leq [\rho(m)]^n \text{ for every } n \in \mathbb{N}. \tag{1.15}$$

In particular (take $m = 1$),

$$\rho(n) \leq [\rho(1)]^n \text{ for every } n \in \mathbb{N}. \tag{1.16}$$

By (1.15), for the given (strictly stationary) Markov chain $X := (X_k, k \in \mathbb{Z})$, the following three conditions are equivalent:
(i) there exists $m \geq 1$ such that $\rho(m) < 1$;

3

(ii) X is ρ-mixing;

(iii) $\rho(n) \to 0$ at least exponentially fast as $n \to \infty$.

Part 5 (*Strictly stationary, reversible Markov chains*). A given strictly stationary Markov chain $X := (X_k, k \in \mathbb{Z})$ is said to be "reversible" if the distribution (on $\mathbb{R}^{\mathbb{Z}}$) of the "time-reversed" sequence $(X_{-k}, k \in \mathbb{Z})$ is identical to that of the sequence X itself.

(A) By a well known argument, a given strictly stationary Markov chain $X := (X_k, k \in \mathbb{Z})$ is reversible if and only if the distribution (on $\mathbb{R}^2$) of the random vector (X_1, X_0) is identical to that of the random vector (X_0, X_1).

(B) If a given strictly stationary Markov chain $X := (X_k, k \in \mathbb{Z})$ is reversible, then

$$\rho(n) = [\rho(1)]^n \text{ for every } n \in \mathbb{N}. \tag{1.17}$$

That was shown by Longla [[13], Lemma 2.1] in the context of strictly stationary, reversible Markov chains involving certain types of copulas (see e.g. [16]); his argument extends beyond that context. It can also be seen as an application of a certain theorem in functional analysis involving self-adjoint bounded linear operators from a Hilbert space to itself. (This latter fact, explained in more detail in [4], was pointed out and contributed there by an anonymous referee of that paper. A presentation of the proof in probabilistic terminology was also included in the exposition in that paper.) Compare (1.17) to (1.16) (where reversibility was not assumed).

(C) The paper [3] constructed some strictly stationary, countable-state, reversible Markov chains that are ρ-mixing but fail to satisfy ρ^*-mixing — the "interlaced" variant of ρ-mixing with the two index sets not restricted to "past" and "future". In those examples, $\rho(\sigma(X_0), \sigma(X_{-n}, X_n)) = 1$ for every $n \in \mathbb{N}$.

(D) As is well known, any strictly stationary *two-state* Markov chain $X := (X_k, k \in \mathbb{Z})$ is reversible.

The explanation is simple. Let us label the two states as 0 and 1. Then by strict stationarity,

$$
\begin{aligned}
P(\{X_0 = 0\} \cap \{X_1 = 1\}) &= P(X_0 = 0) - P(\{X_0 = 0\} \cap \{X_1 = 0\}) \\
&= P(X_1 = 0) - P(\{X_0 = 0\} \cap \{X_1 = 0\}) = P(\{X_0 = 1\} \cap \{X_1 = 0\}).
\end{aligned}
$$

That is, $P((X_0, X_1) = (0, 1)) = P((X_1, X_0) = (0, 1))$. Now it trivially follows that $P((X_0, X_1) = (i, j)) = P((X_1, X_0) = (i, j))$ for all ordered pairs $(i, j) \in \{0, 1\}^2$. Hence by (A) above, one has that (D) holds.

(E) There exist strictly stationary, reversible Markov chains $X := (X_k, k \in \mathbb{Z})$ that satisfy (i) ρ-mixing, and (ii) $\beta(n) = 1$ for all $n \in \mathbb{N}$. Condition (ii) is (for strictly stationary Markov chains) equivalent to the condition that for every positive integer n, the n-step transition distributions are almost surely totally singular with respect to the (invariant) marginal distribution. Such examples cannot be countable-state.

Using "random rotations", Rosenblatt [[22], pp. 214-215] constructed some strictly stationary Markov chains that satisfy conditions (i) and (ii) above. It seems clear that some of those examples (the ones for which the underlying "random rotations" satisfy an appropriate symmetry) are reversible. As a slight variant of those examples of Rosenblatt, the author [[2], Examples 7.16 and 7.17] constructed examples of strictly stationary, reversible Markov chains that satisfy (i) and (ii) above, and showed that those examples in fact satisfy the ρ^*-mixing condition alluded to in (C) above. Reversibility was not mentioned there (for those examples in [2]), but is easily verified as a consequence of the reversibility of the "building blocks" for those examples — strictly stationary, two-state (hence reversible) Markov chains (see (D) above).

Comments. As motivation for the study of strictly stationary, reversible Markov chains, we shall summarize here a few comments from the papers of Longla [12] [13]. In the use of the Metropolis-Hastings algorithm for some aspects of statistical inference for unknown probability distributions, it is common practice to "simulate" (conditional on the choice of a particular "starting point" at a particular "starting time") a (say) strictly stationary Markov chain whose (invariant) marginal distribution is uniform on the unit interval $[0, 1]$. In that process, one deals with copulas (see e.g. [16]) — functions $C : [0, 1] \times [0, 1] \to [0, 1]$ which are the joint (cumulative) distribution functions of random vectors (U, V) for which both marginal distributions are uniform on the unit interval. Such a copula C "induces" (in distribution) a (not necessarily reversible) strictly stationary Markov chain $X := (X_k, k \in \mathbf{Z})$ for which, for each integer k, the joint distribution function of the random

4

vector (X_k, X_{k+1}) is C. In the case where C is symmetric $[C(u,v) = C(v,u)]$, the induced strictly stationary Markov chain is reversible. The statistical inference alluded to above can be simplified and facilitated by the use of appropriate symmetric copulas for which the resulting strictly stationary, reversible Markov chains satisfy $\rho(1) < 1$ (and are consequently ρ-mixing with exponential mixing rate). Longla [12] [13] explained all that with considerable detail, and also obtained bounds on $\rho(1)$ for the strictly stationary Markov chains induced by several special classes of copulas. One class that Longla examined explicitly in those two papers is the Frechet (Mardia) family of copulas given by $C_{a,b}(u,v) = aM(u,v) + (1 - a - b)P(u,v) + bW(u,v)$ where (i) $a \geq 0$, $b \geq 0$, and $a + b \leq 1$, and (ii) $M(u,v) = \min\{u,v\}$ (the "Hoeffding upper-bound copula"), $W(u,v) = \max\{u + v - 1, 0\}$ (the "Hoeffding lower-bound copula"), and $P(u,v) = uv$ (the "independence copula"). Those copulas are all symmetric. By [[13], Theorem 2.16], in the case where $a + b < 1$, a strictly stationary, reversible Markov chain induced by the copula $C_{a,b}$ satisfies $\rho(1) < 1$ and hence ρ-mixing (with exponential mixing rate), but in the case where $a + b = 1$ that Markov chain does not satisfy any mixing properties at all.

The following known theorem will help focus the discussion on the material in the rest of this paper.

Theorem 1.2. *Suppose $X := (X_k, \ k \in \mathbb{Z})$ is a strictly stationary, reversible Markov chain, and r is a real number such that $0 < r < 1$. Then the following four conditions are equivalent:*
(i) $\rho_X(1) \leq r$.
(ii) For all $n \in \mathbb{N}$, $\rho_X(n) \leq r^n$.
(iii) For all $n \in \mathbb{N}$, $\alpha_X(n) \leq r^n$.
(iv) $\alpha_X(n) = O(r^n)$ as $n \to \infty$.

Except for one superficial simplification here, this theorem is [[4],Corollary 5.7]. Obviously (i) $\Leftrightarrow$ (ii) by Part 5(B) of Setting 1.1; and (see (1.10)) trivially (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). The proof of the "remaining implication" (iv) $\Rightarrow$ (i) is the argument that was given for [[4], Lemma 5.5] — an argument that was just a slight variant of calculations in the papers of Roberts and Rosenthal [19] and Roberts and Tweedie [20].

The theory developed by Roberts, Rosenthal, and Tweedie in those two papers [19] and [20] contains (in greater generality, and together with other things) the "key result" alluded to earlier. Without formally defining all of the necessary terminology, one can at least state that "key result" informally as follows: *If a given strictly stationary Markov chain is reversible and satisfies a certain "irreducibility" condition (equivalent to Harris recurrence), then the "geometric ergodicity" condition holds if and only if a certain "L^2 spectral gap" condition from functional analysis holds.* As noted earlier, it is well known that that "key result" and other related results can be transcribed into the terminology of dependence coefficients. See for example [11], [12], [13], and [14], or the detailed exposition in [4]. The papers of Nummelin and Tweedie [18] and Nummelin and Tuominen [17] together had (among other things) already shown (in different terminology) that for a given strictly stationary Markov chain (reversible or not), the geometric ergodicity condition is equivalent to absolute regularity with $\beta(n) \to 0$ at least exponentially fast. (For more on geometric ergodicity, see e.g. [15] or [[2], v.2, Chapter 21].) It is elementary and long well known that the "L^2 spectral gap" condition alluded to above is, for a given strictly stationary Markov chain (reversible or not) equivalent to the condition $\rho(1) < 1$. The "key result" of Roberts, Rosenthal, and Tweedie from above, together with a "comparison of rates of convergence" calculation of theirs connected with it, can together be formulated in a way that is "almost" (but "not quite") an analog of Theorem 1.2 with $\beta(n)$ in place of $\alpha(n)$.

All that is reviewed in a bit of detail in [4], where in essence Theorem 1.2 was developed mainly in order to assist an expository effort to provide a gentle introduction to the "key result" of Roberts, Rosenthal, and Tweedie. In contrast to that "key result", Theorem 1.2 does not require any assumption of "irreducibility" (or Harris recurrence); it applies just as well to the peculiar ρ-mixing examples alluded to in Part 5(E) of Setting 1.1, where the property $\beta(n) = 1$ for all $n \in \mathbb{N}$ actually prevents such "irreducibility" or Harris recurrence (and prevents the state space from being countable).

In this paper here, with Theorem 1.2 as the background starting point, we shall investigate the question of what combinations of behavior of the dependence coefficients $\alpha(n)$ and $\rho(n)$ are possible for strictly stationary, reversible Markov chains.

In addition to the sources cited above in connection with the "key result" of Roberts, Rosenthal, and Tweedie, the main sources of inspiration for this paper here are some earlier papers that (among other

things) established some "slower than exponential" mixing rates for the strong mixing condition for certain classes of strictly stationary (not necessarily reversible) Markov chains with a "renewal structure". To illustrate certain limitations in connection with central limit theorems for strictly stationary (not necessarily Markovian), strongly mixing sequences, Davydov [6], Tikhomirov [23], and Doukhan, Massart, and Rio [7] each constructed strictly stationary Markov chains for which, for some appropriate positive number c, $\alpha(n) \asymp \beta(n) \asymp n^{-c}$ as $n \to \infty$. In those papers, the main focus was on $\alpha(n)$. For the examples in [6] and [23], the inclusion of $\beta(n)$ was implicitly established by the calculations of Davydov [[6], pp. 327-328]; and in the paper [7] the inclusion of $\beta(n)$ was established explicitly. Kesten and O'Brien [[9], pp. 412-414] constructed a broad class of examples of strictly stationary, strongly mixing Markov chains with a quite arbitrary "slower than exponential" mixing rate. The Markov chains there (as well as those in [6] and [23]) were countable-state. The Markov chains constructed in [7] are reversible (and have a "continuum" state space, namely the unit interval $[0, 1]$).

This paper here is focused partly on extending the observations of Kesten and O'Brien [9] cited above (displaying a broad spectrum of possible "sub-exponential" mixing rates for $\alpha(n)$ for strictly stationary Markov chains) to strictly stationary, Markov chains that are reversible. This paper is also intended to give, for strictly stationary Markov chains that are reversible and ρ-mixing, similar results on a "broad spectrum of pairs of mixing rates" (to the narrow extent permitted by Part 5(B) of Setting 1.1 and by Theorem 1.2) involving the dependence coefficients $\rho(n)$ and $\alpha(n)$ together. The Markov chains constructed in this paper will be countable-state; and they will include absolute regularity, with $\alpha(n) \asymp \beta(n)$ as $n \to \infty$.

By (1.17) and the second sentence after (1.5), for a given strictly stationary, reversible Markov chain, the dependence coefficients $\alpha(n)$, $\beta(n)$, and $\rho(n)$, $n \in \mathbb{N}$, will *all* be positive – except in the case of a sequence of independent, identically distributed random variables (where those dependence coefficients are all zero). In the "dependent" cases, i.e. where $0 < \rho(1) \leq 1$ (the value 1 is included here), the ratio $\alpha(n)/\rho(n)$ is bounded above by $1/4$, by (1.10). The focus in this paper will be on cases where that ratio converges to 0 as $n \to \infty$.

Let it first be noted in passing that [see Part 5(D) of Setting 1.1] for some well known strictly stationary, two-state, (hence) reversible Markov chains, the strong mixing, ρ-mixing, and absolute regularity conditions hold with the same (exponential) mixing rate modulo a constant factor. In Section 2, that will be pointed out in concrete simple detail in Lemma 2.5 and its subsequent Remark.

Now (in the "dependent" case) as a consequence of Theorem 1.2 (with no changes), one has that whether $0 < \rho_X(1) < 1$ or $\rho_X(1) = 1$, if the ratio $\alpha(n)/\rho(n)$ converges to 0, it cannot do so at an (at least) exponential rate. The class of examples that will be presented in this paper, will show that that ratio can converge to 0 at a quite arbitrary "slower than exponential" rate, at least within a mild "log convexity" condition — regardless of whether ρ-mixing holds or instead $\rho(n) = 1$ for all $n \in \mathbb{N}$ [see (1.17) or again the last sentence (i.e. the equivalence) of Part 4 of Setting 1.1].

Here is our main result (stated here with some redundancy):

Theorem 1.3. *Suppose r is a real number such that $0 < r \leq 1$. Suppose $f : [0, \infty) \to (0, 1/2]$ is a continuous, strictly decreasing function with the following three properties:*
(i) $f(x) \to 0$ as $x \to \infty$;
(ii) for every $u \in (0, 1)$, $u^x = o(f(x))$ as $x \to \infty$; and
(iii) the mapping $x \mapsto \log f(x)$, for $x \in [0, \infty)$, is convex on $[0, \infty)$.

Then there exists (on the probability space $(\Omega, \mathcal{F}, P)$) a strictly stationary, countable-state Markov chain $X := (X_k, \ k \in \mathbb{Z})$ such that X is reversible and for every positive integer n,

$$\rho_X(n) = r^n \quad \text{and} \quad (1/2)\, r^n f(n) \leq \alpha_X(n) \leq \beta_X(n) \leq 12\, r^n f(n). \tag{1.18}$$

Here the numbers $1/2$ (twice) and 12 are not of special significance, but are instead just an artifact of an endeavor to make the statement and proof of Theorem 1.3 a "little nicer". [Again recall (1.10).] Theorem 1.3 will be proved in Section 3, after some preliminary work is done in Section 2. The rest of Section 1 here will be devoted to certain specific illustrations of Theorem 1.3.

Remark 1.4. Here we shall take a quick look at a couple of related specific applications of Theorem 1.3. In the first, with $r = 1$, ρ-mixing fails to hold; in the second, with $0 < r < 1$, ρ-mixing holds.

Example 1. This example is in spirit an adaptation, to reversibility, of an illustration in the work of Kesten and O'Brien [9] alluded to above.

Suppose that $0 < a < 1$ and $q > 0$, and also that b and c are each an arbitrary real number. Define the function $\eta : (1, \infty) \to \mathbb{R}$ as follows:

$$\text{for all } x \in (1, \infty), \quad \eta(x) := \big(\exp(-qx^a)\big) \cdot x^b \cdot (\log x)^c.$$

Then $\eta(x) > 0$ for all $x > 1$. Define the function $h : (1, \infty) \to \mathbb{R}$ as follows:

$$\text{for all } x \in (1, \infty), \quad h(x) := \log(\eta(x)) = -qx^a + b(\log x) + c\log(\log x).$$

Then by elementary calculations, the following statements hold:
(i) $\eta(x) \to 0$ as $x \to \infty$; and for all $t > 0$, $\eta(t+x)/\eta(x) \to 1$ as $x \to \infty$;
(ii) $h'(x) \to 0$ as $x \to \infty$; and
(iii) for all sufficiently large $x > 1$, one has that $h'(x) < 0$ and $h''(x) > 0$.

From all of the above observations on the functions η and h, one has the following: If $T > 0$ is chosen sufficiently large, the function $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = \eta(T + x)$ for $x \in [0, \infty)$ will be well defined and will satisfy all of the hypotheses (for f) in Theorem 1.3. Consequently, by that theorem (with $r = 1$ there), there exists a strictly stationary, reversible Markov chain $X := (X_k, k \in \mathbb{Z})$ (with countable state space), such that $\rho(n) = 1$ for all $n \in \mathbb{N}$, and strong mixing and absolute regularity both hold with mixing rates

$$\alpha_X(n) \asymp \beta_X(n) \asymp \big(\exp(-qn^a)\big) \cdot n^b \cdot (\log n)^c \quad \text{as } n \to \infty.$$

Example 2. Now for a given r such that $0 < r < 1$, and for a, q, b, and c with the same restrictions as in Example 1 above, this time applying Theorem 1.3 with the given r and (again) $f(x) = \eta(T + x)$ where η and T are as above, one has that there exists a strictly stationary, countable-state, reversible Markov chain $X := (X_k, k \in \mathbb{Z})$ such that ρ-mixing holds with $\rho_X(n) = r^n$ for all $n \in \mathbb{N}$, and the strong mixing and absolute regularity conditions hold with mixing rates

$$\alpha_X(n) \asymp \beta_X(n) \asymp r^n \cdot \big(\exp(-qn^a)\big) \cdot n^b \cdot (\log n)^c \quad \text{as } n \to \infty.$$

2   Preliminaries

Notations 2.1. The construction (in Section 3) of the strictly stationary, countable-state, reversible Markov chain for Theorem 1.3 will involve as "building blocks" a countably infinite collection of strictly stationary, 2-state, reversible Markov chains that are independent of each other. (Again recall Part 5(D) of Setting 1.1.) The arithmetic for that process apparently will be slightly less "cluttered" if in those 2-state "building block" Markov chains, the states are labeled 0 and 1 (instead of, say, 1 and 2). Accordingly, the 2×2 matrices involved in that process (for example, the n-step transition probability matrices) will be set up in the form $M := (m_{ij})_{i,j \in \{0,1\}}$, with the top row being $[m_{00}, m_{01}]$ and the bottom row being $[m_{10}, m_{11}]$.

Part 1. *Joint probability matrices.* First we shall spell out a class of 2×2 matrices that will be used for *joint* (not transition) probabilities.

For any $\varepsilon \in (0, 1/2]$ and any $\theta \in (0, 1)$, define the 2×2 matrix $\Lambda^{(\varepsilon,\theta)} := (\lambda_{ij}^{(\varepsilon,\theta)})_{i,j \in \{0,1\}}$ as follows:

$$\lambda_{00}^{(\varepsilon,\theta)} := (1 - \varepsilon)^2 + (1 - \varepsilon)\varepsilon\theta; \tag{2.1}$$

$$\lambda_{01}^{(\varepsilon,\theta)} = \lambda_{10}^{(\varepsilon,\theta)} := (1 - \varepsilon)\varepsilon - (1 - \varepsilon)\varepsilon\theta; \quad \text{and} \tag{2.2}$$

$$\lambda_{11}^{(\varepsilon,\theta)} := \varepsilon^2 + (1 - \varepsilon)\varepsilon\theta. \tag{2.3}$$

Note that (under the stipulated conditions on ε and θ), these four entries are all positive, and their sum is 1. We shall return to that class of matrices in Lemma 2.2 below.

Part 2. *Transition probability matrices.* The matrices below will play a key role in transition probabilities (of some strictly stationary, 2-state Markov chains).

The 2×2 identity matrix $(\delta_{ij})_{i,j \in \{0,1\}}$ (where $\delta_{00} = \delta_{11} = 1$ and $\delta_{01} = \delta_{10} = 0$) will be denoted simply as I_2.

For any $\varepsilon \in (0, 1/2]$, let $A^{(\varepsilon)} := (a_{ij}^{(\varepsilon)})_{i,j \in \{0,1\}}$ be the 2×2 matrix in which each of the two rows is $[1 - \varepsilon, \varepsilon]$.

For any $\varepsilon \in (0, 1/2]$ and any $\theta \in (0, 1)$, define the 2×2 matrix $\mathbb{P}^{(\varepsilon, \theta)} := (p_{ij}^{(\varepsilon, \theta)})_{i,j \in \{0,1\}}$ as follows:

$$\mathbb{P}^{(\varepsilon, \theta)} := \theta I_2 + (1 - \theta) A^{(\varepsilon)}. \tag{2.4}$$

By simple arithmetic, for any $\varepsilon \in (0, 1/2]$ and any $\theta \in (0, 1)$, the entries of the 2×2 matrix $\mathbb{P}^{(\varepsilon, \theta)}$ are as follows:

$$p_{00}^{(\varepsilon, \theta)} = (1 - \varepsilon) + \varepsilon\theta \quad \text{and} \quad p_{01}^{(\varepsilon, \theta)} = \varepsilon - \varepsilon\theta; \tag{2.5}$$

$$p_{10}^{(\varepsilon, \theta)} = (1 - \varepsilon) - (1 - \varepsilon)\theta \quad \text{and} \quad p_{11}^{(\varepsilon, \theta)} = \varepsilon + (1 - \varepsilon)\theta. \tag{2.6}$$

Note that in each of the two rows of $\mathbb{P}^{(\varepsilon, \theta)}$, the two entries are positive and their sum is 1.

By trivial matrix multiplication, $(A^{(\varepsilon)})^2 = A^{(\varepsilon)}$ for any $\varepsilon \in (0, 1/2]$. Hence by simple matrix multiplication, for any $\varepsilon \in (0, 1/2]$ and any pair of numbers $\theta, \tau \in (0, 1)$, one has that

$$\mathbb{P}^{(\varepsilon, \theta)} \mathbb{P}^{(\varepsilon, \tau)} = \mathbb{P}^{(\varepsilon, \theta\tau)}. \tag{2.7}$$

We shall return to those notations in Lemma 2.3 below.

Part 3. *Alternative notations.* In the case where the number $\theta \in (0, 1)$ itself involves "smaller print", the matrices $\Lambda^{(\varepsilon, \theta)}$ and $\mathbb{P}^{(\varepsilon, \theta)}$ may be written as $\Lambda(\varepsilon, \theta)$ and $\mathbb{P}(\varepsilon, \theta)$ respectively for typographical convenience.

Lemma 2.2. *Suppose $\varepsilon \in (0, 1/2]$ and $\theta \in (0, 1)$. Suppose Y and Z are $\{0, 1\}$-valued random variables such that the (joint) probability function of the random vector (Y, Z) is the matrix $\Lambda^{(\varepsilon, \theta)}$ in (2.1)-(2.3), that is,*

$$\text{for all } (i, j) \in \{0, 1\}^2, \quad P\big((Y, Z) = (i, j)\big) = \lambda_{ij}^{(\varepsilon, \theta)}. \tag{2.8}$$

Then the following statements hold:
(I) The (joint) probability function of the random vector (Z, Y) is the same as that of (Y, Z).
(II) $P(Y = 0) = P(Z = 0) = 1 - \varepsilon$ and $P(Y = 1) = P(Z = 1) = \varepsilon$.
(III) $\text{Corr}(Y, Z) = \theta$.
(IV) $\rho(\sigma(Y), \sigma(Z)) = \theta$.
(V) $\alpha(\sigma(Y), \sigma(Z)) = (1 - \varepsilon)\varepsilon\theta \geq \varepsilon\theta/2$.
(VI) $\beta(\sigma(Y), \sigma(Z)) = 2(1 - \varepsilon)\varepsilon\theta \leq 2\varepsilon\theta$.

Proof. To verify statement (I), it in fact suffices to note that $P((Y, Z) = (0, 1)) = P((Y, Z) = (1, 0))$ by (2.8) and (2.2).

Statement (II) holds by (2.8) and (2.1)-(2.3) and simple arithmetic.

To verify statement (III), use statement (II), eq. (2.8), eqs. (2.1)-(2.3), and simple arithmetic to show that $\text{Var}(Y) = \text{Var}(Z) = (1 - \varepsilon)\varepsilon$ and $\text{Cov}(Y, Z) = (1 - \varepsilon)\varepsilon\theta$. Statement (III) then follows.

Proof of (IV). Suppose V and W are any two random variables such that V (resp. W) is measurable with respect to $\sigma(Y)$ (resp. $\sigma(Z)$). Since Y and Z each take just two values (0, and 1), it is easy to show that there exist real numbers a, b, c, d such that $V = aY + b$ and $W = cZ + d$. One then has by elementary properties of correlation that $\text{Corr}(V, W) = \text{Corr}(Y, Z)$ (resp. 0, resp. $-\text{Corr}(Y, Z)$) if $ac > 0$ (resp. $ac = 0$, resp. $ac < 0$). Then (IV) follows by (III) and eq. (1.3) (and the fact that $\theta > 0$).

Proof of (V)-(VI). If A and B are events, and either one of them has probability 0 or 1, then trivially $P(A \cap B) - P(A)P(B) = 0$. Hence in using (1.1)-(1.2) to calculate the dependence coefficients $\alpha(\sigma(Y), \sigma(Z))$ and $\beta(\sigma(Y), \sigma(Z))$, one can restrict to pairs of events (say A and B) whose probabilities are each strictly between 0 and 1. By statement (II), eq. (2.8), eqs. (2.1)-(2.3), and simple arithmetic, one has that for every ordered pair $(i, j) \in \{0, 1\}^2$,

$$\big|P(\{Y = i\} \cap \{Z = j\}) - P(Y = i)P(Z = j)\big| = (1 - \varepsilon)\varepsilon\theta. \tag{2.9}$$

From (1.1)-(1.2), one has that statements (V) and (VI) both follow immediately from (2.9) and the fact that $1/2 \leq 1 - \varepsilon < 1$. That completes the proof of Lemma 2.2.

8

Lemma 2.3. *Suppose $\varepsilon \in (0, 1/2]$ and $\theta \in (0, 1)$. Suppose V and W are $\{0, 1\}$-valued random variables such that*

$$P(V = 0) = 1 - \varepsilon \quad \text{and} \quad P(V = 1) = \varepsilon. \tag{2.10}$$

and (see (2.5)-(2.6)) for each ordered pair $(i, j) \in \{0, 1\}$,

$$P(W = j | V = i) = p_{ij}^{(\varepsilon, \theta)} \tag{2.11}$$

(the ij-entry of the matrix $\mathbb{P}^{(\varepsilon, \theta)}$).

Then the (joint) probability function of the random vector (V, W) is the matrix $\Lambda^{(\varepsilon, \theta)}$ (see (2.1)-(2.3)); and also in particular,

$$P(W = 0) = 1 - \varepsilon \quad \text{and} \quad P(W = 1) = \varepsilon. \tag{2.12}$$

Proof. For each ordered pair $(i, j) \in \{0, 1\}^2$, one uses (2.10) and (2.11) together to show that (see (2.1)-(2.3)) $P((V, W) = (i, j)) = \lambda_{ij}^{(\varepsilon, \theta)}$. After that, (2.12) follows by trivial arithmetic (or from Lemma 2.2(II)). Thus Lemma 2.3 holds.

Definition 2.4. Suppose $\varepsilon \in (0, 1/2]$ and $\theta \in (0, 1)$. A given random sequence $X := (X_k, k \in \mathbb{Z})$ is is said to satisfy "Condition $\mathcal{S}(\varepsilon, \theta)$" if X is a strictly stationary Markov chain with the following three properties:
   (1) The state space of X is $\{0, 1\}$.
   (2) The (invariant) marginal distribution is given by

$$P(X_0 = 0) = 1 - \varepsilon \quad \text{and} \quad P(X_0 = 1) = \varepsilon. \tag{2.13}$$

   (3) The one-step transition probability matrix for X is $\mathbb{P}^{(\varepsilon, \theta)}$ [see (2.4) and (2.5)-(2.6)].

Condition $\mathcal{S}(\varepsilon, \theta)$ is well defined. By Lemma 2.3, the one-step transition probability matrix $\mathbb{P}^{(\varepsilon, \theta)}$ in (3) is compatible with the (invariant) marginal distribution in (2.13).

Lemma 2.5. *Suppose $\varepsilon \in (0, 1/2]$ and $r \in (0, 1)$. Suppose $X := (X_k, k \in \mathbb{Z})$ is a strictly stationary Markov chain that satisfies Condition $\mathcal{S}(\varepsilon, r)$. Then the following statements hold:*
   (I) The Markov chain X is reversible.
   (II) For each positive integer n, the n-step transition probability matrix for X is $\mathbb{P}(\varepsilon, r^n)$ (see eq. (2.4), eqs. (2.5)-(2.6), and Part 3 of Notations 2.1).
   (III) For each positive integer n, the following holds:

$$\rho_X(n) = r^n; \tag{2.14}$$
$$\alpha_X(n) = (1 - \varepsilon)\varepsilon r^n \geq \varepsilon r^n / 2; \quad \text{and} \tag{2.15}$$
$$\beta_X(n) = 2(1 - \varepsilon)\varepsilon r^n \leq 2\varepsilon r^n. \tag{2.16}$$

Remark. By (III), for any choice of parameters $\varepsilon \in (0, 1/2]$ and $r \in (0, 1)$, the (strictly stationary, two-state, reversible) Markov chain X here in Lemma 2.5 has the property that the dependence coefficients $\rho(n)$, $\alpha(n)$, and $\beta(n)$ all converge to 0 with the same (exponential) mixing rate, modulo constant factors.

Proof of Lemma 2.5. By Definition 2.4, the (strictly stationary) Markov chain X has state space $\{0, 1\}$, with the marginal distribution given by (2.13). Statement (I) holds by Part 5(D) of Setting 1.1.

Proof of (II). By condition (3) in Definition 2.4, one has that for each positive integer n, the n-step transition probability matrix for the Markov chain X is $(\mathbb{P}^{(\varepsilon, r)})^n$ (the nth power with matrix multiplication). By (2.7) and induction, one has that (see Part 3 of Notations 2.1) $(\mathbb{P}^{(\varepsilon, r)})^n = \mathbb{P}(\varepsilon, r^n)$ for every positive integer n. Thus (II) holds.

Proof of (III). By statement (II) and Lemma 2.3, one has that for each positive integer n, the (joint) probability function of the random vector (X_0, X_n) is (again see Part 3 of Notations 2.1) the matrix $\Lambda(\varepsilon, r^n)$. By (1.13) and Lemma 2.2(IV), for each positive integer n, $\rho_X(n) = \rho(\sigma(X_0), \sigma(X_n)) = r^n$. Thus (2.14) holds. Eqs. (2.15)-(2.16) follow similarly from (1.11)-(1.12) and Lemma 2.2(V)(VI). Thus (III) holds. That completes the proof of Lemma 2.5.

We shall return to Definition 2.4 and Lemma 2.5 in Section 3. Here in Section 2, we shall finish with a technical statement that will be employed in Section 3 as part of the process of assembling together a countably infinite collection of strictly stationary, two-state Markov chains (all independent of each other) from Definition 2.4 in order to form the Markov chain for Theorem 1.3.

Lemma 2.6. *Suppose I is a nonempty, countable (index) set, and $\mathcal{A}_i$, $i \in I$ and $\mathcal{B}_i$, $i \in I$ are σ-fields such that the σ-fields $\mathcal{A}_i \vee \mathcal{B}_i$, $i \in I$ are independent. Then*

$$\beta\Big(\bigvee_{i \in I} \mathcal{A}_i, \bigvee_{i \in I} \mathcal{B}_i \Big) \;\leq\; \sum_{i \in I} \beta(\mathcal{A}_i, \mathcal{B}_i); \quad \text{and} \tag{2.17}$$

$$\rho\Big(\bigvee_{i \in I} \mathcal{A}_i, \bigvee_{i \in I} \mathcal{B}_i \Big) \;=\; \sup_{i \in I} \rho(\mathcal{A}_i, \mathcal{B}_i). \tag{2.18}$$

The latter equality (2.18) is due to Csáki and Fischer [5]. Both (2.17) and (2.18) (and under the same assumptions, the inequality (2.17) with β replaced on both sides by α, an inequality that will not be used here) can be found in [[2], v.1, Theorem 6.2]. [Actually, by elementary arguments, all results in that particular theorem, including the ones cited here, extend (pretty frivolously) to the case of an arbitrary (not necessarily countable) nonempty index set; that will not be needed here.]

3   Proof of Theorem 1.3

In this proof, in order to avoid some unnecessary clutter, the use of properties of the function f in Theorem 1.3 (such as the "log convexity" assumption) will sometimes be at "less than full strength", resulting in some crude inequalities and leaving unsaid some natural "extra" or "stronger" observations that will not be needed in the argument.

In this proof, just in an attempt to "make the arithmetic a little nicer", base-2 logarithms will be used, instead of natural logarithms.

The proof will be spelled out here in a series of (mostly) small "steps", identified by consecutive letters ("Step A", "Claim B", "Step C", and so on).

A large portion of this proof here will involve setting up some elementary but tedious "scaffolding" involving lines in connection with the "log convexity" assumption in the theorem. Essentially the same scaffolding, with just minor differences, was used by the author [1] to prove a quite different but somewhat related theorem (involving more mixing conditions, in a non-Markovian, "non-reversible" context). There the scaffolding was spelled out tersely. In the presentation of that theorem and its proof (again) in [[2], v.3, Theorem 26.5], that scaffolding was spelled out in detail, partly in the proof itself and partly in the Appendix at the end of that book. As compared to the context there, the context here will involve some small but nontrivial differences in some information that is developed and used. For the reader's convenience, we shall spell out the scaffolding here in detail, in a self-contained form that is designed to cleanly fit our context here.

Step A. As in the statement of Theorem 1.3, suppose $0 < r \leq 1$, and $f : [0, \infty) \to (0, 1/2]$ is a continuous, strictly decreasing function satisfying assumptions (i), (ii), and (iii) in the statement of that theorem.

Define the continuous, strictly decreasing (and negative) function $g : [0, \infty) \to (-\infty, -1]$ as follows:

$$\text{For each } x \in [0, \infty), \quad g(x) \;:=\; \log_2\big[r^x f(x) \big] \;=\; x(\log_2 r) + \log_2 f(x). \tag{3.1}$$

As a trivial consequence of hypothesis (iii) (which involved the natural logarithm) in Theorem 1.3, the function $x \mapsto \log_2 f(x)$ is convex, and hence the function g in (3.1) is convex.

Claim B. *The following three statements hold:*
(1) $(\log_2 r)x - g(x) \to \infty$ *as* $x \to \infty$.
(2) *For every* $t < \log_2 r$, *one has that* $g(x) - tx \to \infty$ *as* $x \to \infty$.

Proof. To verify (1), note that by (3.1) and hypothesis (i) in Theorem 1.3, $g(x) - (\log_2 r)x = \log_2 f(x) \to -\infty$ as $x \to \infty$. Thus (1) holds.

10

To verify (2), suppose t is any number such that $t < \log_2 r$. Then $2^t < r$ and hence $2^t/r < 1$. Hence by hypothesis (ii) in Theorem 1.3, $(2^t/r)^x/f(x) \to 0$ as $x \to \infty$. That is, $2^{tx}/[r^x f(x)] \to 0$ as $x \to \infty$. Hence by the first equality in (3.1), $tx - g(x) \to -\infty$ as $x \to \infty$. Thus (2) holds. That completes the proof.

Step C. For convenient later reference, we shall just list here again, with substantial redundancy, the properties of the function $g : [0, \infty) \to (-\infty, -1]$ in (3.1) [and the related number $\log^2 r$] that will be used in the proof. In this list, item (0) holds trivially by the hypothesis that $0 < r \leq 1$; items (1), (2), and (4) hold by (0) and (3.1) and the hypothesis of Theorem 1.3; item (3) repeats the sentence after (3.1); items (5) and (6) simply repeat the statements in Claim B; and item (7) holds by (3.1). Here is the list:

(0) $-\infty < \log_2 r \leq 0$.
(1) g is negative; in fact $g(x) \leq -1$ for all $x \in [0, \infty)$.
(2) g is continuous on $[0, \infty)$.
(3) g is convex on on $[0, \infty)$.
(4) g is strictly decreasing on $[0, \infty)$.
(5) $\lim_{x \to \infty}[(\log_2 r)x - g(x)] = \infty$.
(6) For every $t < (\log_2 r)$, $\lim_{x \to \infty}[g(x) - tx] = \infty$.
(7) $2^{g(x)} = r^x f(x)$ for all $x \in [0, \infty)$.

Step D. Suppose v and y are any two real numbers such that $0 \leq v < y$.
Define the real number $\zeta_{v,y}$ by

$$\zeta_{v,y} := \frac{g(y) - g(v)}{y - v}. \tag{3.2}$$

By Step C(4), $\zeta_{v,y} < 0$.
Define the (affine) function $C_{v,y} : [v, y] \to \mathbb{R}$ as follows:

$$\text{For each } x \in [v, y], \quad C_{v,y}(x) := g(v) + \zeta_{v,y}(x - v). \tag{3.3}$$

Then $C_{v,y}(v) = g(v)$ and $C_{v,y}(y) = g(y)$. This function $C_{v,y}$ simply specifies the "chord" in $\mathbb{R}^2$ from the point $(v, g(v))$ to the point $(y, g(y))$. By (3.3), the inequality $\zeta_{v,y} < 0$ above (after (3.2)), and then Step C(3)(4), one has that

$$g(v) \geq C_{v,y}(x) \geq g(x) \geq g(y) \quad \text{for all } x \in [v, y]. \tag{3.4}$$

Define the nonnegative real number $M_{v,y}$ by

$$M_{v,y} := \sup_{x \in [v,y]} [C_{v,y}(x) - g(x)]. \tag{3.5}$$

Claim E. *For any element $v \in [0, \infty)$, there exists $w \in (v, \infty)$ such that $M_{v,w} = 1$.*

Proof. Suppose $v \in [0, \infty)$.
By Step C(5), $[g(v) + (\log_2 r)(x - v) - g(x)] = [(\log_2 r)x - g(x)] - [(\log_2 r)v - g(v)] \to \infty$ as $x \to \infty$. Accordingly, let $z \in (v, \infty)$ be such that $[g(v) + (\log_2 r)(z - v) - g(z)] \geq 3$ (say). Then let $t < \log_2 r$ be such that $[g(v) + t(z - v) - g(z)] \geq 2$.
By Step C(6), $[g(v) + t(x - v) - g(x)] = [tx - g(x)] - [tv - g(v)] \to -\infty$ as $x \to \infty$. Applying that and Step C(2) and the last sentence of the preceding paragraph, let $y \in (z, \infty)$ be such that $g(v) + t(y - v) - g(y) = 0$. Then $g(v) + t(y - v) = g(y)$.
From above, one has that $v < z < y$. By (3.2) and the last sentence of the preceding paragraph just above, $\zeta_{v,y} = t$; and hence for each $x \in [v, y]$, $C_{v,y}(x) = g(v) + t(x - v)$ by (3.3). By the preceding two sentences, followed by the last sentence of the next-to-last paragraph above, $C_{v,y}(z) = g(v) + t(z - v) \geq 2 + g(z)$. Hence by (3.5), $M_{v,y} \geq 2$.
Trivially by (3.5), (3.4), and Step C(2), $\lim_{u \to v+} M_{v,u} = 0$. Also, as a consequence of Step C(2)(3), the mapping $u \mapsto M_{v,u}$, for $u \in (v, \infty)$, is (nondecreasing and) continuous. By the preceding two sentences here and the last sentence of the preceding paragraph, there exists $w \in (v, y)$ such that $M_{v,w} = 1$. Thus Claim E holds.

Recursion F. We shall recursively define an infinite sequence $(y_0, y_1, y_2, \dots)$ of elements of $[0, \infty)$, as follows:

To start off, define $y_0 := 0$.

Now suppose n is a nonnegative integer, and y_n has already been defined in $[0, \infty)$. Applying Claim E, and writing y_j also as $y(j)$ for typographical convenience, let y_{n+1} be such that

$$y_{n+1} > y_n \quad \text{and} \quad M_{y(n),y(n+1)} = 1. \tag{3.6}$$

That completes Recursion F.

Claim G. $y_n \to \infty$ *as* $n \to \infty$.

Proof. Suppose Claim G is false. Then by the "first half" of (3.6), there is a positive number z such that $y_n \uparrow z$ as $n \to \infty$. By Step C(2), $g(y_n) \to g(z)$ as $n \to \infty$. Hence $g(y_{n-1}) - g(y_n) \to 0$ as $n \to \infty$. Hence by (3.4) and (3.5), $M_{y(n-1),y(n)} \to 0$ as $n \to \infty$. But that contradicts (3.6). Thus Claim G holds after all.

Step H. In this "step", we shall simplify some notations, "extend certain key chords to full lines", and define some useful reference points $w_n \in (y_{n-1}, y_n)$ for $n \in \mathbb{N}$.

Refer to the (3.6) (its "first half"), to (3.2), and to (3.3). For each positive integer n, first define the real number $s_n := \zeta_{y(n-1),y(n)}$ [and keep in mind that $s_n < 0$ by the comment after (3.2)], then define the (affine) function $L_n : \mathbb{R} \to \mathbb{R}$ by

$$L_n(x) := g(y_{n-1}) + s_n \cdot (x - y_{n-1}) \quad \text{for} \quad x \in \mathbb{R}, \tag{3.7}$$

and then from (3.7) define the real number $a_n := L_n(0)$. From all three of those definitions and some basic algebra, along with (3.3) and the sentence right after it, one has that for each positive integer n,

$$L_n(x) = a_n + s_n x \quad \text{for all} \quad x \in \mathbb{R}, \tag{3.8}$$
$$L_n(x) = C_{y(n-1),y(n)}(x) \quad \text{for all} \quad x \in [y_{n-1}, y_n], \tag{3.9}$$

and (hence) in particular, $L_n(y_{n-1}) = g(y_{n-1})$ and $L_n(y_n) = g(y_n)$.

From (3.9), (3.5), and (3.6), one has that for each positive integer n,

$$\sup_{x \in [y(n-1),y(n)]} [L_n(x) - g(x)] = M_{y(n-1),y(n)} = 1. \tag{3.10}$$

Of course for any given $n \in \mathbb{N}$, by Step C(2) and (say) (3.8), the expression in the main brackets in the left side of (3.10) is a continuous function of x on (at least) the closed interval $[y_{n-1}, y_n]$. Accordingly, for each $n \in \mathbb{N}$, let $w_n \in [y_{n-1}, y_n]$ be such that $L_n(w_n) - g(w_n) = 1$. For any given $n \in \mathbb{N}$, by the two equations right after (3.9), one in fact has that $y_{n-1} < w_n < y_n$.

Step I. Here we shall display for convenient reference seven technical facts that will be employed in (the rest of) the proof of Theorem 1.3.

Here are the first two:

$$0 = y_0 < w_1 < y_1 < w_2 < y_2 < w_3 < y_3 < \ldots; \quad \text{and} \tag{3.11}$$
$$y_n \to \infty \quad \text{as} \quad n \to \infty. \tag{3.12}$$

Here (3.11) comes from Recursion F and (for every positive integer n) the final pair of inequalities in Step H. Eq. (3.12) is simply Claim G.

Next, for each positive integer n,

$$L_n(y_{n-1}) = g(y_{n-1}); \tag{3.13}$$
$$L_n(w_n) = g(w_n) + 1; \quad \text{and} \tag{3.14}$$
$$L_n(y_n) = g(y_n). \tag{3.15}$$

Eqs. (3.13) and (3.15) were pointed out right after (3.9), and eq. (3.14) comes from the next to last sentence in Step H. Also, for each positive integer n,

$$g(x) \le L_n(x) \le g(x) + 1 \quad \text{for all} \quad x \in [y_{n-1}, y_n]. \tag{3.16}$$

12

Here the first inequality comes from (3.11), (3.13), (3.15), and Step C(3) (convexity of g); and the second inequality comes from (3.10).

Now in fact (3.11), (3.13), (3.14), (3.15), and Step C(3) (again, convexity of g) yield that for each positive integer n, one has that $g(x) < L_n(x)$ for all $x \in (y_{n-1}, y_n)$, and that, again for each $n \in \mathbb{N}$,

$$L_n(x) < g(x) \quad \text{for all} \ \ x \in [0, \infty) - [y_{n-1}, y_n]. \tag{3.17}$$

Remark J. For any given positive integer n, the following comments (1)-(7) hold:

(1) By (3.11), $0 \le y_{n-1} < w_n < y_n < w_{n+1} < y_{n+1}$; and hence $w_n \in [0, \infty) - [y_n, y_{n+1}]$ and $w_{n+1} \in [0, \infty) - [y_{n-1}, y_n]$.

(2) By (3.13) and (3.15), $L_{n+1}(y_n) - L_n(y_n) = g(y_n) - g(y_n) = 0$.

(3) By (1) above and (3.17), and then (3.14), $L_{n+1}(w_n) - L_n(w_n) < g(w_n) - L_n(w_n) = -1$.

(4) By (1), (2), and (3) above, the function $x \mapsto L_{n+1}(x) - L_n(x)$ for $x \in \mathbb{R}$ [with constant derivative $s_{n+1} - s_n$ — recall (3.8)] is strictly increasing (and in particular, $s_{n+1} - s_n > 0$).

(5) By (1), (3), and (4) above, $L_{n+1}(x) - L_n(x) < -1$ for all $x \le y_{n-1}$ (in fact for all $x \le w_n$).

(6) By (1) above and (3.17), and then (3.14), $L_{n+1}(w_{n+1}) - L_n(w_{n+1}) > L_{n+1}(w_{n+1}) - g(w_{n+1}) = 1$.

(7) By (1), (6), and (4) above, $L_{n+1}(x) - L_n(x) > 1$ for all $x \ge y_{n+1}$ (in fact for all $x \ge w_{n+1}$).

Remark K. For any two given positive integers h and j such that $h < j$, the following comments (1)-(5) hold:

(1) For each $x \in \mathbb{R}$, one has the telescoping sum $L_j(x) - L_h(x) = \sum_{i=h}^{i=j-1} [L_{i+1}(x) - L_i(x)]$.

(2) For any $x \ge y_j$ and any $i \in \{h, h+1, \ldots, j-1\}$, one has that $x \ge y_{i+1}$ by (3.11) and hence $L_{i+1}(x) - L_i(x) > 1$ by Remark J(7).

(3) By (1) and (2) above, $L_j(x) - L_h(x) > j - h$ for all $x \ge y_j$.

(4) For any $x \le y_{h-1}$ and any $i \in \{h, h+1, \ldots, j-1\}$, one has that $x \le y_{i-1}$ by (3.11) and hence $L_{i+1}(x) - L_i(x) < -1$ by Remark J(5).

(5) By (1) and (4) above, $L_j(x) - L_h(x) < -(j - h)$ for all $x \le y_{h-1}$.

Remark L. By (3.8), (3.11), (3.13). and Step C(1),

$$a_1 = L_1(0) = L_1(y_0) = g(y_0) = g(0) \le -1. \tag{3.18}$$

For any integer $j \ge 2$, by (3.8) and Remark K(5) (with $h = 1$ and $x = 0$), $a_j - a_1 = L_j(0) - L_1(0) < -(j-1)$, and hence by (3.18), $a_j < -j$. Combining that with (3.18) itself, one has that

$$a_n \le -n \quad \text{for every} \ \ n \in \mathbb{N}. \tag{3.19}$$

Remark M. For any given positive integer n, the following comments (1)-(6) hold:

(1) By (3.16) and (3.17), $L_n(x) \le g(x) + 1$ for all $x \in [0, \infty)$.

(2) By (3.17) (for $0 \le x < y_n$) and (3.13) (for $x = y_n$), $L_{n+1}(x) \le g(x)$ for all $x \in [0, y_n]$.

(3) By (2) above and Remark K(5), for any integer $j \ge n + 2$ and any $x \in [0, y_n]$, one has that $L_j(x) = L_{n+1}(x) + [L_j(x) - L_{n+1}(x)] \le g(x) - [j - (n+1)]$.

(4) If $n \ge 2$, then by (3.17) (for $x > y_{n-1}$) and (3.15) (for $x = y_{n-1}$), $L_{n-1}(x) \le g(x)$ for all $x \ge y_{n-1}$.

(5) If $n \ge 3$, then by (4) above and Remark K(3), for any positive integer $h \le n - 2$ and any $x \ge y_{n-1}$, $L_h(x) = L_{n-1}(x) - [L_{n-1}(x) - L_h(x)] \le g(x) - [(n-1) - h]$.

(6) For the given $n \in \mathbb{N}$, comments (1)-(5) apply to all $x \in [y_{n-1}, y_n]$, and [see also (3.8)] can be repeated together in one convenient display, as follows:

$$\text{For all } x \in [y_{n-1}, y_n] \text{ and all } j \in \mathbb{N}, \quad a_j + s_j x \ = \ L_j(x) \ \le \ g(x) + 1 - |n - j|. \tag{3.20}$$

Claim N below will involve sums whose summands are respectively 2 to the power $L_j(x)$ for $j \in \mathbb{N}$. The left side of (3.20) was included because it may make slightly less cumbersome the notations in both the display in Claim N itself and the subsequent application of Claim N later on.

Claim N. *Writing the coefficients a_n and s_n in (3.8) also as $a(n)$ and $s(n)$ respectively, one has the following: For every $x \in [0, \infty)$,*

$$\sum_{j=1}^{\infty} \left[2^{a(j)} \cdot (2^{s(j)})^x \right] \ \le \ 6 \cdot r^x f(x). \tag{3.21}$$

13

Proof. Suppose $x \in [0, \infty)$. Applying (3.11) and (3.12), let $n \in \mathbb{N}$ be such that $x \in [y_{n-1}, y_n]$. By Remark M(6), for every positive integer j, $a_j + s_j x \leq g(x) + 1 - |n - j|$. Hence by Step C(7),

$$\sum_{j=1}^{\infty} \left[2^{a(j)} \cdot (2^{s(j)})^x \right] \leq \sum_{j=1}^{\infty} 2^{g(x)+1-|n-j|} = 2^{g(x)+1} \sum_{j=1}^{\infty} 2^{-|n-j|}$$

$$\leq 2 \cdot r^x f(x) \cdot \sum_{j=-\infty}^{\infty} 2^{-|n-j|} = 2 \cdot r^x f(x) \cdot 3.$$

Thus (3.21) holds. That completes the proof.

Step O. For each positive integer n, one has that $s_n < 0$ as was noted right before (3.7), and $s_{n+1} - s_n > 0$ as was noted at the end of Remark J(4). Thus

$$s_1 < s_2 < s_3 < \cdots < 0. \tag{3.22}$$

Define the number $\eta \in (-\infty, 0]$ by

$$\eta := \sup_{n \in \mathbb{N}} s_n = \lim_{n \to \infty} s_n. \tag{3.23}$$

Our next task is to establish the value of this number η.

Recall from Step C(5) that $g(x) - (\log_2 r)x \to -\infty$ as $x \to \infty$. It follows that for any given $n \in \mathbb{N}$, by (3.8) and (3.17), $a_n + s_n x - (\log_2 r)x = L_n(x) - (\log_2 r)x \to -\infty$ as $x \to \infty$, hence $(s_n - \log_2 r)x \to -\infty$ as $x \to \infty$, hence $s_n - \log_2 r < 0$ must hold, that is, $s_n < \log_2 r$. Since that holds for all $n \in \mathbb{N}$, it follows from (3.23) that $\eta \leq \log_2 r$.

Next, for any given $n \in \mathbb{N}$ and any given $x \in [y_{n-1}, y_n]$, one has that $[x \geq 0$ by (3.11) and$]$ $\eta x \geq s_n x > a_n + s_n x = L_n(x) \geq g(x)$ by (3.23), (3.19), (3.8), and (3.16). Hence $\eta x > g(x)$ for all $x \in [0, \infty)$ by (3.11) and (3.12). Hence $\eta \geq \log_2 r$ [for otherwise a contradiction would result from Step C(6)].

By the final sentence of each of the preceding two paragraphs, $\eta = \log_2 r$. Hence by (3.23),

$$\log_2 r = \sup_{n \in \mathbb{N}} s_n = \lim_{n \to \infty} s_n. \tag{3.24}$$

Step P. Now the Markov chains in Definition 2.4 will be brought into play.

(1) First recall from (3.19) and (3.22) that for every positive integer j, one has that $a_j \leq -j$ and $s_j < 0$. Thus

$$\text{for every } j \in \mathbb{N}, \quad 0 < 2^{a(j)} \leq 2^{-j} \leq 1/2 \quad \text{and} \quad 0 < 2^{s(j)} < 1 \tag{3.25}$$

where a_j and s_j are (again) written also as $a(j)$ and $s(j)$ for typographical convenience.

(2) Now for each positive integer j, referring to (3.25) and Definition 2.4, let $W^{(j)} := (W_k^{(j)}, k \in \mathbb{Z})$ be a strictly stationary Markov chain that satisfies Condition $\mathcal{S}(2^{a(j)}, 2^{s(j)})$. Let these Markov chains $W^{(1)}, W^{(2)}, W^{(3)}, \ldots$ be constructed in such a way that they are independent of each other.

(3) By Definition 2.4(I)(II), for any given positive integer j and any given $k \in \mathbb{Z}$, the random variable $W_k^{(j)}$ takes only the values 0 and 1, with probabilities

$$P(W_k^{(j)} = 0) = 1 - 2^{a(j)} \quad \text{and} \quad P(W_k^{(j)} = 1) = 2^{a(j)}. \tag{3.26}$$

(4) By (3.25), $\sum_{j=1}^{\infty} 2^{a(j)} < \infty$. Hence by (3.26) and the Borel-Cantelli Lemma, for any given $k \in \mathbb{Z}$,

$$P(W_k^{(j)} = 1 \text{ for infinitely many j} \in \mathbb{N}) = 0. \tag{3.27}$$

Just for technical convenience, deleting a set of probability 0 from the probability space Ω if necessary, we assume without loss of generality that for each integer k, the event in the left side of (3.27) is the empty set.

(5) In what follows, here (and in Step R below), when the notation $W^{(j)}$ appears as a subscript, it will be written as $W(j)$ for typographical convenience.

By Lemma 2.5, for each positive integer j, the following statements hold:

$$\text{The Markov chain } W^{(j)} \text{ is reversible.} \tag{3.28}$$

$$\text{For each } m \in \mathbb{N}, \quad \rho_{W(j)}(m) = (2^{s(j)})^m. \tag{3.29}$$

$$\text{For each } m \in \mathbb{N}, \quad \alpha_{W(j)}(m) \geq (1/2) \cdot (2^{a(j)}) \cdot (2^{s(j)})^m. \tag{3.30}$$

$$\text{For each } m \in \mathbb{N}, \quad \beta_{W(j)}(m) \leq 2 \cdot (2^{a(j)}) \cdot (2^{s(j)})^m. \tag{3.31}$$

Step Q. Let $X := (X_k, k \in \mathbb{Z})$ be the sequence of random variables defined as follows:

$$\text{For each } k \in \mathbb{Z}, \quad X_k := \sum_{j=1}^{\infty} 2^{j-1} W_k^{(j)}. \tag{3.32}$$

By Step P(3)(4), (recall the sentence after (3.27)), for each $k \in \mathbb{Z}$, the sum in (3.32) has, at any given sample point $\omega \in \Omega$, at most finitely many non-zero terms (and therefore trivially converges).

By Step P(3)(4), and (3.32), the random variables $X_k, k \in \mathbb{Z}$ take their values in the (countably infinite) set $\{0, 1, 2, \dots\}$ of all nonnegative integers. In fact (3.32) involves a standard one-to-one correspondence between that set $\{0, 1, 2, \dots\}$ and the set of all sequences $(z_1, z_2, z_3, \dots)$ of elements of $\{0, 1\}$ with at most finitely many 1's. As a consequence,

$$\text{for each } k \in \mathbb{Z}, \quad \sigma(X_k) = \sigma(W_k^{(j)}, j \in \mathbb{N}). \tag{3.33}$$

Further, by (3.25) and (3.26), for any given $k \in \mathbb{Z}$ and any given sequence $(z_1, z_2, z_3, \dots) \in \{0, 1\}^{\mathbb{N}}$ with at most finitely many 1's, the numbers $P(W_k^{(j)} = z_j)$, $j \in \mathbb{N}$ are each an element of the open interval $(0, 1)$, and they satisfy $\sum_{j=1}^{\infty}[1 - P(W_k^{(j)} = z_j)] < \infty$, and hence

$$P\left(\bigcap_{j=1}^{\infty} \{W_k^{(j)} = z_j\} \right) = \prod_{j=1}^{\infty} P(W_k^{(j)} = z_j) \; > \; 0,$$

where the equality comes from the second sentence of Step P(2). It follows from (3.32) that for any given $k \in \mathbb{Z}$, and any given $m \in \{0, 1, 2, \dots\}$, $P(X_k = m) > 0$.

With elementary (if tedious) arguments, using (3.32), (3.33), Step P(2), and (3.28), one can verify that the random sequence X is strictly stationary, that X is a Markov chain (with state space $\{0, 1, 2, \dots\}$), and that X is reversible.

Step R. Recall the last paragraph of Step P and the last sentence of Step Q. To complete the proof of Theorem 1.3, what remains is to verify (1.18).

Let m be an arbitrary fixed positive integer. For convenient reference, taking (1.9) into account, what remains to prove are the following three items:
(1) $\rho_X(m) = r^m$;
(2) $\alpha_X(m) \geq (1/2) \cdot r^m f(m)$; and
(3) $\beta_X(m) \leq 12 \cdot r^m f(m)$.
In the proofs of (1)-(3) given below, when the notation $W^{(j)}$ is itself a subscript, it will be written as $W(j)$ for typographical convenience.

Proof of (1). By eq. (3.33), Step P(2), Lemma 2.6, eq. (3.29), and eq. (3.24), one has that

$$\rho_X(m) \;=\; \sup_{j \in \mathbb{N}} \rho_{W(j)}(m) \;=\; \sup_{j \in \mathbb{N}} (2^{s(j)})^m \;=\; r^m.$$

Thus (1) holds.

Proof of (2). Applying (3.11) and (3.12), let $j \in \mathbb{N}$ be such that $m \in [y_{j-1}, y_j]$. Then by (3.8) and (3.16), $a_j + s_j m = L_j(m) \geq g(m)$. Taking 2 to the powers on the left and right sides and applying Step C(7), one obtains that $2^{a(j)} \cdot (2^{s(j)})^m \geq 2^{g(m)} = r^m f(m)$ (where again a_j and s_j are written as $a(j)$ and $s(j)$). Hence by (3.33) and (3.30), one has that $\alpha_X(m) \geq \alpha_{W(j)}(m) \geq (1/2) \cdot r^m f(m)$. Thus (2) holds.

Proof of (3). By eq. (3.33), Step P(2), Lemma 2.6, eq. (3.31), and finally Claim N, one has that

$$\beta_X(m) \ \leq \ \sum_{j=1}^{\infty} \beta_{W(j)}(m) \ \leq \ \sum_{j=1}^{\infty} [2 \cdot 2^{a(j)} \cdot (2^{s(j)})^m] \ \leq \ 12 \cdot r^m f(m).$$

Thus (3) holds. That completes the proof of Theorem 1.3.

References

[1] R.C. Bradley, *Identical Mixing Rates*, Probab. Theory Related Fields **74** (1987), 497-503.

[2] R.C. Bradley, *Introduction to Strong Mixing Conditions* (Vols. 1, 2, and 3), Kendrick Press, Heber City (Utah), 2007.

[3] R.C. Bradley, *On mixing properties of reversible Markov chains*, New Zealand J. Math. **45** (2015), 71-87.

[4] R.C. Bradley, *On some basic features of strictly stationary, reversible Markov chains*, J. Time Series Anal. **42** (2021), 499-533.

[5] P. Csáki and J. Fischer, *On the general notion of maximal correlation*, Magyar Tud. Akad. Mat. Kutato Int. Kozl. **8** (1963), 27-51.

[6] Yu.A. Davydov, *Mixing conditions for Markov chains*, Theory Probab. Appl. **18** (1973), 312-328.

[7] P. Doukhan, P. Massart, and E. Rio, *The functional central limit theorem for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), 63-82.

[8] H.O. Hirschfeld, *A connection between correlation and contingency*, Proc. Cambridge Phil. Soc. **31** (1935), 520-524.

[9] H. Kesten and G.L. O'Brien, *Examples of mixing sequences*, Duke Math. J. **43** (1976), 405-415.

[10] A.N. Kolmogorov and Yu.A. Rozanov, *On strong mixing conditions for stationary Gaussian processes*, Theory Probab. Appl. **5** (1960), 204-208.

[11] I. Kontoyiannis and S.P. Meyn, *Geometric ergodicity and the spectral gap of non-reversible Markov chains*, Probab. Theory Related Fields **154** (2012), 327-339.

[12] M. Longla, *Remarks on the speed of convergence of mixing coefficients and applications*, Statist. Probab. Lett. **83** (2013), 2439-2445.

[13] M. Longla, *On dependence structure of copula-based Markov chains*, ESIAM: Probab. Statist. **18** (2014), 570-583.

[14] M. Longla and M. Peligrad, *Some aspects of modeling dependence in copula-based Markov chains*, J. Multivariate Anal. **111** (2012), 234-240.

[15] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability* (2nd edn.), Cambridge University Press, Cambridge, 2009.

[16] R.B. Nelsen, *An introduction to copulas* (2nd edn.), Springer, New York, 2006.

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

[18] E. Nummelin and R.L. Tweedie, *Geometric ergodicity and R-positivity for general Markov chains*, Ann. Probab. **6** (1978), 404-420.

[19] G.O. Roberts and J.S. Rosenthal, *Geometric ergodicity and hybrid Markov chains*, Electronic Commun. Probab. **2** (1997), 13-25.

[20] G.O. Roberts and R.L. Tweedie, *Geometric L^2 and L^1 convergence are equivalent for reversible Markov chains*, J. Appl. Probab. **38A** (2001), 37-41.

[21] M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Natl. Acad. Sci. USA **42** (1956), 43–47.

[22] M. Rosenblatt, *Markov Processes. Structure and Asymptotic Behavior,* Springer, Berlin, 1971.

[23] A.N. Tikhomirov, *On the convergence rate in the central limit theorem for weakly dependent random variables*, Theory Probab. Appl. **25** (1980), 790-809.

[24] V.A. Volkonskii and Yu.A. Rozanov, *Some limit theorems for random functions I*, Theory Probab. Appl. **4** (1959), 178-197.