

ON A “REPLICATING CHARACTER STRING” MODEL

Richard C. Bradley
Department of Mathematics
Indiana University
Bloomington
Indiana 47405
USA

E-mail address: bradleyr@indiana.edu

Abstract. In a paper of Chaudhuri and Dasgupta published in 2006, a certain stochastic model for “replicating character strings” (such as in DNA sequences) was studied. In their model, a random “input” sequence was subjected to random mutations, insertions, and deletions, resulting in a random “output” sequence. In this note, their model will be set up in a slightly different way, in an effort to facilitate further development of the theory for their model. In their 2006 paper, Chaudhuri and Dasgupta showed that under certain conditions, strict stationarity of the “input” sequence would be preserved by the “output” sequence, and they proved a similar “preservation” result for the property of strong mixing with exponential mixing rate. In our setup, we shall in spirit slightly extend their “preservation of stationarity” result, and also prove a “preservation” result for the property of absolute regularity with summable mixing rate.

AMS 2010 Mathematics Subject Classification: 60G10

Key words and phrases: Replicating character string, stationarity, absolute regularity

1. Introduction

Chaudhuri and Dasgupta [6] formulated and studied a certain stochastic model for “replicating character strings”. In that paper, they cited numerous other references where other related models had been studied, and in particular, they cited the book by Waterman [17] for the possible application of central limit theory under strong mixing conditions in the use of such models for the statistical analysis of data from biology (e.g. involving DNA sequences). In this note, we shall contribute further results and techniques to the theory for the particular model in [6], and suggest a way of setting up their model that may allow slightly easier handling of certain technical details.

Let \mathbf{N} (resp. \mathbf{Z}) denote the set of all positive integers (resp. the set of all integers).

The model studied by Chaudhuri and Dasgupta [6] can be briefly described as follows: It starts with an “input” sequence $X := (X_k, k \in \mathbf{N})$ of random variables taking their values in some finite “alphabet” — for example the set $\{A, C, G, T\}$ of letters that represent the nucleotides in a DNA sequence. There is another sequence $Z := (Z_k, k \in \mathbf{N})$ of random variables taking their values in the set $\{M, I, D\}$ — to indicate that at a given “time” (or “location”) k , there should be a “mutation” (M), “insertion” (I), or “deletion” (D). (This sequence Z is informally referred to below as the “MID-sequence”.) Probabilities are assigned for what letter of the “alphabet” is inserted when an insertion occurs, or what letter of the alphabet results from a mutation. (The — perhaps high — probability of “no mutation” is formally represented in this scheme as the probability of “replacing a letter by itself” when a “mutation” occurs.) At the end, the result is an “output” sequence $Y := (Y_k, k \in \mathbf{N})$ of random variables, with the same “alphabet” (e.g. $\{A, C, G, T\}$) as the “input” sequence X .

In their paper, Chaudhuri and Dasgupta [6, Theorems 3.1 and 3.2] established certain conditions under which certain properties of the “input” sequence — specifically, strict stationarity, and strong mixing with exponential mixing rate — would be retained by the “output” sequence. Chaudhuri and Dasgupta [6] set up their model using (“one-sided”) random sequences indexed by \mathbf{N} , as described above. In Section 3, we shall set up their model again, but using (“two-sided”) random sequences indexed by \mathbf{Z} . This will hopefully make it a little easier to handle various technical details, such as keeping track of relevant σ -fields when estimating mixing rates.

In the statements of their main results (though not in the initial formulation of their model), Chaudhuri and Dasgupta [6] dealt with the case where the MID-sequence is an (irreducible, aperiodic) Markov chain that is independent of the “input” sequence. We shall retain that “independence” assumption, but allow the MID-sequence itself to satisfy a somewhat more flexible dependence assumption than a “Markov” property.

Instead of studying the mixing rates for the “input” and “output” sequences for the strong mixing condition, we shall do so for the absolute regularity condition, which is stronger than strong mixing. That will provide an opportunity to illustrate the use of a particularly handy “coupling” property (due to Berbee [1]) that is possessed by the absolute regularity condition but not by the strong mixing condition. However, along the way, we shall also give information that may be relevant to the further development of the theory for this model under the strong mixing condition.

Instead of studying the case of exponential mixing rates (for strong mixing) as in [6], we shall focus on a certain slower (“summable”) mixing rate (for absolute regularity) that is natural in central limit theory for bounded random variables (under either strong mixing or absolute regularity).

In the model in [6], one somewhat tricky facet of keeping track of relevant σ -fields was keeping track of the changes in the “clock” resulting from “deletions”. We shall adopt an alternative technical procedure — switching to a new probability measure based on conditioning on a certain event — in the hope of slightly simplifying that task.

In their model, Chaudhuri and Dasgupta [6] assumed a finite state space (for the “input” and “output” sequences), as described above. That is the case of primary interest; but for convenience, we shall relax that assumption and allow the “input” and “output” sequences to consist in essence of real-valued random variables. We shall actually treat those random variables as taking their values in $(0, \infty)$ (think of “coding” a real number x by the positive number e^x), and in an intermediate stage reserve the value 0 as a temporary “place holder” where an “insertion” will ultimately occur.

The model in [6] directly involved probability mass functions for what happens when a “mutation” or “insertion” occurs. As a measure-theoretic convenience, we shall handle that in a slightly different way, using independent random variables uniformly distributed on the unit interval as “randomizers”.

In making these modifications, we shall not change the actual model studied by Chaudhuri and Dasgupta [6] in any significant way. The modifications here only involve how their model is set up. Our quest here is in part to facilitate further development of the theory for their model. There is of course the practical question, not addressed here, of to what extent inaccuracy may occur when, say, a “long but finite” DNA sequence is modeled as a “two-sided” random sequence.

In Section 2, some preliminary information on both the strong mixing and absolute regularity conditions will be given. In Section 3, the model in [6] will be spelled out with the modifications in the setup described above. Then in Section 4, the main result of this note will be stated and proved.

2. Preliminary information on two mixing conditions

In the development of the material in Sections 3 and 4 below, we shall start with a probability space, and then switch to a new probability measure (on the same measurable space) obtained by conditioning on a certain key event. Accordingly, in the notations in definitions below, the relevant probability measure will be specified explicitly. If only one probability measure, say P , is specified, then the notation $E(\dots)$ will be tacitly understood to mean the expected value with respect to that particular probability measure P .

Suppose (Ω, \mathcal{F}) is a measurable space. Suppose $W := (W_i, i \in I)$ is a random variable/vector or stochastic process indexed by a nonempty set I — that is, $W : \Omega \rightarrow \mathbf{R}^I$ is a function which is measurable with respect to the σ -field \mathcal{F} on Ω and the Borel σ -field on \mathbf{R}^I . The σ -field ($\subset \mathcal{F}$) of subsets of Ω generated by W will be denoted $\sigma(W)$ or $\sigma(W_i, i \in I)$.

Definition 2.1. Suppose (Ω, \mathcal{F}) is a measurable space, and P is a probability measure

on (Ω, \mathcal{F}) .

For any two σ -fields \mathcal{A} and $\mathcal{B} \subset \mathcal{F}$, define the following two measures of dependence:

$$\alpha(\mathcal{A}, \mathcal{B}; P) := \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A)P(B)|; \quad \text{and} \quad (2.1)$$

$$\beta(\mathcal{A}, \mathcal{B}; P) := \sup \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J |P(A_i \cap B_j) - P(A_i)P(B_j)| \quad (2.2)$$

where in (2.2) the supremum is taken over all pairs of partitions $\{A_1, A_2, \dots, A_I\}$ and $\{B_1, B_2, \dots, B_J\}$ of Ω such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j . It is easy to see that for any two σ -fields \mathcal{A} and \mathcal{B} , one has that

$$\alpha(\mathcal{A}, \mathcal{B}; P) \leq \beta(\mathcal{A}, \mathcal{B}; P). \quad (2.3)$$

Suppose $X := (X_k, k \in \mathbf{Z})$ is, with respect to P , a strictly stationary sequence of random variables. For each $n \in \mathbf{N}$ define the dependence coefficients

$$\alpha(X, n; P) := \alpha(\sigma(X_k, k \leq 0), \sigma(X_k, k \geq n); P); \quad \text{and} \quad (2.4)$$

$$\beta(X, n; P) := \beta(\sigma(X_k, k \leq 0), \sigma(X_k, k \geq n); P). \quad (2.5)$$

One trivially has that each of the sequences of numbers $(\alpha(X, n; P), n \in \mathbf{N})$ and $(\beta(X, n; P), n \in \mathbf{N})$ is nonincreasing. Also, by (2.3), $\alpha(X, n; P) \leq \beta(X, n; P)$ for every positive integer n . The sequence X is (with respect to the probability measure P) “strongly mixing” [14] if $\alpha(X, n; P) \rightarrow 0$ as $n \rightarrow \infty$, and “absolute regular” [16] if $\beta(X, n; P) \rightarrow 0$ as $n \rightarrow \infty$. By (2.3), absolute regularity implies strong mixing.

To motivate the results later in this paper, we shall state a classic theorem of Ibragimov, from Ibragimov and Linnik [10, Theorem 18.5.4].

Theorem 2.1 (Ibragimov). *Suppose that on a probability space (Ω, \mathcal{F}, P) , $X := (X_k, k \in \mathbf{Z})$ is a strictly stationary sequence of bounded, centered random variables such that $\sum_{n=1}^{\infty} \alpha(X, n; P) < \infty$.*

Then $\sigma^2 := EX_0^2 + 2 \cdot \sum_{n=1}^{\infty} EX_0 X_n$ exists in $[0, \infty)$, with this sum being absolutely convergent. If further $\sigma^2 > 0$, then $(X_1 + X_2 + \dots + X_n)/(n^{1/2}\sigma)$ converges in distribution to the $N(0, 1)$ law as $n \rightarrow \infty$.

That theorem will not be used anywhere in what follows, but it will provide the motivation for the mathematical development in this paper. For example, in a statistical analysis of DNA data, one might deal with indicator functions ($\{0, 1\}$ -valued random variables) marking the locations of a particular pattern of nucleotides along a DNA sequence. Thus if strong mixing is assumed as part of the statistical model, then it might be natural to apply a central limit theorem for *bounded* strongly mixing sequences of random variables, such as Theorem 2.1. Now the (summable) mixing rate in Theorem 2.1 is practically sharp. That was shown to be true even under absolute regularity, by counterexamples of Davydov [7, Example 2] and the author [4]. The counterexample in the latter paper is

a strictly stationary, 3-state sequence that satisfies absolute regularity with (“not quite summable”) mixing rate $\beta(X, n; P) = O(1/n)$. Theorem 2.1 seems to be a natural one to use when either strong mixing or absolute regularity is assumed in the modeling of DNA sequences; and it is the summable mixing rate in that theorem that we shall focus on in this note.

It is worth noting that Merlevède and Peligrad [11] proved (as a special case of the main result in their paper) a modified, refined version of Theorem 2.1 with the barely slower mixing rate $\alpha(X, n; P) = o(1/n)$ and an explicit extra assumption on the rate of growth of the variances of partial sums.

As was mentioned in Section 1, instead of dealing with the strong mixing condition, we shall deal with absolute regularity. That will provide an opportunity to illustrate the use — in Steps 5 and 6 of the proof of Lemma 4.4 in Section 4 — of a handy “coupling” property (from [1]) of the absolute regularity condition. That property does not exist, at least in as strong a form, under just strong mixing. The next three lemmas will facilitate that particular application of that coupling property.

Lemma 2.1. *Suppose (Ω, \mathcal{F}, P) is a probability space, N is a positive integer, \mathcal{A}_n and \mathcal{B}_n , $n \in \{1, 2, \dots, N\}$ are σ -fields $\subset \mathcal{F}$, and the σ -fields $\mathcal{A}_n \vee \mathcal{B}_n$, $n \in \{1, 2, \dots, N\}$ are independent (under P). Then*

$$\beta\left(\bigvee_{n=1}^N \mathcal{A}_n, \bigvee_{n=1}^N \mathcal{B}_n; P\right) \leq \sum_{n=1}^N \beta(\mathcal{A}_n, \mathcal{B}_n; P).$$

In one form or another, this has long been part of the folklore; see e.g. Pinsker [12, p. 73]. One reference for the particular formulation here is [5, v1, Theorem 6.2]. (That reference also gives the exactly analogous inequality for the dependence coefficient $\alpha(\cdot, \cdot)$.)

The next lemma has long been part of the folklore; but a reference for it seems hard to find. In this lemma, the random variables X and Y are not assumed to be identically distributed. In this lemma, the term “Borel space” means a measurable space (S, \mathcal{S}) that is bimeasurably isomorphic to the space $(\mathbf{R}, \mathcal{R})$ where \mathcal{R} denotes the Borel σ -field on \mathbf{R} . It is well known that $\mathbf{R}^{\mathbf{N}}$ (or $\mathbf{R}^{\mathbf{Z}}$), accompanied by its Borel σ -field, is a Borel space.

Lemma 2.2. *Suppose (S, \mathcal{S}) is a Borel space. Suppose (Ω, \mathcal{F}, P) is a probability space, X and Y are random variables on this space which take their values in (S, \mathcal{S}) , and \mathcal{A} is a σ -field $\subset \mathcal{F}$. Then*

$$|\beta(\mathcal{A}, \sigma(X); P) - \beta(\mathcal{A}, \sigma(Y); P)| \leq 2 \cdot P(X \neq Y). \quad (2.6)$$

Proof. By symmetry, it suffices to prove that $\beta(\mathcal{A}, \sigma(X); P) \leq \beta(\mathcal{A}, \sigma(Y); P) + 2 \cdot P(X \neq Y)$. Suppose $\{A_1, A_2, \dots, A_I\}$ and $\{B_1, B_2, \dots, B_J\}$ are each a partition of Ω , with $A_i \in \mathcal{A}$ for each i and $B_j \in \sigma(X)$ for each j . It suffices to show that

$$\frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J |P(A_i \cap B_j) - P(A_i)P(B_j)| \leq \beta(\mathcal{A}, \sigma(Y); P) + 2P(X \neq Y). \quad (2.7)$$

By a well known measure-theoretic fact (a standard elementary generalization of [2, Theorem 20.1]), there exists a partition $\{S_1, S_2, \dots, S_J\}$ of S with $S_j \in \mathcal{S}$ for each j , such that $B_j = \{X \in S_j\}$ for each j .

For any event A ,

$$\begin{aligned}
& \sum_{j=1}^J |P(A \cap \{X \in S_j\}) - P(A \cap \{Y \in S_j\})| \\
& \leq \sum_{j=1}^J |P(A \cap \{X \in S_j\} \cap \{X = Y\}) + P(A \cap \{X \in S_j\} \cap \{X \neq Y\}) \\
& \quad - P(A \cap \{Y \in S_j\} \cap \{X = Y\}) - P(A \cap \{Y \in S_j\} \cap \{X \neq Y\})| \\
& = \sum_{j=1}^J |P(A \cap \{X \in S_j\} \cap \{X \neq Y\}) - P(A \cap \{Y \in S_j\} \cap \{X \neq Y\})| \\
& \leq 2P(A \cap \{X \neq Y\}).
\end{aligned}$$

Applying that with $A = A_i$ and then also with $A = \Omega$, one has that

$$\begin{aligned}
& \sum_{i=1}^I \sum_{j=1}^J |P(A_i \cap B_j) - P(A_i)P(B_j)| \\
& \leq \sum_{i=1}^I \sum_{j=1}^J |P(A_i \cap \{X \in S_j\}) - P(A_i \cap \{Y \in S_j\})| \\
& \quad + \sum_{i=1}^I \sum_{j=1}^J |P(A_i \cap \{Y \in S_j\}) - P(A_i)P(Y \in S_j)| \\
& \quad + \sum_{i=1}^I \sum_{j=1}^J |P(A_i)P(Y \in S_j) - P(A_i)P(X \in S_j)| \\
& \leq \left[\sum_{i=1}^I 2P(A_i \cap \{X \neq Y\}) \right] + 2\beta(\mathcal{A}, \sigma(Y); P) \\
& \quad + \sum_{i=1}^I \left[P(A_i) \sum_{j=1}^J |P(Y \in S_j) - P(X \in S_j)| \right] \\
& \leq 2P(X \neq Y) + 2\beta(\mathcal{A}, \sigma(Y); P) + \sum_{i=1}^I [P(A_i) \cdot 2P(X \neq Y)] \\
& = 4P(X \neq Y) + 2\beta(\mathcal{A}, \sigma(Y); P).
\end{aligned}$$

Thus (2.7) holds. That completes the proof.

The following lemma, the final item of Section 2 here, will play a key role (in Section 4) in the comparison of the mixing rates for the “input” and “output” sequences in the

model in [6]. It will be applied for absolute regularity; but it holds under strong mixing (as stated and proved here) as well.

Lemma 2.3. *Suppose $(H(0), H(1), H(2), H(3), \dots)$ is a nonincreasing sequence of non-negative numbers such that $\sum_{n=0}^{\infty} H(n) < \infty$. Suppose that on some probability space (Ω, \mathcal{F}, P) , $X := (X_k, k \in \mathbf{Z})$ is a nondegenerate, strictly stationary sequence of random variables taking only the values 0 and 1, such that $\sum_{n=1}^{\infty} \alpha(X, n; P) < \infty$. Then*

$$\sum_{n=1}^{\infty} EH(X_1 + X_2 + \dots + X_n) < \infty. \quad (2.8)$$

Proof. Referring to the hypothesis (of Lemma 2.3), define the number

$$p := P(X_0 = 1) = EX_0 > 0 \quad (2.9)$$

Define the constant random variable $S_0 := 0$; and for each positive integer n , define the partial sum $S_n := X_1 + X_2 + \dots + X_n$. By the hypothesis (of Lemma 2.3), the sequence X is strongly mixing and hence ergodic. Hence from (2.9), one has that $S_n \rightarrow \infty$ (monotonically) a.s. as $n \rightarrow \infty$. For technical convenience, without loss of generality (redefining the random variables X_k on a P -null set if necessary), we assume that that happens at literally every $\omega \in \Omega$.

For each nonnegative integer j , define the random variable

$$\eta_j := \text{card}\{n \in \mathbf{N} : S_n = j\}. \quad (2.10)$$

Then for every integer $J \geq 0$,

$$\sum_{j=0}^J \eta_j = \max\{n \geq 0 : S_n = J\}. \quad (2.11)$$

In what follows, for any real number x , let $[x]$ denote the greatest integer $\leq x$. Also, in the calculations below, by the hypothesis (of Lemma 2.3), all sums and summands (“numerical” or random) take their values in $[0, \infty] := [0, \infty) \cup \{\infty\}$, and hence one can change the orders of summations arbitrarily.

Recall that for any nonnegative integer-valued random variable W , $EW = \sum_{n=1}^{\infty} P(W \geq n)$. For each integer $J \geq 0$, by (2.11) and the trivial inequality $P(S_n \leq J) \leq 1$, one has that

$$E\left(\sum_{j=0}^J \eta_j\right) = \sum_{n=1}^{\infty} P\left(\sum_{j=0}^J \eta_j \geq n\right) = \sum_{n=1}^{\infty} P(S_n \leq J) \leq 2J/p + \sum_{n=[2J/p]+1}^{\infty} P(S_n \leq J). \quad (2.12)$$

Let us examine the last sum in (2.12). For each integer $J \geq 0$ and each integer $n > 2J/p$, one has that $J < np/2$ and hence $J - np < -np/2$. Hence for each integer

$J \geq 0$,

$$\begin{aligned} \sum_{n=[2J/p]+1}^{\infty} P(S_n \leq J) &\leq \sum_{n=[2J/p]+1}^{\infty} P(S_n - np \leq -np/2) \leq \sum_{n=[2J/p]+1}^{\infty} P(|S_n - np| \geq np/2) \\ &\leq \sum_{n=[2J/p]+1}^{\infty} (np/2)^{-4} E(S_n - np)^4 \leq (16/p^4) \sum_{n=1}^{\infty} n^{-4} E(S_n - np)^4. \end{aligned} \quad (2.13)$$

Extend the definition (2.4) to include $n = 0$ there. Then here for each positive integer n ,

$$E(S_n - np)^4 \leq (20,000)n \cdot \sum_{m=0}^{n-1} (m+1)^2 \alpha(X, m; P) + 24n^2 \left[\sum_{m=0}^{n-1} \alpha(X, m; P) \right]^2. \quad (2.14)$$

This is a simple direct application of [5, v2, Theorem 14.63], which in turn is a convenient but crude version of a much sharper and more general inequality due to Rio [13, Théorème 2.1]. Keep in mind that since the random variables X_k take only the values 0 and 1, the (“upper-tail”) quantile functions in those particular statements in both references take only the values 0 and 1. Eq. (2.14) can also be obtained directly, in a sharper form, from a careful calculation from Ibragimov’s proof in [10, Theorem 18.5.4] of Theorem 2.1 (examine carefully the argument for [10, Lemma 18.5.2]).

Our next task is to use (2.14) to show that the last sum in (2.13) is finite. From simple calculus, let C_1 be a positive number such that $\sum_{n=q}^{\infty} n^{-3} \leq C_1 q^{-2}$ for every positive integer q . By the hypothesis (of Lemma 2.3),

$$\begin{aligned} \sum_{n=1}^{\infty} \left[n^{-4} \cdot n \cdot \sum_{m=0}^{n-1} (m+1)^2 \alpha(X, m; P) \right] &= \sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty} n^{-3} (m+1)^2 \alpha(X, m; P) \\ &\leq \sum_{m=0}^{\infty} C_1 \alpha(X, m; P) < \infty. \end{aligned}$$

Also, trivially by the hypothesis, $\sum_{n=1}^{\infty} \left[n^{-4} \cdot n^2 \left[\sum_{m=0}^{n-1} \alpha(X, m; P) \right]^2 \right] < \infty$. Applying those two inequalities to (2.14), one obtains that the last sum in (2.13) is finite.

Accordingly, defining the finite numbers $C_2 := (16/p^4) \sum_{n=1}^{\infty} n^{-4} E(S_n - np)^4$ and $C_3 := (2/p) + C_2$, one has by (2.12) and (2.13) that for every integer $J \geq 0$,

$$E\left(\sum_{j=0}^J \eta_j\right) \leq 2J/p + C_2 \leq C_3(J+1). \quad (2.15)$$

Now refer to the function H in the statement of Lemma 2.3. By the hypothesis, $H(n) \downarrow 0$ as $n \rightarrow \infty$. Using the notation $S(n)$ for S_n in subscripts, one has the equality of

nonnegative random variables (possibly taking the value ∞)

$$\begin{aligned}\sum_{n=1}^{\infty} H(S_n) &= \sum_{j=0}^{\infty} \sum_{\{n \in \mathbf{N} : S(n)=j\}} H(j) = \sum_{j=0}^{\infty} H(j) \cdot \eta_j \\ &= \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} \eta_j \cdot [H(i) - H(i+1)] = \sum_{i=0}^{\infty} \sum_{j=0}^i [H(i) - H(i+1)] \cdot \eta_j.\end{aligned}$$

Hence by (2.15),

$$\begin{aligned}E \sum_{n=1}^{\infty} H(S_n) &\leq \sum_{i=0}^{\infty} \left[[H(i) - H(i+1)] \cdot C_3(i+1) \right] = C_3 \sum_{i=0}^{\infty} \sum_{j=0}^i [H(i) - H(i+1)] \\ &= C_3 \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} [H(i) - H(i+1)] = C_3 \sum_{j=0}^{\infty} H(j) < \infty.\end{aligned}$$

Thus (2.8) holds. That completes the proof of Lemma 2.3.

3. The model of Chaudhuri and Dasgupta, in “two-sided” form

In this section, we shall spell out, step by step, the “replicating character string” model studied by Chaudhuri and Dasgupta [6]. As explained in Section 1, essentially the only changes here will be in the “style”: (i) the use of (“two-sided”) random sequences indexed by \mathbf{Z} , rather than (“one-sided”) random sequences indexed by \mathbf{N} , and (ii) the trivial allowing of the “alphabet” or “state space” to be $(0, \infty)$ instead of just a finite set. In the presentation here, the essential mathematical substance of their model will not be changed at all. Much of the notations below will be taken directly from their paper. For convenient reference, the stages in this construction will be referred to as paragraphs (P1), (P2), etc.

(P1). Suppose (Ω, \mathcal{F}, P) is a probability space. All random variables defined below, will be understood to be defined on this space.

(P2). Suppose $X := (X_k, k \in \mathbf{Z})$ is (under P) a strictly stationary sequence of random variables taking their values in the open half line $(0, \infty)$.

(This is the “input” sequence, as in the model in [6].)

(P3). Suppose $Z := (Z_k, k \in \mathbf{Z})$ is (under P) a strictly stationary, ergodic sequence of random variables taking their values in the set $\{M, I, D\}$, with this sequence Z being independent of the sequence X . Assume further that $P(Z_0 = s) > 0$ for all three elements $s \in \{M, I, D\}$. For technical convenience, without loss of generality (by ergodicity), assume that for *every* $\omega \in \Omega$ and all three elements $s \in \{M, I, D\}$, $Z_k(\omega) = s$ for infinitely many negative integers k and infinitely many positive integers k .

(Again, the letters M , I , and D , stand for “mutation”, “insertion”, and “deletion”; Z is the “MID-sequence”, as in [6].)

(P4). Define the strictly increasing sequence $\zeta := (\zeta_j, j \in \mathbf{Z})$ of integer-valued random variables as follows: For every $\omega \in \Omega$,

$$\dots < \zeta_{-2}(\omega) < \zeta_{-1}(\omega) < \zeta_0(\omega) \leq 0 < 1 \leq \zeta_1(\omega) < \zeta_2(\omega) < \zeta_3(\omega) < \dots \quad \text{and} \\ \{j \in \mathbf{Z} : Z_j(\omega) \in \{M, D\}\} = \{\zeta_k(\omega) : k \in \mathbf{Z}\}. \quad (3.1)$$

The random variables ζ_k will sometimes be written $\zeta(k)$ for typographical convenience.

(P5). Define the sequence $\overline{X} := (\overline{X}_k, k \in \mathbf{Z})$ of random variables (taking their values in the closed half line $[0, \infty)$) as follows: For every $\omega \in \Omega$,

$$\forall k \in \mathbf{Z}, \overline{X}_{\zeta(k)(\omega)}(\omega) := X_k(\omega); \quad \text{and} \quad \forall j \notin \{\zeta_k(\omega) : k \in \mathbf{Z}\}, \overline{X}_j(\omega) := 0. \quad (3.2)$$

The state 0 is used here only as a “temporary placeholder” for a spot where an “insertion” will eventually occur (in eq. (3.5) below, in the case where $\overline{Y}_\ell(\omega) = 0$ there). That was the sole motivation for choosing for the original sequence X a state space, namely $(0, \infty)$, that does not include 0.

(P6). Define the strictly increasing sequence $\xi := (\xi_j, j \in \mathbf{Z})$ of integer-valued random variables as follows: For every $\omega \in \Omega$,

$$\dots < \xi_{-2}(\omega) < \xi_{-1}(\omega) < \xi_0(\omega) \leq 0 < 1 \leq \xi_1(\omega) < \xi_2(\omega) < \xi_3(\omega) < \dots \quad \text{and} \\ \{j \in \mathbf{Z} : Z_j(\omega) \in \{M, I\}\} = \{\xi_k(\omega) : k \in \mathbf{Z}\}. \quad (3.3)$$

These random variables ξ_k will sometimes be written $\xi(k)$.

(P7). Define the sequence $\overline{Y} := (\overline{Y}_\ell, \ell \in \mathbf{Z})$ of random variables (taking their values in $[0, \infty)$) as follows: For every $\omega \in \Omega$,

$$\forall \ell \in \mathbf{Z}, \overline{Y}_\ell(\omega) := \overline{X}_{\xi(\ell)(\omega)}(\omega). \quad (3.4)$$

(P8). (i) Let $U := (U_k, k \in \mathbf{Z})$ be (under P) a sequence of independent, identically distributed random variables, each uniformly distributed on the interval $[0, 1]$, with this sequence U being independent of the pair of sequences (X, Z) .

(ii) Let $g : (0, \infty) \times [0, 1] \rightarrow (0, \infty)$ be a Borel function.

(iii) Let $h : [0, 1] \rightarrow (0, \infty)$ be a Borel function.

(iv) Define the sequence $Y := (Y_\ell, \ell \in \mathbf{Z})$ of random variables (taking their values in $(0, \infty)$) as follows: For every $\omega \in \Omega$,

$$\forall \ell \in \mathbf{Z}, Y_\ell(\omega) := \begin{cases} g(\overline{Y}_\ell(\omega), U_\ell(\omega)) & \text{if } \overline{Y}_\ell(\omega) \in (0, \infty) \\ h(U_\ell(\omega)) & \text{if } \overline{Y}_\ell(\omega) = 0. \end{cases} \quad (3.5)$$

(This sequence Y is the “output” sequence, as in [6]. As described in more detail in Remark 3.1(D) below, the functions g and h deal with mutations and insertions respectively.)

The final two “paragraphs” below give a few more items that were not needed in the formulation of the “output” sequence Y but will be needed in the formulation of the main result (Theorem 4.1 in Section 4).

(P9). Let P_0 denote the probability measure on (Ω, \mathcal{F}) defined as follows:

$$\forall F \in \mathcal{F}, P_0(F) := P(F|Z_0 \in \{M, I\}) = P(F|\xi_0 = 0). \quad (3.6)$$

(The second equality follows from paragraph (P6).)

(P10). Define the sequence $V := (V_k, k \in \mathbf{Z})$ of $(\{M, I\} \times \mathbf{N})$ -valued random variables as follows:

$$\forall k \in \mathbf{Z}, V_k := (Z_{\xi(k)}, \xi_k - \xi_{k-1}). \quad (3.7)$$

Also, define the sequence $\Upsilon := (\Upsilon_k, k \in \mathbf{Z})$ of $(\{M, I\} \times \mathbf{N} \times (0, \infty))$ -valued random variables as follows:

$$\forall k \in \mathbf{Z}, \Upsilon_k := (V_k, Y_k) = (Z_{\xi(k)}, \xi_k - \xi_{k-1}, Y_k). \quad (3.8)$$

That completes the (“two-sided”) presentation of the model in [6].

Remark 3.1. Here are several comments pertaining to the model from [6] as spelled out in paragraphs (P1)-(P10) above.

(A) It is well known from renewal theory that even though the sequence Z is (under the original probability measure P) strictly stationary, the sequence $(Z_{\xi(k)}, k \in \mathbf{Z})$ is in general *not* strictly stationary under P . As a consequence, under P the “output” sequence Y will in general not be strictly stationary. To obtain the stationarity of Y , Chaudhuri and Dasgupta [6, Theorem 3.1] assumed directly that that sequence $(Z_{\xi(k)}, k \in \mathbf{Z})$ (though not necessarily the entire sequence Z) is strictly stationary, with the original MID-sequence Z itself being a Markov chain with certain properties. Theorem 4.1 in Section 4 (the main result of this note) will employ the procedure, common in renewal theory, of formally switching to the new probability measure P_0 in (3.6), which under our assumptions will yield the stationarity of Y (and of the entire sequence Υ in (3.8)), with no “Markov” assumption. That is the role here of the probability measure P_0 .

(B) By paragraphs (P4), (P6), and (P10), one has that $\sigma(\zeta, \xi, V) \subset \sigma(Z)$. Recall from paragraphs (P3) and (P8) that under P , the random sequences X , Z , and U are independent of each other. By (3.6) and a trivial argument, those three sequences are independent of each other under P_0 as well. Also by (3.6), the random sequences X and U (but in general not Z or even V) each have the same distribution under P_0 as they do under P . In particular,

$$\forall n \in \mathbf{N}, \beta(X, n; P_0) = \beta(X, n; P) \quad (3.9)$$

(and the analogous equality holds for $\alpha(\cdot, \cdot)$).

(C) In paragraph (P3), it was implicitly understood in the phrase “without loss of generality” that on a certain “bad” event F with $P(F) = 0$, one might need to redefine certain random variables $Z_k, k \in \mathbf{Z}$. By (3.6) $P_0(F) = 0$ as well, and hence that phrase “without loss of generality” applies under P_0 as well as under P .

(D) In [6] (with finite “alphabet”), the probabilities involving “mutations” and “insertions” were specified directly. In [6] it was also pointed out how the context of “mutation” could include, as part of the model, “high probability of no mutation”. Paragraph (P8) just gives an alternative way to set all that up, using the independent random variables U_k uniformly distributed on the interval $[0, 1]$ as “randomizers”, and using appropriate choices of the Borel functions g (to determine “mutations”) and h (to determine “insertions”).

For example, suppose (again in the spirit of [6]) the “input” sequence X is a DNA sequence, with the four nucleotides represented by the letters A, C, G , and T . To “fit” paragraph (P2) and the subsequent paragraphs, one can respectively “code” those letters as the numbers 1, 2, 3, and 4 in $(0, \infty)$. In (3.5), suppose that whenever an “insertion” is to occur, the respective probabilities of inserting the nucleotides A, C, G , and T (the numbers 1, 2, 3, and 4) are to be .15, .5, .25, and .1. One can set that up by using in (3.5) the function h on $[0, 1]$ defined by

$$h(u) := \begin{cases} 1 & \text{if } 0 \leq u < .15 \\ 2 & \text{if } .15 \leq u < .65 \\ 3 & \text{if } .65 \leq u < .9 \\ 4 & \text{if } .9 \leq u \leq 1. \end{cases}$$

Similarly, suppose that whenever a nucleotide A (the state 1) occurs, it is to be “left alone” with probability .95 (“high probability of no mutation”) or to mutate to C, G , or T (2, 3, or 4) with probability .01, .03, or .01 respectively. One can set that up by using in (3.5) a function g that satisfies

$$g(1, u) := \begin{cases} 1 & \text{if } 0 \leq u < .95 \\ 2 & \text{if } .95 \leq u < .96 \\ 3 & \text{if } .96 \leq u < .99 \\ 4 & \text{if } .99 \leq u \leq 1. \end{cases}$$

One can similarly define $g(2, u)$, $g(3, u)$, and $g(4, u)$, $0 \leq u \leq 1$, in an appropriate way to model particular probabilities of specific mutations — or “no mutation” — whenever the nucleotide C, G , or T (the state 2, 3, or 4) occurs.

4. The main result and its proof

This section is devoted to the proof of the following theorem, the main result of this note:

Theorem 4.1. *Assume the entire context of paragraphs (P1)-(P10), with all assumptions there satisfied.*

(I) *Under the probability measure P_0 in (3.6), the sequence Υ (in paragraph (P10)) is strictly stationary (and hence under P_0 the sequences V and Y are each strictly stationary).*

(II) If also $\sum_{n=1}^{\infty} \beta(X, n; P) < \infty$ (see also (3.9)) and $\sum_{n=1}^{\infty} \beta(V, n; P_0) < \infty$, then $\sum_{n=1}^{\infty} \beta(Y, n; P_0) \leq \sum_{n=1}^{\infty} \beta(\Upsilon, n; P_0) < \infty$.

Statement (I) is in spirit a slight extension of [6, Theorem 3.1], which in their setup was a corresponding “preservation of strict stationarity” result involving a “Markov” assumption on the MID-sequence. Statement (II) was inspired by [6, Theorem 3.2], which in their setup was a corresponding “preservation of mixing rate” result involving strong mixing with exponential mixing rate. It seems clear that the setup here in paragraphs (P1)-(P10), involving “two-sided” random sequences, can facilitate the proofs of such “preservation of mixing rates” results involving absolute regularity, such as Statement (II) here; but it is yet to be determined to what extent the setup here might facilitate such results involving strong mixing. The emphasis on summable mixing rates in Statement (II) is motivated by Theorem 2.1, the very sharp central limit theorem of Ibragimov involving summable mixing rates (for strong mixing); recall the comments right after that theorem.

We shall first prove Statement (I). The proof given below will be a somewhat modified version of the argument for [6, Theorem 3.1]. The argument will proceed through a series of lemmas. The first lemma is of a standard form. (In closely related contexts, a very similar fact was used in [3, proof of Lemma 5] and in [4, pp. 7-8]; see also [5, v3, Theorem 26.4(I)].)

Lemma 4.1. *In the context of paragraphs (P1)-(P10) (with all assumptions there satisfied), the sequence $((Z_k, \overline{X}_k), k \in \mathbf{Z})$ is, under the probability measure P , strictly stationary.*

Sketch of proof. Suppose j is any integer. Define the integer-valued random variable $T := \max\{k \in \mathbf{Z} : \zeta_k \leq j\}$. The entire array $((Z_k, k \geq j+1), (\overline{X}_k, k \geq j+1))$ can be represented as $\phi((Z_k, k \geq j+1), (X_{T+1}, X_{T+2}, X_{T+3}, \dots))$, where the (measurable) function $\phi : \{M, I, D\}^{\mathbf{N}} \times (0, \infty)^{\mathbf{N}} \rightarrow \{M, I, D\}^{\mathbf{N}} \times [0, \infty)^{\mathbf{N}}$ does not depend on j . Under P , regardless of j , by the assumptions in paragraphs (P2)-(P5) and an elementary argument, the sequence $(X_{T+1}, X_{T+2}, X_{T+3}, \dots)$ is independent of $\sigma(Z, T)$ ($= \sigma(Z)$) and has the same distribution as the sequence (X_1, X_2, X_3, \dots) , and Lemma 4.1 then follows easily.

Lemma 4.2. *Suppose $L \geq 3$ is an integer. Suppose that for each $\ell \in \{1, 2, \dots, L\}$, $s_\ell \in \{M, I\}$, N_ℓ is a positive integer, and B_ℓ is a Borel subset of $[0, \infty)$. For each $J \in \{1, 2, \dots, L-1\}$, define the event (see (3.7) and (3.3))*

$$F_J := \left\{ Z_0 \in \{M, I\} \right\} \cap \left[\bigcap_{\ell=1}^L \left(\{V_{-J+\ell} = (s_\ell, N_\ell)\} \cap \{\overline{Y}_{-J+\ell} \in B_\ell\} \right) \right]. \quad (4.1)$$

Then $P(F_1) = P(F_2) = \dots = P(F_{L-1})$.

Proof. Define the integers $m_0 := 0$ and $m_\ell := N_1 + N_2 + \dots + N_\ell$ for $\ell \in \{1, 2, \dots, L\}$. These integers m_ℓ will sometimes be written below as $m(\ell)$. Define the set $S := \{1, 2, \dots, m_L\} - \{m_1, m_2, \dots, m_L\}$.

Suppose $J \in \{1, 2, \dots, L-1\}$.

By (3.3), $\{Z_0 \in \{M, I\}\} = \{\xi_0 = 0\}$. As a consequence, one has the equality of events

$$\begin{aligned} \{Z_0 \in \{M, I\}\} \cap \left[\bigcap_{\ell=1}^L \{\xi_{-J+\ell} - \xi_{-J+\ell-1} = N_\ell\} \right] &= \bigcap_{\ell=0}^L \{\xi_{-J+\ell} = -m_J + m_\ell\} \\ &= \left[\bigcap_{\ell=0}^L \{Z_{-m(J)+m(\ell)} \in \{M, I\}\} \right] \cap \left[\bigcap_{u \in S} \{Z_{-m(J)+u} = D\} \right]. \end{aligned} \quad (4.2)$$

Referring to (4.1) and applying both equalities in (4.2) carefully, one obtains

$$\begin{aligned} F_J &= \{Z_0 \in \{M, I\}\} \cap \\ &\quad \left[\bigcap_{\ell=1}^L \left(\{Z_{\xi(-J+\ell)} = s_\ell\} \cap \{\xi_{-J+\ell} - \xi_{-J+\ell-1} = N_\ell\} \cap \{\overline{X}_{\xi(-J+\ell)} \in B_\ell\} \right) \right] \\ &= \left[\bigcap_{\ell=0}^L \{\xi_{-J+\ell} = -m_J + m_\ell\} \right] \cap \left[\bigcap_{\ell=1}^L \left(\{Z_{-m(J)+m(\ell)} = s_\ell\} \cap \{\overline{X}_{-m(J)+m(\ell)} \in B_\ell\} \right) \right] \\ &= \{Z_{-m(J)} \in \{M, I\}\} \cap \left[\bigcap_{\ell=1}^L \left(\{Z_{-m(J)+m(\ell)} = s_\ell\} \cap \{\overline{X}_{-m(J)+m(\ell)} \in B_\ell\} \right) \right] \\ &\quad \cap \left[\bigcap_{u \in S} \{Z_{-m(J)+u} = D\} \right]. \end{aligned} \quad (4.3)$$

By Lemma 4.1, the probability (under P) of the last expression in (4.3) does not depend on $J \in \{1, 2, \dots, L-1\}$. Thus Lemma 4.2 holds.

Lemma 4.3. *The sequence $((V_\ell, \overline{Y}_\ell), \ell \in \mathbf{Z})$ of $(\{M, I\} \times \mathbf{N} \times [0, \infty))$ -valued random variables is strictly stationary under P_0 .*

Proof. Suppose $j \in \mathbf{Z}$ and $n \in \mathbf{N}$. It suffices to prove that under P_0 , the “random vectors” $((V_\ell, \overline{Y}_\ell), \ell \in \{j+1, j+2, \dots, j+n\})$ and $((V_\ell, \overline{Y}_\ell), \ell \in \{j+2, j+3, \dots, j+n+1\})$ have the same distribution (on $(\{M, I\} \times \mathbf{N} \times [0, \infty))^n$).

Let J and L be positive integers such that $\{J-1, J\} \subset \{1, 2, \dots, L-1\}$ and $\{j+1, j+2, \dots, j+n\} \subset \{-J+1, -J+2, \dots, -J+L\}$. It suffices to prove that under P_0 , the “random vectors” $((V_\ell, \overline{Y}_\ell), \ell \in \{-J+1, -J+2, \dots, -J+L\})$ and $((V_\ell, \overline{Y}_\ell), \ell \in \{-J+2, -J+3, \dots, -J+L+1\})$ have the same distribution. But that holds by (3.6), Lemma 4.2, and a trivial calculation. Thus Lemma 4.3 holds.

Proof of Statement (I) in Theorem 4.1. By paragraph (P8) (see Remark 3.1(B)), the sequence U is, under P_0 , independent of the sequence $((V_\ell, \overline{Y}_\ell), \ell \in \mathbf{Z})$. It follows from paragraph (P8) and Lemma 4.3 (again see Remark 3.1(B)) that under P_0 , the sequence $((V_\ell, \overline{Y}_\ell, U_\ell), \ell \in \mathbf{Z})$ is strictly stationary. Now Statement (I) in Theorem 4.1 holds by (3.5) and (3.8).

The proof of Statement (II) in Theorem 4.1 will be based on the following lemma. In what follows, $E_0(\dots)$ denotes expected value with respect to the probability measure P_0 . The indicator function of a given event A will be denoted $I(A)$.

Lemma 4.4. *In the context of (P1)-(P10) (with all assumptions there satisfied), suppose also that $\sum_{n=1}^{\infty} \beta(X, n; P) < \infty$. Define the sequence $(H(n), n \in \{0\} \cap \mathbf{N})$ of nonnegative numbers as follows: For each $n \geq 0$, $H(n) := \beta(X, n+1; P)$.*

Suppose N is an integer such that $N \geq 2$. Then

$$\beta(\Upsilon, N; P_0) \leq \beta(V, N; P_0) + 2E_0 H\left(\sum_{i=1}^{N-1} I(Z_{\xi(i)} = M)\right). \quad (4.4)$$

Proof. The proof of this lemma will proceed through a series of “steps”.

Step 1. Refer to the integer $N \geq 2$ in the hypothesis (of Lemma 4.4). Define the nonnegative integer-valued random variable T by

$$\begin{aligned} T &:= \text{card}\left\{k \in \mathbf{N} : 1 \leq k \leq \xi_N - 1 \text{ and } Z_k \in \{M, D\}\right\} \\ &= \max\{j \in \{0\} \cup \mathbf{N} : \zeta_j < \xi_N\}. \end{aligned} \quad (4.5)$$

(The second equality in (4.5) holds by (3.1).) Define the (“one-sided”) sequence $X^* := (X_1^*, X_2^*, X_3^*, \dots)$ of random variables as follows:

$$\forall k \geq 1, \quad X_k^* := X_{k+T}. \quad (4.6)$$

Step 2. Now let us first look at the random variable \overline{Y}_N . Suppose $\omega \in \Omega$. If $Z_{\xi(N)(\omega)}(\omega) = I$, then $\xi_N(\omega) \notin \{\zeta_k(\omega) : k \in \mathbf{Z}\}$ by (3.1), and $\overline{Y}_N(\omega) = 0$ by (3.4) and (3.2). If instead $Z_{\xi(N)(\omega)}(\omega) = M$ (the only other possibility, by (3.3)), then for some $q \geq 1$, $\xi_N(\omega) = \zeta_q(\omega)$, hence $q = T(\omega) + 1$ by (4.5), and hence $\overline{Y}_N(\omega) = \overline{X}_{\zeta(q)(\omega)}(\omega) = X_q(\omega) = X_{T(\omega)+1}(\omega) = X_1^*(\omega)$ by (3.4), (3.2) and (4.6).

Thus $\overline{Y}_N = 0 \cdot I(Z_{\xi(N)} = I) + X_1^* \cdot I(Z_{\xi(N)} = M)$. Hence by (3.7),

$$\sigma(\overline{Y}_N) \subset \sigma(V_N, X_1^*). \quad (4.7)$$

Step 3. Now suppose ℓ is any integer such that $\ell > N$. Our task here in Step 3 is to obtain some sort of analog of (4.7) for \overline{Y}_ℓ .

First define the random variable

$$\tau := \text{card}\left\{k \in \mathbf{N} : \xi_N \leq k \leq \xi_\ell \text{ and } Z_k \in \{M, D\}\right\}. \quad (4.8)$$

Then $\tau = [\sum_{i=N}^{\ell} I(Z_{\xi(i)} = M)] + [\sum_{i=N+1}^{\ell} (\xi_i - \xi_{i-1} - 1)]$. Here in the right hand side, for a given $\omega \in \Omega$, by (3.3), the first sum is simply the number of indices k in the set in (4.8) such that $Z_k(\omega) = M$, and the second sum is simply the number of indices k in that

set such that $Z_k(\omega) = D$. (Either sum can be zero.) From that expression for τ , one has by (3.7) that

$$\sigma(\tau) \subset \sigma(V_N, V_{N+1}, \dots, V_\ell). \quad (4.9)$$

Now suppose $\omega \in \Omega$. Consider first the case where $Z_{\xi(\ell)(\omega)}(\omega) = M$. Then for some $q \geq 1$, $\xi_\ell(\omega) = \zeta_q(\omega)$; and by (4.5), (4.8), and (3.1), $q = T(\omega) + \tau(\omega)$. Hence by (3.4), (3.2), and (4.6), $\bar{Y}_\ell(\omega) = \bar{X}_{\zeta(q)(\omega)}(\omega) = X_q(\omega) = X_{\tau(\omega)}^*(\omega)$. Also in the case where $Z_{\xi(\ell)(\omega)}(\omega) = M$, one has that $\tau(\omega) \geq 1$ by (4.8). If instead $Z_{\xi(\ell)(\omega)}(\omega) = I$ (the only other possibility), then $\bar{Y}_\ell(\omega) = 0$ by (3.4) and (3.2).

Thus (for our given $\ell > N$), putting all those pieces together,

$$\bar{Y}_\ell = 0 \cdot I(Z_{\xi(\ell)} = I) + X_\tau^* \cdot I(Z_{\xi(\ell)} = M) = 0 + \sum_{t=1}^{\infty} [X_t^* \cdot I(\tau = t) \cdot I(Z_{\xi(\ell)} = M)],$$

and hence by (4.9) and (3.7),

$$\sigma(\bar{Y}_\ell) \subset \sigma(V_N, V_{N+1}, \dots, V_\ell) \vee \sigma(X_1^*, X_2^*, X_3^*, \dots). \quad (4.10)$$

Step 4. Combining (4.7) and (4.10), one now has that

$$\sigma(\bar{Y}_N, \bar{Y}_{N+1}, \bar{Y}_{N+2}, \dots) \subset \sigma(V_N, V_{N+1}, V_{N+2}, \dots) \vee \sigma(X_1^*, X_2^*, X_3^*, \dots). \quad (4.11)$$

Define the two random arrays **A** and **B** by

$$\mathbf{A} := \left((X_k, k \leq 0); (V_k, k \leq 0); (U_k, k \leq 0) \right) \quad \text{and} \quad (4.12)$$

$$\mathbf{B} := \left(X^*; (V_k, k \geq N); (U_k, k \geq N) \right). \quad (4.13)$$

By (3.5), $\sigma(Y_\ell) \subset \sigma(\bar{Y}_\ell, U_\ell)$ for each integer ℓ . One now has by (3.8) and (4.11) that

$$\sigma(\Upsilon_\ell, \ell \geq N) \subset \sigma(\mathbf{B}). \quad (4.14)$$

We need some sort of analog of (4.14) for $\sigma(\Upsilon_\ell, \ell \leq 0)$ and $\sigma(\mathbf{A})$.

For that purpose, we will need to work with the probability measure P_0 (in (3.6)). Some more notations will be needed. For events A and B , the notation $A \doteq B$ will mean that $P_0(A \triangle B) = 0$, where \triangle denotes the symmetric difference. For an event A and a σ -field \mathcal{B} , the notation $A \dot{\in} \mathcal{B}$ will mean that $A \doteq B$ for some $B \in \mathcal{B}$. For σ -fields \mathcal{A} and \mathcal{B} , the notation $\mathcal{A} \dot{\subset} \mathcal{B}$ will mean that $A \dot{\in} \mathcal{B}$ for every $A \in \mathcal{A}$, and the notation $\mathcal{A} \doteq \mathcal{B}$ will mean that $\mathcal{A} \dot{\subset} \mathcal{B}$ and $\mathcal{B} \dot{\subset} \mathcal{A}$.

Refer to (3.7) and both equalities in (3.6). For $\omega \in \{\xi_0 = 0\}$, the “ordered pairs” $(V_k(\omega), k \leq 0)$ determine (“measurably”) the set of integers $\{\xi_k(\omega), k \leq 0\}$ as well as $Z_j(\omega)$ (M or I) for j in that set, and hence determine $Z_j(\omega)$ for all $j \leq 0$ (since $Z_j(\omega) = D$ for integers $j \leq 0$ that are not in that set). Combining that with (3.3), one obtains that $\sigma(V_k, k \leq 0) \doteq \sigma(Z_k, k \leq 0)$. Now $\sigma(\bar{X}_k, k \leq 0) \subset \sigma(X_k, Z_k, k \leq 0)$ by (3.1)

and (3.2); hence $\sigma(\bar{Y}_k, k \leq 0) \subset \sigma(X_k, Z_k, k \leq 0)$ by (3.3) and (3.4); and hence also $\sigma(Y_k, k \leq 0) \subset \sigma(X_k, Z_k, U_k, k \leq 0)$ by (3.5). Hence by (4.12) and (3.8),

$$\sigma(\Upsilon_k, k \leq 0) \dot{\subset} \sigma(\mathbf{A}). \quad (4.15)$$

Step 5. On the original probability space (Ω, \mathcal{F}, P) let Γ be a random variable which (under P) is uniformly distributed on the interval $[0, 1]$ and is independent of $\sigma(X, Z, U)$ (recall paragraphs (P2), (P3), and (P8)). Now of course by the hypothesis (of Lemma 4.4), the sequence $\beta(X, n; P) \rightarrow 0$ as $n \rightarrow \infty$ (absolute regularity). At this point, we shall apply the “coupling” result of Berbee [1, p. 104, Theorem 4.4.7 and p. 106, the Note], which is closely related to the “maximal coupling” result of Goldstein [9]. As a convenient reference for Berbee’s result, we cite [5, v2, p. 277, Theorem 20.7 and pp. 477-478, Lemma A1651]. The latter “lemma” in that reference simply involves the use of the random variable Γ above as a “randomizer”, and it is simply (a special case of) the version in Dudley and Philipp [8, Lemma 2.11] of a theorem of Skorohod [15]. Thereby there exists a sequence $X' := (X'_k, k \in \mathbf{Z})$ of random variables with the following properties (under P):

$$\text{the distributions of the sequences } X \text{ and } X' \text{ are identical (under } P); \quad (4.16)$$

$$\text{the } \sigma\text{-fields } \sigma(X') \text{ and } \sigma(X_k, k \leq 0) \text{ are independent (under } P); \quad (4.17)$$

$$\forall n \in \mathbf{N}, P(\exists k \geq n \text{ such that } X'_k \neq X_k) = \beta(X, n; P); \text{ and} \quad (4.18)$$

$$\sigma(X') \subset \sigma(X, \Gamma). \quad (4.19)$$

By paragraphs (P1)-(P8) and the properties of the random variable Γ , one has that under P the σ -fields $\sigma(\Gamma)$, $\sigma(X)$, $\sigma(Z)$, and $\sigma(U)$ are independent. By (3.6) and a trivial argument, that holds under P_0 as well. Further, by (3.6) and a trivial argument, the distribution of the random array (Γ, X, X', U) (on $[0, 1] \times \mathbf{R}^{\mathbf{Z}} \times \mathbf{R}^{\mathbf{Z}} \times [0, 1]^{\mathbf{Z}}$) is the same under P_0 as it is under P . In particular, (4.16)-(4.18) all hold with P replaced with P_0 (see also (3.9)).

Thus under P_0 , the following statements hold: By (4.19) the σ -fields, $\sigma(X, X')$, $\sigma(Z)$, and $\sigma(U)$ are independent; hence by (4.17) the σ -fields $\sigma(X')$, $\sigma(X_k, k \leq 0)$, $\sigma(Z)$, and $\sigma(U)$ are independent; and hence the σ -field $\sigma(X')$ is independent of the σ -field $\sigma(U, Z) \vee \sigma(X_k, k \leq 0)$.

Step 6. Now referring to (4.5), define analogously to (4.6) the (“one-sided”) sequence $X'^* := (X'^*_1, X'^*_2, X'^*_3, \dots)$ of random variables as follows:

$$\forall k \geq 1, X'^*_k := X'_{k+T}. \quad (4.20)$$

Now consider an arbitrary event $A \subset \sigma(U, Z) \vee \sigma(X_k, k \leq 0)$ such that $P_0(A) > 0$. For each integer $t \geq 0$ such that $P_0(A \cap \{T = t\}) > 0$, one now has (note that $\sigma(T) \subset \sigma(Z)$) that

$$\begin{aligned} \mathcal{L}_0(X'^* | A \cap \{T = t\}) &= \mathcal{L}_0((X'_{t+1}, X'_{t+2}, X'_{t+3}, \dots) | A \cap \{T = t\}) \\ &= \mathcal{L}_0(X'_{t+1}, X'_{t+2}, X'_{t+3}, \dots) = \mathcal{L}_0(X'_1, X'_2, X'_3, \dots), \end{aligned} \quad (4.21)$$

where $\mathcal{L}_0(\dots)$ (resp. $\mathcal{L}_0(\dots | \dots)$) denotes the distribution (resp. conditional distribution) under P_0 . Since the last term in (4.21) is “constant” (not depending on A or t), it follows by

a simple standard calculation that for each such event A , $\mathcal{L}_0(X'^*|A) = \mathcal{L}_0(X'_1, X'_2, X'_3, \dots)$, and also (consider the case $A = \Omega$) $\mathcal{L}_0(X'^*) = \mathcal{L}_0(X'_1, X'_2, X'_3, \dots)$. Consequently, the sequence X'^* is (under P_0) independent of the σ -field $\sigma(U, Z) \vee \sigma(X_k, k \leq 0)$.

Analogously to (4.13), define the random array \mathbf{B}' as follows:

$$\mathbf{B}' := \left(X'^*; (V_k, k \geq N); (U_k, k \geq N) \right). \quad (4.22)$$

Referring to the last sentence of the preceding paragraph and the third sentence after (4.19) (which together with paragraph (P8)(i) yields $\beta(U, N; P_0) = 0$), one has by (4.12), (4.22), Remark 3.1(B), and Lemma 2.1 that

$$\beta(\sigma(\mathbf{A}), \sigma(\mathbf{B}'); P_0) = \beta(V, N; P_0). \quad (4.23)$$

Also, by (4.5), (4.13), (4.22), and (4.18) (and the fact $\sigma(T) \subset \sigma(Z)$), with the sums below taken over all nonnegative integers t such that $P_0(T = t) > 0$, one has (recall the sequence $H(\cdot)$ in the statement of Lemma 4.4) that

$$\begin{aligned} P_0(\mathbf{B}' \neq \mathbf{B}) &= P_0(X'^* \neq X^*) = \sum P_0(X'^* \neq X^* | T = t) \cdot P_0(T = t) \\ &= \sum P_0((X'_{t+1}, X'_{t+2}, X'_{t+3}, \dots) \neq (X_{t+1}, X_{t+2}, X_{t+3}, \dots) | T = t) \cdot P_0(T = t) \\ &= \sum P_0((X'_{t+1}, X'_{t+2}, X'_{t+3}, \dots) \neq (X_{t+1}, X_{t+2}, X_{t+3}, \dots)) \cdot P_0(T = t) \\ &= \sum \beta(X, t+1; P_0) \cdot P_0(T = t) \\ &= E_0 H(T). \end{aligned}$$

Hence by (4.12)-(4.15), (4.23), and Lemma 2.2 (and Theorem 4.1(I), proved above)

$$\begin{aligned} \beta(\Upsilon, N; P_0) &\leq \beta(\sigma(\mathbf{A}), \sigma(\mathbf{B}); P_0) \leq \beta(\sigma(\mathbf{A}), \sigma(\mathbf{B}'); P_0) + 2E_0 H(T) \\ &= \beta(V, N; P_0) + 2E_0 H(T). \end{aligned} \quad (4.24)$$

Now by (4.5), $T \geq \sum_{i=1}^{N-1} I(Z_{\xi(i)} = M)$. Also, the sequence $H(n)$ (in the statement of Lemma 4.4) is nonincreasing as n increases. It follows that $E_0 H(T) \leq E_0 H(\sum_{i=1}^{N-1} I(Z_{\xi(i)} = M))$. Combining this with (4.24) one obtains Lemma 4.4.

Proof of Statement (II) in Theorem 4.1. Define the sequence $\Theta := (I(Z_{\xi(i)} = M), i \in \mathbf{Z})$ of random indicator functions. Under P_0 this sequence is strictly stationary by (3.7) and Theorem 4.1(I) (proved above). By (3.6), paragraph (P3), and a trivial argument, the sequence Θ is also nondegenerate under P_0 . Also, by (3.7) and the hypothesis (of Theorem 4.1(II)), one has that $\sum_{n=1}^{\infty} \beta(\Theta, n; P_0) < \infty$. Also, by the hypothesis (of Theorem 4.1(II)), the (nonincreasing) sequence $(H(n), n \in \{0\} \cap \mathbf{N})$ of nonnegative numbers in Lemma 4.4 is summable. Hence for that sequence $H(\cdot)$, by Lemma 2.3, $\sum_{n=2}^{\infty} E_0 H(\sum_{i=1}^{n-1} I(Z_{\xi(i)} = M)) < \infty$. Hence by (3.8), Lemma 4.4, and the hypothesis (of Theorem 4.1(II)), the conclusion of Theorem 4.1(II) holds. That completes the proof.

Acknowledgement. The author thanks the referee for a correction and for helpful comments that improved the exposition.

References

- [1] Berbee, H.C.P. (1979). *Random Walks with Stationary Increments and Renewal Theory*. Mathematical Centre, Amsterdam.
- [2] Billingsley, P. (1995). *Probability and Measure*, 3rd edn. Wiley, New York.
- [3] Bradley, R.C. (1980). A remark on the central limit question for dependent random variables. *J. Appl. Probab.* **17**, 94-101.
- [4] Bradley, R.C. (1989). A stationary, pairwise independent, absolutely regular sequence for which the central limit theorem fails. *Probab. Th. Rel. Fields* **81**, 1-10.
- [5] Bradley, R.C. (2007). *Introduction to Strong Mixing Conditions*, Volumes 1, 2, and 3. Kendrick Press, Heber City (Utah).
- [6] Chaudhuri, P. and Dasgupta, A. (2006). Stationarity and mixing properties of replicating character strings. *Statistica Sinica* **16**, 29-43.
- [7] Davydov, Yu.A. (1973). Mixing conditions for Markov chains. *Theor. Probab. Appl.* **18**, 312-328.
- [8] Dudley, R.M. and Philipp, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw. Gebiete* **62**, 509-552.
- [9] Goldstein, S. (1979). Maximal coupling. *Z. Wahrsch. verw. Gebiete* **46**, 193-204.
- [10] Ibragimov, I.A. and Linnik, Yu.V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.
- [11] Merlevède, F. and Peligrad, M. (2000). The functional central limit theorem under the strong mixing condition. *Ann. Probab.* **28**, 1336-1352.
- [12] Pinsker, M.S. (1964). *Information and Information Stability of Random Variables and Processes*. Holden-Day, San Francisco.
- [13] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*. Mathématiques & Applications 31. Springer, Berlin.
- [14] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42**, 43-47.
- [15] Skorohod, A.V. (1976). On a representation of random variables. *Theor. Probab. Appl.* **21**, 628-632.
- [16] Volkonskii, V.A. and Rozanov, Yu.A. (1959). Some limit theorems for random functions I. *Theor. Probab. Appl.* **4**, 178-197.
- [17] Waterman, M.S. (1995). *Introduction to Computational Biology*. Chapman and Hall, New York.