Statistics and Probability Letters 78 (2008) 2910–2915

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Constructing processes with prescribed mixing coefficients

Leonid (Aryeh) Kontorovich *

Department of Mathematics, Weizmann Institute of Science, Rehovot, 76100, Israel

ARTICLE INFO

Article history:
Received 7 November 2007
Received in revised form 16 February 2008
Accepted 28 April 2008
Available online 4 May 2008

ABSTRACT

The rate at which dependencies between future and past observations decay in a random process may be quantified in terms of mixing coefficients. The latter in turn appear in strong laws of large numbers and concentration of measure inequalities for dependent random variables. Questions regarding what rates are possible for various notions of mixing have been posed since the 1960's, and have important implications for some open problems in the theory of strong mixing conditions.

This paper deals with η-mixing, a notion defined in [Kontorovich, Leonid, Ramanan, Kavita, 2008. Concentration inequalities for dependent random variables via the martingale method. Ann. Probab. (in press)], which is closely related to ϕ-mixing. We show that there exist measures on finite sequences with essentially arbitrary η-mixing coefficients, as well as processes with arbitrarily slow mixing rates.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Preliminaries

Strong mixing conditions deal with quantifying the decaying dependence between blocks of random variables in a stochastic process. These have been traditionally used to establish strong laws of large numbers for non-independent processes. Bradley (2007a,b,c) is an encyclopedic source on the matter; see also his survey paper Bradley (2005). In Bradley (2007c, Chapter 26), the early research on mixing rates is traced to Volkonskiĭ and Rozanov (1961) and a comprehensive account of progress since then is given.

Our interest in strong mixing was motivated by the desire for a concentration of measure bounds for non-independent random sequences. Given the excellent survey papers and monographs dealing with a concentration of measure (in particular, Ledoux (2001), Lugosi (2003), and Schechtman (2003)), we will give only the briefest summary here.

Suppose Ω is a finite[1] set and let μ be an arbitrary probability measure on Ω^n. We proceed to define a type of strong mixing used throughout this note. For $1 \le i < j \le n$ and $x \in \Omega^i$, let

$$\mathcal{L}(X_j^n \mid X_1^i = x)$$

be the distribution of $X_j^n \equiv (X_j, \dots, X_n)$ conditioned on $X_1^i = x$. For $y \in \Omega^{i-1}$ and $w, w' \in \Omega$, define

$$\eta_{ij}(y, w, w') = \left\| \mathcal{L}(X_j^n \mid X_1^i = yw) - \mathcal{L}(X_j^n \mid X_1^i = yw') \right\|_{\mathrm{TV}}, \tag{1}$$

* Tel.: +972 52 846 0383.
E-mail address: aryeh.kontorovich@weizmann.ac.il.

[1] The quoted results hold verbatim for countable sets, and extend naturally to $\mathbb{R}$ under mild assumptions; see Kontorovich (2007); Kontorovich and Brockwell (in preparation).

0167-7152/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2008.04.016

where $\|\cdot\|_{\mathrm{TV}} \equiv \frac{1}{2}\|\cdot\|_1$ is the total variation norm; likewise, define

$$\bar{\eta}_{ij} = \max_{y \in \Omega^{i-1}, w, w' \in \Omega} \eta_{ij}(y, w, w'). \tag{2}$$

Remark 1. The definition in (1) is flawed as stated, as it does not adequately handle the case where sets of measure zero are conditioned on. One could attempt to perturb μ slightly so as to make it have full support and appeal to continuity arguments, but here too care must be taken. Taking $\mathcal{P}_+^n(\Omega)$ to be the set of all strictly positive probability measures μ on Ω^n (i.e., $\mu(x) > 0$ for all $x \in \Omega^n$), it is not hard to show that the functional $\bar{\eta}_{ij} : \mathcal{P}_+^n(\Omega) \to \mathbb{R}$ is continuous on $\mathcal{P}_+^n(\Omega)$ with respect to $\|\cdot\|_{\mathrm{TV}}$. However, this continuity can break down on the boundary of $\mathcal{P}_+^n(\Omega)$. Thus, we may arbitrarily define $\bar{\eta}_{ij}$ on any probability measure μ by

$$\bar{\eta}_{ij}(\mu) = \inf_{\{\mu_k\}} \lim_{k\to\infty} \bar{\eta}_{ij}(\mu_k) \tag{3}$$

where the infimum is taken over all sequences $\{\mu_k : \mu_k \in \mathcal{P}_+^n(\Omega), \|\mu_k - \mu\|_{\mathrm{TV}} \to 0\}$.

See Section 5.4 of Kontorovich (2007) for a discussion of the continuity of $\bar{\eta}$ and conditioning on sets of measure zero, as well as motivation for the definition in (3).

The notion of mixing defined in (2) is by no means new; it can be traced (at least implicitly) to Marton (1998) and is quite explicit in Samson (2000) and Chazottes et al. (2007). We are not aware of a standardized term for this type of mixing, and have referred to it as *η-mixing* in Kontorovich and Ramanan (in press). It was observed in Samson (2000) that the ϕ-mixing coefficients bound the η-mixing ones:

$$\bar{\eta}_{ij} \le 2\phi_{j-i},$$

and conjectured in Kontorovich (2007) that

$$\frac{1}{2}\sum_{i=1}^{n-1}\phi_i \le 1 + \max_{1\le i<n}\left[\sum_{j=i+1}^{n}\bar{\eta}_{ij}\right];$$

the latter remains open.

In all instances, η-mixing has come up in the context of concentration of measure. In particular, define Γ and Δ to be upper-triangular $n \times n$ matrices, with $\Gamma_{ii} = \Delta_{ii} = 1$ and

$$\Gamma_{ij} = \sqrt{\bar{\eta}_{ij}}, \qquad \Delta_{ij} = \bar{\eta}_{ij} \tag{4}$$

for $1 \le i < j \le n$.

Samson (2000) proved that any distribution μ on $[0, 1]^n$ and any convex $f : [0, 1]^n \to \mathbb{R}$ with $\|f\|_{\mathrm{Lip}} \le 1$ (with respect to ℓ_2) satisfy

$$\mu\{|f - \mu f| > t\} \le 2\exp\left(-\frac{t^2}{2\|\Gamma\|_2^2}\right) \tag{5}$$

where $\|\Gamma\|_2$ is the ℓ_2 operator norm.

Chazottes et al. (2007) and independently, Kontorovich and Ramanan (in press) showed that any distribution μ on Ω^n and any $f : \Omega^n \to \mathbb{R}$ with $\|f\|_{\mathrm{Lip}} \le n^{-1/2}$ (with respect to the Hamming metric) satisfy

$$\mu\{|f - \mu f| > t\} \le 2\exp\left(-\frac{t^2}{2\|\Delta\|_\infty^2}\right) \tag{6}$$

where $\|\Delta\|_\infty$ is the ℓ_∞ operator norm ($\|\Delta\|_\infty$ may be replaced by $\|\Delta\|_2$ and Chazottes et al. (2007) achieve a better constant in the exponent).

Results of type (5) and (6) are known as *concentration of measure* inequalities; broadly, they assert that any "sufficiently continuous" function is tightly concentrated about its mean. Such bounds have a remarkable range of applications, spanning abstract fields such as asymptotic Banach space theory (Ball, 1997; Schechtman, 2003) as well as more practical ones such as randomized algorithms (Dubhashi and Panconesi, 1998) and machine learning (Boucheron et al., 2005). Strong laws of large numbers are readily obtained from concentration bounds (Kontorovich and Brockwell, in preparation).

Having motivated the study of mixing and measure concentration, let us turn to the behavior of the η-mixing coefficients. It is immediate from the construction that $\bar{\eta}_{ij}$ is an upper-triangular $n \times n$ matrix satisfying

(H1) $\bar{\eta}_{ij} = 0$ for $i \ge j$
(H2) $0 \le \bar{\eta}_{ij} \le 1$ for $1 \le i < j \le n$.

It is also simple to show (as we shall do below in Lemma 2) that

(H3) $\bar{\eta}_{ij_2} \le \bar{\eta}_{ij_1}$ for $i < j_1 < j_2$.

1.2. Main results

A natural question (first posed in Kontorovich (2007)) is whether the conditions (H1)–(H3) completely characterize the possible $(\bar{\eta}_{ij})$ matrices, or if there are some other constraints that the η-mixing coefficients must satisfy. The main technical result of this note is Theorem 6, which resolves this question in the affirmative. Thus, for any "valid" (i.e., satisfying (H1)–(H3)) $n \times n$ matrix $H = (h_{ij})$, there is a finite set Ω and a probability measure μ on Ω^n such that $\bar{\eta}_{ij}(\mu) = h_{ij}$ for $1 \le i < j \le n$.

More broadly, it is of interest to characterize the possible mixing rates that various processes may have. Bradley (2007c, Chapter 26) deals with this question and gives several intricate constructions of random processes having prescribed mixing rates, under various types of strong mixing. Following the work of Kesten and O'Brien (1976), it emerged that essentially arbitrary mixing rates are possible for various mixing notions. Thus it is not surprising that the same holds true for η-mixing; this is an easy consequence of our main result (Corollary 8).

Along the way, we collect various other observations regarding the η-mixing coefficients—some of which are auxiliary in proving our main results, and others may be of independent interest.

1.3. Notation

We use the indicator variable $\mathbb{1}_{\{\cdot\}}$ to assign 0-1 truth values to the predicate in $\{\cdot\}$.

Random variables are capitalized (X), specified sequences are written in lowercase ($x \in \Omega^n$), the shorthand $X_i^j = (X_i, \dots, X_j)$ is used for all sequences, and sequence concatenation is denoted multiplicatively: $x_i^j x_{j+1}^k = x_i^k$. Sums will range over the entire space of the summation variable; thus $\sum_{x_i^j} f(x_i^j)$ stands for

$$\sum_{x_i^j \in \Omega_i^j} f(x_i^j),$$

where Ω_i^j is just Ω^{j-i+1}, re-indexed for convenience. For $y \in \Omega_1^i$ and $x \in \Omega_j^n$, we will write $\mu(x \mid y)$ as a shorthand for $\mu\left\{X_j^n = x \mid X_1^i = y\right\}$; no confusion should arise.

The *total variation* norm of a signed measure ν on Ω^n (i.e., vector $\nu \in \mathbb{R}^{\Omega^n}$) is defined by

$$\|\nu\|_{\mathrm{TV}} = \frac{1}{2}\|\nu\|_1 = \frac{1}{2}\sum_{x \in \Omega^n} |\nu(x)|$$

(the factor of 1/2 is not entirely standard). Unless otherwise stated, Ω is a finite set. Whenever we wish to be explicit about the dependence of η_{ij} and $\bar{\eta}_{ij}$ on a given measure μ, we will write $\eta_{ij}(\mu; y, w, w')$ and $\bar{\eta}_{ij}(\mu)$, respectively.

2. Constructions and proofs

Let us begin with an easy verification that (H3) holds for all $(\bar{\eta}_{ij})$:

Lemma 2. *Let $(\bar{\eta}_{ij})_{1 \le i < j \le n}$, be the η-mixing matrix associated with a probability measure μ on Ω^n. Then, for all $1 \le i < j_1 < j_2 \le n$, we have*

$$\bar{\eta}_{ij_2} \le \bar{\eta}_{ij_1}.$$

Proof. Fix $1 \le i < j_1 < j_2 \le n$ and $y \in \Omega_1^{i-1}$, $w, w' \in \Omega_i^i$. Then

$$\begin{aligned}
\eta_{ij_2}(y, w, w') &= \frac{1}{2}\sum_{x \in \Omega_{j_2}^n} \left|\mu(x \mid yw) - \mu(x \mid yw')\right| \\
&= \frac{1}{2}\sum_{x \in \Omega_{j_2}^n} \left|\sum_{u \in \Omega_{j_1}^{j_2-1}} [\mu(ux \mid yw) - \mu(ux \mid yw')]\right| \\
&\le \frac{1}{2}\sum_{x \in \Omega_{j_2}^n} \sum_{u \in \Omega_{j_1}^{j_2-1}} \left|\mu(ux \mid yw) - \mu(ux \mid yw')\right| \\
&= \frac{1}{2}\sum_{z \in \Omega_{j_1}^n} \left|\mu(z \mid yw) - \mu(z \mid yw')\right| \\
&= \eta_{ij_1}(y, w, w'). \quad \square
\end{aligned}$$

Our construction of a measure with the desired mixing coefficients will proceed in stages, the final object being composed of intermediate ones. The building blocks will be measures of a certain simple form. For $1 \le k < n$, let $h \in \mathbb{R}^n_{k+1}$ be a vector of length $n - k$, satisfying

$$0 \le h_{j+1} \le h_j \le 1$$

for $k < j < n$; any such h will be called a *valid* kth *row*. We say that the measure μ on Ω^n is *pure* kth *row* (with respect to h) if its η-mixing matrix $(\bar{\eta}_{ij})_{1 \le i < j \le n}$ satisfies

$$\bar{\eta}_{ij} = \mathbb{1}_{\{i=k\}} h_j.$$

Our first technical result is the existence of arbitrary pure kth row measures:

Lemma 3. *Fix $1 \le k < n$ and suppose $h \in \mathbb{R}^n_{k+1}$ is a valid kth row vector. Then there exists a measure μ on $\{-1, 1\}^n$ which is pure kth row with respect to h.*

Remark 4. The proof given here is due to an anonymous reviewer; it captures the probabilistic intuition much better than the original.

Proof. Let a valid kth row vector $h = (h_{k+1}, h_{k+1}, \dots, h_n) \in [0, 1]^{n-k}$ be fixed. For a given $v = (v_{k+1}, v_{k+1}, \dots, v_n) \in [0, 1]^{n-k}$, we shall define two sequences of random variables, $X_1, \dots, X_n$ and $U_{k+1}, U_{k+2}, \dots, U_n$ taking values in $\{-1, 1\}$. The distribution of these variables is specified by the following relations:

(a) the variables $X_1, \dots, X_k, U_{k+1}, U_{k+2}, \dots, U_n$ are mutually independent
(b) for $1 \le i \le k$, $\mathbf{P}\{X_i = 1\} = \mathbf{P}\{X_i = -1\} = 1/2$
(c) for $k < i \le n$, $\mathbf{P}\{U_i = 1\} = v_i$ and $\mathbf{P}\{U_i = -1\} = 1 - v_i$
(d) for $k < j \le n$, $X_j = U_j \cdot X_k$.

The random variables $X_1, \dots, X_n$ induce the (nonproduct) measure μ_v on $\{-1, 1\}^n$, parametrized by the vector v.

We shall make a few simple observations. Some are immediate while others are straightforward (if tedious) exercises; detailed proofs are omitted due to space constraints.

(e) For each $x \in \{-1, 1\}^n$, the number $\mathbf{P}\{X_1^n = x\} \equiv \mu_v(x)$ is a continuous function of the parameter v in the closed unit cube $[0, 1]^{n-k}$.
(f) For each $z \in \{-1, 1\}^k$, we have $\mu_v\{X_1^k = z\} = 2^{-k}$, regardless of the choice of v.
(g) Given any $y \in \{-1, 1\}^{k-1}$ and $w, w' \in \{-1, 1\}$, and any $k < j \le n$, the mixing coefficient $\eta_{kj}(\mu_v; y, w, w')$ is a continuous function of v and depends *only* on $v_j, v_{j+1}, \dots, v_n$ (and not on $v_{k+1}, \dots, v_{j-1}$).
(h) For each $k < j \le n$, the mixing coefficient $\bar{\eta}_{kj}(\mu_v)$ is a continuous function of v and depends *only* on $v_j, v_{j+1}, \dots, v_n$.
(i) The random variables $X_{k+1}, X_{k+2}, \dots, X_n$ are conditionally independent given $X_1, X_2, \dots, X_k$.
(j) The random vectors $(X_1, X_2, \dots, X_{k-1})$ and $(X_k, X_{k+1}, \dots, X_n)$ are independent of each other.

We are going to choose the parameters $v_n, v_{n-1}, \dots, v_{k+1}$ inductively so that μ_v is pure kth row with respect to the given $h = (h_{k+1}, h_{k+2}, \dots, h_n)$.

Define the function $f_n : [0, 1] \to [0, 1]$ by $f_n(v_n) = \bar{\eta}_{kn}(\mu_v)$, and recall from above that f_n is continuous and only depends on v_n. Observe that $f_n(1/2) = 0$ and $f_n(1) = 1$. Thus, there exists a $v_n^* \in [1/2, 1]$ such that $\bar{\eta}_{kn}(\mu_v) = h_n$, regardless of the choice of $v_{k+1}, v_{k+1}, \dots, v_{n-1}$. Let us henceforth fix $v_n := v_n^*$.

Now consider the function $f_{n-1} : [0, 1] \to [0, 1]$ defined by $f_{n-1}(v_{n-1}) = \bar{\eta}_{k,n-1}(\mu_v)$. It formally depends on v_{n-1} and v_n (and is independent of the choice of $v_{k+1}, v_{k+1}, \dots, v_{n-2}$), but since we have fixed $v_n = v_n^*$, f_{n-1} is indeed only a function of v_{n-1}. Observe that $f_{n-1}(1/2) = h_n$ and $f_{n-1}(1) = 1$, and so there is a $v_{n-1}^* \in [1/2, 1]$ such that $\bar{\eta}_{k,n-1} = h_{n-1}$.

In a similar manner, we can choose $v_{n-2}^*, v_{n-3}^*, \dots, v_{k+1}^*$ (crucially, in that order) so that $\bar{\eta}_{kj}(\mu_{v^*}) = h_j$ for each $k < j \le n$. It is a simple (though somewhat tedious) matter to verify that $\bar{\eta}_{ij}(\mu_{v^*}) = 0$ for $i \ne k$. □

Next we turn to product measures. There are (at least) two natural ways to form products of probability measures; we shall refer to them as *series* and *parallel*. Let $\mathcal{X}, \mathcal{Y}$ be finite sets and $m, n \in \mathbb{N}$. If μ is a measure on $\mathcal{X}^m$ and ν a measure on $\mathcal{X}^n$, we define their series product, denoted by $\mu \oplus \nu$, to be the following measure on $\mathcal{X}^{m+n}$:

$$(\mu \oplus \nu)(z) = \mu(x)\nu(y), \qquad z = xy \in \mathcal{X}^{m+n}, \quad x \in \mathcal{X}^m, \ y \in \mathcal{X}^n. \tag{7}$$

If μ is a measure on $\mathcal{X}^n$ and ν a measure on $\mathcal{Y}^n$, we define their parallel product, denoted by $\mu \otimes \nu$, to be the following measure on $(\mathcal{X} \times \mathcal{Y})^n$:

$$(\mu \otimes \nu)(z) = \mu(x)\nu(y), \qquad z = (x, y) \in (\mathcal{X} \times \mathcal{Y})^n. \tag{8}$$

As our main construction will involve parallel products of measures, the following simple result is useful.

Lemma 5. *Let μ and ν be probability measures on $\mathcal{X}^n$ and $\mathcal{Y}^n$, respectively, and let $\bar{\eta}_{ij}(\mu)$, $\bar{\eta}_{ij}(\nu)$ and $\bar{\eta}_{ij}(\mu \otimes \nu)$ be the corresponding η-mixing matrices. Then we have*

$$\max\left\{\bar{\eta}_{ij}(\mu), \bar{\eta}_{ij}(\nu)\right\} \leq \bar{\eta}_{ij}(\mu \otimes \nu) \leq \bar{\eta}_{ij}(\mu) + \bar{\eta}_{ij}(\nu) \quad (9)$$

for all $1 \leq i < j \leq n$.

Proof. The lower bound is obtained via an elementary (albeit tedious) calculation, similar to the one employed in the proof of Lemma 2, involving breaking up summations and moving absolute values to inner sums; it is omitted due to space constraints. The upper bound follows from the easily verified (and well-known) fact that $\|p \otimes q - p' \otimes q'\|_{\mathrm{TV}} \leq \|p - p'\|_{\mathrm{TV}} + \|q - q'\|_{\mathrm{TV}}$. □

The interested reader may consult Kontorovich (2007, Lemma 3.2.1) for some observations regarding the behavior of η-mixing coefficients under series products.

We are now ready to prove the main result of this note.

Theorem 6. *Let $H = (h_{ij})$ be any $n \times n$ matrix satisfying* (H1)–(H3). *Then there exists a finite set Ω and a probability measure μ on Ω^n such that*

$$\bar{\eta}_{ij}(\mu) = h_{ij} \quad (10)$$

for $1 \leq i < j \leq n$.

Proof. For $k = 1, \dots, n-1$, let $h^{(k)} \in \mathbb{R}^n_{k+1}$ be the vector $(h_{k,k+1}, h_{k,k+2}, \dots, h_{k,n})$. Then Lemma 3 provides a measure $\mu^{(k)}$ on $\{-1, 1\}^n$ which is pure kth row with respect to $h^{(k)}$. Let μ be the (parallel) product of these pure kth row measures:

$$\mu = \mu^{(1)} \otimes \mu^{(2)} \otimes \cdots \otimes \mu^{(n-1)};$$

note that μ is a measure on Ω^n, where $\Omega = \{-1, 1\}^{n-1}$. By definition of pure kth row measures and by Lemma 5, we have that (10) holds. □

Remark 7. Our construction requires an exponential state space, $|\Omega| = 2^{n-1}$. Are there analogous constructions using fewer states? In Kontorovich (2007, Section 5.7) we constructed a measure μ on a binary state space satisfying (10) for the special case where the rows of H are constant: $h_{i,i+1} = h_{i,i+2} = \cdots = h_{i,n}$; it seems unlikely that the general case is achievable with a constant number of states.

Up to this point, we have been discussing the η-mixing coefficients of probability measures on finite sequences. This notion extends quite naturally to random processes—i.e., probability measures μ on $\Omega^{\mathbb{N}}$. Let μ_n be the marginal distribution of X_1^n and denote by $\bar{\eta}_{ij}^{(n)}$ the η-mixing matrix of μ_n. It is straightforward to verify that in general, $\bar{\eta}_{ij}^{(n)}$ depends on n and that

$$\bar{\eta}_{ij}^{(n)} \leq \bar{\eta}_{ij}^{(n+1)}$$

for $1 \leq i < j \leq n$. Let $\Delta_n(\mu)$ be the $n \times n$ matrix Δ corresponding to μ_n, as defined in (4). Recall that the ℓ_∞ operator norm of a nonnegative matrix is its maximal row sum. Thus we can define the η-mixing rate of the process μ as the function $R_\mu : \mathbb{N} \to \mathbb{R}$:

$$R_\mu(n) = \|\Delta_n(\mu)\|_\infty .$$

It is clear that (i) R_μ is nondecreasing and (ii) $1 \leq R_\mu(n) \leq n$; any function satisfying these properties will be called a *valid* rate function.

Corollary 8. *Let $r : \mathbb{N} \to \mathbb{N}$ be a valid rate function. Then for $\Omega = \mathbb{N}$ there is a measure μ on $\Omega^{\mathbb{N}}$ such that*

$$\limsup_{n\to\infty} \frac{R_\mu(n)}{r(n)} = 1. \quad (11)$$

Remark 9. To accommodate countably infinite state spaces, the max in (2) needs to be replaced with sup. The original proof used an uncountable state space and had measure-theoretic gaps; the present countable-state-space version was suggested by the reviewer.

Proof. We begin with the simple observation that if r is a valid rate function then for all $k \geq 1$ and all $0 < \varepsilon < 1$, there is an $n = n(k, \varepsilon) > k$ and an $b = b(k, \varepsilon) \in [0, 1]$ such that

$$1 - \varepsilon \leq \frac{b(k, \varepsilon)(n - k)}{r(n)} \leq 1. \quad (12)$$

Let $1 > \varepsilon_1 > \varepsilon_2 > \cdots > 0$ be a sequence decreasing to 0. Pick a $k \geq 1$ and let $n(k) = n(k, \varepsilon_k)$ and $b(k) = b(k, \varepsilon_k)$, as stipulated in (12). Define $h^{(k)} \in \mathbb{R}^n_{k+1}$ by

$$h_j^{(k)} = b(k), \quad k < j \leq n(k),$$

and let $\mu^{(k)}$ be the measure on $\{-1, 1\}^{n(k)}$ which is pure kth row with respect to $h^{(k)}$, as constructed in Lemma 3. Let us slightly modify the construction in Lemma 3, firstly by renaming the state space from $\{-1, 1\}$ to $\{0, 1\}$ and secondly by requiring in step (b) that $X_i \equiv 0$ for $1 \leq i < k$ (for $i = k$ we still have $\mathbf{P}\{X_k = 0\} = \mathbf{P}\{X_k = 1\} = 1/2$). This affects the observation in (f) but nothing else; this modified measure $\tilde{\mu}^{(k)}$ on $\{0, 1\}^{n(k)}$ is still pure kth row with respect to $h^{(k)}$.

Let β be the symmetric Bernoulli measure on $\{0, 1\}$ (i.e., $\beta(0) = \beta(1) = 1/2$) and define the measure $\hat{\mu}^{(k)}$ on $\{0, 1\}^{\mathbb{N}}$ by

$$\hat{\mu}^{(k)} = \tilde{\mu}^{(k)} \oplus \beta \oplus \beta \oplus \cdots$$

where the operation $\oplus$ is defined in (7). In this way, we have obtained a countable collection of measures $\left\{\hat{\mu}^{(k)} : k \in \mathbb{N}\right\}$ on $\{0, 1\}^{\mathbb{N}}$; note that by construction, we have for each k

$$1 - \varepsilon_k \leq \frac{\left\| \Delta_{n(k)}(\hat{\mu}^{(k)}) \right\|_\infty}{r(n(k))} \leq 1. \tag{13}$$

Now let μ be the measure on $(\{0, 1\}^{\mathbb{N}})^{\mathbb{N}}$ obtained by taking the (parallel) product of all the $\hat{\mu}^{(k)}$'s:

$$\mu = \hat{\mu}^{(1)} \otimes \hat{\mu}^{(2)} \otimes \cdots$$

(the $\otimes$ operator is defined in (8)). It follows from Lemma 5 that (13) also holds with μ in place of $\hat{\mu}^{(k)}$.

Although μ and its induced random process $\{X_i : i \in \mathbb{N}, X_i \in \{0, 1\}^{\mathbb{N}}\}$ are formally defined on the uncountable state space $\{0, 1\}^{\mathbb{N}}$, observe that the modified construction of $\tilde{\mu}^{(k)}$ ensures that for all $i \in \mathbb{N}$, we have $\mu\{X_i \in \Omega_{\mathrm{FIN}}\} = 1$, where $\Omega_{\mathrm{FIN}} \subset \{0, 1\}^{\mathbb{N}}$ is defined by

$$\Omega_{\mathrm{FIN}} = \left\{ x \in \{0, 1\}^{\mathbb{N}} : \sum_{u \in \mathbb{N}} x_u < \infty \right\}$$

(that is, Ω_{FIN} consists of all sequences $x \in \{0, 1\}^{\mathbb{N}}$ which contain finitely many 1s). Since Ω_{FIN} is countable and the support of μ is contained in $\Omega_{\mathrm{FIN}}^{\mathbb{N}}$, we may take the state space $\Omega = \Omega_{\mathrm{FIN}}$ (or equivalently, $\Omega = \mathbb{N}$). It is readily verified that $\hat{\mu}^{(k)}$ and μ are well-defined measures on $\{0, 1\}^{\mathbb{N}}$ and $\Omega_{\mathrm{FIN}}^{\mathbb{N}}$, respectively— for example, via the Ionescu Tulcea theorem (Kallenberg, 2002, Theorem 6.17). □

Remark 10. Is there a construction achieving (11) with lim in place of lim sup?

Acknowledgments

My thesis advisor John Lafferty encouraged me to explore the question of measures having prescribed mixing coefficients. I also thank Gideon Schechtman for hosting and guidance at the Weizmann Institute. A special thanks to Richard Bradley for the very helpful correspondence. This paper owes much of its current form to an anonymous referee, who not only read it closely, weeding out typos and measure-theoretic inaccuracies, but also effectively rewrote (and much improved) the proofs of Lemma 3 and Corollary 8. My sincere gratitude for his diligence and generosity. The author was supported in part by the Israel Science Foundation.

References

Ball, Keith, 1997. In: Levy, S. (Ed.), An Elementary Introduction to Modern Convex Geometry. Flavors of Geometry. In: MSRI Publications, vol. 31. Cambridge University Press, New York, pp. 1–58.
Boucheron, Stéphane, Bousquet, Olivier, Lugosi, Gábor, 2005. Theory of classification: A survey of recent advances. ESAIM Probab. Stat. 9, 323–375.
Bradley, Richard C., 2005. Basic properties of strong mixing conditions. a survey and some open questions. Probab. Surv. 2, 107–144.
Bradley, Richard C., 2007a. Introduction to Strong Mixing Conditions. vol. 1. Kendrick Press.
Bradley, Richard C., 2007b. Introduction to Strong Mixing Conditions. vol. 2. Kendrick Press.
Bradley, Richard C., 2007c. Introduction to Strong Mixing Conditions. vol. 3. Kendrick Press.
Chazottes, Jean-René, Collet, Pierre, Külske, Christof, Redig, Frank, 2007. Concentration inequalities for random fields via coupling. Probab. Theory Related Fields 137 (1–2), 201–225.
Dubhashi, Devdatt, Panconesi, Alessandro, 1998. Concentration of measure for the analysis of randomised algorithms, book draft.
Kallenberg, Olav, 2002. Foundations of modern probability. In: Probability and its Applications, second ed. Springer-Verlag.
Kesten, Harry, O'Brien, George L., 1976. Examples of mixing sequences. Duke Math. J. 40 (2), 405–415.
Kontorovich, Leonid, 2007. Measure Concentration of Strongly Mixing Processes with Applications. Ph.D. Thesis, Carnegie Mellon University.
Kontorovich, Leonid, Brockwell, Anthony E., 2008. A Strong Law of Large Numbers for Strongly Mixing Processes (in preparation).
Kontorovich, Leonid, Ramanan, Kavita, 2008. Concentration inequalities for dependent random variables via the martingale method. Ann. Probab. (in press).
Ledoux, Michel, 2001. The Concentration of Measure Phenomenon. In: Mathematical Surveys and Monographs, vol. 89. American Mathematical Society.
Lugosi, Gábor, 2003. Concentration-of-measure inequalities, http://www.econ.upf.es/~lugosi/anu.ps.
Marton, Katalin, 1998. Measure concentration for a class of random processes. Probability Theory and Related Fields 110 (3), 427–439.
Samson, Paul-Marie, 2000. Concentration of measure inequalities for Markov chains and Φ-mixing processes. Ann. Probab. 28 (1), 416–461.
Schechtman, Gideon, 2003. Concentration, Results and Applications. In: Handbook of the Geometry of Banach Spaces, vol 2. North-Holland.
Volkonskiĭ, V.A., Rozanov, Ju.A., 1961. Some limit theorems for random functions. II. Teor. Verojatnost. i Primenen 6, 202–215 (in Russian).