Duke Math. J. **120**, no. 3 (2003), 515–575. Version of 15 Nov. 2019

Stationary Determinantal Processes: Phase Multiplicity, Bernoullicity, Entropy, and Domination

by Russell Lyons and Jeffrey E. Steif

Abstract. We study a class of stationary processes indexed by $\mathbb{Z}^d$ that are defined via minors of d-dimensional (multilevel) Toeplitz matrices. We obtain necessary and sufficient conditions for phase multiplicity (the existence of a phase transition) analogous to that which occurs in statistical mechanics. Phase uniqueness is equivalent to the presence of a strong K property, a particular strengthening of the usual K (Kolmogorov) property. We show that all of these processes are Bernoulli shifts (isomorphic to i.i.d. processes in the sense of ergodic theory). We obtain estimates of their entropies and we relate these processes via stochastic domination to product measures.

Contents

2000 Mathematics Subject Classification. Primary 82B26, 28D05, 60G10. Secondary 82B20, 37A05, 37A25, 37A60, 60G25, 60G60, 60B15.
Key words and phrases. Random field, Kolmogorov mixing, Toeplitz determinants, negative association, fermionic lattice model, stochastic domination, entropy, Szegő infimum, Fourier coefficients, insertion tolerance, deletion tolerance, geometric mean.
Research partially supported by NSF grants DMS-9802663, DMS-0103897 (Lyons), and DMS-0103841 (Steif), the Swedish Natural Science Research Council (Steif) and the Erwin Schrödinger Institute for Mathematical Physics, Vienna (Lyons and Steif).

§1. Introduction.

Determinantal probability measures and point processes arise in numerous settings, such as mathematical physics (where they are called fermionic point processes), random matrix theory, representation theory, and certain other areas of probability theory. See Soshnikov (2000) for a survey and Lyons (2003b) for additional developments in the discrete case. We present here a detailed analysis of the discrete stationary case. After this paper was first written, we learned of independent but slightly prior work of Shirai and Takahashi (2003), announced in Shirai and Takahashi (2000). We discuss the (small) overlap between their work and ours in the appropriate places below. See also Shirai and Takahashi (2002) and Shirai and Yoo (2002) for related contemporaneous work.

Stationary determinantal processes are interesting from several viewpoints. First, they have interesting relations with the theory of Toeplitz determinants. As in that theory, the **geometric mean** of a nonnegative function f, defined as

$$\mathsf{GM}(f) := \exp \int \log f \,,$$

will play an important role in some of our results. (In fact, the arithmetic mean and the harmonic mean of f will also characterize certain properties of our processes.) Second, such processes arise in certain combinatorial models, such as uniform spanning trees and dimer models. Third, these systems have a rich infinite-dimensional parameter space, consisting, in the case of a $\mathbb{Z}^d$ action, of all measurable functions f from the d-dimensional torus to $[0, 1]$. We illustrate the variety of possible behaviors with examples throughout the paper. Fourth, they have the unusual property of negative association. Though unusual, negative association occurs in various places and a fair amount is known about it (see Joag-Dev and Proschan (1983), Pemantle (2000), Newman (1984), Shao and Su (1999), Shao (2000), Zhang and Wen (2001), Zhang (2001), and the references therein, for example). We offer a whole new class of examples of negatively associated stationary processes. As such, determinantal processes provide an easy way to construct examples of many kinds of behavior that might otherwise be difficult to construct, such as negatively associated (stationary) processes with slow decay of correlations, or with the even sites independent of the odd sites (in one dimension, say), or with the property of being finitely dependent. Fifth, all our processes are Bernoulli shifts, i.e., isomorphic to i.i.d. processes. This may be surprising in view of the fact that only measurability, rather than smoothness, of the parameter f is required. Sixth, in one dimension, some of the processes are strong K, while others are not. Namely, strong K is equivalent to $f(\mathbf{1} - f)$ having a positive geometric mean. Similarly, we characterize exactly, in all dimensions, which f give strong *full* K

systems. It turns out that not only the rate at which f approaches 0 or 1 matters, but also where. For example, in two dimensions, if f is real analytic, then the system is strong full K iff the (possibly empty) sets $f^{-1}(0)$ and $f^{-1}(1)$ belong to nontrivial algebraic varieties. The strong full K property is analogous to phase uniqueness in statistical physics, as we explain in Section 7.

We now state our results somewhat more precisely and present several examples. Let $f : \mathbb{T}^d \to [0,1]$ be a Lebesgue-measurable function on the d-dimensional torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. Define a $\mathbb{Z}^d$-invariant probability measure $\mathbf{P}^f$ on the Borel sets of $\{0,1\}^{\mathbb{Z}^d}$ by defining the probabilities of the cylinder sets

$$\begin{aligned}\mathbf{P}^f[\eta(e_1)=1,\ldots,\eta(e_k)=1] &:= \mathbf{P}^f[\{\eta\in\{0,1\}^{\mathbb{Z}^d}\,;\ \eta(e_1)=1,\ldots,\eta(e_k)=1\}]\\ &:= \det[\widehat{f}(e_j-e_i)]_{1\le i,j\le k}\end{aligned}$$

for all $e_1,\ldots,e_k\in\mathbb{Z}^d$, where $\widehat{f}$ denotes the Fourier coefficients of f. We shall prove in Section 2 that this does indeed define a probability measure. Note that when $d=1$ and when $e_1,\ldots,e_k$ are chosen to be k consecutive integers, the right-hand side above is the usual $k\times k$ Toeplitz determinant of f denoted $D_{k-1}(f)$. In particular, we have $\mathbf{P}^f\big[\eta(e)=1\big]=\widehat{f}(\mathbf{0})=\int_{\mathbb{T}^d} f$ for every $e\in\mathbb{Z}^d$.

Example 1.1. As a simple example, if $f\equiv p$, then $\mathbf{P}^f$ is i.i.d. Bernoulli(p) measure.

Example 1.2. For a more interesting example, consider

$$f(x,y) := \frac{\sin^2 \pi x}{\sin^2 \pi x + \sin^2 \pi y}\,. \tag{1.1}$$

A portion of a sample of a configuration from $\mathbf{P}^f$ is shown in Figure 1, where a square with upper left corner (i,j) is colored black iff $\eta(i,j)=1$. These correspond to the horizontal edges of a uniform spanning tree in the square lattice. That is, if T is a spanning tree of $\mathbb{Z}^2$, then $\eta(i,j)$ is the indicator that the edge from (i,j) to $(i+1,j)$ belongs to T. The portion of the spanning tree from which Figure 1 was constructed is shown in Figure 2. When T is chosen "uniformly" (see Pemantle (1991), Lyons (1998), or Benjamini, Lyons, Peres, and Schramm (2001) for definitions and information on this), then η has the law $\mathbf{P}^f$. This follows from the Transfer Current Theorem of Burton and Pemantle (1993) and the representation of the Green function as an integral; see Lyons with Peres (2004) for more details. Similarly, the edges of the uniform spanning forest (it is a tree only for $d\le 4$, as shown by Pemantle (1991)) parallel to the x_1-axis in d dimensions correspond to the function

$$f(x_1,x_2,\ldots,x_d) := \frac{\sin^2 \pi x_1}{\sum_{j=1}^d \sin^2 \pi x_j}\,. \tag{1.2}$$

We remark that the uniform spanning tree is the so-called random-cluster model when one takes the limit $q \downarrow 0$, then $p \downarrow 0$, and finally the thermodynamic limit, the latter shown to exist by Pemantle (1991).

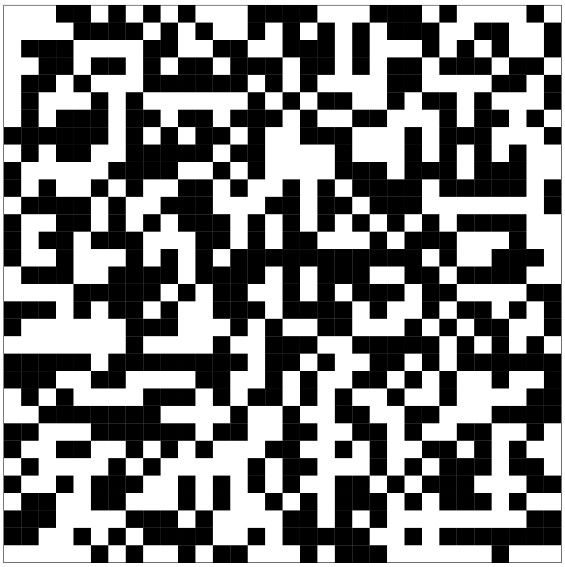

Figure 1. A sample from $\mathbf{P}^f$ of (1.1).

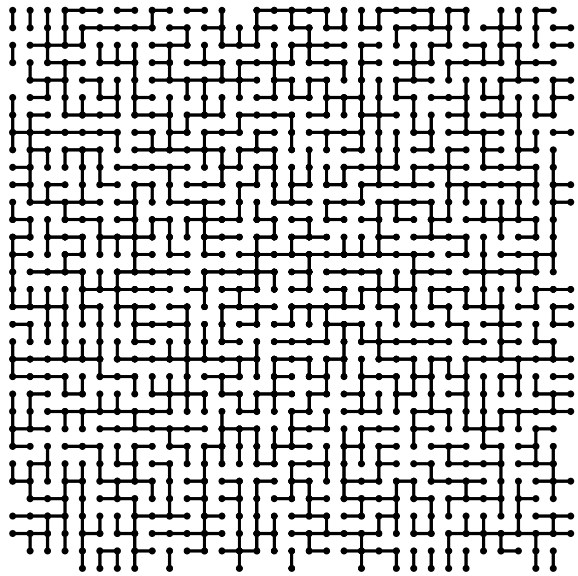

Figure 2. A uniform spanning tree.

EXAMPLE 1.3. Let

$$g(x) := \frac{\sin \pi x}{\sqrt{1 + \sin^2 \pi x}} .$$

Then the edges of the uniform spanning tree in the plane that lie on the x-axis have the law $\mathbf{P}^g$, as shown in Example 5.5 below.

EXAMPLE 1.4. Let

$$f(x) := \frac{1}{2} + \frac{|\sin 2\pi x| - 1}{2 \cos 2\pi x} .$$

An elementary calculation shows that

$$\widehat{f}(k) = \begin{cases} 1/2 & \text{if } k = 0, \\ 0 & \text{if } k \neq 0 \text{ is even}, \\ (-1)^{(k-1)/2} \left(-\frac{1}{2} + \frac{2}{\pi} \sum_{j=0}^{(k-1)/2} \frac{(-1)^j}{2j+1} \right) - \frac{1}{\pi k} & \text{if } k \text{ is odd}. \end{cases}$$

We shall show in Example 5.21 that the measure $\mathbf{P}^f$ arises as follows. Given a spanning tree T of the square lattice, let $\eta(n)$ be the indicator that $e_n \in T$, where e_n is the edge

$$e_n := \begin{cases} [(n/2, n/2), (n/2 + 1, n/2)] & \text{if } n \text{ is even}, \\ [((n+1)/2, (n-1)/2), ((n+1)/2, (n+1)/2)] & \text{if } n \text{ is odd}. \end{cases}$$

The collection of edges $\{e_n \, ; \, n \in \mathbb{Z}\}$ is a zig-zag path in the plane. The law of η is $\mathbf{P}^f$ when T is chosen as a uniform spanning tree. (Although it is not hard to see that the law of η is $\mathbb{Z}$-invariant, by using planar duality, and although the law of η must be a determinantal probability measure because the law of T is, it is not apparent *a priori* that the law of η has the form $\mathbf{P}^f$ for some f.)

EXAMPLE 1.5. Fix a horizontal edge of the hexagonal lattice (also known as the honeycomb lattice) and index all its vertical translates by $\mathbb{Z}$. If one considers the standard measure of maximal entropy on perfect matchings of the hexagonal lattice, also called the dimer model and equivalent to lozenge tilings of the plane, and looks only at the edges indexed as above by $\mathbb{Z}$, then the law is $\mathbf{P}^f$ for $f := \mathbf{1}_{[1/3,2/3]}$, as shown by Kenyon (1997).

EXAMPLE 1.6. It is interesting that the function $f := \mathbf{1}_{[0,1/2]}$ for $d = 1$ also arises from a combinatorial model. In this case,

$$\widehat{f}(n) = \begin{cases} 1/2 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0 \text{ is even}, \\ 1/(\pi i n) & \text{if } n \text{ is odd}. \end{cases}$$

The measure $\mathbf{P}^f$ is the zig-zag process of Johansson (2002) derived from uniform domino tilings in the plane. For the definition of "uniform" in this case, see Burton and Pemantle (1993). A picture of a portion of such a tiling is shown in Figure 3. Consider the squares on a diagonal from upper left to lower right. The domino covering any such square also covers a second square. If this second square is above the diagonal, then we color the original square black, as shown in Figure 4. Johansson (2002) showed that the law of this process is $\mathbf{P}^f$ when the diagonal squares are indexed by $\mathbb{Z}$ in the natural way. More generally, the processes $\mathbf{P}^f$ for f the indicator of any arc of $\mathbb{T}$ are used by Borodin, Okounkov, and Olshanski (2000), Theorem 3, to describe the typical shape of Young diagrams.

EXAMPLE 1.7. Let $0 < a < 1$ and $d = 1$. If $f(x) := (1-a)^2/|e^{2\pi i x} - a|^2$, then $\mathbf{P}^f$ is a renewal process (Soshnikov, 2000). The number of 0s between successive 1s has the same distribution as the number of tails until 2 heads appear for a coin that has probability a of coming up tails. More explicitly, for $n \geq 1$,

$$\mathbf{P}^f[\eta(1) = \cdots = \eta(n-1) = 0, \eta(n) = 1 \mid \eta(0) = 1] = n(1-a)^2 a^{n-1} \,. \tag{1.3}$$

Since

$$f(x) = \frac{1-a}{1+a} \left(\frac{a e^{2\pi i x}}{1 - a e^{2\pi i x}} + \frac{1}{1 - a e^{-2\pi i x}} \right) ,$$

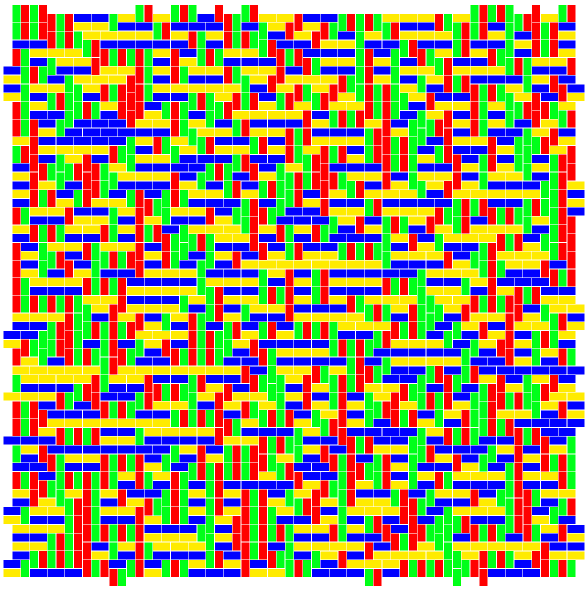

Figure 3. A uniform domino tiling.

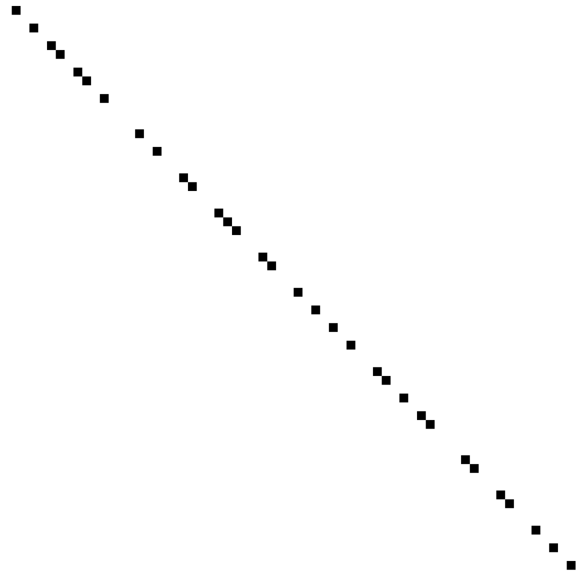

Figure 4. A sample from $\mathbf{P}^f$ of Example 1.6.

expansion in a geometric series shows that

$$\widehat{f}(k) = \frac{1-a}{1+a} a^{|k|}\,.$$

We prove that $\mathbf{P}^f$ is indeed this explicit renewal process after we prove Proposition 2.10, in which we extend this example to other regenerative processes.

Example 1.8. If $0 < p < 1$ and $f : \mathbb{T}^d \to [0,1]$ is measurable, then a fair sample of $\mathbf{P}^{pf}$ can be obtained from a fair sample of $\mathbf{P}^f$ simply by independently changing each 1 to a 0 with probability $1 - p$.

For general systems, note that the covariance of $\eta(\mathbf{0})$ and $\eta(k)$ for $k \in \mathbb{Z}^d$ is $-|\widehat{f}(k)|^2$. This is summable in k since $f \in L^2(\mathbb{T}^d)$, but that is essentially the most one can say for its rate of decay. That is, given any $\langle a_k \rangle \in \ell^1(\mathbb{Z}^d)$, there is some (even continuous) $f : \mathbb{T}^d \to [0,1]$ and some constant $c > 0$ such that $|\widehat{f}(k)|^2 \geq c|a_k|$ for all $k \in \mathbb{Z}^d$, as shown by de Leeuw, Katznelson, and Kahane (1977). Observe also that as is the case for Gaussian processes, the processes studied here have the property that if the random variables are uncorrelated, then they are mutually independent.

It is shown (in a much more general context) by Lyons (2003b) that these measures, as well as these measures conditioned on the values of η restricted to any finite subset of $\mathbb{Z}^d$, have the following negative association property: If A and B are increasing events that are measurable with respect to the values of η on disjoint subsets of $\mathbb{Z}^d$, then A and B are negatively correlated.

In the order presented in this paper, our principal results are the following.

• For all f, the process $\mathbf{P}^f$ is a Bernoulli shift. This was shown in dimension 1 by Shirai and Takahashi (2003) for those f such that $\sum_{n\geq 1} n|\widehat{f}(n)|^2 < \infty$ by showing that those $\mathbf{P}^f$ are weak Bernoulli.

• For all f, the process $\mathbf{P}^f$ stochastically dominates product measure $\mathbf{P}^{\mathsf{GM}(f)}$ and is stochastically dominated by product measure $\mathbf{P}^{1-\mathsf{GM}(\mathbf{1}-f)}$, and these bounds are optimal. This is rather unexpected for the process of Example 1.2 related to the uniform spanning tree; explicit calculations are given below in Example 5.13 for this particular process. We give similar optimal bounds for full domination (uniform insertion and deletion tolerance) in terms of harmonic means.

• We present methods to estimate the entropy of $\mathbf{P}^f$. For example, we show that for the function g in Example 1.3, the entropy of the system $\mathbf{P}^g$ lies in the interval $[0.69005, 0.69013]$; see Example 6.15.

• The process $\mathbf{P}^f$ is strong full K iff there is a nonzero trigonometric polynomial T such that $\frac{|T|^2}{f(\mathbf{1}-f)} \in L^1(\mathbb{T}^d)$. This is equivalent to phase uniqueness in the sense that no conditioning at infinity can change the measure.

• In one dimension, $\mathbf{P}^f$ is strong K iff $f(\mathbf{1}-f)$ has a positive geometric mean. This is equivalent to phase uniqueness when conditioning on one side only. Higher-dimensional versions of this will also be obtained.

We shall give full definitions as they become needed. Some general background on determinantal probability measures is presented in Section 2, where we also exhibit a key representation of certain conditional probabilities as Szegő infima. The property of being a Bernoulli shift is proved in Section 3, while the auxiliary result that all $\mathbf{P}^f$ have full support (except in two degenerate cases) is shown in Section 4. Properties concerning stochastic domination are proved in Section 5. These are used to estimate entropy in Section 6. More sophisticated methods of estimating entropy are also developed and illustrated in Section 6. The main result about phase multiplicity is proved in Section 7, while the one-sided case for $d = 1$ and its higher-dimensional generalizations are treated in Section 8. Finally, we end with some open questions in Section 9.

The above definitions can be generalized to any countable abelian group with the discrete topology. For example, if we use $\mathbb{Z} \times \mathbb{Z}_2$, we obtain nontrivial joinings of the above systems with themselves. That is, suppose $f : \mathbb{T} \to [0,1]$ and $h : \mathbb{T} \to [0,1]$ are such that $f \pm h : \mathbb{T} \to [0,1]$. Then the function $f_h : \mathbb{T} \times \{-1,1\} \to [0,1]$ defined by $(x,\epsilon) \mapsto f(x) + \epsilon h(x)$ gives a system that, when restricted to each copy of $\mathbb{Z}$, is just $\mathbf{P}^f$, but has correlations between the two copies that are given by h (the case $h = \mathbf{0}$ gives the independent joining). We can obtain slow decay of correlations between the two copies and negative associations in the joining itself; this is perhaps something that is not easy to construct directly.

§2. Background.

We first quickly review the probability measures studied in Lyons (2003b); see that paper for complete details.

Let E be a finite or countable set and consider the complex Hilbert space $\ell^2(E)$. Given any closed subspace $H \subseteq \ell^2(E)$, let P_H denote the orthogonal projection onto H. There is a unique probability measure $\mathbf{P}^H$ on $2^E := \{0,1\}^E$ defined by

$$\mathbf{P}^H[\eta(e_1) = 1, \ldots, \eta(e_k) = 1] = \det[(P_H e_i, e_j)]_{1 \le i,j \le k} \tag{2.1}$$

for all $k \ge 1$ and any set of distinct $e_1, \ldots, e_k \in E$; see, e.g., Lyons (2003b) or Daley and Vere-Jones (1988), Exercises 5.4.7–5.4.8. On the right-hand side, we are identifying each $e \in E$ with the element of $\ell^2(E)$ that is 1 in coordinate e and 0 elsewhere. In case H is finite-dimensional, then $\mathbf{P}^H$ is concentrated on subsets of E of cardinality equal to the dimension of H.

More generally, let Q be a positive contraction, meaning that Q is a self-adjoint operator on $\ell^2(E)$ such that for all $u \in \ell^2(E)$, we have $0 \le (Qu, u) \le (u, u)$. There is a unique probability measure $\mathbf{P}^Q$ such that

$$\mathbf{P}^Q[\eta(e_1) = 1, \ldots, \eta(e_k) = 1] = \det[(Q e_i, e_j)]_{i,j \le k} \tag{2.2}$$

for all $k \ge 1$ and distinct $e_1, \ldots, e_k \in E$. When Q is the orthogonal projection onto a closed subspace H, then $\mathbf{P}^Q = \mathbf{P}^H$. In fact, properties of $\mathbf{P}^Q$ can be deduced from the special case of orthogonal projections. Since this will be useful for our analysis, we review this reduction procedure.

Note first that uniqueness follows from the fact that (2.2) determines all finite-dimensional marginals via the inclusion-exclusion theorem. Indeed, we have the following formula for any disjoint pair of finite sets $A, B \subseteq E$ (see, e.g., Lyons (2003b)):

$$\mathbf{P}^Q[\eta\restriction A \equiv 1,\, \eta\restriction B \equiv 0] = \det\Big[\big(\mathbf{1}_B(e)e + (-1)^{\mathbf{1}_B(e)} Qe, e'\big)\Big]_{e,e' \in A \cup B}\,. \tag{2.3}$$

To show existence, let P_H be any orthogonal projection that is a dilation of Q. This means that H is a closed subspace of $\ell^2(E')$ for some $E' \supseteq E$ and that for all $u \in \ell^2(E)$, we have $Qu = P_{\ell^2(E)} P_H u$, where we regard $\ell^2(E') = \ell^2(E) \oplus \ell^2(E' \setminus E)$. (In this case, Q is also called the compression of P_H to $\ell^2(E)$.) The existence of a dilation is standard and is easily constructed; see, e.g., Lyons (2003b). Having chosen a dilation, we simply define $\mathbf{P}^Q$ as the law of η restricted to E when η has the law $\mathbf{P}^H$. Then (2.2) is a special case of (2.1).

A probability measure $\mathbf{P}$ on 2^E is said to have **negative associations** if for all pairs A and B of increasing events that are measurable with respect to the values of η on disjoint subsets of E, we have that A and B are negatively correlated with respect to $\mathbf{P}$. The following **conditional negative association** (CNA) property is proved in Lyons (2003b) and a consequence of this (see Proposition 2.6) will be used frequently here.

Theorem 2.1. *If Q is any positive contraction on $\ell^2(E)$, A is a finite subset of E, and $\eta_0 \in 2^A$, then $\mathbf{P}^Q[\,\cdot \mid \eta{\restriction}A = \eta_0]$ has negative associations.*

We now assume that $E = \mathbb{Z}^d$. Then the group structure of $\mathbb{Z}^d$ allows $\mathbb{Z}^d$ to act naturally on $\ell^2(\mathbb{Z}^d)$ and on $2^{\mathbb{Z}^d}$. The proof of the following lemma is straightforward and therefore skipped.

Lemma 2.2. *If Q is a $\mathbb{Z}^d$-equivariant positive contraction on $\ell^2(\mathbb{Z}^d)$, then $\mathbf{P}^Q$ is also $\mathbb{Z}^d$-invariant.*

As is well known, there exists a complex Hilbert-space isomorphism between $L^2(\mathbb{T}^d, \lambda_d)$ and $\ell^2(\mathbb{Z}^d)$, where $\mathbb{T}^d$ is the d-dimensional torus $\mathbb{R}^d/\mathbb{Z}^d$ and λ_d is unit Lebesgue measure on $\mathbb{T}^d$. This isomorphism is given by the Fourier transform $f \mapsto \widehat{f}$, where for $f \in L^2(\mathbb{T}^d, \lambda_d)$, we have $\widehat{f}(k) = \int_{\mathbb{T}^d} f(x) e^{-2\pi i k\cdot x}\, d\lambda_d(x)$ for $k \in \mathbb{Z}^d$. If $\mathbf{e}_k$ denotes the function $x \mapsto e^{2\pi i k\cdot x}$, then the isomorphism takes the set $\{\mathbf{e}_k\,;\ k \in \mathbb{Z}^d\}$ to the standard basis for $\ell^2(\mathbb{Z}^d)$. From now on, we shall abbreviate $L^2(\mathbb{T}^d, \lambda_d)$ by $L^2(\mathbb{T}^d)$.

The following is well known.

Theorem 2.3.

(i) *Let $A \subseteq \mathbb{T}^d$ be measurable and consider the operator $T_A : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ given by*

$$T_A(g) = g\mathbf{1}_A\,.$$

Then these projections (as A varies over the measurable subsets of $\mathbb{T}^d$) correspond (via the Fourier isomorphism) to the $\mathbb{Z}^d$-equivariant projections on $\ell^2(\mathbb{Z}^d)$.

(ii) *More generally, let $f : \mathbb{T}^d \to [0,1]$ be measurable and consider the operator $M_f : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ given by*

$$M_f(g) = fg\,.$$

Then these positive contractions (as f varies over the measurable functions from $\mathbb{T}^d$ to $[0,1]$) correspond (via the Fourier isomorphism) to the $\mathbb{Z}^d$-equivariant positive contractions on $\ell^2(\mathbb{Z}^d)$. More specifically, M_f corresponds to convolution with $\widehat{f}$.

As in the above theorem, an $f : \mathbb{T}^d \to [0,1]$ yields a $\mathbb{Z}^d$-equivariant positive contraction Q_f on $\ell^2(\mathbb{Z}^d)$, which in turn yields a translation-invariant probability measure $\mathbf{P}^{Q_f}$ on $2^{\mathbb{Z}^d}$, which we denote more simply by $\mathbf{P}^f$.

Lemma 2.4. *Given $f : \mathbb{T}^d \to [0,1]$ measurable and $e_1, \ldots, e_k \in \mathbb{Z}^d$,*

$$\mathbf{P}^f[\eta(e_1) = 1, \ldots, \eta(e_k) = 1] = \det[\widehat{f}(e_j - e_i)]_{1 \le i,j \le k} \,.$$

Proof. By definition, the left-hand side is $\det[(Q_f e_i, e_j)]_{1 \le i,j \le k}$. By Theorem 2.3(ii),

$$(Q_f e_i, e_j) = \big(M_f \mathbf{e}_{e_i}, \, \mathbf{e}_{e_j}\big) = \widehat{f}(e_j - e_i) \,. \qquad \blacksquare$$

Remark 2.5. Lemma 2.4 says that for $d = 1$, the probability of having 1s on some finite collection of elements of $\mathbb{Z}$ is a particular minor of the Toeplitz matrix associated to f.

Equation (2.3) shows a symmetry of $\mathbf{P}^f$ and $\mathbf{P}^{\mathbf{1}-f}$, namely, if η has the distribution $\mathbf{P}^f$, then $\mathbf{1} - \eta$ has the distribution $\mathbf{P}^{\mathbf{1}-f}$. Shirai and Takahashi (2003) prove the existence of $\mathbf{P}^f$ by a different method and also note this symmetry.

Although the last lemma gives us a formula for $\mathbf{P}^f$ directly in terms of f without reference to any projections, it is still useful to know a specific projection of which M_f is a compression. Let $f : \mathbb{T}^d \to [0,1]$ be measurable. Identifying $\mathbb{T}^{d+1}$ with $\mathbb{T}^d \times [0,1)$, we let $A_f \subseteq \mathbb{T}^{d+1}$ be the set $\{(x,y) \in \mathbb{T}^{d+1} \, ; \; y \le f(x)\}$. Consider the projection T_{A_f} of $L^2(\mathbb{T}^{d+1})$ given in Theorem 2.3(i). We view $L^2(\mathbb{T}^d)$ as a subspace of $L^2(\mathbb{T}^{d+1})$ by identifying $g \in L^2(\mathbb{T}^d)$ with $g \otimes \mathbf{1} \in L^2(\mathbb{T}^{d+1})$, where $(g \otimes \mathbf{1})(x,y) := g(x)$ for $x \in \mathbb{T}^d$, $y \in \mathbb{T}$. The orthogonal projection P of $L^2(\mathbb{T}^{d+1})$ onto $L^2(\mathbb{T}^d)$ is then given by $g \mapsto (x \mapsto \int_{\mathbb{T}} g(x,y)\, dy)$. A very simple calculation, left to the reader, shows that M_f is a compression of T_{A_f}; i.e.,

$$M_f = P T_{A_f} \tag{2.4}$$

on $L^2(\mathbb{T}^d)$ viewed as a subspace of $L^2(\mathbb{T}^{d+1})$. For later use, let

$$H_f := \{g \in L^2(\mathbb{T}^{d+1}) \, ; \; g = 0 \text{ a.e. on } (A_f)^c\}$$

be the image of T_{A_f}.

We next remind the reader of the notion of stochastic domination between two probability measures on 2^E. First, if η, δ are elements of 2^E, we write $\eta \preccurlyeq \delta$ if $\eta(e) \le \delta(e)$ for all $e \in E$. A subset A of 2^E is called **increasing** if $\eta \in A$ and $\eta \preccurlyeq \delta$ imply that $\delta \in A$. If ν and μ are two probability measures on 2^E, we write $\nu \preccurlyeq \mu$ if $\nu(A) \le \mu(A)$ for all increasing sets A. A theorem of Strassen (1965) says that this is equivalent to the existence of a probability measure m on $2^E \times 2^E$ that has ν and μ as its first and second marginals (i.e., m is a **coupling** of ν and μ) and such that m is concentrated on the set $\{(\eta, \delta) : \eta \preccurlyeq \delta\}$ (i.e., m is **monotone**).

Throughout the paper, we shall use the following consequence of conditional negative association (Theorem 2.1), sometimes called **joint negative regression dependence** (see Pemantle (2000)).

Proposition 2.6. *Assume that $\langle X_i \rangle_{i \in I}$ has conditional negative association, I is the disjoint union of I_1 and I_2, and $a, b \in \{0,1\}^{I_1}$ with $a_i \le b_i$ for each $i \in I_1$. Then*

$$[\langle X_i \rangle_{i \in I_2} \mid X_i = b_i, i \in I_1] \preccurlyeq [\langle X_i \rangle_{i \in I_2} \mid X_i = a_i, i \in I_1],$$

where $[Y \mid A]$ stands for the law of Y conditional on A.

Lemma 2.7. *Let $f_1, f_2 : \mathbb{T}^d \to [0,1]$ with $f_1 \le f_2$ a.e. Then $\mathbf{P}^{f_1} \preccurlyeq \mathbf{P}^{f_2}$.*

Proof. This follows from a more general result (see Lyons (2003b)) that says that if Q_1 and Q_2 are two commuting positive contractions on $\ell^2(E)$ such that $Q_1 \le Q_2$ in the sense that $Q_2 - Q_1$ is positive, then $\mathbf{P}^{Q_1} \preccurlyeq \mathbf{P}^{Q_2}$. However, here is a more concrete proof in our case. Since $f_1 \le f_2$ a.e., it follows that $H_{f_1} \subseteq H_{f_2}$, which implies by Theorem 6.2 in Lyons (2003b) that the projection measures $\mathbf{P}^{H_{f_1}}$ and $\mathbf{P}^{H_{f_2}}$ on $\mathbb{Z}^{d+1}$ satisfy $\mathbf{P}^{H_{f_1}} \preccurlyeq \mathbf{P}^{H_{f_2}}$ and hence their restrictions to $\mathbb{Z}^d$ satisfy the same relationship; i.e., $\mathbf{P}^{f_1} \preccurlyeq \mathbf{P}^{f_2}$. ∎

We close this section with a key representation of certain conditional probabilities and an application. The minimum in (2.5) below is often referred to as a Szegő infimum. For an infinite set $B \subseteq \mathbb{Z}^d \setminus \{\mathbf{0}\}$, write

$$\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction B \equiv 1] := \lim_{n \to \infty} \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction B_n \equiv 1],$$

where B_n is any increasing sequence of finite subsets of B whose union is B. This is a decreasing limit by virtue of Proposition 2.6. Let $(\cdot, \cdot)_f$ denote the usual inner product in the complex Hilbert space $L^2(f)$. For any set $B \subset \mathbb{Z}^d$, write $[B]_f$ for the closure in $L^2(f)$ of the linear span of the complex exponentials $\{\mathbf{e}_k ;\ k \in B\}$.

Theorem 2.8. *Let $f : \mathbb{T}^d \to [0,1]$ be measurable and $B \subset \mathbb{Z}^d$ with $\mathbf{0} \notin B$. Then*

$$\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction B \equiv 1] = \min\left\{ \int |\mathbf{1} - u|^2 f \, d\lambda_d \,;\ u \in [B]_f \right\}. \tag{2.5}$$

Proof. It suffices to prove the theorem for B finite since the infinite case then follows by a simple limiting argument. So assume that B is finite. Note that $\widehat{f}(k-j) = (\mathbf{e}_j,\, \mathbf{e}_k)_f$, so that

$$\mathbf{P}^f[\eta\restriction B \equiv 1] = \det[(\mathbf{e}_j,\, \mathbf{e}_k)_f]_{j,k\in B}\,,$$

and similarly for $\mathbf{P}^f[\eta\restriction(B \cup \{\mathbf{0}\}) \equiv 1]$. Thus, the left-hand side of (2.5) is a quotient of determinants. The fact that such a quotient has the form of the right-hand side is sometimes called Gram's formula. We include the proof for the convenience of the reader. Since $\mathbf{e_0} = \mathbf{1}$, it follows by row operations on the matrix $[(\mathbf{e}_j,\, \mathbf{e}_k)]_{j,k\in B\cup\{\mathbf{0}\}}$ that

$$\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction B \equiv 1] = \|P^\perp_{[B]_f}\mathbf{1}\|_f^2\,, \tag{2.6}$$

where $P^\perp_{[B]_f}$ denotes orthogonal projection onto the orthogonal complement of $[B]_f$ in $L^2(f)$. Since this is the squared distance from $\mathbf{1}$ to $[B]_f$, the equation (2.5) now follows. ∎

An extension of the above reasoning, given in Lyons (2003b), provides the entire conditional probability measure:

Theorem 2.9. *Let $f : \mathbb{T}^d \to [0,1]$ be measurable and $B \subset \mathbb{Z}^d$. Then the law of $\eta\restriction(\mathbb{Z}^d \setminus B)$ conditioned on $\eta\restriction B \equiv 1$ is the determinantal probability measure corresponding to the positive contraction on $\ell^2(\mathbb{Z}^d \setminus B)$ whose (j,k)-matrix entry is*

$$\big(P^\perp_{[B]_f}\mathbf{e}_j,\; P^\perp_{[B]_f}\mathbf{e}_k\big)_f$$

for $j, k \notin B$.

The Szegő infimum that appears in Theorem 2.8 involves trigonometric approximation, a classical area that has strong connections to the topics of prediction and interpolation for wide-sense stationary processes. We briefly discuss these topics now. In later sections, we shall describe more explicit connections to our results. Recall that a mean-0 wide-sense stationary process is a (not necessarily stationary) process $\langle Y_n\rangle_{n\in\mathbb{Z}^d}$ for which all the variables have finite variance and mean 0, and such that for each $k \in \mathbb{Z}^d$, the covariance $\mathrm{Cov}(Y_{n+k}, Y_n) = \mathbf{E}[Y_{n+k}\overline{Y_n}]$ does not depend on n. There is then a positive measure G (called the spectral measure) on $\mathbb{T}^d$ satisfying

$$\widehat{G}(k) = \mathrm{Cov}(Y_{n+k}, Y_n)$$

for $n, k \in \mathbb{Z}^d$. It turns out that for a one-dimensional wide-sense stationary process, if G is absolutely continuous with density g, then $\mathsf{GM}(g) = 0$ iff perfect linear prediction is possible, which means that Y_0 is in the closed linear span of $\{Y_n\}_{n\le -1}$ in $L^2(\Omega)$, where

Ω is the underlying probability space. This was proved in various versions by Szegő, Kolmogorov and Kreĭn. Since the covariance with respect to $\mathbf{P}^f$ of $\eta(\mathbf{0})$ and $\eta(k)$ for $k \in \mathbb{Z}^d$ is also given by a Fourier coefficient (namely $-|\widehat{f}(k)|^2$ for $k \neq \mathbf{0}$) and since, as we shall see, the geometric mean of f will play an important role in classifying the behavior of $\mathbf{P}^f$ as well, one might wonder about the relationship between our determinantal processes and wide-sense stationary processes. It is not hard to show that $\mathbf{P}^f$ (viewed as a wide-sense stationary process) has a spectral measure G that is absolutely continuous with density g given by the formula

$$g := \widehat{f}(\mathbf{0})\mathbf{1} - f * \widetilde{f},$$

where $*$ denotes convolution and $\widetilde{f}(t) := f(1-t)$. Other than the trivial cases $f = \mathbf{0}$ and $f = \mathbf{1}$, it is easy to check that g is bounded away from 0 and so, in particular, its geometric mean is always strictly positive. This suggests that our results are perhaps not so connected to prediction and interpolation. However, it turns out that some of the questions concerning $\mathbf{P}^f$ that we deal with here (such as phase multiplicity and domination) do have interpretations in terms of prediction and interpolation for wide-sense stationary processes whose spectral density is f (which does not include $\mathbf{P}^f$). More specifics will be given in the relevant sections.

Special attention is often devoted to stationary Gaussian processes, one reason being that their distribution is determined entirely by their spectral measure. It is known that a stationary Gaussian process with no deterministic component is a multistep Markov chain iff its spectral density is the reciprocal of a trigonometric polynomial; see Doob (1944). The analogous property for determinantal processes is regeneration:

Proposition 2.10. *If $d = 1$, then f is the reciprocal of a trigonometric polynomial of degree at most n iff $\mathbf{P}^f$ is a regenerative process that regenerates after n successive* 1*s appear.*

This last property means that for any k, given that $\eta\restriction[k+1, k+n] \equiv 1$, the future, $\eta\restriction[k+n+1, \infty)$, is conditionally independent of the past, $\eta\restriction(-\infty, k]$.

Proof. Note first that because of Theorem 2.9, this regenerative property holds for $\mathbf{P}^f$ iff for all $j \geq 0$ and all $C \subset (-\infty, -n-1]$, we have $P^{\perp}_{[B]_f}\mathbf{e}_j = P^{\perp}_{[B\cup C]_f}\mathbf{e}_j$, where $B := [-n, -1]$. (Here, we are relying on the fact that $\|P^{\perp}_{[B]_f}\mathbf{e}_j\|_f > \|P^{\perp}_{[B\cup C]_f}\mathbf{e}_j\|_f$ if the vectors are not equal.) This is the same as $P^{\perp}_{[B]_f}\mathbf{e}_j \perp [C]_f$ for all $j \geq 0$, or, in other words, there exists some $u_j \in [B]_f$ with $\mathbf{e}_j - u_j \perp [B \cup C]_f$. As this would have to hold for all C (and u_j is independent of C), it is the same as the existence of some $u_j \in [B]_f$ such that $T_j := (\mathbf{e}_j - u_j)f$ is analytic, i.e., $\widehat{T_j}(k) = 0$ for all $k < 0$. Now this implies that

$\overline{T_0} = (\mathbf{1} - \overline{u_0})f$, whence $T_0(\mathbf{1} - \overline{u_0}) = \overline{T_0}(\mathbf{1} - u_0)$. The left side of this last equation is an analytic function, while the right side is the conjugate of an analytic function (i.e., its Fourier coefficients vanish on $\mathbb{Z}^+$). Therefore, both equal some constant, c. Hence

$$f = \frac{T_0}{\mathbf{1} - u_0} = \frac{c}{(\mathbf{1} - u_0)(\mathbf{1} - \overline{u_0})} = \frac{c}{|\mathbf{1} - u_0|^2}\,,$$

which is indeed the reciprocal of a trigonometric polynomial of degree at most n.

Conversely, if $1/f$ is the reciprocal of a trigonometric polynomial of degree at most n, then since $f \geq 0$, the theorem of Fejér and Riesz (see Grenander and Szegő (1984), p. 20) allows us to write $f = c/|\mathbf{1} - u_0|^2$ for some conjugate-analytic polynomial $u_0 \in [B]_f$ such that the analytic extension of $\mathbf{1} - \overline{u_0}$ to the unit disc has no zeroes. We may rewrite this as $(\mathbf{1} - u_0)f = T_0$ for $T_0 := c/(\mathbf{1} - \overline{u_0})$. Since T_0 has an extension to the unit disc as the reciprocal of an analytic polynomial with no zeroes, it follows that T_0 is also analytic. Multiplying both sides of this equation by $\mathbf{e}_1$ and rewriting $\mathbf{e}_1 u_0 f = c_1(\mathbf{1} - u_1')f = c_1(u_0 - u_1')f + c_1 T_0$ for some constant c_1 and some $u_1' \in [B]_f$, we see that $(\mathbf{e}_1 - u_1)f = T_1$ for $u_1 := c_1(u_0 - u_1') \in [B]_f$ and $T_1 := \mathbf{e}_1 T_0 + c_1 T_0$, an analytic function. Similarly, we may establish by induction that for each $j \geq 0$, there is some $u_j \in [B]_f$ such that $T_j := (\mathbf{e}_j - u_j)f$ is analytic. This proves the equivalence desired. ∎

We now use this proof to establish the explicit probabilistic form of the renewal process in Example 1.7. In this case, $B = \{-1\}$ and one easily verifies that $u_j = a^{j+1}\mathbf{e}_{-1}$, as is standard in the theory of linear prediction. Therefore,

$$\mathbf{P}^f[\eta(j) = 1 \mid \eta(-1) = 1] = \|P^{\perp}_{[B]_f}\mathbf{e}_j\|_f^2 = \|\mathbf{e}_j - u_j\|_f^2 = \|\mathbf{e}_j\|_f^2 - \|u_j\|_f^2 = \frac{1-a}{1+a}(1 - a^{2j+2})\,.$$

It now suffices to verify that this is also true for the explicit renewal process described in Example 1.7. First, it is well known from basic renewal theory that for a renewal process, the probabilities $\mathbf{P}^f[\eta(j) = 1 \mid \eta(-1) = 1]$ determine the distribution of the number of 0s between two 1s. Hence to verify the above statement, one simply needs to check that these latter probabilities are related to the interrenewal distribution via the appropriate convolution-type equation. In this case, it comes down to verifying that for all $j \geq 0$,

$$\frac{1-a}{1+a}(1 - a^{2j+2}) = (j+1)(1-a)^2 a^j + \frac{1-a}{1+a}\sum_{k=1}^{j} k(1-a)^2 a^{k-1}(1 - a^{2(j-k)+2})\,.$$

This identity is easy to check.

§3. The Bernoulli Shift Property.

In this section, we prove that the stationary determinantal processes studied here are Bernoulli shifts. We assume the reader is familiar with the basic notions of a Bernoulli shift (see, e.g., Ornstein (1974)).

Theorem 3.1. *Let $f : \mathbb{T}^d \to [0,1]$ be measurable. Then $\mathbf{P}^f$ is a Bernoulli shift; i.e., it is isomorphic (in the sense of ergodic theory) to an i.i.d. process.*

Before beginning the proof, we present a few preliminaries. We first recall the definition of the $\overline{d}$-metric.

Definition 3.2. If μ and ν are $\mathbb{Z}^d$-invariant probability measures on $2^{\mathbb{Z}^d}$, then

$$\overline{d}(\mu,\nu) := \inf_m m\Big[\big\{(\eta,\delta) \in 2^{\mathbb{Z}^d} \times 2^{\mathbb{Z}^d}\ ;\ \eta(\mathbf{0}) \ne \delta(\mathbf{0})\big\}\Big]\,,$$

where the infimum is taken over all couplings m of μ and ν that are $\mathbb{Z}^d$-invariant.

The following is a slight generalization of a well-known result (see, for example, page 75 in Liggett (1985)). The proof is an immediate consequence of the existence of a monotone coupling and is left to the reader.

Lemma 3.3. *Suppose that σ_1, σ_2, ν and μ are $\mathbb{Z}^d$-invariant probability measures on $2^{\mathbb{Z}^d}$ such that $\nu \preccurlyeq \sigma_i \preccurlyeq \mu$ for both $i = 1, 2$. Then*

$$\overline{d}(\sigma_1,\sigma_2) \le \mu[\eta(\mathbf{0}) = 1] - \nu[\eta(\mathbf{0}) = 1]\,.$$

Proposition 3.4. *Let $f, g : \mathbb{T}^d \to [0,1]$ be measurable. Then*

$$\overline{d}(\mathbf{P}^f,\mathbf{P}^g) \le \int_{\mathbb{T}^d} |f - g|\, d\lambda_d\,.$$

Proof. Because of Lemma 2.7, we may apply Lemma 3.3 to $\sigma_1 := \mathbf{P}^f$, $\sigma_2 := \mathbf{P}^g$, $\nu := \mathbf{P}^{f\wedge g}$, and $\mu := \mathbf{P}^{f\vee g}$. We obtain that

$$\overline{d}(\mathbf{P}^f,\mathbf{P}^g) \le \int_{\mathbb{T}^d} f\vee g\, d\lambda_d - \int_{\mathbb{T}^d} f\wedge g\, d\lambda_d = \int_{\mathbb{T}^d} |f-g|\, d\lambda_d\,. \qquad \blacksquare$$

Proof of Theorem 3.1. The first step is to approximate f by trigonometric polynomials. Let K_r be the rth Fejér kernel for $\mathbb{T}$,

$$K_r := \sum_{|j|\le r} \Big(1 - \frac{|j|}{r+1}\Big)\mathbf{e}_j\,,$$

and define $K_r^d(x_1,\ldots,x_d) := \prod_{i=1}^d K_r(x_i)$. It is well known that K_r^d is a positive summability kernel, so that if we define g_r by

$$g_r := f * K_r^d\,,$$

then $0 \le g_r \le 1$ and $\lim_{r\to\infty} g_r = f$ a.e. and in $L^1(\mathbb{T}^d)$.

Next, since each K_r^d is a trigonometric polynomial, so is each g_r. From this and Lemma 2.4, it is easy to see that there is a constant C such that $\mathbf{P}^{g_r}(A\cap B) = \mathbf{P}^{g_r}(A)\,\mathbf{P}^{g_r}(B)$ if A is of the form $\eta \equiv 1$ on S_1 and B is of the form $\eta \equiv 1$ on S_2 with S_1 and S_2 being finite sets and having distance at least C between them. From this, a standard argument in probability (a π-λ argument) shows that if A is any event depending only on locations S_1 and if B is any event depending only on locations S_2 with S_1 and S_2 possibly infinite sets having distance at least C between them, then $\mathbf{P}^{g_r}(A\cap B) = \mathbf{P}^{g_r}(A)\,\mathbf{P}^{g_r}(B)$. A process with this property is called a **finitely dependent** process and this property implies it is a Bernoulli shift (e.g., the so-called very weak Bernoulli condition is immediately verified; see Ornstein (1974) for this definition for $d = 1$ and, e.g., Steif (1991) for general d).

Since the processes that are Bernoulli shifts are closed in the $\overline{d}$ metric (see Ornstein (1974)) and Proposition 3.4 tells us that $\lim_{r\to\infty} \overline{d}(\mathbf{P}^{g_r}, \mathbf{P}^f) = 0$, we conclude that $\mathbf{P}^f$ is a Bernoulli shift. ∎

Remark 3.5. An important property of 1-dimensional stationary processes is the weak Bernoulli (WB) property. This is also referred to as "β-mixing" and "absolute regularity" in the literature. Despite its name, it is known that WB is strictly stronger than Bernoullicity. It is easy to check that "aperiodic" regenerative processes and finitely dependent processes are WB. Hence, by our earlier results, if f is a trigonometric polynomial or the inverse of a trigonometric polynomial (or if $\mathbf{1} - f$ is the inverse of a trigonometric polynomial), then $\mathbf{P}^f$ is WB. This is subsumed by the independent work of Shirai and Takahashi (2003), who showed that $\mathbf{P}^f$ is WB whenever $\sum_{n\ge1} n|\widehat{f}(n)|^2 < \infty$. The precise class of f for which $\mathbf{P}^f$ is WB is not known. We also note that it follows from Smorodinsky (1992) that if f is a trigonometric polynomial, then $\mathbf{P}^f$ is finitarily isomorphic to an i.i.d. process.

§4. Support of the Measures.

In this section, we show that all of the probability measures $\mathbf{P}^f$ have full support except in two degenerate cases. We begin with the following lemma.

Lemma 4.1. *Let H be a closed subspace of $\ell^2(E)$ and let A and B be finite disjoint subsets of E. Then*

$$\mathbf{P}^H[\eta(e) = 1 \text{ for } e \in A,\ \eta(e) = 0 \text{ for } e \in B] > 0$$

if and only if $\{P_H(e)\}_{e\in A} \cup \{P_{H^\perp}(e)\}_{e\in B}$ is linearly independent.

Proof. In the special case that A or B is empty, the result follows from Lyons (2003b). In general, $\{P_H(e)\}_{e\in A} \cup \{P_{H^\perp}(e)\}_{e\in B}$ is linearly independent if and only if $\{P_H(e)\}_{e\in A}$ and $\{P_{H^\perp}(e)\}_{e\in B}$ are each linearly independent sets. Also, $\mathbf{P}^H[\eta\restriction A \equiv 1,\ \eta\restriction B \equiv 0] > 0$ if and only if $\mathbf{P}^H[\eta\restriction A \equiv 1] > 0$ and $\mathbf{P}^H[\eta\restriction B \equiv 0] > 0$ since

$$\mathbf{P}^H[\eta\restriction A \equiv 1,\ \eta\restriction B \equiv 0] \ge \mathbf{P}^H[\eta\restriction A \equiv 1]\,\mathbf{P}^H[\eta\restriction B \equiv 0]$$

by the negative association property, Theorem 2.1. Thus, the general case follows from the special case. ∎

Theorem 4.2. *$\mathbf{P}^f$ has full support for any function $f : \mathbb{T}^d \to [0,1]$ other than $\mathbf{0}$ or $\mathbf{1}$.*

Proof. Since marginals of probability measures with full support clearly have full support, it suffices to prove the result for $\mathbf{P}^H$ when H is a closed $\mathbb{Z}^d$-invariant subspace of $\ell^2(\mathbb{Z}^d)$ other than $\ell^2(\mathbb{Z}^d)$ or 0. If we translate over to $L^2(\mathbb{T}^d)$, we see, by Lemma 4.1, that it suffices to show that if $A \subseteq \mathbb{T}^d$ with $0 < \lambda_d(A) < 1$ and $n_1, \ldots, n_k, m_1, \ldots, m_\ell$ are all distinct elements of $\mathbb{Z}^d$, then

$$\{\mathbf{e}_{n_j}\mathbf{1}_A\}_{1\le j\le k} \cup \{\mathbf{e}_{m_r}\mathbf{1}_{A^c}\}_{1\le r\le \ell}$$

is linearly independent. Suppose that $c_1, \ldots, c_k, d_1, \ldots, d_\ell$ are complex numbers such that

$$\sum_{j=1}^{k} c_j \mathbf{e}_{n_j}\mathbf{1}_A + \sum_{r=1}^{\ell} d_r \mathbf{e}_{m_r}\mathbf{1}_{A^c} = \mathbf{0}\,.$$

From this it follows that $\sum_{j=1}^k c_j \mathbf{e}_{n_j} = 0$ a.e. on A. Since $\lambda_d(A) > 0$, we have that $\sum_{j=1}^k c_j \mathbf{e}_{n_j}$ is 0 on a set of positive measure. It is well known that this implies that $c_1, \ldots, c_k$ vanish (the proof uses induction on d and Fubini's theorem). Similarly, $d_1, \ldots, d_\ell$ vanish. ∎

§5. Domination Properties.

In this section, we study the question of which product measures are stochastically dominated by $\mathbf{P}^f$ and which product measures stochastically dominate $\mathbf{P}^f$. For simplicity, we give our first results for $d = 1$, and only afterwards describe how these results extend to higher dimensions. We also treat a different notion of "full domination" at the end of the section. Recall that for $f : \mathbb{T}^d \to [0, 1]$, we define

$$\mathsf{GM}(f) := \exp \int_{\mathbb{T}^d} \log f \, d\lambda_d \, .$$

We introduce an auxiliary stronger notion of domination than $\preccurlyeq$, but only in the case where one of the measures is a product measure. Let μ_p denote product measure with density p.

Definition 5.1. A stationary process $\{\eta_n\}_{n \in \mathbb{Z}}$ with distribution ν **strongly dominates** μ_p, written $\mu_p \preccurlyeq_{\rm s} \nu$, if for any n and any $a_1, \ldots, a_n \in \{0, 1\}$,

$$\mathbf{P}[\eta_0 = 1 \mid \eta_i = a_i, i = 1, \ldots, n] \geq p \, .$$

Similarly, we define $\nu \preccurlyeq_{\rm s} \mu_p$ if the above inequality holds when $\geq$ is replaced by $\leq$.

The following lemma is easy and well known; it is sometimes referred to as Holley's lemma.

Lemma 5.2. *If $\mu_p \preccurlyeq_{\rm s} \nu$, then $\mu_p \preccurlyeq \nu$.*

The converse is not true, as we shall see later (Remark 6.4). In light of Lemma 2.7, we have that if $p \leq f \leq q$, then

$$\mu_p \preccurlyeq \mathbf{P}^f \preccurlyeq \mu_q \, .$$

The optimal improvement of these stochastic bounds is as follows.

Theorem 5.3. *For any $f : \mathbb{T} \to [0, 1]$, we have $\mu_p \preccurlyeq \mathbf{P}^f$ iff $p \leq \mathsf{GM}(f)$ iff $\mu_p \preccurlyeq_{\rm s} \mathbf{P}^f$. Similarly, $\mathbf{P}^f \preccurlyeq \mu_q$ iff $q \geq 1 - \mathsf{GM}(\mathbf{1} - f)$ iff $\mathbf{P}^f \preccurlyeq_{\rm s} \mu_q$. In addition, for any stationary process μ that has conditional negative association, we have $\mu_p \preccurlyeq \mu$ iff $\mu_p \preccurlyeq_{\rm s} \mu$.*

Proof. Let $d_n := D_{n-1}(f)$ be the probability of having n 1s in a row. According to Szegő's theorem (Grenander and Szegő (1984), pp. 44, 66), d_{n+1}/d_n is decreasing in n and

$$\lim_{n \to \infty} d_{n+1}/d_n = \mathsf{GM}(f) = \lim_{n \to \infty} d_n^{1/n} \, .$$

In particular, $d_{n+1}/d_n \geq \mathsf{GM}(f)$ for all n.

Proposition 2.6 implies that for any fixed n,

$$\mathbf{P}^f[\eta_0 = 1 \mid \eta_i = a_i, i = 1, \ldots, n]$$

is minimized among all $a_1, \ldots, a_n \in \{0, 1\}$ when $a_1 = a_2 = \cdots = a_n = 1$. In this case, the value is d_{n+1}/d_n. Since this is at least $\mathsf{GM}(f)$, we deduce that if $p \le \mathsf{GM}(f)$, then $\mu_p \preccurlyeq_{\rm s} \mathbf{P}^f$ and hence that $\mu_p \preccurlyeq \mathbf{P}^f$.

Conversely, if $\mu_p \preccurlyeq \mathbf{P}^f$, then certainly $p^n \le d_n$ for all n. Hence $p \le \mathsf{GM}(f)$. The second-to-last statement can be proved in the same way or can be concluded by symmetry. Finally, the last statement can be proved in a similar fashion. ∎

Remark 5.4. The above theorem gives us two interesting examples. First, if we take $f := \mathbf{1}_A$ where A has Lebesgue measure $1 - \epsilon$, then $\overline{d}(\mathbf{P}^f, \delta_{\mathbf{1}}) \le \epsilon$ by Proposition 3.4, but nonetheless, $\mathbf{P}^f$ does not dominate any nontrivial product measure since $\mathsf{GM}(f) = 0$. Second, if we take $f := \mathbf{1}_{[0,1/2]} + .4 \cdot \mathbf{1}_{[1/2,1]}$, then $f < 1/2$ on a set of positive measure, but nonetheless $\mathbf{P}^f$ dominates $\mathbf{P}^{1/2} = \mu_{1/2}$ since $\mathsf{GM}(f) = (.4)^{1/2} > 1/2$.

Example 5.5. Let f be as in (1.1), so that $\mathbf{P}^f$ is the law of the horizontal edges of the uniform spanning tree in the plane. In order to examine a 1-dimensional process, let us consider only the edges lying on the x-axis. If we let

$$g(x) := \int_{\mathbb{T}} f(x, y)\, d\lambda_1(y)\,,$$

then the edges lying on the x-axis have the law $\mathbf{P}^g$ since $\widehat{g}(k) = \widehat{f}(k, 0)$ for all $k \in \mathbb{Z}$. Since an antiderivative of $1/(1 + a\sin^2 \pi y)$ is

$$\frac{\arctan\left(\sqrt{1+a}\tan \pi y\right)}{\pi\sqrt{1+a}}\,,$$

we have

$$g(x) = \frac{\sin \pi x}{\sqrt{1 + \sin^2 \pi x}}\,,$$

as given in Example 1.3. We claim that $\mathbf{P}^g$ strongly dominates μ_p for $p := \sqrt{2} - 1$, and this is optimal. In order to show this, we calculate $\mathsf{GM}(g)$. Write $g_1(x) := \sin^2 \pi x$. Then $(\mathsf{GM}(g))^2 = \mathsf{GM}(g_1)/\mathsf{GM}(\mathbf{1} + g_1)$. Let G_1 be the harmonic extension of $\log g_1$ from the circle to the unit disc. Then $\int \log g_1\, d\lambda_1 = G_1(0)$ by the mean value property of harmonic functions. Since $g_1 = |(1 - \mathbf{e}_1)/2|^2$, we see that G_1 is the real part of the analytic function $z \mapsto 2\log[(1 - z)/2]$ in the unit disc, from which we conclude that $G_1(0) = \log(1/4)$. Therefore, $\mathsf{GM}(g_1) = 1/4$. Similarly, $\mathbf{1} + g_1 = |[\sqrt{2} + 1 - (\sqrt{2} - 1)\mathbf{e}_1]/2|^2$, whence $\mathsf{GM}(\mathbf{1}+g_1) = [(\sqrt{2}+1)/2]^2$. Therefore $\mathsf{GM}(g) = \sqrt{2}-1 = 0.4142^+$, as desired. It turns out that $q := 1 - \mathsf{GM}(\mathbf{1}-g) = 1 - 2(\sqrt{2}-1)e^{-2\mathrm{G}/\pi} = 0.5376^+$, where

$$\mathrm{G} := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2} = 0.9160^- \tag{5.1}$$

is **Catalan's constant**. Thus, $\mathbf{P}^g \preccurlyeq_{\mathrm{s}} \mu_q$. It is interesting how close p and q are.

We now introduce some new mixing conditions. Given a positive integer r, we may restrict $\mathbf{P}^f$ to $2^{r\mathbb{Z}}$. If we identify $r\mathbb{Z}$ with $\mathbb{Z}$, then we obtain the process $\mathbf{P}^{f_r}$, where

$$f_r(t) := \frac{1}{r} \sum_{x \in r^{-1}t} f(x)\,,$$

where $r^{-1}t := \{x \in \mathbb{T};\ rx = t\}$. The reason that this restriction is equal to $\mathbf{P}^{f_r}$ is that for all $k \in \mathbb{Z}$, we have $\widehat{f_r}(k) = \widehat{f}(rk)$, as is easy to check. Because of this relation, we have that $\int |f_r(t) - \widehat{f}(0)|^2\, d\lambda_1(t) \to 0$ as $r \to \infty$, so that by Proposition 3.4, it follows that $\overline{d}(\mathbf{P}^{f_r}, \mathbf{P}^{\widehat{f}(0)}) \to 0$ as $r \to \infty$. (It is not hard to show that a similar property holds for any Kolmogorov automorphism (see Definition 7.1), while the example in Remark 6.4 shows that this property can occur in other cases as well.) In fact, we often have a stronger convergence for our determinantal processes, as we show next.

Theorem 5.6. *Let $f : \mathbb{T} \to [0,1]$ be measurable. If $\mathsf{GM}(f) > 0$ or if f is bounded away from 0 on an interval of positive length, then there exist constants $p_r \to \widehat{f}(0)$ such that $\mathbf{P}^{f_r} \succcurlyeq \mathbf{P}^{p_r}$ for all r. Therefore, if $\mathsf{GM}\big(f(\mathbf{1}-f)\big) > 0$ or if f is continuous and not equal to $\mathbf{0}$ nor $\mathbf{1}$, then there exist constants $p_r, q_r \to \widehat{f}(0)$ such that $\mathbf{P}^{p_r} \preccurlyeq \mathbf{P}^{f_r} \preccurlyeq \mathbf{P}^{q_r}$ for all r.*

Of course, in view of Theorem 5.3, we use $p_r := \mathsf{GM}(f_r)$ and $q_r := 1 - \mathsf{GM}(\mathbf{1} - f_r)$. Thus, the theorem is an immediate consequence of the following lemma.

Lemma 5.7. *Let $f : \mathbb{T} \to [0,1]$ be measurable. If $\mathsf{GM}(f) > 0$ or f is bounded away from 0 on an interval of positive length, then $\mathsf{GM}(f_r) \to \widehat{f}(0)$ as $r \to \infty$.*

Proof. We have seen that $f_r \to \widehat{f}(0) \cdot \mathbf{1}$ in L^2, whence also in measure. Therefore $\log f_r \to (\log \widehat{f}(0)) \cdot \mathbf{1}$ in measure. Thus, it remains to show that $\{\log f_r\}$ is uniformly integrable (at least, for all large r). Suppose first that $\mathsf{GM}(f) > 0$.

Given $h \in L^1(\mathbb{T})$, write $h^{(r)}(t) := h(rt)$, so that $\widehat{h^{(r)}}(k)$ is 0 when k is not a multiple of r and is $\widehat{h}(k/r)$ when k is a multiple of r. For any $g, h \in L^2(\mathbb{T})$, we have

$$\begin{aligned}\int_{\mathbb{T}} g_r(t)\overline{h(t)}\, d\lambda_1(t) &= \sum_{k\in\mathbb{Z}} \widehat{g_r}(k)\overline{\widehat{h}(k)} = \sum_{k\in\mathbb{Z}} \widehat{g}(rk)\overline{\widehat{h}(k)} = \sum_{k\in\mathbb{Z}} \widehat{g}(k)\overline{\widehat{h^{(r)}}(k)} \\ &= \int_{\mathbb{T}} g(t)\overline{h^{(r)}(t)}\, d\lambda_1(t) = \int_{\mathbb{T}} g(t)\overline{h(rt)}\, d\lambda_1(t)\,. \end{aligned} \tag{5.2}$$

Therefore, for any set $A \subseteq \mathbb{T}$, we have

$$\int_A \log(f_r)\,d\lambda_1 \ge \int_A (\log f)_r\,d\lambda_1 = \int_{r^{-1}A} \log f\,d\lambda_1\,;$$

the inequality follows by concavity of log, while the equality follows from (5.2) applied to $g := \log f$ and $h := \mathbf{1}_A$. Given any $\epsilon \in \big(0, \widehat{f}(0)\big)$, let $A^r_\epsilon := \{t\,;\ f_r(t) < \epsilon\}$. Note that for any measurable $A \subset \mathbb{T}$, we have $\lambda_1(r^{-1}A) = \lambda_1(A)$. Since $f_r \to \widehat{f}(0) \cdot \mathbf{1}$ in measure, it follows that

$$\lim_{r\to\infty} \lambda_1(r^{-1}A^r_\epsilon) = \lim_{r\to\infty} \lambda_1(A^r_\epsilon) = 0\,.$$

Since $\log f$ is integrable, it follows that

$$\lim_{r\to\infty} \int_{r^{-1}A^r_\epsilon} \log f\,d\lambda_1 = 0\,.$$

This establishes the uniform integrability in the first case.

In the second case where $f \ge c > 0$ on an interval of length $\epsilon > 0$, we have that $f_r \ge c\epsilon$ on all of $\mathbb{T}$ for all $r > 1/\epsilon$. It is then obvious that $\{\log f_r\}$ is uniformly integrable for all $r > 1/\epsilon$. ∎

Remark 5.8. There are functions f with $f > 0$ a.e., yet $\mathsf{GM}(f_r) = 0$ for all r. For example, enumerate the rationals $\{x_j\,;\ j \ge 1\}$ in $(0,1)$ and choose $\epsilon_j > 0$ such that $x_j + \epsilon_j < 1$ and $\sum_j \epsilon_j < 1$. Define $A_n := [0,1] \setminus \bigcup_{j>n} [x_j, x_j + \epsilon_j]$ and $f_0(x) := e^{-1/x}$. Then one can show that

$$f(x) := f_0(x)\mathbf{1}_{A_0}(x) + \sum_{n=1}^{\infty} f_0(x - x_n)\mathbf{1}_{[x_n, x_n+\epsilon_n]\cap A_n}(x)$$

is such an example.

We turn next to the extension of the prior results to higher dimensions. This is basically straightforward, but has interesting applications, as we shall see. First, we recall the usual notion of the **past σ-field** $\mathsf{Past}(k)$ at a point $k \in \mathbb{Z}^d$. We use lexicographic ordering on $\mathbb{Z}^d$, i.e., we write $(k_1, k_2, \ldots, k_d) \prec (l_1, l_2, \ldots, l_d)$ if $k_i < l_i$ when i is the smallest index such that $k_i \ne l_i$. Then define $\mathsf{Past}(k)$ to be the σ-field generated by $\eta(j)$ for all $j \prec k$. More generally, the past σ-field could be defined with respect to any **ordering** of $\mathbb{Z}^d$, which means the selection of a set $\Pi \subset \mathbb{Z}^d$ that has the properties $\Pi \cup (-\Pi) = \mathbb{Z}^d \setminus \{\mathbf{0}\}$, $\Pi \cap (-\Pi) = \varnothing$, and $\Pi + \Pi \subset \Pi$. The associated ordering is that where $k \prec l$ iff $l - k \in \Pi$. $\mathsf{Past}_\Pi(u)$ will denote the past of u with respect to the ordering Π, i.e., $\{v\,;\ v \prec u\}$. Thus, $\mathsf{Past}_\Pi(\mathbf{0})$ is just $-\Pi$. (For a characterization of all orders, see Teh (1961), Zaĭceva (1953), or Trevisan (1953).) As before, let μ_p denote product measure with density p.

Definition 5.9. Given an ordering Π, a stationary process $\{\eta_n\}_{n\in\mathbb{Z}^d}$ with distribution ν **strongly dominates** μ_p, written $\mu_p \preccurlyeq_{\mathrm{s}} \nu$, if

$$\mathbf{P}[\eta_{\mathbf{0}} = 1 \mid \mathsf{Past}_\Pi(\mathbf{0})] \ge p \quad \nu\text{-a.s.}$$

Similarly, we define $\nu \preccurlyeq_{\mathrm{s}} \mu_p$ if the above inequality holds when $\ge$ is replaced by $\le$. (Note that the ordering Π here is suppressed in the notation.)

Although we have phrased it differently, in one dimension, this is equivalent to Definition 5.1 when $\Pi = \{-1, -2, \ldots\}$. Again, we have a version of Holley's lemma:

Lemma 5.10. *Given any ordering* Π*, if* $\mu_p \preccurlyeq_{\mathrm{s}} \nu$*, then* $\mu_p \preccurlyeq \nu$*.*

(Note that to produce a monotone coupling for a general ordering with respect to a set Π, it is enough to do so for the measures restricted to any finite $B \subset \mathbb{Z}^d$. Given such a B, order B by the restriction of $\prec$ to B and couple the measures by adding sites from B in this order.)

We now prove

Theorem 5.11. *Fix any ordering* Π*. For any measurable* $f : \mathbb{T}^d \to [0,1]$*, we have* $\mu_p \preccurlyeq \mathbf{P}^f$ *iff* $p \le \mathsf{GM}(f)$ *iff* $\mu_p \preccurlyeq_{\mathrm{s}} \mathbf{P}^f$*. Similarly,* $\mathbf{P}^f \preccurlyeq \mu_q$ *iff* $q \ge 1 - \mathsf{GM}(\mathbf{1} - f)$ *iff* $\mathbf{P}^f \preccurlyeq_{\mathrm{s}} \mu_q$*. In addition, for any stationary process* μ *that has conditional negative association, we have* $\mu_p \preccurlyeq \mu$ *iff* $\mu_p \preccurlyeq_{\mathrm{s}} \mu$*.*

Proof. Proposition 2.6 and Theorem 2.8 imply that

$$\begin{aligned}\operatorname{ess\,inf} \mathbf{P}^f\Big[\eta(\mathbf{0}) = 1 \mid \mathsf{Past}_\Pi(\mathbf{0})\Big] &= \inf\Big\{\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}A \equiv 1]\,;\ A \subset -\Pi \text{ is finite}\Big\}\\ &= \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}(-\Pi) \equiv 1]\\ &= \min\left\{\int |\mathbf{1} - T|^2 f\, d\lambda_d\,;\ T \in [-\Pi]_f\right\}.\end{aligned}$$

By Helson and Lowdenslager (1958), the latter quantity equals $\mathsf{GM}(f)$. Therefore, $p \le \mathsf{GM}(f)$ iff $\mu_p \preccurlyeq_{\mathrm{s}} \mathbf{P}^f$. On the other hand, if $\mu_p \preccurlyeq \mathbf{P}^f$, then let $A(r)$ be the finite subset of $-\Pi$ consisting of all points within some large radius r about $\mathbf{0}$. Let $a_1 \prec a_2 \prec \cdots \prec a_n$ be the elements of $A(r)$ in order, where $n := |A(r)|$, and write $A_j := A(r) \cap (-\Pi + a_j)$. Then

$$p^n \le \mathbf{P}^f[\eta{\restriction}A(r) \equiv 1] = \prod_{j=1}^n \mathbf{P}^f[\eta(a_j) = 1 \mid \eta{\restriction}A_j \equiv 1]\,.$$

Most terms in this product are quite close to $\mathsf{GM}(f)$, while all lie in $[\mathsf{GM}(f), 1]$. It follows that

$$\lim_{r\to\infty} \mathbf{P}^f[\eta{\restriction}A(r) \equiv 1]^{1/n} = \mathsf{GM}(f)\,,$$

so that $p \le \mathsf{GM}(f)$. Finally, the remaining statements can be proved in a similar fashion. ∎

Remark 5.12. It follows from Theorem 5.11 that the choice of ordering Π does not determine whether $\mu_p \preccurlyeq_{\rm s} \mathbf{P}^f$. This is not true for general stationary processes, even in one dimension. We are grateful to Olle Häggström for the following simple example. Let μ be the distribution on $\{0,1\}^2$ given by

$$\mu = (1/7)(\delta_{(0,0)} + \delta_{(0,1)}) + (2/7)\delta_{(1,0)} + (3/7)\delta_{(1,1)} \,.$$

Let $\eta \in \{0,1\}^{\mathbb{Z}}$ be such that with probability 1/2, (η_{2n}, η_{2n+1}) are chosen independently each with distribution μ and otherwise (η_{2n-1}, η_{2n}) are chosen independently each with distribution μ. Let ν be the law of η. If $\Pi := \mathbb{Z}^-$, then $\mu_p \preccurlyeq_{\rm s} \nu$ iff $p \le 1/2$, while if $\Pi := \mathbb{Z}^+$, then $\mu_p \preccurlyeq_{\rm s} \nu$ iff $p \le 4/7$.

Example 5.13. Let f be as in (1.1), so that $\mathbf{P}^f$ is the law of the horizontal edges of a uniform spanning tree in the square lattice $\mathbb{Z}^2$. By Kasteleyn (1961) or Montroll (1964), we have

$$\int_{\mathbb{T}^2} \log 4(\sin^2 \pi x + \sin^2 \pi y)\, d\lambda_2(x,y) = \frac{4{\rm G}}{\pi} = 1.1662^+ \,, \tag{5.3}$$

where G is again Catalan's constant (5.1). As we have shown in Example 5.5, $\mathsf{GM}(\sin^2 \pi x) = 1/4$, whence by (5.3), $\mathsf{GM}(f) = e^{-4{\rm G}/\pi} = \mathsf{GM}(\mathbf{1} - f)$, where we are using the observation that $1 - f(x,y) = f(y,x)$. (This last identity has a combinatorial reason arising from planar dual trees.) Therefore $\mu_p \preccurlyeq \mathbf{P}^f \preccurlyeq \mu_{1-p}$ for $p := e^{-4{\rm G}/\pi} = 0.3115^+$ and this p is optimal (on each side). This result in itself is rather surprising and it would be fascinating to see an explicit monotone coupling.

For $f : \mathbb{T}^d \to [0,1]$ measurable and $r = (r_1, r_2, \ldots, r_d) \in \mathbb{Z}^d$ with all $r_j > 0$, define

$$f_r(t) := \frac{1}{\prod_{j=1}^d r_j} \sum_{x \in r^{-1} t} f(x) \,,$$

where $r^{-1}(t_1, t_2, \ldots, t_d) := \{(x_1, x_2, \ldots, x_d) \in \mathbb{T}^d \,;\ \forall j \quad r_j x_j = t_j\}$. Write $r \to \infty$ to mean that $\min r_j \to \infty$. The following is a straightforward extension of Theorem 5.6. We leave its proof to the reader.

Theorem 5.14. *Let $f : \mathbb{T}^d \to [0,1]$ be measurable. If $\mathsf{GM}(f) > 0$ or if f is positive on a non-empty open set, then there exist constants $p_r \to \widehat{f}(\mathbf{0})$ as $r \to \infty$ such that $\mathbf{P}^{f_r} \succcurlyeq \mathbf{P}^{p_r}$ for all r. Therefore, if $\mathsf{GM}\big(f(\mathbf{1} - f)\big) > 0$ or if f is continuous and not equal to $\mathbf{0}$ nor $\mathbf{1}$, then there exist constants $p_r, q_r \to \widehat{f}(\mathbf{0})$ such that $\mathbf{P}^{p_r} \preccurlyeq \mathbf{P}^{f_r} \preccurlyeq \mathbf{P}^{q_r}$ for all r.*

Consider now a domination property even stronger than our previously defined strong domination. Let $\mathcal{Z}$ be the σ-field generated by $\eta(k)$ for $k \ne \mathbf{0}$.

Definition 5.15. A stationary process $\{\eta_n\}_{n\in\mathbb{Z}^d}$ with distribution ν **fully dominates** μ_p, written $\mu_p \preccurlyeq_{\mathrm{f}} \nu$, if

$$\mathbf{P}[\eta_{\mathbf{0}} = 1 \mid \mathcal{Z}] \ge p \quad \nu\text{-a.s.}$$

In this situation, we also say that ν is **uniformly insertion tolerant at level** p. Similarly, we define $\nu \preccurlyeq_{\mathrm{f}} \mu_p$ if the above inequality holds when $\ge$ is replaced by $\le$, and say that ν is **uniformly deletion tolerant at level** $1-p$. We say that ν is **uniformly insertion tolerant** if $\mu_p \preccurlyeq_{\mathrm{f}} \nu$ for some $p > 0$ and that ν is **uniformly deletion tolerant** if $\nu \preccurlyeq_{\mathrm{f}} \mu_p$ for some $p < 1$.

We show that the optimal level of uniform insertion tolerance of a determinantal process $\mathbf{P}^f$ is the **harmonic mean** of f, defined as

$$\mathsf{HM}(f) := \left(\int_{\mathbb{T}^d} \frac{d\lambda_d}{f}\right)^{-1} .$$

Note that $\mathsf{HM}(f) = 0$ iff $1/f$ is not integrable.

Theorem 5.16. *For any measurable $f : \mathbb{T}^d \to [0,1]$, we have $\mu_p \preccurlyeq_{\mathrm{f}} \mathbf{P}^f$ iff $p \le \mathsf{HM}(f)$. Similarly, $\mathbf{P}^f \preccurlyeq_{\mathrm{f}} \mu_q$ iff $q \ge 1 - \mathsf{HM}(\mathbf{1} - f)$.*

Proof. By Proposition 2.6 and Theorem 2.8, we have

$$\begin{aligned}\operatorname{ess\,inf} \mathbf{P}^f\Big[\eta(\mathbf{0}) = 1 \mid \mathcal{Z}\Big] &= \inf\Big\{\, \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}A \equiv 1]\,;\ A \subset \mathbb{Z}^d \setminus \{\mathbf{0}\} \text{ is finite } \Big\} \\ &= \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}B \equiv 1] = \min\left\{ \int |\mathbf{1} - T|^2 f \, d\lambda_d\,;\ T \in [B]_f \right\},\end{aligned}$$

where $B := \mathbb{Z}^d \setminus \{\mathbf{0}\}$. As shown by Kolmogorov (1941a, 1941b) for $d = 1$, the latter equals $\mathsf{HM}(f)$. The proof extends immediately to general d. This proves the first assertion. The second follows by symmetry. ∎

Remark 5.17. Since a proof of Kolmogorov's theorem that

$$\min\left\{ \int |\mathbf{1} - T|^2 f \, d\lambda_d\,;\ T \in [B]_f \right\} = \mathsf{HM}(f)\,,$$

where $B := \mathbb{Z}^d \setminus \{\mathbf{0}\}$, is difficult to find in readily accessible sources, we provide one here. We have

$$\min\left\{ \int |\mathbf{1} - T|^2 f \, d\lambda_d\,;\ T \in [B]_f \right\} = \|u\|_f^2\,,$$

where

$$u := P^{\perp}_{[B]_f} \mathbf{1}\,.$$

Now $g \perp [B]_f$ iff $g \in L^2(f)$ and $\widehat{gf}(k) = 0$ for all $k \in B$. The latter condition holds iff gf is a constant. If $g \neq \mathbf{0}$, then we deduce that $1/f \in L^2(f)$, i.e., $\mathsf{HM}(f) > 0$. Therefore,

$$[B]_f^\perp = \begin{cases} 0 & \text{if } \mathsf{HM}(f) = 0, \\ \mathbb{C}/f & \text{if } \mathsf{HM}(f) > 0. \end{cases}$$

Hence $u = \mathbf{0}$ if $\mathsf{HM}(f) = 0$, while otherwise,

$$\|u\|_f^2 = |(\mathbf{1}, \sqrt{\mathsf{HM}(f)}/f)_f|^2 = \mathsf{HM}(f)\,,$$

as desired.

REMARK 5.18. It is easy to find f such that $\mathsf{GM}(f) > 0$ and $\mathsf{HM}(f) = 0$. Indeed, a natural such example is the function g of Example 1.3. For any such f, the corresponding process $\mathbf{P}^f$ strongly dominates a nontrivial product measure, but does not fully dominate any nontrivial product measure. In addition, for any function f such that $\mathsf{HM}(f) > 0$ and f is not constant a.e., $\mathsf{GM}(f) > \mathsf{HM}(f)$ (as a consequence of Jensen's inequality), so that $\mathbf{P}^f$ will strongly dominate strictly more product measures than it will fully dominate.

EXAMPLE 5.19. Let f be as in (1.2), so that $\mathbf{P}^f$ is the law of the edges of the uniform spanning forest in $\mathbb{Z}^d$ that lie parallel to the x_1-axis. Then $\mathbf{P}^f$ is uniformly deletion tolerant iff $d \geq 4$. This is because $1/(\mathbf{1} - f)$ is integrable iff $d \geq 4$. For example, when $d = 4$, we obtain full domination by μ_p with $p := 0.66425^-$, where we have calculated $\mathsf{HM}(\mathbf{1} - f)$ as follows. First, we have

$$1/(\mathbf{1} - f) = 1 + \frac{\sin^2 \pi x_1}{\sum_{j=2}^4 \sin^2 \pi x_j}\,,$$

whence

$$\begin{aligned}\int_{\mathbb{T}^4} 1/(\mathbf{1} - f)\, d\lambda_4 &= 1 + \frac{1}{2}\int_{\mathbb{T}^3} \frac{1}{\sum_{j=2}^4 \sin^2 \pi x_j}\, d\lambda_3(x_2, x_3, x_4) \\ &= 1 + \frac{1}{2}\int_{\mathbb{T}^2} \frac{1}{\sqrt{\left(\sum_{j=2}^3 \sin^2 \pi x_j\right)\left(1 + \sum_{j=2}^3 \sin^2 \pi x_j\right)}}\, d\lambda_2(x_2, x_3)\,.\end{aligned}$$

This last integral has no simple form and is calculated numerically. This gives the value reported for $1 - \mathsf{HM}(\mathbf{1} - f)$. For large d, we have $1 - \mathsf{HM}(\mathbf{1} - f) \sim 1/d$. Indeed, write λ_{d-1}^* for Lebesgue measure on $\mathbb{R}^{d-1}$. Letting

$$A_d := \int_{\mathbb{T}^{d-1}} \frac{d-1}{\sum_{j=2}^d 2\sin^2 \pi x_j}\, d\lambda_{d-1}(x_2, \ldots, x_d)\,,$$

we have that

$$(d-1)(1-\mathsf{HM}(\mathbf{1}-f)) = \frac{A_d}{1+A_d/(d-1)} \,.$$

Hence it suffices to show that $\lim_{d\to\infty} A_d = 1$. To show this, note that

$$\begin{aligned} A_d &= \int_0^\infty \lambda_{d-1}\Big[\sum_{j=2}^d 2\sin^2 \pi x_j < (d-1)/t\Big]\,dt \\ &\le \int_0^2 \lambda_{d-1}\Big[\sum_{j=2}^d 2\sin^2 \pi x_j < (d-1)/t\Big]\,dt \\ &\quad + \int_2^{10(d-1)} \lambda_{d-1}\Big[\sum_{j=2}^d 2\sin^2 \pi x_j < (d-1)/2\Big]\,dt \\ &\quad + \int_{10(d-1)}^\infty \lambda_{d-1}^*\Big[\forall j \ \ |x_j| < 1/2 \text{ and } \sum_{j=2}^d 8x_j^2 < (d-1)/t\Big]\,dt \\ &\to \int_0^2 \mathbf{1}_{[0,1]}(t)\,dt = 1 \end{aligned}$$

as $d \to \infty$, where we have used the weak law of large numbers and the bounded convergence theorem for the first piece, a standard large-deviation result for the second piece, and an easy estimate on the third piece. The reverse inequality is obtained by using only the first piece. Thus, $A_d \to 1$, as desired. This value of $1-\mathsf{HM}(\mathbf{1}-f)$, which gives a full domination upper bound on $\mathbf{P}^f$, should be compared to $\widehat{f}(\mathbf{0}) = 1/d$ (by symmetry), which is the probability of a 1 at a site. On the other hand, for no d is the process uniformly insertion tolerant since $1/f$ is never integrable. We remark that for the full uniform spanning forest measure on $\mathbb{Z}^d$ (considering edges in all directions), we have change intolerance, meaning that $\mathbf{P}[\eta(\mathbf{0}) = 1 \mid \mathscr{Z}] \in \{0,1\}$ a.s. This follows from a result of Benjamini, Lyons, Peres, and Schramm (2001), as explained by Heicklen and Lyons (2003).

Example 5.20. Let $g(x)$ be as in Example 1.3, so that the edges of the uniform spanning tree in the plane that lie on the x-axis have the law $\mathbf{P}^g$. We have $\mathbf{P}^g \preccurlyeq_{\mathrm{f}} \mu_p$ for $p := (1+\pi)/(1+2\pi) = 0.56865^+$, and this is optimal. This is because

$$\begin{aligned} \int_{\mathbb{T}} 1/(1-g)\,d\lambda &= \int_{\mathbb{T}} \Big(1+\sin^2 \pi x + \sin \pi x \sqrt{1+\sin^2 \pi x}\Big)\,d\lambda(x) \\ &= 3/2 + \int_{\mathbb{T}} \sin \pi x \sqrt{1+\sin^2 \pi x}\,d\lambda(x)\,. \end{aligned}$$

An antiderivative of this integrand is

$$-\frac{1}{\pi}\arctan\left(\frac{\cos \pi x}{\sqrt{1+\sin^2 \pi x}}\right) - \frac{1}{2\pi}\cos \pi x \sqrt{1+\sin^2 \pi x}\,,$$

whence the remaining integral is $1/2 + 1/\pi$. Therefore $\int_{\mathbb{T}} 1/(1-g)\,d\lambda = 2 + 1/\pi$ and $1 - \mathsf{HM}(\mathbf{1} - g) = (1+\pi)/(1+2\pi)$, as desired. Observe that $\int 1/g = \infty$ and so the process is not uniformly insertion tolerant. Similarly, for the edges lying on the x-axis of the uniform spanning tree in 3 dimensions, we have full domination by μ_p with $p := 0.37732^+$. Here, the law of these edges is $\mathbf{P}^h$, where

$$h(x) := \int_{\mathbb{T}^2} f(x,y,z)\,d\lambda_2(y,z) = \int_{\mathbb{T}} \frac{\sin^2 \pi x}{\sqrt{\left(\sin^2 \pi x + \sin^2 \pi y\right)\left(1 + \sin^2 \pi x + \sin^2 \pi y\right)}}\,d\lambda_1(y)\,.$$

This integral has no simpler form, so to compute $p := 1 - \mathsf{HM}(\mathbf{1} - h)$, we calculated h numerically and used the result to calculate the harmonic mean numerically. Since $\int 1/h = \infty$, as is easily checked, this process is not uniformly insertion tolerant. Similarly, one can check that the process of edges lying on the x-axis of the uniform spanning tree in $d \geq 4$ dimensions is not uniformly insertion tolerant.

Example 5.21. Let f be as in Example 1.4. It turns out that $\mathsf{GM}(f) = e^{-2\mathrm{G}/\pi}/\sqrt{2} = 1 - \mathsf{GM}(\mathbf{1} - f) = 0.39467^+$. This gives strong domination inequalities. It is easy to see that $\mathsf{HM}(f) = \mathsf{HM}(\mathbf{1} - f) = 0$, so there are no nontrivial full domination inequalities. The interest of this function is that it describes the process of edges along a zig-zag path in the plane for the uniform spanning tree, as claimed in Example 1.4. We now sketch how to prove this. The Transfer Current Theorem of Burton and Pemantle (1993) allows one to calculate, via determinants, the law for any set of possible edges of the uniform spanning tree. Thus, it is enough to verify that for the edges belonging to the zig-zag path, the matrix entries are those of the Toeplitz matrix associated to the Fourier coefficients given in Example 1.4. For even k, the values of $\widehat{f}(k)$ are given in Lyons with Peres (2004). (These values imply the astonishing fact that the edges in the plane that lie along a diagonal, e.g., the horizontal edges with left endpoints (n,n) $(n \in \mathbb{Z})$, are independent, i.e., have law $\mu_{1/2}$.) Thus, it remains to treat the case of odd k. A straightforward application of the Transfer Current Theorem gives that the matrix entry corresponding to the edges e_0 and e_{2k-1} is

$$\int \frac{e(s) + e(t) - e(s+t) - 1}{4 - e(s) - e(-s) - e(t) - e(-t)} e(ks + kt)\,d\lambda_2(s,t)$$
$$= \int \frac{e(s) + e(x-s) - e(x) - 1}{4 - e(s) - e(-s) - e(x-s) - e(s-x)} e(kx)\,d\lambda_2(s,x)\,,$$

where $e(x) := e^{2\pi i x}$. Evaluate the integral in s for fixed x by a contour integral

$$\frac{1}{2\pi i} \oint \frac{z + e(x)z^{-1} - e(x) - 1}{4 - z - z^{-1} - e(x)z^{-1} - e(-x)z} \frac{dz}{z}$$

over the contour $|z| = 1$. The integrand has poles inside the unit disc at $z = 0$ and

$$z = \frac{2 - \sqrt{4 - |1 + e(x)|^2}}{1 + e(-x)} .$$

After use of the residue theorem and integrating in x, one obtains $\widehat{f}(2k-1)$, as desired.

We close this section by describing how our domination results can be interpreted in terms of prediction and interpolation questions for wide-sense stationary processes. In view of Theorem 2.8 and the well-known correspondence between prediction and Szegő infima, for $d = 1$, $\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction\{-1, -2, \ldots\} \equiv 1]$ is exactly the mean squared error for the best linear predictor of Y_0 given $\langle Y_n \rangle_{n \le -1}$, where $\langle Y_n \rangle$ is a wide-sense stationary process with spectral density f. Similarly, $\mathbf{P}^f[\eta(\mathbf{0}) = 0 \mid \eta\restriction\{-1, -2, \ldots\} \equiv 0]$ is exactly the mean squared error for the best linear predictor of Y_0 given $\langle Y_n \rangle_{n \le -1}$, where $\langle Y_n \rangle$ is now a wide-sense stationary process with spectral density $\mathbf{1} - f$. This gives us a correspondence between strong domination and prediction. An analogous correspondence holds between full domination and interpolation, where one instead looks at the mean squared error for the best linear predictor of Y_0 given $\langle Y_n \rangle_{n \ne 0}$.

§6. Entropy.

We assume the reader is familiar with the definition of the entropy $H(\mu)$ of a process μ, as well as basic results concerning entropy (see Walters (1982) and Katznelson and Weiss (1972)). Because of Ornstein's theorem (and its generalizations, see Katznelson and Weiss (1972), Conze (1972/73), Thouvenot (1972), and Ornstein and Weiss (1987)) that entropy characterizes Bernoulli shifts up to isomorphism, the following question is particularly interesting:

Question 6.1. What is $H(\mathbf{P}^f)$?

We know the answer only in the trivial case where f is a constant and in case f or $\mathbf{1} - f$ is the reciprocal of a trigonometric polynomial of degree 1. In principle, as we shall see, one can also determine the entropy when f or $\mathbf{1} - f$ is the reciprocal of any trigonometric polynomial, but the formula would be rather unwieldy.

In general, then, we shall discuss how to estimate the entropy of $\mathbf{P}^f$. The definition of entropy always provides upper bounds, due to subadditivity, so the harder bound is the lower bound. As we shall see, reciprocals of trigonometric polynomials can be used to get arbitrarily close lower (and upper) bounds. Unfortunately, that method is not practical for precise computation. Nevertheless, in many cases, we have another method that appears to work quite well. Indeed, our method seems to provide upper bounds that converge more quickly than does use of the definition.

Shirai and Takahashi (2003) proved that $H(\mathbf{P}^f) > 0$ for all $f \neq \mathbf{0}, \mathbf{1}$. Of course, this also follows immediately from our Theorem 3.1.

Let

$$H[p] := H(\mu_p) = -p \log p - (1-p)\log(1-p)\,.$$

Theorem 5.11 yields easy lower bounds on the entropy of those processes $\mathbf{P}^f$ such that $f(\mathbf{1}-f)$ has a strictly positive geometric mean. We shall obtain more refined lower bounds later in the one-dimensional case. It is easy to see that

$$\mu_p \preccurlyeq_{\mathrm{s}} \mu \preccurlyeq_{\mathrm{s}} \mu_{1-p} \quad\Longrightarrow\quad H(\mu) \geq H[p]\,. \tag{6.1}$$

By Theorem 5.11 and (6.1), we deduce the following lower bound on entropy:

Proposition 6.2. *For any measurable $f : \mathbb{T}^d \to [0,1]$, we have*

$$H(\mathbf{P}^f) \geq \min\Big\{H\big[\mathsf{GM}(f)\big], H\big[\mathsf{GM}(\mathbf{1}-f)\big]\Big\}\,.$$

Remark 6.3. This can be compared to the trivial upper bound $H(\mathbf{P}^f) \leq H\big[\widehat{f}(0)\big]$. Note that $\widehat{f}(0)$ is the arithmetic mean of f. Shirai and Takahashi (2003) also note this upper bound and provide two lower bounds: $H(\mathbf{P}^f) \geq H\big[\widehat{f}(0)\big]/2$ and $H(\mathbf{P}^f) \geq \int_{\mathbb{T}^d} H[f(x)]\,d\lambda_d(x)$.

Remark 6.4. It is interesting that (6.1) is not true if $\preccurlyeq_{\mathrm{s}}$ is replaced by $\preccurlyeq$. For example, let (X_{2i}, X_{2i+1}) be, independently for different i, $(1,1)$ or $(0,0)$, each with probability $1/2$. If ν is the distribution of this process, then ν is not stationary, but $\mu := \big(\nu + T(\nu)\big)/2$ is stationary, where T is the shift. Next, $H(\mu) = \log 2/2 < H[1/\sqrt{2}]$, even though $\mu_{1-1/\sqrt{2}} \preccurlyeq \mu \preccurlyeq \mu_{1/\sqrt{2}}$, as is easily verified. We also observe that this measure does not even have full support.

Example 6.5. Let f be as in (1.1), so that $\mathbf{P}^f$ is the law of the horizontal edges of a uniform spanning tree in the square lattice $\mathbb{Z}^2$. It is known that the entropy of the entire uniform spanning tree measure (both horizontal and vertical edges included) is (5.3); see Burton and Pemantle (1993), corrected in Lyons (2003a). This can be compared to the bound on $H(\mathbf{P}^f)$ that Proposition 6.2 provides, together with the calculations of Example 5.13, namely, $H(\mathbf{P}^f) \geq H[e^{-4\mathrm{G}/\pi}] = 0.6203^+$. (The results are much worse if one uses the bounds of Shirai and Takahashi (2003) reported in Remark 6.3 above.) Direct calculation using cylinder events corresponding to a 4-by-4 block gives an upper bound for the entropy

$H(\mathbf{P}^f) \le 0.68864$. Of course, the vertical edges of the uniform spanning tree measure have the same entropy as the horizontal edges.

Example 6.6. Let $g(x)$ be as in Example 1.3, so that the edges of the uniform spanning tree in the plane that lie on the x-axis have the law $\mathbf{P}^g$. In view of Example 5.5, Proposition 6.2 implies a lower bound of $H(\mathbf{P}^g) \ge H[\sqrt{2}-1] \ge 0.67835$. Also, direct calculation using cylinder events of length 16 gives an upper bound $H(\mathbf{P}^g) \le 0.69034$. In Example 6.15, we shall obtain more refined bounds on the entropy. It is not hard to see that the horizontal edges associated to the y-axis have law $\mathbf{P}^{\mathbf{1}-g}$.

Although not relevant for our determinantal probability measures (where we have seen that $\preccurlyeq_{\mathrm{s}}$ and $\preccurlyeq$ are equivalent), it is interesting to ask whether $\mu_p \preccurlyeq \mu \preccurlyeq \mu_{1-p}$ with $p > 0$ yields any lower bound on the entropy $H(\mu)$ for general μ. We first ask the question whether $\mu_p \preccurlyeq \mu$ with $p > 0$ implies that $H(\mu) > 0$ provided $\mu \ne \delta_{\mathbf{1}}$. The answer to this question is affirmative since it is known (see Furstenberg (1967) for $d = 1$ and Glasner, Thouvenot, and Weiss (2000) for $d > 1$) that zero-entropy processes are **disjoint** from i.i.d. processes (meaning that there are no stationary couplings of them other than independent couplings). The next result provides an explicit lower bound on the entropy for processes trapped between two i.i.d. processes. The proof was obtained jointly with Chris Hoffman.

Proposition 6.7. *For any $d \ge 1$, if $\mu_p \preccurlyeq \mu \preccurlyeq \mu_{1-p}$ with $p > 0$, then*

$$H(\mu) \ge \max\{a_p, b_p\} > 0\,,$$

where

$$a_p := (1-p)\log\left(\frac{1}{1-p}\right) - \frac{1-2p}{2}\log\left(\frac{1}{1-2p}\right)$$

and

$$b_p := 2(1-p)\log\left(\frac{1}{1-p}\right) - (1-2p)\log\left(\frac{1}{1-2p}\right) - (1-2p)\log 2.$$

Remark 6.8. Observe that b_p approaches $\log 2$ as $p \to 1/2$.

Proof. Let X, Y and Z denote processes with respective distributions μ_p, μ and μ_{1-p}. We construct a joining (stationary coupling) of all three processes as follows. Consider any joining m_1 of X and Y with $X_i \le Y_i$ a.s. for all i and any joining m_2 of Y and Z with $Y_i \le Z_i$ a.s. for all i. We now pick a realization for Y according to μ and then choose X and Z (conditionally) independently using m_1 and m_2 respectively (this is called the fibered product of m_1 and m_2 over Y). This gives us a joining (X', Y', Z') of X, Y and Z. We may now assume that $(X, Y, Z) = (X', Y', Z')$.

We now use standard facts about entropy. First, $H(X,Z) \geq H(X,Y,Z) - H(Y)$. We next note, using the conditional independence of X and Z given Y, that

$$\begin{aligned} H(X,Y,Z) &= H(Y) + H(X,Z \mid Y) = H(Y) + H(X \mid Y) + H(Z \mid Y) \\ &\geq H(X) + H(Z) - H(Y) \,. \end{aligned}$$

Hence

$$H(X,Z) \geq H(X) + H(Z) - 2H(Y) = 2H[p] - 2H(Y) \,.$$

Since $X \leq Z$, the one-dimensional marginal of (X,Z) necessarily has 3 atoms of weights $p, 1-2p$ and p. It follows that

$$H(X,Z) \leq 2p \log\left(\frac{1}{p}\right) + (1-2p) \log\left(\frac{1}{1-2p}\right) .$$

Combining this with the previous inequality shows that $H(\mu) \geq a_p$.

We modify the above proof to show that b_p is also a lower bound. Rather than using $H(X,Z) \geq H(X,Y,Z) - H(Y)$, we use $H(X,Z) = H(X,Y,Z) - H(Y \mid X,Z)$. The earlier argument then gives

$$H(X,Z) \geq 2H[p] - H(Y) - H(Y \mid X,Z) \,.$$

Using the same upper bound on $H(X,Z)$ as before and the trivial upper bound for $H(Y \mid X,Z)$ of $(1-2p)\log 2$ (obtained by noting that Y_i is determined by X_i and Z_i when $X_i = Z_i$) yields the lower bound of b_p.

The easiest way to check the strict positivity of a_p is to observe that

$$2H[p] > 2p \log\left(\frac{1}{p}\right) + (1-2p) \log\left(\frac{1}{1-2p}\right) .$$

This, in turn, follows from the fact that the left-hand side is the entropy of the independent joining $\mu_p \times \mu_{1-p}$, while the right-hand side is the entropy of the monotonic coordinatewise-independent joining of μ_p and μ_{1-p}, whose existence follows from the inequality $p \leq 1-p$. ∎

There are few cases where we know the entropy $H(\mathbf{P}^f)$ exactly. Besides the case when f is constant, we can calculate the entropy when f is the reciprocal of a trigonometric polynomial of degree 1. We shall illustrate this in the simplest case, Example 1.7. Thus, let $0 < a < 1$, $d = 1$, and $f(x) := (1-a)^2/|e^{2\pi i x} - a|^2$. Recall that for any invariant measure μ in one dimension, we have

$$H(\mu) = \int H\big[\mu[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots]\big] \, d\mu(\eta_{-1}, \eta_{-2}, \ldots) \,. \tag{6.2}$$

Now by (1.3), it is easy to check that

$$P^f[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] = \frac{(1-a)^2 N}{N-(N-1)a}\,,$$

where $N := \min\{k \geq 1\,;\ \eta_{-k} = 1\}$. In addition, we may easily show that for $n \geq 1$,

$$\mathbf{P}^f[N = n] = \frac{(1-a)(n-(n-1)a)a^{n-1}}{1+a}\,.$$

Therefore

$$\begin{aligned} H(\mathbf{P}^f) &= \sum_{n=1}^{\infty} \frac{(1-a)(n-(n-1)a)a^{n-1}}{1+a}\Big(- \frac{(1-a)^2 n}{n-(n-1)a} \log \frac{(1-a)^2 n}{n-(n-1)a} \\ &\qquad\qquad - \frac{n-(n-1)a-(1-a)^2 n}{n-(n-1)a} \log \frac{n-(n-1)a-(1-a)^2 n}{n-(n-1)a}\Big) \\ &= \frac{1-a}{1+a}\sum_{n=1}^{\infty} a^{n-1}\Big(-(1-a)^2 n \log[(1-a)^2 n] \\ &\qquad\qquad - [a(1-a)n + a]\log[a(1-a)n+a] \\ &\qquad\qquad + [(1-a)n+a]\log[(1-a)n+a]\Big)\,. \end{aligned}$$

In principle, one can also write an infinite series of positive terms for the entropy of $\mathbf{P}^f$ when f is equal to the reciprocal of any trigonometric polynomial, since, by Proposition 2.10, the process $\mathbf{P}^f$ is then a regenerative process. Of course, the answer will be much more unwieldy than the above formula. However, it can be used to get arbitrarily good lower bounds on the entropy of any process $\mathbf{P}^f$, in theory. To see this, we use the following lemma, which is more or less Lemma 5.10 in Rudolph (1990). As our proof is somewhat shorter and gives a more precise bound, we include it. See also Burton and Pemantle (1993), Lemma 6.2, for a similar proof.

Lemma 6.9. *For any stationary processes μ and ν, we have $|H(\mu) - H(\nu)| \leq H[\bar{d}(\mu,\nu)]$.*

Proof. Consider a stationary process $(X, Y) = \langle (X_i, Y_i)\rangle_{i\in\mathbb{Z}^d}$, where $X = \langle X_i\rangle_{i\in\mathbb{Z}^d}$ has distribution μ, $Y = \langle Y_i\rangle_{i\in\mathbb{Z}^d}$ has distribution ν, and $P(X_0 \neq Y_0) = \bar{d}(\mu,\nu)$. Let $Z_i := X_i + Y_i \bmod 2$. Then

$$H(\mu) \leq H(X,Z) = H(Y,Z) \leq H(\nu) + H(Z) \leq H(\nu) + H[\bar{d}(\mu,\nu)]\,.$$

Since the same holds with μ and ν switched, the result follows. ∎

Thus, given any measurable $f : \mathbb{T} \to [0,1]$ and any $\epsilon > 0$, define δ to be the smallest positive number so that $H[\delta] = \epsilon$. Then choose a trigonometric polynomial $T \geq 1$ such that $\int_{\mathbb{T}} |T^{-1} - \max(f, \delta/2)|\, d\lambda_1 < \delta/2$. We can calculate $H(\mathbf{P}^{1/T})$ as closely as desired. Since $\overline{d}(\mathbf{P}^{1/T}, \mathbf{P}^f) < \delta$ by Proposition 3.4, it follows by Lemma 6.9 that $|H(\mathbf{P}^f) - H(\mathbf{P}^{1/T})| < \epsilon$.

Unfortunately, this method of calculation is hopeless in practice since when T has a high degree, it will take a very long time to see a renewal, which means that cylinder events of great length will be needed for the estimation. When a cylinder event of length n is used, one must calculate 2^n probabilities. Thus, this is completely impractical.

However, we now exhibit an alternative method that works extremely well in practice, although we cannot prove that it works well *a priori* in all cases. We begin with two simple examples,

$$f(x) := \mathbf{1}_{[0,1/2]}(x) \tag{6.3}$$

(which also occurred in Example 1.6) and

$$g(x) := \sin^2 \pi x = (1 - \cos 2\pi x)/2\,. \tag{6.4}$$

Since $\widehat{f}(k) = \widehat{g}(k) = 0$ for all even $k \neq 0$, both $\mathbf{P}^f$ and $\mathbf{P}^g$ have the property that looking at only the even coordinates, we see independent fair coin flips, i.e., $\mu_{1/2}$. Therefore the entropy of both processes is at least $(1/2)\log 2$. In either case, we know of no direct method to prove that strict inequality holds, but it does. We first show this for $\mathbf{P}^g$. In Example 5.5 we showed that $\mathsf{GM}(g) = 1/4$. By symmetry, $\mathsf{GM}(1-g) = 1/4$. Since $H[1/4] = 0.56^+ > (1/2)\log 2 = 0.35^-$, we obtain by Proposition 6.2 that $H(\mathbf{P}^g) > (1/2)\log 2$.

In order to show that $H(\mathbf{P}^f) > (1/2)\log 2$ for f as in (6.3), we need a method to obtain more refined entropy bounds. We illustrate such a method beginning with this simple function g. While we shall explain afterwards a method to obtain results for more general functions for the case $d = 1$, this first proof contains the essential idea of this method while at the same time relying on a more elementary calculation, and therefore, we feel, is worth including.

PROPOSITION 6.10. *With g as in (6.4), we have*

$$0.63^- = \frac{3}{8}H[1/4] + \frac{5}{8}H[11/28] \leq H(\mathbf{P}^g) \leq \frac{3}{8}H[7/20] + \frac{5}{8}H[5/12] = 0.67^-\,.$$

LEMMA 6.11. *Let $h : \mathbb{T} \to [0,1]$ be a trigonometric polynomial of degree at most 1, i.e., $\widehat{h}(k) = 0$ for $|k| \geq 2$. Fix $n > 0$ and $A \subseteq \{1, 2, \ldots, n\}$. For $C \subseteq \{1, 2, \ldots, n\}$, let $\lambda(C)$ denote the sequence of lengths of consecutive intervals in $\{1, 2, \ldots, n\} \setminus C$. Then*

$$\mathbf{P}^h[\eta\restriction A \equiv 0,\, \eta\restriction(\{1,2,\ldots,n\} \setminus A) \equiv 1] = \sum_{C \subseteq A} (-1)^{|A \setminus C|} \prod_{i \in \lambda(C)} D_{i-1}(h)\,.$$

Proof. Whenever two sets $F_1, F_2 \subset \mathbb{Z}$ neither overlap nor come within distance 1 of each other, the configurations on F_1 and F_2 are $\mathbf{P}^h$-independent because $\widehat{h}(k) = 0$ for $|k| \ge 2$. Therefore $\mathbf{P}^h[\eta{\restriction}(\{1, 2, \ldots, n\} \setminus C) \equiv 1] = \prod_{i \in \lambda(C)} D_{i-1}(h)$. Now the desired formula follows from the inclusion-exclusion principle. ∎

Proof of Proposition 6.10. Write $d_n := D_{n-1}(g)$. Recall that

$$H(\mathbf{P}^g) = \int H\big[\,\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots]\big]\, d\,\mathbf{P}^g(\eta_{-1}, \eta_{-2}, \ldots)\,. \tag{6.5}$$

By the negative association property of $\mathbf{P}^g[\,\cdot \mid \eta_{-1} = \eta_{-2} = 1]$, we have that on the event $\{\eta_{-1} = \eta_{-2} = 1\}$,

$$\begin{aligned}
\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] &\ge \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1} = \eta_{-2} = \cdots = \eta_{-n} = 1] \\
&= \lim_{n\to\infty} \frac{\mathbf{P}^g[\eta_0 = \eta_{-1} = \eta_{-2} = \cdots = \eta_{-n} = 1]}{\mathbf{P}^g[\eta_{-1} = \eta_{-2} = \cdots = \eta_{-n} = 1]} \\
&= \lim_{n\to\infty} d_{n+1}/d_n = \mathsf{GM}(g) = 1/4\,.
\end{aligned}$$

Likewise, on the same event, we have

$$\begin{aligned}
\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] &\le \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1} = \eta_{-2} = 1,\ \eta_{-3} = \cdots = \eta_{-n} = 0] \\
&= \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 0 \mid \eta_{-1} = \eta_{-2} = 0,\ \eta_{-3} = \cdots = \eta_{-n} = 1] \\
&\qquad\qquad [\text{by symmetry}] \\
&= 1 - \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1} = \eta_{-2} = 0,\ \eta_{-3} = \cdots = \eta_{-n} = 1] \\
&= 1 - \lim_{n\to\infty} \frac{\mathbf{P}^g[\eta_{-1} = \eta_{-2} = 0,\ \eta_0 = \eta_{-3} = \cdots = \eta_{-n} = 1]}{\mathbf{P}^g[\eta_{-1} = \eta_{-2} = 0,\ \eta_{-3} = \cdots = \eta_{-n} = 1]} \\
&= 1 - \lim_{n\to\infty} \frac{d_1 d_{n-2} - d_2 d_{n-2} - d_1 d_{n-1} + d_{n+1}}{d_{n-2} - d_1 d_{n-2} - d_{n-1} + d_n} \\
&\qquad\qquad [\text{by Lemma 6.11}] \\
&= 1 - \frac{d_1 - d_2 - d_1 \mathsf{GM}(g) + \mathsf{GM}(g)^3}{1 - d_1 - \mathsf{GM}(g) + \mathsf{GM}(g)^2} \\
&= 7/20
\end{aligned}$$

since $d_1 = 1/2$ and $d_2 = 3/16$. We may conclude that on the event $\{\eta_{-1} = \eta_{-2} = 1\}$, we have

$$H\big[\,\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots]\big] \in \big[H[1/4], H[7/20]\big]\,.$$

By symmetry, the same holds on the event $\{\eta_{-1} = \eta_{-2} = 0\}$.

Similarly, we have that on the event $\{\eta_{-1} = 1,\, \eta_{-2} = 0\}$,

$$\begin{aligned}\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] &\ge \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-2} = 0,\, \eta_{-1} = \eta_{-3} = \cdots = \eta_{-n} = 1] \\ &= \lim_{n\to\infty} \frac{\mathbf{P}^g[\eta_{-2} = 0,\, \eta_0 = \eta_{-1} = \eta_{-3} = \cdots = \eta_{-n} = 1]}{\mathbf{P}^g[\eta_{-2} = 0,\, \eta_{-1} = \eta_{-3} = \cdots = \eta_{-n} = 1]} \\ &= \lim_{n\to\infty} \frac{d_2 d_{n-2} - d_{n+1}}{d_1 d_{n-2} - d_n} \\ &= \frac{d_2 - \mathsf{GM}(g)^3}{d_1 - \mathsf{GM}(g)^2} \\ &= 11/28\,.\end{aligned}$$

Likewise, on the same event, we have

$$\begin{aligned}\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] &\le \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1} = 1,\, \eta_{-2} = \eta_{-3} = \cdots = \eta_{-n} = 0] \\ &= \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 0 \mid \eta_{-1} = 0,\, \eta_{-2} = \eta_{-3} = \cdots = \eta_{-n} = 1] \\ &= 1 - \lim_{n\to\infty} \mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1} = 0,\, \eta_{-2} = \eta_{-3} = \cdots = \eta_{-n} = 1] \\ &= 1 - \lim_{n\to\infty} \frac{\mathbf{P}^g[\eta_{-1} = 0,\, \eta_0 = \eta_{-2} = \cdots = \eta_{-n} = 1]}{\mathbf{P}^g[\eta_{-1} = 0,\, \eta_{-2} = \cdots = \eta_{-n} = 1]} \\ &= 1 - \lim_{n\to\infty} \frac{d_1 d_{n-1} - d_{n+1}}{d_{n-1} - d_n} \\ &= 1 - \frac{d_1 - \mathsf{GM}(g)^2}{1 - \mathsf{GM}(g)} \\ &= 5/12\,.\end{aligned}$$

We may conclude that on the event $\{\eta_{-1} = 1,\, \eta_{-2} = 0\}$, we have

$$H\big[\,\mathbf{P}^g[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots]\big] \in \big[H[11/28], H[5/12]\big]\,.$$

By symmetry, the same holds on the event $\{\eta_{-1} = 0,\, \eta_{-2} = 1\}$.

Putting all these bounds into (6.5) gives the claimed bounds on the entropy of $\mathbf{P}^g$. ∎

For more general functions, we can get lower bounds on the entropy by using the asymptotics of Bump and Diaconis (2002) or of Tracy and Widom (2002), which serve as replacements for our use of the special Lemma 6.11. In fact, we use an extension of the formula of Tracy and Widom (2002) due to Lyons (2003c). Denote by ν_f the measure on $2^{\mathbb{N}}$ obtained by the limiting condition $\lim_{n\to\infty} \mathbf{P}^f[\,\cdot \mid \eta_{-1} = \eta_{-2} = \cdots = \eta_{-n} = 1]$, which exists by Proposition 2.6. If $\mathsf{GM}(f) > 0$, define

$$\Phi_f(z) := \exp \frac{1}{2} \int_{\mathbb{T}} \frac{e^{2\pi i t} + z}{e^{2\pi i t} - z} \log f(t)\, d\lambda_1(t) \tag{6.6}$$

for $|z| < 1$. The **outer** function

$$\varphi_f(t) := \lim_{r \uparrow 1} \Phi_f(re^{2\pi i t}) \tag{6.7}$$

exists for λ_1-a.e. $t \in \mathbb{T}$ and satisfies $|\varphi_f|^2 = f$ λ_1-a.e. This limit also holds in $L^1(\mathbb{T})$. See Rudin (1987), Theorem 17.11, p. 340, and Theorem 17.16, p. 343. As is easily verified,

$$\log \Phi_f(z) = \widehat{F}(0) + 2\sum_{k=1}^{\infty} \widehat{F}(k) z^k \tag{6.8}$$

when $F := (1/2)\log f$. The mean value theorem for analytic functions and the $L^1(\mathbb{T})$ convergence above show that

$$\widehat{\varphi_f}(0) = \Phi_f(0) = \sqrt{\mathsf{GM}(f)}\,.$$

Define $\Phi_f := \varphi_f := \mathbf{0}$ if $\mathsf{GM}(f) = 0$. An analytic trigonometric polynomial is (a constant times) an outer function iff its extension to the unit disc has no zeroes in the open disc. [On the one hand, the extension of no outer function has any zeroes in the open disc since it is an exponential. On the other hand, by factoring a polynomial, one has to check merely that $z \mapsto z - \zeta$ is a constant times an outer function for $|\zeta| \ge 1$. Indeed, $\Phi_f(z) = (-|\zeta|/\zeta)(z - \zeta)$ for $f(t) := |e^{2\pi i t} - \zeta|^2$, as can be seen by the Poisson integral formula for $|\zeta| > 1$ (Rudin (1987), Theorem 11.9, p. 235) and by a limiting procedure for $|\zeta| = 1$.] The results of Lyons (2003c) show that Theorem 2.9 reduces to the following formula:

THEOREM 6.12. *For any measurable $f : \mathbb{T} \to [0, 1]$, the measure ν_f is equal to the determinantal probability measure corresponding to the positive contraction on $\ell^2(\mathbb{N})$ whose (j, k)-matrix entry is*

$$\sum_{l=0}^{j \wedge k} \overline{\widehat{\varphi_f}(j-l)}\,\widehat{\varphi_f}(k-l)\,.$$

This can be substituted in the appropriate places in the proof of Proposition 6.10. For example, on the event $\{\eta_{-1} = \eta_{-2} = 0\}$, one has that

$$\begin{aligned}\mathbf{P}^f[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots] &\ge \lim_{n\to\infty} \mathbf{P}^f[\eta_0 = 1 \mid \eta_{-1} = \eta_{-2} = 0,\ \eta_{-3} = \eta_{-4} = \cdots = \eta_{-n} = 1] \\ &= \nu_f[\eta_2 = 1, \eta_1 = \eta_0 = 0]/\nu_f[\eta_1 = \eta_0 = 0]\,.\end{aligned}$$

One can use $\nu_{\mathbf{1}-f}$ in a similar fashion for estimating such probabilities above by η that terminate in repeating 0s, rather than in repeating 1s. For example, if

$$f := |1 + \mathbf{e}_1 + \mathbf{e}_2|^2/9\,,$$

then $\varphi_f = (1 + \mathbf{e}_1 + \mathbf{e}_2)/3$ since the polynomial $(1 + z + z^2)/3$ has no zeroes in the open unit disc. Similarly, one can show that

$$\varphi_{\mathbf{1}-f} = \frac{\sqrt{6}+\sqrt{2}}{6} - \frac{\sqrt{2}}{3}\mathbf{e}_1 + \frac{\sqrt{2}-\sqrt{6}}{6}\mathbf{e}_2 \,.$$

With these and Theorem 6.12, one can show by following the method of proof of Proposition 6.10 that

$$0.53 \le H(\mathbf{P}^f) \le 0.61 \,.$$

This method of estimation is relatively easy to program on a computer; decomposing by cylinder events of length 8 instead of the length 2 used in the proof of Proposition 6.10 and using Mathematica, we find that

$$0.601992 \le H(\mathbf{P}^f) \le 0.602433 \,.$$

Similarly, using cylinder events of length 15, we obtain that

$$0.65907716 \le H(\mathbf{P}^g) \le 0.65907733 \tag{6.9}$$

for the function g in (6.4).

Having illustrated our method for estimating entropy in case $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1}-f) > 0$, we may now prove that the bounds it gives do indeed converge to the actual entropy provided that f is bounded away from 0 and 1. (Although we cannot prove it, the same probably holds whenever $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1}-f) > 0$.) Our method relies on being able to approximate uniformly $\mathbf{P}^f[\eta_0 = 1 \mid \eta_{-1}, \eta_{-2}, \ldots]$ arbitrarily well on each cylinder event of the form $\{\eta_{-1} = a_{-1}, \ldots, \eta_{-n} = a_{-n}\}$. In other words, our method gives arbitrarily good approximations to the entropy as long as $\mathbf{P}^f$ is one-sided quasi-local, a concept defined as follows and also sometimes referred to as the uniform martingale property (see Kalikow (1990)). We begin with a relevant property from statistical physics. Let $\mathscr{Z}$ be the σ-field in Definition 5.15. A stationary process ν is called **quasi-local** if $\nu[\eta(\mathbf{0}) = 1 \mid \mathscr{Z}]$ has a continuous version, where the product topology is used on $2^{\mathbb{Z}^d \setminus \{\mathbf{0}\}}$. When $\mathscr{Z}$ is replaced by the σ-field $\mathsf{Past}(\mathbf{0})$ in this definition, we add the word "one-sided" to the name. Now Shirai and Takahashi (2003), in their Theorem 6.2, have shown that on any $\ell^2(E)$, for any positive contraction Q with spectrum in $(0,1)$, $\mathbf{P}^Q[\eta(e_0) = 1 \mid \eta_e, e \ne e_0]$ has a continuous version on $2^{E \setminus \{e_0\}}$ for any $e_0 \in E$. Since the spectrum of the multiplication operator M_f on $L^2(\mathbb{T})$ is the essential range of f, the same holds for the corresponding convolution operator Q_f on $\ell^2(\mathbb{Z})$ and thus $\mathbf{P}^f$ is quasi-local if f is bounded away from 0 and 1. Moreover, the compression of Q_f to $\ell^2(-\mathbb{N})$ then also has spectrum in $(0,1)$, whence $\mathbf{P}^f$ is also one-sided quasi-local, which proves our claim.

We now return to the problem of showing that $H(\mathbf{P}^f) > (1/2)\log 2$ for f as in (6.3).

Proposition 6.13. *If* $f := \mathbf{1}_{[0,1/2]}$, *then* $H(\mathbf{P}^f) > (1/2)\log 2$.

Proof. Let $\widetilde{f} := .99\mathbf{1}_{[0,1/2]} + .01\mathbf{1}_{(1/2,1)}$. Proposition 3.4 tells us that $\overline{d}(\mathbf{P}^f, \mathbf{P}^{\widetilde{f}}) \le .01$, and hence from Lemma 6.9, $H(\mathbf{P}^f) \ge H(\mathbf{P}^{\widetilde{f}}) - H[.01]$. Now $H[.01] < 0.0561$. Write $\delta := .01$ and $L := \log(\delta^{-1} - 1) = \log 99$. A simple integration shows that

$$\widehat{\widetilde{f}}(k) = \begin{cases} 1/2 & \text{if } k = 0, \\ \dfrac{(-1)^k - 1}{2k\pi} i(1-2\delta) & \text{if } k \ne 0 \end{cases}$$

(where $i := \sqrt{-1}$). To find $\varphi_{\widetilde{f}}$, we proceed as follows. Let $F := (1/2)\log \widetilde{f}$. Then

$$\widehat{F}(k) = \begin{cases} (1/4)\log \delta(1-\delta) & \text{if } k = 0, \\ \dfrac{(-1)^k - 1}{4k\pi} iL & \text{if } k \ne 0. \end{cases}$$

By using (6.8) and exponentiating, we obtain

$$\Phi_{\widetilde{f}}(z) = [\delta(1-\delta)]^{1/4} \left(1 - \frac{i}{\pi} Lz - \frac{1}{2\pi^2} L^2 z^2 - \frac{i}{6\pi^3}(2\pi^2 - L^2)Lz^3 + \cdots \right).$$

Since the coefficients of this power series are the Fourier coefficients of $\varphi_{\widetilde{f}}$, we have a procedure to compute the Fourier coefficients $\varphi_{\widetilde{f}}$. Since $\widetilde{f}(1-x) = 1 - \widetilde{f}(x)$, we have $\overline{\Phi_{\widetilde{f}}(\overline{z})} = \Phi_{\mathbf{1}-\widetilde{f}}(z)$. Thus, the kth Fourier coefficient of $\varphi_{\mathbf{1}-\widetilde{f}}$ equals the complex conjugate of the kth Fourier coefficient of $\varphi_{\widetilde{f}}$. Using the method of Proposition 6.10 with cylinder events of length 3, which requires knowledge of Fourier coefficients only for $|k| \le 2$, we obtain that $H(\mathbf{P}^{\widetilde{f}}) \ge 0.4105$. Therefore $H(\mathbf{P}^f) > 0.4105 - 0.0561 = 0.3544 > (1/2)\log 2$. (If we use cylinder events of length 12, then we obtain the bound $H(\mathbf{P}^f) > 0.4442$. However, we believe that the true entropy is significantly larger still.) ∎

Example 6.14. We can also use these bounds to prove that $H(\mathbf{P}^f)$ does *not* depend *only* on the distribution of f. For example, consider the function

$$f(t) := \sin^2(\pi t/2)$$

on $[0,1]$, which has the same distribution as the function g of (6.4). An elementary integration shows that

$$\widehat{f}(k) = \begin{cases} 1/2 & \text{if } k = 0, \\ \dfrac{2ki}{(2k-1)(2k+1)\pi} & \text{if } k \ne 0 \end{cases}$$

(where $i := \sqrt{-1}$). To find φ_f, we proceed as follows. Write $F := (1/2)\log f$. The real parts of the integrals giving $\widehat{F}(k)$ can be found in standard tables, while the imaginary parts are derived in Lyons, Paule, and Riese (2002), giving

$$\widehat{F}(k) = \begin{cases} -\log 2 & \text{if } k = 0, \\ -\dfrac{1}{4k} + \dfrac{i}{k\pi}\displaystyle\sum_{j=1}^{k} \dfrac{1}{2j-1} & \text{if } k > 0. \end{cases}$$

As in the proof of Proposition 6.13, we compute

$$\begin{aligned} \Phi_f(z) = 1/2 &+ (-1/4 + i/\pi)z + (-1/16 + i/6\pi - 1/\pi^2)z^2 \\ &- (1/32 - 19i/360\pi + 5/6\pi^2 + 2i/3\pi^3)z^3 \\ &+ (-5/256 + 89i/5040\pi - 27/40\pi^2 - i/\pi^3 + 1/3\pi^4)z^4 + \cdots, \end{aligned}$$

whose coefficients are the Fourier coefficients of φ_f. Since $1-\sin^2(\pi t/2) = \sin^2(\pi(1-t)/2)$, it follows from the definition (6.6) that $\overline{\Phi_f(\overline{z})} = \Phi_{\mathbf{1}-f}(z)$. Since the power series coefficients of $\Phi_{\mathbf{1}-f}$ are the Fourier coefficients of $\varphi_{\mathbf{1}-f}$, we obtain that the kth Fourier coefficient of $\varphi_{\mathbf{1}-f}$ equals the complex conjugate of the kth Fourier coefficient of φ_f. Using the method of Proposition 6.10 with cylinder events of length 8, which requires knowledge of Fourier coefficients only for $|k| \le 7$, we obtain that

$$0.659648 \le H(\mathbf{P}^f) \le 0.684021\,.$$

Comparing with (6.9), we see that $H(\mathbf{P}^f) > H(\mathbf{P}^g)$.

Note that even when neither $\widehat{f}$ nor $\widehat{\varphi_f}$ can be found explicitly, they can always be found by numerical integration and exponentiation, and then one can follow the procedure we have used in this last example. Indeed, that will be done for part of our final example.

EXAMPLE 6.15. As in Example 6.6, let

$$g(x) := \frac{\sin \pi x}{\sqrt{1 + \sin^2 \pi x}}\,.$$

Then the edges of the uniform spanning tree measure in the plane that lie on the x-axis have the law $\mathbf{P}^g$. Write $g_1(x) := \sin^2 \pi x$. The calculations in Example 1.3 show that $\Phi_{g_1}(z) = (1-z)/2$ and $\Phi_{\mathbf{1}+g_1}(z) = [\sqrt{2}+1-(\sqrt{2}-1)z]/2$, whence

$$\Phi_g(z) = \sqrt{\Phi_{g^2}(z)} = \sqrt{\frac{\Phi_{g_1}}{\Phi_{\mathbf{1}+g_1}}} = \sqrt{\frac{1-z}{\sqrt{2}+1-(\sqrt{2}-1)z}}\,.$$

Expansion of $\Phi_g(z)$ in a Maclaurin series gives

$$\frac{1}{(1+\sqrt{2})^{1/2}} - \frac{z}{(1+\sqrt{2})^{3/2}} - \frac{(-1+2\sqrt{2})\,z^2}{2\,(1+\sqrt{2})^{5/2}} - \frac{(5-2\,\sqrt{2})\,z^3}{2\,(1+\sqrt{2})^{7/2}} - \frac{(-27+28\,\sqrt{2})\,z^4}{8\,(1+\sqrt{2})^{9/2}} + \cdots\,,$$

which tells us $\widehat{\varphi_g}$. The transfer currents, which can be calculated by the method in Burton and Pemantle (1993), tell us $\widehat{g}$; for example, $\widehat{g}(k)$ for $k = 0, 1, 2, 3, 4$ is

$$\frac{1}{2}\,, \qquad \frac{1}{2} - \frac{2}{\pi}\,, \qquad \frac{5}{2} - \frac{8}{\pi}\,, \qquad \frac{25}{2} - \frac{118}{3\,\pi}\,, \qquad \frac{129}{2} - \frac{608}{3\,\pi}\,.$$

We use numerical integration to find $\widehat{\varphi_{\mathbf{1}-g}}$. Then if we use cylinder events of length 8, we find that

$$0.69005 \le H(\mathbf{P}^g) \le 0.69013\,.$$

It is interesting how close this is to $\log 2 = 0.69315^-$.

We close our treatment of entropy with some elementary observations. If $f_n \to 1/2$ weak* (meaning that we have weak* convergence of the measures having these densities), must $H(\mathbf{P}^{f_n}) \to \log 2$? The answer is "no", as we now demonstrate. Given any $f : \mathbb{T} \to [0, 1]$ and any integer n, let $M_{\times n} f$ denote the function

$$M_{\times n} f(x) := f(nx)\,.$$

Let $\langle \eta_k^{(n)}\,;\ k \in \mathbb{Z}\rangle$ have the distribution $\mathbf{P}^{M_{\times n} f}$. Since $f(x) = \sum_{k\in\mathbb{Z}} \widehat{f}(k) e^{2\pi i k x}$ in $L^2(\mathbb{T})$, the Fourier expansion of $M_{\times n} f$ is $M_{\times n} f(x) = \sum_{k\in\mathbb{Z}} \widehat{f}(k) e^{2\pi i k n x}$ for $n \ne 0$. In particular, for any $n \ne 0$ and any r, the processes $\langle \eta_{nk+r}^{(n)}\,;\ k \in \mathbb{Z}\rangle$ each have distribution $\mathbf{P}^f$ and are independent of each other as r ranges from 0 to $n-1$. Therefore $H(\mathbf{P}^{M_{\times n} f}) = H(\mathbf{P}^f)$. Note that unless f is constant, $H(\mathbf{P}^{M_{\times n} f}) = H(\mathbf{P}^f) < H\big[\widehat{f}(0)\big]$, even though $M_{\times n} f$ tends weak* to the constant function $\widehat{f}(0) \cdot \mathbf{1}$. One can show that a similar phenomenon holds in higher dimensions, that is, if $f : \mathbb{T}^d \to [0, 1]$ and $A : \mathbb{T}^d \to \mathbb{T}^d$ is a group epimorphism, then $H(\mathbf{P}^f) = H(\mathbf{P}^{f\circ A})$.

§7. Phase Multiplicity.

In this section, we classify exactly the set of functions f for which $\mathbf{P}^f$ satisfies a strong full K property. This property, which we now describe, is an essential strengthening of the usual Kolmogorov or K property. One of the reasons this property is interesting is that it is closely connected to the notion of multiplicity for Gibbs states in statistical mechanics; in particular, it corresponds to *uniqueness*. In the next section, we classify exactly the set of functions f for which $\mathbf{P}^f$ satisfies a (1-sided) strong K property.

In this section and the next, we always assume that f is not identically 0 nor identically 1.

To begin, we first recall the K property for stationary processes indexed by $\mathbb{Z}$. There are many equivalent formulations, of which we choose an appropriate one.

For $S \subseteq \mathbb{Z}^d$, write $\mathcal{F}(S)$ for the σ-field on $2^{\mathbb{Z}^d}$ generated by $\eta(e)$ for $e \in S$.

Definition 7.1. A translation-invariant probability measure μ on $2^{\mathbb{Z}}$ is K (or a **Kolmogorov automorphism**) if for any finite cylinder event E and for all $\epsilon > 0$ there exists an N such that

$$\mu\Big[\big|\mu\big(E \mid \mathcal{F}((-\infty, -N] \cap \mathbb{Z})\big) - \mu(E)\big| \geq \epsilon\Big] \leq \epsilon\,.$$

It is well known that the above definition of K is equivalent to having a trivial (1-sided) tail σ-algebra in the sense that the σ-algebra $\bigcap_{m\geq 1} \mathcal{F}\big((-\infty, -m]\big)$ is trivial (see page 120 of Georgii (1988) for a version of this). The K property is known to be an isomorphism invariant and is also known to be equivalent to the property that all nontrivial factors have strictly positive entropy. The latter implies that a process is K iff its time reversal is K, something which is not immediate from the definition.

In order to give a complete discussion of the points that we wish to make, we need to introduce three further properties, which are respectively called the full K property, the strong full K property, and the (1-sided) strong K property, the last property given only for $d = 1$.

While the notion of the K property generalizes to $\mathbb{Z}^d$ (see Conze (1972/73)), there is a slight strengthening of the definition that has a more aesthetic extension to $\mathbb{Z}^d$. To give this, define $B_n^d := [-n, n]^d \cap \mathbb{Z}^d$.

Definition 7.2. A translation-invariant probability measure μ on $2^{\mathbb{Z}^d}$ is **full** K if for any finite cylinder event E and for all $\epsilon > 0$ there exists an N such that

$$\mu\Big[\big|\mu\big(E \mid \mathcal{F}((B_N^d)^c)\big) - \mu(E)\big| \geq \epsilon\Big] \leq \epsilon\,.$$

Analogously to an earlier statement, the full K property is equivalent (see again page 120 of Georgii (1988)) to having a **trivial full tail**, which means that $\bigcap_{m\geq 1} \mathcal{F}\big((B_m^d)^c\big)$ is trivial. It is well known that K does not imply full K even for $d = 1$ (see, for example, the bilaterally deterministic Bernoulli shift processes constructed in Ornstein and Weiss (1975)). However, for Markov random fields, the two notions are equivalent (see den Hollander and Steif (1997, 2000)). Lyons (2003b) proved that all (not necessarily $\mathbb{Z}^d$-invariant) determinantal probability measures satisfy the full K property. For Gibbs states arising in statistical mechanics, this property is equivalent to an "extremality" property of the Gibbs state (see page 118 of Georgii (1988) for precisely what this means and this equivalence).

The following definitions strengthen the K property in an even more essential way. We begin with the full version, which seems to us more natural.

Definition 7.3. A translation-invariant probability measure μ on $2^{\mathbb{Z}^d}$ is **strong full** K if for any finite cylinder event E and for all $\epsilon > 0$ there exists an N such that

$$\left|\mu\big(E \mid \mathcal{F}((B_N^d)^c)\big) - \mu(E)\right| < \epsilon \quad \mu\text{-a.e.}$$

For the full K property, "μ-most" conditionings far away have little effect on a "local event", while in the strong full K case, *all* conditionings far away have little effect on a "local event". This is a substantial difference. An example that illustrates this difference is the Ising model in $\mathbb{Z}^2$. The plus state for the Ising model at high temperatures is strong full K, while at low temperatures, it is full K (because it is extremal) but not strong full K.

Finally, if $d = 1$, we define strong K in the following way (which the reader can presumably anticipate). One extension of this definition to $\mathbb{Z}^d$ (among various possible) will be given in Section 8.

Definition 7.4. A translation-invariant probability measure μ on $2^{\mathbb{Z}}$ is **strong** K if for any finite cylinder event E and for all $\epsilon > 0$ there exists an N such that

$$\left|\mu\big(E \mid \mathcal{F}((-\infty, -N] \cap \mathbb{Z})\big) - \mu(E)\right| < \epsilon \quad \mu\text{-a.e.}$$

We note that this definition is closely related to, but weaker than, the ψ-mixing property (see Bradley (1986)). The difference is that the event E here is specified in advance, rather than having a uniformity for all events E, provided only that they depend on the random variables with positive index. If we extend the notion of strong K in the obvious way to the case when $\{0, 1\}$ is replaced by a countable set, then the example in Bradley (1986) of a process that is ψ-mixing, but whose time reversal is not ψ-mixing, yields an example of a strong K process whose time reversal is not. This cannot occur for our measures $\mathbf{P}^f$ since they are all time reversible.

The following development has an analogy with the “plus and minus states” in the Ising model from statistical mechanics; however, such knowledge is not needed by the reader.

Consider a function $f : \mathbb{T}^d \to [0,1]$ and consider the corresponding probability measure $\mathbf{P}^f$ on $2^{\mathbb{Z}^d}$. We shall define a probability measure $(\mathbf{P}^f)^+$ on $2^{\mathbb{Z}^d}$ which will be “$\mathbf{P}^f$ conditioned on all 1s at ∞”; more specifically (but still not precisely), we want to define $(\mathbf{P}^f)^+$ by

$$(\mathbf{P}^f)^+ := \lim_{n\to\infty} \mathbf{P}^f[\,\cdot \mid \eta \equiv 1 \text{ on } (B_n^d)^c]\,.$$

To make sure this is well defined, we proceed in stages. Let

$$((\mathbf{P}^f)^+)_n := \lim_{k\to\infty} \mathbf{P}^f[\,\cdot \mid \eta \equiv 1 \text{ on } B_{n+k}^d \backslash B_n^d]\,.$$

This limit is taken in the weak*-topology. Proposition 2.6 implies that, when restricted to B_n^d, this sequence is stochastically decreasing and hence necessarily converges. $((\mathbf{P}^f)^+)_n$ is therefore well defined. One next defines

$$(\mathbf{P}^f)^+ := \lim_{n\to\infty} ((\mathbf{P}^f)^+)_n\,,$$

where again the limit is taken in the weak*-topology. Proposition 2.6 again implies that for fixed k, $((\mathbf{P}^f)^+)_n$ restricted to B_k^d is stochastically increasing in n for $n > k$ and hence converges. This implies that its limit $(\mathbf{P}^f)^+$ is well defined and completes the definition of $(\mathbf{P}^f)^+$. The stochastic monotonicity results also imply that $(\mathbf{P}^f)^+ \preccurlyeq \mathbf{P}^f$ and that $(\mathbf{P}^f)^+$ is translation invariant.

In exactly the same way (using 0 instead of 1 boundary conditions), one defines $(\mathbf{P}^f)^-$, which satisfies $\mathbf{P}^f \preccurlyeq (\mathbf{P}^f)^-$. Analogy with the ferromagnetic Ising model in statistical mechanics leads us to the following definition.

Definition 7.5. The probability measure $\mathbf{P}^f$ has **phase multiplicity** if $(\mathbf{P}^f)^- \neq (\mathbf{P}^f)^+$, and otherwise has **phase uniqueness**.

Lemma 7.6. *$\mathbf{P}^f$ has phase uniqueness if and only if it is strong full K.*

The reason this is true is that Proposition 2.6 implies that the most extreme boundary conditions are when all 1s or all 0s are used and the measures corresponding to any other boundary conditions are “stochastically trapped” between these two special cases. A detailed proof follows straightforwardly from the stochastic monotonicity arguments above and is left to the reader.

We shall now obtain necessary and sufficient conditions for the equality of $\mathbf{P}^f$ and $(\mathbf{P}^f)^+$. From these, a necessary and sufficient condition for the strong full K property will easily emerge.

Let $\mathcal{T}$ denote the set of trigonometric polynomials on $\mathbb{T}^d$. Let $L^2(1/f)$ denote the set

$$\left\{h : \mathbb{T}^d \to \mathbb{C}\,;\ \int_{\mathbb{T}^d} \frac{|h|^2}{f}\, d\lambda_d < \infty\right\}.$$

Here we use the convention that $0/0 := 0$. Note also that h needs to vanish where f does.

Our main result on phase multiplicity is the following.

Theorem 7.7. *Assume that $f : \mathbb{T}^d \to [0, 1]$ is measurable. The following are equivalent.*

(i) $(\mathbf{P}^f)^+ = \mathbf{P}^f$*;*

(ii) *f is in the closure in $L^2(1/f)$ of $\mathcal{T} \cap L^2(1/f)$;*

(iii) *There exists a nonzero trigonometric polynomial T such that $\frac{|T|^2}{f} \in L^1(\mathbb{T}^d)$; i.e., $\mathcal{T} \cap L^2(1/f) \neq 0$.*

Moreover, if $(\mathbf{P}^f)^+ \neq \mathbf{P}^f$, then $(\mathbf{P}^f)^+ = \delta_{\mathbf{0}}$.

Proof. We shall first show that (i) and (ii) are equivalent and then that (ii) and (iii) are equivalent.

(i) implies (ii): Let u_n be the element in $[B]_f$ achieving the minimum in (2.5) for $B := (B_n^d)^c$. Then

$$\|\mathbf{1}\|_f^2 = \|\mathbf{1} - u_n\|_f^2 + \|u_n\|_f^2\,.$$

By (i) and (2.5), we have $\|\mathbf{1} - u_n\|_f \to \|\mathbf{1}\|_f$ as $n \to \infty$, or in other words, $\|u_n\|_f \to 0$ as $n \to \infty$. Furthermore, $\mathbf{1} - u_n \perp [(B_n^d)^c]_f$ in $L^2(f)$, which is the same as $(\mathbf{1} - u_n)f$ being a trigonometric polynomial T_n with $\widehat{T_n}$ supported in B_n^d. We have that

$$\|f - T_n\|_{(1/f)}^2 = \int |f - T_n|^2 \frac{1}{f}\, d\lambda_d = \int \left|\mathbf{1} - \frac{T_n}{f}\right|^2 f\, d\lambda_d = \int |u_n|^2 f\, d\lambda_d = \|u_n\|_f^2 \tag{7.1}$$

tends to 0 as $n \to \infty$, which proves (ii).

(ii) implies (i): Given $\epsilon > 0$, let T be a trigonometric polynomial with $\|f - T\|_{(1/f)} < \epsilon$. Let n be such that $\widehat{T}$ is supported in B_n^d. Define $v := T/f$, which is in $L^2(f)$. Then $v \perp [(B_n^d)^c]_f$ in $L^2(f)$ and $\|\mathbf{1} - v\|_f = \|f - T\|_{(1/f)} < \epsilon$, so $\|P^{\perp}_{[(B_n^d)^c]_f}\mathbf{1}\|_f^2 > \|\mathbf{1}\|_f^2 - \epsilon^2$. Combining this with (2.6), we see that $(\mathbf{P}^f)^+[\eta(\mathbf{0}) = 1] > \mathbf{P}^f[\eta(\mathbf{0}) = 1] - \epsilon^2$. Since this holds for any $\epsilon > 0$, we obtain that $(\mathbf{P}^f)^+[\eta(\mathbf{0}) = 1] \geq \mathbf{P}^f[\eta(\mathbf{0}) = 1]$. However, since $(\mathbf{P}^f)^+ \preccurlyeq \mathbf{P}^f$ and both are translation invariant, the two measures are actually equal by Lemma 3.3, which proves (i).

(ii) implies (iii): This is immediate.

(iii) implies (ii): Assume there exists a nonzero trigonometric polynomial $T \in L^2(1/f)$. Let $\epsilon > 0$. Choose $B \subseteq \mathbb{T}^d$ such that

$$\int_B \left(\frac{|T|^2}{f} + 2|T| + f\right) < \epsilon$$

and B is an open set containing the zero set of T (which is a set of measure 0, as remarked in the proof of Theorem 4.2). Let $g := T\mathbf{1}_B + f\mathbf{1}_{B^c}$. Then $g/T \in L^\infty(\mathbb{T}^d)$ since T is bounded away from 0 on B^c and $0 \le f \le 1$. We can now choose trigonometric polynomials p_n on $\mathbb{T}^d$ such that $p_n \to g/T$ a.e. on $\mathbb{T}^d$ with $\|p_n\|_\infty \le \|g/T\|_\infty$ for all n (just use $p_n := (g/T) * K_n^d$, where K_n^d is, as before, the Fejér kernel for $\mathbb{T}^d$). Since $|T|^2/f \in L^1(\mathbb{T}^d)$, it follows from the Lebesgue dominated convergence theorem that

$$\lim_{n\to\infty} \int_{\mathbb{T}^d} \frac{|T|^2}{f} \left| \frac{g}{T} - p_n \right|^2 = 0 \,.$$

Hence there exists a trigonometric polynomial h such that

$$\int_{\mathbb{T}^d} \frac{|T|^2}{f} \left| \frac{g}{T} - h \right|^2 < \epsilon \,. \tag{7.2}$$

Minkowski's inequality (applied to $L^2(|T|^2/f)$) yields

$$\left(\int_{\mathbb{T}^d} \frac{|T|^2}{f} \left| \frac{f}{T} - h \right|^2 \right)^{\frac{1}{2}} \le \left(\int_{\mathbb{T}^d} \frac{|T|^2}{f} \left| \frac{f}{T} - \frac{g}{T} \right|^2 \right)^{\frac{1}{2}} + \left(\int_{\mathbb{T}^d} \frac{|T|^2}{f} \left| \frac{g}{T} - h \right|^2 \right)^{\frac{1}{2}} .$$

The second summand is at most $\epsilon^{1/2}$ by (7.2). The first summand is

$$\left(\int_B \frac{|T|^2}{f} \left| \frac{f}{T} - 1 \right|^2 \right)^{\frac{1}{2}} \le \left(\int_B f + 2|T| + \frac{|T|^2}{f} \right)^{\frac{1}{2}} < \epsilon^{\frac{1}{2}}$$

by choice of B. Hence $\int_{\mathbb{T}^d} |f - Th|^2 \frac{1}{f}\, dx < 4\epsilon$. Since Th is a trigonometric polynomial, (ii) is proved.

Finally, assume that $(\mathbf{P}^f)^+ \ne \mathbf{P}^f$. Since (iii) fails, $[(B_n^d)^c]_f^\perp = \mathbf{0}$ for all n, which means by (2.6) that $(\mathbf{P}^f)^+[\eta(\mathbf{0}) = 1] = 0$, or, in other words, $(\mathbf{P}^f)^+ = \delta_{\mathbf{0}}$. ∎

EXAMPLE 7.8. If $f : \mathbb{T} \to [0, 1]$ is continuous and has a finite number of 0s with f approaching each of these 0s at most polynomially quickly, then $(\mathbf{P}^f)^+ = \mathbf{P}^f$ since it is easy to construct a trigonometric polynomial T with the same 0s as f and approaching 0 at least as quickly as f (for example, T could be of the form $\prod_{i=1}^k \sin^n 2\pi(x - a_i)$). In particular, if f is real analytic and not $\mathbf{0}$, then $(\mathbf{P}^f)^+ = \mathbf{P}^f$.

EXAMPLE 7.9. If f vanishes on a set of positive measure, then $(\mathbf{P}^f)^+ = \delta_{\mathbf{0}}$.

EXAMPLE 7.10. If $f : \mathbb{T} \to [0, 1]$ is a continuous function with a single zero at $x_0 \in \mathbb{T}$ and $f(x) = e^{-1/|x - x_0|}$ in some neighborhood of x_0, then $(\mathbf{P}^f)^+ = \delta_{\mathbf{0}}$. Indeed, there is no nonzero trigonometric polynomial T with $|T|^2/f \in L^1(\mathbb{T})$ since the rate at which a trigonometric polynomial approaches 0 is at most polynomially quickly.

Example 7.11. We give an example to show that the property $(\mathbf{P}^f)^+ = \mathbf{P}^f$ does *not* depend only on the distribution of f, even among real-analytic functions f. In 2 dimensions, the function

$$f(x,y) := \sin^2(2\pi y - \cos 2\pi x)$$

generates a system for which $(\mathbf{P}^f)^+ \neq \mathbf{P}^f$. This is because f vanishes (even to second order) on a curve

$$y = \frac{1}{2\pi}\cos 2\pi x + \mathbb{Z} \tag{7.3}$$

that is not in the zero set of any trigonometric polynomial in two variables except the zero polynomial. However, f has the same distribution as the trigonometric polynomial $g(x,y) := \sin^2 2\pi y$ and $(\mathbf{P}^g)^+ = \mathbf{P}^g$. To see that the curve (7.3) does not lie in the zero set of any nonzero trigonometric polynomial, suppose that $T(x,y)$ is a trigonometric polynomial that vanishes there. Write $w := e^{2\pi ix}$ and $z := e^{2\pi iy}$. By multiplying by a suitable complex exponential, we may assume that $P(w,z) = T(x,y)$ is an analytic polynomial. Our assumption is that

$$h(w) := P(w, e^{i(w+w^{-1})/2})$$

satisfies $h(w) = 0$ for all $|w| = 1$. Since h is an analytic function for $w \neq 0$, it follows that $h(w) = 0$ for all $w \neq 0$. Now for each $z \neq 0$, there are an infinity of w such that $e^{i(w+w^{-1})/2} = z$. Therefore for each $z \neq 0$, there are an infinity of w such that $P(w,z) = 0$. Since a polynomial has only a finite number of zeroes if it is not identically zero, this means that for each $z \neq 0$, $P(w,z) = 0$ for all $w \in \mathbb{C}$. Hence $P \equiv 0$, as desired.

Example 7.12. For $d = 2$, if f is real analytic on a neighborhood of its zero set, then $(\mathbf{P}^f)^+ = \mathbf{P}^f$ iff its zero set is contained in a (nontrivial) algebraic variety, where we view $\mathbb{T}^2$ as $\{(z_1, z_2) \in \mathbb{C}^2\,;\ |z_1| = |z_2| = 1\}$. This is because the slowest f can vanish at a point is of order $x^2 + y^2$ (since the constant and linear terms of f must vanish) and $1/(x^2+y^2)$ is not integrable. Therefore, all the zeroes of f must be cancelled by those of T. Conversely, if the zero set of f is contained in the zero set of a trigonometric polynomial T, then by Łowasiewicz's inequality (see, e.g., Bierstone and Milman (1988), Theorem 6.4), for a sufficiently large n, we have $|T|^n/f$ is bounded, whence integrable.

Example 7.13. Here, we give a 1-dimensional example showing that the property $(\mathbf{P}^f)^+ = \mathbf{P}^f$ does *not* depend only on the distribution of f, even among continuous f. Let $f(x) :=$

$\sqrt{x}|\sin(\pi/x)|$ on $[0,1]$. Then $(\mathbf{P}^f)^+ \neq \mathbf{P}^f$ since there are an infinite number of first-order zeroes. However, let g be the increasing rearrangement of f on $[0,1]$ and define

$$h(x) := \begin{cases} g(2x) & \text{if } x \leq 1/2, \\ g(2-2x) & \text{if } x \geq 1/2. \end{cases}$$

Then h is continuous and has the same distribution as f, yet $(\mathbf{P}^h)^+ = \mathbf{P}^h$ by Example 7.8 together with an easy computation. Of course, there is no such example for real-analytic f in one dimension by Example 7.8.

We finally state and prove our necessary and sufficient condition for the strong full K property.

Corollary 7.14. *Consider $f : \mathbb{T}^d \to [0,1]$ and the corresponding probability measure $\mathbf{P}^f$. Then $\mathbf{P}^f$ is strong full K if and only if there is a nonzero trigonometric polynomial T such that $\frac{|T|^2}{f(\mathbf{1}-f)} \in L^1(\mathbb{T}^d)$.*

Proof. According to Theorem 7.7, $(\mathbf{P}^f)^+ = \mathbf{P}^f$ iff there exists a nonzero trigonometric polynomial T_1 such that $\frac{|T_1|^2}{f} \in L^1(\mathbb{T}^d)$. The same argument applied to $\mathbf{1} - f$ tells us that $(\mathbf{P}^f)^- = \mathbf{P}^f$ iff there exists a nonzero trigonometric polynomial T_2 such that $\frac{|T_2|^2}{\mathbf{1}-f} \in L^1(\mathbb{T}^d)$. Lemma 7.6 now completes the proof. (Take $T := T_1 T_2$.) ∎

Remark 7.15. Observe that by trivial scaling, we could have an f whose Fourier coefficients go to zero slowly but which is bounded away from 0 and 1. Hence slow decay of Fourier coefficients does not imply phase multiplicity. On the other hand, if the coefficients decay exponentially, then $\sum_{n\in\mathbb{Z}} \widehat{f}(n) z^n$ is complex-analytic in an annulus around the unit circle, which implies that its restriction to the unit circle is real-analytic and hence there is phase uniqueness.

Remark 7.16. We mention that if $(\mathbf{P}^f)^+ = \mathbf{P}^f$, then this measure is Følner independent in the sense of Adams (1992) and if in addition $(\mathbf{P}^f)^- = \mathbf{P}^f$, then this measure is even strong Følner independent in the sense that any conditioning outside a box yields a measure which is $\overline{d}$ close to the unconditioned process. The arguments for proving these facts are analogous to those of Ornstein and Weiss (1973) (see Adams (1992) for a published version), where it is proved that the plus state for the Ising model is a Bernoulli shift. (The concept of Følner independence has been used also in den Hollander and Steif (1997) and Hoffman (1999).)

We close this section with an interpretation of our phase multiplicity results in terms of interpolation questions for wide-sense stationary processes. The following comments can be proved from Theorem 2.8 together with the well-known correspondence between interpolation and Szegő infima. First, $(\mathbf{P}^f)^+ = \delta_{\mathbf{0}}$ is equivalent to being able to interpolate perfectly from information far away for a wide-sense stationary process $\langle Y_n \rangle$ with spectral density f, which means that for every n, Y_0 is in the closed linear subspace spanned by $\{Y_k\}_{|k|\geq n}$. For $d = 1$, it is stated on p. 102 of Rozanov (1967) that this latter condition is equivalent to the negation of condition (iii) in Theorem 7.7. A similar statement holds for $(\mathbf{P}^f)^-$ with f replaced by $\mathbf{1} - f$. The proof of Theorem 7.7 implies (via this whole correspondence) that, for spectral measures that are absolutely continuous with a bounded nonnegative density, if perfect linear interpolation fails, then our ability to interpolate as n gets large goes to 0; i.e., the length of the projection of Y_0 onto the closed linear subspace spanned by $\{Y_k\}_{|k|\geq n}$ goes to 0 as $n \to \infty$.

§8. 1-Sided Phase Multiplicity.

We first study the notion of strong K for $d = 1$. This is natural since the structure of "one-sided" behavior can be very different from "two-sided" behavior in various situations; for example, the existence of bilaterally deterministic Bernoulli shift processes demonstrates this. In contrast with Example 7.13, we shall see that whether $\mathbf{P}^f$ is strong K depends only on the distribution of f.

Consider a function $f : \mathbb{T} \to [0, 1]$ and the corresponding probability measure $\mathbf{P}^f$ on $2^{\mathbb{Z}}$. We define a probability measure $(\mathbf{P}^f)^{+,1}$ on $2^{\mathbb{Z}}$ thought of as "$\mathbf{P}^f$ conditioned on all 1s at $-\infty$" (the superscript "1" refers to the fact that we are doing this on 1 side); namely,

$$(\mathbf{P}^f)^{+,1} := \lim_{n\to\infty} \lim_{k\to\infty} \mathbf{P}^f \big[\, \cdot \mid \eta \equiv 1 \text{ on } [-n-k, -n] \cap \mathbb{Z} \big] \,.$$

The existence of the limit and its translation invariance follow from stochastic monotonicity, as did that of $(\mathbf{P}^f)^+$. As before, we also have $(\mathbf{P}^f)^{+,1} \preccurlyeq \mathbf{P}^f$. In the analogous way (using 0 instead of 1 boundary conditions), one defines $(\mathbf{P}^f)^{-,1}$, which then satisfies $\mathbf{P}^f \preccurlyeq (\mathbf{P}^f)^{-,1}$.

Definition 8.1. The probability measure $\mathbf{P}^f$ has **1-sided phase multiplicity** if $(\mathbf{P}^f)^{-,1} \neq (\mathbf{P}^f)^{+,1}$, and otherwise has **1-sided phase uniqueness**.

Note the following analogue (whose proof is left to the reader) of Lemma 7.6 for the 1-sided case.

Lemma 8.2. $\mathbf{P}^f$ *has 1-sided phase uniqueness if and only if it is strong* K.

Our main result for 1-sided phase multiplicity is the following.

Theorem 8.3. *If* $f : \mathbb{T} \to [0, 1]$*, then* $(\mathbf{P}^f)^{+,1} = \mathbf{P}^f$ *iff* $\mathsf{GM}(f) > 0$*. Moreover, if* $(\mathbf{P}^f)^{+,1} \neq \mathbf{P}^f$*, then* $(\mathbf{P}^f)^{+,1} = \delta_{\mathbf{0}}$*.*

Proof. By Theorem 6.12, we have

$$\lim_{k\to\infty} \mathbf{P}^f[\eta(0) = 1 \mid \eta_{-n} = \eta_{-n-1} = \cdots = \eta_{-n-k} = 1] = \sum_{l=0}^{n-1} |\widehat{\varphi_f}(l)|^2 .$$

(In fact, via Theorem 2.8, this special case of Theorem 6.12 is due to Kolmogorov and Wiener; see Grenander and Szegő (1984), Section 10.9.) Taking $n \to \infty$, we find that

$$(\mathbf{P}^f)^{+,1}\big[\eta(0) = 1\big] = \|\widehat{\varphi_f}\|_2^2 = \|\varphi_f\|_2^2 .$$

If $\mathsf{GM}(f) > 0$, then we obtain

$$(\mathbf{P}^f)^{+,1}\big[\eta(0) = 1\big] = \|f\|_1 = \widehat{f}(0) = \mathbf{P}^f\big[\eta(0) = 1\big] .$$

Since $(\mathbf{P}^f)^{+,1} \preccurlyeq \mathbf{P}^f$ and both probability measures are $\mathbb{Z}$-invariant, it follows by Lemma 3.3 that $(\mathbf{P}^f)^{+,1} = \mathbf{P}^f$. On the other hand, if $\mathsf{GM}(f) = 0$, then

$$(\mathbf{P}^f)^{+,1}\big[\eta(0) = 1\big] = 0 ,$$

so $(\mathbf{P}^f)^{+,1} = \delta_{\mathbf{0}}$. ∎

Our necessary and sufficient condition for the strong K property now follows immediately.

Corollary 8.4. *For* $f : \mathbb{T} \to [0, 1]$*, the corresponding probability measure* $\mathbf{P}^f$ *is strong* K *if and only if* $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1} - f) > 0$*.*

We now construct a process $\mathbf{P}^f$ that is strong K but not strong full K.

Theorem 8.5. *There exists an* $f : \mathbb{T} \to [0, 1]$ *such that* $\mathbf{P}^f$ *is strong* K *but not strong full* K*.*

Proof. Choose f such that f is continuous, bounded away from 1, vanishes at a single point x_0, and such that $f(x) = e^{-1/(|x-x_0|)^{1/2}}$ in some neighborhood of x_0. Since a trigonometric polynomial vanishes at its zeroes at most polynomially quickly, there cannot exist a nonzero trigonometric polynomial T such that $|T|^2/f \in L^1(\mathbb{T})$. Hence by Corollary 7.14, $\mathbf{P}^f$ is not strong full K. On the other hand, it is clear that $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1}-f) > 0$, so by Corollary 8.4, we know that $\mathbf{P}^f$ is strong K. ∎

Remark 8.6. By combining Theorems 5.3 and 8.3, we see that $\mathbf{P}^f \succcurlyeq \mathbf{P}^g$ implies $(\mathbf{P}^f)^{+,1} \succcurlyeq (\mathbf{P}^g)^{+,1}$. However, it is not necessarily the case that $(\mathbf{P}^f)^+ \succcurlyeq (\mathbf{P}^g)^+$. For example, let $g \equiv 1/2$ and f be a function as described in the last result, but which has a geometric mean larger than 1/2 (recall Theorem 5.3). This example also shows that if $\mathbf{P}^g \preccurlyeq \mathbf{P}^f$, there do not necessarily exist $g' \le f'$ such that $\mathbf{P}^{g'} = \mathbf{P}^g$ and $\mathbf{P}^{f'} = \mathbf{P}^f$. To see this, let f and g be as above. Then $\mathbf{P}^{g'} = \mathbf{P}^g$ easily implies that $g' \equiv 1/2$, which in turn (using $g' \le f'$) implies that $(\mathbf{P}^{f'})^+ = \mathbf{P}^{f'}$. Since $(\mathbf{P}^f)^+ \neq \mathbf{P}^f$, this contradicts the fact that $\mathbf{P}^f = \mathbf{P}^{f'}$.

It may be of interest to see another proof (which we only sketch) of Theorem 8.3 that does not depend on the asymptotics of Theorem 6.12, but rather follows the lines of the proof of Theorem 7.7. The appropriate replacement of the trigonometric polynomials for this question is the set

$$\mathcal{A} := \{h \in L^2(\mathbb{T});\ \text{ there exists } \ell \text{ such that } \widehat{h}(k) = 0 \text{ for } k < \ell\}\,.$$

Note that when ℓ is taken to be 0, we get the **Hardy space** of analytic functions, denoted $H^2(\mathbb{T})$.

Theorem 8.7. *Assume that $f : \mathbb{T} \to [0,1]$ and let $\mathcal{A}$ be as defined as above. The following are equivalent.*

(i) $(\mathbf{P}^f)^{+,1} = \mathbf{P}^f$;

(ii) *f is in the closure in $L^2(1/f)$ of $\mathcal{A} \cap L^2(1/f)$;*

(iii) *There exists a nonzero element $T \in \mathcal{A}$ such that $\frac{|T|^2}{f} \in L^1(\mathbb{T})$; i.e., $\mathcal{A} \cap L^2(1/f) \neq 0$;*

(iv) $\mathsf{GM}(f) > 0$.

Moreover, if $(\mathbf{P}^f)^{+,1} \neq \mathbf{P}^f$, then $(\mathbf{P}^f)^{+,1} = \delta_{\mathbf{0}}$.

Proof. **(i) iff (ii)** and **(ii) implies (iii):** These are proved exactly as in Theorem 7.7.

(iii) implies (ii): Fix a nonzero $T \in \mathcal{A}$ with $\frac{|T|^2}{f} \in L^1(\mathbb{T})$. Given any ϵ, there exists $\delta > 0$ such that if B is any subset of $\mathbb{T}$ with measure less than δ, then

$$\int_B \left(\frac{|T|^2}{f} + 2|T| + f \right) < \epsilon\,.$$

Since T vanishes only on a set of measure 0 (see Rudin (1987), Theorem 17.18, p. 345), there is a $\gamma > 0$ such that $\{x \in \mathbb{T};\ |T(x)| < \gamma\}$ has measure less than δ. If we take B to be this latter set, then we can proceed exactly as in Theorem 7.7 since now T is bounded away from 0 on B^c.

(iii) implies (iv): Any function in $\mathcal{A}$ is a complex exponential times a function in the Hardy space $H^2(\mathbb{T})$. Hence there is a nonzero function $T \in H^2(\mathbb{T})$ with $|T|^2/f \in L^1(\mathbb{T})$.

Now $\mathsf{GM}(|T|) > 0$ (see Rudin (1987), Theorem 17.17, p. 344). Letting $g := |T|^2/f$, we have that $g \in L^1(\mathbb{T})$ and hence $\mathsf{GM}(g) \le \int g < \infty$. Therefore $\mathsf{GM}(f) = \mathsf{GM}(|T|^2/g) = \mathsf{GM}(|T|)^2/\mathsf{GM}(g) > 0$, as desired.

(iv) implies (iii): If $\mathsf{GM}(f) > 0$, then take $T := \varphi_f$, defined in Section 5. For this T, we have $|T|^2/f = \mathbf{1}$, which is trivially integrable.

The final statement is proved as in Theorem 7.7. ∎

We now extend Theorem 8.3 and Corollary 8.4 to higher dimensions. Fix a choice of ordering of $\mathbb{Z}^d$ given by a set Π and a sequence $\langle k_n \rangle \in \mathbb{Z}^d$ such that the distance from $\mathbf{0}$ to $-(\Pi + k_n)$ goes to ∞.

Definition 8.8. A translation-invariant probability measure μ on $2^{\mathbb{Z}^d}$ **has phase uniqueness relative to the ordering induced by** Π **and the sequence** $\langle k_n \rangle$ as above if for any finite cylinder event E and for all $\epsilon > 0$ there exists an N such that for all $n \ge N$

$$\left|\mu\big(E \mid \mathcal{F}(-(\Pi + k_n))\big) - \mu(E)\right| < \epsilon \quad \mu\text{-a.e.}$$

It is easy to check using Proposition 2.6 that this definition *does not depend* on the choice of the sequence $\langle k_n \rangle$. However, it turns out that the choice of ordering, Π, of $\mathbb{Z}^d$ can make a difference. In particular, if the ordering is archimedean (meaning that for any two positive elements a and b, there exists an integer n such that $na > b$), such as when the ordering is induced by $\Pi := \{k \in \mathbb{Z}^d ;\ k \cdot x > 0\}$, where $x \in \mathbb{R}^d$ is a fixed vector having two coordinates with an irrational ratio,* then we have a complete characterization of 1-sided phase multiplicity in terms of the geometric mean. For the standard lexicographic ordering, however, 1-sided phase multiplicity cannot be characterized by the geometric mean alone, as shown by the example $f := \mathbf{1}_{[0,1]\times[0,1/2]}$ in two dimensions. Here, both f and $\mathbf{1} - f$ have 0 geometric mean, but the columns of a configuration are independent under $\mathbf{P}^f$, so that conditioning on the remote past has no effect on the present. On the other hand, positivity of the geometric means will still *suffice* for phase uniqueness relative to the ordering induced by Π, as we now prove. Analogously to the one-dimensional case, we let, for our given ordering Π,

$$(\mathbf{P}^f)^{+,1} := \lim_{n\to\infty} \mathbf{P}^f\big[\,\cdot \mid \eta \equiv 1 \text{ on } -(\Pi + k_n)\big]\,.$$

By Proposition 2.6, this limit exists and is independent of $\langle k_n \rangle$. As before, $(\mathbf{P}^f)^{-,1}$ is defined analogously using boundary conditions of 0s, one-sided phase multiplicity is defined by $(\mathbf{P}^f)^{-,1} \ne (\mathbf{P}^f)^{+,1}$, and the property in Definition 8.8 is equivalent to one-sided phase uniqueness.

* In fact, all archimedean orders arise in this way; see Teh (1961), Zaĭceva (1953), or Trevisan (1953).

Theorem 8.9. *Let* $f : \mathbb{T}^d \to [0,1]$*. If* $\mathsf{GM}(f) > 0$*, then* $(\mathbf{P}^f)^{+,1} = \mathbf{P}^f$*. If* $\mathsf{GM}(f) = 0$ *and the ordering is archimedean, then* $(\mathbf{P}^f)^{+,1} = \delta_{\mathbf{0}}$*. In particular, if the ordering is archimedean, then* $\mathbf{P}^f$ *has phase uniqueness (relative to this ordering) if and only if* $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1} - f) > 0$*.*

We begin with the background results on which we rely for the proof. For an ordering induced by Π, define the corresponding **Helson-Lowdenslager space**

$$\mathsf{HL}^2 := \mathsf{HL}^2(\mathbb{T}^d, \Pi) := \left\{ f \in L^2(\mathbb{T}^d) \, ; \, \operatorname{supp} \widehat{f} \subset \Pi \cup \{\mathbf{0}\} \right\} .$$

For $0 \le f \in L^1(\lambda)$, let $[\Pi \cup \{\mathbf{0}\}]$ be the linear span of $\{\mathbf{e}_k \, ; \, k \in \Pi \cup \{\mathbf{0}\}\}$ and $[\Pi \cup \{\mathbf{0}\}]_f$ be its closure in $L^2(f)$. The replacement for outer functions is the class of **spectral factors**, which are the functions $\varphi \in \mathsf{HL}^2$ with the additional properties

$$\widehat{\varphi}(\mathbf{0}) > 0 \tag{8.1}$$

and

$$1/\varphi \in [\Pi \cup \{\mathbf{0}\}]_{|\varphi|^2} \, . \tag{8.2}$$

Helson and Lowdenslager (1958) show that for any Π and for any $0 \le f \in L^1(\mathbb{T}^d)$, the condition $\mathsf{GM}(f) > 0$ is equivalent to the existence of a spectral factor φ_f such that $|\varphi_f|^2 = f$ a.e. [More precisely, they prove $\mathsf{GM}(f) > 0$ iff $\exists \varphi \in \mathsf{HL}^2$ satisfying $|\varphi|^2 = f$ a.e. and (8.1). Their proof shows that in this case, φ can be chosen so that also (8.2) holds.] Furthermore, Lyons (2003c) shows that φ_f is then unique.

Denote by ν_f the measure on $2^{\Pi \cup \{\mathbf{0}\}}$ given by $\mathbf{P}^f[\,\cdot \mid \eta\restriction(-\Pi) \equiv 1]$. This is, as usual, defined by conditioning on more and more 1s in $-\Pi$ and using Proposition 2.6. The results of Lyons (2003c) imply the following.

Theorem 8.10. *Fix an ordering induced by a set* Π*. Let* $f : \mathbb{T}^d \to [0,1]$ *be measurable. If* $\mathsf{GM}(f) > 0$*, define* φ_f *to be its spectral factor (with respect to* Π*). Otherwise, define* $\varphi_f := \mathbf{0}$*. If* $\mathsf{GM}(f) > 0$ *or the ordering of* $\mathbb{Z}^d$ *given by* Π *is archimedean, then the measure* ν_f *is equal to the determinantal probability measure corresponding to the positive contraction on* $\ell^2(\Pi \cup \{\mathbf{0}\})$ *whose* (j,k)*-matrix entry is*

$$\sum_{l \in \Pi \cup \{\mathbf{0}\}, \; l \preccurlyeq j,k} \overline{\widehat{\varphi_f}(j-l)} \widehat{\varphi_f}(k-l) \, .$$

Proof of Theorem 8.9. Suppose first that $\mathsf{GM}(f) > 0$ or the ordering given by Π is archimedean. By Theorem 8.10, we have

$$\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta\restriction-(\Pi + k_n) \equiv 1] = \mathbf{P}^f[\eta(k_n) = 1 \mid \eta\restriction(-\Pi) \equiv 1] = \sum_{l \in \Pi \cup \{\mathbf{0}\}, \; l \preccurlyeq k_n} |\widehat{\varphi_f}(k_n - l)|^2 \, .$$

Taking $n \to \infty$, we find that

$$(\mathbf{P}^f)^{+,1}\big[\eta(\mathbf{0}) = 1\big] = \|\widehat{\varphi_f}\|_2^2 = \|\varphi_f\|_2^2 \,.$$

If $\mathsf{GM}(f) > 0$, then we obtain

$$(\mathbf{P}^f)^{+,1}\big[\eta(\mathbf{0}) = 1\big] = \|f\|_1 = \widehat{f}(\mathbf{0}) = \mathbf{P}^f\big[\eta(\mathbf{0}) = 1\big] \,.$$

Since $(\mathbf{P}^f)^{+,1} \preccurlyeq \mathbf{P}^f$ and both probability measures are $\mathbb{Z}^d$-invariant, it follows that $(\mathbf{P}^f)^{+,1} = \mathbf{P}^f$. However, if $\mathsf{GM}(f) = 0$ and the ordering is archimedean, then

$$(\mathbf{P}^f)^{+,1}\big[\eta(\mathbf{0}) = 1\big] = 0 \,,$$

so $(\mathbf{P}^f)^{+,1} = \delta_{\mathbf{0}}$. The last statement of the theorem follows as before. ∎

Analogously to Section 7, our one-sided phase multiplicity results for $d = 1$ can be translated to known results for prediction (rather than interpolation) questions for wide-sense stationary processes. We leave these observations to the reader.

§9. Open Questions.

Question 9.1. Calculate $H(\mathbf{P}^f)$.

We conjecture that entropy is concave:

Conjecture 9.2. *For any f and g, we have $H\big(\mathbf{P}^{(f+g)/2}\big) \ge \big(H(\mathbf{P}^f) + H(\mathbf{P}^g)\big)/2$.*

Question 9.3. Letting A vary over all sets of measure $1/2$, how do we get the largest and the smallest entropies for $\mathbf{P}^{\mathbf{1}_A}$? More generally, given f, which g with the same distribution as f maximize or minimize $H(\mathbf{P}^g)$?

Question 9.4. If $d = 1$ and $\mathsf{GM}(f)\mathsf{GM}(\mathbf{1}-f) > 0$, do we get arbitrarily close lower bounds (in principle) on $H(\mathbf{P}^f)$ by the method in Section 5? How can one get close lower bounds on the entropy for higher-dimensional processes?

Question 9.5. Suppose $f : \mathbb{T} \to [0,1]$ is a trigonometric polynomial of degree m. Then $\mathbf{P}^f$ is m-dependent, as are all $(m+1)$-block factors of independent processes (as defined by Hedlund (1969)). Is it the case that $\mathbf{P}^f$ is an $(m+1)$-block factor of an i.i.d. process? Broman (2003) has shown this when $m = 1$. If one could find such block factors sufficiently explicitly for trigonometric polynomials, then one could find explicit factors of i.i.d.

processes that give any process $\mathbf{P}^f$. This would enable one to use more standard probabilistic techniques to study $\mathbf{P}^f$. More generally, in higher dimensions, if $f : \mathbb{T}^d \to [0, 1]$ is a trigonometric polynomial, is $\mathbf{P}^f$ a block factor?

Question 9.6. Given an ordering Π on $\mathbb{Z}^d$ and an increasing sequence of finite sets $\Pi_n \subset \Pi$ whose union is all of Π, when is there phase multiplicity in the sense that

$$\lim_{n\to\infty} \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}(\Pi \setminus \Pi_n) \equiv 1] \neq \lim_{n\to\infty} \mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \eta{\restriction}(\Pi \setminus \Pi_n) \equiv 0]\,?$$

This clearly does not depend on the choice of Π_n. In one dimension, this is the same as 1-sided phase multiplicity, where we know the answer.

Question 9.7. Suppose that $d = 1$. Note that translation and flip of f yield the same measure $\mathbf{P}^f$. Does $\mathbf{P}^f$ determine f up to translation and flip?

We now ask about some properties that are important for models in statistical physics. Let $\mathscr{Z}$ be as in Definition 5.15. A stationary process ν is called **almost surely quasi-local** if $\nu[\eta(\mathbf{0}) = 1 \mid \mathscr{Z}]$ has an almost surely continuous version, where the product topology is used on $2^{\mathbb{Z}^d \setminus \{\mathbf{0}\}}$.

Question 9.8. For which f is $\mathbf{P}^f$ quasi-local or almost surely quasi-local? As we remarked in Section 6, Theorem 6.2 of Shirai and Takahashi (2003) implies that $\mathbf{P}^f$ is both quasi-local and one-sided quasi-local if f is bounded away from 0 and 1. When is $\mathbf{P}^f$ (a.s.) one-sided quasi-local? When there is phase multiplicity, then $\mathbf{P}^f$ is as far as possible from this since then each version of $\mathbf{P}^f[\eta(\mathbf{0}) = 1 \mid \mathsf{Past}(\mathbf{0})]$ is nowhere continuous.

We say that a stationary process ν on $2^{\mathbb{Z}^d}$ is **insertion tolerant** if $\nu[\eta(\mathbf{0}) = 1 \mid \mathscr{Z}] > 0$ ν-a.s. Similarly, ν is **deletion tolerant** if $\nu[\eta(\mathbf{0}) = 0 \mid \mathscr{Z}] > 0$ ν-a.s.

Conjecture 9.9. *Every $\mathbf{P}^f$ is insertion and deletion tolerant, other than the trivial cases $f = \mathbf{0}$ and $f = \mathbf{1}$.*

Theorem 5.16 implies a positive solution to Conjecture 9.9 when $1/[f(\mathbf{1} - f)]$ is integrable. Indeed, we see by Theorem 5.16 that this is the exact criterion for uniform insertion tolerance and uniform deletion tolerance.

Acknowledgements. We thank Chris Hoffman for his help with Proposition 6.7. We thank Eric Bedford and Jeff Geronimo for important references and Jean-Paul Thouvenot for discussions. We are grateful to David Wilson for permission to include Figures 2 and 3 and to Peter Duren and J.M. Landsberg for helpful remarks. Figure 3 was created using the linear algebraic techniques of Wilson (1997) for generating domino tilings; the needed

matrix inversion was accomplished using the formulas of Kenyon (1997). The resulting tiling gives dual spanning trees by the bijection of Temperley (see Kenyon, Propp, and Wilson (2000)). One of the trees is Figure 2.

REFERENCES

ADAMS, S. (1992). Følner independence and the amenable Ising model. *Ergodic Theory Dynam. Systems* **12**, 633–657.

BENJAMINI, I., LYONS, R., PERES, Y., and SCHRAMM, O. (2001). Uniform spanning forests. *Ann. Probab.* **29**, 1–65.

BIERSTONE, E. and MILMAN, P.D. (1988). Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, 5–42.

BORODIN, A., OKOUNKOV, A., and OLSHANSKI, G. (2000). Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13**, 481–515 (electronic).

BRADLEY, R.C. (1986). Basic properties of strong mixing conditions. In Eberlein, E. and Taqqu, M.S., editors, *Dependence in Probability and Statistics*, pages 165–192. Birkhäuser Boston, Boston, MA. A survey of recent results, Papers from the conference held at the Mathematical Research Institute Oberwolfach, Oberwolfach, April 1985.

BROMAN, E. (2003). One-dependent trigonometric determinantal processes are two-block-factors. *Ann. Probab.* To appear.

BUMP, D. and DIACONIS, P. (2002). Toeplitz minors. *J. Combin. Theory Ser. A* **97**, 252–271.

BURTON, R.M. and PEMANTLE, R. (1993). Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.* **21**, 1329–1371.

CONZE, J.P. (1972/73). Entropie d'un groupe abélien de transformations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **25**, 11–30.

DALEY, D.J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York.

DOOB, J.L. (1944). The elementary Gaussian processes. *Ann. Math. Statistics* **15**, 229–282.

FURSTENBERG, H. (1967). Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1**, 1–49.

GEORGII, H.O. (1988). *Gibbs Measures and Phase Transitions*. Walter de Gruyter & Co., Berlin-New York.

GLASNER, E., THOUVENOT, J.P., and WEISS, B. (2000). Entropy theory without a past. *Ergodic Theory Dynam. Systems* **20**, 1355–1370.

GRENANDER, U. and SZEGŐ, G. (1984). *Toeplitz Forms and Their Applications*. Chelsea Publishing Co., New York, second edition.

HEDLUND, G.A. (1969). Endormorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory* **3**, 320–375.

HEICKLEN, D. and LYONS, R. (2003). Change intolerance in spanning forests. *J. Theoret. Probab.* **16**, 47–58.

HELSON, H. and LOWDENSLAGER, D. (1958). Prediction theory and Fourier series in several variables. *Acta Math.* **99**, 165–202.

HOFFMAN, C. (1999). A Markov random field which is K but not Bernoulli. *Israel J. Math.* **112**, 249–269.

DEN HOLLANDER, F. and STEIF, J.E. (1997). On K-automorphisms, Bernoulli shifts and Markov random fields. *Ergodic Theory Dynam. Systems* **17**, 405–415.

DEN HOLLANDER, F. and STEIF, J.E. (2000). On the equivalence of certain ergodic properties for Gibbs states. *Ergodic Theory Dynam. Systems* **20**, 231–239.

JOAG-DEV, K. and PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11**, 286–295.

JOHANSSON, K. (2002). Non-intersecting paths, random tilings and random matrices. *Probab. Theory Related Fields* **123**, 225–280.

KALIKOW, S. (1990). Random Markov processes and uniform martingales. *Israel J. Math.* **71**, 33–54.

KASTELEYN, P. (1961). The statisics of dimers on a lattice I. The number of dimer arrangements on a quadratic lattice. *Physica* **27**, 1209–1225.

KATZNELSON, Y. and WEISS, B. (1972). Commuting measure-preserving transformations. *Israel J. Math.* **12**, 161–173.

KENYON, R. (1997). Local statistics of lattice dimers. *Ann. Inst. H. Poincaré Probab. Statist.* **33**, 591–618.

KENYON, R.W., PROPP, J.G., and WILSON, D.B. (2000). Trees and matchings. *Electron. J. Combin.* **7**, Research Paper 25, 34 pp. (electronic).

KOLMOGOROV, A. (1941a). Stationary sequences in Hilbert space. *Bull. Moscow State Univ.* **2**, 1–40. In Russian.

KOLMOGOROV, A. (1941b). Interpolation and extrapolation of stationary random sequences. *Izv. Akad. Nauk SSSR. Ser. Mat.* **5**, 3–14. In Russian.

DE LEEUW, K., KATZNELSON, Y., and KAHANE, J.P. (1977). Sur les coefficients de Fourier des fonctions continues. *C. R. Acad. Sci. Paris Sér. A-B* **285**, A1001–A1003.

LIGGETT, T.M. (1985). *Interacting Particle Systems.* Springer-Verlag, New York-Berlin.

LYONS, R. (1998). A bird's-eye view of uniform spanning trees and forests. In Aldous, D. and Propp, J., editors, *Microsurveys in Discrete Probability*, volume 41 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 135–162. Amer. Math. Soc., Providence, RI. Papers from the workshop held as part of the Dimacs Special Year on Discrete Probability in Princeton, NJ, June 2–6, 1997.

LYONS, R. (2003a). Asymptotic enumeration of spanning trees. Preprint.

LYONS, R. (2003b). Determinantal probability measures. *Publ. Math. Inst. Hautes Études Sci.* To appear.

LYONS, R. (2003c). Szegő limit theorems. *Geom. Funct. Anal.* **13**, 574–590.

LYONS, R., PAULE, P., and RIESE, A. (2002). A computer proof of a series evaluation in terms of harmonic numbers. *Appl. Algebra Engrg. Comm. Comput.* **13**, 327–333.

LYONS, R. with PERES, Y. (2004). *Probability on Trees and Networks.* Cambridge University Press. In preparation. Current version available at http://mypage.iu.edu/~rdlyons/.

MONTROLL, E. (1964). Lattice statistics. In Beckenbach, E., editor, *Applied Combinatorial Mathematics*, pages 96–143. John Wiley and Sons, Inc., New York-London-Sydney. University of California Engineering and Physical Sciences Extension Series.

NEWMAN, C.M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. In Tong, Y.L., editor, *Inequalities in Statistics and Probability*, pages 127–140. Inst. Math. Statist., Hayward, CA. Proceedings of the symposium held at the University of Nebraska, Lincoln, Neb., October 27–30, 1982.

ORNSTEIN, D.S. (1974). *Ergodic Theory, Randomness, and Dynamical Systems.* Yale University Press, New Haven, Conn. James K. Whittemore Lectures in Mathematics given at Yale University, Yale Mathematical Monographs, No. 5.

ORNSTEIN, D.S. and WEISS, B. (1973). $\mathbf{Z}^d$-actions and the Ising model. Unpublished manuscript.

ORNSTEIN, D.S. and WEISS, B. (1975). Every transformation is bilaterally deterministic. *Israel J. Math.* **21**, 154–158. Conference on Ergodic Theory and Topological Dynamics (Kibbutz Lavi, 1974).

ORNSTEIN, D.S. and WEISS, B. (1987). Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.* **48**, 1–141.

PEMANTLE, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19**, 1559–1574.

PEMANTLE, R. (2000). Towards a theory of negative dependence. *J. Math. Phys.* **41**, 1371–1390. Probabilistic techniques in equilibrium and nonequilibrium statistical physics.

ROZANOV, Y.A. (1967). *Stationary Random Processes.* Holden-Day Inc., San Francisco, Calif.

RUDIN, W. (1987). *Real and Complex Analysis.* McGraw-Hill Book Co., New York, third edition.

RUDOLPH, D.J. (1990). *Fundamentals of Measurable Dynamics.* The Clarendon Press Oxford University Press, New York. Ergodic theory on Lebesgue spaces.

SHAO, Q.M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoret. Probab.* **13**, 343–356.

SHAO, Q.M. and SU, C. (1999). The law of the iterated logarithm for negatively associated random variables. *Stochastic Process. Appl.* **83**, 139–148.

SHIRAI, T. and TAKAHASHI, Y. (2000). Fermion process and Fredholm determinant. In Begehr, H., Gilbert, R., and Kajiwara, J., editors, *Proceedings of the Second ISAAC Congress*, volume 1, pages 15–23. Kluwer Academic Publ. International Society for Analysis, Applications and Computation Volume 7.

SHIRAI, T. and TAKAHASHI, Y. (2002). Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. Preprint.

SHIRAI, T. and TAKAHASHI, Y. (2003). Random point fields associated with certain Fredholm determinants II: fermion shifts and their ergodic and Gibbs properties. *Ann. Probab.* **31**, 1533–1564.

Shirai, T. and Yoo, H.J. (2002). Glauber dynamics for fermion point processes. *Nagoya Math. J.* **168**, 139–166.

Smorodinsky, M. (1992). Finitary isomorphism of m-dependent processes. In Walters, P., editor, *Symbolic Dynamics and Its Applications*, pages 373–376. Amer. Math. Soc., Providence, RI. Proceedings of the AMS Conference held at Yale University, New Haven, Connecticut, July 28–August 2, 1991.

Soshnikov, A. (2000). Determinantal random point fields. *Uspekhi Mat. Nauk* **55**, 107–160.

Steif, J.E. (1991). Space-time Bernoullicity of the lower and upper stationary processes for attractive spin systems. *Ann. Probab.* **19**, 609–635.

Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist* **36**, 423–439.

Teh, H.H. (1961). Construction of orders in Abelian groups. *Proc. Cambridge Philos. Soc.* **57**, 476–482.

Thouvenot, J.P. (1972). Convergence en moyenne de l'information pour l'action de $\mathbf{Z}^2$. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24**, 135–137.

Tracy, C.A. and Widom, H. (2002). On the limit of some Toeplitz-like determinants. *SIAM J. Matrix Anal. Appl.* **23**, 1194–1196.

Trevisan, G. (1953). Classificazione dei semplici ordinamenti di un gruppo libero commutativo con n generatori. *Rend. Sem. Mat. Univ. Padova* **22**, 143–156.

Walters, P. (1982). *An Introduction to Ergodic Theory.* Springer-Verlag, New York.

Wilson, D.B. (1997). Determinant algorithms for random planar structures. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997)*, pages 258–267, New York. ACM. Held in New Orleans, LA, January 5–7, 1997.

Zaĭceva, M.I. (1953). On the set of ordered Abelian groups. *Uspehi Matem. Nauk (N.S.)* **8**, 135–137. In Russian.

Zhang, L.X. (2001). Strassen's law of the iterated logarithm for negatively associated random vectors. *Stochastic Process. Appl.* **95**, 311–328.

Zhang, L.X. and Wen, J. (2001). A weak convergence for negatively associated fields. *Statist. Probab. Lett.* **53**, 259–267.

Department of Mathematics, Indiana University, Bloomington, IN 47405-5701
rdlyons@indiana.edu
http://php.indiana.edu/~rdlyons/

and

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160
rdlyons@math.gatech.edu

Department of Mathematics, Chalmers University of Technology, S-41296 Gothenburg, Sweden
steif@math.chalmers.se

and

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160
steif@math.gatech.edu