The Annals of Probability
1997, Vol. 25, No. 1, 71–95

DOMINATION BY PRODUCT MEASURES[1]

BY T. M. LIGGETT,[1] R. H. SCHONMANN[1] AND A. M. STACEY

University of California, Los Angeles

We consider families of $\{0, 1\}$-valued random variables indexed by the vertices of countable graphs with bounded degree. First we show that if these random variables satisfy the property that conditioned on what happens outside of the neighborhood of each given site, the probability of seeing a 1 at this site is at least a value p which is large enough, then this random field dominates a product measure with positive density. Moreover the density of this dominated product measure can be made arbitrarily close to 1, provided that p is close enough to 1. Next we address the issue of obtaining the critical value of p, defined as the threshold above which the domination by positive-density product measures is assured. For the graphs which have as vertices the integers and edges connecting vertices which are separated by no more than k units, this critical value is shown to be $1 - k^k/(k + 1)^{k+1}$, and a discontinuous transition is shown to occur. Similar critical values of p are found for other classes of probability measures on $\{0, 1\}^{\mathbb{Z}}$. For the class of k-dependent measures the critical value is again $1 - k^k/(k + 1)^{k+1}$, with a discontinuous transition. For the class of two-block factors the critical value is shown to be $1/2$ and a continuous transition is shown to take place in this case. Thus both the critical value and the nature of the transition are different in the two-block factor and 1-dependent cases.

0. Introduction. The investigation reported in this paper started with a question deriving from the following: let $(X_s)_{s \in \mathbb{Z}^d}$ be a collection of random variables which take values 0 and 1 and which satisfy the following condition: for each $s \in \mathbb{Z}^d$, the conditional probability that $X_s = 1$ given any information about the values of the X_u for indices u which are at a (l_∞, say) distance larger than k from s is bounded below by p. Then is it the case that if p is very close to 1, this random field $(X_s)_{s \in \mathbb{Z}^d}$ can be guaranteed to be stochastically above (in the usual sense of the existence of a coupling which puts no mass on one side of the diagonal; see Section 1 for a precise definition) a translation invariant product random field with a density ρ which is also large (in the sense that ρ can be taken arbitrarily close to 1, provided that p is close enough to 1)?

This question is answered in the affirmative in Section 1 of this paper (see Corollary 1.4, which is stated and proven in a more general setting) and we explain why this is of use. For future reference we will denote by $C(d, k, p)$ the class of random fields $(X_s)_{s \in \mathbb{Z}^d}$ which have the property defined in the

Received September 1995; revised April 1996.
[1]Supported by NSF Grants DMS-91-00725 and DMS-94-00644.
AMS 1991 *subject classifications*. Primary 60G60, 60G10; secondary 60K35.
Key words and phrases. Stochastic domination, random fields, product measures, critical points, rescaling.

previous paragraph. In a typical abuse of language, we will sometimes say that a random field is in $C(d, k, p)$ and sometimes say that its law is in this class. A very important subclass of $C(d, k, p)$ is that of k-dependent random fields whose marginals are at least p. We recall that $(X_s)_{s \in \mathbb{Z}^d}$ is said to be a k-dependent random field if, for each pair, $A, B \subset \mathbb{Z}^d$ such that all indices in A are at a distance greater than k from all indices in B, the families of random variables $(X_s)_{s \in A}$ and $(X_s)_{s \in B}$ are independent of each other.

The question raised in the first paragraph of the Introduction comes up frequently in connection with block (rescaling) arguments, which are quite useful in the study of percolation processes, statistical mechanics models and interacting particle systems. A typical setup is as follows. One starts with an underlying random field $(V_i)_{i \in \mathbb{Z}^d}$ in which the random variables V_i may take values in a very general set. The lattice $\mathbb{Z}^d$ is then covered with cubic blocks of the form $\Gamma_s = \{1, \ldots, AN\}^d + Ns$, where $A, N \in \{1, 2, \ldots\}$ are fixed and s runs over $\mathbb{Z}^d$. (We are using the standard definition $S + x = \{y \in \mathbb{Z}^d\colon\ y - x \in S\}$ for $S \subset \mathbb{Z}^d$ and $x \in \mathbb{Z}^d$.) Each one of these blocks is regarded as a site of a "rescaled" lattice, which can also be identified with $\mathbb{Z}^d$. Note that unless $A = 1$, each block overlaps a few neighboring ones, with this number depending on A, but not on N. One declares each block Γ_s to be in one of two possible states, depending on the realization of the underlying random field (V_i) restricted to $i \in \Gamma_s$. The block is said to be "good" in one of these two cases and "bad" otherwise. A random field $(X_s)_{s \in \mathbb{Z}^d}$ is then introduced with X_s being the indicator of the event that the block Γ_s is good. In the applications that we have in mind, this is done in such a way that A is kept fixed, while N can be varied, and the random field $(X_s)_{s \in \mathbb{Z}^d}$ can be shown to be in $C(d, k, p)$ for some finite k which depends only on A and with p as close to 1 as desired, provided N is large enough. The fact that the probability of each block's being good can be made large depends on specific knowledge about the random field (V_i) and the definition of the good event, but if this random field has a product law, it is clear that (X_s) will have a finite range of dependence, which depends on A but not on N. For more general underlying random fields (V_i), as for instance Gibbs measures, one usually does not obtain a rescaled random field (X_s) which has a finite range of dependence, but one can sometimes specify a worst case scenario for what happens with (V_i) when i lies outside of the fixed block Γ_s, and in this fashion establish that $(X_s) \in C(d, k, p)$. When applying this sort of rescaling argument one wants to argue that the bad blocks are rare enough so that certain estimates can be carried out.

For many purposes one can easily obtain the desired estimate directly for random fields in $C(d, k, p)$, without the need for any comparison with product measures [a typical example can be found, for instance, in Section 11a of Durrett (1988), with the relevant computation appearing in Section 5a, in the last paragraph starting on page 86]. On the other hand, this is not always the case, and in a significant number of papers special techniques and constructions were introduced to deal with rescaled random fields which were not themselves product measures. Of special relevance is the recent sequence

of papers: Pisztora (1996), Antal and Pisztora (1996), Penrose and Pisztora (1996), in which ingenious techniques are used. These papers prompted our investigation, and the techniques introduced there were a source of inspiration for our approach in Section 1. Another example of the usefulness of our result can be found by looking at Section 2 in Andjel (1993). There a special argument was used to couple site and bond percolation processes, in such a way that results about bond percolation with large density could be extended to site percolation with large density. Our general result can replace this particular construction: starting with a site percolation process one declares each bond to be occupied if its endpoints are both occupied. The resulting bond percolation process, while not independent, has nevertheless a finite range of dependence and a density which approaches 1 as the density of the underlying site percolation process approaches 1. Therefore we know that this bond percolation process dominates an independent bond percolation process which also has a large density, and this is enough for the purposes in Andjel (1993), as well as other similar applications, in which one wants to import results from bond to site percolation under a large density condition. (Of course, a similar argument allows one to go in the opposite direction and extend to bond percolation results proven for site percolation with large density.) Finally, in Schonmann (1994), the lack of an answer at the time to the question that we are discussing caused the author (and hence also the readers of the paper) a great deal of extra work in proving Theorem 7. The solution found there involved checking that a long proof of a general shape theorem, provided in Section 11c of Durrett (1988) can be adapted to cases in which a finite-range-of-dependence condition replaces independence. The comparison provided here eliminates this need altogether. Since rescaling is a standard technique, we believe that the result we are discussing will continue to be a relevant tool.

Having obtained domination from below by product measures when the p defined above is large, it became natural, for its own sake, to ask whether for smaller values of p one can still assure that every random field $(X_s)_{s \in \mathbb{Z}^d} \in C(d, k, p)$ is dominated from below by a product measure with some positive density. Interestingly enough, it is not hard to provide an example which shows that even for translation invariant 1-dependent Bernoulli random fields on $\mathbb{Z}$, if $p = 1/2$ there may be no product measure with positive density below it: let $(U_s)_{s \in \mathbb{Z}}$ be a family of i.i.d. random variables which are uniform on $[0, 1]$ and set

$$X_s = 1_{\{U_s < U_{s+1}\}}. \tag{0.0}$$

Then $P(X_1 = 1) = 1/2$, and more generally,

$$P(X_s = 1 \ \forall\ 1 \le s \le n) = P(U_1 \le U_2 \le \cdots \le U_{n+1}) = \frac{1}{(n+1)!}.$$

Since this quantity decays faster than any exponential, $(X_s)_{s \in \mathbb{Z}}$ cannot dominate a product random field with positive density.

This suggests the problem of finding the critical value $p_c = p_c(d, k)$, defined as the infimum of the values of p which are large enough to guarantee the existence of a product measure with positive density below the law of every random field $(X_s)_{s \in \mathbb{Z}^d} \in C(d, k, p)$. It turns out that in the one-dimensional case we were able to compute the value of this critical point as being $1 - k^k/(k+1)^{(k+1)}$; for instance, when $k = 1$, $p_c = 3/4$. Moreover, if we modify the question, by replacing the condition on the random field $(X_s)_{s \in \mathbb{Z}^d}$ with the stronger condition that it be a k-dependent random field, this value for the critical point in the $d = 1$ case remains unchanged.

The proof that the value $1 - k^k/(k+1)^{(k+1)}$ is an upper bound for p_c results from the estimates in Section 1. For the proof that this value is also a lower bound for p_c we construct in Section 2 examples of random fields in $C(1, k, p)$ which are not dominated from below by any product measure with positive density. This is done for p taking an increasing sequence of values p_n, with $p_n \to 1 - k^k/(k+1)^{(k+1)}$ as $n \to \infty$. For each n we can find a probability distribution on $\{0, 1\}^{\{1, \cdots, n\}}$ which gives no mass to the configuration with all variables equal to 1, while the properties of being k-dependent and having density bounded below by p_n are satisfied. Such a distribution can then trivially be extended to a distribution on $\{0, 1\}^{\mathbb{Z}}$ in the form of a product random field of density p_n outside of this interval. The impossibility that any product measure with positive density dominates this one from below is clear.

Of course, one then naturally asks whether the absence of translation invariance of these examples is of fundamental relevance. In other words, would the value of p_c be modified if we restricted ourselves to translation invariant random fields $(X_s)_{s \in \mathbb{Z}^d}$? We only have some partial answers to this question. First of all, the answer may depend on whether in addition to translation invariance we require (X_s) to be in $C(1, k, p)$ or to be k-dependent and have density p. In the former case, the answer is that p_c is unchanged, and this can be easily seen by using the result described in the previous paragraph. For each n, one constructs a measure which is k-dependent and has period n, by simply repeating indefinitely the basic building block of length n (alluded to in the previous paragraph) in an independent fashion. By then mixing n shifts of this measure, a translation invariant random field is produced, which is no longer k-dependent, but still belongs to $C(1, k, p_n)$ and again clearly cannot be dominated from below by any product measure with positive density.

We have not succeeded in finding the value of p_c when (X_s) is assumed to be k-dependent and translation invariant. Nevertheless for a natural subclass of this with $k = 1$, we did find that $p_c = 1/2$, in contrast with the value $3/4$ mentioned earlier. This subclass is that of two-block factors, which are defined as follows. A measure μ is said to be a two-block factor if there is a $\{0, 1\}$-valued measurable function $f(x, y)$ on $[0, 1]^2$ so that the process

$$X_s = f(U_s, U_{s+1}) \tag{0.1}$$

has distribution μ when the U_s's are i.i.d. random variables which are uniform on $[0, 1]$. Two-block factors are natural, since most 1-dependent translation invariant measures on $\{0, 1\}^{\mathbb{Z}}$ which arise in applications are actually of this kind. In fact, for many years it was not known whether there were any others. The first proof of the existence of stationary, 1-dependent random processes which are not two-block factors was given by Aaronson, Gilat, Keane and de Valk (1989).

The example (0.0) shows the easy half of our claim that for two-block factors $p_c = 1/2$, and the other half is shown in Section 3. It is interesting to note that not only the value of p_c is different from that in the other one-dimensional cases discussed earlier, but also the nature of the transition is different. In the earlier cases, it turns out that for

$$p = p_c = 1 - k^k/(k+1)^{(k+1)}$$

we can assure domination of (X_s) from below by a product measure with positive density ("discontinuous transition"). In contrast, in the case of (translation invariant) two-block factors, we have a "continuous transition," in that for p above but close to $1/2$, (X_s) can only be assured to be dominated from below by product measures with small densities (more precisely, the supremum of such densities is a function of p which vanishes as $p \to 1/2 +$).

There is one other case where we can say what the value of the critical point is, and which may be worth mentioning. Here the class is those probability measures on $\{0, 1\}^{\mathbb{Z}}$ which are 1-dependent, have density at least p at each site, and are positively correlated in the sense of Liggett (1985), Chapter II, Definition 2.11. Measures which arise from a rescaling argument do often have this positive correlations property. A slight generalization of Proposition 3.4 can be used to show that for every $p > 0$ all of these measures are dominated from below by a product measure with positive density. In other words, the critical value of p is 0 in this case.

We finish this introduction by stating a theorem which summarizes the main results discussed above. (As pointed out, some of these results will be presented in greater generality in the paper.)

THEOREM 0.0.

(i) *For each d and k, when p is large enough, the random fields in $C(d, k, p)$ are dominated from below by the product random field with density ρ, where ρ is a positive constant depending on d, k and p. One can make the density of these product random fields become arbitrarily close to* 1 *by taking p large enough.*

In the case $d = 1$, the critical value of p beyond which the result stated in (i) *holds, is*

$$p_c(d, k) = 1 - \frac{k^k}{(k+1)^{(k+1)}}.$$

The transition is discontinuous, in the sense that all the random fields in $C(1, k, p_c)$ *are dominated from below by product random fields with positive density* $\rho = k/(k+1)^2$.

(ii) *All these results remain true if the class of random fields* $C(d, k, p)$ *is replaced by the class of k-dependent fields.*

(iii) *All two-block factors with density p are dominated from below by product measures with positive density if and only if p is above the critical value*

$$p_c^F = \tfrac{1}{2}.$$

In this case, the transition is continuous in the sense that for arbitrary $\varepsilon > 0$ *one can find a two-block factor with density* $p > 1/2$ *which is not dominated from below by the product random field with density* ε.

In parts (i) and (ii) of this theorem, we have identified two types of critical behavior in a class of measures. It is natural to ask whether for other classes, one might have a different behavior, or perhaps rule out other possibilities. In some cases we are indeed able to reduce the possibilities. For instance, for the class $C(d, k, p)$ in an arbitrary dimension d and with an arbitrary value of k, either all the random fields in $C(d, k, p_c)$ are dominated from below by a common product random field with positive density, as in part (i) of Theorem 0.0, or else, for arbitrary $\varepsilon > 0$, one can find a random field in $\bigcup_{p > p_c} C(d, k, p)$ which is not dominated from below by the product random field with density ε, as in part (iii) of Theorem 0.0. To prove this, it is enough to observe that if $(X_s)_{s \in \mathbb{Z}^d} \in C(d, k, p)$ and the random field $(X'_s)_{s \in \mathbb{Z}^d}$ is defined by setting $X'_s = X_s \vee Y_s$, where $(Y_s)_{s \in \mathbb{Z}^d}$ is a product random field with density δ, which is independent of (X_s), then $(X'_s)_{s \in \mathbb{Z}^d} \in C(d, k, p + \delta(1-p))$. An analogous remark can be made for the class $T(p)$, discussed in Section 3.

On the other hand, one can artificially, but easily, introduce classes of measures indexed by a parameter $p \in [0, 1]$, for which the transition could be different from the two alternatives above. Given a function $\theta\colon [0, 1] \to [0, 1]$ let $A_\theta(p)$ be the class of random fields which are dominated from below by the product random field with density $\theta(p)$. By choosing $\theta(p) = 0$ for $p \le p_c$ and $\theta(p) > a > 0$ for $p > p_c$, for some $0 < p_c < 1$, we would obviously violate both alternatives above.

1. Domination by product measures under weak independence conditions. Suppose that we have a countable (possibly finite) set of sites, S, and a state space $\Omega = \{0, 1\}^S$, with the product topology and the corresponding Borel σ-algebra, B. We begin with the definition of stochastic domination. Given $\omega_1, \omega_2 \in \Omega$, we define the usual partial order by saying that $\omega_1 \le \omega_2$ if

$$\forall\, s \in S \quad \omega_1(s) \le \omega_2(s).$$

Given a function $f\colon \Omega \to \mathbb{R}$, we say that f is *increasing* if

$$\omega_1 \le \omega_2 \quad \Rightarrow \quad f(\omega_1) \le f(\omega_2).$$

Now, given two Borel probability measures on Ω, μ and ν, we shall say that μ *stochastically dominates* ν—and write $\mu \succcurlyeq \nu$— if, for any continuous increasing function f,

$$\int f\,d\mu \ge \int f\,d\nu.$$

This is a very strong condition which amounts to saying that in every possible way, μ puts more mass on bigger elements of Ω than ν does. See, for example, Liggett (1985), Chapter II, Section 2, for further details. One can, of course, define stochastic domination for more general partially ordered state spaces, but we will only be interested in the case we have just described.

It is easy to show by a simple limiting argument that, in order to prove $\mu \succcurlyeq \nu$, it is sufficient to verify that this inequality holds for the finite-dimensional distributions of μ and ν. The following is also well known [see Liggett (1985), Chapter II, Theorem 2.4].

LEMMA 1.0. *Let S be a countable set. Then $\mu \succcurlyeq \nu$ if and only if there are two random fields, $(X_s)_{s\in S}$ and $(Y_s)_{s\in S}$, defined on the same probability space such that, for each $s \in S$, $X_s \ge Y_s$ almost surely and such that the joint laws of $(X_s)_{s\in S}$ and of $(Y_s)_{s\in S}$ are μ and ν, respectively.*

Given $0 \le p \le 1$ and a set of sites S, we let π_p^S (or simply π_p when there is no danger of confusion) be the measure in which, for each site $s \in S$, $\pi_p(\{\omega\colon \omega(s) = 1\}) = p$ and each of the sites behaves independently, so π_p is a product measure.

Where convenient, we will work with random variables $(X_s)_{s\in S}$ whose joint law is μ, rather than directly with μ itself. Suppose we know that $\mathbb{P}(X_s = 1) \ge p$ for all s. If the family (X_s) is independent then it is easy to conclude that μ stochastically dominates π_p (it follows from Lemma 1.1, which is a more general result), but if we have no independence information then we cannot draw any such conclusion (except in the trivial cases $p = 0$ and $p = 1$). In this paper we examine some conditions, weaker than full independence, under which we can conclude than μ dominates some nontrivial product measure.

Let $G = (S, E)$ be a graph with vertex set S and edge set E. We say that two subsets of S, S_1 and S_2, are *mutually independent* if $S_1 \cap S_2 = \varnothing$ and there is no edge joining any vertex in S_1 to a vertex in S_2.

A collection of random variables $(X_s)_{s\in S}$, is said to have *dependence graph* G if, whenever S_1 and S_2 are mutually independent, the σ-algebras $\sigma(X_s\colon s \in S_1)$ and $\sigma(X_s\colon s \in S_2)$ are independent. We shall be concerned with the case in which the X_s are $\{0, 1\}$-valued with joint law μ; in this case we shall say that μ has dependence graph G whenever the family $(X_s)_{s\in S}$ does. Given a graph G, we shall use $G(p)$ to denote the class of Borel probability

measures μ on $\{0,1\}^S$ with dependence graph G and which satisfy $\mu(\{\eta: \eta(s) = 1\}) \geq p$ for all $s \in S$.

If G is the usual graph with vertex set $\mathbb{Z}$, in which vertex n is joined by edges to $n-1$ and $n+1$, then $G(p)$ is just the class of 1-dependent measures on $\{0,1\}^{\mathbb{Z}}$ in which the probability of a 1 at any vertex is at least p. By adding edges we can obtain the class of k-dependent measures for any $k \in \mathbb{N}$, and this easily generalizes to higher dimensions with a given neighborhood of dependence.

We define

$$\sigma^G(p) = \sup\{\rho: \mu \succcurlyeq \pi_\rho \ \forall \mu \in G(p)\}$$

and a critical value

$$p_c^G = \inf\{p: \sigma^G(p) > 0\}.$$

It is not hard to see that (provided G has at least one edge) $p_c^G \geq 1/2$. Indeed, let s_1 and s_2 be two of the vertices of G between which there is an edge. Let $(X_s)_{s \in S}$ be a random field such that with probability $1/2$ the event $\{X_{s_1} = 0,\ X_{s_2} = 1\}$ happens, with probability $1/2$ the event $\{X_{s_1} = 1,\ X_{s_2} = 0\}$ happens, and the random variables $(X_s)_{s \in S \setminus \{s_1, s_2\}}$ have a product law with density $1/2$ and are independent of $(X_s)_{s \in \{s_1, s_2\}}$. Clearly this random field is in $G(1/2)$, but does not dominate any product measure with positive density, since the probability of the event $\{X_{s_1} = X_{s_2} = 1\}$ is 0. This example is a prototype of the construction that will be developed in Section 2, in order to obtain much finer lower bounds on the critical point p_c^G for some special graphs.

It does not seem a priori obvious that $p_c^G < 1$, for any graphs with infinite components. We shall prove in this section that $p_c^G < 1$ if G has bounded degree and obtain upper bounds on p_c^G which, as we shall see in Section 2, are tight in certain cases. It will also be a consequence of our results that $\lim_{p \to 1} \sigma^G(p) = 1$ if G has bounded degree.

In order to obtain results in a somewhat more general context, we introduce a broader class of measures, $G^W(p)$ (W for "weak") in which the dependence conditions are relaxed but in which we preserve properties we will need in this section. We say that $\mu \in G^W(p)$ if μ is a Borel probability measure on $\{0,1\}^S$ and

$$\mathbb{E}(X_s | \sigma(\{X_t: st \notin E\})) \geq p \quad \text{a.s.} \tag{1.0}$$

for any vertex $s \in S$. Here again, $(X_s)_{s \in S}$ have joint law μ. As usual, an edge joining vertices s and t is abbreviated st. Statement (1.0) asserts that if one conditions on information about the states of the vertices other than those which are neighbors of s, then the probability that s is in state 1 is still at least p.

This condition can be reformulated in terms of elementary conditional probabilities as follows. For any $n \in \mathbb{N}$ and any choice of distinct $s, s_1, \ldots, s_n$

$\in S$ with $ss_i \notin E$ for each i, and for any choice of $\varepsilon_1, \dots \varepsilon_n \in \{0,1\}$,

$$\mathbb{P}\big(X_s = 1 | X_{s_1} = \varepsilon_1, \dots, X_{s_n} = \varepsilon_n\big) \geq p, \tag{1.1}$$

whenever the event on which we condition has positive probability.

The following notation will be used:

$$\sigma_W^G(p) = \sup\big\{\rho : \mu \succcurlyeq \pi_\rho \ \forall\, \mu \in G^W(p)\big\}$$

and

$$p_{W,c}^G = \inf\big\{p : \sigma_W^G(p) > 0\big\}.$$

Note that $\sigma_W^G(p) \leq \sigma^G(p)$ and $p_{W,c}^G \geq p_c^G$. We will show that if G has bounded degree, then $\lim_{p \to 1} \sigma_W^G(p) = 1$, so that in particular $p_{W,c}^G < 1$.

Before we actually prove anything of substance, we recall a rather simple lemma which will enable later proofs to go more smoothly. This lemma can be proved by a reasonably standard use of the coupling technique. However, the details, which involve a step-by-step construction, are a little involved, so we omit the proof. Results of this kind are, however, well known. See, for example, Russo [(1982), Lemma 1] for a proof of a more general result. Note that the condition on μ is a weaker version of the condition that $\mu \in G^W(p)$ where $G = (S, \varnothing)$.

LEMMA 1.1. *Suppose that $(X_s)_{s \in S}$ is a family of random variables, indexed by a countable set S, with joint law, μ. Suppose S can be totally ordered in such a way that, given any finite subset of S, $s_1 < s_2 < \cdots < s_j < s_{j+1}$, and any choice of $\varepsilon_1, \dots, \varepsilon_j \in \{0,1\}$, then, whenever $\mathbb{P}(X_{s_1} = \varepsilon_1, \dots, X_{s_j} = \varepsilon_j) > 0$,*

$$\mathbb{P}\big(X_{s_{j+1}} = 1 | X_{s_1} = \varepsilon_1, \dots, X_{s_j} = \varepsilon_j\big) \geq \rho.$$

Then $\mu \succcurlyeq \pi_\rho^S$.

We are now in a position to present the main results of this section. We start with a proposition which leads up to the main general theorem of this section, Theorem 1.3.

PROPOSITION 1.2. *Let $G = (S, E)$ be a graph with maximum degree at most $\Delta \geq 1$. Let $0 < p < 1$, $q = 1 - p$ and suppose $\mu \in G^W(p)$. Suppose there exist α, $r \in (0,1)$ with*

$$(1 - \alpha)(1 - r)^{\Delta - 1} \geq q, \tag{1.2}$$

$$(1 - \alpha)\alpha^{\Delta - 1} \geq q. \tag{1.3}$$

Let $(X_s)_{s \in S}$ have joint law μ, let (Y_s) be a family independent of (X_s) and with joint law π_r and, for each $s \in S$, let $Z_s = X_s Y_s$. Let $s_0, s_1, \dots s_j \in S$ be distinct, and suppose that s_0 is adjacent in G to at most $\Delta - 1$ of the vertices $s_1, \dots, s_j$. Then, for any choice of $\varepsilon_1, \dots, \varepsilon_j \in \{0,1\}$, we have

$$\mathbb{P}\big(X_{s_0} = 1 | Z_{s_1} = \varepsilon_1, \dots, Z_{s_j} = \varepsilon_j\big) \geq \alpha \tag{1.4}$$

or equivalently,

$$\mathbb{P}\big(Z_{s_0} = 1 | Z_{s_1} = \varepsilon_1, \ldots, Z_{s_j} = \varepsilon_j\big) \geq \alpha r. \tag{1.5}$$

PROOF. Note that the equivalence of the two statements is clear from the independence of the family (Y_s) and the fact $r > 0$. Take G, μ, p, α and r satisfying the conditions of the theorem. For ease of notation let us write X_i for X_{s_i} and likewise for Z_i. We prove the claim by induction on j.

If $j = 0$ then the l.h.s. of (1.4) is simply the probability that $X_0 = 1$. We know, from the fact that $\mu \in G^W(p)$, that this is at least p. Since (1.3) implies that $p \geq \alpha$, (1.4) holds in this case.

Now suppose that we have $j \geq 1$ and we know that (1.4) holds for smaller values. Fix a choice of $s_0, s_1, \ldots, s_j$ and $\varepsilon_1 \cdots \varepsilon_j$. Then we split $s_1, \ldots, s_j$ into three classes:

$$N_0 = \{s_i : 1 \leq i \leq j,\ s_0 s_i \in E \text{ and } \varepsilon_i = 0\},$$
$$N_1 = \{s_i : 1 \leq i \leq j,\ s_0 s_i \in E \text{ and } \varepsilon_i = 1\},$$
$$M = \{s_i : 1 \leq i \leq j \text{ and } s_0 s_i \notin E\}.$$

The condition that $Z_s = 1$ is the same as the condition that both $X_s = 1$ and $Y_s = 1$, so we can rewrite the probability in (1.4) as

$$\mathbb{P}\big(X_0 = 1 | (\forall\, s \in N_0,\ Z_s = 0), (\forall\, s \in N_1,\ X_s = 1 \text{ and } Y_s = 1), (\forall s_i \in M,\ Z_i = \varepsilon_i)\big).$$

By the independence of the (Y_s) family, we can simplify a little to give

$$\mathbb{P}\big(X_0 = 1 | (\forall\, s \in N_0,\ Z_s = 0), (\forall\, s \in N_1,\ X_s = 1), (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big).$$

We will actually work with the conditional probability that $X_0 = 0$ which can be written out as

$$\frac{\mathbb{P}\big(X_0 = 0, (\forall\, s \in N_0,\ Z_s = 0), (\forall\, s \in N_1,\ X_s = 1), (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}{\mathbb{P}\big((\forall\, s \in N_0,\ Z_s = 0), (\forall\, s \in N_1,\ X_s = 1), (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}.$$

By enlarging the event in the numerator and shrinking the event in the denominator, we see that this is at most

$$\frac{\mathbb{P}\big(X_0 = 0, (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}{\mathbb{P}\big((\forall\, s \in N_0,\ Y_s = 0), (\forall\, s \in N_1,\ X_s = 1), (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}.$$

By dividing both numerator and denominator by $\mathbb{P}(\forall\, s_i \in M,\ Z_i = \varepsilon_i)$, this becomes

$$\frac{\mathbb{P}\big(X_0 = 0 | (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}{\mathbb{P}\big((\forall\, s \in N_0,\ Y_s = 0), (\forall\, s \in N_1,\ X_s = 1) | (\forall\, s_i \in M,\ Z_i = \varepsilon_i)\big)}. \tag{1.6}$$

Since M contains no neighbors of s_0 and $\mu \in G^W(p)$, the numerator is at most q. The denominator is equal to

$$(1 - r)^{|N_0|}\mathbb{P}(\forall\, s \in N_1,\ X_s = 1 | \forall\, s_i \in M,\ Z_i = \varepsilon_i).$$

We shall use the induction hypothesis to bound this conditional probability from below. By relabeling if necessary, let $N_1 = \{s_1, \ldots, s_k\}$. Then

$$\begin{aligned} &\mathbb{P}\big(\forall\, s \in N_1,\ X_s = 1 | \forall\, s_i \in M,\ Z_i = \varepsilon_i\big) \\ &\quad = \mathbb{P}\big(X_1 = 1 | \forall\, s_i \in M,\ Z_i = \varepsilon_i\big) \\ &\qquad \cdots \mathbb{P}\big(X_k = 1 | (\forall\, s_i \in M,\ Z_i = \varepsilon_i),\ X_1 = \cdots = X_{k-1} = 1\big). \end{aligned} \tag{1.7}$$

Now, since G has maximum degree Δ, each of the neighbors of s_0 (in particular, each of the elements of N_1) has at most $\Delta - 1$ neighbors in $M \cup N_1$, since this set does not include s_0. Hence, applying the induction hypothesis, each of the $|N_1| (= k)$ terms in the product on the r.h.s. of (1.7) is at least α. (In this step, we have used the fact that the Y_s are independent of each other and of the X's.) Hence the denominator in (1.6) is at least $(1-r)^{|N_0|}\alpha^{|N_1|}$. So (1.6), which is an upper bound for the conditional probability that $X_0 = 0$, is bounded above by $q/((1-r)^{|N_0|}\alpha^{|N_1|})$. Since $|N_0| + |N_1|$ is at most $\Delta - 1$, the inequalities (1.2) and (1.3) imply that this upper bound is at most $1 - \alpha$ and so the conditional probability that $X_0 = 1$ is indeed at least α which completes the proof of the proposition.

We now combine Proposition 1.2 and Lemma 1.1 to give some conditions which ensure that a measure in $G^W(p)$ dominates a nontrivial product measure.

THEOREM 1.3. *Let $G = (S, E)$ be a graph with a countable vertex set in which every vertex has degree at most $\Delta \geq 1$, and in which every finite connected component of G contains a vertex of degree strictly less than Δ. Let $p, \alpha, r \in [0, 1]$, $q = 1 - p$, and suppose inequalities* (1.2) *and* (1.3) *are satisfied. Let $\mu \in G^W(p)$. Then $\mu \succcurlyeq \pi_{\alpha r}$. In particular, if $q \leq (\Delta - 1)^{\Delta - 1}/\Delta^{\Delta}$, then $\mu \succcurlyeq \pi_\rho$ where*

$$\rho = \left(1 - \frac{q^{1/\Delta}}{(\Delta - 1)^{(\Delta-1)/\Delta}}\right)\big(1 - (q(\Delta - 1))^{1/\Delta}\big).$$

PROOF. We may assume that p, α and r are strictly between 0 and 1, since in the other cases the theorem is trivial. The theorem is also rather easy in the case $\Delta = 1$ (when, in fact, G cannot have any edges), although the argument below does also work for this case; note that we interpret 0^0 as 1.

Now let (X_s), (Y_s) and (Z_s) be as in Proposition 1.2, and let the joint law of (Z_s) be ν. It is clear from Lemma 1.0 that $\mu \succcurlyeq \nu$. We will show that ν satisfies the conditions of Lemma 1.1, with $\rho = \alpha r$. This will imply (due to Lemma 1.1) that $\nu \succcurlyeq \pi_{\alpha r}$, which will complete the proof of the first part of the theorem.

Let F be any finite subset of S with, say, k vertices. The condition on G ensures that every finite subgraph of G contains a vertex of degree at most $\Delta - 1$. We label the vertices of F recursively as follows: having chosen vertices $u_1, \ldots, u_i$ (for some $0 \leq i < k$), pick u_{i+1} to be one of the vertices of degree at most $\Delta - 1$ in the induced subgraph $G[F \setminus \{u_1 \ldots u_i\}]$. Having

labeled all the vertices of F, we order them in reverse, $u_1 \geq u_2 \geq \cdots \geq u_k$. This ordering is designed to ensure that each vertex of F is adjacent to at most $\Delta - 1$ of its predecessors.

Now, with this ordering on F, the random variables $(Z_s)_{s \in F}$ satisfy the condition of Lemma 1.1 with $\rho = \alpha r$; this follows immediately by applying (1.5) of Proposition 1.2. So Lemma 1.1 implies that, if ν^F is the measure induced by ν on $\{0,1\}^F$, then $\nu^F \succcurlyeq \pi_{\alpha r}^F$. This is true for any finite subset F of S, and hence $\nu \succcurlyeq \pi_{\alpha r}^S$.

To establish the final part of the conclusion in the case when $q \leq (\Delta - 1)^{\Delta - 1}/\Delta^{\Delta}$, let

$$\alpha = \left(1 - \frac{q^{1/\Delta}}{(\Delta - 1)^{(\Delta - 1)/\Delta}}\right) \quad \text{and} \quad r = \left(1 - \left(q(\Delta - 1)\right)^{1/\Delta}\right).$$

Then $(1 - \alpha)(1 - r)^{\Delta - 1} = q$, so (1.2) holds. The condition that $q \leq (\Delta - 1)^{\Delta - 1}/\Delta^{\Delta}$ gives

$$q^{1/\Delta} \leq \frac{(\Delta - 1)^{(\Delta - 1)/\Delta}}{\Delta},$$

which implies by a simple substitution that

$$\alpha \geq 1 - \frac{1}{\Delta} \quad \text{and} \quad r \geq \frac{1}{\Delta}.$$

These inequalities demonstrate that $\alpha, r \in [0,1]$; they also show that $\alpha \geq 1 - r$ and hence (1.3) holds too. Consequently, $\mu \succcurlyeq \pi_{\alpha r}$ with this choice of α and r, which is precisely the final statement of the theorem. □

Let us comment that the condition of Lemma 1.1, although sufficient, is certainly not necessary for μ to dominate π_{ρ}^S. Indeed, the family (X_s), used above, need not satisfy this condition, although, when "weakened," we obtain a family (Z_s) which does satisfy it. It is, however, true that the family (X_s) satisfies this condition in the case $\varepsilon_1 = \cdots = \varepsilon_n = 1$; in other words, given $s_1 < \cdots < s_j < s_{j+1}$,

$$\mathbb{P}\left(X_{s_{j+1}} = 1 | X_{s_1} = 1, \ldots, X_{s_j} = 1\right) \geq \rho.$$

This follows from Proposition 1.2, but it is more natural to regard this fact as a form of the local lemma of Erdös and Lovász (1975), and Proposition 1.2 can be regarded as an extension of that result. The use of induction in our proof of Proposition 1.2 is very similar to its use in the proof of the local lemma. Details of this result can also be found in Bollobás [(1985), pages 20–22].

Theorem 1.3 has the following corollary, which only requires a couple of simple checks about the behavior of ρ as a function of q.

COROLLARY 1.4. *Let $G = (S, E)$ be a graph with a countable vertex set in which every vertex has degree at most $\Delta \geq 1$, and in which every finite*

connected component of G contains a vertex of degree strictly less than Δ. Then

$$\lim_{p\to 1}\sigma_W^G(p)=1 \quad \textit{and} \quad p_{W,c}^G \le 1-\frac{(\Delta-1)^{\Delta-1}}{\Delta^\Delta}.$$

In particular,

$$\lim_{p\to 1}\sigma^G(p)=1 \quad \textit{and} \quad p_c^G \le 1-\frac{(\Delta-1)^{\Delta-1}}{\Delta^\Delta}.$$

The best possible choices of α and r are roots of certain polynomials which are not quite the same as the values used in Corollary 1.4. However, it is easy to see that we cannot improve on Corollary 1.4 simply by making a slightly better choice. For if $q > (\Delta-1)^{\Delta-1}/\Delta^\Delta$, then there is no value of α which satisfies (1.3).

However, in one important case, we can improve on Theorem 1.3 and Corollary 1.4. Let $\mathbb{Z}_k$ be the graph with vertex set $\mathbb{Z}$ in which two vertices, m and n, are joined by an edge if $|n-m|\le k$. So $\mathbb{Z}_k(p)$ consists of k-dependent measures on subsets of $\mathbb{Z}$ of density at least p, while $\mathbb{Z}_k^W(p)$ is what we called $C(1,k,p)$ in the Introduction. The following theorem improves Theorem 1.3 in the case $k\ge 2$ (it gives the same result if $k=0$ or $k=1$).

THEOREM 1.5. *Let $p,\alpha,r\in[0,1]$, let $k\ge 0$ and let $q=1-p$. Suppose that*

$$(1-\alpha)(1-r)^k \ge q \tag{1.8}$$

and

$$(1-\alpha)\alpha^k \ge q. \tag{1.9}$$

If $\mu\in\mathbb{Z}_k^W(p)$, then $\mu\succcurlyeq\pi_{\alpha r}$. In particular, if $q\le k^k/(k+1)^{k+1}$ then $\mu\succcurlyeq\pi_\rho$ where

$$\rho=\left(1-\frac{q^{1/(k+1)}}{k^{k/(k+1)}}\right)\left(1-(qk)^{1/(k+1)}\right).$$

OUTLINE OF PROOF. Theorem 1.5 is proved in a very similar way to Proposition 1.2 and Theorem 1.3. We set up random variables $(Z_s)_{s\in\mathbb{Z}}$ in exactly the same way as in Proposition 1.2. Instead of (1.4), we use the following as our induction hypothesis. Let $s_1<\cdots<s_{j+1}\in\mathbb{Z}$. Then for any choice of $\varepsilon_1,\dots,\varepsilon_j$,

$$\mathbb{P}\big(X_{s_{j+1}}=1 \,|\, Z_{s_1}=\varepsilon_1,\dots,Z_{s_j}=\varepsilon_j\big)\ge\alpha. \tag{1.10}$$

Note that this is a weaker hypothesis than (1.4) because it is one-sided; we can only condition on information about sites to the left of the site of interest, s_{j+1}.

This weaker hypothesis can be proved by induction on j in almost exactly the same way as in the proof of Proposition 1.2. The only place where a little extra care is needed is when we relabel the vertices of N_1, $s_1,\dots,s_k$. In the

proof of Proposition 1.2 it did not matter how they were ordered. Here, however, they must be arranged so that $s_1 < \cdots < s_k$. This ordering, together with the fact that all the sites in M (nonneighbors of s_{j+1}) are less than all the sites in N_1 (which are neighbors of s_{j+1}), ensures that each of the k-conditional probabilities appearing on the right-hand side of (1.7) can be bounded using the induction hypothesis.

Having used induction to prove (1.10), one applies Lemma 1.1 (with, simply, the usual ordering on the elements of $\mathbb{Z}$) to establish the equivalent of the first part of Theorem 1.3. The remainder follows precisely as in Theorem 1.3. □

We observe that the hypothesis $\mu \in \mathbb{Z}_k^W(p)$ in Theorem 1.5 can be weakened, by replacing (1.0) with a one-sided conditional expectation, much as the induction hypothesis in Theorem 1.5 is weaker than that in Proposition 1.2.

We can use Theorem 1.5 to improve on Corollary 1.4 in the case of $\mathbb{Z}_k$.

COROLLARY 1.6. *For each value of k,*

$$p_{W,c}^{\mathbb{Z}_k} \leq 1 - \frac{k^k}{(k+1)^{k+1}}.$$

In particular, the critical value for k-dependent measures on subsets of $\mathbb{Z}$ satisfies also the same bound,

$$p_c^{\mathbb{Z}_k} \leq 1 - \frac{k^k}{(k+1)^{k+1}}.$$

As we shall see in the next section, these last results are best possible.

2. *k*-dependent measures which do not dominate product measures. In this section we show how to construct measures, on a state space $\Omega = \{0,1\}^S$, which satisfy certain independence conditions, have $\mathbb{P}(\{\eta\colon \eta(s) = 1\})$ fairly close to 1 which yet do not dominate any nontrivial product measure. These examples provide a partial converse to the results of the previous section, and they show that some of the inequalities given in that section are best possible.

The principal ideas of this section are mostly due to Shearer (1985), where a similar question is addressed with a somewhat different motivation.

We shall start by considering a finite graph, and then deal with the extension to the infinite case. Given a finite graph $G = (S, E)$, and $0 < p < 1$, we will try to define a probability measure on $\Omega = \{0,1\}^S$ in such a way that G is a dependence graph for the events $(\eta(s) = 1)_{s \in S}$, so that each of these events has probability p, and such that the probability that all of these events occur simultaneously is zero; the last property, needless to say, would ensure that the measure we have defined does not dominate π_ρ for any positive value of ρ.

As in Shearer (1985), we shall define a signed measure as follows [for a proof that this approach is best possible, a fact we shall not need, see Shearer (1985)]. Given a set $Q \subseteq S$, then, as usual, we say that Q is an *independent* set if no edge joins two vertices of Q. If Q is independent, we wish to try to define the probability that $(\forall\, s \in Q,\, \eta(s) = 0)$ to be $q^{|Q|}$ (where $q = 1 - p$); if Q is not independent, we proceed by defining this probability to be zero. All the remaining probabilities can then be obtained by the use of the inclusion–exclusion principle. These probabilities may turn out to be negative, although we will be able to get around that difficulty. In any case, they will have the independence properties we require. More formally, given two disjoint sets, $Q, R \subseteq S$, we define

$$P_p(Q, R) = \sum_{\substack{T:\, Q \subseteq T \subseteq Q \cup R \\ T \text{ independent}}} (-1)^{|T|-|Q|} q^{|T|}. \tag{2.0}$$

($T = \varnothing$ is included in this sum if $Q = \varnothing$.) Where there is no danger of confusion, we shall write $P(Q, R)$ for $P_p(Q, R)$. For any set $Q \subseteq S$, the quantity $P(Q, S \setminus Q)$ is the value that the inclusion–exclusion formula and the other comments of the previous paragraph lead us to assign to

$$\mathbb{P}\big(\{\eta: (\forall\, s \in Q,\, \eta(s) = 0) \text{ and } (\forall\, s \in S \setminus Q,\, \eta(s) = 1)\}\big).$$

The way in which these quantities are defined, via inclusion–exclusion (or a direct double summation argument) easily implies that for any disjoint sets Q and R, setting $S' = S \setminus (Q \cup R)$,

$$P(Q, R) = \sum_{Q' \subseteq S'} P\big(Q \cup Q',\, R \cup (S' \setminus Q')\big),$$

so the quantity $P(Q, R)$ is the value our attempt assigns to

$$\mathbb{P}\big(\{\eta: (\forall\, s \in Q,\, \eta(s) = 0) \text{ and } (\forall\, s \in R,\, \eta(s) = 1)\}\big).$$

We continue to refer to this simply as an attempt to define probabilities since some of these quantities may well be negative. We do note, however, that $P(\varnothing, \varnothing) = 1$ so the quantities $P(Q, S \setminus Q)$ do sum to 1 [by a special case of the above expression for $P(\cdot\,, \cdot\,)$].

Before moving on to deal with the question of the signs of the quantities we have defined, we observe that they have the required independence properties. We need only check that given pairwise disjoint vertex sets, Q_1, Q_2, R_1 and R_2 such that $Q_1 \cup R_1$ and $Q_2 \cup R_2$ are mutually independent, then $P(Q_1 \cup Q_2, R_1 \cup R_2) = P(Q_1, R_1)\, P(Q_2, R_2)$. This is straightforward.

In order to have a genuine probability distribution, all quantities $P(Q, S \setminus Q)$ must be nonnegative, and we would like $P(\varnothing, S) = 0$ (so that the measure does not dominate any π_p). However, if some quantities are negative, then it is not difficult to obtain a measure with the required properties, by increasing the value of p, as the following theorem states.

THEOREM 2.0. *Let $G = (S, E)$ be a finite graph and let $0 < p_0 < 1$. Suppose that for some (possibly empty) independent set of vertices, Q, $P_{p_0}(Q,$*

$S \setminus Q) \le 0$. *Then there exists* $p_1 \ge p_0$ *such that* $P_{p_1}(\varnothing, S) = 0$ *and, for all* Q, $P_{p_1}(Q, S \setminus Q) \ge 0$. $R \to P_{p_1}(S \setminus R, R)$ *is then a probability distribution on the subsets of* S, *and it has dependence graph* G.

PROOF. To begin, note that if Q is not an independent set of vertices then it is always the case that $P_p(Q, S \setminus Q) = 0$ for any value of p.

If Q is independent then, as $p \to 1$ (i.e., as $q \to 0$), the positive term $q^{|Q|}$ dominates the other terms in (2.0), so for p sufficiently close to 1, $P_p(Q, S \setminus Q)$ is strictly positive. Therefore, if we set

$$p_1 = \sup\{p < 1: \exists\ Q \text{ independent, with } P_p(Q, S \setminus Q) \le 0\},$$

then $p_1 < 1$. The conditions of the theorem imply that $p_1 \ge p_0$. The fact that, for each Q, $P_p(Q, S \setminus Q)$ is a continuous function of p, implies that, for some independent Q, $P_{p_1}(Q, S \setminus Q) = 0$, and also implies that, for all Q, $P_{p_1}(Q, S \setminus Q) \ge 0$. This last fact, together with the remarks preceding Theorem 2.0 tells us that $R \to P_{p_1}(S \setminus R, R)$ defines a genuine probability distribution on subsets of S, and that this distribution has independence graph G. It just remains to show that $P_{p_1}(\varnothing, S) = 0$; in other words, there is zero probability of a 1 at every site in S simultaneously.

We know that for some independent set Q, $P_{p_1}(Q, S \setminus Q) = 0$. Let W be the set of vertices in $S \setminus Q$ which have some neighbor in Q. Let X be those vertices neither in Q nor W. So S is a disjoint union, $S = Q \cup W \cup X$. Now, the fact that we cannot have zeros at neighboring sites implies that if all the vertices in Q have the value 0, then none of the vertices in W does. Also, X behaves independently of Q. Hence

$$0 = P_{p_1}(Q, S \setminus Q) = P_{p_1}(Q, \varnothing)\, P_{p_1}(\varnothing, X) = q_1^{|Q|} P_{p_1}(\varnothing, X), \tag{2.1}$$

where, of course, $q_1 = 1 - p_1$. Since $p_1 < 1$, (2.1) implies that $P_{p_1}(\varnothing, X) = 0$. So under the distribution given by P_{p_1}, the event $\{\eta: \forall\ x \in X,\ \eta(x) = 1\}$ has zero probability. This event contains the event $\{\eta: \forall\ x \in S,\ \eta(x) = 1\}$, so this latter event also has zero probability, that is, $P_{p_1}(\varnothing, S) = 0$, as required. □

To apply this theorem, we need to be able to evaluate quantities $P(Q, R)$ which, as Shearer (1985) remarks, is not always easy in practice. However, this can be done efficiently for certain graphs. In Shearer (1985), the necessary calculations are carried out for certain trees, which essentially include our first case (1-dependent measures on $\mathbb{Z}$). Of course, the graphs of greatest interest to us are infinite. However, if $G = (S, E)$ is an infinite graph and $0 < p < 1$, then, for any disjoint finite subsets of vertices, $Q, R \subseteq S$, we can still define the quantities $P_p(Q, R)$. We shall use the following variant of Theorem 2.0.

THEOREM 2.1. *Let* $G = (S, E)$ *be a (not necessarily finite) graph and let* $0 < p_0 < 1$. *Suppose that for some finite set* $R \subseteq V$, $P_{p_0}(\varnothing, R) \le 0$. *Then, for some* $p_1 \ge p_0$, *there exists a probability distribution on the subsets of* S *which*

has dependence graph G, in which the events $\{\eta\colon \eta(v) = 1\}$ all have probability p_1, and which does not stochastically dominate any nontrivial product measure.

PROOF. Apply Theorem 2.0 to the induced subgraph $G[R]$ and then extend to a distribution on all the subsets of S in which the vertices outside R behave independently on R and of each other. □

Let us now consider the case of k-dependent measures on $\mathbb{Z}$. So the graph under consideration is $\mathbb{Z}_k$, defined in Section 1: the vertex set is $\mathbb{Z}$ and two vertices are joined by an edge if they differ by at most k. Let $[n]$ be the set $\{1, \ldots, n\}$ and let

$$a_n = P_p(\varnothing, [n]).$$

It is easy to see that

$$a_{-} = 1, \qquad a_1 = 1 - q, \qquad a_2 = 1 - 2q, \ldots, a_{k+1} = 1 - (k+1)q$$

and for $n \geq k + 1$,

$$a_n = a_{n-1} - qa_{n-1-k}.$$

The difference equation follows from an argument similar to that used to complete the proof of Theorem 2.0: $P(\varnothing, [n]) = P(\varnothing, [n-1]) - P(\{n\}, [n-1]) = P(\varnothing, [n-1]) - P(\{n\}, \varnothing)P(\varnothing, [n-1-k]) = a_{n-1} - qa_{n-1-k}$.

Now suppose that $a_n > 0$ for all n. Consider the ratio of consecutive terms, a_n/a_{n-1}. We can see that for $n \leq k + 1$, this ratio is decreasing. By induction, we show that, for all n, $a_n/a_{n-1} < a_{n-1}/a_{n-2}$. Suppose that this is true for all values of n up to n_0 for some $n_0 \geq k + 1$. We have, using the above difference equation,

$$\frac{a_{n_0}}{a_{n_0-1}} = 1 - q\frac{a_{n_0-1-k}}{a_{n_0-1}} \quad \text{and} \quad \frac{a_{n_0+1}}{a_{n_0}} = 1 - q\frac{a_{n_0-k}}{a_{n_0}}.$$

By the induction hypothesis, together with the assumption that all the terms are positive,

$$\frac{a_{n_0-1-k}}{a_{n_0-1}} \Big/ \frac{a_{n_0-k}}{a_{n_0}} = \frac{a_{n_0}}{a_{n_0-1}} \Big/ \frac{a_{n_0-k}}{a_{n_0-k-1}} < 1.$$

Hence $a_{n_0+1}/a_{n_0} < a_{n_0}/a_{n_0-1}$, completing the induction. So the ratios of successive terms decrease to some limit, $a_n/a_{n-1} \to \lambda$, say.

Dividing the difference equation by a_{n-1-k}, we see that $\lambda^{k+1} = \lambda^k - q$, so, for some $0 \leq \lambda < 1$, we have $q = \lambda^k - \lambda^{k+1}$. This expression is maximized when $\lambda = k/(k+1)$, which gives $q = k^k/(k+1)^{k+1}$. If q is greater than this value then our assumption that a_n, is positive for all n must be false. Then applying Theorem 2.1 gives the following result.

COROLLARY 2.2. *Let $k \geq 1$. Then the critical value for domination from below by product measures of k-dependent measures on $\mathbb{Z}$ satisfies the following inequality*:

$$p_c^{\mathbb{Z}_k} \geq 1 - \frac{k^k}{(k+1)^{k+1}}.$$

Combining this with Corollary 1.6, we see that we know the precise value of the critical value for k-dependent measures in one dimension is given by

$$p_c^{\mathbb{Z}_k} = 1 - \frac{k^k}{(k+1)^{k+1}}.$$

Furthermore, we see that there is a discontinuous transition. if

$$p < 1 - \frac{k^k}{(k+1)^{k+1}}$$

than $\sigma^{\mathbb{Z}_k}(p) = 0$, yet Theorem 1.5 implies that

$$\sigma^{\mathbb{Z}_k}\left(1 - \frac{k^k}{(k+1)^{k+1}}\right) \geq \frac{k}{(k+1)^2}.$$

However, we do not know everything about the behavior of the function $\sigma^{\mathbb{Z}_k}$. In particular, we do not know the exact size of the discontinuous jump at $p_c^{\mathbb{Z}_k}$. The argument leading up to Corollary 2.2 does show that

$$\sigma^{\mathbb{Z}_k}\left(1 - \frac{k^k}{(k+1)^{k+1}}\right) \leq \frac{k}{(k+1)},$$

but this still leaves a substantial gap, especially for large k. Even when $k = 1$ (so $p_c = 3/4$) our best bounds give

$$\tfrac{1}{4} \leq \sigma^{\mathbb{Z}_1}\left(\tfrac{3}{4}\right) \leq \tfrac{1}{2}.$$

It seems a reasonable conjecture that the upper bound given by the counterexamples of this section is tight, whereas the lower bound is not best possible because we lose information when we introduce the random variables (Y_s) in Proposition 1.2.

Observe that it also follows from Corollaries 1.6 and 2.2 that

$$p_{W,c}^{\mathbb{Z}_k} = 1 - \frac{k^k}{(k+1)^{k+1}},$$

and yet

$$\sigma_W^{\mathbb{Z}_k}\left(1 - \frac{k^k}{(k+1)^{k+1}}\right) \geq \frac{k}{(k+1)^2}.$$

We also remark that the measures used to establish Corollary 2.2 are not translation invariant. It is not difficult to use them to construct translation invariant examples: given an n and a p_1 so that $P_{p_1}(\varnothing, [n]) = 0$, we break up

$\mathbb{Z}$ into intervals of length n. We can easily define a measure in which the blocks are mutually independent and in which there is zero probability that all the sites in any given block have state 1. Taking the average of n shifts of this measure, we obtain a translation invariant measure. However, this new measure is no longer k-dependent, although it is in $\mathbb{Z}_k^W(p_1) = C(1, k, p_1)$. It is rather surprising that, if we require the measure to be both translation invariant and k-dependent, then examples are rather harder to come by. This matter is discussed more fully in Section 3.

The calculations necessary to apply Theorem 2.1 to obtain higher-dimensional analogues of Corollary 2.2 do not seem to be so straightforward. One can obtain certain examples rather trivially by restricting to a subgraph which is essentially one-dimensional. For example, consider the dependence graph, G, with vertex set $\mathbb{Z}^2$ and two vertices joined by an edge if their l_∞ distance is 1, so each vertex has eight neighbors. This graph contains a subgraph which is isomorphic to the one-dimensional case $k = 2$: the graph spanned by all vertices of the form (n, n) and $(n, n + 1)$. So the examples used in that case directly translate to this case and show that $p_c^G \geq 23/27 = 1 - 2^2/3^3 = 0.8518\ldots$. In contrast, Theorem 1.3 shows that $p_c^G \leq 1 - 7^7/8^8 = 0.9509\ldots$.

With a little more work, it is possible to do slightly less trivial calculations which are not directly lifted from the one-dimensional case. For example, consider the case of a dependence graph with vertex set again $\mathbb{Z}^2$ and two vertices joined by an edge if their l_1-distance is 1; we denote this graph simply $\mathbb{Z}^2$ since it is the most usual graph with this vertex set. Let W_n be a set of n vertices defined as follows:

$$W_{2n} = \{(i, 0), (i, 1): 0 \leq i < n\}$$
$$W_{2n+1} = \{(i, 0), (i, 1): 0 \leq i < n\} \cup \{(n, 0)\}.$$

Using a similar approach to that employed in the one-dimensional case, let $a_n = P_p(\varnothing, W_n)$. Then we see that for $n \geq 1$, $a_{2n+1} = a_{2n} - qa_{2n-1}$ and $a_{2n+2} = a_{2n+1} - qa_{2n-1}$. With a little work, one can show that some of these terms must be negative unless $q \leq 3 - 2\sqrt{2}$, showing that $p_c^{\mathbb{Z}^2} \geq 2\sqrt{2} - 2 \approx 0.828$.

Although ad hoc methods such as this can be used to obtain certain numerical results on a case-by-case basis in higher dimensions, we do not believe that these results are best possible, nor does it seem easy to obtain good general results.

3. Domination by product measures of two-block factors. In the previous two sections we showed, for instance, that $p_c^{\mathbb{Z}_1} = 3/4$, and the corresponding transition is discontinuous in the sense that $\sigma^{\mathbb{Z}_1}(p_c^{\mathbb{Z}_1}) > 0$. It is natural to try to evaluate the critical p and ask whether or not the transition is continuous when $\mathbb{Z}_1(p)$ is replaced by other natural classes of measures.

For example, one could look at

$$T(p) = \{\mu \in \mathbb{Z}_1(p): \mu \text{ is translation invariant}\}.$$

We have not been able to answer these questions in this case, but will make some comments about this at the end of this section.

Another natural class of measures which could be considered in this context is

$$F(p) = \{\mu \in T(p): \mu \text{ is a two-block factor}\}.$$

Recall that the definition of two-block factors is given by (0.1), where we also point out that it is known that F is a proper subset of T.

The main result of this section is given below. For its statement, we define

$$\sigma^F(p) = \sup\{\rho: \mu \succcurlyeq \pi_\rho \ \forall \mu \in F(p)\},$$

and the critical value

$$p_c^F = \inf\{p: \sigma^F(p) > 0\}.$$

THEOREM 3.0. *For two-block factors, $p_c^F = \frac{1}{2}$, and the transition is continuous, in the sense that $\sigma^F(\frac{1}{2} +) = 0$.*

The proof of this result is broken down into a sequence of propositions. For the first of them, define a two-block factor to be symmetric if one can take f in the representation (0.1) to be a symmetric function of x and y.

PROPOSITION 3.1. *If $p \geq \frac{1}{2}$, then any measure in $F(p)$ dominates a symmetric measure in $F(2p - 1)$.*

PROOF. Take $\mu \in F(p)$, and let f be a function which represents it as in (0.1). Define

$$g(x, y) = \min\{f(x, y), f(y, x)\}.$$

Then g is symmetric, and satisfies

$$f(x, y) \geq g(x, y) \geq f(x, y) + f(y, x) - 1.$$

Therefore, the distribution of the process $\{X_n\}$ corresponding to f dominates the distribution of $\{Y_n\}$ corresponding to g, and

$$\mathbb{P}(Y_n = 1) = \int_0^1 \int_0^1 g(x, y)\, dx\, dy \geq \int_0^1 \int_0^1 [f(x, y) + f(y, x) - 1]\, dx\, dy$$
$$\geq 2p - 1. \qquad \square$$

PROPOSITION 3.2. *If μ is a symmetric two-block factor, then there is a random variable W satisfying $|W| \leq 1$ so that*

$$\mu(\{\eta: \eta(k) = 1 \ \forall\ 1 \leq k \leq n\}) = \mathbb{E}W^n$$

for all $n \geq 1$.

PROOF. Let f be the symmetric function which represents μ as in (0.1), and assume for the moment that for some integer N, f is constant on each square of the form $(j/N, (j + 1)/N] \times (k/N, (k + 1)/N]$. Define an operator

on $L_2[0,1]$ by

$$Au(x) = \int_0^1 f(x,y)\,u(y)\,dy. \tag{3.0}$$

This is a compact (even finite-dimensional) symmetric operator, so by the spectral theorem, it can be represented in the form

$$Au = \sum_k \lambda_k(\phi_k, u)\phi_k,$$

where the λ_k's are the eigenvalues of A (which are real) and the ϕ_k's are the corresponding complete, orthonormal family of eigenfunctions:

$$A\phi_k = \lambda_k \phi_k. \tag{3.1}$$

Combining (3.0) and (3.1) and taking absolute values, we see that

$$|\lambda_k||\phi_k(x)| \le \int_0^1 |f(x,y)||\phi_k(y)|dy \le \int_0^1 |\phi_k(y)|\,dy.$$

Integrating this with respect to x gives $|\lambda_k| \le 1$.

Using (0.1), write

$$\begin{aligned}\mu(\{\eta\colon \eta(k) = 1\ \forall\ 1 \le k \le n\}) &= \mathbb{E}\prod_{k=1}^n X_k \\ &= \mathbb{E}\prod_{k=1}^n f(U_k, U_{k+1}) \qquad (3.2)\\ &= \int_0^1 \cdots \int_0^1 \prod_{k=1}^n f(x_k, x_{k+1})\,dx_1 \ldots dx_{n+1}.\end{aligned}$$

To evaluate the right-hand side of this expression, we will show by induction that

$$\int_0^1 \cdots \int_0^1 \prod_{k=1}^n f(x_k, x_{k+1})\,dx_2 \ldots dx_{n+1} = \sum_j \lambda_j^n(\phi_j, 1)\phi_j(x_1). \tag{3.3}$$

For $n = 1$, the statement

$$\int_0^1 f(x,y)\,dy = \sum_j \lambda_j(\phi_j, 1)\phi_j(x)$$

is a consequence of (3.0) and the representation of $A1$ which follows it. For the induction step, write the left-hand side of (3.3) [using (3.3) for $n-1$] as

$$\begin{aligned}\int_0^1 f(x_1,x_2)\left[\sum_j \lambda_j^{n-1}(\phi_j,1)\phi_j(x_2)\right]dx_2 &= \sum_j \lambda_j^{n-1}(\phi_j,1)\,A\phi_j(x_1) \\ &= \sum_j \lambda_j^n(\phi_j,1)\phi_j(x_1).\end{aligned}$$

This proves (3.3). Now integrate (3.3) with respect to x_1 to conclude that the expression in (3.2) is

$$\sum_j \lambda_j^n(\phi_j,1)^2.$$

But this is the nth moment of a random variable W with distribution

$$\mathbb{P}\big(W = \lambda_j\big) = \big(\phi_j, 1\big)^2.$$

That these numbers define a genuine probability distribution follows from Parseval's relation:

$$\sum_j \big(\phi_j, 1\big)^2 = \|1\|_2^2 = 1.$$

Since $|\lambda_k| \le 1$ for each k, it follows that $|W| \le 1$.

Now take f to be a general symmetric indicator function, and let f_N be symmetric indicator functions which satisfy

$$\lim_{N\to\infty} \int_0^1 \int_0^1 |f_N(x, y) - f(x, y)|\, dx\, dy = 0,$$

and such that f_N is constant on each of the squares of the form $(j/N, (j+1)/N] \times (k/N, (k+1)/N]$. By the first part of the proof, there are random variables W_N which satisfy the conclusion of the theorem for the measure μ_N corresponding to the function f_N. Since the W_N's are uniformly bounded, we can extract a weakly convergent subsequence of their distributions, and hence construct a W corresponding to the given f, since the finite-dimensional μ_N-probabilities converge. □

PROPOSITION 3.3. *In the context of Proposition* 3.2, *if the X_n's have joint distribution μ, then*

$$\mathbb{P}\big(X_{n+1} = 1 | X_k = 1 \ \forall\ 1 \le k \le n\big) \ge \mathbb{E}W^2 \tag{3.4}$$

for each $n \ge 1$, where W is the random variable appearing in that proposition.

PROOF. Define α_n to be the left-hand side of (3.4). Since W^2 and W^{2n} are increasing functions of W^2, they are positively correlated [see Liggett (1985), page 65],

$$\mathbb{E}W^{2n+2} \ge \mathbb{E}W^{2n}\mathbb{E}W^2.$$

Since

$$\alpha_n = \frac{\mathbb{E}W^{n+1}}{\mathbb{E}W^n},$$

this implies that

$$\alpha_{2n}\alpha_{2n+1} \ge \mathbb{E}W^2.$$

The result follows from this, together with the fact that $0 \le \alpha_n \le 1$ for all n. □

PROPOSITION 3.4. *Suppose that μ is stationary and* 1*-dependent and that the corresponding process $\{X_n\}$ has conditional probabilities satisfying*

$$\mathbb{P}\big(X_{n+1} = 1 | X_k = 1 \ \forall\ 1 \le k \le n\big) \ge p$$

for all $n \ge 0$. Then $\mu \succcurlyeq \pi_\rho$ for

$$\rho = \big(1 - \sqrt{1-p}\big)^2.$$

PROOF. Let $\{Y_n\}$ be a process with distribution $\pi_{p'}$ which is independent of $\{X_n\}$, and put $Z_n = X_n Y_n$. Then $\{Z_n\}$ is again 1-dependent and stationary, and μ dominates the distribution of $\{Z_n\}$. We will find an upper bound for

$$\mathbb{P}\big(X_{n+1} = 0 | Z_j = \varepsilon_j \ \forall\ 1 \le j \le n\big) \tag{3.5}$$

for any choice of ε_j's in $\{0, 1\}$. If $\varepsilon_j = 1$ for all j, then this conditional probability is at most $q = 1 - p$ by hypothesis. So we can assume that at least one of the ε_j's is zero, and then choose the index j so that $\varepsilon_j = 0$ and $\varepsilon_j = 1$ for $j < i \le n$. Then the conditional probability (3.5) is bounded above by

$$\frac{\mathbb{P}(Z_i = \varepsilon_i \ \forall\ 1 \le i \le n,\ i \ne j;\ X_{n+1} = 0)}{\mathbb{P}\big(Z_i = \varepsilon_i \ \forall\ 1 \le i \le n,\ i \ne j;\ Y_j = 0\big)}.$$

Multiply and divide this fraction by

$$\mathbb{P}(Z_i = \varepsilon_i \ \forall\ 1 \le i \le n,\ i \ne j).$$

The result is the ratio of two conditional probabilities, which can be written as

$$\frac{\mathbb{P}\big(X_{n+1} = 0 | Z_i = 1 \ \forall\ j < i \le n\big)}{\mathbb{P}\big(Y_j = 0\big)}.$$

In doing so, we have used the 1-dependence of the $\{(X_n, Y_n)\}$ process in simplifying the numerator, and the fact that Y_j is independent of $\{Z_i,\ i \ne j\}$ in simplifying the denominator. The last fraction is bounded above by q/q', where $q' = 1 - p'$. Therefore, in all cases,

$$\mathbb{P}\big(Z_{n+1} = 0 | Z_i = \varepsilon_i \ \forall\ 1 \le i \le n\big) \le \frac{q}{q'} + q' - q.$$

Choosing $q' = \sqrt{q}$ makes this bound equal to $2\sqrt{q} - q$. By Lemma 1.1, it follows that the distribution of $\{Z_n\}$ (and hence μ) dominates π_ρ, where

$$\rho = 1 - 2\sqrt{q} + q = \big(1 - \sqrt{q}\big)^2 = \big(1 - \sqrt{1 - p}\big)^2. \qquad \square$$

Putting together the preceding propositions, we have the following result:

$$\sigma^F(p) \ge \Big(1 - \sqrt{4p(1-p)}\Big)^2$$

for $p \ge 1/2$. To see this, take $p \ge 1/2$ and $\mu \in F(p)$. By Proposition 3.1, there is a symmetric $\mu' \in F(p')$ with $p' = 2p - 1$ so that $\mu \succcurlyeq \mu'$. Applying Propositions 3.2 and 3.3 to μ', and letting W be the random variable appearing there, we have the conditional probabilities in (3.4) bounded below by $\mathbb{E}W^2 \ge (\mathbb{E}W)^2 \ge (p')^2$. Now simply apply Proposition 3.4.

Recalling the example in equation (0.0) we see now that $\sigma^F(\frac{1}{2}) = 0$, and hence the critical value is $1/2$. Our next objective is to obtain an upper bound for $\sigma^F(p)$ for $p > 1/2$ in order to show that the transition at $1/2$ is

continuous. Our construction is motivated by example (0.0). This example is the two-block factor with

$$f(x, y) = 1_{\{(x, y):\, x \le y\}}.$$

Here is an extension of this observation, which implies that $\sigma^F(\frac{1}{2} +) = 0$:

PROPOSITION 3.5.

$$\sigma^F\left(\frac{N+1}{2N}\right) \le \frac{1}{N}.$$

for each integer $N \ge 1$.

PROOF. Let μ be the distribution of the process given in (0.1) where

$$f(x, y) = 1_{\{(x, y):\, \lfloor xN \rfloor \le \lfloor yN \rfloor\}},$$

and $\lfloor \cdot \rfloor$ is the greatest integer function. Then

$$\mathbb{P}(X_k = 1 \ \forall\ 1 \le k \le n) = \mathbb{P}(\lfloor NU_1 \rfloor \le \lfloor NU_2 \rfloor \le \cdots \le \lfloor NU_{n+1} \rfloor).$$

To evaluate the right-hand side, note that for any $0 \le j_1 \le j_2 \le \cdots \le j_{N-1} \le n + 1$,

$$\mathbb{P}(\lfloor NU_k \rfloor = 0 \ \forall\ 0 < k \le j_1, \lfloor NU_k \rfloor = 1 \ \forall\ j_1 < k \le j_2, \ldots, \lfloor NU_k \rfloor = N - 1 \ \forall\ j_{N-1} < k \le n + 1) = \frac{1}{N^{n+1}},$$

and the number of choices of such $j_1, \ldots, j_{N-1}$ is $\binom{n+N}{n+1}$. Therefore,

$$\mathbb{P}(X_k = 1 \ \forall\ 1 \le k \le n) = \binom{n+N}{n+1}\frac{1}{N^{n+1}} \sim \frac{n^{N-1}}{N!\,N^n}$$

as $n \to \infty$, from which we conclude that $\rho \le 1/N$ whenever $\mu \succcurlyeq \pi_\rho$. Since

$$\mathbb{P}(X_1 = 1) = \frac{N+1}{2N},$$

the result follows. □

We conclude this section with some remarks about the class $T(p)$. Since

$$F(p) \subset T(p) \subset \mathbb{Z}_1(p),$$

it is an immediate consequence of the results of this paper that

$$\tfrac{1}{2} \le p_c^T \le \tfrac{3}{4}. \tag{3.6}$$

it would be interesting to know whether these inequalities are strict. Here is an argument which suggests that the first is strict, though we have not been able to carry out the details. Define

$$T_n(p) = \{\mu \in T(p): \mu(\{\eta: \eta(0) = \cdots = \eta(n) = 1\}) = 0\},$$

$$F_n(p) = \{\mu \in F(p): \mu(\{\eta: \eta(0) = \cdots = \eta(n) = 1\}) = 0\},$$

$$t_n = \sup\{p \in [0,1]: T_n(p) \ne \varnothing\} \quad \text{and} \quad f_n = \sup\{p \in [0,1]: F_n(p) \ne \varnothing\}.$$

It is clear that no measures in $T_n(p)$ or $F_n(p)$ can dominate a nontrivial product measure. Aaronson, Gilat, Keane and de Valk (1989) proved the following facts:

$$f_1 = t_1 = \tfrac{1}{4}, \qquad f_2 = \tfrac{1}{3} \quad \text{and} \quad t_2 \geq \tfrac{9}{25} > \tfrac{1}{3}.$$

It seems likely from this that

$$\sup_n t_n > \sup_n f_n. \tag{3.7}$$

Since it is not hard to show that

$$f_n \geq \frac{n}{2n+2},$$

if one could prove (3.7) it would follow that the first inequality in (3.6) is strict.

Acknowledgments. We thank Enrique Andjel for very useful discussions concerning this paper. We also thank the referee for having read the paper carefully and for suggesting the inclusion of a discussion along the lines of that which is presented after the statement of Theorem 0.0.

REFERENCES

AARONSON, J., GILAT, D., KEANE, M. and DE VALK, V. (1989). An algebraic construction of a class of one-dependent processes. *Ann. Probab.* **17** 128–143.

ANDJEL, E. (1993). Characteristic exponents for two-dimensional bootstrap percolation. *Ann. Probab.* **21** 926–935.

ANTAL, P. and PISZTORA, A. (1996). On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** 1036–1048.

BOLLOBÁS, B. (1985). *Random Graphs*. Academic Press, London.

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth & Brooks/Cole, Pacific Grove, CA.

ERDŐS, P. and LOVÁSZ, L. (1975). Problems and results on 3-chromatic hypergraphs and some related results. In *Infinite and Finite Sets* (A. Hajnal, R. Rado and V. T. Sós, eds.) *Coll. Math. Soc. János Bolyai* **11** 609–627.

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

PENROSE, M. and PISZTORA, A. (1996). Large deviations for discrete and continuous percolation. *Adv. in Appl. Probab.* **28** 29–52.

PISZTORA, A. (1996). Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields* **104** 427–466.

RUSSO, L. (1982). An approximate zero-one law. *Z. Wahrsch. Verw. Gebiete* **61** 129–139.

SCHONMANN, R. H. (1994). Theorems and conjectures on the droplet-driven relaxation of stochastic Ising models. In *Probability and Phase Transition* (G. Grimmett, ed.) 265–301. Kluwer, Dordrecht.

SHEARER, J. B. (1985). On a problem of Spencer. *Combinatorica* **5** 241–245.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024
E-MAIL: tml@math.ucla.edu
rhs@math.ucla.edu
stacey@math.ucla.edu