The Annals of Probability
2000, Vol. 28, No. 4, 1645–1665

MONOTONICITY OF CONDITIONAL DISTRIBUTIONS AND GROWTH MODELS ON TREES[1]

BY THOMAS M. LIGGETT

University of California, Los Angeles

We consider a sequence of probability measures ν_n obtained by conditioning a random vector $X = (X_1, \dots, X_d)$ with nonnegative integer valued components on

$$X_1 + \cdots + X_d = n - 1$$

and give several sufficient conditions on the distribution of X for ν_n to be stochastically increasing in n. The problem is motivated by an interacting particle system on the homogeneous tree in which each vertex has $d + 1$ neighbors. This system is a variant of the contact process and was studied recently by A. Puha. She showed that the critical value for this process is 1/4 if $d = 2$ and gave a conjectured expression for the critical value for all d. Our results confirm her conjecture, by showing that certain ν_n's defined in terms of d-ary Catalan numbers are stochastically increasing in n. The proof uses certain combinatorial identities satisfied by the d-ary Catalan numbers.

1. Introduction. For an integer $d \geq 2$, let $X = (X_1, \dots, X_d)$ be a random vector with values in

$$S = \{0, 1, 2, \dots\}^d = \bigcup_{n=1}^{\infty} S_n,$$

which we have written as the union of discrete simplices

$$S_n = \big\{x = (x_1, \dots, x_d) \in S : x_1 + \cdots + x_d = n - 1\big\}.$$

Suppose that

$$P\big(X_1 = x_1, \dots, X_d = x_d\big) > 0$$

for all choices of $x_1, \dots, x_d$. For each $n \geq 1$, define ν_n as the following conditional distribution on S_n:

$$\nu_n(x_1, \dots, x_d) = P\big(X_1 = x_1, \dots, X_d = x_d \mid X_1 + \cdots + X_d = n - 1\big).$$

The purpose of this paper is to find sufficient conditions for the measures ν_n to be stochastically increasing in n.

Specifically, we can regard S as a partially ordered set, with the partial order

$$(x_1, \dots, x_d) \leq (y_1, \dots, y_d) \quad \text{iff } x_i \leq y_i \text{ for each } 1 \leq i \leq d.$$

Received August 1999; revised March 2000.
[1]Supported in part by NSF Grant DMS-97-03830.
AMS 1991 *subject classification*. 60K35.
Key words and phrases. Stochastic monotonicity, contact process, growth models on trees, critical values, critical exponents, coupling, Catalan numbers.

A function f on S is increasing if $x \le y$ implies $f(x) \le f(y)$. Two probability measures μ and ν on S are said to be stochastically ordered, $\mu \le \nu$, if

$$\int f\, d\mu \le \int f\, d\nu \tag{1.1}$$

for all increasing functions f on S. A necessary and sufficient condition for this to be the case is that there exist a probability measure on $\{(x, y) \in S \times S : x \le y\}$ with marginals μ and ν respectively. [See Theorem 2.4 in Chapter II of Liggett (1985).] Such a measure is called a coupling measure. Our question then is to determine when $\nu_n \le \nu_{n+1}$ for each $n \ge 1$.

A lot of work has been done in which stochastic monotonicity is proved or used in various contexts. A recent book on the subject is Shaked and Shanthikumar (1994). The particular problem we are concerned with has apparently come up only a few times. In perhaps the first of these, Efron (1965) found a sufficient condition for $\nu_n \le \nu_{n+1}$ (see the remarks following the statement of Theorem 1.9 below). Joag-Dev and Proschan (1983) showed that in this situation, ν_n is negatively associated. In the opposite direction, Pemantle (2000) showed recently that certain measures on $\{0, 1\}^d$ that are negatively correlated in an appropriate sense have the property that $\nu_n \le \nu_{n+1}$. In his question 10, he asks the same question we do, but in the context of $\{0, 1\}^d$, and explains why a good answer to the question would advance his program of better understanding negative dependence.

A common way of checking stochastic monotonicity on partially ordered sets is to apply Holley's (1974) theorem, which appears as Theorem 2.9 of Chapter II of Liggett (1985), or one of its extensions. See Preston (1974), Karlin and Rinott (1980) and Theorem 4.E.5 of Shaked and Shanthikumar (1994), for example. It states that a sufficient condition for $\mu \le \nu$ is

$$\mu(x)\nu(y) \le \mu(x \wedge y)\nu(x \vee y), \qquad x, y \in S,$$

where $x \wedge y$ and $x \vee y$ denote the coordinatewise minimum and maximum of x and y, respectively. Note that this condition cannot be applied in the present context, since if $x \in S_n$ and $y \in S_{n+1}$, then $x \wedge y$ and $x \vee y$ will typically not be in $S_n \cup S_{n+1}$, so the right side above will be zero when applied to $\mu = \nu_n$ and $\nu = \nu_{n+1}$.

Before proceeding with a statement of results, we will describe the example that motivates this paper, and the application of our results to interacting particle systems. Define a sequence $\{c_n, n \ge 0\}$ recursively by $c_0 = 1$ and, for $n \ge 1$,

$$c_n = \sum_{x \in S_n} c_{x_1} \cdots c_{x_d}. \tag{1.2}$$

This sequence has various combinatorial interpretations, including the following, which is especially relevant to this paper: let T_d be the homogeneous tree in which each vertex has $d+1$ neighbors, and T_d^* be the tree whose root e has degree d and all other vertices have degree $d+1$. Then c_n is the number of connected subtrees of T_d^* containing e that have n vertices. In fact, (1.2) is simply the recursion that one gets if one tries to compute this number of subtrees recursively: e has d neighbors in T_d^*, and the summand on the right side of (1.2) is the number of connected subtrees of T_d^* with n vertices that contain e, and for which those of the d neighbors that are in the subtree have $x_1 - 1, \dots, x_d - 1$ offspring in that subtree, respectively.

Puha (1999) shows that

$$c_n = \frac{1}{(d-1)n+1}\binom{dn}{n}. \tag{1.3}$$

This also follows from Corollaries 2.4 and 2.7 of Hilton and Pedersen (1991). An alternate direct proof is contained in Proposition 6.2 below—see the remark following its proof. If $d = 2$, $\{c_n, n \geq 0\}$ are called the Catalan numbers. We will refer to $\{c_n, n \geq 0\}$ for general d as d-ary Catalan numbers. Note that this is not the only type of generalization of the Catalan numbers that has been studied. For example, the q-Catalan numbers discussed in Fürlinger and Hoffbauer (1985) are quite different.

Using the recursion (1.2), we can define a probability measure ν_n on S_n by

$$\nu_n(x) = \frac{c_{x_1} \cdots c_{x_d}}{c_n}. \tag{1.4}$$

To regard these as conditional measures in the present context, it suffices to let $X_1, \dots, X_d$ be i.i.d. random variables with distribution

$$P(X_i = k) = \frac{c_k u^k}{\sum_{j \geq 0} c_j u^j}$$

for some positive u that is sufficiently small so that the series in the denominator converges. We are interested in showing that $\nu_n \leq \nu_{n+1}$ for this choice, because of its application to a problem in interacting particle systems. We describe this problem next.

Puha (1999, 2000) are devoted to the study of a particular interacting particle system A_t on T_d. The state of the system at time t is a finite, connected subset of T_d. Neighbors of A_t are added at rate β each, and leaves of A_t are deleted at rate 1 each. (A leaf of A_t is a vertex in A_t with only one neighbor in A_t. If A_t is a singleton, its single member is considered to be a leaf.) The empty set is absorbing. The process A_t can be thought of as the contact process with infection parameter β, modified so as not to allow recoveries that would disconnect the infected set. [A treatment of the contact process on T_d can be found in Section 4 of Part I of Liggett (1999).] Just like the contact process, this system has a critical value β_c, which is defined by the requirement that

$$P^A(A_t = \varnothing \text{ for some } t) \begin{cases} = 1, & \text{if } \beta < \beta_c, \\ < 1, & \text{if } \beta > \beta_c, \end{cases}$$

for each finite, connected, nonempty $A \subset T_d$. Unlike the contact process, this process is reversible (away from $\varnothing$), and that suggests the possibility of explicit evaluation of the critical value. In her 1999 paper, Puha proves that

$$\beta_c \geq \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1} \tag{1.5}$$

for any d. She proves equality for $d = 2$, and conjectures equality for $d \geq 3$. The process dies out when β is equal to the right side of (1.5), so establishing the conjecture implies in particular that the critical process dies out. The right side of (1.5) becomes a bit less mysterious when one applies Stirling's formula to (1.3), and observes that

$$c_n \sim \frac{C}{n^{3/2}}\left[d\left(\frac{d}{d-1}\right)^{d-1}\right]^n.$$

Puha proved her results by giving a sufficient condition for equality to hold in (1.5). It is valid for all d and is stated at the beginning of Section 7 of this paper. It turns out that this sufficient condition is equivalent to $\nu_n \leq \nu_{n+1}$, $n \geq 1$ for the ν_n's given in (1.4), as we will prove in that section. Puha verified her condition for $d = 2$, but was unable to check it for larger d. From our current perspective, her verification of the condition for $d = 2$ can be viewed as checking $\nu_n \leq \nu_{n+1}$ by explicitly constructing a coupling measure. This appears not to be possible for larger d. We prove $\nu_n \leq \nu_{n+1}$ without constructing a coupling measure, and hence a consequence of our results is that Puha's condition is satisfied for $d \geq 3$ as well, and so equality holds in (1.5) for general d.

Proving the monotonicity of ν_n in n for $d = 2$ is relatively simple. In fact, one can easily give a necessary and sufficient condition for $\nu_n \leq \nu_{n+1}$ in this case [see (2.8) below]. This appears not to be possible for larger d, so we consider other approaches for $d = 2$ that yield only sufficient conditions, but that have a chance of generalizing to $d > 2$.

We begin with coupling, which is frequently a useful technique for proving stochastic monotonicity. Examples of its use can be found in Chapter II of Liggett (1985). In the present context, this approach would involve finding a continuous time, irreducible Markov chain (X_t, Y_t) on

$$\{(x, y) \in S_n \times S_{n+1} : x \leq y\}$$

with the property that the marginal processes X_t and Y_t are irreducible Markov chains with invariant measures ν_n and ν_{n+1}, respectively. A coupling measure is provided by the stationary distribution of (X_t, Y_t). A natural first attempt, then, is to check when simple choices of marginal processes can be coupled to yield useful sufficient conditions for $\nu_n \leq \nu_{n+1}$. Using this approach, we will show in Section 2 that a sufficient condition for $\nu_n \leq \nu_{n+1}$ in case $d = 2$ is

$$\frac{\nu_{n+1}(k+2, n-k-2)}{\nu_{n+1}(k+1, n-k-1)} \leq \frac{\nu_n(k+1, n-k-2)}{\nu_n(k, n-k-1)} \leq \frac{\nu_{n+1}(k+1, n-k-1)}{\nu_{n+1}(k, n-k)} \tag{1.6}$$

for $0 \leq k \leq n-2$.

Using a quite different argument, we will show there (again if $d = 2$) that another sufficient condition for $\nu_n \le \nu_{n+1}$ is

$$\nu_n(k, n-k-1) \ge \max\{\nu_{n+1}(k, n-k), \nu_{n+1}(k+1, n-k-1)\}, \quad 0 \le k \le n-1. \tag{1.7}$$

This is equivalent to saying that the ratios

$$\frac{\nu_n(k, n-k-1)}{\nu_{n+1}(k, n-k)} \quad \text{and} \quad \frac{\nu_n(k, n-k-1)}{\nu_{n+1}(k+1, n-k-1)} \tag{1.8}$$

are greater than or equal to 1, while (1.6) can be interpreted as saying that the first of these ratios is decreasing in k, while the second is increasing in k. Condition (1.7) will be used in the proof of our main result, Theorem 1.12 below.

If $d = 2$ and ν_n is given by (1.4), the ratios in (1.8) are

$$\frac{c_{n+1}}{c_n}\frac{c_{n-k-1}}{c_{n-k}} \quad \text{and} \quad \frac{c_{n+1}}{c_n}\frac{c_k}{c_{k+1}},$$

respectively. Thus in this case, (1.7) is equivalent to $c_k/c_{k+1} \downarrow$, while (1.6) is equivalent to $c_k/c_{k+1} \uparrow$. However,

$$\frac{c_k}{c_{k+1}} = \frac{k+2}{4k+2},$$

which is decreasing in k. Thus (1.7) is satisfied in this case, but (1.6) is not.

The following simple example should help the reader to understand the differences between (1.6), (1.7) and the necessary and sufficient condition (2.8): take $d = 2$ and $n = 2$ and suppose ν_2 and ν_3 are given by

$$\nu_2(0,1) = \nu_2(1,0) = \tfrac{1}{2}, \quad \nu_3(0,2) = \nu_3(2,0) = a, \ \nu_3(1,1) = 1-2a, \quad 0 \le a \le \tfrac{1}{2}.$$

Then (2.8) is satisfied for all such a, while (1.6) is satisfied if and only if $0 \le a \le \frac{1}{3}$, and (1.7) is satisfied if and only if $\frac{1}{4} \le a \le \frac{1}{2}$. Thus the sufficient conditions (1.6) and (1.7) are not comparable.

The balance of this paper is devoted to determining the extent to which the simple approaches that led to sufficient conditions (1.6) and (1.7) can be made to work for $d > 2$. The coupling technique used for (1.6) generalizes easily to all d, though it becomes necessary to assume that the X_i's are independent. The following result is proved in Section 3. The process used in the proof is the zero range process introduced by Spitzer (1970) and studied by many authors; see Kipnis and Landim (1999), for example. The assumption of the theorem amounts to saying that this process is attractive.

THEOREM 1.9. *Suppose that $X_1, \dots, X_d$ are independent. If*

$$\frac{P(X_i = k)}{P(X_i = k+1)} \quad \uparrow \ \textit{in } k \textit{ for each } i,$$

then $\nu_n \le \nu_{n+1}$ for each n.

This result is an almost direct consequence of attractiveness of the appropriate zero range processes. The result itself was first proved by Efron (1965), though of course his proof used neither coupling nor the zero process. We state the theorem here in order to be able to contrast it to Theorem 1.12 below, which is our main result. We include the proof partly because it is different from Efron's, but more importantly in order to explain that one cannot do much better by considering more general couplings: we start with a fairly general coupling and see why we are essentially forced to use the zero range process. The word "essentially" is important here. Once we know that $\nu_n \le \nu_{n+1}$, there is always a coupled Markov process that preserves the needed inequality: simply jump into the (naturally coupled) stationary distributions after an exponential time.

Note again that Theorem 1.9 does not apply to example (1.4), since

$$\rho(n) = \frac{c_n}{c_{n+1}} = \frac{1}{d}\prod_{j=1}^{d-1}\frac{(d-1)n+j+1}{dn+j} \tag{1.10}$$

is decreasing in n. Thus we are motivated to look for other sufficient conditions that do apply to this example. It is not enough to simply reverse the direction of the monotonicity assumption in Theorem 1.9, as the following example shows.

EXAMPLE. If $d = 2$, X_1, X_2 are i.i.d., and $n = 3$, the necessary and sufficient condition (2.8) becomes

$$P(X_1 = 1)P(X_1 = 3) \le 2\big[P(X_1 = 2)\big]^2.$$

Therefore, if we choose

$$P(X_1 = 0) = \frac{25}{32}, \qquad P(X_1 = 1) = \frac{5}{32}, \qquad P(X_1 = n) = \frac{1}{2^{n+3}} \quad \text{for } n \ge 2,$$

it follows that $\nu_3 \not\le \nu_4$, even though

$$\frac{P(X_1 = k)}{P(X_1 = k+1)} \quad \text{decreases in } k.$$

The extensions of the technique that led to (1.7) are somewhat harder than the coupling approach, but of greater interest because they do apply to example (1.4). We give two types of extension. The first applies only to the case $d = 3$, but the second applies to all d. Here is the first result, which is proved in Section 4. The first assumption in Theorem 1.11 is an obvious analogue of (1.7). The theorem states that given this, in order to show $\nu_n(B)\uparrow$ for every increasing set B, it suffices to check it for all increasing sets of the special form $B = \{(j,k,l): j \ge j_0, k \ge k_0, l \ge l_0\}$.

THEOREM 1.11. *Suppose that* $d = 3$. *If*

$$\nu_n(j,k,l) \ge \max\big\{\nu_{n+1}(j+1,k,l), \nu_{n+1}(j,k+1,l), \nu_{n+1}(j,k,l+1)\big\} \quad \textit{for } (j,k,l) \in S_n$$

and

$$\sum_{j\geq j_0,\,k\geq k_0,\,l\geq l_0} \nu_n(j,k,l) \leq \sum_{j\geq j_0,\,k\geq k_0,\,l\geq l_0} \nu_{n+1}(j,k,l) \quad \textit{for } j_0, k_0, l_0 \geq 0,$$

then $\nu_n \leq \nu_{n+1}$.

In Section 5, we prove the following result, which works for all d. It is the main result of this paper and the one that leads to our solution of the motivating problem for the growth model A_t. Note that if $d = 2$, the assumptions of Theorem 1.12 reduce to (1.7) (for independent X_1, X_2).

THEOREM 1.12. *Suppose that* $X_1, \dots, X_d$ *are independent. If*

$$\begin{aligned} &P(X_{i+1} = m)P(X_1 + \cdots + X_{i+1} = l + m + 1) \\ &\qquad \geq P(X_{i+1} = m + 1)P(X_1 + \cdots + X_{i+1} = l + m) \end{aligned}$$

and

$$\begin{aligned} &P(X_1 + \cdots + X_i = l)P(X_1 + \cdots + X_{i+1} = l + m + 1) \\ &\qquad \geq P(X_1 + \cdots + X_i = l + 1)P(X_1 + \cdots + X_{i+1} = l + m) \end{aligned}$$

for $l, m \geq 0$ *and* $1 \leq i \leq d - 1$. *Then* $\nu_n \leq \nu_{n+1}$ *for* $n \geq 1$.

REMARK. The independence assumption above can be weakened somewhat, as explained in the course of the proof. Also, in the proof of this result, sufficient condition (1.7) (for $d = 2$) is used in a critical way. Using the necessary and sufficient condition (2.8) in its place would lead to a further weakening of the hypotheses, but they would then become more cumbersome.

One situation in which $\nu_n \leq \nu_{n+1}$ is obvious is that in which $X_1, \dots, X_d$ are independent Poisson distributed random variables with parameters $\lambda_1, \dots, \lambda_d$, respectively, since then ν_n is the multinomial distribution with parameters $n - 1$ and $p_1, \dots, p_d$, where

$$p_i = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_d}.$$

It is perhaps instructive to see what the hypotheses of Theorems 1.9 and 1.12 become in this case. In the context of Theorem 1.9,

$$\frac{P(X_i = k)}{P(X_i = k + 1)} = \frac{k + 1}{\lambda_i},$$

so the assumption of that theorem is automatically satisfied. The assumptions of Theorem 1.12, on the other hand, become

$$\frac{m}{l + 1} \leq \frac{\lambda_{i+1}}{\lambda_1 + \cdots + \lambda_i} \leq \frac{m + 1}{l},$$

which are not satisfied. This should not be regarded as implying that Theorem 1.9 is more useful or important than Theorem 1.12. Rather, they provide quite different types of sufficient conditions, with the latter being more subtle.

Section 6 is devoted to applying Theorem 1.12 to the ν_n given in (1.4). The observation that makes this possible is that, in this case, the distributions of the partial sums $X_1 + \cdots + X_i$ can be computed explicitly. The result is the following.

COROLLARY 1.13. *If* ν_n *is given by* (1.4), *then* $\nu_n \le \nu_{n+1}$ *for* $n \ge 1$.

In Section 7 we check that Puha's sufficient condition for equality in (1.5) is equivalent to $\nu_n \le \nu_{n+1}$ for the measures given in example (1.4). Combining this with Corollary 1.13 and Puha's work in her two papers, we obtain the following.

THEOREM 1.14. (a) *The critical value* β_c *for the process* A_t *is given by*

$$\beta_c = \frac{1}{d}\left(\frac{d-1}{d}\right)^{d-1}.$$

(b) *The survival probability for* A_t *starting from a singleton lies asymptotically between constant multiples of*

$$(\beta - \beta_c)^{7/2} \quad \text{and} \quad (\beta - \beta_c)^{5/2}$$

as $\beta \downarrow \beta_c$.

REMARK. In her paper, Puha uses the explicit coupling measure available in case $d = 2$ to improve the power 7/2 significantly. Since our coupling is not explicit, it does not appear to lead to a similar improvement for $d \ge 3$. There is some numerical evidence for the critical exponent to be $\frac{5}{2}$ [Tretyakov and Konno (2000)], and it would certainly be of interest to prove this, even for $d = 2$.

2. The case $d = 2$.

As mentioned in the introduction, it is easy to give a necessary and sufficient condition for the monotonicity of ν_n in n in this case, as we will do at the end of this section. Since this does not extend to larger d, we begin by considering several different approaches to proving $\nu_n \le \nu_{n+1}$ in the case $d = 2$ in the hopes that they will be more generally useful. We begin by carrying out the coupling proof that leads to condition (1.6). To simplify notation, record only the first coordinate of a point in S_n, identifying $(k, n-k-1)$ with k, so write $S_n = \{0, \dots, n-1\}$. We will require that the marginal chains move only to nearest neighbors, say for the process on S_n,

$$k \to k+1 \quad \text{at rate } \alpha_n(k),$$
$$k \to k-1 \quad \text{at rate } \beta_n(k).$$

Of course, $\alpha_n(n-1) = \beta_n(0) = 0$. In order that the chain have the right stationary distribution, these rates must satisfy the following conditions:

$$\alpha_n(k)\nu_n(k) = \beta_n(k+1)\nu_n(k+1). \tag{2.1}$$

We will not make any special assumptions about the coupled chain, other than those needed to make it couple the marginal processes correctly. The coupled chain has state space

$$\{(k,k), 0 \le k \le n-1\} \cup \{(k,k+1), 0 \le k \le n-1\},$$

and possible transitions

$$\begin{aligned}
(k,k) &\to (k,k+1) \text{ at rate } a_k,\\
(k,k) &\to (k-1,k-1) \text{ at rate } b_k,\\
(k,k) &\to (k+1,k+1) \text{ at rate } c_k,\\
(k,k) &\to (k-1,k) \text{ at rate } d_k,\\
(k,k+1) &\to (k,k) \text{ at rate } e_k,\\
(k,k+1) &\to (k+1,k+1) \text{ at rate } f_k,\\
(k,k+1) &\to (k-1,k) \text{ at rate } g_k,\\
(k,k+1) &\to (k+1,k+2) \text{ at rate } h_k.
\end{aligned}$$

In order for the transition rates of the marginal processes to be correct, the following relations must be satisfied:

$$\begin{aligned}
\alpha_n(k) &= c_k = f_k + h_k,\\
\beta_n(k) &= g_k = b_k + d_k,\\
\alpha_{n+1}(k) &= h_{k-1} = a_k + c_k,\\
\beta_{n+1}(k) &= b_k = e_{k-1} + g_{k-1}.
\end{aligned}$$

Solving these equations leads to

$$\begin{gathered}
c_k = \alpha_n(k), \qquad g_k = \beta_n(k), \qquad h_k = \alpha_{n+1}(k+1), \qquad b_k = \beta_{n+1}(k),\\
f_k = \alpha_n(k) - \alpha_{n+1}(k+1), \qquad d_k = \beta_n(k) - \beta_{n+1}(k),\\
a_k = \alpha_{n+1}(k) - \alpha_n(k), \qquad e_k = \beta_{n+1}(k+1) - \beta_n(k).
\end{gathered}$$

These will be nonnegative provided that

$$\alpha_{n+1}(k+1) \le \alpha_n(k) \le \alpha_{n+1}(k), \qquad \beta_{n+1}(k) \le \beta_n(k) \le \beta_{n+1}(k+1). \tag{2.2}$$

It is then not hard to check that it is possible to choose the α's and β's so that (2.1) and (2.2) are satisfied if and only if

$$\frac{\nu_{n+1}(k+2)}{\nu_{n+1}(k+1)} \le \frac{\nu_n(k+1)}{\nu_n(k)} \le \frac{\nu_{n+1}(k+1)}{\nu_{n+1}(k)}. \tag{2.3}$$

This is exactly *not* the case of primary interest in this paper, as explained in the introduction. Let us consider then another approach to proving $\nu_n \le \nu_{n+1}$ that will lead to the sufficient condition (1.7). To prove (1.1), it is enough to prove

$$\nu_n(B) \le \nu_{n+1}(B) \tag{2.4}$$

for all increasing $B \subset S$. Since ν_n concentrates on S_n, only $B \cap S_n$ and $B \cap S_{n+1}$ are relevant in (2.4). Letting $A = B \cap S_n$ and

$$A^* = \bigcup_{k \in A} \{k, k+1\} = A \cup (A+1) \subset B \cap S_{n+1},$$

we see that (2.4) is implied by

$$\nu_n(A) \le \nu_{n+1}(A^*), \qquad A \subset S_n. \tag{2.5}$$

$\big[$Since A^* is a subset of S_{n+1} rather than of S_n, perhaps one should clarify its definition by writing

$$A^* = \big\{(k, n-k), (k+1, n-k-1): (k, n-k-1) \in A\big\}.\big]$$

Writing A as a disjoint union of maximal intervals, we see that (2.5) will hold for all A if it holds for all intervals. The important observation in checking this is that if $A = A_1 \cup A_2$ where A_1 and A_2 are separated in that $k \in A_1$, $l \in A_2$ implies that $|k - l| > 1$, then $A_1^* \cap A_2^* = \varnothing$, so that

$$\nu_n(A_1 \cup A_2) = \nu_n(A_1) + \nu_n(A_2) \quad \text{and} \quad \nu_{n+1}(A_1^* \cup A_2^*) = \nu_{n+1}(A_1^*) + \nu_{n+1}(A_2^*).$$

Thus, it is enough to check

$$\nu_n(A^c) \ge \nu_{n+1}\big((A^*)^c\big) \tag{2.6}$$

for all intervals A. But the complement of an interval is a union of (one or) two intervals, so (2.6) will hold for all intervals if

$$\nu_k(k) \ge \max\big\{\nu_{n+1}(k), \nu_{n+1}(k+1)\big\}, \qquad 0 \le k \le n-1. \tag{2.7}$$

But this is just (1.7) in the present notation.

The two techniques explained thus far are proposed as approaches that might suggest ideas for use for larger d. When $d = 2$, it is easy to give a necessary and sufficient condition for $\nu_n \le \nu_{n+1}$, as we now explain. Let the coupling measure put mass

$$\begin{aligned} &\lambda_k \quad \text{on } (k, k), \\ &\rho_k \quad \text{on } (k, k+1) \end{aligned}$$

for $0 \le k \le n-1$. The marginals will be correct provided that

$$\nu_n(k) = \lambda_k + \rho_k \quad \text{and} \quad \nu_{n+1}(k) = \rho_{k-1} + \lambda_k,$$

where we have set $\rho_{-1} = \lambda_n = 0$. Solving for the λ's and ρ's leads to

$$\lambda_k = \sum_{j=0}^{k} \nu_{n+1}(j) - \sum_{j=0}^{k-1} \nu_n(j),$$

$$\rho_k = \sum_{j=0}^{k} \nu_n(j) - \sum_{j=0}^{k} \nu_{n+1}(j).$$

Thus, the necessary and sufficient condition for $\nu_n \leq \nu_{n+1}$ is

$$\sum_{j=0}^{k-1} \nu_n(j) \leq \sum_{j=0}^{k} \nu_{n+1}(j) \leq \sum_{j=0}^{k} \nu_n(j) \tag{2.8}$$

for $0 \leq k \leq n-1$.

3. The coupling condition. In this section, we take a general $d \geq 2$ and ask when one can prove $\nu_n \leq \nu_{n+1}$ via a natural coupling argument. The result will be a proof of Theorem 1.9. Let $b_i(x) = P(X_i = x)$ so that

$$\nu_n(x) = B_n b_1(x_1) \cdots b_d(x_d),$$

where

$$B_n = \left[\sum_{x \in S_n} b_1(x_1) \cdots b_d(x_d) \right]^{-1}.$$

Let $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ be the ith unit vector: the 1 appears as the ith coordinate. Consider the continuous time Markov chain X_t on S that has the transition

$$x \to x - e_i + e_j$$

at rate $\sigma_{i,j}(x)$ for $x = (x_1, \dots, x_d) \in S$ and $1 \leq i \neq j \leq d$. Note that if we interpret the coordinates x_i of x as the numbers of particles at the sites $\{1, \dots, d\}$, then this transition corresponds to moving a particle from site i to site j. Initially, we will let the transition rates depend on the full configuration x in an arbitrary way, and will not assume that ν_n has the product form given above. Assume that $\sigma_{i,j}(x) > 0$ except when $x_i = 0$, when it must be zero since there is no particle at i to move. This guarantees that X_t is irreducible when restricted to each S_n.

We will have proved that $\nu_n \leq \nu_{n+1}$ if we can choose $\sigma_{i,j}(x)$ so that the following properties hold:

1. X_t is reversible with respect to ν_n for each n, that is,

$$\nu_n(x)\sigma_{i,j}(x) = \nu_n(x - e_i + e_j)\sigma_{j,i}(x - e_i + e_j) \tag{3.1}$$

for $x \in S_n$, $x_i \geq 1$.

2. If $x \le y$, then two copies X_t and Y_t with initial states x and y, respectively, can be coupled to preserve the relation $X_t \le Y_t$. It is enough to carry out this coupling in case y is obtained from x by adding 1 to a single component, say $y = x + e_k$. Then the most natural coupling would have transitions for the joint process (X_t, Y_t),

$$(x, x+e_k) \to (x - e_i + e_j, x + e_k - e_i + e_j) \quad \text{if } i, j \ne k,$$
$$(x, x+e_k) \to (x - e_k + e_j, x + e_j) \text{ or } (x, x + e_j) \quad \text{if } j \ne k,$$

and

$$(x, x+e_k) \to (x - e_i + e_k, x - e_i + 2e_k) \text{ or } (x - e_i + e_k, x + e_k) \quad \text{if } i \ne k.$$

To carry out this coupling so that the marginals have the correct law, we need

$$\begin{aligned} \sigma_{i,j}(x) &= \sigma_{i,j}(x+e_k) \quad \text{for } i, j \ne k, \\ \sigma_{k,j}(x) &\le \sigma_{k,j}(x+e_k) \quad \text{for } j \ne k, \\ \sigma_{i,k}(x) &\ge \sigma_{i,k}(x+e_k) \quad \text{for } i \ne k. \end{aligned} \tag{3.2}$$

The first constraint in (3.2) says that $\sigma_{i,j}(x)$ depends on x only through x_i and x_j, so write

$$\sigma_{i,j}(x) = \sigma_{i,j}(x_i, x_j),$$

and then the other two constraints in (3.2) become

$$\sigma_{i,j}(u, v) \text{ is increasing in } u \text{ and decreasing in } v. \tag{3.3}$$

In these terms, (3.1) becomes

$$\nu_n(x)\sigma_{i,j}(x_i, x_j) = \nu_n(x - e_i + e_j)\sigma_{j,i}(x_j + 1, x_i - 1). \tag{3.4}$$

This forces σ to take a special form. To see this, take i, j, k distinct, and write the following two relations of the form (3.4):

$$\nu_n(x)\sigma_{i,k}(x_i, x_k) = \nu_n(x - e_i + e_k)\sigma_{k,i}(x_k + 1, x_i - 1),$$
$$\nu_n(x - e_i + e_k)\sigma_{k,j}(x_k + 1, x_j) = \nu_n(x - e_i + e_j)\sigma_{j,k}(x_j + 1, x_k).$$

Combining these with (3.4) leads to

$$\frac{\sigma_{j,i}(x_j + 1, x_i - 1)}{\sigma_{i,j}(x_i, x_j)} = \frac{\sigma_{k,i}(x_k + 1, x_i - 1)}{\sigma_{i,k}(x_i, x_k)} \frac{\sigma_{j,k}(x_j + 1, x_k)}{\sigma_{k,j}(x_k + 1, x_j)}. \tag{3.5}$$

This says that the left side of (3.5) can be written in the form

$$\frac{\sigma_{j,i}(x_j + 1, x_i - 1)}{\sigma_{i,j}(x_i, x_j)} = \frac{\lambda_j(x_j)}{\lambda_i(x_i - 1)} \tag{3.6}$$

for some functions $\lambda_i(u)$. Using this in (3.4) leads to

$$\nu_n(x)\lambda_i(x_i - 1) = \nu_n(x - e_i + e_j)\lambda_j(x_j).$$

Iterating this relation gives

$$\nu_n(x) = \frac{B_n}{\prod_{i=1}^{d}\left[\prod_{j=0}^{x_i-1}\lambda_i(j)\right]},$$

where B_n is the normalizing constant

$$B_n = \sum_{x\in S_n}\left[\prod_{i=1}^{d}\left[\prod_{j=0}^{x_i-1}\lambda_i(j)\right]\right]^{-1}.$$

This is of the form given at the beginning of this section, with

$$b_i(k) = \left[\prod_{j=0}^{k-1}\lambda_i(j)\right]^{-1}.$$

Given the λ's, we can set

$$\sigma_{i,j}(k,l) = \lambda_i(k-1).$$

This makes (3.6) hold, and then (3.3) is equivalent to

$$\lambda_i(k) \quad \text{increases in } k.$$

Since

$$\lambda_i(k) = \frac{b_i(k)}{b_i(k+1)},$$

the proof of Theorem 1.9 is complete. The fact that the rate for a particle to move from i to j depends only on the number of particles at i makes the marginal processes zero range.

4. A sufficient condition when $d = 3$. This section is devoted to the proof of Theorem 1.11. We assume throughout the hypotheses of that theorem. The first part of the proof follows that of the sufficiency of (2.7) in $d = 2$. In particular, in order to prove $\nu_n \le \nu_{n+1}$ in case $d = 3$, it suffices to check that

$$\nu_n(A) \le \nu_{n+1}(A^*) \tag{4.1}$$

for all connected $A \subset S_n$. Here connectedness is to be interpreted in terms of the graph structure on S_n in which two points $x, y \in S_n$ are neighbors if $\sum_{i=1}^{d}|x_i - y_i| = 2$, that is, if y can be obtained from x by reducing one coordinate by one and increasing another coordinate by one. Given A, A^* is now defined by

$$A^* = \bigcup_{(j,k,l)\in A}\{(j+1,k,l),(j,k+1,l),(j,k,l+1)\}.$$

We begin with the part of the argument that requires $d = 3$. By an edge of S_n, we will mean one of the sets

$$\{(j,k,l)\in S_n: j=0\}, \{(j,k,l)\in S_n: k=0\}, \{(j,k,l)\in S_n: l=0\}.$$

The corners of S_n are the points $(0,0,n-1)$, $(0,n-1,0)$ and $(n-1,0,0)$.

LEMMA 4.2. *Suppose A and B are disjoint connected subsets of S_n. Then it is not possible for both A and B to contain a point from each of the three edges of S_n.*

PROOF. Assume that A and B are disjoint, connected, and contain a point on each of the three edges of S_n. We want to reach a contradiction. First, assume that one of the sets contains one of the corners of S_n, say $(0, 0, n-1) \in A$. Then A must contain a point of the form $(j, n-j-1, 0)$. Since A is connected, there must be a connected path through A joining $(0, 0, n-1)$ and $(j, n-j-1, 0)$. But B cannot contain $(0, 0, n-1)$, so it contains two points of the form $(0, k, n-k-1)$ and $(l, 0, n-l-1)$ with $k, l \geq 1$. Since B is connected, there must be a connected path through B joining $(0, k, n-k-1)$ and $(l, 0, n-l-1)$. But these two paths must intersect, and this gives the required contradiction.

In the other case, neither A nor B contains a corner of S_n. Therefore, there is a point $(j, k, l) \in A$ that is connected via connected paths in A to points in the interiors of each of the three edges of S_n. Without loss of generality, these paths can be taken to be nonself-intersecting, and disjoint except for the common point (j, k, l). Then, we may take A to be the union of these paths. In this case, $S_n \backslash A$ breaks up into three components, each including exactly one of the corners of S_n. Since B is connected, it is contained in one of these components, say the one that contains $(0, 0, n-1)$. But then B cannot contain any point whose third coordinate is zero. $\square$

By a subsimplex of S_n, we will mean a subset of S_n of the form

$$\{(j, k, l) \in S_n : j \geq j_0, k \geq k_0, l \geq l_0\}$$

for some $j_0, k_0, l_0 \geq 0$. Note that this is naturally isomorphic to $S_{n-j_0-k_0-l_0}$. Also, if A is a subsimplex of S_n, then A^* is a subsimplex of S_{n+1} (corresponding to the same j_0, k_0, l_0).

PROPOSITION 4.3. *If* (4.1) *holds for all subsimplices of S_n, then it holds for all connected subsets of S_n.*

PROOF. Let A be a connected subset of S_n, and define B to be the smallest subsimplex of S_n that contains A. By assumption, (4.1) holds with (A, A^*) replaced by (B, B^*), so it will be enough to show that

$$\nu_{n+1}(B^* \backslash A^*) \leq \nu_n(B \backslash A). \tag{4.4}$$

This is an analogue of (2.6).

We will prove (4.4) by finding an injective mapping

$$\phi: B^* \backslash A^* \to B \backslash A$$

with the property that

$$\phi(j, k, l) \in \{(j-1, k, l), (j, k-1, l), (j, k, l-1)\}$$

for every $(j,k,l) \in B^*\backslash A^*$. By the first assumption of Theorem 1.11, it will follow that

$$\nu_{n+1}(j,k,l) \le \nu_n(\phi(j,k,l)), \qquad (j,k,l) \in B^*\backslash A^*.$$

Summing this expression leads to (4.4), since ϕ is injective.

To define ϕ, let $C_1,\dots,C_N$ be the maximal connected components of $B^*\backslash A^*$. A given C_i cannot contain a point on each of the edges of B^*. To see this, note that by the minimality of B, A^* has a point on each edge of B^*, so C_i cannot have that property by Lemma 4.2. (Recall that B^* is isomorphic to S_m for some m.) If C_i has no point whose first coordinate is zero, define $\phi(j,k,l) = (j-1,k,l)$ for $(j,k,l) \in C_i$. If it does contain a point with zero first coordinate, but no point with zero second coordinate, define $\phi(j,k,l) = (j,k-1,l)$ for $(j,k,l) \in C_i$. Otherwise, C_i contains no point with zero third coordinate, and then define $\phi(j,k,l) = (j,k,l-1)$ for $(j,k,l) \in C_i$. Clearly, the $\phi(j,k,l)$ defined in this way is in $B\backslash A$, since if it were in A, we would have $(j,k,l) \in A^*$.

We need to check that ϕ is injective. So, suppose that $\phi(j_1,k_1,l_1) = \phi(j_2,k_2,l_2)$ for two distinct points in $B^*\backslash A^*$. Since ϕ is injective on each C_i, there must be $i_1 \neq i_2$ so that $(j_1,k_1,l_1) \in C_{i_1}$ and $(j_2,k_2,l_2) \in C_{i_2}$. Since $\phi(j_1,k_1,l_1) = \phi(j_2,k_2,l_2)$, (j_1,k_1,l_1) and (j_2,k_2,l_2) are neighbors in B^*, and this means that $C_{i_1} \cup C_{i_2}$ is connected, which contradicts the maximality of the C_i's. $\square$

Theorem 1.11 follows from Proposition 4.3, together with the remarks at the beginning of this section.

5. A sufficient condition for general d. Here we prove Theorem 1.12. Let $B \subset S$ be an increasing set. We need to show that the conditional probability

$$\nu_n(B) = P(X \in B \mid X_1 + \cdots + X_d = n-1) \tag{5.1}$$

is an increasing function of n. To do so, for $1 \le i \le d$, let

$$f_i(k; j_1,\dots,j_{d-1}) = P\big(X \in B \mid X_1 + \cdots + X_i = k, X_{i+1} = j_1,\dots,X_d = j_{d-i}\big).$$

We will show by induction on i that f_i is an increasing function of all of its arguments. Note that

$$f_1(k; j_1,\dots,j_{d-i}) = \begin{cases} 1, & \text{if } (k, j_1,\dots,j_{d-1}) \in B, \\ 0, & \text{otherwise,} \end{cases}$$

so f_1 is increasing in its arguments because B is an increasing set. This gives the basis step for the induction proof. On the other hand,

$$f_d(k) = P(X \in B \mid X_1 + \cdots + X_d = k),$$

so that its monotonicity in k is exactly the statement that needs to be proved. Thus it will suffice to carry out the induction argument.

To carry out the induction step, write

$$
\begin{aligned}
f_{i+1}&(k; j_1, \dots, j_{d-i-1})\\
&= \sum_m P\big(X \in B, X_{i+1} = m \mid X_1 + \cdots + X_{i+1} = k,\\
&\qquad\qquad\qquad X_{i+2} = j_1, \dots, X_d = j_{d-i-1}\big)\\
&= \sum_m \Big[f_i(k-m; m, j_1, \dots, j_{d-i-1})\\
&\qquad \times P(X_{i+1} = m \mid X_1 + \cdots + X_{i+1} = k, X_{i+2} = j_1, \dots, X_d = j_{d-i-1})\Big]\\
&= \sum_m f_i(k-m; m, j_1, \dots, j_{d-i-1}) \mu_{k;\, j_1, \dots, j_{d-i-1}}(k-m, m),
\end{aligned}
$$

where $\mu_{k;\, j_1,\dots,j_{d-i-1}}(\cdot,\cdot)$ is the probability measure on $\{(p,q)\colon p+q = k\}$ defined by the final identity

$$
\begin{aligned}
\mu_{k;\, j_1,\dots,j_{d-i-1}}&(k-m, m)\\
&= P\big(X_{i+1} = m \mid X_1 + \cdots + X_{i+1} = k, X_{i+2} = j_1, \dots, X_d = j_{d-i-1}\big).
\end{aligned}
$$

By the inductive hypothesis, f_i is increasing in all of its variables, so it suffices to show that the measure $\mu_{k;\, j_1,\dots,j_{d-i-1}}(\cdot,\cdot)$ is increasing in $k, j_1, \dots, j_{d-i-1}$.

Up to this point, we have not used the independence of $X_1, \cdots, X_d$ in order to explain why we assume this, and to show that the argument does work under slightly weaker hypotheses. When the j's vary, the measures $\mu_{k;\, j_1,\dots,j_{d-i-1}}(\cdot,\cdot)$ are supported by the same set $\{(p,q)\colon p+q = k\}$. However, the only way that two probability measures on this set can be stochastically ordered is for them to be the same. To see this, note that a coupling measure would have to concentrate on

$$\big\{((p_1, q_1), (p_2, q_2))\colon p_1 + q_1 = p_2 + q_2 = k,\ p_1 \le p_2,\ q_1 \le q_2\big\},$$

and any probability measure on this set has equal marginals. Thus one must assume that these measures are independent of the j_l variables. This is weaker than independence of $X_1, \dots, X_d$, but we will not pursue extensions based on this observation.

Assuming now the independence of $X_1, \dots, X_d$, we may eliminate the j subscripts, and define

$$\mu_k(l, m) = \frac{P(X_{i+1} = m) P(X_1 + \cdots + X_i = l)}{P(X_1 + \cdots + X_{i+1} = l + m)}, \qquad l + m = k.$$

To complete the proof, we need to check that $\mu_k \le \mu_{k+1}$ for all k. But the hypotheses of Theorem 1.12 are exactly (1.7) for the measures μ_k. So the proof of the induction step is complete, using the implication (2.7) [which is the same as (1.7)] $\Rightarrow$ (2.4) that was proved in Section 2.

6. Application to example (1.4). In this section, d is general and c_n and ν_n are given by (1.2) [equivalently, (1.3)] and (1.4), respectively. We will prove Corollary 1.13 by checking the hypotheses of Theorem 1.12 in this case. Since these hypotheses involve the distributions of partial sums of X_i's, we need to consider convolutions of c_n's.

Define the i-fold convolutions $c_n^{(i)}$ of c_n with itself by

$$c_n^{(0)} = \begin{cases} 1, & \text{if } n = 0, \\ 0, & \text{if } n \geq 1, \end{cases}$$

$$c_n^{(1)} = c_n,$$

$$c_n^{(i+1)} = \sum_{k=0}^{n} c_n^{(i)} c_{n-k}, \qquad i \geq 0.$$

Note that by (1.2),

$$c_n^{(d)} = c_{n+1}. \tag{6.1}$$

It turns out that these convolutions can be computed explicitly.

PROPOSITION 6.2.

$$c_n^{(i)} = i \frac{\binom{nd+i-1}{i-1}}{\binom{n(d-1)+i}{i-1}} c_n, \qquad i \geq 1. \tag{6.3}$$

PROOF. We will prove (6.3) by induction on i. It is clearly true for $i = 1$. For the induction step, it suffices to check that the ratio of the left side of (6.3) for two successive values of i is the same as the ratio of the right side of (6.3) for the same successive values of i. There is a lot of cancellation in the ratio of the right sides, and this leads to the following: it suffices to show that

$$\frac{c_n^{(i+1)}}{c_n^{(i)}} = \frac{i+1}{i} \frac{nd+i}{n(d-1)+i+1}. \tag{6.4}$$

Let

$$C(u) = \sum_{n=0}^{\infty} c_n u^n$$

be the generating function of the sequence c_n. The radius of convergence of the series is strictly positive by (1.3). Multiplying (1.2) by u^{n-1} and summing for $n \geq 1$ gives

$$uC^d(u) - C(u) + 1 = 0. \tag{6.5}$$

[See equation (4.3) of Puha (1999).] Since the generating function of a convolution is the product of the corresponding generating functions,

$$\sum_{n=0}^{\infty} c_n^{(i)} u^n = C^i(u). \tag{6.6}$$

Cross-multiplying (6.4), and then multiplying the result by u^n and summing on n, we see that (6.4) is equivalent to

$$i \sum_{n=0}^{\infty} [n(d-1)+i+1] c_n^{(i+1)} u^n = (i+1) \sum_{n=0}^{\infty} (nd+i) c_n^{(i)} u^n \tag{6.7}$$

for all sufficiently small u. Let D denote differentiation with respect to u:

$$Df(u) = f'(u).$$

By (6.6), (6.7) can be written as

$$i(d-1)uDC^{i+1}(u) + i(i+1)C^{i+1}(u) = (i+1)duDC^i(u) + i(i+1)C^i(u).$$

By carrying out the differentiations and cancelling common factors, we see that this is equivalent to

$$(d-1)uC(u)C'(u) + C^2(u) = duC'(u) + C(u). \tag{6.8}$$

To check that (6.8) is true, differentiate (6.5),

$$C^d(u) + udC^{d-1}(u)C'(u) - C'(u) = 0.$$

Then replace $C^d(u)$ both places it appears in this expression by $[C(u)-1]/u$, which also comes from (6.5). The result is (6.8). $\square$

Remark. We have used (1.3) only to conclude that $C(u)$ has positive radius of convergence. Puha (1999) gives a simple direct argument for this fact; see (4.4) there. Given this, (6.1) and Proposition 6.2 provide an alternate proof of (1.3),

$$c_{n+1} = c_n^{(d)} = d \frac{\binom{nd+d-1}{d-1}}{\binom{nd+d-n}{d-1}} c_n = \frac{(d-1)n+1}{(d-1)(n+1)+1} \frac{\binom{(n+1)d}{n+1}}{\binom{nd}{n}} c_n.$$

It follows that

$$c_n[(d-1)n+1] \Big/ \binom{dn}{n}$$

is independent of n, and hence always equal to 1, since that is its value for $n = 1$. Of course, once we know (1.3), it can be combined with (6.3) to write

$$c_n^{(i)} = \frac{i}{nd+i} \binom{nd+i}{n}. \tag{6.9}$$

This equation [and working backward, (6.3) as well] can also be obtained from Theorems 2.3 and 2.5 of Hilton and Pedersen (1991). Their results are proved via combinatorial arguments. We have included our simple analytic proof in order to keep this paper self-contained.

PROOF OF COROLLARY 1.13. In this case,

$$P(X_1 + \cdots + X_i = n) = \frac{c_n^{(i)} u^n}{C^i(u)},$$

where $u > 0$ is fixed, and small enough that $C(u) < \infty$. Therefore, the assumptions of Theorem 1.12 become

$$c_m c_{l+m+1}^{(i+1)} \geq c_{m+1} c_{l+m}^{(i+1)} \quad \text{and} \quad c_l^{(i)} c_{l+m+1}^{(i+1)} \geq c_{l+1}^{(i)} c_{l+m}^{(i+1)},$$

respectively. These will both follow from

$$c_m^{(i)} c_{n+1}^{(j)} \geq c_{m+1}^{(i)} c_n^{(j)}, \qquad 1 \leq i \leq j \leq d, 0 \leq m \leq n.$$

This is just the statement that the ratios

$$\frac{c_m^{(i)}}{c_{m+1}^{(i)}} \tag{6.10}$$

are decreasing in m for $m \geq 0$ and decreasing in i for $1 \leq i \leq d$. Using (6.9), we have

$$\begin{aligned}\frac{c_m^{(i)}}{c_{m+1}^{(i)}} &= (m+1)\frac{[(m+1)(d-1)+i]\cdots[m(d-1)+i+1]}{[(m+1)d+i-1]\cdots[md+i]} \\ &= \frac{1}{d}\prod_{j=1}^{i-1}\frac{(m+1)(d-1)+j+1}{(m+1)d+j}\prod_{j=i}^{d-1}\frac{m(d-1)+j+1}{md+j}.\end{aligned}$$

Note that this generalizes (1.10). Each of the factors in the above products is decreasing in m, so the monotonicity in m of the ratios in (6.10) is immediate. The monotonicity in i is equivalent to

$$\frac{m(d-1)+i+1}{md+i} \geq \frac{(m+1)(d-1)+i+1}{(m+1)d+i},$$

which is easily checked. So, Corollary 1.13 follows from Theorem 1.12. □

7. Equivalent form of Puha's condition.

Puha (1999) proved that the following is a sufficient condition for equality in (1.5); see her Lemma 15. There exists nonnegative, permutation invariant functions $\alpha_i(n, x)$ defined for $1 \leq i \leq d$, $n \geq 1$, $x \in S_n$ that satisfy

$$\sum_{i=1}^{d} \alpha_i(m, x) = 1 \quad \text{and} \quad \sum_{i=1}^{d} \alpha_i(n, y - e_i)\frac{c_{y_i} - 1}{c_{y_i}} = \frac{c_n}{c_{n+1}} \tag{7.1}$$

for all $x \in S_m$, $y \in S_{n+1}$, $m \geq 1$, $n \geq 1$, where the c_n's are given by (1.3) for $n \geq 0$, and $c_{-1} = 0$. Here permutation invariant means that

$$\alpha_{\sigma(i)}(n, \sigma(x)) = \alpha_i(n, x)$$

for every permutation σ of $\{1, \ldots, d\}$, and $\sigma(x)$ is the vector obtained by applying σ to the coordinates of x. [Actually, her condition is a bit weaker in that the nonnegativity assumption is replaced by $|\alpha_i(n, x)| \le 1$. However, this is implied by nonnegativity and the first equation in (7.1).]

PROPOSITION 7.2. *Let ν_n be defined by* (1.4). *Then there exists a nonnegative, permutation invariant solution of* (7.1) *if and only if $\nu_n \le \nu_{n+1}$ for each $n \ge 1$.*

PROOF. By Theorem 2.4 in Chapter II of Liggett (1985), $\nu_n \le \nu_{n+1}$ is equivalent to the existence of a probability measure μ on

$$\{(x, y) \in S_n \times S_{n+1} : x \le y\}$$

that satisfies

$$\nu_n(x) = \sum_{y \in S_{n+1}} \mu(x, y) \quad \text{and} \quad \nu_{n+1}(x) = \sum_{x \in S_n} \mu(x, y). \tag{7.3}$$

Since ν_n and ν_{n+1} are permutation invariant, the μ can also be taken to be permutation invariant. (If μ is not already permutation invariant, replace it by an appropriate average of permutations of μ.) Any (x, y) in the support of μ must be of the form $(x, x + e_i)$ for some $1 \le i \le d$, so given either μ or the α's we can define the other via

$$\alpha_i(n, x) = \frac{\mu(x, x + e_i)}{\nu_n(x)}.$$

With this identification,

$$\sum_{y \in S_{n+1}} \mu(x, y) = \sum_{i=1}^{d} \mu(x, x + e_i) = \nu_n(x) \sum_{i=1}^{d} \alpha_i(n, x)$$

and

$$\begin{aligned} \sum_{x \in S_n} \mu(x, y) &= \sum_{i=1}^{d} \mu(y - e_i, y) = \sum_{i=1}^{d} \alpha_i(n, y - e_i) \nu_n(y - e_i) \\ &= \sum_{i=1}^{d} \alpha_i(n, y - e_i) \frac{c_{y_i} - 1}{c_{y_i}} \nu_{n+1}(y) \frac{c_{n+1}}{c_n}. \end{aligned}$$

Therefore, (7.1) and (7.3) are equivalent. □

REFERENCES

EFRON, B. (1965). Increasing properties of Pólya frequency functions. *Ann. Math. Statist.* **36** 272–279.

FÜRLINGER, J. and HOFBAUER, J. (1985). q-Catalan numbers. *J. Combin. Theory Ser. A* **40** 248–264.

HILTON, P. and PEDERSEN, J. (1991). Catalan numbers, their generalization and their uses. *Math. Intelligencer* **13** 64–75.

HOLLEY, R. (1974). Remarks on the FKG inequalities. *Comm. Math. Phys.* **36** 227–231.

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

KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities I: multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467–498.

KIPNIS, C. and LANDIM, C. (1999). *Scaling Limits of Interacting Particle Systems*. Springer, Berlin.

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

LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, Berlin.

PEMANTLE, R. (2000). Towards a theory of negative dependence. Unpublished manuscript.

PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Comm. Math. Phys.* **36** 233–241.

PUHA, A. L. (1999). A reversible nearest particle system on the homogeneous tree. *J. Theoret. Probab.* **12** 217–254.

PUHA, A. L. (2000). Critical exponents for a reversible nearest particle system on the binary tree. *Ann. Probab.* **28** 395–415.

SHAKED, M. and SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, New York.

SPITZER, F. (1970). Interaction of Markov processes. *Adv. Math.* **5** 246–290.

TRETYAKOV, A. Y. and KONNO, N. (2000). Survival probability for uniform model on binary tree: Critical behavior and scaling. In *Soft Computing in Industrial Applications*. Springer, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90095
E-MAIL: tml@math.ucla.edu