Realizable Monotonicity and Inverse Probability Transform

James Allen Fill (jimfill@jhu.edu) *

Department of Mathematical Sciences, The Johns Hopkins University

Motoya Machida (machida@math.usu.edu)

Department of Mathematics and Statistics, Utah State University

July 19, 2000

Abstract. A system $(P_\alpha : \alpha \in A)$ of probability measures on a common state space S indexed by another index set A can be "realized" by a system $(X_\alpha : \alpha \in A)$ of S-valued random variables on some probability space in such a way that each X_α is distributed as P_α. Assuming that A and S are both partially ordered, we may ask when the system $(P_\alpha : \alpha \in A)$ can be realized by a system $(X_\alpha : \alpha \in A)$ with the monotonicity property that $X_\alpha \le X_\beta$ almost surely whenever $\alpha \le \beta$. When such a realization is possible, we call the system $(P_\alpha : \alpha \in A)$ "realizably monotone." Such a system necessarily is stochastically monotone, that is, satisfies $P_\alpha \le P_\beta$ in stochastic ordering whenever $\alpha \le \beta$. In general, stochastic monotonicity is not sufficient for realizable monotonicity. However, for some particular choices of partial orderings in a finite state setting, these two notions of monotonicity are equivalent. We develop an inverse probability transform for a certain broad class of posets S, and use it to explicitly construct a system $(X_\alpha : \alpha \in A)$ realizing the monotonicity of a stochastically monotone system when the two notions of monotonicity are equivalent.

Keywords: Realizable monotonicity, stochastic monotonicity, monotonicity equivalence, perfect sampling, partially ordered set, Strassen's theorem, marginal problem, inverse probability transform, synchronizing function, synchronizable.

AMS subject classification: Primary 60E05; secondary 06A06, 60J10, 05C05, 05C38.

1. Introduction

1.1. Two notions of monotonicity

We will discuss two notions of monotonicity for probability measures on a finite partially ordered set (poset). Let $\mathcal{S}$ be a finite poset and let (P_1, P_2) be a pair of probability measures on S. (We use a calligraphic letter $\mathcal{S}$ in order to distinguish the set S from the same set equipped with a partial ordering $\le$.) A subset U of S is said to be an *up-set* in

* Research for both authors was supported by NSF grant DMS-98-03780.

$\mathcal{S}$ (or *increasing set*) if $y \in U$ whenever $x \in U$ and $x \leq y$. We say that P_1 is *stochastically smaller* than P_2, denoted $P_1 \preceq P_2$, if

$$P_1(U) \leq P_2(U) \quad \text{for every up-set } U \text{ in } \mathcal{S}. \tag{1.1}$$

An important characterization of stochastic ordering was established by Strassen (1965) and fully investigated by Kamae et al. (1977). They show that (1.1) is necessary and sufficient for the existence of a pair $(\mathbf{X}_1, \mathbf{X}_2)$ of S-valued random variables [defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$] satisfying the properties that $\mathbf{X}_1 \leq \mathbf{X}_2$ and that $\mathbb{P}(\mathbf{X}_i \in \cdot) = P_i(\cdot)$ for $i = 1, 2$.

Now let $\mathcal{A}$ be a finite poset. Let $(P_\alpha : \alpha \in A)$ be a system of probability measures on S. We call $(P_\alpha : \alpha \in A)$ a *realizably monotone* system if there exists a system $(\mathbf{X}_\alpha : \alpha \in A)$ of S-valued random variables such that

$$\mathbf{X}_\alpha \leq \mathbf{X}_\beta \quad \text{whenever } \alpha \leq \beta \tag{1.2}$$

and

$$\mathbb{P}(\mathbf{X}_\alpha \in \cdot) = P_\alpha(\cdot) \quad \text{for every } \alpha \in A. \tag{1.3}$$

In such a case we shall say that $(\mathbf{X}_\alpha : \alpha \in A)$ *realizes the monotonicity* of $(P_\alpha : \alpha \in A)$. The (easier half of the) characterization of stochastic ordering applied pairwise implies

$$P_\alpha \preceq P_\beta \quad \text{whenever } \alpha \leq \beta. \tag{1.4}$$

The system $(P_\alpha : \alpha \in A)$ is said to be *stochastically monotone* if it satisfies (1.4). Thus, stochastic monotonicity is necessary for realizable monotonicity.

In light of Strassen's characterization of stochastic ordering, one might guess that stochastic monotonicity is also sufficient for realizable monotonicity. It is perhaps surprising that the conjecture is false in general. Various counterexamples are given by Fill and Machida (2000), including one independently discovered by Ross (1993). Given a pair $(\mathcal{A}, \mathcal{S})$ of posets, if the two notions of monotonicity—stochastic and realizable—are equivalent, then we say that *monotonicity equivalence* holds for $(\mathcal{A}, \mathcal{S})$.

1.2. Inverse probability transform

Suppose that $\mathcal{S}$ is linearly ordered. Then, for a given probability measure P on S, we can define its *inverse probability transform* P^{-1} by

$$P^{-1}(t) := \min\{x \in S : t < F(x)\}, \quad t \in [0, 1), \tag{1.5}$$

where $F(x) := P(\{z \in S : z \leq x\})$ is the distribution function for P. Furthermore, let $\mathcal{A}$ be any poset, and let $(P_\alpha : \alpha \in A)$ be a stochastically monotone system of probability measures on S. Given a single uniform random variable $\mathbf{U}$ on $[0, 1)$, we can construct a system $(\mathbf{X}_\alpha : \alpha \in A)$ of S-valued random variables via $\mathbf{X}_\alpha := P_\alpha^{-1}(\mathbf{U})$ which realizes the monotonicity. This proves that monotonicity equivalence always holds for $(\mathcal{A}, \mathcal{S})$ when $\mathcal{S}$ is linearly ordered.

In Section 2 we generalize the definition of inverse probability transform to a certain class of posets $\mathcal{S}$ which are not necessarily linearly ordered. We then extend the construction in the preceding paragraph and present Theorems 2.2 and 2.3, thereby establishing monotonicity equivalence under certain additional assumptions. A further extension of Theorem 2.3 is discussed briefly in Section 3, which culminates in Theorem 3.1. We will not discuss the proofs of Theorems 2.2, 2.3, and 3.1 in the present brief paper, but rather refer the reader to Machida (1999) for (the highly technical) proofs and more extensive discussion.

1.3. Importance in perfect sampling algorithms

Of particular interest in our general study of realizable monotonicity is the case $\mathcal{A} = \mathcal{S}$. Here the system $(P(x, \cdot) : x \in S)$ of probability measures can be considered as a Markov transition matrix $\mathbf{P}$ on the state space S. Propp and Wilson (1996) and Fill (1998) introduced algorithms to produce observations distributed *perfectly* according to the long-run distribution of a Markov chain. Both algorithms apply most readily and operate most efficiently when the state space $\mathcal{S}$ is a poset and a suitable monotonicity condition holds. Of the many differences between the two algorithms, one is that the appropriate notion of monotonicity for the Propp–Wilson algorithm is realizable monotonicity, while for Fill's algorithm it is stochastic monotonicity; see Remark 4.5 in Fill (1998). Here the properties (1.2)–(1.3) are essential for the Propp–Wilson algorithm to be able to generate transitions simultaneously from every state in such a way as to preserve ordering relations. For further discussion of these perfect sampling algorithms in the monotone setting, see Fill (1998) and Propp and Wilson (1996); for further discussion of perfect sampling in general, consult the annotated bibliography at `http://www.dbwilson.com/exact/`. Fill and Machida (2000) show that the two notions of monotonicity are equivalent if and only if the poset $\mathcal{S}$ is acyclic; see Section 2.1 herein for the definition of this term.

2. A generalization of inverse probability transform

2.1. Distribution functions on an acyclic poset

We begin with a notion of acyclic poset, and its use in introducing a distribution function on such a poset. Most of the basic poset terminology adopted here can be found in Stanley (1986) or Trotter (1992), and most of the graph-theoretic terminology in West (1996). Let $\mathcal{S}$ be a poset. For $x, y \in S$, we say that y *covers* x if $x < y$ in $\mathcal{S}$ and no element z of S satisfies $x < z < y$. We define the *cover graph* $(S, \mathcal{E}_{\mathcal{S}})$ of $\mathcal{S}$ to be the undirected graph with edge set $\mathcal{E}_{\mathcal{S}}$ consisting of those unordered pairs $\{x, y\}$ such that either x covers y or y covers x in $\mathcal{S}$. A poset $\mathcal{S}$ is said to be *acyclic* when its cover graph $(S, \mathcal{E}_{\mathcal{S}})$ is acyclic in the usual graph-theoretic sense (i.e., the graph has no cycle).

Throughout the sequel we assume that the cover graph $(S, \mathcal{E}_{\mathcal{S}})$ is acyclic and also connected, that is, that the graph $(S, \mathcal{E}_{\mathcal{S}})$ is a tree. Let τ be a fixed leaf of $(S, \mathcal{E}_{\mathcal{S}})$, that is, an element τ in S such that there exists a unique edge $\{\tau, z\}$ in $\mathcal{E}_{\mathcal{S}}$ (for some $z \in S$). Then, declare $x \le_\tau y$ for $x, y \in S$ if the (necessarily existent and unique) path $(\tau, \dots, x)$ in the graph from τ to x contains the path $(\tau, \dots, y)$ from τ to y as a segment. This introduces a partial ordering $\le_\tau$ on the ground set S [Bogart (1996)], which may be different from $\le$ for the original poset $\mathcal{S}$. We call this new poset $(S, \le_\tau)$ a *rooted tree* (rooted at τ).

For each $x \in S$, set

$$\mathcal{C}(x) := \{z \in S : x \text{ covers } z \text{ in } (S, \le_\tau)\}.$$

Then a linear extension $(S, \le_\psi)$ of $(S, \le_\tau)$ can be obtained by choosing a linear ordering on $\mathcal{C}(x)$ for every $x \in S$. Explicitly, we define $x \le_\psi y$ if either (i) $x \le_\tau y$, or (ii) there exist some $z \in S$ and some $w, w' \in \mathcal{C}(z)$ such that $x \le_\tau w$, $y \le_\tau w'$, and w has been chosen to be smaller than w' in $\mathcal{C}(z)$. See Section 2.3 for an example.

DEFINITION 2.1. For a given probability measure P on S, we define the *distribution function* $F(\cdot)$ of P by

$$F(x) := P(\{z \in S : z \le_\tau x\}) \quad \text{for each } x \in S,$$

and the *distribution function* $F\lceil\cdot\rceil$ *of linear extension* by

$$F\lceil x\rceil := P(\{z \in S : z \le_\psi x\}) \quad \text{for each } x \in S.$$

In particular, when $(S, \mathcal{E}_{\mathcal{S}})$ is a path from one end point τ to the other end point, the rooted tree $(S, \le_\tau)$ is linearly ordered, and therefore $F(\cdot) \equiv F\lceil\cdot\rceil$.

2.2. Inverse probability transform

For a given distribution function $F\lceil\cdot\rceil$ of linear extension on S, we define the *inverse probability transform* P^{-1}, a map from $[0,1)$ to S, by

$$P^{-1}(t) := \min\{x \in S : t < F\lceil x\rceil\} \quad \text{for } t \in [0,1), \tag{2.6}$$

where the minimum is given in terms of the linearly ordered set $(S, \leq_\psi)$. When $\mathcal{S}$ is linearly ordered, the two definitions of inverse probability transform in (1.5) and (2.6) are the same. This equivalence can be extended to the case that the cover graph $(S, \mathcal{E}_\mathcal{S})$ is a path, because of the fact that then $F(\cdot) \equiv F\lceil\cdot\rceil$. Moreover, the property of inverse probability transform discussed in Section 1.2 remains true in that case:

THEOREM 2.2. *Let* $\mathbf{U}$ *be a uniform random variable on* $[0,1)$. *Suppose that* $(S, \mathcal{E}_\mathcal{S})$ *is a path. Then, a stochastically monotone system* $(P_\alpha : \alpha \in A)$ *is always realizably monotone via* $\mathbf{X}_\alpha := P_\alpha^{-1}(\mathbf{U})$.

Theorem 2.2 reiterates a result presented by Fill and Machida (2000), namely, Theorem 6.1 in their paper. An acyclic poset $\mathcal{S}$ is called a *poset of Class* Z if the cover graph $(S, \mathcal{E}_\mathcal{S})$ is a path. Otherwise, the acyclic (connected) poset $\mathcal{S}$ has a multiple-element $\mathcal{C}(x)$ for some $x \in S$. An example in Section 2.3 will demonstrate that Theorem 2.2 can fail when an acyclic poset $\mathcal{S}$ is not in Class Z. Besides the result for Class Z, Fill and Machida gave a complete answer to the monotonicity equivalence problem [i.e., the question whether monotonicity equivalence holds for given $(\mathcal{A}, \mathcal{S})$] when there exists some $x \in S$ such that (i) $\mathcal{C}(x)$ contains at least two elements, and (ii) x is neither minimal nor maximal in $\mathcal{S}$, that is, when an acyclic poset $\mathcal{S}$ falls into either *Class* B or *Class* Y, in their terms. [In their investigation, a construction of random variables with the desired properties (1.2)–(1.3) was reduced to application of Strassen's characterization of stochastic ordering if monotonicity equivalence holds for $(\mathcal{A}, \mathcal{S})$ with $\mathcal{S}$ a poset either of Class B or of Class Y.] However, when $\mathcal{S}$ is a poset satisfying the property that x is either maximal or minimal in $\mathcal{S}$ whenever $\mathcal{C}(x)$ contains at least two elements, which they (and we) call a *poset of Class* W, we do not know a complete answer to the monotonicity equivalence problem. But for a poset $\mathcal{S}$ of Class W our generalization of inverse probability transform can, for some posets $\mathcal{A}$, be used to establish monotonicity equivalence:

THEOREM 2.3. *Let* $\mathbf{U}$ *be a uniform random variable on* $[0,1)$. *Suppose that* $\mathcal{S}$ *is a poset of Class* W, *and that* $\mathcal{A}$ *is a poset having a minimum element and a maximum element. Then, given a stochastically monotone system* $(P_\alpha : \alpha \in A)$, *there exists a system* $(\phi_\alpha : \alpha \in A)$

of $\mathbf{U}$*-invariant maps from* $[0,1)$ *to* $[0,1)$ *[i.e.,* $\phi(\mathbf{U}) \stackrel{\mathcal{L}}{=} \mathbf{U}$*] such that*

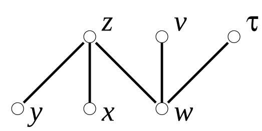
$$\mathbf{X}_\alpha := P_\alpha^{-1}(\phi_\alpha(\mathbf{U})), \quad \alpha \in A,$$

realizes the monotonicity.

We call the $\mathbf{U}$-invariant maps ϕ_α in Theorem 2.3 *synchronizing functions*. Discussion about how we can practically construct the desired synchronizing functions can be found in Machida (1999).

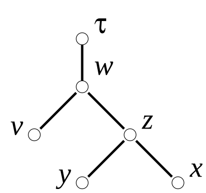
2.3. An example

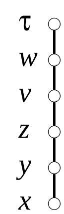
Consider the poset $\mathcal{S}$ of Class W with the following Hasse diagram:

$\mathcal{S} :=$ 

;

the downward arc from z to y, for example, indicates that z covers y in $\mathcal{S}$. Then the rooted tree $(S, \leq_\tau)$ rooted at τ and the corresponding linear extension $(S, \leq_\psi)$ are given respectively by

$(S, \leq_\tau) =$ 

and $(S, \leq_\psi) =$ 

,

where we have chosen the linear ordering $z < v$ in $\mathcal{C}(w)$ and the linear ordering $x < y$ in $\mathcal{C}(z)$. Let P_1 and P_2 be the two probability measures on S in the following table:

ξ	x	y	z	v	w	τ
$P_1(\{\xi\})$	3/15	2/15	1/15	1/15	7/15	1/15
$P_2(\{\xi\})$	1/15	1/15	6/15	3/15	2/15	2/15

.

We can easily check that $P_1 \preceq P_2$. However, $P_1^{-1}(t) \leq_{\mathcal{S}} P_2^{-1}(t)$ does not hold for all $t \in [0,1)$: For $t \in [\frac{1}{15}, \frac{2}{15})$, $x = P_1^{-1}(t)$ is incomparable with $y = P_2^{-1}(t)$, and again for $t \in [\frac{6}{15}, \frac{7}{15})$, $v = P_1^{-1}(t)$ is incomparable with $z = P_2^{-1}(t)$.

Figure 1 displays the synchronizing functions ϕ_1 and ϕ_2, which are both $\mathbf{U}$-invariant from $[0,1)$ to $[0,1)$. Then consider the map $P_k^{-1} \circ \phi_k$ from

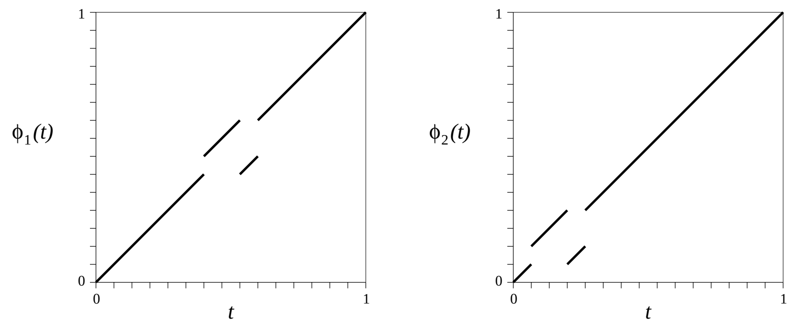


Figure 1. The synchronizing functions ϕ_1 and ϕ_2

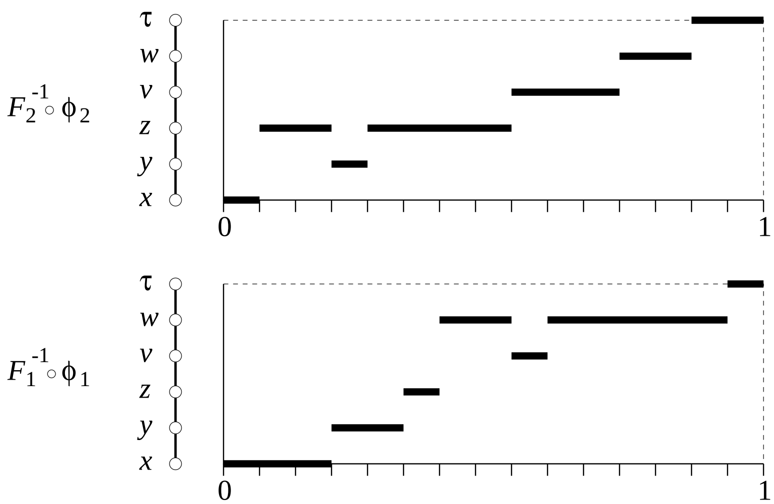


Figure 2. The synchronized inverse probability transforms $P_1^{-1} \circ \phi_1$ and $P_2^{-1} \circ \phi_2$

$[0, 1)$ to S for $k = 1$ and $k = 2$, as in Figure 2. It is clear from Figure 2 that $P_1^{-1}(\phi_1(t)) \leq_{\mathcal{S}} P_2^{-1}(\phi_2(t))$ for all $t \in [0, 1)$, as desired.

3. More on the monotonicity problem

3.1. Synchronizable posets

Given a poset $\mathcal{S}$ of Class W, Theorem 2.3 implies that if $\mathcal{A}$ is a poset having a minimum element and a maximum element, then monotonicity equivalence holds for $(\mathcal{A}, \mathcal{S})$. This section introduces without detail a further extension of monotonicity equivalence to a synchronizable poset $\mathcal{A}$ (as defined below).

Let $D_{\mathcal{A}}$ be the set of all the minimal elements in $\mathcal{A}$. Then we define a graph $(D_{\mathcal{A}}, \mathcal{I}_{\mathcal{A}})$ on the vertex set $D_{\mathcal{A}}$ by including $\{\alpha, \alpha'\}$ as an edge in

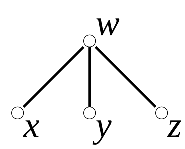
$\mathcal{I}_{\mathcal{A}}$ if $\alpha \neq \alpha'$ and there exists some $\beta \in A$ such that $\alpha, \alpha' < \beta$ in $\mathcal{A}$. We define in analogous fashion a graph $(D^*_{\mathcal{A}}, \mathcal{I}^*_{\mathcal{A}})$ on the set $D^*_{\mathcal{A}}$ of all the maximal elements in $\mathcal{A}$. We call these *graphs of interlacing relation*. Let $(D_{\mathcal{A}}, \mathcal{I}_0)$ be a spanning tree of $(D_{\mathcal{A}}, \mathcal{I}_{\mathcal{A}})$, that is, let $(D_{\mathcal{A}}, \mathcal{I}_0)$ be a tree with $\mathcal{I}_0 \subseteq \mathcal{I}_{\mathcal{A}}$. We will say that $(D_{\mathcal{A}}, \mathcal{I}_0)$ is a *locally connected spanning tree* of $(D_{\mathcal{A}}, \mathcal{I}_{\mathcal{A}})$ if for every $\alpha \in A$, the subgraph of $(D_{\mathcal{A}}, \mathcal{I}_0)$ induced by $D_{\mathcal{A}}(\alpha) := \{\beta \in D_{\mathcal{A}} : \beta \leq \alpha \text{ in } \mathcal{A}\}$ is connected. Finally, we call $\mathcal{A}$ a *synchronizable poset* if there exist respective locally connected spanning trees of $(D_{\mathcal{A}}, \mathcal{I}_{\mathcal{A}})$ and of $(D^*_{\mathcal{A}}, \mathcal{I}^*_{\mathcal{A}})$.

THEOREM 3.1. *If $\mathcal{S}$ is a poset of Class* W *and $\mathcal{A}$ is a synchronizable poset, then monotonicity equivalence holds for $(\mathcal{A}, \mathcal{S})$.*

Theorem 3.1 [which is Theorem 6.2 in Machida (1999)] is the most general positive result we know for the monotonicity equivalence problem when $\mathcal{S}$ is a poset of Class W.

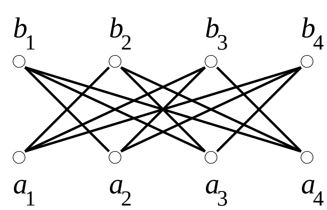
3.2. Open problem

Let

$\mathcal{S} :=$ 

,

which is a poset of Class W, and let

$\mathcal{A} :=$ 

.

Then the poset $\mathcal{A}$ is not synchronizable. However, we can show that monotonicity equivalence holds for $(\mathcal{A}, \mathcal{S})$ [Machida (1999)].

Theorem 3.1 has shown that synchronizability of the poset $\mathcal{A}$ is sufficient for monotonicity equivalence when $\mathcal{S}$ is a poset of Class W. But the above example disproves the assertion that synchronizability is necessary for monotonicity equivalence. Furthermore, let $\mathcal{M}(\mathcal{S})$ denote the class of all posets $\mathcal{A}$ of monotonicity equivalence for $\mathcal{S}$. Then we can also demonstrate [cf. Example 6.33 in Machida (1999)] that $\mathcal{M}(\mathcal{S})$ is *not* the same for all posets $\mathcal{S}$ of Class W. Thus, the interesting question raised but not settled by the present paper is how to completely characterize posets $\mathcal{A}$ of monotonicity equivalence given a poset $\mathcal{S}$ of Class W, that is, to determine $\mathcal{M}(\mathcal{S})$ exactly for each poset $\mathcal{S}$ of Class W.

Acknowledgments. The second author carried out research leading to this paper while he was a doctoral student in the Department of Mathematical Sciences at the Johns Hopkins University. We thank Keith Crank, Alan Goldman, Leslie Hall, and Edward Scheinerman for providing helpful comments.

References

Bogart, K. P. (1996). *Introductory Combinatorics.* Harcourt Brace Jovanovich, New York.

Fill, J. A. (1998). An interruptible algorithm for perfect sampling via Markov chains. *Ann. Appl. Probab.* **8** 131–162.

Fill, J. A. and Machida, M. (2000). Stochastic and realizable monotonicity. Preprint. Available from `http://www.mts.jhu.edu/~fill/`.

Kamae, T., Krengel, U., and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered state spaces. *Ann. Probab.* **5** 899–912.

Machida, M. (1999). Stochastic monotonicity and realizable monotonicity. Ph.D. dissertation, The Johns Hopkins University.

Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms* **9** 223–252.

Ross, D. A. (1993). A coherence theorem for ordered families of probability measures on a partially ordered space. Unpublished manuscript.

Stanley, R. P. (1986). *Enumerative Combinatorics*, Volume 1. Wadsworth & Brooks/Cole, Monterey, California.

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

Trotter, W. T. (1992). *Combinatorics and Partially Ordered Sets: Dimension Theory.* The Johns Hopkins University Press, Baltimore, Maryland.

West, D. B. (1996). *Introduction to Graph Theory.* Prentice Hall, New Jersey.