# NEGATIVE CORRELATION AND LOG-CONCAVITY

BY MICHAEL NEIMAN

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Jeff Kahn

and approved by

New Brunswick, New Jersey

May, 2009

**ABSTRACT OF THE DISSERTATION**

## Negative correlation and log-concavity

**by Michael Neiman**

**Dissertation Director: Jeff Kahn**

This thesis is concerned with negative correlation and log-concavity properties and relations between them, with much of our motivation provided by [40], [46], and [12]. Our main results include a proof that "almost exchangeable" measures satisfy the "Feder-Mihail" property; counterexamples and a few positive results related to several conjectures of Pemantle [40], Wagner [46], and Choe and Wagner [7] concerning negative correlation and log-concavity properties for probability measures and relations between them; a proof that a conditional version of the "antipodal pairs property" implies a strong form of log-concavity, which yields some partial results on a well-known conjecture of Mason [38]; a proof that "competing urn" measures satisfy "conditional negative association"; and proofs that certain classes of measures introduced by Srinivasan [42] and Pemantle [40] satisfy a strong form of negative association.

## Acknowledgements

I thank my thesis advisor, Jeff Kahn, for introducing me to many interesting correlation questions, for several years of guidance and encouragement, for collaborating with me on the work reported here, and for carefully reading several drafts of this thesis. I thank Daniel Ocone, Michael Saks, and Paul Seymour for their time and thoughts in their service on my thesis committee. I thank David Wagner for helpful discussions and some pointers to the literature. I also thank József Beck, Rob van den Berg, Stephen Greenfield, Van Vu, and Doron Zeilberger. Finally, I thank my parents, Leon and Linda, and my fiancée, Tiffany, for their constant love and support.

# Table of Contents

# Chapter 1

# Introduction

In this chapter, we state our main results, provide a bit of context, background, and motivation, and mention several open problems. Since this subject seems to generate a lot of acronyms and terminology, a short glossary is included at the end of this thesis.

We begin with some terminology. Given a finite set $S$, denote by $\mathcal{M} = \mathcal{M}_S$ the set of probability measures on $\Omega = \Omega_S = \{0,1\}^S$. As a default we take $S = [n] = \{1,\ldots,n\}$ (which for us is simply a generic $n$-set), using $\Omega_n$ in place of $\Omega_{[n]}$. We will occasionally identify $\Omega$ with the Boolean algebra $2^{[n]}$ (the collection of subsets of $[n]$ ordered by inclusion) in the natural way (namely, identifying a set with its indicator). An event $\mathcal{A} \subseteq \Omega$ is *increasing* (really, nondecreasing) if $x \geq y \in \mathcal{A}$ implies $x \in \mathcal{A}$ (where we give $\Omega$ the product order), and similarly for *decreasing*. While our concern here is with negative dependence properties, for perspective we first recall one or two points regarding their better understood positive counterparts.

## 1.1 Positive correlation and association

Events $\mathcal{A}$ and $\mathcal{B}$ in a probability space are *positively correlated*—we write $\mathcal{A} \uparrow \mathcal{B}$—if $\Pr(\mathcal{AB}) \geq \Pr(\mathcal{A})\Pr(\mathcal{B})$. The joint distribution of random variables $X_1, ..., X_n$—here always $\{0,1\}$-valued—is said to be *positively associated* (PA) if any two events both increasing in the $X_i$'s are positively correlated. (This is easily seen to be equivalent to the property that for any two increasing functions $f, g$ of the $X_i$'s one has $\mathsf{E}fg \geq \mathsf{E}f\mathsf{E}g$.)

The seminal result here is *Harris' Inequality* [25], which says that product measures are PA. (The special case of uniform measure on $\Omega$ was rediscovered in [30], and in combinatorial circles has often been called *Kleitman's Lemma*.) The best known—and most useful—extension of Harris' Inequality is the *FKG Inequality* of Fortuin, Kasteleyn, and Ginibre [15], *viz.*

**Theorem 1.1** *If $\mu \in \mathcal{M}$ satisfies*

$$\mu(\eta)\mu(\tau) \leq \mu(\eta \wedge \tau)\mu(\eta \vee \tau) \qquad \forall\ \eta, \tau \in \Omega \tag{1.1}$$

*(where $\wedge, \vee$ denote meet and join in the product order on $\Omega$), then $\mu$ is PA.*

(Stronger still, and also very useful, is the *Ahlswede-Daykin* or *"Four Functions" Theorem* [1], whose statement we omit.)

The *positive lattice condition* (1.1) (a.k.a. the *FKG lattice condition* or *log supermodularity*) is equivalent to *conditional positive association*, the property that every measure obtained from $\mu$ by conditioning on the values of some of the variables is PA; this follows easily from Theorem 1.1 and is a good way to make sense of (1.1). One also says that $\mu$ with (1.1) is an *FKG measure.* See, e.g., [3], [18], [32], [17], [13], [14] for a small sample of applications of these notions in combinatorics, probability, statistical mechanics, statistics and computer science.

## 1.2 Negative association and related properties

While negative correlation has the obvious meaning ($\mu(\mathcal{AB}) \leq \mu(\mathcal{A})\mu(\mathcal{B})$, denoted $\mathcal{A} \downarrow \mathcal{B}$), *negative association* requires a little care (for instance, $\mathcal{A} \uparrow \mathcal{A}$ holds strictly for any $\mathcal{A}$ with $\mu(\mathcal{A}) \notin \{0, 1\}$). Say $i \in [n]$ *affects* event $\mathcal{A}$ if there are $\eta \in \mathcal{A}$ and $\tau \in \Omega \setminus \mathcal{A}$ with $\eta_j = \tau_j$ for all $j \neq i$, and write $\mathcal{A} \perp \mathcal{B}$ if no coordinate affects both $\mathcal{A}$ and $\mathcal{B}$. (Note that we sometimes use "variable" in place of "coordinate.") Then $\mu \in \mathcal{M}$ is *negatively associated* (or has *negative association*; we use "NA" for both) if $\mathcal{A} \downarrow \mathcal{B}$ whenever $\mathcal{A}, \mathcal{B}$ are increasing and $\mathcal{A} \perp \mathcal{B}$. We say $\mu$ has *negative correlations* (or *is NC*) if $\eta_i \downarrow \eta_j$ (that is, $\{\eta_i = 1\} \downarrow \{\eta_j = 1\}$) whenever $i \neq j$.

Negative association turns out to be a much subtler property than PA. Pemantle [40] proposes a number of questions regarding conditions related to NA, and possible implications among them; we sketch what we need from this, and refer to [40] for a more thorough discussion (and more motivation). The properties of interest for us are those obtained from NC and NA by requiring closure under either conditioning or imposition of external fields. We first define these operations.

Here *conditioning* always means fixing the values of some variables (and this specification is always assumed to have positive probability); thus a measure obtained from $\mu \in \mathcal{M}$ by conditioning is one of the form $\mu(\cdot|\eta_i = \xi_i \ \forall i \in I)$ (which we regard as a measure on $\Omega_{[n]\setminus I}$) for some $I \subseteq [n]$ and $\xi \in \{0,1\}^I$. (If we think of $\Omega$ as $2^{[n]}$, then conditioning amounts to restricting our measure to some interval $[J,K]$ of $2^{[n]}$ (and normalizing).)

For $W = (W_1, \ldots, W_n) \in \mathbb{R}_+^n$ and $\mu \in \mathcal{M}$, define $W \circ \mu \in \mathcal{M}$ by

$$W \circ \mu(\eta) \ \propto \ \mu(\eta) \prod W_i^{\eta_i} \tag{1.2}$$

(meaning, as usual, that the left side is the right side multiplied by the appropriate normalizing constant). Borrowing Ising terminology, one says that $W \circ \mu$ is obtained from $\mu$ *by imposing the external field* $W$ (though to make this specialize correctly to the Ising model, we should really take the "field" to be $h$ given by $h_i = \ln W_i$). It will be convenient to allow $W_i = \infty$, which we interpret as conditioning on $\{\eta_i = 1\}$; similarly, we interpret $W_i = 0$ as conditioning on $\{\eta_i = 0\}$.

A third standard operation is projection: the *projection* of $\mu$ on $J \subseteq [n]$ is the measure $\mu'$ on $\{0,1\}^J$ obtained by integrating out the variables of $[n] \setminus J$; that is,

$$\mu'(\xi) = \sum \{\mu(\eta) : \eta \in \Omega, \eta_i = \xi_i \ \forall i \in J\} \qquad (\xi \in \{0,1\}^J).$$

A basic motivation for much of [40] was the desire for a natural and robust notion (or notions) of negative dependence, one measure of naturalness (and also of usefulness) being invariance under some or all of the preceding operations (and a few others that we will not discuss here). This leads in particular to the following classes, which were alluded to above.

We say that $\mu \in \mathcal{M}$ is *conditionally negatively correlated* (CNC) if every measure obtained from $\mu$ by conditioning is NC, and NC+ if every measure obtained from $\mu$ by imposition of an external field is NC. *Conditional negative association* (CNA) and NA+ are defined analogously. Of course NC+ and NA+ imply CNC and CNA respectively.

(This would be true even if we didn't allow $W_i \in \{0,\infty\}$ in (1.2), since a limit of NC measures is again NC, and similarly for NA; but there are properties of interest—in particular the Feder-Mihail property below—for which things go a little more smoothly with the present convention.)

Note that Pemantle uses CNA+ where we use NA+, but it is easy to see that the two notions coincide. In general he uses "+" for closure under both projections and external fields, but for the properties we are considering, this collapses to the definitions above: it is easy to see that all of the properties NC, CNC, NC+, NA, CNA, NA+ are preserved by projections.

Following [7], [46], we will also sometimes use the term *Rayleigh* for NC+. (The reference is to Rayleigh's monotonicity law for electric networks; see the second paragraph following Conjecture 1.13 below or e.g. [11] or [7].) We should also say a little more about the relation between our usage and that of [40], for which we need the *negative lattice condition* (NLC) for $\mu \in \mathcal{M}$:

$$\mu(\eta)\mu(\tau) \geq \mu(\eta\wedge\tau)\mu(\eta\vee\tau) \qquad \forall \eta,\tau\in\Omega. \tag{1.3}$$

This is of course the analogue of (1.1), but turns out to be not nearly as useful, a crucial difference being that, unlike (1.1), it is not preserved by projections. Following [40], we say that $\mu$ has the *hereditary negative lattice condition* (h-NLC) if all projections of $\mu$ satisfy the NLC, and that $\mu$ is *h-NLC+* if every measure obtained from $\mu$ by imposition of an external field is h-NLC. It is not hard to see

**Proposition 1.2** (a) *The properties CNC and h-NLC are equivalent.*

(b) *The properties NC+ and h-NLC+ are equivalent.*

This has also been observed in [4] (see their Proposition 2.2 for (b) and Remark 2.2 for a statement equivalent to (a)), so we will not prove it here, but briefly: (a) clearly implies (b); h-NLC trivially implies CNC; and the reverse implication follows easily from the observation that the support of a CNC measure is convex (i.e. $\mu(\eta),\mu(\tau)>0$ implies $\mu(\sigma)>0$ whenever $\eta\leq\sigma\leq\tau$), proof of which is identical to that of [46, Theorem 4.2].)

Extremely interesting would be

**Conjecture 1.3 ([40])** (a) *The properties CNC and CNA are equivalent.*

(b) *The properties NC+ and NA+ are equivalent.*

See [40, Conjectures 2 and 3]. Note that in each case it's enough to show that the first named property implies NA. As shown in [40], CNA does not imply NC+; we will later (Theorem 1.21) see a "naturally occurring" example of this. See also Conjecture 1.20 below for one approach to proving Conjecture 1.3(b). A measure $\mu \in \mathcal{M}$ is *exchangeable* if it is invariant under permutations of the coordinates (that is, $\mu(\eta_{\sigma(1)}, \ldots, \eta_{\sigma(n)}) = \mu(\eta_1, \ldots, \eta_n)$ for any $\eta \in \Omega$ and permutation $\sigma$ of $[n]$), or, equivalently, if $\mu(\eta)$ depends only on $|\eta| := \sum \eta_i$. We say $\mu$ is *almost exchangeable* if it is invariant under permutations of some subset of $n-1$ of the variables.

Pemantle shows [40, Theorem 2.7] that for exchangeable measures the properties CNC, NC+, CNA and NA+ are equivalent, while [4] proves Conjecture 1.3 for almost exchangeable measures, that is,

**Theorem 1.4 ([4], Corollary 6.6)** *For almost exchangeable measures*
(a) *the properties CNC and CNA are equivalent, and*
(b) *the properties NC+ and NA+ are equivalent.*

In Chapter 2 we give quick proofs of both these results. (Note that, in contrast to the exchangeable case, CNA and NC+ are not equivalent for almost exchangeable measures; see Theorem 1.21 and Example 5.13.) It may be worth noting that, despite its apparent simplicity, the class of almost exchangeable measures is considerably richer than the class of exchangeable measures; in particular, the examples proving Theorem 1.6 below (and also those of [4]) are almost exchangeable.

## 1.3 Log-concavity

A sequence $a = (a_0, \ldots, a_n)$ of real numbers is *unimodal* if there is some $k \in \{0, \ldots, n\}$ for which $a_0 \le a_1 \le \cdots \le a_k \ge \cdots \ge a_n$, and is *log-concave* (LC) if $a_i^2 \ge a_{i-1}a_{i+1}$ for $1 \le i \le n-1$. Of course a nonnegative LC sequence with no internal zeros is unimodal (where "no internal zeros" means $\{i : a_i \neq 0\}$ is an interval). Following [40] we say that

$a$ (as above) is *ultra-log-concave* (ULC) if the sequence $(a_i/\binom{n}{i})_{i=0}^n$ is log-concave and has no internal zeros.

We also say that $\mu \in \mathcal{M}$ is ULC if its *rank sequence*, $(\mu(|\eta| = i))_{i=0}^n$, is ULC. We define "$\mu$ is LC" and "$\mu$ is unimodal" similarly, except that for the former we add the stipulation that the rank sequence has no internal zeros. (It would be convenient to also make this a requirement for "LC" for sequences, but we politely adhere to the standard definition.)

Pemantle shows [40, Theorem 2.7] that for exchangeable measures, ULC coincides with CNC, NC+, CNA and NA+. He conjectures (see his Conjecture 4) that each of the latter properties implies ULC for general $\mu$; more precisely, this is a set of four conjectures, the weakest of which is

**Conjecture 1.5** *NA+ implies ULC.*

(He also conjectures that NA implies ULC, but, as noted in [4], this is easily seen to be incorrect, even for exchangeable measures.)

One of the stronger versions of Pemantle's conjecture—that the Rayleigh property (i.e. NC+) implies ULC—was separately proposed by Wagner in [46], where it was called the "Big Conjecture." (The overlap seems due to the failure in [40], [46] to notice Proposition 1.2(b).) We will say more about Wagner's motivation below. In Section 3.1 we show

**Theorem 1.6** *Conjecture 1.5 is false; in fact NA+ does not even imply unimodality.*

The first part of this was discovered a little earlier by Borcea *et al.* [4]. The present examples are slightly smaller (12 variables as opposed to 20 for violation of ULC) and simpler, and also disprove more, as the example of [4] is LC.

The examples for Theorem 1.6 also turn out to disprove Conjectures 8 and 9 of [40]; again, the first of these is also disproved by the example of [4]. Statements of these conjectures are deferred to Section 3.1.

A more particular notion than ULC, from [45] and [7], is as follows. Say $\mu \in \mathcal{M}$ is BLC[$m$] if every measure gotten from $\mu$ by imposing an external field and then projecting onto a set of size at most $m$ is ULC (the acronym is for "binomial log-concavity"), and BLC if it is BLC[$m$] for all $m$. (In [7], BLC[$m$] is called LC[$m$].) Choe and Wagner [7, Theorem 4.8] show that the three properties NC+, BLC[2], and BLC[3] are equivalent. (Strictly speaking, [7] is confined to a smaller class of $\mu$'s, but the proof is valid in the present generality.) They ask whether NC+ implies BLC[4]. (Of course, since projections preserve NC+, Wagner's conjecture above would say that NC+ implies BLC.) In Section 3.2 we will show

**Theorem 1.7** *NC+ implies BLC[5]*

(so also BLC[4]), whereas the examples for Theorem 1.6 will show

$$\text{NC+ does not imply BLC[12].} \tag{1.4}$$

We don't know what happens between 5 and 12. Of course Theorem 1.7 is now less interesting than formerly, when it was thought to be a step in the direction of Conjecture 1.5.

## 1.4 The antipodal pairs property

For $\mu \in \mathcal{M}$ set

$$\alpha_i(\mu) = \binom{n}{i}^{-1} \sum \{\mu(\eta)\mu(\underline{1} - \eta) : \eta \in \Omega, |\eta| = i\} \tag{1.5}$$

(where $\underline{1} = (1, \ldots, 1)$). Say $\mu \in \mathcal{M}_{2k}$ has the *antipodal pairs property* (APP) if $\alpha_k(\mu) \geq \alpha_{k-1}(\mu)$, and that $\mu \in \mathcal{M}$ has the *conditional antipodal pairs property* (CAPP) if any measure obtained from $\mu$ by conditioning on the values of some $n - 2k$ variables (for some $k$) has the APP. (Note that these properties are not affected by strictly positive and finite external fields. That is, if $\mu$ has the APP then so does $W \circ \mu$ for all $W$ with strictly positive and finite entries, and similarly for the CAPP.) In Chapter 4 we prove

**Theorem 1.8** *For measures without internal zeros in their rank sequences, the CAPP implies ULC.*

(A somewhat more general version of Theorem 1.8 is stated and proved in Section 4.1.)

Theorem 1.8 is useful for establishing ULC in some settings; here we describe a few applications. (See also Section 1.5 and Theorem 5.11.) A first, easy consequence is improvement of some of the results of [45], for which we need to recall some terminology from that paper. Given a positive integer $k$ and positive real number $\lambda$, say $\mu$ satisfies $\lambda$-Ray$[k]$ if every measure $\nu$ gotten from $\mu$ by imposing an external field and then projecting onto a set $S$ of $2k$ variables satisfies

$$\sum\{\nu(\eta)\nu(\underline{1}-\eta) : \eta \in \Omega, |\eta| = k\} \geq \lambda \sum\{\nu(\eta)\nu(\underline{1}-\eta) : \eta \in \Omega, |\eta| = k-1\}. \quad (1.6)$$

With the notation of (1.5), the above condition is

$$\alpha_k(\nu) \geq \frac{\lambda k}{k+1}\alpha_{k-1}(\nu);$$

thus $(1+1/k)$-Ray$[k]$ says that each $\nu$ as above has the APP. (As observed by Wagner [45]—see his Proposition 4.6—$\lambda = (1+1/k)$ is an "especially natural strength for these conditions"; there as here, this is essentially because $1+1/k$ is the ratio of the numbers of summands on the two sides of (1.6).) Note also that 2-Ray[1] is precisely the Rayleigh property. Wagner proved

**Theorem 1.9 ([45], Theorem 4.3)** *If a measure satisfies* 2*-Ray*[1] *and* $(1+1/k)^2$*-Ray*$[k]$ *for all* $2 \leq k \leq m$*, then it is BLC*$[2m+1]$*.*

(In [45] this is stated only for uniform measure on the bases of a matroid, but the proof is valid for general measures in $\mathcal{M}$.) Theorem 1.8 easily implies the following strengthening (see Section 4.2).

**Corollary 1.10** *If a measure satisfies* $(1+1/k)$*-Ray*$[k]$ *for all* $1 \leq k \leq m$*, then it is BLC*$[2m+1]$*.*

Using Corollary 1.10 in place of Theorem 1.9 improves Corollary 4.5(b) and Theorem 5.2 of [45] by substituting BLC for the weaker property $\sqrt{\text{BLC}}$; see [45] for definitions and statements.

It is easy to see that if $\mu \in \mathcal{M}_S$, $\nu \in \mathcal{M}_T$ are NA+ then the product measure $\mu \times \nu$ (given by $\mu \times \nu(\xi,\eta) = \mu(\xi)\nu(\eta)$ for $(\xi,\eta) \in \{0,1\}^S \times \{0,1\}^T$)) is also NA+.

Note that the rank sequence of $\mu \times \nu$ is the convolution of the rank sequences for $\mu$ and $\nu$. One consequence of Conjecture 1.5 would have been that the convolution of two ULC sequences is ULC or, equivalently, that the product of two ULC measures is ULC. (The implication follows from a result of Pemantle [40, Theorem 2.7] stating that for exchangeable measures the properties NA+ and ULC coincide.) Surprisingly—given that the analogous statement for ordinary log-concavity is fairly trivial—preservation of ULC under convolution turns out not to be so obvious; it was conjectured by Pemantle [40] (motivated by the preceding considerations) and proved by Liggett:

**Theorem 1.11 ([33], Theorem 2)** *The convolution of two ULC sequences is ULC.*

In Section 4.3 we derive this from Theorem 1.8 and also discuss a potentially interesting strengthening of ULC for measures that is again preserved by products.

## 1.5 Mason's conjecture

Here we want to say a little about the motivation for Wagner's ("big") conjecture and mention a few related questions. For this discussion we regard a *matroid* as a collection $\mathcal{I}$ of independent sets, subsets of some ground set $E$. We will not go into matroid definitions; see e.g. [49] or [39]. Prototypes are the collection of (edge sets of) forests of a graph (with edge set $E$)—this is a *graphic* matroid—and (as it turns out, more generally) the collection of linearly independent subsets of some finite subset $E$ of some (not necessarily finite) vector space; for present purposes not too much is lost by thinking only of graphic matroids.

We are interested in the *independence numbers* of a matroid $\mathcal{I}$, that is, the numbers

$$a_k = a_k(\mathcal{I}) = |\{I \in \mathcal{I} : |I| = k\}| \quad k = 0, \ldots, n,$$

concerning which we have a celebrated conjecture of J. Mason [38]:

**Conjecture 1.12** *For any matroid $\mathcal{I}$ on a ground set of size $n$, the sequence of independence numbers $a = a(\mathcal{I}) = (a_0, \ldots, a_n)$ is ULC.*

(Note that $a$ will typically end with some 0's, and also that in the graphic case $n$ counts edges, not vertices.) Of course one can relax Conjecture 1.12 by asking for LC or unimodality in place of ULC. In fact unimodality, first suggested by Welsh [47], was the original conjecture in this direction, and even this, even for graphic matroids, remains open. (See [43] or [5] for much more on log-concavity in combinatorial settings.)

From the present viewpoint, Mason's Conjecture asks for ultra-log-concavity of uniform measure on $\mathcal{I}$ (regarded in the usual way as a subset of $\{0,1\}^E$). In case $\mathcal{I}$ is graphic such a measure is a *uniform spanning forest* (USF) measure ("spanning" because we think of a member of $\mathcal{I}$ as a subgraph that includes all vertices). These measures are also very interesting from a correlation standpoint; in particular we have

**Conjecture 1.13** *USF measures are Rayleigh.*

This natural guess was perhaps first proposed in [29] (which was circulated in the combinatorial community as early as 1993, but took a while to get to press). It is also, for example, Conjecture 5.11.2 in [46]. (The statement in [29] is (in present language) that USF measures are NC, but it's not hard to see that this is equivalent.)

As essentially shown by Brooks *et al.* [6], the analogue of Conjecture 1.13 for uniform measure on the spanning *trees* of a (finite) graph amounts to Rayleigh's monotonicity law for electric networks (again, see [11]). This was extended by Feder and Mihail ([14], to which we will return shortly) to say that such measures are in fact NA+ (more precisely, this is what their proof gives).

Let us call a measure obtained from a USF measure by imposition of an external field—equivalently, a measure $\mu$ on the spanning forests of some finite graph $G$ with, for some $W : E(G) \to \mathbb{R}_+$, $\mu(F) \propto \prod_{e\in F} W(e)$—a *weighted spanning forest (weighted SF)* measure, and define *weighted spanning tree (WST)* measures and *weighted matroid measures* (replace "forest" by "independent set") similarly. (We avoid "WSF" since it means *wired* SF; see e.g. [36].)

One should note that, while the intuition for Conjecture 1.13 may seem clear—presence of a given edge $e$ makes it easier for a second edge $f$ to complete a cycle—this may be misleading, since the same intuition applies to uniform measure on the independent sets of a general matroid, for which NC need *not* hold (as can be derived from an example of Seymour and Welsh [41]). Some evidence for Conjecture 1.13, and its analogue for spanning connected subgraphs, is given in [21]. Also worth mentioning here—though without definitions; see [20]—is the following far-reaching extension of Conjecture 1.13, which has been "in the air" for a while (e.g. [40], [19]).

**Conjecture 1.14** *Any random cluster measure with $q < 1$ is NA+.*

(Equivalently, such measures are NA.) Limiting cases include the aforementioned uniform measures on forests, spanning trees and connected subgraphs of a graph; again see [20]. Conjecture 1.14 with NC+ in place of NA+ is proved for series-parallel graphs (part of a more general matroid statement) in [46] (see Example 5.1 and Theorem 5.8(d)).

Of course Wagner's "Big Conjecture," if true, would have implied Mason's Conjecture for any class of matroids for which one could establish the Rayleigh property (meaning, of course, for uniform measure on independent sets). Conjecture 1.13 says that graphic matroids should be such a class, and Wagner [46, Conj. 5.11] suggests a sequence of strengthenings of this. (Mason's Conjecture also partly motivated Conjecture 1.13 in [29], though at the time the connection wasn't much more than a feeling that the issues underlying the two were similar.)

According to Theorem 1.8, Mason's conjecture would follow from

**Conjecture 1.15** *Uniform measure on the independent sets of a matroid has the CAPP.*

(See also the remark following Corollary 4.4 for a possible strengthening.) Of course here it's enough to show APP, since each conditional measure is just uniform measure on the independent sets of some minor.

Though probably not for lack of effort, progress on Mason's conjecture has been fairly modest. Dowling [10] proved that for each $\mathcal{I}$ the sequence $(a_0, \ldots, a_8)$ is LC; Mahoney [37] proved that for graphic matroids corresponding to *outerplanar* graphs, the full sequence of independence numbers is LC; and Hamidoune and Salaün [22] proved that for any matroid on a ground set of size $n$ the sequence $(a_i/\binom{n}{i})_{i=0}^{4}$ is LC, i.e. the sequence $(a_i)$ is "ULC up to 4".

Here we adapt one of Dowling's arguments to prove Conjecture 1.15 for small matroids:

**Theorem 1.16** *For every matroid on a ground set of size at most 11, uniform measure on independent sets has the CAPP.*

This is proved in Section 4.4. Combined with Theorem 1.8 (for (a)) or the more general Theorem 4.1 below (for (b)) it gives

**Theorem 1.17** (a) *Every matroid on a ground set of size at most 11 satisfies Conjecture 1.12.*
(b) *For any matroid on a ground set of size $n$ with independence numbers $a_i$, the sequence $(a_i/\binom{n}{i})_{i=0}^{6}$ is LC (a.k.a. the sequence $(a_i)$ is "ULC up to 6").*

## 1.6 Feder-Mihail

Say $\mu \in \mathcal{M}$ has the *Feder-Mihail property* (or *is FM*) if

$$\textit{for any increasing } \mathcal{A} \subseteq \Omega, \ \{\eta_i = 1\} \uparrow \mathcal{A} \textit{ for some } i \in [n],$$

and extend this to CFM and FM+ in the usual way. (Of course FM+ trivially implies CFM, but note that this implication requires that we explicitly include conditioning in our definition of "+" (i.e. we allow $W_i \in \{0, \infty\}$ in (1.2)), since there are situations where $W \circ \mu$ is FM for all $W$ with positive entries but $\mu(\cdot|\eta_1 = 1)$ is not FM (e.g. if $\mu$ is the product of a measure on $\{0,1\}$ and a non-FM measure on $\{0,1\}^{\{2,\ldots,n\}}$).) The following simple but powerful observation is essentially from [14], though given there only in a special case.

**Theorem 1.18** (a) *If $\mu \in \mathcal{M}$ is both CNC and CFM then it is CNA.*

(b) *If $\mu \in \mathcal{M}$ is both NC+ and FM+ then it is NA+.*

(A statement equivalent to (a) is proved in [40, Theorem 1.3], and (b) follows easily from (a).)

Given the power of Theorem 1.18, it would be useful to identify situations where the FM property holds. This is trivially the case for $\mu$ concentrated on a level (that is,

$\{\eta \in \Omega : |\eta| = k\}$ for some $k$; see e.g. [12, Corollary 3.2]), e.g. for WST measures (a key to [14]), and it is fairly easy to show that exchangeable measures and, more generally, "rescalings" of product measures, satisfy the stronger "normalized matching property" (NMP; see Chapter 2 for definitions). But in general FM seems hard to establish, and indeed we're not aware of any interesting classes of non-NMP measures that are known to be FM. Thus the following result, which is proved in Chapter 2, may be of some interest.

**Theorem 1.19** *Almost exchangeable measures are FM+.*

Note that this combined with Theorem 1.18 gives Theorem 1.4. (This is not quite the "quick" proof of Theorem 1.4 promised earlier, since Theorem 1.19 requires some effort; but, as observed in Chapter 2, FM (resp. FM+) for almost exchangeable measures that are also NC (resp. NC+) is much easier.)

Despite the (apparent) difficulty of proving FM, the property seems to tend to hold for measures not deliberately constructed to violate it. Thus we propose, perhaps optimistically,

**Conjecture 1.20** *The Feder-Mihail property holds for*

(a) *Rayleigh measures,*

(b) *weighted SF measures, and (more generally)*

(c) *weighted matroid measures.*

Note that, in view of Theorem 1.18, (a) would imply Conjecture 1.3(b) (the corresponding approach to Conjecture 1.3(a) fails because CNC measures need not be FM), while (b) together with Conjecture 1.13 would say that USF measures are NA+. (Extending this to matroids via (c) fails because Conjecture 1.13 does.)

## 1.7 Competing urns

One of the principal motivating examples for [40] is *competing urns*, which refers to the experiment in which $m$ balls are dropped, randomly and independently, into urns $1, \ldots, n$. Formally, we have a random $\sigma : [m] \to [n]$ with the $\sigma(i)$'s independent. We then take $\mathbf{x}_j$ to be the indicator for occupation of urn $j$ (i.e. $\mathbf{x}_j = \mathbf{1}_{\{\sigma^{-1}(j)\neq\emptyset\}}$) and are interested in the law, $\mu$, of $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ (a measure on $\{0,1\}^n$). In the traditional case where the balls are identical (i.e. the $\sigma(i)$'s are i.i.d.) we call $\mu$ an *urn measure*, or, for emphasis, an *ordinary* urn measure. More generally, setting $B_j = |\sigma^{-1}(j)|$, we may consider thresholds $t_1, \ldots, t_n$, and let $\mathbf{x}_j$ be the indicator of $\{B_j \geq t_j\}$; for i.i.d. balls, we then call the law of $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ a *threshold* urn measure. When the balls are not required to be identical we speak of *generalized urn measures* and *generalized threshold urn measures*.

The competing urns model was explored in some detail by Dubhashi and Ranjan [13][1], who proved *inter alia* that threshold urn measures are NA. Another proof of this is given in [40]. Actually the argument of [13]—which proves the stronger statement that the (law of the) random variables

$$\xi_{ij} = \mathbf{1}_{\{\sigma(i)=j\}} \tag{1.7}$$

is NA—does not require identical balls. (The argument of [40] does not work for nonidentical balls.) In Chapter 5 we prove (a slight generalization of)

**Theorem 1.21** *Threshold urn measures are CNA.*

In contrast, it's not hard to give examples (see Example 5.13) showing that even ordinary urn measures need not be Rayleigh. Thus, as mentioned earlier, we have a natural class of measures for which CNA does not imply Rayleigh.

The question of whether Theorem 1.21 extends to nonidentical balls seems very interesting, and at this point we don't even have an opinion on what the answer should be. As far as we can see, even the following very general statement could be true.

**Question 1.22** *Suppose $T_0 \cup T_1 \cup \cdots \cup T_s$ is a partition of $[m] \times [n]$, and $a_r, b_r \in \mathbb{N}$ for $r = 1, \ldots, s$. Is it necessarily true that the $\xi_{ij}$'s in (1.7) are NA given*

$$\{\xi(T_r) \in [a_r, b_r] \ \forall r \in [s]\}$$

*(where $\xi(T) = \sum_{(i,j)\in T} \xi_{ij}$)?*

---

[1] They say "bins" rather than "urns."

This would be a considerable strengthening of CNA for generalized threshold urn measures.

That the weaker (than CNA) CNC, at least, does hold for generalized threshold urn measures is a special case of the following result, a somewhat more general version of Corollary 34 of [13], which implies CNC for generalized threshold urn measures. For $\mathcal{A} \subseteq 2^{[m]}$ and $a, b \in \mathbb{N}^{n-1}$, set $p(\mathcal{A}, a, b) = \Pr(\sigma^{-1}(n) \in \mathcal{A} | B_j \in [a_j, b_j] \ \forall j \in [n-1])$.

**Theorem 1.23** *For any increasing $\mathcal{A}$ (and any generalized threshold urn measure), $p(\mathcal{A}, a, b)$ is decreasing in $(a, b)$; that is, $p(\mathcal{A}, a, b) \leq p(\mathcal{A}, r, s)$ whenever $a \geq r$ and $b \geq s$ (in the product order on $\mathbb{N}^{n-1}$).*

(The proof of Corollary 34 in [13] is not quite correct, since it depends on the incorrect Proposition 24. Theorem 1.23 is proved at the end of Section 5.2.) Thus one reason to be interested in whether generalized urn measures are CNA is that a negative answer would provide a counterexample to Conjecture 1.3(a).

Pemantle [40] suggests

**Conjecture 1.24** *Every ordinary urn measure is ULC.*

In fact he conjectures something more general that we will not state, since unfortunately

**Proposition 1.25** *Conjecture* 1.24 *is not true; ordinary urn measures need not be LC.*

(See Example 5.12.) Proposition 1.25 and Theorem 1.21 provide a (natural) counterexample to the strengthening of Conjecture 1.5 obtained by replacing NA+ by CNA (which is again one of the versions of Conjecture 4 of [40]).

Finally we turn to some conjectures of Farr (unpublished, circa 2004) and Welsh [48]. To put these in our framework, we add an urn $\Lambda$ and assume

$$\Pr(\sigma(i) = j) = q \quad \forall i \in [m], j \in [n].$$

(So $\Pr(\sigma(i) = \Lambda) = 1 - nq$.) Let $\mathcal{I} \subseteq 2^{[m]}$ be decreasing and set $\mathcal{A}_j = \{\sigma^{-1}(j) \in \mathcal{I}\}$ and $\mathcal{A}_J = \bigcap\{\mathcal{A}_j : j \in J\}$. Then Farr's conjecture (somewhat rephrased) is

**Conjecture 1.26** *If $G$ is a graph on $[m]$ and $\mathcal{I}$ is the collection of independent sets of $G$, then for any disjoint $I, J, K \subseteq [n]$, $\mathcal{A}_I \downarrow \mathcal{A}_J$ given $\mathcal{A}_K$.*

It's not clear why this should require that $\mathcal{I}$ be of the type described, and Welsh's conjecture was that the same conclusion holds for an arbitrary $\mathcal{I}$, *viz.*

**Conjecture 1.27 ([48])** *For any decreasing $\mathcal{I} \subseteq 2^{[m]}$ and disjoint $I, J, K \subseteq [n]$, $\mathcal{A}_I \downarrow \mathcal{A}_J$ given $\mathcal{A}_K$.*

We disprove this stronger version (see Example 5.14). At present we don't see how to extend to a counterexample to Conjecture 1.26, though we feel that this too is likely to be false.

## 1.8 Srinivasan's sampling process and Pemantle measures

Srinivasan [42] introduced a procedure for generating a random $\eta \in \Omega$ concentrated on a level, having prescribed marginal probabilities, and satisfying certain negative dependence properties; in particular, he proved that for such measures

$$\Pr(\eta_i = 1 \ \forall\ i \in I) \leq \prod_{i \in I} \Pr(\eta_i = 1)$$

for every $I \subseteq [n]$ (which is a special case of NA). Srinivasan used these measures to create a new type of randomized rounding scheme and give improved approximation algorithms for several NP-hard problems; see [42], [16] for details.

For our purposes (see the remark following Proposition 6.5) we will need something a little more general than Srinivasan's procedure. We first describe this generalization. A *pairing tree* for $I \subseteq [n]$ is a rooted binary tree $T$ with set of leaves $I$ and some additional structure, as follows. (N.B. this differs somewhat from the "pairing tree" in [42].) The two children of each internal (i.e. non-leaf) vertex are distinguished: one is the left child, and the other the right child. (As usual, one may think of a plane drawing of the tree.) Each internal vertex $u$ is assigned parameters $t_u \in \{0, 1\}$ and $\beta_u \in [0, 1]$, and there is one additional parameter $\alpha = \alpha(T) \in [0, 1]$. (Srinivasan's procedure corresponds to the special case that $\alpha \in \{0, 1\}$; see Proposition 6.5 and the discussion preceding it.)

Each pairing tree for $[n]$ yields a measure in $\mathcal{M}$, the law of $\eta \in \Omega$ generated by the following procedure. (In our generalized setting, $\eta$ will be concentrated on two consecutive levels of $\Omega$, rather than a single level, whenever $\alpha \in (0,1)$.) The values $\eta_i$ ($\in \{0,1\}$) are set sequentially. Assume that we have already fixed the values of entries indexed by coordinates in $J$ (initially $J = \emptyset$), and that we have inherited from the previous step a pairing tree for $I := [n] \setminus J$. Pick an internal vertex $u$, both of whose children are leaves, to "process" at the current step; say the left child is $i$ and the right child is $j$. (We use common terminology for rooted trees: $v$ is a *descendant* of $w$ if the (unique) path from $v$ to the root contains $w$, and a *child* of $w$ if in addition $w$ is adjacent to $v$. Also, for a vertex $v$, $T_v$ is the subtree of $T$ rooted at $v$, where descendants of $v$ retain their $t, \beta$ parameters.) With probability $\beta_u$ fix $\eta_i = t_u$, and otherwise (so with probability $1-\beta_u$) fix $\eta_j = t_u$; let $k \in \{i,j\}$ be the coordinate fixed at this step. Create a pairing tree for $I \setminus \{k\}$ from the one for $I$ by removing leaves $i, j$ and relabeling vertex $u$ (now a leaf) by whichever of $i, j$ has not been fixed. (The remaining parameters ($t_w$, $\beta_w$ for internal vertices $w$ and $\alpha$) are unchanged.)

We continue this process until all but one coordinate, say $i$, have been fixed, at which point the current pairing tree consists of a single vertex labeled by the unfixed coordinate. At this point we make use of $\alpha$ to fix the value of $\eta_i$, setting $\eta_i$ to 1 with probability $\alpha$ (so to 0 with probability $1-\alpha$).

For consistency with [12] (and for lack of a better name), we call this procedure a super-generalized Srinivasan sampling process (SGSSP), and the resulting measure an SGSSP measure. In [12], the special case that $\alpha \in \{0,1\}$ (i.e. the measure is concentrated on a single level) is called a generalized Srinivasan sampling process (GSSP) measure. (The measures initially introduced by Srinivasan [42], called Srinivasan sampling process (SSP) measures in [12], appear to be a subclass of the GSSP measures; however, as we will see below (Theorem 6.5), the two classes coincide.)

Dubhashi *et al.* [12] conjectured that GSSP measures satisfy strong negative dependence conditions, *viz.*

**Conjecture 1.28** *Every GSSP measure is CNA,* and proved it in the special case that the initial pairing tree has the property that every internal vertex has at least one leaf as a child (i.e. the internal vertices are simply a path from the root). In Chapter 6 we prove Conjecture 1.28 and a bit more:

**Theorem 1.29** *Every SGSSP measure is NA+.*

Pemantle defined and proved some negative dependence properties for a class of measures (defined shortly) that, as shown here (see Lemma 6.1 and Corollary 6.3), includes all SGSSP measures. Recall that the product of $\mu \in \mathcal{M}_S$ and $\nu \in \mathcal{M}_T$ (with $S \cap T = \emptyset$) is $\mu \times \nu \in \mathcal{M}_{S \cup T}$ with $\mu \times \nu(\eta) = \mu(\eta|_S)\nu(\eta|_T)$. Given a nonnegative LC sequence $q = (q_0, \ldots, q_n)$ with no internal zeros, we follow [40] and define the *rank rescaling of* $\mu \in \mathcal{M}$ *by* $q$ to be the measure $q \otimes \mu \in \mathcal{M}$ with

$$q \otimes \mu(\eta) \propto q_{|\eta|}\mu(\eta).$$

(To be precise, we only make this definition when the right side is not identically zero.) Let $\mathcal{P}$ be the smallest class of measures containing all Bernoulli measures (i.e. measures on $\{0, 1\}$) and closed under imposition of external fields, products, and rank rescaling; since they were introduced in [40], we call measures in $\mathcal{P}$ *Pemantle measures*. In Chapter 6 we give an improvement of an earlier result of Pemantle [40]:

**Theorem 1.30** *Pemantle measures are NA+.*

(Following [40], say $\mu \in \mathcal{M}$ is *jointly negative regression dependent* (JNRD) if for every measure gotten from $\mu$ by conditioning on the values of some of the coordinates we have $\mathcal{A} \downarrow \{\eta_i = 1\}$ whenever $\mathcal{A}$ is an increasing event that is not affected by coordinate $i$, and extend to JNRD+ in the usual way; it is easy to see that CNA implies JNRD and JNRD implies CNC, but we don't know if the reverse implications hold (cf. Conjecture 1.3). Pemantle [40, Theorem 3.1] proved that Pemantle measures are JNRD+.) The core of our proof of Theorem 1.30 is a result stating that several properties of measures are preserved by products (Theorem 6.6).

Of course, in view of the aforementioned Corollary 6.3, Theorem 1.30 contains Theorem 1.29; but, as we will see in Section 6.1, there is a simple derivation of Theorem 1.29 from Lemma 6.1 that does not require Theorem 1.30.

# Chapter 2

# Exchangeable and almost exchangeable measures

In this chapter we examine the Feder-Mihail property and some negative dependence properties for exchangeable and almost exchangeable measures; in particular, we prove Theorem 1.19 and give short proofs of Theorems 1.4 and 2.2.

We begin with a few more definitions. We extend the definitions of exchangeable and almost exchangeable measures (given following Conjecture 1.3) to general functions on $\Omega$ in the obvious way ($f : \Omega \to \mathbb{R}$ is *almost exchangeable* if it is invariant under permutations of some subset of $n-1$ of the variables and *exchangeable* if it is invariant under permutations of all the variables). We also extend our notation for positive and negative correlation to functions: for $f, g : \Omega \to \mathbb{R}$, we write $f \uparrow g$ if $\mathsf{E}fg \geq \mathsf{E}f\mathsf{E}g$ (and similarly for $f \downarrow g$); we will also write, e.g., $\mathcal{A} \uparrow f$ for $\mathbf{1}_{\mathcal{A}} \uparrow f$. A stronger statement is that $\mathcal{A}$ is *stochastically increasing in* $f$, that is, that $\Pr(\mathcal{A}|f = t)$ is increasing in $t$, where we restrict to values of $t$ for which $\Pr(f = t)$ is positive. (N.B. our use of the notation $\mathcal{A} \uparrow f$ differs from that in [40].) Following [12], for a function $f : \Omega \to \mathbb{R}$ and measure $\mu \in \mathcal{M}$, we say $i \in [n]$ is a *variable of positive influence* for the pair $(f, \mu)$ (or $(\mathcal{A}, \mu)$ if $f = \mathbf{1}_{\mathcal{A}}$) if $\eta_i \uparrow f$. Thus the FM property for $\mu$ says that for every increasing $\mathcal{A}$ there is a variable of positive influence for $(\mathcal{A}, \mu)$. In [12], Dubhashi *et al.* prove the (easy) result that $(f, \mu)$ has a variable of positive influence if $f$ is increasing and at least one of $f$, $\mu$ is exchangeable, and ask for other classes of function-measure pairs having variables of positive influence. Here we show

**Theorem 2.1** *If there is a variable $l$ for which*

$$\mathsf{E}[f|\eta_l = j, \sum_{i \neq l} \eta_i = k] \text{ is increasing in } j \text{ and } k \tag{2.1}$$

*(for pairs $(j, k)$ for which the conditioning event has positive probability, and where*

$\mathsf{E}$ *is expectation with respect to* $\mu$*), then* $(f, \mu)$ *has a variable of positive influence. In particular, if* $f$ *is increasing and almost exchangeable then* $(f, \mu)$ *has a variable of positive influence.*

As we will see shortly, this implies Theorem 1.19.

As observed earlier, Theorem 1.4 is an immediate consequence of Theorems 1.18 and 1.19; but it does not require the full strength of Theorem 1.19, and before proving Theorem 2.1 we will give an easier argument, together with a quick proof of

**Theorem 2.2 ([40], Theorem 2.7)** *For exchangeable measures the properties CNC, CNA, NC+, NA+, and ULC are equivalent.*

For these arguments and the derivation of Theorem 1.19 we need a little background. We first recall Chebyshev's Inequality, which in our terminology says

**Proposition 2.3** *Any probability measure on a totally ordered set is PA*

(where, of course, PA is as for measures on $\{0,1\}^S$: $f \uparrow g$ for any two increasing functions $f, g$).

For probability measures $\mu$ and $\nu$, $\mu$ *stochastically dominates* $\nu$ (written $\mu \succeq \nu$) if $\mu(\mathcal{A}) \geq \nu(\mathcal{A})$ for every increasing event $\mathcal{A}$. Writing $\mu_k$ for the conditional measure $\mu(\cdot|\sum \eta_i = k)$ (defined only when $\mu(\sum \eta_i = k) > 0$), we say $\mu$ has the *normalized matching property* (NMP) if $\mu_l \succeq \mu_k$ whenever $l \geq k$ and both conditional measures are defined. (This generalizes the usual definition, for which see e.g. [2].) The NMP is equivalent to the property that every increasing event $\mathcal{A}$ is stochastically increasing in $\sum \eta_i$, which implies (easily and directly, by Proposition 2.3, or essentially by Proposition 1.2 in [40]) that $\mathcal{A} \uparrow \sum \eta_i$, and thus (since expectation is linear) that $\mathcal{A} \uparrow \eta_i$ for some $i$. This proves

**Proposition 2.4** *The NMP implies FM.*

Conjecture 8 of [40] says NA+ implies NMP, but this is false; see Conjecture 3.1 and Theorem 3.3 in Section 3.1.

Given $\mu \in \mathcal{M}$ and a nonnegative sequence $a = (a_i)_{i=0}^n$, the *generalized rank rescaling of* $\mu$ *by* $a$ is the measure $a \otimes \mu \in \mathcal{M}$ with

$$a \otimes \mu(\eta) \propto a_{|\eta|}\mu(\eta).$$

(Again, we only make this definition when the right side is not identically zero.) This generalizes the rank rescaling operation defined in Section 1.8, which required that $a$ be LC with no internal zeros. Observe that (since $\mu_k = (a \otimes \mu)_k$ whenever $a_k > 0$) generalized rank rescalings preserve the NMP.

**Lemma 2.5** *Product measures (and, consequently, generalized rank rescalings of product measures) have the NMP.*

(A proof is sketched in [12, Section 4.2]. See also Theorem 6.6(b) for a more general result.) For the proof of Theorem 1.4 we also need the following standard observation, which is an easy consequence of Proposition 2.3.

**Lemma 2.6** *Let* $f, g : \Omega \to \mathbb{R}$*, and suppose for some event* $\mathcal{B}$

(i) *each of* $f$*,* $g$ *is positively correlated with* $\mathcal{B}$*, and*

(ii) $f$ *and* $g$ *are conditionally positively correlated given each of* $\mathcal{B}$*,* $\Omega \setminus \mathcal{B}$*.*

*Then* $f$ *and* $g$ *are positively correlated.*

*Proof of Theorem 1.19.* Let $\mu' \in \mathcal{M}$ be invariant under permutations of the variables $1, \ldots, n-1$, and $\mu = W \circ \mu'$ for some $W \in \mathbb{R}_+^n$. We may assume all $W_i$ are finite and strictly positive, since otherwise we can reduce the number of variables (note that any measure gotten from an almost exchangeable measure by conditioning is again almost exchangeable). Then, with $\nu \in \mathcal{M}_{n-1}$ the product measure satisfying

$$\nu(\eta) \propto \prod_{i \in [n-1]} W_i^{\eta_i},$$

we have

$$\mu(\cdot | \eta_n = 0, \sum_{i \in [n-1]} \eta_i = k) = \mu(\cdot | \eta_n = 1, \sum_{i \in [n-1]} \eta_i = k) = \nu_k,$$

so (by Lemma 2.5) $f = \mathbf{1}_{\mathcal{A}}$ satisfies (2.1) with $l = n$ for every increasing event $\mathcal{A}$. $\square$

*Proof of Theorem 1.4.* By Theorem 1.18, it suffices to show that *every NC measure that can be gotten by applying an external field to an almost exchangeable measure is FM.* Let $\mu \in \mathcal{M}$ be such a measure, obtained by imposing an external field on a measure invariant under permutations of coordinates $\{2, \ldots, n\}$, and let $\mathcal{A} \subseteq \Omega$ be increasing. We should show that $\mathcal{A} \uparrow \eta_i$ for some $i \in [n]$. We may assume that all coordinates of the external field are finite and strictly positive (or we can reduce the number of variables), and that $\mathcal{A} \downarrow \eta_1$ (or we are done). For $j \in \{0, 1\}$, the conditional measure $\mu(\cdot|\eta_1 = j)$ is a generalized rank rescaling of a product measure, so by Lemma 2.5 and the discussion in the paragraph preceding Proposition 2.4 we have $\mathcal{A} \uparrow f := \sum_{i \neq 1} \eta_i$ conditionally given either of the events $\{\eta_1 = 0\}$, $\{\eta_1 = 1\}$. Since $\mu$ is NC, we have $\eta_1 \downarrow \eta_i$ for all $i \in \{2, \ldots, n\}$, so that $\eta_1 \downarrow f$. But then applying Lemma 2.6 with $g = \mathbf{1}_{\mathcal{A}}$ and $\mathcal{B} = \{\eta_1 = 0\}$ gives $\mathcal{A} \uparrow f$, whence $\mathcal{A} \uparrow \eta_i$ for some $i \in \{2, \ldots, n\}$. $\square$

*Proof of Theorem 2.2.* It suffices to show CNC implies ULC and ULC implies NA+. The first implication is easy: CNC implies NLC (cf. Proposition 1.2), which for exchangeable measures is equivalent to ULC. For the second implication, our main point is that we can eliminate much of the work in [40] by observing that exchangeable measures are FM+ (by Lemma 2.5 and Proposition 2.4; of course this is also an instance of Theorem 1.19, but not one that requires the less trivial Theorem 2.1), so that by Theorem 1.18(b) it is enough to show that ULC implies NC+. This is a special case of the observation, proved in Lemma 2.8 of [40], that a measure obtained from an exchangeable ULC measure by imposing an external field that is identically 1 on $J \subseteq [n]$, followed by projection on $J$, is exchangeable and ULC. $\square$

*Proof of Theorem 2.1.* First observe that if $f : \Omega_n \to \mathbb{R}$ is increasing and invariant under permutations of coordinates in $[n] \setminus \{l\}$ then (2.1) is satisfied for every $\mu \in \mathcal{M}$, so the last part of Theorem 2.1 follows from the first.

To prove the first part, suppose, without loss of generality, that $\mu \in \mathcal{M}$ and $f : \Omega_n \to \mathbb{R}$ satisfy (2.1) with $l = 1$, and set $h(\eta) = \sum_{i \neq 1} \eta_i$ ($\eta \in \Omega_n$). It suffices to show that either

$$f \uparrow \eta_1 \quad \text{or} \quad f \uparrow h. \tag{2.2}$$

For $i \in \{0, \ldots, n-1\}$, let

$$\alpha_i = \mu(h = i, \eta_1 = 1) \quad \text{and} \quad \beta_i = \mu(h = i, \eta_1 = 0).$$

Choose increasing, nonnegative sequences $\gamma = (\gamma_0, \ldots, \gamma_{n-1})$ and $\delta = (\delta_0, \ldots, \delta_{n-1})$ such that

$$\gamma_i = \mathsf{E}[f|h = i, \eta_1 = 1] \quad \text{and} \quad \delta_i = \mathsf{E}[f|h = i, \eta_1 = 0] \tag{2.3}$$

whenever the conditioning events have positive probability and $\gamma_i \geq \delta_j$ whenever $i \geq j$. (Existence of $\gamma, \delta$ is guaranteed by (2.1). This extension to values not given by (2.3) is convenient, but not really necessary, as these values play no role; see (2.4) and Lemma 2.7.)

Assume $f \downarrow \eta_1$, i.e.,

$$\frac{\sum \alpha_i \gamma_i}{\sum \alpha_i} \leq \frac{\sum \beta_i \delta_i}{\sum \beta_i}$$

(all sums in this proof are over $\{0, \ldots, n-1\}$ unless otherwise specified). We want to show $\mathsf{E}fh \geq \mathsf{E}f\mathsf{E}h$, that is,

$$\sum(\alpha_i\gamma_i + \beta_i\delta_i)\sum i(\alpha_i + \beta_i) \leq \sum i(\alpha_i\gamma_i + \beta_i\delta_i)\sum(\alpha_i + \beta_i). \tag{2.4}$$

(Of course the last sum is 1.) This will follow from

$$\sum \alpha_i\gamma_i \sum i\alpha_i \leq \sum i\alpha_i\gamma_i \sum \alpha_i,$$

$$\sum \beta_i\delta_i \sum i\beta_i \leq \sum i\beta_i\delta_i \sum \beta_i,$$

and

$$\sum i\alpha_i \sum \beta_i\delta_i + \sum i\beta_i \sum \alpha_i\gamma_i \leq \sum i\alpha_i\gamma_i \sum \beta_i + \sum i\beta_i\delta_i \sum \alpha_i. \tag{2.5}$$

The first two of these are instances of Proposition 2.3 (since $\gamma$ and $\delta$ are increasing), so we only need

**Lemma 2.7** *Let $\alpha = (\alpha_i)_{i=0}^{n-1}$, $\beta = (\beta_i)_{i=0}^{n-1}$, $\gamma = (\gamma_i)_{i=0}^{n-1}$, and $\delta = (\delta_i)_{i=0}^{n-1}$ be nonnegative sequences (with neither of $\alpha$, $\beta$ identically zero). If $\gamma$ and $\delta$ are increasing, $\gamma_i \geq \delta_j$ whenever $i \geq j$, and $(\sum \alpha_i\gamma_i)/(\sum \alpha_i) \leq (\sum \beta_i\delta_i)/(\sum \beta_i)$, then (2.5) holds.*

*Proof.* Since scaling $\alpha$, $\beta$ affects neither our hypotheses nor (2.5), we may assume $\sum \alpha_i = \sum \beta_i$. It suffices to show

$$\sum_{i\geq s}\alpha_i\sum\beta_i\delta_i+\sum_{i\geq s}\beta_i\sum\alpha_i\gamma_i\leq\sum_{i\geq s}\alpha_i\gamma_i\sum\beta_i+\sum_{i\geq s}\beta_i\delta_i\sum\alpha_i \tag{2.6}$$

for $s \in [n-1]$ (since summing (2.6) over $s$ yields (2.5)).

Fix $s \in [n-1]$. Obviously,

$$\text{if (2.6) is true, then it remains true when any } \delta_i \text{ with } i \geq s \text{ is increased.} \tag{2.7}$$

We define $\delta' = (\delta'_i)_{i=0}^{n-1}$ by

$$\delta'_i = \begin{cases} \delta_i & \text{if } i < s \\ \delta_s & \text{if } i \geq s \end{cases}$$

and consider two cases.

<u>Case 1</u>: $\sum \alpha_i\gamma_i > \sum \beta_i\delta'_i$. Then there is an increasing sequence $\delta'' = (\delta''_i)_{i=0}^{n-1}$ with $\delta'_i \leq \delta''_i \leq \delta_i$ for all $i$ and $\sum \alpha_i\gamma_i = \sum \beta_i\delta''_i$ (note $\sum \alpha_i\gamma_i \leq \sum \beta_i\delta_i$, since we've normalized to $\sum \alpha_i = \sum \beta_i$). Since $\gamma$ and $\delta''$ are increasing, we have

$$\sum_{i\geq s}\alpha_i\sum\alpha_i\gamma_i\leq\sum_{i\geq s}\alpha_i\gamma_i\sum\alpha_i$$

and

$$\sum_{i\geq s}\beta_i\sum\beta_i\delta''_i\leq\sum_{i\geq s}\beta_i\delta''_i\sum\beta_i.$$

This yields (2.6) with $\delta$ replaced by $\delta''$, and then (2.6) (for $\delta$) follows from (2.7).

<u>Case 2</u>: $\sum \alpha_i\gamma_i \leq \sum \beta_i\delta'_i$. By (2.7), it suffices to prove (2.6) with $\delta$ replaced by $\delta'$; this is a straightforward computation:

$$\begin{aligned}
\sum_{i\geq s}\alpha_i\sum\beta_i\delta'_i+\sum_{i\geq s}\beta_i\sum\alpha_i\gamma_i &\leq \sum_{i\geq s}(\alpha_i+\beta_i)\sum\beta_j\delta'_j \\
&= \sum_{i\geq s}\sum_j(\alpha_i\beta_j\delta'_j+\beta_i\beta_j\delta'_j) \\
&\leq \sum_{i\geq s}\sum_j(\alpha_i\beta_j\delta'_i+\beta_i\beta_j\delta'_i) \\
&= \sum_{i\geq s}(\alpha_i\delta'_i+\beta_i\delta'_i)\sum\beta_j \\
&\leq \sum_{i\geq s}\alpha_i\gamma_i\sum\beta_i+\sum_{i\geq s}\beta_i\delta'_i\sum\alpha_i,
\end{aligned}$$

where we used: $\sum \alpha_i \gamma_i \leq \sum \beta_i \delta'_i$ for the first inequality; $\delta'_i \geq \delta'_j$ $\forall i \geq s$ (and $\forall j$) for the second; and $\gamma_i \geq \delta'_i$ $\forall i$ for the third. This completes the proofs of Lemma 2.7 and Theorem 2.1. $\square$

# Chapter 3

# The Rayleigh property and ultra-log-concavity

In this chapter we prove Theorem 1.6, disprove two other conjectures from [40], and (in Section 3.2) prove Theorem 1.7.

## 3.1 Counterexamples

Here we give the construction for Theorem 1.6. As mentioned earlier, two further conjectures from [40] turn out to be disproved by the same examples, and we begin by stating these.

**Conjecture 3.1 ([40], Conjecture 8)** *NA+ implies the NMP.*

(Recall from Chapter 2 that $\mu \in \mathcal{M}$ has the NMP if $\mu(\cdot|\sum \eta_i = k)$ stochastically dominates $\mu(\cdot|\sum \eta_i = l)$ whenever $k \geq l$.)

For $\eta, \zeta \in \Omega$, we say $\eta$ *covers* $\zeta$ ($\eta \cdot\!> \zeta$) if there is an $i \in [n]$ for which $\eta_i = 1$, $\zeta_i = 0$ and $\eta_j = \zeta_j$ for all $j \neq i$. Following [40] we say that $\mu \in \mathcal{M}$ *stochastically covers* $\nu \in \mathcal{M}$ (written $\mu \succ\!\!\cdot\, \nu$) if we can couple r.v.'s $\eta$, $\zeta$ having laws $\mu$ and $\nu$ so that with probability 1, $\eta = \zeta$ or $\eta \cdot\!> \zeta$; and that $\mu$ has the *stochastic covering property* (SCP) if $\mu(\cdot|\eta_i = 0) \succ\!\!\cdot\, \mu(\cdot|\eta_i = 1)$ for every $i$ (where, again, we regard these as measures on $\Omega_{[n]\setminus\{i\}}$). Observe that if $\mu$ is NA+ (or even just NA) then $\mu(\cdot|\eta_i = 0)$ stochastically dominates $\mu(\cdot|\eta_i = 1)$; a possible strengthening suggested by Pemantle is

**Conjecture 3.2 ([40], Conjecture 9)** *NA+ implies the SCP.*

Again, our examples will give

**Theorem 3.3** *Conjectures 3.1 and 3.2 are false.*

Conjecture 3.1 was also disproved in [4]; see also the note at the end of this section.

We now describe the examples. For a positive integer $k \geq 2$ and real number $\beta \in (0,1)$, let $\nu^{k,\beta}$ be the measure on $\Omega_{2k}$ with

$$\nu^{k,\beta}(\eta) \propto \begin{cases} 1 & \text{if } [|\eta| = k-1 \text{ and } \eta_1 = 1] \\ \beta^2 & \text{if } [|\eta| = k-1 \text{ and } \eta_1 = 0] \\ \beta & \text{if } |\eta| = k \\ \beta^2 & \text{if } [|\eta| = k+1 \text{ and } \eta_1 = 1] \\ 1 & \text{if } [|\eta| = k+1 \text{ and } \eta_1 = 0] \\ 0 & \text{otherwise} \end{cases}$$

Note that (clearly) $\nu^{k,\beta}$ is almost exchangeable.

**Proposition 3.4** *The measure $\nu^{k,\beta}$ satisfies*

(a) *NA+ if and only if $\beta \geq \frac{1}{\sqrt{2}}$*

(b) *ULC if and only if $\beta \geq 1 - \frac{2}{k+1}$*

(c) *unimodality (and also LC) if and only if $\beta \geq 1 - \sqrt{\frac{2}{k+1}}$*

(d) *NMP if and only if $\beta \geq \sqrt{1 - \frac{2}{k+1}}$*

(e) *SCP if and only if $\beta \geq \sqrt{1 - \frac{2}{k+1}}$*

For example, for $\beta = 0.71$: $\nu^{6,\beta}$ is NA+ but not ULC, giving the first part of Theorem 1.6 (i.e. disproving Conjecture 1.5); $\nu^{23,\beta}$ is NA+ but not unimodal (proving Theorem 1.6); and $\nu^{4,\beta}$ is NA+ but not NMP or SCP (proving Theorem 3.3).

*Proof.* We will mainly prove what we need for Theorems 1.6 and 3.3, namely "if" in (a) and "only if" in (b)-(e). The other direction in (b),(c) will come for free, but we omit the (not very difficult) verifications of the remaining implications.

Fix $k$ and $\beta$, write $\nu$ for $\nu^{k,\beta}$, and set $r_i = \nu(|\eta| = i)$. We have, for some $C$,

$$r_k = C\beta\binom{2k}{k} \quad \text{and} \quad r_{k-1} = r_{k+1} = C\Big(\frac{k-1}{2k} + \frac{k+1}{2k}\beta^2\Big)\binom{2k}{k-1}.$$

Unimodality and LC for $\nu$ are equivalent to (each other and) $r_k \geq r_{k-1}$, which reduces to

$$\beta^2 - 2\beta + \frac{k-1}{k+1} \leq 0,$$

giving (c). ULC for $\nu$ is equivalent to $k^2 r_k^2 \geq (k+1)^2 r_{k-1} r_{k+1}$, which reduces to

$$(k+1)\beta^2 - 2k\beta + (k-1) \leq 0,$$

giving (b). The NMP requires that

$$\nu(\eta_1 = 1 \| \eta| = k) = \frac{1}{2}$$

be at least as large as

$$\nu(\eta_1 = 1 \| \eta| = k-1) = \frac{k-1}{(k-1) + (k+1)\beta^2},$$

from which the forward direction of (d) follows. The SCP requires

$$\nu(|\eta| = k-1 | \eta_1 = 1) \leq \nu(|\eta| = k-1 | \eta_1 = 0),$$

which reduces to

$$\binom{2k-1}{k-2} \leq \beta^2 \binom{2k-1}{k-1}$$

and yields the forward direction of (e).

It remains to prove the backward direction of (a); that is, we assume $\beta \geq 1/\sqrt{2}$ and should show $\nu$ is NA+. Since $\nu$ is almost exchangeable, Theorem 1.4 says we only need to show NC+, which, by symmetry, will follow if we show $\eta_1 \downarrow \eta_2$ and $\eta_2 \downarrow \eta_3$ with respect to $W \circ \nu$, for any external field $W$. (Our original proof of this has been shortened using some ideas from [4].) Observe that, since a limit of NC measures is NC, it suffices to consider the case when all entries of $W$ are finite and strictly positive.

Let $W' = (W_1, 1, \ldots, 1)$, and let $\nu'$ be the projection of $W' \circ \nu$ on $\Omega_{\{2,\ldots,2k\}}$. To prove $\eta_2 \downarrow \eta_3$ for $W \circ \nu$, it suffices to show $\nu'$ is NC+, which, since $\nu'$ is exchangeable, will follow *via* Theorem 2.2 if we show $\nu'$ has a ULC rank sequence. The nonzero part of the normalized rank sequence $(a_i := \nu'(|\eta| = i)/\binom{2k-1}{i})_{i=0}^{2k-1}$ is

$$(a_{k-2}, \ldots, a_{k+1}) \propto (W_1, W_1\beta + \beta^2, W_1\beta^2 + \beta, 1),$$

which a straightforward computation shows to be LC when $\beta \geq 1/\sqrt{2}$.

That $\eta_1 \downarrow \eta_2$ for $W \circ \nu$ will follow immediately from

$$W \circ \nu(\cdot | \eta_1 = 0) \text{ stochastically dominates } W \circ \nu(\cdot | \eta_1 = 1). \tag{3.1}$$

Set

$$\begin{aligned}
\pi_1 &= W \circ \nu(\cdot|\eta_1 = 0, |\eta| = k+1), \\
\pi_2 &= W \circ \nu(\cdot|\eta_1 = 0, |\eta| \in \{k-1, k\}), \\
\pi_3 &= W \circ \nu(\cdot|\eta_1 = 1, |\eta| \in \{k, k+1\}), \text{ and} \\
\pi_4 &= W \circ \nu(\cdot|\eta_1 = 1, |\eta| = k-1).
\end{aligned}$$

It follows readily from Lemma 2.5 (since $\beta < 1$ and the two measures appearing in (3.1) are rank rescalings of a common product measure, namely the measure $\mu \in \mathcal{M}_{\{2,\ldots,2k\}}$ with $\mu(\tau) \propto \prod W_i^{\tau_i}$) that each of $\pi_1$, $\pi_2$ stochastically dominates each of $\pi_3$, $\pi_4$. Consequently, every convex combination of $\pi_1$ and $\pi_2$ stochastically dominates every convex combination of $\pi_3$ and $\pi_4$, which in particular gives (3.1). $\square$

Before closing this section, let us just mention that a more natural class of counterexamples to Conjecture 3.1 is *probably* provided by the following simple construction, which, as far as we know, first appeared in [9]. Given $k$, let $G$ be the graph with $V(G) = \{x, y, z_1, \ldots, z_k\}$ and $E(G) = \{xy, xz_1, \ldots, xz_k, yz_1, \ldots, yz_k\}$. It is well known and easy to see (consider the event $\{\eta_{xy} = 1\}$) that for $k \geq 5$, the USF measure for $G$ fails the NMP, so is a counterexample to Conjecture 3.1 if the USF measure for $G$ is NA+. The latter would follow from Conjecture 1.20(b) (USF measures are FM+) for $G$, since Conjecture 1.13 for these graphs is contained in the result from [46] mentioned following Conjecture 1.14. (We can prove FM+ for $k \leq 5$, and even this is not so easy).

## 3.2 Rayleigh implies BLC[5]

The main point here is the following lemma, stating that NC+ (i.e. the Rayleigh property) implies the APP (defined in Section 1.4) for measures in $\mathcal{M}_4$. For notational simplicity, we set $\alpha_X = \mu(X)$ for $X \subseteq [n]$ (where we now treat $\Omega_n$ as $2^{[n]}$), often omit commas and set braces in subscripts (e.g. $\alpha_{134} = \mu(\{1,3,4\})$), and write $\alpha_0$ for $\mu(\emptyset)$. Let $\Sigma_{r,s}^t = \sum \alpha_X \alpha_Y$, with the sum over unordered pairs $\{X, Y\}$ of subsets of $[n]$ with $|X| = r$, $|Y| = s$, and $|X \cap Y| = t$.

**Lemma 3.5** *If $\mu \in \mathcal{M}_4$ is NC+, then*

$$3\Sigma^0_{1,3} \leq 4\Sigma^0_{2,2}. \tag{3.2}$$

We first prove this and then give the easy derivation of Theorem 1.7. (Notice that we could also get Theorem 1.7 from Lemma 3.5 via Theorem 1.8, (3.2) being the only part of the CAPP that is not immediate from NC+; but, as the proof of Theorem 1.8 is relatively difficult, we give a direct proof of Theorem 1.7 here.)

For convenience, we now work with unnormalized (nonnegative) measures on $\Omega$, and say that such a measure $\mu$ with $\mu(\Omega) > 0$ has a property (CNC, NC+, etc.) iff its normalization $\mu'$ (given by $\mu'(\eta) = \mu(\eta)/\mu(\Omega)$) does. Observe that $\mathcal{A} \downarrow \mathcal{B}$ under $\mu$ if and only if

$$\mu(\mathcal{A}\bar{\mathcal{B}})\mu(\bar{\mathcal{A}}\mathcal{B}) \geq \mu(\mathcal{A}\mathcal{B})\mu(\bar{\mathcal{A}}\bar{\mathcal{B}}) \tag{3.3}$$

(where $\bar{\mathcal{A}} = \Omega \setminus \mathcal{A}$).

*Proof of Lemma 3.5.* Let $A = \Sigma^0_{2,2}$ $(= \alpha_{12}\alpha_{34} + \alpha_{13}\alpha_{24} + \alpha_{14}\alpha_{23})$. We may assume $\alpha_0 = \alpha_{1234} = 0$, since decreasing $\alpha_0$ or $\alpha_{1234}$ preserves NC+ and has no effect on (3.2). Furthermore, by renaming variables, applying a uniform external field, and scaling (none of which affect (3.2)), we may assume

$$\alpha_{123} = \alpha_4 = 1 \quad \text{and} \quad \alpha_1\alpha_{234}, \alpha_2\alpha_{134}, \alpha_3\alpha_{124} \leq 1. \tag{3.4}$$

(First, rename coordinates so $\alpha_i\alpha_{[4]\setminus\{i\}}$ is largest for $i = 4$ (and observe we may assume this largest value is strictly positive). Second, impose a uniform external field (one of the form $(w, w, w, w)$ for some $w \in \mathbb{R}_+$) to get $\alpha_{123} = \alpha_4$. Third, divide all $\alpha_X$ by $\alpha_{123}$.)

Since $\mu$ is NC+ (so in particular CNC), (3.4) gives $\alpha_1 \leq \alpha_{12}\alpha_{13}$ (and similarly for $\alpha_2$, $\alpha_3$) and $\alpha_{124} \leq \alpha_{14}\alpha_{24}$ (and similarly for $\alpha_{134}$, $\alpha_{234}$). It thus suffices to show

$$3(1 + xy + xz + yz) \leq 4A, \tag{3.5}$$

where

$$x = \alpha_{12}\alpha_{34}, \; y = \alpha_{13}\alpha_{24}, \; \text{ and } \; z = \alpha_{14}\alpha_{23} \quad (\text{so } A = x + y + z).$$

For fixed $A$, the left hand side of (3.5) is maximized when $x = y = z$; thus (3.5) holds whenever $A \in [1, 3]$ (as can be seen by examining the quadratic polynomial $A^2 - 4A + 3$). In view of (3.4) we can assume $A \leq 3$, so we just need $A \geq 1$.

Assume, for a contradiction, that $A < 1$. Negative correlation of $\eta_2$ and $\eta_3$ for the measure $(0, 1, 1, W) \circ \mu$ implies (use (3.3))

$$P^1(W) := (\alpha_{24}\alpha_{34} - \alpha_{234})W^2 + (\alpha_2\alpha_{34} + \alpha_3\alpha_{24} - \alpha_{23})W + \alpha_2\alpha_3 \geq 0 \qquad \text{for } W > 0.$$

Similarly, negative correlation of $\eta_1$ and $\eta_2$ for $(\infty, 1, 1, W) \circ \mu$ implies

$$P_1(W) := \alpha_{124}\alpha_{134}W^2 + (\alpha_{12}\alpha_{134} + \alpha_{13}\alpha_{124} - \alpha_{14})W + (\alpha_{12}\alpha_{13} - \alpha_1) \geq 0 \qquad \text{for } W > 0.$$

Similarly (interchanging 1 with either 2 or 3) we have, again for $W > 0$,

$$P^2(W) := (\alpha_{14}\alpha_{34} - \alpha_{134})W^2 + (\alpha_1\alpha_{34} + \alpha_3\alpha_{14} - \alpha_{13})W + \alpha_1\alpha_3 \geq 0,$$

$$P_2(W) := \alpha_{124}\alpha_{234}W^2 + (\alpha_{12}\alpha_{234} + \alpha_{23}\alpha_{124} - \alpha_{24})W + (\alpha_{12}\alpha_{23} - \alpha_2) \geq 0,$$

$$P^3(W) := (\alpha_{14}\alpha_{24} - \alpha_{124})W^2 + (\alpha_1\alpha_{24} + \alpha_2\alpha_{14} - \alpha_{12})W + \alpha_1\alpha_2 \geq 0, \text{ and}$$

$$P_3(W) := \alpha_{134}\alpha_{234}W^2 + (\alpha_{13}\alpha_{234} + \alpha_{23}\alpha_{134} - \alpha_{34})W + (\alpha_{13}\alpha_{23} - \alpha_3) \geq 0.$$

We pause to show

$$\alpha_X > 0 \text{ for } X \neq \emptyset, [4]. \tag{3.6}$$

First we show $\alpha_X > 0$ if $|X| = 2$. Suppose for example that $\alpha_{12} = 0$. Since $P^1(W) \geq 0$ for all $W > 0$ and $\alpha_2 = 0$ (since $\alpha_2 \leq \alpha_{12}\alpha_{23}$) the coefficient of $W$ in $P^1$ must be nonnegative, and thus (using $\alpha_3 \leq \alpha_{13}\alpha_{23}$) $y\alpha_{23} \geq \alpha_{23}$. If $\alpha_{23} > 0$, this gives $A \geq y \geq 1$; thus (since we are assuming $A < 1$) $\alpha_{23} = 0$, and similar reasoning shows $\alpha_1 = \alpha_3 = \alpha_{13} = 0$. Hence, $\alpha_X = 0$ unless $X = \{1, 2, 3\}$ or $4 \in X$; but then nonnegativity of $P_1$, $P_2$, and $P_3$ gives $\alpha_{14} = \alpha_{24} = \alpha_{34} = 0$. Thus $\alpha_X > 0$ if and only if $X = \{1, 2, 3\}$ or $X = \{4\}$; but for any such measure $\eta_1$ and $\eta_2$ are strictly positively correlated. This contradiction shows $\alpha_{12} > 0$, and similar arguments (or symmetry) give $\alpha_X > 0$ whenever $|X| = 2$.

If $\alpha_2 = 0$, then nonnegativity of the linear term in $P^1$ gives, as in the preceding paragraph (and using $\alpha_{23} > 0$), $A \geq 1$; thus $\alpha_2 > 0$. Similar arguments (or, again, symmetry) show $\alpha_1$, $\alpha_3$, $\alpha_{124}$, $\alpha_{134}$, and $\alpha_{234}$ are positive, and we have (3.6).

Set

$$a = \frac{\alpha_1}{\alpha_{12}\alpha_{13}}, \; b = \frac{\alpha_2}{\alpha_{12}\alpha_{23}}, \; c = \frac{\alpha_3}{\alpha_{13}\alpha_{23}}, \; d = \frac{\alpha_{124}}{\alpha_{14}\alpha_{24}}, \; e = \frac{\alpha_{134}}{\alpha_{14}\alpha_{34}}, \; \text{ and } \; f = \frac{\alpha_{234}}{\alpha_{24}\alpha_{34}}.$$

Note $a, b, c, d, e, f \in (0, 1]$. If the coefficient of W in $P^1$ is nonnegative, then, as above, $A \geq 1$; thus this coefficient is negative, whence the discriminant of $P^1$ is nonpositive. This yields

$$1 - \frac{\alpha_2\alpha_{34}}{\alpha_{23}} - \frac{\alpha_3\alpha_{24}}{\alpha_{23}} \leq 2\sqrt{\frac{\alpha_2\alpha_3}{\alpha_{23}^2}(\alpha_{24}\alpha_{34} - \alpha_{234})},$$

which in the notation introduced above becomes

$$1 - bx - cy \leq 2\sqrt{bxcy(1-f)} \leq (bx + cy)\sqrt{1-f}.$$

Thus

$$x + y \geq [(1 + \sqrt{1-f}) \max\{b, c\}]^{-1},$$

and a similar argument using $P_1$ gives

$$x + y \geq [(1 + \sqrt{1-a}) \max\{d, e\}]^{-1},$$

so that

$$x + y \geq \max\{[(1 + \sqrt{1-f}) \max\{b, c\}]^{-1}, [(1 + \sqrt{1-a}) \max\{d, e\}]^{-1}\}.$$

Similar arguments using $P^2$, $P_2$, $P^3$, and $P_3$ yield

$$x + z \geq \max\{[(1 + \sqrt{1-e}) \max\{a, c\}]^{-1}, [(1 + \sqrt{1-b}) \max\{d, f\}]^{-1}\}$$

and

$$y + z \geq \max\{[(1 + \sqrt{1-d}) \max\{a, b\}]^{-1}, [(1 + \sqrt{1-c}) \max\{e, f\}]^{-1}\}.$$

In particular, we have $x + y, x + z, y + z \geq 1/2$.

The proof is now an easy consequence of

$$\inf\{[(1 + \sqrt{1-v})u]^{-1} + [(1 + \sqrt{1-u})v]^{-1} : 0 < u, v \leq 1\} = \frac{27}{16}, \tag{3.7}$$

verification of which is a straightforward calculus exercise which we omit. Assuming (3.7) and, without loss of generality, $a \geq b$ and $d \geq e$, we have

$$(x + y) + (y + z) \geq [(1 + \sqrt{1-a})d]^{-1} + [(1 + \sqrt{1-d})a]^{-1} \geq \frac{27}{16},$$

which, combined with $x + z \geq 1/2$, gives the final contradiction $2A > 2$. $\square$

*Proof of Theorem 1.7.* First observe that BLC$[m]$ is equivalent to having

$$\text{NC+ implies ULC for measures in } \mathcal{M}_n \tag{3.8}$$

for all $n \le m$. Suppose $\mu \in \mathcal{M}_n$ has rank sequence $(r_i)_{i=0}^n$. Notice that in general for (3.8) it is enough to show that NC+ implies

$$r_k^2 \binom{n}{k}^{-2} \ge r_{k-1}r_{k+1}\binom{n}{k-1}^{-1}\binom{n}{k+1}^{-1} \tag{3.9}$$

for $1 \le k \le \lfloor n/2 \rfloor$, since the measure $\mu^* \in \mathcal{M}_n$ with $\mu^*(X) = \mu([n] \setminus X)$ has rank sequence $(r_{n-i})_{i=0}^n$ and is NC+ if and only if $\mu$ is. In fact, Choe and Wagner [7] show that (3.9) holds for $k = 1$ and *any* $n$ (assuming NC+). This gives (3.8) for $n \le 3$ (and hence BLC[3]) and for the cases of interest here—that is, $n = 4, 5$—reduces the problem to proving (3.9) when $k = 2$.

Assume $n \in \{4, 5\}$. Using inequalities of the form

$$\alpha_i \alpha_{ijl} \le \alpha_{ij}\alpha_{il}$$

(which follow from NC+) we obtain

$$\Sigma^1_{1,3} \le \Sigma^1_{2,2}. \tag{3.10}$$

It follows from Lemma 3.5 (for $n = 4$ this is the conclusion of the lemma, and for $n = 5$ we apply the lemma to each of the five conditional measures $\mu(\cdot|\eta_i = 0)$) that

$$3\Sigma^0_{1,3} \le 4\Sigma^0_{2,2}. \tag{3.11}$$

Note also that Cauchy-Schwarz implies that the average size of a term in $\Sigma^2_{2,2}$ is at least the average size of a term in either of $\Sigma^0_{2,2}$, $\Sigma^1_{2,2}$; that is,

$$\Sigma^2_{2,2} \ge \frac{4}{(n-2)(n-3)}\Sigma^0_{2,2} \quad \text{and} \quad \Sigma^2_{2,2} \ge \frac{1}{n-2}\Sigma^1_{2,2}. \tag{3.12}$$

Thus, finally, we have (3.9) for $k = 2$:

$$9r_1r_3 = 9\Sigma^0_{1,3} + 9\Sigma^1_{1,3} \le 12\Sigma^0_{2,2} + 9\Sigma^1_{2,2} \le 8\Sigma^0_{2,2} + 8\Sigma^1_{2,2} + 4\Sigma^2_{2,2} = 4r_2^2 \quad \text{if } n = 4$$

and

$$2r_1r_3 = 2\Sigma^0_{1,3} + 2\Sigma^1_{1,3} \le \frac{8}{3}\Sigma^0_{2,2} + 2\Sigma^1_{2,2} \le 2\Sigma^0_{2,2} + 2\Sigma^1_{2,2} + \Sigma^2_{2,2} = r_2^2 \quad \text{if } n = 5,$$

where in each case we used (3.10) and (3.11) for the first inequality and (3.12) for the second. $\square$

# Chapter 4

# The antipodal pairs property and ultra-log-concavity

The main point of this chapter is the proof of (a generalization of) Theorem 1.8, given in Section 4.1. In the remaining sections we deduce several consequences (including Corollary 1.10 and Theorems 1.11 and 1.17) by using Theorem 1.8 (or Theorem 4.1) to establish ultra-log-concavity in certain situations. (See Theorem 5.11 for another application of Theorem 1.8.)

## 4.1 The CAPP implies ULC

In this section we prove the following generalization of Theorem 1.8, which will be used in Section 4.4.

**Theorem 4.1** *Suppose $\mu \in \mathcal{M}$ has the property that, for every $k \in [t]$, every measure gotten from $\mu$ by conditioning on the values of $n-2k$ coordinates has the APP. Then the sequences $(\mu(|\eta| = i)/\binom{n}{i})_{i=0}^{t+1}$ and $(\mu(|\eta| = i)/\binom{n}{i})_{i=n-t-1}^{n}$ are LC.*

For the rest of this section it will be convenient to treat $\Omega$ as $2^{[n]}$, so that (1.5) becomes

$$\alpha_i(\mu) = \binom{n}{i}^{-1} \sum \{\mu(X)\mu([n] \setminus X) : X \in \binom{[n]}{i}\}$$

(where $\binom{[n]}{i} = \{X \subseteq [n] : |X| = i\}$).

We will use some properties of the Johnson association scheme; this material (up to (4.2)) is taken from chapter 30 of [34]. Fix positive integers $n$ and $l$ with $l \leq n/2$, let $\mathfrak{X} = \binom{[n]}{l}$, and, for $i = 0, 1, \ldots, l$, let $A_i$ be the $\mathfrak{X} \times \mathfrak{X}$ adjacency matrix of $i$th associates, *viz.*

$$A_i(X, Y) = \begin{cases} 1 & \text{if } |X \cap Y| = l - i \\ 0 & \text{otherwise.} \end{cases}$$

We write elements of $\mathbb{R}^{\mathfrak{X}}$ as row vectors. For $T \subseteq [n]$ with $|T| \le l$, let $e_T$ be the vector in $\mathbb{R}^{\mathfrak{X}}$ with

$$e_T(S) = \begin{cases} 1 & \text{if } S \supseteq T \\ 0 & \text{otherwise,} \end{cases}$$

and let $U_i$ be the span of $\{e_T : T \in \binom{[n]}{i}\}$. Then $U_0 \subseteq U_1 \subseteq \cdots \subseteq U_l = \mathbb{R}^{\mathfrak{X}}$ and $\dim U_i = \binom{n}{i}$. Set $V_0 = U_0$ and $V_i = U_i \cap U_{i-1}^{\perp}$ for $i = 1, 2, \ldots, l$, and let $E_i$ be the projection of $\mathbb{R}^{\mathfrak{X}}$ onto $V_i$. Then

$$\mathbb{R}^{\mathfrak{X}} = V_0 \oplus V_1 \oplus \cdots \oplus V_l$$

is an orthogonal decomposition,

$$E_i E_j = \begin{cases} E_i & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

and

$$E_0 + E_1 + \cdots + E_l = I.$$

Note that $V_0$ consists of the constant vectors. The span, $\mathfrak{A}$, of $A_0, \ldots, A_l$ is an algebra under matrix multiplication (the *Bose-Mesner algebra*). The set of matrices $\{E_0, E_1, \ldots, E_l\}$ is also a basis for $\mathfrak{A}$, with

$$A_i = \sum_{j=0}^{l} P_i(j) E_j \quad (i = 0, 1, \ldots, l), \tag{4.1}$$

where

$$P_i(j) = \sum_{k=0}^{i} (-1)^{i-k} \binom{l-k}{i-k} \binom{l-j}{k} \binom{n-l+k-j}{k}. \tag{4.2}$$

The next lemma is presumably well-known.

**Lemma 4.2** *For any $\gamma_0, \gamma_1, \ldots, \gamma_l \in \mathbb{R}$, the $\mathfrak{X} \times \mathfrak{X}$ real symmetric matrix $M = \sum_{i=0}^{l} \gamma_i A_i$ is positive semidefinite if and only if $\sum_{i=0}^{l} \gamma_i P_i(j) \ge 0$ for $j = 0, 1, \ldots, l$.*

*Proof.* Since $M = \sum_{j=0}^{l} \sum_{i=0}^{l} \gamma_i P_i(j) E_j$ and $E_j$ is the orthogonal projection of $\mathbb{R}^{\mathfrak{X}}$ onto $V_j$, the eigenvalues of $M$ are $\{\sum_{i=0}^{l} \gamma_i P_i(j) : j = 0, 1, \ldots, l\}$. □

*Remark.* It is not hard to show, using some additional properties of the Johnson scheme, that the condition appearing in Lemma 4.2 is equivalent to the statement that the vector $(\binom{l}{i}\binom{n-l}{l-i}\gamma_i)_{i=0}^{l}$ satisfies *Delsarte's inequalities* ([8] or [34, p. 416]).

We also need one technical lemma:

**Lemma 4.3** *For all positive integers $M, N$ and real numbers $a, b$,*

$$\sum_{t=0}^{N}(-1)^t \frac{at+b}{t+M}\binom{N}{t} = (\frac{b}{M} - a)\binom{M+N}{M}^{-1}.$$

*Proof.* It suffices to prove either of the equivalent

$$\sum_{t=0}^{N}(-1)^t \frac{t}{t+M}\binom{N}{t}\binom{M+N}{M} = -1, \tag{4.3}$$

$$\sum_{t=0}^{N}(-1)^t \frac{M}{t+M}\binom{N}{t}\binom{M+N}{M} = 1,$$

since the desired identity is a linear combination of these. We prove (4.3), fixing $M$ and proceeding by induction on $N$, with the base case $N = 1$ trivial. For the induction step, we just check that the left side of (4.3) does not change when we replace $N$ by $N + 1$; indeed, the difference is

$$\begin{aligned}(-1)^{N+1}&\frac{N+1}{M+N+1}\binom{M+N+1}{M} \\ &+ \sum_{t=1}^{N}(-1)^t \frac{t}{t+M}\left[\binom{N+1}{t}\binom{M+N+1}{M} - \binom{N}{t}\binom{M+N}{M}\right] \\ &= (-1)^{N+1}\binom{M+N}{M} + \sum_{t=1}^{N}(-1)^t\binom{N}{t-1}\binom{M+N}{M},\end{aligned}$$

which is zero. $\square$

*Proof of Theorem 4.1.* Let $\mu$ be a measure on $2^{[n]}$ satisfying the hypotheses of the theorem, with rank sequence $(a_i)_{i=0}^{n}$. Our goal is to show

$$l(n-l)a_l^2 \geq (l+1)(n-l+1)a_{l-1}a_{l+1} \tag{4.4}$$

for $l \in \{1, \ldots, t\} \cup \{n-t, \ldots, n-1\}$; but, since $\mu' \in \mathcal{M}$ given by $\mu'(X) = \mu([n] \setminus X)$ again satisfies the hypotheses of Theorem 4.1 and has rank sequence $(a_{n-i})_{i=0}^{n}$, it suffices to prove (4.4) when $l \leq \min\{t, n/2\}$. To this end, fix such an $l$ and set

$$Z_{j,k}^{i} = \sum\{\mu(X)\mu(Y) : (X,Y) \in \binom{[n]}{j} \times \binom{[n]}{k} \text{ and } |X \cap Y| = i\};$$

with this notation, (4.4) is

$$l(n-l)\sum_{i=0}^{l} Z_{l,l}^{i} \geq (l+1)(n-l+1)\sum_{i=0}^{l-1} Z_{l-1,l+1}^{i}. \tag{4.5}$$

For each $i \in \{0,1,\ldots,l-1\}$ and $I \subseteq J \subseteq [n]$ with $|I| = i$ and $|J| = 2l-i$, let $\mu_{I,J} \in \mathcal{M}_{J\setminus I}$ be the conditional measure with

$$\mu_{I,J}(X) \propto \mu(X \cup I) \qquad (X \subseteq J \setminus I)$$

(or $\mu_{I,J} \equiv 0$ if $\mu([I,J]) = 0$). By hypothesis (or trivially if $\mu_{I,J} \equiv 0$) $\mu_{I,J}$ has the APP, i.e.

$$\alpha_{l-i}(\mu_{I,J}) \geq \alpha_{l-i-1}(\mu_{I,J}). \tag{4.6}$$

With

$$Z_{j,k}(I,J) = \sum\{\mu(X)\mu(Y) : (X,Y) \in \binom{[n]}{j} \times \binom{[n]}{k},\ X \cup Y = J,\ \text{and } X \cap Y = I\},$$

we have

$$\alpha_j(\mu_{I,J}) = \left(\sum_{I\subseteq X\subseteq J} \mu(X)\right)^{-1} \binom{2l-2i}{j}^{-1} Z_{i+j,2l-i-j}(I,J),$$

and (4.6) becomes

$$\frac{l-i}{l-i+1} Z_{l,l}(I,J) \geq Z_{l-i.l+1}(I,J). \tag{4.7}$$

Summing (4.7) over $I, J$ with $|I| = i$ and $|J| = 2l-i$ gives

$$\frac{l-i}{l-i+1} Z_{l,l}^{i} \geq Z_{l-1,l+1}^{i} \tag{4.8}$$

(since each pair $(X,Y)$ contributing to $Z_{l,l}^{i}$ contributes to the left side of (4.7) for exactly one choice of $(I,J)$, and similarly for pairs contributing to $Z_{l-1,l+1}^{i}$). Replacing each $Z_{l-1,l+1}^{i}$ in (4.5) by the (corresponding) left side of (4.8), we find that it is enough to show

$$\sum_{i=0}^{l} \beta_i Z_{l,l}^{i} \geq 0, \tag{4.9}$$

where

$$\beta_i = \frac{i(n+1) - l(l+1)}{l-i+1}.$$

In fact, we will show that (4.9) holds for *every* $\mu \in \mathcal{M}$. Let $\psi = \psi_\mu$ be the (row) vector in $\mathbb{R}^{\mathfrak{X}}$ with $\psi(X) = \mu(X)$ for $X \in \binom{[n]}{l}$, and recall the matrices $A_i$ defined before Lemma 4.2. Since

$$\psi A_i \psi^T = Z_{l,l}^{l-i},$$

the left side of (4.9) is

$$\psi(\sum_{i=0}^{l} \beta_{l-i} A_i)\psi^T \tag{4.10}$$

and (4.9) will follow from Lemma 4.2 once we show

$$\sum_{i=0}^{l} \beta_{l-i} P_i(j) \geq 0 \tag{4.11}$$

for $j = 0, 1, \ldots, l$.

Fix $j \in [l]$. (We deal with the case $j = 0$ separately below.) The left side of (4.11) is

$$\sum_{i=0}^{l} \frac{(l-i)(n+1) - l(l+1)}{i+1} \sum_{k=0}^{i} (-1)^{i-k} \binom{l-k}{i-k} \binom{l-j}{k} \binom{n-l+k-j}{k}, \tag{4.12}$$

which we want to show is nonnegative. Interchanging the order of summation and making the substitution $t = i - k$, we may rewrite (4.12) as

$$\sum_{k=0}^{l} \binom{l-j}{k} \binom{n-l+k-j}{k} \sum_{t=0}^{l-k} (-1)^t \frac{(l-t-k)(n+1) - l(l+1)}{t+k+1} \binom{l-k}{t}. \tag{4.13}$$

It is thus enough to show

$$\sum_{t=0}^{l-k} (-1)^t \frac{(l-t-k)(n+1) - l(l+1)}{t+k+1} \binom{l-k}{t} \geq 0 \tag{4.14}$$

whenever $k \leq l-1$ (since the $k = l$ term in (4.13) is zero). But Lemma 4.3, with $N = l-k$, $M = k+1$, $a = -(n+1)$, and $b = N(n+1) - l(l+1)$, says that the left side of (4.14) is

$$\frac{(n-l+1)(l+1)}{k+1} \binom{l+1}{k+1}^{-1},$$

which is positive since $l \leq n/2$.

Finally we show that when $j = 0$, (4.11) holds with equality. To see this, notice that when $\mu$ is uniform measure on $2^{[n]}$, we have equality in (4.5) and (4.8), and consequently

(4.9), from which it follows (see (4.10)) that

$$\psi(\sum_{i=0}^{l} \beta_{l-i} A_i)\psi^T = 0. \tag{4.15}$$

But, since $\psi E_j \psi^T$ is $2^{-2n}\binom{n}{l}$ if $j = 0$ and zero otherwise (recall that $E_0$ is projection onto the span of $(1, 1, \ldots, 1)$), the left side of (4.15) is (by (4.1))

$$2^{-2n}\binom{n}{l}\sum_{i=0}^{l} \beta_{l-i} P_i(0),$$

which gives the promised equality in (4.11). □

## 4.2 Proof of Corollary 1.10

In this short section we use Theorem 1.8 to prove Corollary 1.10.

*Proof of Corollary 1.10.* The statement is: if $\mu \in \mathcal{M}$ satisfies $(1+1/k)$-Ray$[k]$ $\forall k \in [m]$, $T \subseteq [n]$, $|T| \leq 2m+1$, and $\nu \in \mathcal{M}_T$ is obtained from $\mu$ by imposing an external field and projecting on $T$, then $\nu$ is ULC. By Theorem 1.8, it suffices to show $\nu$ has the CAPP (since it's not hard to see that the hypothesis implies the rank sequence of $\nu$ has no internal zeros). But any measure gotten from $\nu$ by conditioning on the values of the variables in some set $T \setminus S$ is the limit of a sequence of measures, each gotten from $\mu$ by imposing an external field and projecting on $S$; the CAPP for $\nu$ thus follows from our assumption on $\mu$. □

## 4.3 Convolution of ULC sequences

In this section we define a property of measures which is stronger than ULC, prove it is preserved by products, and show that this implies Theorem 1.11.

We begin with some definitions. With $\alpha_i(\mu)$ as in (1.5), say $\mu \in \mathcal{M}$ is *antipodal pairs unimodal* (APU) if the sequence $(\alpha_i(\mu))_{i=0}^n$ is unimodal (since $\alpha_i(\mu) = \alpha_{n-i}(\mu)$, this means $\alpha_0(\mu) \leq \cdots \leq \alpha_{\lfloor n/2 \rfloor}(\mu) = \alpha_{\lceil n/2 \rceil}(\mu) \geq \cdots \geq \alpha_n(\mu)$), and say $\mu$ is *conditionally antipodal pairs unimodal* (CAPU) if every measure obtained from $\mu$ by conditioning is APU. Since CAPU trivially implies the CAPP (and no internal zeros in the rank sequence), Theorem 1.8 gives

**Corollary 4.4** *Every CAPU measure is ULC.*

(As far as we know, Conjecture 1.15 can be strengthened by replacing "CAPP" with "CAPU.") We will show

**Theorem 4.5** (a) *The product of two APU measures is APU.*

(b) *The product of two CAPU measures is CAPU.*

Before giving the proof of Theorem 4.5, we show that it implies Theorem 1.11. Recall that $\mu \in \mathcal{M}$ is *exchangeable* if $\mu(\eta)$ depends only on $|\eta| = \sum \eta_i$.

**Lemma 4.6** *For exchangeable measures, the properties ULC and CAPU are equivalent.*

*Remark.* Pemantle shows that for exchangeable measures, ULC and several negative dependence properties coincide (see Theorem 2.2). Lemma 4.6 adds CAPU to this list.

*Proof of Lemma 4.6.* By Corollary 4.4, we need only show that every exchangeable ULC measure is CAPU; in fact, since conditioning preserves both exchangeability and ULC, it suffices to prove that an exchangeable ULC measure is APU. But if $\mu \in \mathcal{M}$ is exchangeable with rank sequence $(a_0, \ldots, a_n)$, then

$$\alpha_i(\mu) = a_i a_{n-i} \binom{n}{i}^{-1} \binom{n}{n-i}^{-1}$$

so that log-concavity of (and absence of internal zeros in) $(a_i / \binom{n}{i})_{i=0}^n$ implies unimodality of $(\alpha_i(\mu))_{i=0}^n$. $\square$

*Proof of Theorem 1.11.* Given ULC sequences $a = (a_0, \ldots, a_n)$ and $b = (b_0, \ldots, b_m)$, let $\mu \in \mathcal{M}_{[n]}$ and $\nu \in \mathcal{M}_{\{n+1,\ldots,n+m\}}$ be the corresponding exchangeable measures; that is,

$$\mu(\eta) = \frac{a_{|\eta|}}{\binom{n}{|\eta|}} \quad \text{and} \quad \nu(\eta) = \frac{b_{|\eta|}}{\binom{m}{|\eta|}}.$$

By Lemma 4.6, $\mu$ and $\nu$ are CAPU, so that Theorem 4.5(b) and Corollary 4.4 give ULC for $\mu \times \nu \in \mathcal{M}_{[n+m]}$, completing the proof (since the rank sequence of $\mu \times \nu$ is the convolution of $a$ and $b$). $\square$

*Remark.* Following [33], say an infinite nonnegative sequence $(a_0, a_1, \ldots)$ is ULC$[\infty]$ if there are no internal zeros and $a_i^2 \geq \frac{i+1}{i} a_{i-1} a_{i+1}$ for $i \geq 1$. The proof of Theorem

1.11 in [33] allows one or both sequences to be ULC$[\infty]$, but an easy limiting argument suffices to get this more general statement from the finite version proved above.

*Proof of Theorem 4.5.* Notice that (a) implies (b), since any measure gotten from $\mu\times\nu$ by conditioning is the product of measures obtained from $\mu$ and $\nu$ by conditioning.

Call a nonnegative sequence $(p_0,\ldots,p_s)$ *symmetric* if $p_i=p_{s-i}$ for $i=0,\ldots,s$ and *ultra-unimodal* if $(p_i/\binom{s}{i})_{i=0}^s$ is unimodal. Let $\mu\in\mathcal{M}_{[n]}$ and $\nu\in\mathcal{M}_{\{n+1,\ldots,n+m\}}$ be APU. Then $(\binom{n}{i}\alpha_i(\mu))_{i=0}^n$ and $(\binom{m}{i}\alpha_i(\nu))_{i=0}^m$ are symmetric and ultra-unimodal, and we want to say that their convolution, $(\binom{n+m}{k}\alpha_k(\mu\times\nu))_{k=0}^{n+m}$, is ultra-unimodal. So we will be done if we show

**Lemma 4.7** *The convolution of two symmetric ultra-unimodal sequences is ultra-unimodal*

(and symmetric). It's easy to see that Lemma 4.7 is not true without the symmetry assumption.

*Proof of Lemma 4.7.* Since every symmetric ultra-unimodal sequence $(p_0,\ldots,p_s)$ is a positive linear combination of sequences of the form $(\binom{s}{i}\mathbf{1}_{\{k\le i\le s-k\}})_{i=0}^s$ (and since convolution is bilinear), it suffices to prove that the convolution of $(\binom{s}{i}\mathbf{1}_{\{k\le i\le s-k\}})_{i=0}^s$ and $(\binom{t}{i}\mathbf{1}_{\{l\le i\le t-l\}})_{i=0}^t$ is ultra-unimodal for all $k,l,s,t$ with $k\le s/2$ and $l\le t/2$. (Of course this is also implied by Theorem 1.11.)

To see this set (for $k,l,s,t$ as above)

$$f_j=\binom{s+t}{j}^{-1}\sum_i\binom{s}{i}\mathbf{1}_{\{k\le i\le s-k\}}\binom{t}{j-i}\mathbf{1}_{\{l\le j-i\le t-l\}};$$

so we should show

$$f_j\le f_{j+1}\quad\forall j<(s+t)/2. \tag{4.16}$$

It's convenient to work with the natural interpretation of $f_j$ as a probability. Let $S$ and $T$ be disjoint sets with $|S|=s$ and $|T|=t$, and let

$$Q=\{Z\subseteq S\cup T: k\le|Z\cap S|\le s-k,\ l\le|Z\cap T|\le t-l\}.$$

Then $f_j=\Pr(X_j\in Q)$, where $X_j$ is chosen uniformly from $\binom{S\cup T}{j}$. To prove (4.16), we consider the usual coupling of $X=X_j$ and $Y=X_{j+1}$; namely, choose $X$ uniformly from $\binom{S\cup T}{j}$ and $y$ uniformly from $(S\cup T)\setminus X$, and set $Y = X\cup\{y\}$. We have

$$f_{j+1} - f_j = \Pr(X \notin Q, Y \in Q) - \Pr(X \in Q, Y \notin Q),$$

so should show that the right side is nonnegative.

We may assume that $j \geq k+l$, since otherwise we cannot have $X \in Q$. Then $\{X \notin Q, Y \in Q\}$ occurs if and only if either (i) $|X\cap S| = k-1$, $y \in S$, and $j-k+1 \leq t-l$, or (ii) $|X\cap T| = l-1$, $y\in T$, and $j-l+1 \leq s-k$; thus,

$$\Pr(X \notin Q, Y \in Q) = \tbinom{s}{k-1}\tbinom{t}{j-k+1}\tfrac{s-k+1}{s+t-j}\mathbf{1}_{\{j-k+1\leq t-l\}} + \tbinom{t}{l-1}\tbinom{s}{j-l+1}\tfrac{t-l+1}{s+t-j}\mathbf{1}_{\{j-l+1\leq s-k\}}.$$

Similarly (noting that $j \leq s-k+t-l$), $\{X\in Q, Y\notin Q\}$ occurs if and only if either (i) $|X\cap S| = s-k$, $y\in S$, and $l \leq j-s+k$ or (ii) $|X\cap T| = t-l$, $y \in T$, and $k \leq j-t+l$, whence

$$\Pr(X \in Q, Y \notin Q) = \tbinom{s}{s-k}\tbinom{t}{j-s+k}\tfrac{k}{s+t-j}\mathbf{1}_{\{l\leq j-s+k\}} + \tbinom{t}{t-l}\tbinom{s}{j-t+l}\tfrac{l}{s+t-j}\mathbf{1}_{\{k\leq j-t+l\}}.$$

Thus, since

$$\tbinom{s}{k-1}(s-k+1) = \tbinom{s}{s-k}k \quad \text{and} \quad \tbinom{t}{l-1}(t-l+1) = \tbinom{t}{t-l}l,$$

we will be done if we show

$$\tbinom{t}{j-k+1}\mathbf{1}_{\{j-k+1\leq t-l\}} \geq \tbinom{t}{j-s+k}\mathbf{1}_{\{l\leq j-s+k\}}. \tag{4.17}$$

and

$$\tbinom{s}{j-l+1}\mathbf{1}_{\{j-l+1\leq s-k\}} \geq \tbinom{s}{j-t+l}\mathbf{1}_{\{k\leq j-t+l\}}.$$

The easy verifications are similar and we just do (4.17): we have $j-k+1 \geq j-s+k$ (since $2k \leq s$) and $(j-k+1)+(j-s+k) \leq t$ (since $2j \leq s+t-1$), implying both $\binom{t}{j-k+1} \geq \binom{t}{j-s+k}$ and $\mathbf{1}_{\{j-k+1\leq t-l\}} \geq \mathbf{1}_{\{l\leq j-s+k\}}$. $\square$

## 4.4 Consequences for Mason's conjecture

In this section we prove Theorem 1.16. As noted at the end of Section 1.5, this with Theorem 4.1 (or, for part (a), Theorem 1.8) immediately implies Theorem 1.17. Here we do assume a (very) few matroid basics—again, [49] and [39] are standard references—and now denote matroids by $M$. Our argument mainly follows that of [10], which, as mentioned in Section 1.5, makes some progress on the "LC version" of Mason's conjecture.

Given a matroid $M$ on ground set $E$, let $\Pi_i = \Pi_i(M)$ be the set of ordered partitions $(A, B)$ of $E$ with $|A| = i$ and each of $A, B$ independent. Notice that when $|E| = 2k$, APP for uniform measure on the independent sets of $M$ is the inequality $|\Pi_{k-1}| \leq \frac{k}{k+1}|\Pi_k|$.

Dowling's point of departure was the observation that if $|\Pi_k(M)| \geq |\Pi_{k-1}(M)|$ for every $k \leq t$ and every $M$ on an $E$ of size $2k$, then for an arbitrary $M$ (on a ground set of any size) the initial portion $(a_0, \ldots, a_{t+1})$ of the sequence of independence numbers is LC. This is, of course, analogous to Theorem 4.1. Note, though, that, in contrast to Theorem 4.1, the implication here is quite straightforward; namely, a natural (and standard) grouping of terms represents the expansion of $a_i^2 \geq a_{k-1}a_{k+1}$ as a positive combination of inequalities $|\Pi_k(M)| \geq |\Pi_{k-1}(M)|$ for various $M$'s. (If, in analogy with (1.5), we set $\beta_i(\nu) = \sum\{\nu(\eta)\nu(\underline{1} - \eta) : \eta \in \Omega, |\eta| = i\}$, then Dowling's argument shows that $\mu \in \mathcal{M}$ is LC provided each $\nu$ obtained from $\mu$ by conditioning on the values of some $n - 2k$ variables satisfies $\beta_k(\nu) \geq \beta_{k-1}(\nu)$.)

Dowling also showed that every matroid on a ground set of size $2k \leq 14$ satisfies $|\Pi_k| \geq |\Pi_{k-1}|$ (which yields the result mentioned in Section 1.5). This is mainly based on Lemma 4.9 below and (a version of) the following easy observation, in which we use $d$ for degree and "$\sim$" for adjacency.

**Lemma 4.8** *Let $G$ be a simple, bipartite graph with bipartition $X \cup Y$. If $d(x) \geq 1$ for all $x \in X$ and $\sum_{x\sim y} d(x)^{-1} \leq C$ for all $y \in Y$, then $|X| \leq C|Y|$.*

*Proof.* This is standard: $|X| = \sum_{x\in X}\sum_{y\sim x} d(x)^{-1} = \sum_{y\in Y}\sum_{x\sim y} d(x)^{-1} \leq C|Y|$. $\square$

*Proof of Theorem 1.16.* Since the class of measures in question is closed under conditioning, it's enough to show that every matroid $M$ on a ground set $E$ of size $2k \leq 10$ satisfies

$$|\Pi_{k-1}(M)| \leq \frac{k}{k+1}|\Pi_k(M)|. \tag{4.18}$$

This is trivial when $k = 1$, so we assume $k \in \{2, 3, 4, 5\}$. Define bipartite graphs $G_1, G_2$ with the common bipartition $\Pi_{k-1} \cup \Pi_k$ by setting, for $(C, D) \in \Pi_{k-1}$ and $(A, B) \in \Pi_k$, $(C, D) \sim (A, B)$ in $G_1$ (resp. $G_2$) if $C \subseteq A$ (resp. $C \subseteq B$). Let $G = G_1 \cup G_2$. Then, writing $r$ for rank and $d_i$ and $d$ for degrees in $G_i$ and $G$, we have (see [10], pp. 24-27)

**Lemma 4.9** *If $r(M) \geq k + 2$ or $r(M) = k + 1$ and $M$ has no coloops, then*

(a) *every $(A, B) \in \Pi_k$ satisfies $2 \leq d_i(A, B) \leq k$ for $i = 1, 2$;*

(b) *every $(A, B) \in \Pi_k$ satisfies*

$$\sum_{(C,D)\sim(A,B)} \frac{1}{d(C,D)} \leq \frac{1}{2}\left(\frac{d_1(A,B)}{d_2(A,B)+1} + \frac{d_2(A,B)}{d_1(A,B)+1}\right);$$

(c) *every $(A, B) \in \Pi_k$ with $d_1(A, B) < d_2(A, B)$ satisfies*

$$\sum_{(C,D)\sim(A,B)} \frac{1}{d(C,D)} \leq \frac{1}{2}\left(\frac{d_1(A,B)-1}{d_1(A,B)+1} + \frac{d_2(A,B)-d_1(A,B)+1}{d_1(A,B)+2} + \frac{d_1(A,B)}{d_2(A,B)+1}\right). \tag{4.19}$$

*Proof of* (4.18). We may assume $r(M) > k$, since otherwise $\Pi_{k-1} = \emptyset$. Also, if $r(M) = k + 1$ and $M$ has a coloop $e$, then $|\Pi_{k-1}(M)| = |\Pi_{k-1}(M \setminus e)|$ (since every basis contains $e$) and $|\Pi_k(M)| = 2|\Pi_{k-1}(M \setminus e)|$, so we have (4.18).

So we may assume we are in the situation of Lemma 4.9 (either $r(M) \geq k + 2$ or $r(M) = k + 1$ and $M$ has no coloops). By Lemma 4.8, it suffices to show that for each $(A, B) \in \Pi_k$,

$$\sum_{(C,D)\sim(A,B)} \frac{1}{d(C,D)} \leq \frac{k}{k+1}. \tag{4.20}$$

Since $d_1(A, B) = d_2(B, A)$, we may assume that $2 \leq d_1(A, B) \leq d_2(A, B) \leq k$ (by Lemma 4.9(a)). If $d_1(A, B) = d_2(A, B)$, then Lemma 4.9(b) bounds the left side of (4.20) by $d_1(A, B)/(d_1(A, B) + 1) \leq k/(k + 1)$. Otherwise (i.e. if $d_1(A, B) < d_2(A, B)$), Lemma 4.9(c) bounds the left side of (4.20) by the right side of (4.19), which a little calculation shows—this is where we use the hypothesis that $k \leq 5$—to be at most $d_2(A, B)/(d_2(A, B) + 1) \leq k/(k + 1)$. □

# Chapter 5

# Competing urns

Our main task in this chapter is to prove Theorem 1.21. We begin with a natural extension of the definition of negative association given in Section 1.2. Call the joint distribution of the random variables $X_1, \dots, X_n$, with each $X_i$ taking values in a totally ordered set $\Omega_i$, *negatively associated* (NA) if $\mathcal{A} \downarrow \mathcal{B}$ for all increasing $\mathcal{A}, \mathcal{B} \subseteq \Omega_1 \times \cdots \times \Omega_n$ satisfying $\mathcal{A} \perp \mathcal{B}$ (where "increasing" is with respect to the product order on $\Omega_1 \times \cdots \times \Omega_n$ and $\mathcal{A} \perp \mathcal{B}$ again means no coordinate affects both $\mathcal{A}$ and $\mathcal{B}$). In particular, two real-valued random variables $X, Y$ are NA if

$$\{X \geq s\} \downarrow \{Y \geq t\} \text{ for all } s, t \in \mathbb{R}, \tag{5.1}$$

which is easily seen to be equivalent to

$$\mathsf{E} f(X) g(Y) \leq \mathsf{E} f(X) \mathsf{E} g(Y) \text{ for all increasing } f, g : \mathbb{R} \to \mathbb{R}.$$

We also call the joint distribution of $X_1, \dots, X_n$ (again, with each $X_i$ taking values in a totally ordered set) *conditionally negatively associated* (CNA) if every measure gotten by conditioning on the values of some of the $X_i$'s is NA.

The proof of Theorem 1.21 actually gives something a little more general, as follows. Suppose that for each $j \in [n]$ we are given a sequence $0 = a_0(j) < \cdots < a_{k_j}(j) = m+1$, and for $\sigma : [m] \to [n]$ set

$$\mathbf{x}_j(\sigma) = t \;\text{ iff }\; a_t(j) \leq B_j < a_{t+1}(j) \tag{5.2}$$

(recall $B_j = |\sigma^{-1}(j)|$).

**Theorem 5.1** *If the $\sigma(i)$'s are i.i.d. then the $\mathbf{x}_j$'s in* (5.2) *are CNA.*

Call the law of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ as in (5.2) a (generalized) *interval urn measure.*

The rest of this chapter is organized as follows. Section 5.1 reduces Theorem 5.1 to either of our two main inequalities, (5.8) and (5.12); these are equivalent in the case of ordinary urns but not in general. Section 5.2 presents a general result, inspired by the approach of [13], that easily specializes to give each of these, and Section 5.3 gives a different proof of (5.12), based on a graph-theoretic approach thought to be of independent interest. (It is only in the derivation of Theorem 5.1 from (5.8) that we need the $\sigma(i)$'s to be i.i.d.) Finally, in Section 5.4 we give examples justifying Proposition 1.25 and the remark following Theorem 1.21, and provide a counterexample to Conjecture 1.27. As before, we always tacitly assume that conditioning events have positive probability.

## 5.1 Setting up

For a nonnegative vector

$$\gamma = (\gamma_{ij} : i \in [m], j \in [n]), \tag{5.3}$$

$A \subseteq [m]$ and $K \subseteq [n]$, the probability measure on $K^A$ (functions from $A$ to $K$) *corresponding to* $\gamma$ is that given by

$$\Pr(\sigma) \propto W(\sigma) := \prod_{i \in A} \gamma_{i,\sigma(i)}. \tag{5.4}$$

Thus the r.v.'s $\sigma(i)$ are independent; they are i.i.d. if $\gamma_{ij}$ does not depend on $i$, in which case we write simply $\gamma_j$. We also use $\Pr^L$ ($L \subseteq [m]$) for the measure on $[n]^L$ corresponding to $\gamma$ (so $\Pr = \Pr^{[m]}$).

Let $\sigma \in [n]^{[m]}$ be as above (that is, given by (5.4)) and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ as in (5.2). Let $I \cup J \cup K$ be a partition of $[n]$ and $t_j \in \{0, \ldots, k_j - 1\}$ for $j \in K$, and set

$$Q := \{\mathbf{x}(\sigma) \equiv t \text{ on } K\} \qquad (= \{a_{t_j}(j) \leq |\sigma^{-1}(j)| < a_{t_j+1}(j)\ \forall j \in K\}), \tag{5.5}$$

$X = |\sigma^{-1}(I)|$ and $Y = |\sigma^{-1}(J)|$. The main point for the proof of Theorem 5.1 is

$$X, Y \text{ are NA given } Q, \tag{5.6}$$

given which we finish easily:

*Proof of Theorem* 5.1. With notation as above, let $\mathcal{A},\mathcal{B} \subseteq \Omega$ be increasing events determined by $I$ and $J$ (more precisely, by the values of the variables $\mathbf{x}_j$ indexed by $I$ and $J$) respectively. For Theorem 5.1 we should show $\mathcal{A} \downarrow \mathcal{B}$ given $Q$. Define $f,g : \mathbb{N} \to \mathbb{R}$ (where $\mathbb{N} = \{0,1,\ldots\}$) by $f(k) = \Pr(\mathcal{A}|X = k)$, $g(l) = \Pr(\mathcal{B}|Y = l)$. A standard coupling argument shows that $f$ and $g$ are increasing, whence, according to (5.6),

$$\Pr(\mathcal{A} \cap \mathcal{B}|Q) = \mathsf{E}[f(X)g(Y)|Q] \leq \mathsf{E}[f(X)|Q]\mathsf{E}[g(Y)|Q] = \Pr(\mathcal{A}|Q)\Pr(\mathcal{B}|Q).$$

(The first equality follows from the observation that $\mathcal{A},\mathcal{B}$ are independent given $X,Y$.)

□

We continue to condition on $Q$ and write $\mu_k$ for the law of $Y$ given $\{X = k\}$; that is,

$$\mu_k(l) = \Pr(Y = l|X = k). \tag{5.7}$$

We will actually prove

$$\frac{\mu_{k+1}(l+1)}{\mu_{k+1}(l)} \leq \frac{\mu_k(l+1)}{\mu_k(l)} \tag{5.8}$$

(whenever neither side is $0/0$, where we agree that $x/0 = \infty$ when $x > 0$), which is a strengthening of (5.6) once we rule out some pathologies. We recall the standard

**Definition 5.2** *$\mathcal{C} \subseteq \mathbb{N}^n$ is* convex *if $a, c \in \mathcal{C}$ and $a \leq b \leq c$ imply $b \in \mathcal{C}$.*

It will follow from Proposition 5.6 below that

$$\text{supp}(\Pr) := \{(k,l) : \Pr(X = k, Y = l) > 0\} \;\; \textit{is convex.} \tag{5.9}$$

Given this convexity, (5.8) implies that $Y$ is stochastically decreasing in $X$—that is,

$$\mu_{k+1}(Y \geq t) \leq \mu_k(Y \geq t) \;\; \forall t$$

(the easy implication is essentially Proposition 1.2 of [40])—which in turn easily implies (5.6).

Let

$$Z = |\sigma^{-1}(I \cup J)|. \tag{5.10}$$

When the $\sigma(i)$'s are i.i.d., an alternate way to specify $X$ and $Y$ is: let $Z$ be as in (5.10), $X \sim \mathrm{Bin}(Z, \alpha)$ and $Y = Z - X$, where $\alpha = \gamma_I/\gamma_{I\cup J}$ (with $\gamma_I = \sum_{i\in I} \gamma_i$) and $\mathrm{Bin}(Z, \alpha)$ is the binomial distribution with parameters $Z$ and $\alpha$.

In general, for $\nu$ the law of an $\mathbb{N}$-valued random variable $Z$ and $\alpha \in [0, 1]$, let $X = X_{\nu,\alpha} \sim \mathrm{Bin}(Z, \alpha)$, $Y = Y_{\nu,\alpha} = Z - X$ and, for lack of a better name, say $\nu$ is *binomially negatively associated* (BNA) if $X \downarrow Y$ for every $\alpha$. Call a nonnegative sequence $a = (a_i)_{i=0}^{\infty}$ *strongly log-concave* (SLC) if

$$ia_i^2 \geq (i + 1)a_{i-1}a_{i+1} \quad \text{for all } i \geq 1 \tag{5.11}$$

(that is, $(i!a_i)_{i=0}^{\infty}$ is log-concave), and say a probability measure $\nu$ on $\mathbb{N}$ is SLC if the sequence $(\nu(i))_{i=0}^{\infty}$ is. (For the equivalence of (5.8) and (5.12) below, it will be convenient to allow SLC sequences and measures to have internal zeros, even though the measures we are interested in here have no internal zeros.) A straightforward calculation shows that this is equivalent to saying that (5.8) holds for any $\alpha$, $X = X_{\nu,\alpha}$ and $Y = Y_{\nu,\alpha}$ (and $\mu_k$ as in (5.7)): since

$$\mu_k(l) = \frac{\nu(k + l)\Pr(X = k|Z = k + l)}{\Pr(X = k)} = \frac{\nu(k + l)\binom{k+l}{k}\alpha^k(1 - \alpha)^l}{\Pr(X = k)},$$

we may rewrite (5.8) as

$$\nu(k + l + 2)\binom{k+l+2}{k+1}\nu(k + l)\binom{k+l}{k} \leq \nu(k + l + 1)\binom{k+l+1}{k+1}\nu(k + l + 1)\binom{k+l+1}{k},$$

which is SLC for $\nu$. (If $\nu$ is Poisson—that is, if (5.11) holds with equality—then $X$ and $Y$ are independent Poisson r.v.'s and the inequalities (5.1) are equalities.) Thus, in the i.i.d. case, (5.8) is equivalent to saying that $Z$ as in (5.10) is SLC. That $Z$ is SLC again turns out to be true at the level of *generalized* urns; that is, for any $\gamma$ as in (5.3), $\sigma \in [n]^{[m]}$ with law given by (5.4), $Q$ as in (5.5) and $Z$ as in (5.10),

$$\textit{the law of } Z \textit{ is SLC.} \tag{5.12}$$

It seems interesting that both (5.8) and (5.12) are valid for generalized urns, though the equivalence that holds for i.i.d. balls disappears in the more general setting.

To repeat, (5.12) gives an alternate proof of (5.8)—and thus of (5.6) and Theorem 5.1—in the i.i.d. case. More generally, the preceding discussion shows that any SLC $\nu$

with no internal zeros is BNA (since it's easy to see that absence of internal zeros in $(\nu(i))$ is equivalent to (5.9) for $X = X_{\nu,\alpha}$, $Y = Y_{\nu,\alpha}$).

## 5.2 First proof

Let Pr be the measure on $[n]^{[m]}$ corresponding to some $\gamma$ (see the beginning of Section 5.1), and for $a, b \in \mathbb{N}^{n-1}$ and $k \in \mathbb{N}$, set

$$p(k, a, b) = \Pr(B_n = k | B_j \in [a_j, b_j] \ \forall j \in [n-1]) \tag{5.13}$$

(recalling that $B_j = |\sigma^{-1}(j)|$). The main result of this section is

**Theorem 5.3** *With notation as above,*

(a) $\dfrac{p(k+1,a,b)}{p(k,a,b)}$ *is decreasing (i.e. nonincreasing) in* $(a,b)$

(b) $\dfrac{p(k+1,a,b)}{p(k,a,b)} \leq \dfrac{k}{k+1} \cdot \dfrac{p(k,a,b)}{p(k-1,a,b)}$

*(where we say nothing about the case* $0/0$ *and agree that* $x/0 = \infty$ *when* $x > 0$*).*

This is somewhat related to Theorem 1.23, whose proof is sketched at the end of this section.

Before proving Theorem 5.3 we observe a few consequences—in particular that it yields both (5.8) and (5.12)—and give the promised Proposition 5.6.

**Corollary 5.4** *For fixed* $a, b \in \mathbb{N}^{[n-1]}$ *the sequence* $\{p(k,a,b)\}$ *is SLC.*

This is just Theorem 5.3(b). Note that it includes (5.12): with notation as in (5.5) define

$$\gamma'_{ij} = \begin{cases} \gamma_{ij} & \text{if } j \in K = [n-1] \\ \sum\{\gamma_{ij} : j \in I \cup J\} & \text{if } j = n; \end{cases}$$

then, for $\sigma$ chosen according to $\gamma'$, the law of $|\sigma^{-1}(n)|$ is the same as that of $Z$ in (5.10), and the conditioning in (5.5) is of the same type as that in (5.13). $\square$

A second application of Theorem 5.3, this time of part (a), gives (5.8). Here we should take

$$\gamma'_{ij} = \begin{cases} \gamma_{ij} & \text{if } j \in K = [n-2] \\ \sum\{\gamma_{ij} : j \in I\} & \text{if } j = n-1 \\ \sum\{\gamma_{ij} : j \in J\} & \text{if } j = n, \end{cases}$$

let the $k$ of Theorem 5.3 be the $l$ of (5.8), and, in (5.13), compare the pairs $(a, b)$ corresponding to $Q \wedge \{B_{n-1} = k\}$ and $Q \wedge \{B_{n-1} = k+1\}$ (where $Q$ is as in (5.5)). $\square$

For $f, a \in \mathbb{N}^n$, let $\mathcal{M}_f(a) = \{\sigma \in [n]^{[m]} : |\sigma^{-1}(j)| \in [a_j, a_j + f_j] \ \forall j \in [n]\}$ and $M_f(a) = \Pr(\mathcal{M}_f(a))$. Though we won't use the next corollary (but see the remark following Corollary 5.9), it seems natural and worth mentioning.

**Corollary 5.5** *For each $f \in \mathbb{N}^n$, $M = M_f$ satisfies the negative lattice condition:*

$$M(a)M(c) \geq M(a \vee c)M(a \wedge c) \quad \forall a, c \in \mathbb{N}^n. \tag{5.14}$$

This is more or less immediate from Theorem 5.3 once we have the next little observation, which, as noted earlier, also gives (5.9) and absence of internal zeros in the law of $Z$ in (5.10).

**Proposition 5.6** *For any $f$ and $M_f$ as above, the support of $M = M_f$ is convex.*

*Proof.* This will follow easily from

*Claim.* For any $\sigma, \tau \in [n]^{[m]}$ with $\Pr(\sigma), \Pr(\tau) > 0$ and $i \in [n]$ with $|\sigma^{-1}(i)| > |\tau^{-1}(i)|$, there are $j \in [n]$ and $\rho \in [n]^{[m]}$ with $\Pr(\rho) > 0$, $|\sigma^{-1}(j)| < |\tau^{-1}(j)|$ and

$$|\rho^{-1}(k)| = \begin{cases} |\sigma^{-1}(i)| - 1 & \text{if } k = i \\ |\sigma^{-1}(j)| + 1 & \text{if } k = j \\ |\sigma^{-1}(k)| & \text{if } k \in [n] \setminus \{i, j\}. \end{cases}$$

This is a standard type of graph-theoretic observation: regarding $\sigma$ and $\tau$ as edge sets of bipartite graphs on $[m] \cup [n]$ in the natural way,[1] we need a path with edges alternately from $\sigma \setminus \tau$ and $\tau \setminus \sigma$ that begins with a $\sigma$-edge at $i$ and ends with a $\tau$-edge at some $j$ as

---

[1] We pretend $[m] \cap [n] = \emptyset$.

above. (We then get $\rho$ by switching $\sigma$ and $\tau$ on this path.) We omit the routine proof that such a path must exist.

To prove Proposition 5.6, we should show that for all $a, b, c \in \mathbb{N}^n$ with $a < b < c$ and $a, c \in \text{supp}(M)$, we also have $b \in \text{supp}(M)$. Of course it suffices to show this when there is some $i \in [n]$ with $b_i = c_i - 1$ and $b_k = c_k$ for all $k \neq i$. Choose $\tau \in \mathcal{M}(a) := \mathcal{M}_f(a)$ and $\sigma \in \mathcal{M}(c)$ with $\Pr(\tau), \Pr(\sigma) > 0$. We assume $|\sigma^{-1}(i)| = c_i + f_i$, since otherwise $\sigma \in \mathcal{M}(b)$ and we are finished. Letting $j, \rho$ be as in the claim (note $|\tau^{-1}(i)| < c_i + f_i$), we have

$$|\rho^{-1}(i)| = b_i + f_i$$

and

$$|\rho^{-1}(j)| = |\sigma^{-1}(j)| + 1 \in [c_j + 1, a_j + f_j] \subseteq [b_j, b_j + f_j],$$

whence $\rho \in \mathcal{M}(b)$ and $b \in \text{supp}(M)$. □

*Proof of Corollary* 5.5. It is easy to see (and standard) that convexity of $M$ (given by Proposition 5.6) implies that it's enough to prove (5.14) when both sides of (5.14) are positive and there are indices $i$ and $j$ with $a_i = c_i - 1$, $a_j = c_j + 1$, and $a_k = c_k$ for all $k \neq i, j$. In this case—assuming, w.l.o.g., that $i = n - 1$ and $j = n$—we set

$$p_1(k) = \Pr(B_n = k | B_l \in [a_l, a_l + f_l] \ \forall l \in [n-1])$$

and

$$p_2(k) = \Pr(B_n = k | B_l \in [c_l, c_l + f_l] \ \forall l \in [n-1]).$$

Then (5.14) is

$$\left(\sum_{k=c_n+1}^{c_n+f_n+1} p_1(k)\right)\left(\sum_{k=c_n}^{c_n+f_n} p_2(k)\right) \geq \left(\sum_{k=c_n+1}^{c_n+f_n+1} p_2(k)\right)\left(\sum_{k=c_n}^{c_n+f_n} p_1(k)\right)$$

and follows immediately from

$$p_1(k)p_2(l) \geq p_1(l)p_2(k) \ \textit{whenever} \ k \geq l,$$

which is a consequence of Theorem 5.3(a) (and Proposition 5.6). □

The proof of Theorem 5.3 resembles that of Theorem 33 in [13], and is based on

**Observation 5.7** *For any $i \in [n]$, $k \in \mathbb{N}$, and event $Q$ determined by $(\sigma^{-1}(j) : j \neq i)$,*

$$\Pr(B_i = k+1, Q) = \frac{1}{k+1} \sum_{l \in [m]} \gamma_{li} \Pr^{[m]\setminus\{l\}}(B_i = k, Q).$$

(Recall $\Pr^L$ was defined at the beginning of Section 5.1.) We also use the trivial

$$\min_i \frac{\alpha_i}{\beta_i} \leq \frac{\alpha_1 + \cdots + \alpha_k}{\beta_1 + \cdots + \beta_k} \leq \max_i \frac{\alpha_i}{\beta_i} \tag{5.15}$$

(for all $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \geq 0$ with $\beta_1 + \cdots + \beta_k > 0$, where $x/0 = \infty$ when $x > 0$).

*Proof of Theorem* 5.3. We proceed by induction on $m$, omitting the easy base cases with $m = 1$. For (a), it's enough to show the ratio in question does not increase when we increase a single entry—w.l.o.g. the $(n-1)$st—of one of $a, b$. Thus, by (5.15), it suffices to show that

$$\frac{\Pr(B_n = k+1 | B_{n-1} = t, R)}{\Pr(B_n = k | B_{n-1} = t, R)} \text{ is nonincreasing in } t,$$

where $R = \{a_j \leq B_j \leq b_j \ \forall j \in [n-2]\}$; by Proposition 5.6, it is enough to show

$$\frac{\Pr(B_n = k+1 | B_{n-1} = t+1, R)}{\Pr(B_n = k | B_{n-1} = t+1, R)} \leq \frac{\Pr(B_n = k+1 | B_{n-1} = t, R)}{\Pr(B_n = k | B_{n-1} = t, R)} \tag{5.16}$$

for all $t$ for which all four probabilities appearing in (5.16) are positive.

Using Observation 5.7, the left side of (5.16) is

$$\frac{\sum_{l \in [m]} \gamma_{l,n-1} \Pr^{[m]\setminus\{l\}}(B_n = k+1, B_{n-1} = t, R)}{\sum_{l \in [m]} \gamma_{l,n-1} \Pr^{[m]\setminus\{l\}}(B_n = k, B_{n-1} = t, R)},$$

which, by (5.15), is at most

$$\max_{l \in [m]} \frac{\Pr^{[m]\setminus\{l\}}(B_n = k+1 | B_{n-1} = t, R)}{\Pr^{[m]\setminus\{l\}}(B_n = k | B_{n-1} = t, R)}.$$

Setting $Q = \{B_{n-1} = t\} \wedge R$ and assuming (w.l.o.g.) that the maximum occurs at $l = m$, we will have (5.16) if we show

$$\frac{\Pr(B_n = k+1 | Q)}{\Pr(B_n = k | Q)} \geq \frac{\Pr^{[m-1]}(B_n = k+1 | Q)}{\Pr^{[m-1]}(B_n = k | Q)}. \tag{5.17}$$

Now

$$\frac{\Pr(B_n = k+1 | Q)}{\Pr(B_n = k | Q)} = \frac{\sum_{j \in [n]} \Pr(\sigma(m) = j | Q) \Pr(B_n = k+1 | Q, \sigma(m) = j)}{\sum_{j \in [n]} \Pr(\sigma(m) = j | Q) \Pr(B_n = k | Q, \sigma(m) = j)},$$

so that (5.17) will follow (again using (5.15)) from

$$\frac{\Pr(B_n = k+1|Q, \sigma(m)=j)}{\Pr(B_n = k|Q, \sigma(m)=j)} \geq \frac{\Pr^{[m-1]}(B_n = k+1|Q)}{\Pr^{[m-1]}(B_n = k|Q)} \qquad \text{for all } j \in [n] \tag{5.18}$$

(where, again, "for all $j \in [n]$" really includes only those for which $\Pr(Q, \sigma(m) = j) > 0$).

There are three cases to consider. If $j = n$, the left side of (5.18) is

$$\frac{\Pr^{[m-1]}(B_n = k|Q)}{\Pr^{[m-1]}(B_n = k-1|Q)},$$

which is at least the right side of (5.18) by (part (b) of) our induction hypothesis. If $j = n-1$, the left side of (5.18) is

$$\frac{\Pr^{[m-1]}(B_n = k+1|B_{n-1} = t-1, R)}{\Pr^{[m-1]}(B_n = k|B_{n-1} = t-1, R)},$$

which is at least the right side of (5.18) by (part (a) of) the induction hypothesis. Finally, if $j \neq n-1, n$, the left side of (5.18) is

$$\frac{\Pr^{[m-1]}(B_n = k+1|B_{n-1} = t, R^*)}{\Pr^{[m-1]}(B_n = k|B_{n-1} = t, R^*)},$$

where $R^*$ is obtained from $R$ by replacing the condition $a_j \leq B_j \leq b_j$ by the condition $a_j - 1 \leq B_j \leq b_j - 1$; again this is at least the right side of (5.18) by part (a) of the induction hypothesis.

We now turn to (b) and set $Q = \{a_j \leq B_j \leq b_j \;\; \forall j \in [n-1]\}$. Then we have, again using Observation (5.7) and (5.15),

$$\begin{aligned}
\frac{\Pr(B_n = k+1|Q)}{\Pr(B_n = k|Q)} &= \frac{k\sum_{l\in[m]} \gamma_{ln} \Pr^{[m]\setminus\{l\}}(B_n = k, Q)}{(k+1)\sum_{l\in[m]} \gamma_{ln} \Pr^{[m]\setminus\{l\}}(B_n = k-1, Q)} \\
&\leq \max_{l\in[m]} \frac{k \Pr^{[m]\setminus\{l\}}(B_n = k, Q)}{(k+1)\Pr^{[m]\setminus\{l\}}(B_n = k-1, Q)} \\
&\overset{\text{w.l.o.g.}}{=} \frac{k \Pr^{[m-1]}(B_n = k, Q)}{(k+1)\Pr^{[m-1]}(B_n = k-1, Q)}
\end{aligned}$$

(where, again by Proposition 5.6, we may assume that $\Pr(B_n = r|Q)$ is positive for $r \in \{k-1, k, k+1\}$); so we will be done if we can show

$$\frac{\Pr^{[m-1]}(B_n = k|Q)}{\Pr^{[m-1]}(B_n = k-1|Q)} \leq \frac{\Pr(B_n = k|Q)}{\Pr(B_n = k-1|Q)}. \tag{5.19}$$

Proceeding as in the proof of part (a), we may rewrite

$$\frac{\Pr(B_n = k|Q)}{\Pr(B_n = k-1|Q)} = \frac{\sum_{j\in[n]} \Pr(\sigma(m) = j|Q)\Pr(B_n = k|Q, \sigma(m) = j)}{\sum_{j\in[n]} \Pr(\sigma(m) = j|Q)\Pr(B_n = k-1|Q, \sigma(m) = j)};$$

so for (5.19) it is enough to show that, for each $j \in [n]$,

$$\frac{\Pr(B_n = k|Q, \sigma(m) = j)}{\Pr(B_n = k-1|Q, \sigma(m) = j)} \geq \frac{\Pr^{[m-1]}(B_n = k|Q)}{\Pr^{[m-1]}(B_n = k-1|Q)},$$

which, as did (5.18), follows easily from our induction hypothesis (here we only need to consider the two cases $j = n$ and $j \neq n$). □

*Proof of Theorem 1.23.* This is similar to the above proof of Theorem 5.3, so we just give a brief sketch. We again proceed by induction on $m$ and note that it's enough to show $p(\mathcal{A}, a, b)$ does not increase when we increase a single entry—w.l.o.g. the $(n-1)$st—of one of $a, b$; that is, with $Q = \{a_j \leq B_j \leq b_j$ for all $j \in [n-2]\}$, we should show

$$\Pr(\sigma^{-1}(n) \in \mathcal{A}|B_{n-1} = t, Q) \text{ is nonincreasing in } t.$$

Using Observation 5.7 and (5.15), we may assume without loss of generality that

$$\Pr(\sigma^{-1}(n) \in \mathcal{A}|B_{n-1} = t+1, Q) \leq \Pr^{[m-1]}(\sigma^{-1}(n) \in \mathcal{A}|B_{n-1} = t, Q),$$

and it suffices to show

$$\Pr^{[m-1]}(\sigma^{-1}(n) \in \mathcal{A}|B_{n-1} = t, Q) \leq \Pr(\sigma^{-1}(n) \in \mathcal{A}|B_{n-1} = t, Q, \sigma(m) = j)$$

for all $j \in [n]$. But for $j \neq n$ this is true by the induction hypothesis and for $j = n$ it follows from the fact that $\mathcal{A}$ is increasing. □

## 5.3 A graphical approach

We will return to (5.12) soon, but begin with a natural and seemingly new graph theoretic statement which we regard as the main point of this section. Given a multigraph $G$ on vertex set $V$ and $a, b \in \mathbb{N}^V$, let $\mathcal{O}(a, b) = \mathcal{O}_G(a, b)$ be the set of orientations of $G$ for which

$$d^+(x) \geq a_x \text{ and } d^-(x) \geq b_x \quad \text{for all } x \in V,$$

and $N(a,b) = N_G(a,b) = |\mathcal{O}(a,b)|$. Here $d^+$ and $d^-$ are, as usual, out- and in-degrees. We will also use $d_x$ for the degree of $x$ in $G$. Note that we regard a loop (at $x$, say) as having two orientations, each of which contributes 1 to each of $d^+(x)$ and $d^-(x)$.

**Lemma 5.8** *If $a, b, r, s \in \mathbb{N}^V$ satisfy*

$$a \geq r, s \quad \textit{and} \quad a + b \geq r + s$$

*(where the inequalities are with respect to the product order on $\mathbb{N}^V$), then*

$$N(a,b) \leq N(r,s). \tag{5.20}$$

Of course the idea is that it's harder to satisfy a set of demands that always requires large out-degrees than one for which these requirements are mixed. For the sake of comparison, let us also mention the specialization of Corollary 5.5 to the present situation:

**Corollary 5.9** *If $a + b = r + s$ then*

$$N(a,b)N(r,s) \geq N(a \vee r, b \wedge s)N(a \wedge r, b \vee s). \tag{5.21}$$

*Proof.* Interpret vertices of $G$ as urns and edges as balls, and assume that for each edge (ball) $e$ we have $\gamma_{ex} = 1$ or 0 according to whether $x$ is or is not an end of $e$. Then (5.21) is just Corollary 5.5 with $f = d - a - b$ $(= d - r - s)$, where $d = (d_x : x \in V(G))$ is the vector of degrees. □

*Remark.* It's possible to derive Lemma 5.8 from Corollary 5.9, but the derivation is about as difficult as the following direct argument.

*Proof of Lemma 5.8.* We proceed by induction on

$$\varphi(G, a, b) := |E(G)| + \sum_{x \in V} (d_x - a_x - b_x),$$

calling $x \in V$ *saturated* if $a_x + b_x = d_x$. Since $N$ is decreasing in each of its arguments, we may assume $a + b = r + s$ (otherwise we increase $r$ or $s$ so as not to increase $\varphi$).

Suppose first that there is at least one saturated vertex, $x$. We may assume there are no loops at $x$, since otherwise (5.20) follows easily from the induction hypothesis applied to the graph gotten from $G$ by deleting such loops. Let $\alpha = a_x, \beta = b_x, \rho = r_x, \sigma = s_x$, and let $X = \{e_1, \ldots, e_{\alpha+\beta}\}$ be the set of edges incident to $x$.

Consider a set $\pi$ consisting of $\beta$ pairs $\{e_i, e_j\} \subseteq X$, with the $2\beta$ edges appearing in $\pi$ distinct, and, say, $y_i$ the vertex joined to $x$ by $e_i$ (so the $y_i$'s need not be distinct). Let $G(\pi)$ be the graph with vertex set $V \setminus \{x\}$ and edge set $E(G) \setminus X \cup \{e_{ij} : \{e_i, e_j\} \in \pi\}$, where $e_{ij}$ joins $y_i$ and $y_j$. Let $U(\pi)$ be the set of edges in $X$ not belonging to pairs from $\pi$, and $U_z(\pi)$ the set of edges of $U(\pi)$ incident to $z$.

Define $a^\pi, b^\pi \in \mathbb{N}^{V(G(\pi))}$ by

$$a^\pi_z = a_z \text{ and } b^\pi_z = \max\{b_z - |U_z(\pi)|, 0\} \text{ for all } z \in V \setminus \{x\}.$$

For each $\pi$ as above and $T \in \binom{U(\pi)}{\rho-\beta}$ (where, as usual, $\binom{S}{k} = \{T \subseteq S : |T| = k\}$), define $r^{\pi,T}, s^{\pi,T} \in \mathbb{N}^{V(G(\pi))}$ by

$$r^{\pi,T}_z = \max\{r_z - |U_z(\pi) \setminus T|, 0\}$$

and

$$s^{\pi,T}_z = \max\{s_z - |U_z(\pi) \cap T|, 0\}$$

for all $z \in V \setminus \{x\}$.

Each $\sigma \in \mathcal{O}_{G(\pi)}(a^\pi, b^\pi)$ maps naturally to a (unique) $\hat{\sigma} \in \mathcal{O}_G(a, b)$, namely: $\hat{\sigma}$ agrees with $\sigma$ on $E(G - x)$; orients all edges of $U(\pi)$ away from $x$; and orients $e_i$ from $y_i$ to $x$ and $e_j$ from $x$ to $y_j$ whenever $\sigma$ orients $e_{ij}$ from $y_i$ to $y_j$ (where, when $y_i = y_j$, we interpret one orientation of the loop $e_{ij}$ as $y_i \to y_j$ and the other as $y_j \to y_i$). Since each $\tau \in \mathcal{O}_G(a, b)$ is in the range of this map for exactly $\binom{\alpha}{\beta}\beta!$ choices of $\pi$, we have

$$N(a, b) = \frac{1}{\binom{\alpha}{\beta}\beta!} \sum_\pi N_{G(\pi)}(a^\pi, b^\pi). \tag{5.22}$$

Similarly,

$$N(r, s) = \frac{1}{\binom{\rho}{\beta}\binom{\sigma}{\beta}\beta!} \sum_\pi \sum_{T \in \binom{U(\pi)}{\rho-\beta}} N_{G(\pi)}(r^{\pi,T}, s^{\pi,T}). \tag{5.23}$$

Since $a^\pi + b^\pi \geq r^{\pi,T} + s^{\pi,T}$ and $a^\pi \geq r^{\pi,T}, s^{\pi,T}$, it follows from the induction hypothesis that

$$N_{G(\pi)}(a^\pi, b^\pi) \leq N_{G(\pi)}(r^{\pi,T}, s^{\pi,T}) \text{ for all } \pi \text{ and } T \in \binom{U(\pi)}{\rho-\beta}. \tag{5.24}$$

(Note that $\varphi(G(\pi), a^\pi, b^\pi) < \varphi(G, a, b)$, since $|E(G(\pi))| < |E(G)|$ and, for $z \in V \setminus \{x\}$, $a_z^\pi + b_z^\pi \geq a_z + b_z - |U_z(\pi)|$, while the degree of $z$ in $G(\pi)$ is $d_z - |U_z(\pi)|$.) Combining (5.23), (5.24) and (5.22), we have

$$N(r,s) \geq \frac{1}{\binom{\rho}{\beta}\binom{\sigma}{\beta}\beta!} \sum_{\pi} \binom{\alpha-\beta}{\rho-\beta} N_{G(\pi)}(a^\pi, b^\pi) = \frac{\alpha!\beta!}{\rho!\sigma!} N(a,b) \geq N(a,b),$$

where the last inequality follows from the assumptions $\alpha \geq \rho, \sigma$ and $\alpha + \beta = \rho + \sigma$.

So we may assume there are no saturated vertices. In this case we fix $x \in V$ with $a_x > b_x$. (Of course if there is no such vertex, then $a = b = r = s$ and (5.20) is an equality.) For $\gamma, \delta \in \mathbb{N}$ let $N'(\gamma, \delta)$ be the number of orientations of $G$ with

$$(d_y^+, d_y^-) \geq \begin{cases} (a_y, b_y) & \text{if } y \neq x \\ (\gamma, \delta) & \text{if } y = x, \end{cases}$$

and let $N''(\gamma, \delta)$ be defined analogously with $(r, s)$ in place of $(a, b)$. Let $\alpha = a_x$, $\beta = b_x$, $\rho = r_x$, and $\sigma = s_x$, so that (5.20) is

$$N'(\alpha, \beta) \leq N''(\rho, \sigma). \tag{5.25}$$

By induction we have

$$N'(\gamma, \delta) \leq N''(\eta, \xi) \text{ whenever } \gamma \geq \eta, \xi \text{ and } \gamma + \delta = \eta + \xi > \alpha + \beta. \tag{5.26}$$

We apply this to the identity

$$N'(\alpha, \beta) = N'(\alpha, \beta + 1) + N'(d_x - \beta, \beta). \tag{5.27}$$

If $\alpha > \sigma$, then, by (5.26), the right side of (5.27) is at most

$$N''(\rho, \sigma + 1) + N''(d_x - \sigma, \sigma) = N''(\rho, \sigma).$$

If $\alpha > \rho$, then, again using (5.26), the right side of (5.27) is at most

$$N''(\rho + 1, \sigma) + N''(\rho, d_x - \rho) = N''(\rho, \sigma).$$

(And, since $\alpha > \beta$, we have at least one of $\alpha > \sigma$, $\alpha > \rho$.) $\square$

The next result isolates (and generalizes) the main point in the derivation of (5.12) from Lemma 5.8. We consider a hypergraph $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ on a set $W$ of size $2l$, where

(i) the edges of $\mathcal{H}_1$ are pairwise disjoint and

(ii) the edges of $\mathcal{H}_2$ are of size 2 and pairwise disjoint.

Let $S$ be the set of vertices of $\mathcal{H}$ not covered by edges of $\mathcal{H}_2$, and $|S| = 2t$. Given $\alpha : \mathcal{H}_1 \to \mathbb{N}$, let $N_i$ be the number of partitions $(X, Y)$ of $W$ with each of $X, Y$ a vertex cover of $\mathcal{H}_2$, each of $|X \cap H|$, $|Y \cap H|$ at least $\alpha_{_H}$ (for all $H \in \mathcal{H}_1$), and $|X| = i$.

**Lemma 5.10** *In the above situation,* $N_l \geq \frac{t+1}{t} N_{l+1}$.

*Proof.* For $i \in \mathbb{N}$ and $\pi$ a collection of $t-1$ disjoint 2-sets contained in $S$, let $\mathcal{N}_i(\pi)$ be the set of partitions as above for which each of $X, Y$ also covers the edges of $\pi$, and set $N_i(\pi) = |\mathcal{N}_i(\pi)|$. We assert that (for each $\pi$)

$$N_l(\pi) \geq 2N_{l+1}(\pi). \tag{5.28}$$

This implies the proposition since (as is easily seen)

$$N_l = \frac{1}{t \cdot t!} \sum_{\pi} N_l(\pi)$$

and

$$N_{l+1} = \frac{1}{\binom{t+1}{2}(t-1)!} \sum_{\pi} N_{l+1}(\pi).$$

For the proof of (5.28) let $x, y$ be the two vertices of $W$ not contained in members of $\mathcal{H}_2' := \mathcal{H}_2 \cup \pi$. Noting that $(X, Y) \in \mathcal{N}_{l+1}(\pi)$ implies $x, y \in X$, we may regard $(X, Y)$ as an orientation of $\mathcal{H}_2'$, where orienting $\{u, v\}$ from $u$ to $v$ corresponds to putting $u$ in $X$ (and $v$ in $Y$). The orientations corresponding to $(X, Y)$'s from $\mathcal{N}_{l+1}$ are those for which, for each $H \in \mathcal{H}_1$,

$$d^+(H) \geq \alpha_{_H} - |H \cap \{x, y\}| \quad \text{and} \quad d^-(H) \geq \alpha_{_H},$$

where, for the given orientation, $d^+(H)$ (resp. $d^-(H)$) is the number of oriented edges whose tails (resp. heads) lie in $H$.

If we let $G$ be the multigraph gotten from $\mathcal{H} \cup \pi$ by collapsing each $H \in \mathcal{H}_1$ to a single vertex (so for example, any $\{u, v\} \in \mathcal{H}_2'$ contained in some $H \in \mathcal{H}_1$ becomes a loop in $G$), then the above discussion says that $N_{l+1}(\pi) = N_G(a,b)$ (see Lemma 5.8 for the notation), where

$$a_z = \alpha_{_H} - |H \cap \{x,y\}| \quad \text{and} \quad b_z = \alpha_{_H}$$

if $z$ is the vertex of $G$ corresponding to $H \in \mathcal{H}_1$, and $a_z = b_z = 0$ if $z$ is not of this type.

A similar discussion shows that $N_l(\pi) = N_G(r,s) + N_G(s,r)$, where

$$r_z = \alpha_{_H} - \mathbf{1}_{\{x \in H\}} \quad \text{and} \quad s_z = \alpha_{_H} - \mathbf{1}_{\{y \in H\}}$$

if $z$ is the vertex of $G$ corresponding to $H \in \mathcal{H}_1$, and $r_z = s_z = 0$ if $z$ is not of this type. (For example, $N_G(r,s)$ counts pairs $(X,Y)$ with $x \in X$ (and $y \in Y$).)

Finally, Lemma 5.8 gives $N_G(r,s), N_G(s,r) \geq N_G(a,b)$, so we have (5.28). (Strictly speaking we may be applying Lemma 5.8 with some negative entries in $b, r$ and/or $s$; but it's easy to see that this slightly more general version follows from the lemma as stated.) $\square$

*Proof of* (5.12). For $A \subseteq [m]$ and $\sigma \in K^A$, say $\sigma \in Q$ if it satisfies the conditions in (5.5), which we now rewrite

$$S_j \leq |\sigma^{-1}(j)| < T_j \quad \forall j \in K, \tag{5.29}$$

where $S_j = a_{t_j}(j)$ and $T_j = a_{t_j+1}(j)$. (This extends the $Q$ of (5.5), which was a subset of $[n]^{[m]}$.)

Write $\sigma \sim A$ if $\sigma \in K^A \cap Q$, and $\sigma \sim l$ if $\sigma \sim A$ for some $A$ of size $l$. For $i \in [m]$ let $\gamma_{i0} = \sum\{\gamma_{ij} : j \in I \cup J\}$ and set $X(A) = \prod\{\gamma_{i0} : i \in [m] \setminus A\}$.

We may then rewrite (5.12) as

$$\begin{aligned}\sum_{|A|=l} \sum_{|B|=l} X(A)X(B) \sum\sum\{W(\sigma)W(\tau) : \sigma \sim A, \tau \sim B\} \\ \geq \frac{m-l+1}{m-l} \sum_{|C|=l+1} \sum_{|D|=l-1} X(C)X(D) \sum\sum\{W(\alpha)W(\beta) : \alpha \sim C, \beta \sim D\}.\end{aligned} \tag{5.30}$$

Again regard each of $\sigma, \tau, \alpha, \beta$ in (5.30) as a bipartite graph on the vertex set $[m] \cup K$ in the natural way (e.g. the edge set for $\sigma \sim A$ is $\{\{i, \sigma(i)\} : i \in A\}$, where we again pretend that $[m] \cap K = \emptyset$). Then for each pair $(\sigma, \tau)$ appearing in (5.30) the *multiset* union $G = \sigma \cup \tau$ is a bipartite multigraph with exactly $2l$ edges and

$$d_G(i) \le 2 \quad \text{for all } i \in [m]$$

(and similarly for pairs $(\alpha, \beta)$). Setting $X(G) = \prod_{i \in [m]} \gamma_{i0}^{2-d_G(i)}$ we may rewrite (5.30) as

$$\sum_G X(G) \sum\sum \{W(\sigma)W(\tau) : \sigma \cup \tau = G, \sigma \sim l, \tau \sim l\}$$
$$\ge \frac{m-l+1}{m-l} \sum_G X(G) \sum\sum \{W(\alpha)W(\beta) : \alpha \cup \beta = G, \alpha \sim l+1, \beta \sim l-1\}.$$

It is thus enough to show that for each fixed $G$ we have the corresponding inequality for the inner double sums, i.e.

$$\sum\sum \{W(\sigma)W(\tau) : \sigma \cup \tau = G, \sigma \sim l, \tau \sim l\}$$
$$\ge \frac{m-l+1}{m-l} \sum\sum \{W(\alpha)W(\beta) : \alpha \cup \beta = G, \alpha \sim l+1, \beta \sim l-1\}. \tag{5.31}$$

This has the advantage that the weights no longer play a role, since for $\sigma, \dots, \beta$ as in (5.31), we have

$$W(\sigma)W(\tau) = W(\alpha)W(\beta);$$

so we will have (5.31) if we show

$$N_l \ge \tfrac{m-l+1}{m-l} N_{l+1}, \tag{5.32}$$

where $N_i = N_i(G)$ is the number of partitions $G = \gamma \cup \delta$ with $\gamma \sim i$.

Now let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ be the hypergraph with vertex set $W = E(G)$,

$$\mathcal{H}_1 = \{H_j : j \in K\},$$

where $H_j = \{e \in W : j \in e\}$, and

$$\mathcal{H}_2 = \{\{e, f\} : e \ne f,\ e \text{ and } f \text{ have the same end in } [m]\}.$$

Define $\alpha : \mathcal{H}_1 \to \mathbb{N}$ by

$$\alpha_{H_j} = \max\{S_j, d_G(j) - T_j + 1\}$$

(recall $S_j, T_j$ were defined following (5.29)). We are then in the situation of Lemma 5.10: a partition $W = X \cup Y$ with $|X| = i$ (and $|Y| = 2l - i$) as in the proposition is the same thing as a partition $G = \gamma \cup \delta$ with $\gamma \sim i$ and $\delta \sim 2l - i$.

Defining $t$ as in the proposition, we have $2t = |\{i \in [m] : d_G(i) = 1\}| \leq 2(m - l)$ (since $d_G(i) = 1$ (for $i \in [m]$) requires that in any partition $\gamma \cup \delta$ of $E(G)$, at least one of $\gamma, \delta$ fails to cover $i$, and the number of such $i$ is at most $2(m-l)$). Thus $\frac{t+1}{t} \geq \frac{m-l+1}{m-l}$ and (5.32) follows. □

*Remark.* It's possible to derive Theorem 5.3 from Lemma 5.8, as follows. First observe that part (b) of Theorem 5.3 is equivalent to (5.12), as in the remark following Corollary 5.4 (even for generalized urns). To get part (a), it suffices to show that the ratio in question does not increase when we increase a single entry—w.l.o.g. the $(n-1)$st—of one of $a, b$, whence, by (5.15) and Proposition 5.6, it is enough to show

$$\begin{aligned} &\Pr(B_n = k+1, B_{n-1} = l, R)\Pr(B_n = k, B_{n-1} = l+1, R) \\ &\qquad \geq \Pr(B_n = k+1, B_{n-1} = l+1, R)\Pr(B_n = k, B_{n-1} = l, R), \end{aligned} \tag{5.33}$$

where $R = \{a_j \leq B_j \leq b_j \ \forall j \in [n-2]\}$. Let $I = \{n\}$, $J = \{n-1\}$, and $K = [n-2]$; and, for $\sigma \in [n]^{[m]}$, write $\sigma \sim (s,t)$ if $|\sigma^{-1}(n)| = s$, $|\sigma^{-1}(n-1)| = t$, and $\sigma \in R$. Then (5.33) is

$$\begin{aligned} &\sum\sum\{W(\sigma)W(\tau) : \sigma \sim (k+1, l), \tau \sim (k, l+1)\} \\ &\qquad \geq \sum\sum\{W(\alpha)W(\beta) : \alpha \sim (k+1, l+1), \beta \sim (k, l)\}. \end{aligned}$$

Proceeding as in the above proof of (5.12) and regarding each of $\sigma, \tau, \alpha, \beta$ as a bipartite graph on $[m] \cup [n]$, it suffices to show that, for each fixed multigraph $G$,

$$\begin{aligned} &|\{(\sigma,\tau) : \sigma \cup \tau = G, \sigma \sim (k+1, l), \tau \sim (k, l+1)\}| \\ &\qquad \geq |\{(\alpha,\beta) : \alpha \cup \beta = G, \alpha \sim (k+1, l+1), \beta \sim (k, l)\}| \end{aligned} \tag{5.34}$$

(since, as earlier, the weights no longer play a role). Let $H$ be the multigraph on $[n]$ with edge set $\{e_x : x \in [m]\}$, where $e_x$ joins $i$ and $j$ if $x$ was joined to $i, j$ in $G$, and identify a partition $E(G) = \gamma \cup \delta$ with the orientation of $H$ gotten by directing $e_x$ from

$i$ to $j$ whenever $\{x,i\} \in \gamma$ and $\{x,j\} \in \delta$. The orientations corresponding to pairs $(\sigma,\tau)$ in (5.34) are those satisfying

$$d^+(j), d^-(j) \geq \max\{a_j, d_H(j) - b_j\} \quad \text{for all } j \in [n-2] \tag{5.35}$$

and

$$d^+(n) \geq k+1, d^-(n) \geq k, d^+(n-1) \geq l, d^-(n-1) \geq l+1,$$

while the orientations corresponding to pairs $(\alpha,\beta)$ are those satisfying (5.35) and

$$d^+(n) \geq k+1, d^-(n) \geq k, d^+(n-1) \geq l+1, d^-(n-1) \geq l;$$

so (5.34) is an instance of Lemma 5.8. □

Log-concavity results being of some interest (see e.g. [43], [5]), we mention one natural consequence of Lemma 5.10 and Theorem 1.8 (and Proposition 5.6). For a bipartite graph $G = (V \cup K, E)$, define a *G-map* to be function $f : A \to K$ with $A \subseteq V$ and $(v, f(v)) \in E$ for all $v \in A$. Given $l, u \in \mathbb{N}^K$ (with $l_j \leq u_j$), call a $G$-map *valid* if

$$|f^{-1}(j)| \in [l_j, u_j]$$

for all $j \in K$. Let $s_k = s_k(G, l, u)$ be the number of valid $G$-maps $f : A \to K$ with ($A \subseteq V$ and) $|A| = k$.

**Theorem 5.11** *For any $G, l, u$ the sequence $(s_0, \ldots, s_{|V|})$ is ultra-log-concave.*

(See Section 1.3 for the definition of ultra-log-concave.)

*Remark.* In the special case that $l \equiv 0$ and $u \equiv 1$, $s_k$ is the number of matchings in $G$ of size $k$ (where, as usual, the *size* of a matching is the number of edges it contains), which we denote here by $\Phi_k = \Phi_k(G)$. Heilman and Lieb [26],[27] and Kunz [31] (see also [35, Chapter 8]) proved that for any (not necessarily bipartite) graph $G$, the *matching generating polynomial*

$$p(x) = \sum_{k=0}^{\nu} \Phi_k x^k$$

(where $\nu = \max\{k : \Phi_k > 0\}$ is the *matching number* of $G$) has all real (negative) roots. This implies, by Newton's inequalities (see e.g. [23, Theorem 51]), that the sequence

$(\Phi_0, \ldots, \Phi_\nu)$ is ULC, which (if $\nu < |V|$) is slightly stronger than Theorem 5.11 in this special case. In contrast, it's not hard to see that for arbitrary $l$ and $u$ the polynomial

$$\sum_{k=0}^{|V|} s_k x^k$$

need not have all real roots. (For example, consider the case $V = \{y, v, w\}$, $K = \{1\}$, $E(G) = \{\{y, 1\}, \{v, 1\}, \{w, 1\}\}$, $l_1 = 1, u_1 = 3$.)

*Proof of Theorem 5.11.* Let $\mu \in \mathcal{M}_V$ be the probability measure with $\mu(A)$ proportional to the number of valid $G$-maps from $A$ to $K$ (for $A \subseteq V$). Since $s := (s_0, \ldots, s_{|V|})$ is proportional to the rank sequence of $\mu$, it suffices (by Theorem 1.8) to show that $\mu$ has the CAPP and that $s$ has no internal zeros.

That $s$ has no internal zeros is an instance of Proposition 5.6: with notation as in Section 5.1 and Proposition 5.6, let $K = [n-1]$, $V = [m]$ (where, again, the copies of $i \in [\min\{m, n\}]$ in $V$ and $[n]$ are considered distinct), set

$$\gamma_{ij} = \begin{cases} 1 & \text{if } ij \in E \text{ or } j = n \\ 0 & \text{otherwise,} \end{cases}$$

and let $\hat{u} = (u_1 - l_1, \ldots, u_{n-1} - l_{n-1}, 0)$; then $s_k = M_{\hat{u}}(l_1, \ldots, l_{n-1}, m - k)$. (In fact, applying (5.12) to this setup shows that $s$ is LC.)

To show that $\mu$ has the CAPP, we should verify the APP for the conditional measures $\mu_{X,Y} \in \mathcal{M}_{Y \setminus X}$, where $X \subseteq Y$, $|Y \setminus X| = 2k$ (for some $k$), and

$$\mu_{X,Y}(A) \propto \mu(A \cup X) \quad (A \subseteq Y \setminus X).$$

(Here we regard $\Omega_V$ as $2^V$, as we did in Chapter 4.) Fix $X, Y$ as above, and for a $G$-map $f$, write $D(f)$ for the domain of $f$. Then, with

$$\mathcal{S}_i = \{(f, g) : f, g \text{ valid } G\text{-maps}, D(f) \cap D(g) = X, D(f) \cup D(g) = Y, |D(f)| = |X| + i\},$$

the APP for $\mu_{X,Y}$ is

$$|\mathcal{S}_k| \geq \frac{k+1}{k} |\mathcal{S}_{k-1}|. \tag{5.36}$$

We regard a $G$-map $f : A \to K$ as a subgraph of $G$ in the usual way (namely, as the graph on $V \cup K$ with edge set $\{(v, f(v)) : v \in A\}$). For each pair $(f, g)$ contributing to the left side of (5.36), the multiset union $F = f \cup g$ is a bipartite multigraph on $V \cup K$ satisfying

$$d_F(v) = \begin{cases} 2 & \text{if } v \in X \cap Y \\ 1 & \text{if } v \in Y \setminus X \\ 0 & \text{if } v \in V \setminus Y \end{cases} \tag{5.37}$$

for all $v \in V$ (and similarly for pairs contributing to the right side). We may rewrite (5.36) as

$$\sum_F |\{(f,g) \in \mathcal{S}_k : f \cup g = F\}| \geq \frac{k+1}{k} \sum_F |\{(h,p) \in \mathcal{S}_{k-1} : h \cup p = F\}|,$$

so it is enough to show

$$|\{(f,g) \in \mathcal{S}_k : f \cup g = F\}| \geq \frac{k+1}{k} |\{(h,p) \in \mathcal{S}_{k-1} : h \cup p = F\}| \tag{5.38}$$

for each fixed $F$.

Now let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ be the hypergraph on vertex set $W := E(F)$ with

$$\mathcal{H}_1 = \{H_j : j \in K\},$$

where $H_j = \{e \in W : j \in e\}$, and

$$\mathcal{H}_2 = \{\{e, e'\} : e \neq e', \ e \text{ and } e' \text{ have the same end in } V\},$$

and define $\alpha : \mathcal{H}_1 \to \mathbb{N}$ by

$$\alpha_{H_j} = \max\{l_j, d_F(j) - u_j\}.$$

Then (5.38) is an instance of Lemma 5.10 (since, by (5.37), the $t$ in Lemma 5.10 is equal to $k$). $\square$

Before closing this section we point out one further consequence of Lemma 5.8 that seems to us interesting in its own right. With notation as in the above proof of (5.12), set $f(A) = \sum\{W(\sigma) : \sigma \in K^A \cap Q\}$ $(A \subseteq [m])$. We assert that

$$f(A \cup B) f(A \cap B) \leq f(A) f(B) \qquad \forall A, B \subseteq [m]. \tag{5.39}$$

While we don't see how to get (5.12) from this, it does give another proof of Theorem 5.1, as follows. Set $f(i) = \sum\{f(A) : |A| = i\}$. If we specialize (5.30) to the case that the

$\sigma(i)$'s are i.i.d., the terms $X(A)X(B)$ and $X(C)X(D)$ cancel, and we are left needing the inequality

$$f^2(l) \geq \tfrac{m-l+1}{m-l} f(l+1)f(l-1).$$

But for i.i.d. $\sigma(i)$'s the measure (not probability measure) $f$ is exchangeable—i.e. $f(A)$ depends only on $|A|$—in which case (5.39) is trivially equivalent to ultra-log-concavity of the sequence $\{f(l\}_{l=0}^{m}$, that is, to

$$f^2(l) \geq \tfrac{m-l+1}{m-l}\tfrac{l+1}{l} f(l+1)f(l-1).$$

□

*Proof of* (5.39). We may rewrite the inequality as

$$\sum\sum\{W(\sigma)W(\tau) : \sigma \sim A \cup B, \tau \sim A \cap B\} \leq \sum\sum\{W(\alpha)W(\beta) : \alpha \sim A, \beta \sim B\}. \tag{5.40}$$

As before we regard $\sigma, \tau, \alpha, \beta$ in (5.40) as bipartite graphs on $[m] \cup K$. For each pair $(\sigma, \tau)$ appearing in (5.40), the (multiset) union $G = \sigma \cup \tau$ is a bipartite multigraph with

$$d_G(i) = \begin{cases} 2 & \text{if } i \in A \cap B \\ 1 & \text{if } i \in A \triangle B \\ 0 & \text{otherwise} \end{cases}$$

(and similarly for pairs $(\alpha, \beta)$), and it's enough to show that, for each such $G$, (5.40) still holds if we restrict to pairs $(\sigma, \tau)$ and $(\alpha, \beta)$ with

$$\sigma \cup \tau = \alpha \cup \beta = G. \tag{5.41}$$

Again the weights ($W(\sigma)$ etc.) cancel and it's enough to show

$$N(A \cup B, A \cap B) \leq N(A)N(B), \tag{5.42}$$

where, for $C, D \subseteq [m]$, $N(C, D) = N_G(C, D)$ is the number of partitions $E(G) = \gamma \cup \delta$ with $\gamma \sim C$ and $\delta \sim D$.

Notice now that we are really counting partitions $\hat{\sigma} \cup \hat{\tau}$ and $\hat{\alpha} \cup \hat{\beta}$ of the edges of $G' := G[(A \cap B) \cup K]$, since for any $\sigma, \ldots, \beta$ (as in (5.40)) satisfying (5.41), any edge of $G$ with an end in $A \setminus B$ (resp. $B \setminus A$) must belong to $\sigma \cap \alpha$ (resp. $\sigma \cap \beta$).

For $j \in K$ and $C \subseteq [m]$ write $d_C(j)$ for the number of edges of $G$ joining $j$ to $C$. In terms of $\hat{\sigma}, \ldots, \hat{\beta}$ the requirement that $\sigma, \ldots, \beta$ satisfy (5.29) becomes the condition that for each $j \in K$,

$$\begin{aligned} S_j - d_{A\Delta B}(j) \leq |\hat{\sigma}^{-1}(j)| < T_j - d_{A\Delta B}(j), \\ S_j \leq |\hat{\tau}^{-1}(j)| < T_j, \\ S_j - d_{A\setminus B}(j) \leq |\hat{\alpha}^{-1}(j)| < T_j - d_{A\setminus B}(j), \text{ and} \\ S_j - d_{B\setminus A}(j) \leq |\hat{\beta}^{-1}(j)| < T_j - d_{B\setminus A}(j). \end{aligned} \tag{5.43}$$

Now let $H$ be the multigraph on $K$ with edge set $\{e_x : x \in A \cap B\}$, where $e_x$ joins $j$ and $k$ if $x$ was joined to $j, k$ in $G'$. We may identify a partition $E(G') = \gamma \cup \delta$ with the orientation of $H$ gotten by directing $e_x$ from $j$ to $k$ whenever $\{x, j\} \in \gamma$ and $\{x, k\} \in \delta$. The orientations corresponding to pairs $(\hat{\sigma}, \hat{\tau})$ as in (5.43) are then those satisfying

$$S_j - d_{A\Delta B}(j) \leq d^+(j) < T_j - d_{A\Delta B}(j) \text{ and } \; S_j \leq d^-(j) < T_j$$

for all $j \in K$, while those corresponding to pairs $(\hat{\alpha}, \hat{\beta})$ are those with

$$S_j - d_{A\setminus B}(j) \leq d^+(j) < T_j - d_{A\setminus B}(j) \;\text{ and }\; S_j - d_{B\setminus A}(j) \leq d^-(j) < T_j - d_{B\setminus A}(j)$$

for all $j \in K$. That the number of orientations of the first type is at most the number of the second type is then an instance of Lemma 5.8. $\square$

## 5.4 Examples

In this short section, we give the easy examples justifying Proposition 1.25 and the remark following Theorem 1.21 (recall these say that log-concavity and the Rayleigh property fail for ordinary urn measures), and provide a counterexample to Conjecture 1.27.

In Examples 5.12 and 5.13 we use $p(j)$ for the probability that any given ball lands in urn $j$.

**Example 5.12** *Suppose we have three balls and urns $0, \ldots, n$, with $p(0) = \varepsilon$ and $p(1) = \cdots = p(n) = (1-\varepsilon)/n$, where $\varepsilon$ is small and $n\varepsilon^{3/2}$ is large. Then for the associated rank sequence, say $a = (a_1, a_2, a_3)$, we have $a_1 \approx \varepsilon^3$, $a_3 \approx (1+2\varepsilon)(1-\varepsilon)^2$,* *and*

$$a_2 = 3\varepsilon^2(1-\varepsilon) + 3\varepsilon(1-\varepsilon)^2/n + 3(1-\varepsilon)^3(n-1)/n^2 \approx 3\varepsilon^2(1-\varepsilon);$$

*so LC fails for a.*

(We don't know what happens if we replace "LC" by "unimodal.")

**Example 5.13** *Suppose we have two balls, urns* $0,1,2$*, and* $p(1) = p(2) = \varepsilon$*, with* $\varepsilon$ *small, and impose the external field* $(\varepsilon, 1, 1)$*. Then for the corresponding urn measure* $\mu$ *on* $\{0,1\}^{\{0,1,2\}}$ *(and* $\eta$ *the random configuration) we have* $\mu(\eta_1 = \eta_2 = 1) \propto 2\varepsilon^2$*,* $\mu(\eta_1 = \eta_2 = 0) \propto (1-2\varepsilon)^2\varepsilon$*, and* $\mu(\eta_1 = 1, \eta_2 = 0), \mu(\eta_1 = 0, \eta_2 = 1) \propto \varepsilon^2 + 2\varepsilon^2(1-2\varepsilon)$*, so that* $\eta_1$ *and* $\eta_2$ *are strictly positively correlated.*

**Example 5.14** *Let* $n = 3$ *and* $q = 1/3$ *(so we don't need* $\Lambda$*). Let* $M \cup A \cup B \cup C$ *be a partition of* $V := [m]$ *with* $|M| = s$ *(large) and* $|A| = |B| = |C| = t = s + 3$*. Let* $\mathcal{I}_1 = 2^{V \setminus M}$*,*

$$\mathcal{I}_2 = \{X \subseteq V : |X \cap M| < .4|M| \text{ and } X \text{ meets at most two of } A, B, C\},$$

*and* $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$*. Then, we assert,*

$$\Pr(\mathcal{A}_{\{3\}}) \Pr(\mathcal{A}_{\{1,2,3\}}) > \Pr(\mathcal{A}_{\{1,3\}}) \Pr(\mathcal{A}_{\{2,3\}}),$$

*which contradicts Conjecture 1.27 (with* $I = \{1\}$*,* $J = \{2\}$ *and* $K = \{3\}$*). We omit the precise calculations; roughly, with* $\alpha = (2/3)^t$ *and* $c = (3/2)^3$*, we have (as* $t \to \infty$*)*

$$\Pr(\mathcal{A}_L) \sim \begin{cases} (c+3)\alpha & \text{if } |L| = 1 \\ 6\alpha^2 & \text{if } |L| = 2 \\ 6\alpha^3 & \text{if } |L| = 3. \end{cases}$$

# Chapter 6

# Srinivasan's sampling process and Pemantle measures

In this chapter we prove Theorems 1.29 and 1.30, and also show that every that every SGSSP measure is a Pemantle measure. (These were defined in Section 1.8.)

## 6.1 Super-generalized Srinivasan sampling process measures are NA+

As we will see shortly, the proof of Theorem 1.29 is almost immediate once we establish an alternative description of SGSSP measures, for which we need some definitions. The *truncation* of $\mu \in \mathcal{M}$ to [r,s] is the conditional measure $\mu(\cdot|r \leq \sum \eta_i \leq s) \in \mathcal{M}$ (where, as usual, we assume the conditioning event has positive probability); it is a $k$-*truncation* if $s-r \leq k-1$ (i.e. the conditioning specifies that $|\eta|$ take one of at most $k$ consecutive values), which here we only use with $k \leq 2$.

Let $\mathcal{S}$ be the smallest class of measures that contains all Bernoulli measures (i.e. measures on $\{0,1\}$) and is closed under external fields, rank rescalings (defined after Theorem 1.29; note that log-concavity of the rescaling sequence $\{q_i\}$ will be irrelevant here), and the following operation:

$$\text{take the product of two measures and apply a 2-truncation.} \tag{6.1}$$

(So, in particular, every $\mu \in \mathcal{S}$ is concentrated on at most two levels, i.e. there is a $k$ so that $\mu(|\eta| \in \{k, k+1\}) = 1$.)

Note that the rank rescaling of a measure $\mu$ concentrated on at most two levels can be gotten from $\mu$ by applying a (uniform) external field, and that external fields commute with products and 2-truncations, so

$$\begin{array}{l}\mathcal{S} \text{ is the smallest class of measures that contains all Bernoulli} \\ \text{measures and is closed under (6.1).}\end{array} \tag{6.2}$$

We will prove

**Lemma 6.1** *$\mathcal{S}$ consists of exactly the SGSSP measures,*

but first give the easy derivation of Theorem 1.29 from Lemma 6.1.

Dubhashi *et al.* [12, Theorem 5.1] give a simple proof that every GSSP measure is NA. In fact, their proof is valid for SGSSP measures:

**Theorem 6.2** *SGSSP measures are NA.*

For completeness, we repeat their argument at the end of this section.

*Proof of Theorem 1.29.* Since $\mathcal{S}$ is trivially closed under imposition of external fields, this is an immediate consequence of Theorem 6.2 and Lemma 6.1. $\square$

Since $\mathcal{S}$ is obviously contained in the class of Pemantle measures (defined before Theorem 1.30), another consequence of Lemma 6.1 is

**Corollary 6.3** *Every SGSSP measure is a Pemantle measure;*

so, as noted earlier, Theorem 1.29 also follows from Theorem 1.30.

*Remark.* In fact, SGSSP measures are *strongly Rayleigh*, an even stronger property than NA+, whose definition we omit (see [4]). This follows from Lemma 6.1 and (6.2), since the class of strongly Rayleigh measures contains Bernoulli measures and is closed under products and 2-truncations [4]. In contrast, even some very simple Pemantle measures are not strongly Rayleigh (e.g. uniform measure on $\{\eta \in \Omega_3 : |\eta| \geq 1\}$).

We now slowly work toward the proof of Lemma 6.1, beginning with a result that relates the parameters of the pairing tree to properties of the corresponding measure. Consider an SGSSP measure $\mu$ corresponding to a pairing tree $T$ for $[n]$. For a vertex $u$, let $L(u)$ be the set of leaves that are descendants of $u$, and let $k(u)$ be the sum of $t_w$ over internal vertices $w$ that are descendants of $u$ (note $u$ is a descendant of itself). Also, for $\eta \in \Omega$ and a vertex $u$, set $\eta(u) = \sum_{i \in L(u)} \eta_i$. Observe that

$$\mu(\eta(u) \in \{k(u), k(u)+1\}) = 1 \quad \text{for all } u \tag{6.3}$$

(since once $u$ is "processed" exactly $k(u)$ of the coordinates in $L(u)$ have been fixed to 1 and there is exactly one unfixed coordinate in $L(u)$; note the final decision involving $\alpha$ is not considered part of the processing of the root); furthermore, if $v$ and $w$ are the two children of $u$, then, trivially,

$$k(u) = k(v) + k(w) + t_u. \tag{6.4}$$

It is also obvious that $\alpha = \mu(|\eta| = k(x) + 1)$, where $x$ is the root of $T$.

**Proposition 6.4** *With notation as above, for each internal vertex $u$ with left child $v$ and right child $w$, we have*

(a) *if $\mu(\eta(u) = k(v) + k(w)) > 0$, then $t_u = 0$,*

(b) *if $\mu(\eta(u) = k(v) + k(w) + 2) > 0$, then $t_u = 1$, and*

(c) *if $\mu(\eta(u) = k(v) + k(w) + 1) > 0$, then*

$$\begin{aligned} \beta_u &= \mu(\eta(v) = k(v) + t_u | \eta(u) = k(v) + k(w) + 1) \\ (&= \mu(\eta(w) = k(w) + 1 - t_u | \eta(u) = k(v) + k(w) + 1)). \end{aligned}$$

Observe that if $\mu(\eta(u) = k(v) + k(w) + 1) = 0$ then the value of $\beta_u$ has no effect on $\mu$ (since e.g. if $\mu(\eta(u) = k(v) + k(w) + 2) = 1$ then the two unfixed coordinates that are the children of $u$ when it is processed are always both eventually fixed to 1). Informally, part (c) says that when vertex $u$ is processed, it is determined which of the two unfixed coordinates that are the children of $u$ at that step of the procedure will be fixed to 1 *if it happens that eventually exactly one of them is fixed to* 1; otherwise (that is, if both coordinates are eventually fixed to the same value), which coordinate gets fixed in the step that $u$ is processed does not affect the random output $\eta$.

*Proof.* Parts (a) and (b) follow immediately from (6.3) and (6.4). To prove (c), first assume that $t_u = 1$. Then, with $\eta$ the (random) output of the process, the event $\{\eta(u) = k(v) + k(w) + 1\}$ occurs when the coordinate in $L(u)$ that remains unfixed after $u$ is processed is eventually fixed to 0. Thus, conditioned on $\{\eta(u) = k(v) + k(w) + 1\}$, $\eta(v) = k(v) + 1$ if and only if the coordinate fixed to 1 when $u$ is processed is in $L(v)$, which happens with probability $\beta_u$. A similar argument proves (c) when $t_u = 0$. $\square$

We now describe (in our terminology) the measures (now called SSP measures) that correspond to Srinivasan's original process [42], and use Proposition 6.4 to show that GSSP measures are in fact no more general than these. (Recall that GSSP measures correspond to pairing trees with $\alpha \in \{0,1\}$.) An SSP measure is any GSSP measure $\mu$ obtained from a pairing tree that satisfies ($\alpha \in \{0,1\}$ and)

$$t_u = \begin{cases} 1 & \text{if } e(v)+e(w) > 1 \\ 0 & \text{if } e(v)+e(w) < 1 \end{cases} \tag{6.5}$$

(so there is no requirement here when $e(v)+e(w)=1$) and

$$\beta_u = \begin{cases} \frac{e(w)}{e(v)+e(w)} & \text{if } t_u = 0 \\ \frac{1-e(w)}{2-e(v)-e(w)} & \text{if } t_u = 1 \end{cases} \tag{6.6}$$

for every internal vertex $u$ with left child $v$ and right child $w$, where

$$e(z) := \mathsf{E}\eta(z) - k(z) \quad (= \mu(\eta(z) = k(z)+1)). \tag{6.7}$$

(If the relevant denominator in (6.6) is zero then the value of $\beta_u$ has no effect on the resulting measure $\mu$.) The purpose of this choice of $t_u$'s and $\beta_u$'s was to produce $\eta$'s with prescribed marginal distributions. We omit a complete explanation (for more see [42], [12]), but briefly: each unfixed coordinate $i$ maintains a (changing) "value," initially set to the marginal probability $\mu(\eta_i = 1)$, which is given in advance; when vertex $u$ is processed, the value of the coordinate that was a child of $u$ and is still unfixed is updated to $e(u)$; and the parameters $t_u$, $\beta_u$ for each internal vertex $u$ with children $v, w$ are thought of as being determined by $e(v)$, $e(w)$ (as in (6.5) and (6.6)) instead of being given as part of the pairing tree. We now prove

**Proposition 6.5** *Every GSSP measure is an SSP measure.*

*Proof.* We need only show that (6.5) and (6.6) hold for any GSSP measure $\mu \in \mathcal{M}$. (In fact, the same argument shows that (6.5) and (6.6) hold for any SGSSP measure.) Let $T$ be a pairing tree for $[n]$ with associated measure $\mu$, and consider an internal vertex $u$ with left child $v$ and right child $w$. It is easy to see that (6.5) follows from Proposition 6.4 (e.g. if $e(v)+e(w) < 1$ then $\mu(\eta(u) = k(v)+k(w)) > 0$), so we just need to prove (6.6).

As noted after Proposition 6.4, the value of $\beta_u$ is irrelevant unless

$$\mu(\eta(u) = k(v) + k(w) + 1) > 0,$$

so assume this is the case. Consider the case when $t_u = 1$. (The case $t_u = 0$ is similar, and we omit it.) Since $t_u = 1$, the events $\{\eta(v) = k(v)\}$, $\{\eta(w) = k(w)\}$ are disjoint and

$$\mu(\eta(u) = k(v) + k(w) + 1) = \mu(\eta(v) = k(v)) + \mu(\eta(w) = k(w)).$$

Thus, by Proposition 6.4(c) and (6.7),

$$\beta_u = \frac{\mu(\eta(w) = k(w))}{\mu(\eta(v) = k(v)) + \mu(\eta(w) = k(w))} = \frac{1 - e(w)}{(1 - e(v)) + (1 - e(w))}.$$

□

*Remark.* Let us say a word about why we generalize from GSSP to SGSSP measures. It's not hard to see that, for any vertex $u$ in a pairing tree that yields an SGSSP measure $\mu$, the projection of $\mu$ onto the coordinates in $L(u)$ is again an SGSSP measure. (A pairing tree that yields the projection can be obtained from a pairing tree $T$ that gave $\mu$ by restricting to $T_u$ (the subtree rooted at $u$) and taking $\alpha(T_u) = e(u)$.) Such projections of GSSP measures need not be GSSP measures (since the projections will typically be concentrated on not one, but two, levels), so it's natural to consider the more general SGSSP measures if we are aiming for a result like Lemma 6.1 that describes a way to inductively build the measures.

*Proof of Lemma 6.1.* We show that every SGSSP measure $\mu \in \mathcal{M}_n$ is in $\mathcal{S}$ by induction on $n$. Since Bernoulli measures are in $\mathcal{S}$ (this is the case $n = 1$), it suffices to show that every SGSSP measure in $\mu \in \mathcal{M}_n$ ($n \geq 2$) can be gotten by applying (6.1), followed by a rank rescaling, to two other SGSSP measures. Let $T$ be a pairing tree for $[n]$ that yields $\mu$, let $u$ be the root of $T$, and let $v$ and $w$ be the left and right children of $u$, respectively. (Note that one or both of $v, w$ may be leaves.) We assume that $t_u = 1$; the proof in the case $t_u = 0$ is similar.

Set $\alpha(T_v) = \frac{1}{2-\beta}$ and $\alpha(T_w) = \frac{1}{1+\beta}$, where $\beta := \beta_u$ (recall that e.g. $T_v$ is the subtree of $T$ rooted at $v$). With $\pi$ and $\psi$ the SGSSP measures corresponding to $T_v$ and $T_w$, $k = k(u)$, $r = k(v)$, and $s = k(w)$ (so $k = r + s + 1$), it suffices to show

*Claim.* Let $\nu \in \mathcal{M}_n$ be the 2-truncation of $\pi \times \psi$ to $[k, k+1]$, and let $q = (q_i)_{i=0}^n$ be the sequence with

$$q_i = \begin{cases} \alpha & \text{if } i = k+1 \\ 1-\alpha & \text{if } i = k \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha = \alpha(T)$. Then $\mu = q \otimes \nu$.

The main points here (to be verified below) are: the parameters $\alpha(T_v), \alpha(T_w)$ are chosen to ensure that

$$\nu(\eta(v) = r+1 | \eta(u) = k) = \mu(\eta(v) = r+1 | \eta(u) = k); \tag{6.8}$$

the 2-truncation is chosen to ensure that $\nu$ is concentrated on levels $k, k+1$; and the rank rescaling is chosen to ensure that

$$q \otimes \nu(|\eta| = k+1) = \mu(|\eta| = k+1). \tag{6.9}$$

Since $\mu$ and $q \otimes \nu$ are each clearly concentrated on levels $k$ and $k+1$, it suffices to show (6.8), (6.9),

$$\text{the projection of } \mu(\cdot | \eta(v) = l) \text{ onto } L(v) \text{ is } \pi_l \text{ for } l \in \{r, r+1\}, \tag{6.10}$$

$$\text{the projection of } \mu(\cdot | \eta(w) = l) \text{ onto } L(w) \text{ is } \psi_l \text{ for } l \in \{s, s+1\}, \tag{6.11}$$

and

$$\text{under } \mu,\ \eta|_{L(v)} \text{ and } \eta|_{L(w)} \text{ are independent given } \eta(v), \eta(w), \tag{6.12}$$

where e.g. $\pi_l = \pi(\cdot | \sum_{i \in L(v)} \sigma_i = l) \in \mathcal{M}_{L(v)}$ and $\eta|_{L(v)} \in \{0,1\}^{L(v)}$ is the restriction of $\eta$ to $L(v)$. For convenience, for a nonnegative integer $l$ we set $\mu(l) = \mu(|\eta| = l)$, and define $\nu(l)$, $\pi(l)$, and $\psi(l)$ similarly.

To prove (6.8), observe that

$$\begin{aligned} \nu(\eta(v) = r+1 | \eta(u) = k) &= \frac{\pi(r+1)\psi(s)}{\pi(r+1)\psi(s) + \pi(r)\psi(s+1)} \\ &= \frac{\frac{1}{2-\beta} \cdot \frac{\beta}{1+\beta}}{\frac{1}{2-\beta} \cdot \frac{\beta}{1+\beta} + \frac{1-\beta}{2-\beta} \cdot \frac{1}{1+\beta}} \\ &= \beta \\ &= \mu(\eta(v) = r+1 | \eta(u) = k); \end{aligned}$$

the final equality is an instance of Proposition 6.4(c).

We also have

$$\begin{aligned} q \otimes \nu(k+1) &= \frac{q_{k+1}\pi(r+1)\psi(s+1)}{q_{k+1}\pi(r+1)\psi(s+1) + q_k(\pi(r+1)\psi(s) + \pi(r)\psi(s+1))} \\ &= \frac{\alpha \cdot \frac{1}{2-\beta} \cdot \frac{1}{1+\beta}}{\alpha \cdot \frac{1}{2-\beta} \cdot \frac{1}{1+\beta} + (1-\alpha) \cdot (\frac{1}{2-\beta} \cdot \frac{\beta}{1+\beta} + \frac{1-\beta}{2-\beta} \cdot \frac{1}{1+\beta})} \\ &= \alpha \\ &= \mu(k+1), \end{aligned}$$

which gives (6.9).

Finally, observe that (under $\mu$), given $\eta(v)$, $\eta|_{L(v)}$ is determined by what happens when the descendants of $v$ are processed, and similarly for $\eta|_{L(w)}$ given $\eta(w)$. This proves (6.10)-(6.12), completing the proof of the claim and the assertion that every SGSSP measure is in $\mathcal{S}$.

We now show that every measure in $\mathcal{S}$ is an SGSSP measure. Every Bernoulli measure is clearly an SGSSP measure (the pairing tree consisting of a single vertex and parameter $\alpha$ yields the measure on $\{0,1\}$ that assigns probability $\alpha$ to 1), so by (6.2) it suffices to prove that

any measure obtained from two SGSSP measures by applying (6.1) is an SGSSP measure. (6.13)

Let $\pi \in \mathcal{M}_I$ be an SGSSP measure gotten from a pairing tree $R$ for $I$, and let $\psi \in \mathcal{M}_J$ ($I \cap J = \emptyset$) be an SGSSP measure gotten from a pairing tree $S$ for $J$. Let $v$ be the root of $R$ and $w$ the root of $S$, and set $r = k(v)$ and $s = k(w)$. We will show that the truncation, $\nu$, of $\pi \times \psi$ to $[r+s, r+s+1]$ is an SGSSP measure by constructing an appropriate pairing tree for $I \cup J$. (Similar arguments show the truncation of $\pi \times \psi$ to $[r+s+1, r+s+2]$ and any 1-truncation of $\pi \times \psi$ are SGSSP measures; note that $\pi \times \psi$ is concentrated on levels $r+s, r+s+1, r+s+2$.) Note that we may assume at least one of $\alpha(R), \alpha(S)$ is strictly less than 1 (else $\nu$ is identically zero).

We construct a pairing tree $T$ for $I \cup J$ from $R$ and $S$, as follows. Join each of $v$, $w$ to a new vertex $u$ (which will be the root of $T$), designating $v$ as the left child of $u$.

Internal vertices of $R$ and $S$ retain their $t, \beta$ values, and we set $t_u = 0$,

$$\beta_u = \frac{(1-\alpha(R))\alpha(S)}{(1-\alpha(R))\alpha(S) + \alpha(R)(1-\alpha(S)},$$

and

$$\alpha(T) = \frac{(1-\alpha(R))\alpha(S) + \alpha(R)(1-\alpha(S))}{1-\alpha(R)\alpha(S)}.$$

(If $\alpha(R) = \alpha(S) = 0$, then, as pointed out following Proposition 6.4, the value of $\beta_u$ is irrelevant.)

Letting $\mu$ be the SGSSP measure gotten from the pairing tree $T$, we should show that $\mu = \nu$. Each of $\mu$, $\nu$ is clearly concentrated on levels $r+s, r+s+1$, and it's easy to see that

$$\mu(r+s+1) = \nu(r+s+1) = \alpha(T)$$

and

$$\mu(\eta(v) = r | \eta(u) = r+s+1) = \nu(\eta(v) = r | \eta(u) = r+s+1) = \beta_u.$$

Thus, that $\mu = \nu$ follows from the following three facts, each of which (as above) is easily seen to be true: the projection of $\mu(\cdot | \eta(v) = l)$ onto $I = L(v)$ is $\pi_l$ (for $l \in \{r, r+1\}$); the projection of $\mu(\cdot | \eta(w) = l)$ onto $J = L(w)$ is $\psi_l$ (for $l \in \{s, s+1\}$); and $\eta|_I$, $\eta|_J$ are independent (under $\mu$) given $\eta(v), \eta(w)$. This finishes the proof of Lemma 6.1. $\square$

*Proof of Theorem 6.2.* As mentioned above, this argument is from [12], though given here using our terminology. We proceed by induction on $n$, letting $\mu \in \mathcal{M}_n$ be an SGSSP measure. (The case $n = 1$ is trivial.) We are given increasing $\mathcal{A}, \mathcal{C} \subseteq \Omega$ with $\mathcal{A} \perp \mathcal{C}$, and should show $\mathcal{A} \downarrow \mathcal{C}$. Let $u$ be the first internal vertex that is processed; we may assume without loss of generality that the children of $u$ are the leaves 1 and 2. If $t_u = 1$, let $\mathcal{B}$ be the event that coordinate 1 is fixed to 1 at the first step. (If $t_u = 0$, let $\mathcal{B}$ be the event that coordinate 2 is fixed to 0 at the first step.) Observe that each of $\mu(\cdot|\mathcal{B})$, $\mu(\cdot|\Omega \setminus \mathcal{B})$ is (equivalent to) an SGSSP measure with $n-1$ variables, so by induction we have $\mathcal{A} \downarrow \mathcal{C}$ conditionally given each of $\mathcal{B}$, $\Omega \setminus \mathcal{B}$. If one of $\mathcal{A}$, $\mathcal{C}$ (say $\mathcal{A}$) is not affected by either of the coordinates 1, 2, then $\mathcal{A}$ is independent of $\mathcal{B}$ (since the distribution of $(\eta_3, \ldots, \eta_n)$ does not depend on $\mathcal{B}$), whence $\mathcal{A} \downarrow \mathcal{C}$ follows via Lemma 2.6. (If $\mathcal{C} \uparrow \mathcal{B}$, take $f = \mathbf{1}_{\Omega \setminus \mathcal{A}}$ and $g = \mathbf{1}_{\mathcal{C}}$; otherwise take $f = \mathbf{1}_{\mathcal{A}}$ and $g = \mathbf{1}_{\Omega \setminus \mathcal{C}}$.)

So we may assume, without loss of generality, that $\mathcal{A}$ is affected by coordinate 1 and $\mathcal{C}$ is affected by coordinate 2. In this case $\mathcal{A} \uparrow \mathcal{B}$ and $\mathcal{C} \downarrow \mathcal{B}$ (since $\mathcal{A}$, $\mathcal{C}$ are increasing and the distribution of $(\eta_3, \ldots, \eta_n)$ is independent of $\mathcal{B}$), and we again finish by Lemma 2.6 (with $f = \mathbf{1}_{\mathcal{A}}$ and $g = \mathbf{1}_{\Omega \setminus \mathcal{C}}$). □

## 6.2 Pemantle measures are NA+

In this section we prove Theorem 1.30, which, as we will see shortly, is an immediate consequence of Theorem 1.18 and the following result. Recall the normalized matching property defined before Proposition 2.4 and the stochastic covering property defined before Conjecture 3.2, and say $\mu \in \mathcal{M}$ is SCP* if every measure gotten from $\mu$ by rank rescaling (defined in Section 1.8) has the SCP; note that any measure which has the SCP is NC.

**Theorem 6.6** (a) *The product of two LC measures is LC; equivalently, the convolution of two LC sequences with no internal zeros is LC (and has no internal zeros).*

(b) *The product of two LC measures with the NMP has the NMP.*

(c) *The product of two measures that are both SCP* and LC and have the NMP is SCP*.*

Part (a) is easy and standard (see e.g. [28, Theorem 1], [33, p. 317], or [2, Exercise 4.7]), so we omit the proof here. Part (b) is an instance of a result of Harper [24, Section I.C] (and can also essentially be gotten from a combinatorial version proved in [28], [2]); for completeness, and since Harper's original proof is not that easy, we include a proof (together with the proof of (c)) below. Before giving these arguments, we show that Theorem 6.6 implies Theorem 1.30.

*Proof of Theorem 1.30.* Since the class of Pemantle measures is closed under imposition of external fields, by Theorem 1.18(b) it suffices to show that every Pemantle measure is both NC and FM. The imposition of external fields commutes with rank rescalings and products, so in fact every Pemantle measure can be obtained from a collection of Bernoulli measures by repeatedly taking products and applying rank rescalings, and it thus suffices to show that every measure obtained in this way is SCP* and has the NMP (since SCP* implies NC and the NMP implies FM); but this follows immediately from Theorem 6.6 (since SCP*, LC, and NMP are all obviously preserved by rank rescalings). □

For the proof of Theorem 6.6 we need two preliminary results.

**Proposition 6.7** (a) *$\mu \in \mathcal{M}$ stochastically dominates $\nu \in \mathcal{M}$ if and only if there exists a "flow" $f : \Omega \times \Omega \to \mathbb{R}_+$ satisfying*

(i) $\sum_{\eta\in\Omega} f(\eta,\tau) = \mu(\tau)$ *for all* $\tau \in \Omega$, $\sum_{\tau\in\Omega} f(\eta,\tau) = \nu(\eta)$ *for all* $\eta \in \Omega$, *and*

(ii) $f(\eta,\tau) > 0$ *implies* $\tau \geq \eta$

(b) *$\mu \in \mathcal{M}$ stochastically covers $\nu \in \mathcal{M}$ if and only if there is a "flow" $f : \Omega \times \Omega \to \mathbb{R}_+$ satisfying* (i) *and*

(ii′) $f(\eta,\tau) > 0$ *implies* $\tau = \eta$ *or* $\tau \gtrdot \eta$.

Part (a) is a special case of the finite version of *Strassen's Theorem* [44] (and also follows easily from the "maxflow-mincut theorem," for which see e.g. [34, p. 64]), and (b) is a restatement of the definition of stochastic covering (given before Conjecture 3.2).

**Lemma 6.8** *Let $a = (a_0,\dots,a_s)$ and $b = (b_0,\dots,b_t)$ be nonnegative LC sequences with no internal zeros, and set*

$$Z_l = \sum_i a_{l-i}b_i \quad \text{and} \quad Z_l(j) = \sum_{i\leq j} a_{l-i}b_i. \tag{6.14}$$

*Then, for each $k$ and $j$ with $Z_k, Z_{k+1} > 0$, we have*

$$\frac{Z_k(j)}{Z_k} \geq \frac{Z_{k+1}(j)}{Z_{k+1}} \geq \frac{Z_k(j-1)}{Z_k}. \tag{6.15}$$

*Proof.* It suffices to prove the first inequality in (6.15) for all relevant $k, j$, since the second then follows by interchanging the roles of $a$ and $b$ in (6.14). Fix $k, j$ with $Z_k, Z_{k+1} > 0$. After cross multiplying and canceling terms that appear on both sides, the first inequality in (6.15) becomes

$$\sum_{i\leq j} a_{k-i}b_i \sum_{m>j} a_{k+1-m}b_m \geq \sum_{i\leq j} a_{k+1-i}b_i \sum_{m>j} a_{k-m}b_m,$$

which clearly follows from

$$a_{k-i}a_{k+1-m} \geq a_{k+1-i}a_{k-m} \text{ whenever } m > i,$$

an easy (and standard) consequence of the fact that $a$ is LC without internal zeros. $\square$

*Proof of Theorem 6.6.* As mentioned above, we omit the standard proof of part (a). To prove (b), assume $\mu \in \mathcal{M}_S$ and $\nu \in \mathcal{M}_T$ (with $S \cap T = \emptyset$) each have the NMP and are LC, and let $\pi = \mu \times \nu$ ($\in \mathcal{M}_{S\cup T}$) and $k \in \{0, \ldots, |S| + |T| - 1\}$; we should show $\pi_{k+1} \succeq \pi_k$, where $\pi_l = \pi(\cdot|\sum \eta_i = l)$. It will be convenient to write elements of $\Omega_{S\cup T}$ as $(\rho, \xi)$ with $\rho \in \Omega_S$ and $\xi \in \Omega_T$. Since $\mu$, $\nu$ have the NMP (by hypothesis), Proposition 6.7(a) gives flows $f_i : \Omega_S \times \Omega_S \to \mathbb{R}_+$ ($i \in \{0, \ldots, |S| - 1\}$) and $g_j : \Omega_T \times \Omega_T \to \mathbb{R}_+$ ($j \in \{0, \ldots, |T| - 1\}$) satisfying

$$\sum_{\rho\in\Omega_S} f_i(\rho,\sigma) = \mu_{i+1}(\sigma), \quad \sum_{\sigma\in\Omega_S} f_i(\rho,\sigma) = \mu_i(\rho), \text{ and } f_i(\rho,\sigma) > 0 \text{ implies } \sigma > \rho \quad (6.16)$$

and

$$\sum_{\xi\in\Omega_T} g_j(\xi,\tau) = \nu_{j+1}(\tau), \quad \sum_{\tau\in\Omega_T} g_j(\xi,\tau) = \nu_j(\xi), \text{ and } g_j(\xi,\tau) > 0 \text{ implies } \tau > \xi. \quad (6.17)$$

Of course the idea is to use the $f_i$'s and $g_j$'s to construct a flow $h$ that shows $\pi_{k+1} \succeq \pi_k$.

Let $a = (a_0, \ldots, a_{|S|})$ and $b = (b_0, \ldots, b_{|T|})$ be the rank sequences of $\mu$ and $\nu$, respectively, define $Z_l$ and $Z_l(j)$ as in (6.14), and define $h : \Omega_{S\cup T} \times \Omega_{S\cup T} \to \mathbb{R}$ by

$$h((\rho,\xi),(\sigma,\tau)) = \begin{cases} g_{|\xi|}(\xi,\tau)\frac{\mu(\rho)}{a_{|\rho|}}\left(\frac{Z_k(k-|\rho|)}{Z_k} - \frac{Z_{k+1}(k-|\rho|)}{Z_{k+1}}\right) & \text{if } |\rho| + |\xi| = k \text{ and } \rho = \sigma \\ \\ f_{|\rho|}(\rho,\sigma)\frac{\nu(\xi)}{b_{|\xi|}}\left(\frac{Z_{k+1}(|\xi|)}{Z_{k+1}} - \frac{Z_k(|\xi|-1)}{Z_k}\right) & \text{if } |\rho| + |\xi| = k \text{ and } \xi = \tau \\ \\ 0 & \text{otherwise.} \end{cases}$$

It clearly follows from (6.16) and (6.17) that

$$h((\rho,\xi),(\sigma,\tau)) > 0 \text{ implies } |\rho| + |\xi| = k,\ |\sigma| + |\tau| = k + 1,\ \sigma \geq \rho \text{ and } \tau \geq \xi, \quad (6.18)$$

so to prove $\pi_{k+1} \succeq \pi_k$ via Proposition 6.7(a) it is enough to show

$$\sum_{(\rho,\xi)} h((\rho,\xi),(\sigma,\tau)) = \pi_{k+1}(\sigma,\tau) \quad \text{for all } (\sigma,\tau), \quad (6.19)$$

$$\sum_{(\sigma,\tau)} h((\rho,\xi),(\sigma,\tau)) = \pi_k(\rho,\xi) \quad \text{for all } (\rho,\xi), \tag{6.20}$$

and

$$h((\rho,\xi),(\sigma,\tau)) \geq 0 \quad \text{for all } (\rho,\xi) \text{ and } (\sigma,\tau). \tag{6.21}$$

The proofs of (6.19) and (6.20) are similar, so we just do (6.19). Observe that, by (6.18), it suffices to consider the case that $|\sigma| + |\tau| = k+1$, whence (6.19) is a straightforward computation:

$$\begin{aligned}
\sum_{(\rho,\xi)} h((\rho,\xi),(\sigma,\tau)) &= \sum_{\xi} g_{|\xi|}(\xi,\tau)\frac{\mu(\sigma)}{a_{|\sigma|}}\left(\frac{Z_k(k-|\sigma|)}{Z_k} - \frac{Z_{k+1}(k-|\sigma|)}{Z_{k+1}}\right) \\
&\quad + \sum_{\rho} f_{|\rho|}(\rho,\sigma)\frac{\nu(\tau)}{b_{\tau|}}\left(\frac{Z_{k+1}(|\tau|)}{Z_{k+1}} - \frac{Z_k(|\tau|-1)}{Z_k}\right) \\
&= \frac{\nu(\tau)}{b_{|\tau|}} \cdot \frac{\mu(\sigma)}{a_{|\sigma|}}\left(\frac{Z_k(k-|\sigma|)}{Z_k} - \frac{Z_{k+1}(k-|\sigma|)}{Z_{k+1}}\right) \\
&\quad + \frac{\mu(\sigma)}{a_{|\sigma|}} \cdot \frac{\nu(\tau)}{b_{|\tau|}}\left(\frac{Z_{k+1}(|\tau|)}{Z_{k+1}} - \frac{Z_k(|\tau|-1)}{Z_k}\right) \\
&= \frac{\mu(\sigma)\nu(\tau)}{a_{|\sigma|}b_{|\tau|}Z_{k+1}}\left(Z_{k+1}(|\tau|) - Z_{k+1}(k-|\sigma|)\right) \\
&= \pi_{k+1}(\sigma,\tau);
\end{aligned}$$

here the first equality follows from the definition of $h$, the second from (6.16) and (6.17), and the third and fourth from the assumption that $|\sigma| + |\tau| = k+1$.

So in order to finish the proof of (b) we only need (6.21), which is an instance of Lemma 6.8 (and so it is here that we use the hypothesis that $\mu$ and $\nu$ are LC).

To prove (c), let $\mu \in \mathcal{M}_S$ and $\nu \in \mathcal{M}_T$ ($S \cap T = \emptyset$) be SCP* and LC and have the NMP, and let $q = (q_0, \ldots, q_n)$ be an LC sequence with no internal zeros (where $n = |S| + |T|$); we should show

$$q \otimes (\mu \times \nu)(\cdot|\eta_i = 0) \succ q \otimes (\mu \times \nu)(\cdot|\eta_i = 1) \tag{6.22}$$

for all $i \in S \cup T$. Assume, without loss of generality, that $i \in S$, and set $S' = S \setminus \{i\}$ and $\mu^k = \mu(\cdot|\eta_i = k) \in \mathcal{M}_{S'}$, whence (6.22) is

$$(q_j)_{j=0}^{n-1} \otimes (\mu^0 \times \nu) \succ (q_j)_{j=1}^{n} \otimes (\mu^1 \times \nu).$$

It will be convenient to write elements of $\Omega_{S'\cup T}$ as $(\rho,\xi)$ with $\rho\in\Omega_{S'}$ and $\xi\in\Omega_T$ and let $\pi^k=(q_j)_{j=k}^{n+k-1}\otimes(\mu^k\times\nu)$ for $k\in\{0,1\}$. For $k\in\{0,1\}$, let $\varepsilon_k>0$ be such that

$$\pi^k(\rho,\xi)=\varepsilon_k q_{|\rho|+|\xi|+k}\mu^k(\rho)\nu(\xi) \tag{6.23}$$

(so $\varepsilon_k$ is just a normalizing factor), and set

$$Z_l=\sum_i b_i q_{l+i} \quad \text{and} \quad Z_l(j)=\sum_{i\le j} b_i q_{l+i},$$

where $b=(b_0,\ldots,b_{|T|})$ is the rank sequence of $\nu$. The sequence $Z=(Z_l)_{l=0}^{|S|}$ is LC and has no internal zeros (by Theorem 6.6(a), since $Z$ is the convolution of $b$ and the reverse of $q$), so our hypotheses include that $Z\otimes\mu$ has the SCP; that is, (using Proposition 6.7 (b)) there exists a flow $f:\Omega_{S'}\times\Omega_{S'}\to\mathbb{R}_+$ satisfying

$$\sum_{\sigma\in\Omega_{S'}} f(\rho,\sigma)=\varepsilon_1\mu^1(\rho)Z_{|\rho|+1}, \qquad \sum_{\rho\in\Omega_{S'}} f(\rho,\sigma)=\varepsilon_0\mu^0(\sigma)Z_{|\sigma|}, \tag{6.24}$$

and

$$f(\rho,\sigma)>0 \text{ implies } \sigma=\rho \text{ or } \sigma\gtrdot\rho.$$

Also, since $\nu$ has the NMP, there are flows $g_j:\Omega_T\times\Omega_T\to\mathbb{R}_+$ $(j\in\{0,\ldots,|T|-1\})$ satisfying

$$\sum_{\xi\in\Omega_T} g_j(\xi,\tau)=\nu_{j+1}(\tau), \quad \sum_{\tau\in\Omega_T} g_j(\xi,\tau)=\nu_j(\xi), \text{ and } g_j(\xi,\tau)>0 \text{ implies } \tau\gtrdot\xi, \tag{6.25}$$

where, again, $\nu_l$ is the conditional measure $\nu(\cdot|\sum_{i\in T}\tau_i=l)$.

We proceed in a manner similar to the proof of (b) above; again, the idea is to use $f$ and the $g_j$'s to construct a flow $h$ that shows $\pi^0\succ\pi^1$. Define $h:\Omega_{S'\cup T}\times\Omega_{S'\cup T}\to\mathbb{R}$ by

$$h((\rho,\xi),(\sigma,\tau))=\begin{cases} f(\rho,\sigma)\frac{q_{|\sigma|+|\xi|}\nu(\xi)}{Z_{|\sigma|}} & \text{if } \xi=\tau \text{ and } \sigma\gtrdot\rho \\[2ex] f(\rho,\rho)\nu_{|\xi|}(\xi)\left(\frac{Z_{|\rho|}(|\xi|)}{Z_{|\rho|}}-\frac{Z_{|\rho|+1}(|\xi|-1)}{Z_{|\rho|+1}}\right) & \text{if } \rho=\sigma \text{ and } \xi=\tau \\[2ex] f(\rho,\rho)g_{|\xi|}(\xi,\tau)\left(\frac{Z_{|\rho|+1}(|\xi|)}{Z_{|\rho|+1}}-\frac{Z_{|\rho|}(|\xi|)}{Z_{|\rho|}}\right) & \text{if } \rho=\sigma \text{ and } \tau\gtrdot\xi \\[2ex] 0 & \text{otherwise.} \end{cases}$$

It is obvious that

$$h((\rho,\xi),(\sigma,\tau)) > 0 \text{ implies } (\sigma,\tau) = (\rho,\xi) \text{ or } (\sigma,\tau) \gtrdot (\rho,\xi),$$

and it follows from Lemma 6.8 that

$$h((\rho,\xi),(\sigma,\tau)) \geq 0 \quad \text{for all } (\rho,\xi) \text{ and } (\sigma,\tau).$$

Thus, by Proposition 6.7(b), in order to complete the proof of (c) we only have to show

$$\sum_{(\rho,\xi)} h((\rho,\xi),(\sigma,\tau)) = \pi^0(\sigma,\tau) \quad \text{for all } (\sigma,\tau)$$

and

$$\sum_{(\sigma,\tau)} h((\rho,\xi),(\sigma,\tau)) = \pi^1(\rho,\xi) \quad \text{for all } (\rho,\xi);$$

again these are similar, so we just do the latter:

$$\begin{aligned}
\sum_{(\sigma,\tau)} h((\rho,\xi),(\sigma,\tau)) &= f(\rho,\rho)\nu_{|\xi|}(\xi)\left(\frac{Z_{|\rho|}(|\xi|)}{Z_{|\rho|}} - \frac{Z_{|\rho|+1}(|\xi|-1)}{Z_{|\rho|+1}}\right) \\
&\qquad + \sum_{\tau \gtrdot \xi} f(\rho,\rho) g_{|\xi|}(\xi,\tau)\left(\frac{Z_{|\rho|+1}(|\xi|)}{Z_{|\rho|+1}} - \frac{Z_{|\rho|}(|\xi|)}{Z_{|\rho|}}\right) \\
&\qquad + \sum_{\sigma \gtrdot \rho} f(\rho,\sigma)\frac{q_{|\rho|+|\xi|+1}\nu(\xi)}{Z_{|\rho|+1}} \\
&= f(\rho,\rho)\nu_{|\xi|}(\xi)\frac{b_{|\xi|}q_{|\rho|+|\xi|+1}}{Z_{|\rho|+1}} + \sum_{\sigma \gtrdot \rho} f(\rho,\sigma)\frac{q_{|\rho|+|\xi|+1}\nu(\xi)}{Z_{|\rho|+1}} \\
&= \varepsilon_1 q_{|\rho|+|\xi|+1}\mu^1(\rho)\nu(\xi) \\
&= \pi^1(\rho,\xi),
\end{aligned}$$

where we used the definition of $h$ for the first equality, (6.25) for the second, (6.24) (and that $\nu(\xi) = b_{|\xi|}\nu_{|\xi|}(\xi)$) for the third, and (6.23) for the fourth. □

## References


[1] R. Ahlswede and D.E. Daykin, An inequality for the weights of two families of sets, their unions and intersections, *Z. Wahrscheinl. Geb.* **43** (1978), 183-185.

[2] I. Anderson, *Combinatorics of Finite Sets*, Oxford Univ. Pr., Oxford, 1989.

[3] B. Bollobás, *Combinatorics*, Cambridge University Press, Cambridge, 1986.

[4] J. Borcea, P. Brändén and T.M. Liggett, Negative dependence and the geometry of polynomials, *J. Amer. Math. Soc.* **22** (2009), 521-567 (arXiv:0707.2340v2).

[5] F. Brenti, Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update, *Contemporary Math.* **178** (1994), 71-89.

[6] R.L. Brooks, C.A.B. Smith, A.H. Stone and W.T. Tutte, The dissection of rectangles into squares, *Duke Math. J.* **7** (1940), 312-340.

[7] Y. Choe and D. Wagner, Rayleigh Matroids, *Combin. Probab. Comput.* **15** (2006), 765-781 (arXiv:math.CO/0307096v3).

[8] Ph. Delsarte, *An Algebraic Approach to the Association Schemes of Coding Theory*, Philips Res. Rep. Suppl. **10**, 1973.

[9] R.P. Dilworth and C. Greene, A counter-example to the generalisation of Sperner's theorem, *J. Comb. Theory* **10** (1971), 18-21.

[10] T.A. Dowling, On the independent set numbers of a finite matroid, *Combinatorics 79 (Proc. Colloq., Univ. Montréal, Montreal, Que., 1979), Part I, Ann. Discrete Math.* **8** (1980), 21-28.

[11] P. Doyle and J.L. Snell, *Random Walks and Electric Networks*, Carus Mathematical Monographs **22**, MAA, Washington DC, 1984.

[12] D. Dubhashi, J. Jonasson and D. Ranjan, Positive influence and negative dependence, *Combin. Probab. Comput.* **16** (2007), 29-41.

[13] D. Dubhashi and D. Ranjan: Balls and bins: a study of negative dependence, *Random Structures & Algorithms* **13** (1998), 99-124.

[14] T. Feder and M. Mihail, Balanced matroids, pp. 26-38 in *Proc. 24th STOC*, ACM, 1992.

[15] C. Fortuin, P. Kasteleyn and J. Ginibre Correlation inequalities on some partially ordered sets, *Comm. Math. Phys.* **22** (1971), 89-103.

[16] R. Gandhi, S. Khuller, S. Parthasarathy and A. Srinivasan, Dependent rounding and its applications to approximation algorithms, *Journal of the ACM* **53** (2006), 324-360.

[17] H.-O. Georgii, O. Häggström and C. Maes, The random geometry of equilibrium phases, pp. 1-142 in *Phase Transitions and Critical Phenomena, Volume 18* (C. Domb and J.L. Lebowitz, eds.), Academic Press, London, 2001.

[18] G. Grimmett, *Percolation*, Springer-Verlag, Berlin, 1999.

[19] G. Grimmett, The random-cluster model, pp. 73-123 in *Probability on Discrete Structures,* H. Kesten ed., *Encyclopedia of Mathematical Sciences, vol. 110,* Springer, 2003.

[20] G. Grimmett, *The Random-Cluster Model*, Springer, 2006.

[21] G. Grimmett and S. Winkler, Negative association in uniform forests and connected graphs *Random structures & Algorithms* **24** (2004) 444-460.

[22] Y.O. Hamidoune and I. Salaün, On the Independence Numbers of a Matroid, *J. Comb. Theory Series B* **47** (1989), 146-152.

[23] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, 2nd edition, Cambridge University Press, Cambridge, 1952.

[24] L.H. Harper, The morphology of partially ordered sets, *J. Combin. Theory Series A* **17** (1974), 44-58.

[25] T.E. Harris, A lower bound on the critical probability in a certain percolation process, *Proc. Cam. Phil. Soc.* **56** (1960), 13-20.

[26] O.J. Heilman and E.H. Lieb, Monomers and dimers, *Phys. Rev. Letters* **24** (1970), 1412-1414.

[27] O.J. Heilman and E.H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Physics* **25** (1972), 190-232.

[28] W.N. Hsieh and D.J. Kleitman, Normalized matching in direct products of partial orders. *Studies in Appl. Math.* **52** (1973), 285-289.

[29] J. Kahn, A normal law for matchings, *Combinatorica* **20** (2000), 339-391.

[30] D.J. Kleitman, Families of non-disjoint subsets, *J. Combinatorial Th.* **1** (1966), 153-155.

[31] H. Kunz, Location of the zeros of the partition function for some classical lattice systems, *Phys. Lett. (A)* (1970), 311-312.

[32] T.M. Liggett, *Interacting Particle Systems*, Springer-Verlag, New York, 1985.

[33] T.M. Liggett, Ultra logconcave sequences and negative dependence, *J. Combinatorial Th. A* **79** (1997), 315-325.

[34] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, 2nd edition, Cambridge University Press, Cambridge, 2001.

[35] L. Lovász and M.D. Plummer, *Matching Theory*, North Holland, Amsterdam, 1986.

[36] R. Lyons and Y. Peres, *Probability on Trees and Networks*, http://mypage.iu.edu/~rdlyons/#book.

[37] C. Mahoney, On the unimodality of the independent set numbers of a class of matroids, *J. Comb. Theory Series B* **39** (1985), 77-85.

[38] J.H. Mason, Matroids: unimodal conjectures and Motzkin's theorem, pp. 207-221 in *Combinatorics* (D.J.A. Welsh and D.R. Woodall, eds.), Inst. Math. & Appl., 1972.

[39] J.G. Oxley, *Matroid Theory*, Oxford Univ. Pr., Oxford, 1992.

[40] R. Pemantle, Towards a theory of negative dependence, *J. Math. Phys.* **41** (2000), 1371-1390.

[41] P.D. Seymour and D.J.A. Welsh, Combinatorial applications of an inequality from statistical mechanics, *Math. Proc. Cam. Phil. Soc.* **77** (1975), 485-495.

[42] A. Srinivasan, Distributions on level sets with applications to approximation algorithms, *Proc. 42nd Annual Symp. Found. Comp. Sci.* IEEE, 2001.

[43] R.P. Stanley, Log-concave and unimodal sequences in algebra, combinatorics, and geometry, pp. 500-535 in *Graph theory and its applications: East and West (Jinan, 1986), Ann. New York Acad. Sci.* **576**, New York Acad. Sci., New York, 1989.

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

[45] D. Wagner, Matroid inequalities from electrical network theory, *Electronic J. Combin.* **11** (2005), #A1.

[46] D. Wagner, Negatively correlated random variables and Mason's conjecture for independent sets in matroids, *Ann. Comb.* **12** (2008), 211-239.

[47] D.J.A. Welsh, Combinatorial problems in matroid theory, pp. 291-307 in *Combinatorial Mathematics and its Applications*, Academic Press, 1971.

[48] D.J.A. Welsh, Harris's inequality and its descendants, lecture at Isaac Newton Inst., June 2008, http://www.newton.ac.uk/programmes/CSM/seminars/062310005.pdf.

[49] D.J.A. Welsh, *Matroid Theory*, Academic Press, London, 1976.

# Glossary

Here we list some acronyms and terminology that are used frequently in this thesis. Each entry includes a brief definition and a pointer to where the terminology is defined in the main text. Recall that $\mathcal{M} = \mathcal{M}_n$ is the set of probability measures on $\Omega = \Omega_n = \{0,1\}^n$, and that measure means probability measure.

**almost exchangeable** Measure $\mu \in \mathcal{M}$ is *almost exchangeable* if it is invariant under permutations of some subset of $n-1$ of the coordinates. Page 5.

**APP** A measure $\mu \in \mathcal{M}_{2k}$ has the *antipodal pairs property* if

$$\binom{2k}{k}^{-1} \sum_{\eta \in \Omega_{2k}, |\eta| = k} \mu(\eta)\mu(\underline{1} - \eta) \geq \binom{2k}{k-1}^{-1} \sum_{\eta \in \Omega_{2k}, |\eta| = k-1} \mu(\eta)\mu(\underline{1} - \eta).$$

Page 7.

**APU** A measure $\mu \in \mathcal{M}$ is *antipodal pairs unimodal* if the sequence $(\alpha_i)_{i=0}^n$ is unimodal, where

$$\alpha_i = \binom{n}{i}^{-1} \sum_{\eta \in \Omega, |\eta| = i} \mu(\eta)\mu(\underline{1} - \eta).$$

Page 40.

**BLC** Measure $\mu$ is BLC$[m]$ if every measure gotten from $\mu$ by imposing an external field and then projecting onto a set of size at most $m$ is ULC, and it is BLC if it is BLC$[m]$ for all $m$. Page 6.

**CAPP** A measure $\mu \in \mathcal{M}$ has the *conditional antipodal pairs property* if every measure obtained from $\mu$ by conditioning on the values of some $n-2k$ variables (for some $k$) has the APP. Page 7.

**CAPU** Measure $\mu$ is *conditionally antipodal pairs unimodal* if every measure gotten from $\mu$ by conditioning on the values of some of the coordinates is APU. Page 40.

**CFM** Measure $\mu$ is CFM if every measure gotten from $\mu$ by conditioning on the values of some of the coordinates is FM. Page 12.

**CNA** Measure $\mu$ is *conditionally negatively associated* if every measure gotten from $\mu$ by conditioning on the values of some of the coordinates is NA. Page 3.

**CNC** Measure $\mu$ is *conditionally negatively correlated* if every measure gotten from $\mu$ by conditioning on the values of some of the coordinates is NC. This is equivalent to the hereditary negative lattice condition (h-NLC). Page 3.

**exchangeable** Measure $\mu$ is *exchangeable* if it is invariant under permutations of the coordinates (equivalently, $\mu(\eta)$ depends only on $|\eta|$). Page 5.

**FM** A measure has the *Feder-Mihail property* (or is FM) if for every increasing event $\mathcal{A}$ there is a coordinate $i$ such that $\mathcal{A} \uparrow \{\eta_i = 1\}$. Page 12.

**FM+** Measure $\mu$ is FM+ if every measure gotten from $\mu$ by imposing an external field is FM. Page 12.

**GSSP** generalized Srinivasan sampling process, Page 17.

**LC** A sequence $(a_i)_{i=0}^n$ is *log-concave* if $a_i^2 \geq a_{i-1}a_{i+1}$ for $1 \leq i \leq n-1$, and a measure is LC if its rank sequence is log-concave with no internal zeros. Page 5.

**NA** Measure $\mu$ is *negatively associated* (or has *negative association*) if $\mu(\mathcal{A}\mathcal{B}) \leq \mu(A)\mu(B)$ for all increasing $\mathcal{A} \perp \mathcal{B}$. Page 2.

**NA+** Measure $\mu$ is NA+ if every measure gotten from $\mu$ by imposing an external field is NA. Page 3.

**NC** A measure $\mu \in \mathcal{M}$ has *negative correlations* if $\{\eta_i = 1\} \downarrow \{\eta_j = 1\}$ (under $\mu$) for all $i \neq j$. Page 2.

**NC+** Measure $\mu$ is NC+ if every measure gotten from $\mu$ by imposing an external field is NC. This is equivalent h-NLC+. We sometimes use *Rayleigh* for NC+. Page 3.

**NLC** Measure $\mu \in \mathcal{M}$ satisfies the *negative lattice condition* if

$$\mu(\eta)\mu(\tau) \geq \mu(\eta \wedge \tau)\mu(\eta \vee \tau)$$

for all $\eta, \tau \in \Omega$. Page 4.

**NMP** A measure $\mu$ has the *normalized matching property* if $\mu(\cdot|\sum \eta_i = l)$ stochastically dominates $\mu(\cdot|\sum \eta_i = k)$ whenever $l \geq k$. Page 20.

**PA** Measure $\mu$ is *positively associated* if $\mu(\mathcal{A}\mathcal{B}) \geq \mu(\mathcal{A})\mu(\mathcal{B})$ for all increasing events $\mathcal{A}, \mathcal{B}$. Page 1.

**Pemantle measure** Any measure that can be obtained from Bernoulli measures (measures on $\{0, 1\}$) by repeatedly taking products, imposing external fields, and rank rescaling. Page 18.

**rank rescaling** Given a nonnegative LC sequence $q = (q_0, \ldots, q_n)$, the *rank rescaling* of $\mu \in \mathcal{M}$ by $q$ is the measure $q \otimes \mu \in \mathcal{M}$ with $q \otimes \mu(\eta) \propto q_{|\eta|}\mu(\eta)$. Page 18.

**rank sequence** The *rank sequence* of $\mu \in \mathcal{M}$ is $(\mu(|\eta| = i))_{i=0}^{n}$. Page 6.

**Rayleigh** Measure $\mu$ is *Rayleigh* if every measure gotten from $\mu$ by imposing an external field is NC. This is another name for NC+. Page 4.

**SCP** Measure $\mu \in \mathcal{M}$ has the *stochastic covering property* if $\mu(\cdot|\eta_i = 0)$ stochastically covers $\mu(\cdot|\eta_i = 1)$ for every $i \in [n]$. Page 26.

**SCP*** Measure $\mu$ is SCP* if every measure gotten from $\mu$ by rank rescaling has the SCP. Page 77.

**SGSSP** super-generalized Srinivasan sampling process, Page 17.

**SLC** A nonnegative sequence $(a_i)_{i=0}^{\infty}$ is *strongly log-concave* if $ia_i^2 \geq (i+1)a_{i-1}a_{i+1}$ for all $i \geq 1$, and a measure $\nu$ on $\mathbb{N}$ is SLC if the sequence $(\nu(i))_{i=0}^{\infty}$ is. Page 49.

**SSP** Srinivasan sampling process, Page 71.

**stochastic domination** Measure $\mu$ *stochastically dominates* measure $\nu$ if $\mu(\mathcal{A}) \geq \nu(\mathcal{A})$ for all increasing $\mathcal{A}$. Page 20.

**truncation** The *truncation* of a measure $\mu \in \mathcal{M}$ to [r,s] is the conditional measure $\mu(\cdot | r \leq \sum \eta_i \leq s) \in \mathcal{M}$; it is a *k-truncation* if $s - r \leq k - 1$. Page 69.

**ULC** A sequence $(a_i)_{i=0}^n$ is *ultra-log-concave* if the sequence $(a_i/\binom{n}{i})_{i=0}^n$ is log-concave and has no internal zeros, and a measure is ULC if its rank sequence is ULC. Page 5.

**variable of positive influence** Coordinate $i \in [n]$ is a *variable of positive influence* for the pair $(f, \mu)$, where $f : \Omega \to \mathbb{R}$ and $\mu \in \mathcal{M}$, if $\eta_i \uparrow f$. Page 19.

# Vita

## Michael Neiman

| | |
|---|---|
| **2009** | Ph.D. in Mathematics, Rutgers University |
| **2003** | M.S. in Chemistry, Lehigh University |
| **2001** | B.A. in Mathematics and Chemistry, Lehigh University |
| **1999** | Graduated from Oley Valley High School |