NEGATIVE DEPENDENCE
AND THE GEOMETRY OF POLYNOMIALS

JULIUS BORCEA, PETTER BRÄNDÉN, AND THOMAS M. LIGGETT

CONTENTS

1. INTRODUCTION

Let $\mu : 2^{[n]} \to \mathbb{R}$, $[n] = \{1, \ldots, n\}$, be a function attaining non-negative values and satisfying $\sum_{S \in 2^{[n]}} \mu(S) = 1$. The function μ is said to satisfy the *positive lattice*

condition [30, 65] if

(PLC) $\mu(S)\mu(T) \leq \mu(S \cup T)\mu(S \cap T)$

for all $S, T \subseteq [n]$. The corresponding probability measure (also denoted by μ) on $2^{[n]}$ defined by $\mu(\mathcal{A}) = \sum_{S \in \mathcal{A}} \mu(S)$ is said to be *positively associated* if

$$\int F d\mu \int G d\mu \leq \int F G d\mu$$

for any pair of increasing functions F, G on $2^{[n]}$. The latter is a strong correlation inequality that yields many others as well as distributional limit theorems for models of statistical mechanics such as the ferromagnetic Ising model or certain urn models [63]. A fundamental result in the theory of positively associated random variables, namely the FKG theorem [30], asserts that PLC implies positive association. This is a powerful tool that allows verification of (global) correlation inequalities from the PLC property, which is a local condition and therefore often easier to check.

There are many important examples of *negatively dependent* "repelling" random variables in probability theory, combinatorics, stochastic processes and statistical mechanics: uniform random spanning tree measures [16], symmetric exclusion processes [52, 55], random cluster models (with $q < 1$) [33, 42, 65], balanced and Rayleigh matroids [20, 29, 68, 70, 72], competing urns models [26], etc; see, e.g., [45, 65] and the references therein for a discussion of some of these examples and several others. To add to this list, in §3 we show that both the inequalities characterizing multi-affine real stable polynomials [8, 9, 15] and Hadamard-Fischer-Kotelyansky type inequalities in matrix theory [28, 40, 39] may in fact be viewed as natural manifestations of negative dependence properties.

The reverse of the PLC inequality gives the so-called *negative lattice condition*

(NLC) $\mu(S)\mu(T) \geq \mu(S \cup T)\mu(S \cap T)$

for all $S, T \subseteq [n]$, while the usual definition of *negative association* for a probability measure μ on $2^{[n]}$ is the "negative" analog of positive association [65], namely the inequality

$$\int F d\mu \int G d\mu \geq \int F G d\mu$$

for any pair of increasing functions F, G on $2^{[n]}$ depending on *disjoint* sets of coordinates (note that the latter condition is important since by the Cauchy-Schwarz inequality F and F are always positively correlated). However, negative dependence is not nearly as robust as positive dependence. In particular, the NLC inequality *fails* to imply negative association (see Example 2.1), and there is no known local-to-global tool comparable to the FKG theorem in the theory of positive association. Thus, no corresponding theory of negatively dependent events exists as yet, but, as explained in [65], there certainly is a need for such a theory. In *op. cit.* Pemantle made a systematic study of measures on Boolean lattices with various negative dependence properties. He proved certain relations between them and conjectured numerous others, so as to pave the way for a theory of negative dependence. However, most open problems and conjectures are unresolved to this day, and the scope of the positive results so far is quite restricted.

In this paper we introduce the class of *strongly Rayleigh* probability measures by means of purely geometric conditions on the zero sets of their generating polynomials that amount to the real stability of the latter (Definitions 2.9–2.10). As we

show in §3, this class contains several important examples such as distributions at time $t \geq 0$ for symmetric exclusion processes generated by product or deterministic measures [55] (or, more generally, by strongly Rayleigh measures), and determinantal probability measures induced by positive contractions [5, 57]. Measures of the latter type include e.g. uniform random spanning tree measures [16, 57] and arise naturally in a variety of contexts pertaining to fermionic/determinantal point processes and their continuous scaling limits, eigenvalues of random matrices, number theory, non-intersecting paths, transfer current matrices, orthogonal polynomial ensembles, etc.; see the discussion in §3.3 and the references therein. Note that the topological dimension of the class of strongly Rayleigh probability measures on $2^{[n]}$ is $2^n - 1$ ([10, Remark 3.1]) whereas the class of determinantal probability measures has dimension n^2, so the former is much larger than the latter. Of course, there are also many examples of measures that satisfy some negative dependence property but are not in the strongly Rayleigh class, e.g. Counterexample 1 in §7 and the random cluster measure [33] whenever the graph contains a cycle; see §3.4.

The geometry underlying strongly Rayleigh measures provides us with powerful tools, such as fundamental complex analytic and geometric results of Gårding and Grace-Szegö-Walsh, combinatorial/probabilistic results of Feder-Mihail, matrix theoretic methods, and recent developments in the theory of stable polynomials [7, 8, 9, 15] that we use in §4 to develop a complete theory of negative dependence for this class of measures. Indeed, we show that strongly Rayleigh measures enjoy all virtues of negative dependence, including the strongest form of negative association (CNA+). In particular, this allows us to prove several conjectures made by Liggett [55], Pemantle [65], and Wagner [72], respectively, and to recover and extend Lyons' main results [57] on negative association and stochastic domination for determinantal probability measures induced by positive contractions. Moreover, we define a partial order on the set of strongly Rayleigh measures (by means of the notion of proper position for multivariate stable polynomials studied in [7, 8, 9, 15]) and use it to settle Pemantle's questions and conjectures on stochastic domination for truncations of "negatively dependent" measures [65].

A series of conjectures, first appearing in print in [65, Conjecture 4] but of obscure folklore origin, states that the various negative dependent properties studied in *loc. cit.* give rise to so-called ultra log-concave rank sequences (see §2.1 below for the definition). Subsequently, in [72] one of these conjectures was coined "The Big Conjecture" since it would imply the validity of Mason's long-standing conjecture in enumerative graph/matroid theory [62] for a large class of matroids. In §7 we construct the first counterexamples in the literature to all of these conjectures except Mason's, which once again confirms the delicate nature of negative dependence. After this work was made publicly available on www.arxiv.org, other counterexamples to some of these conjectures were reported [48, 60]. Markström [60] proved that negative association does not imply unimodality, and Kahn-Neiman [48] later showed that not even strong negative association (CNA+) implies unimodality. Note that the rank sequence of the CNA+ measure constructed in Counterexample 1 of §7 is neither ultra nor strong log-concave but it is log-concave.

The problem of describing natural negative dependence properties that are preserved by symmetric exclusion evolutions has attracted some attention in the theory of interacting particle systems and Markov processes [52, 54, 55]. In §4.4 we prove

that the class of strongly Rayleigh measures is invariant under partial symmetriza-
tion (this and several other properties fail, for example, for Rayleigh measures). As
a consequence, in §5 we provide an answer to the aforementioned problem and show
that if the initial distribution of a symmetric exclusion process is strongly Rayleigh,
then so is the distribution at time $t \geq 0$; therefore, by the results in §4.2, the latter
distribution is strongly negatively associated (CNA+). In particular, this solves an
open problem of Pemantle [65] and Liggett [55] stating that the distribution of a
symmetric exclusion process at time $t \geq 0$ with non-random/deterministic initial
configuration is negatively associated, and shows that the same is actually true
whenever the initial distribution is strongly Rayleigh. In a later paper [56], Liggett
has applied these results to prove convergence to the normal and Poisson laws for
various functionals of the symmetric exclusion process.

In §6 we establish equivalences between several negative dependence properties
for the class of almost symmetric/exchangeable measures that we define in §2.1.
This extends Pemantle's corresponding theorem for symmetric measures and con-
firms his conjecture on strong negative association (CNA+) in the almost exchange-
able case. The results in §6 are also useful in §7, as they allow us to show that the
examples we construct there provide counterexamples to the series of conjectures on
ultra log-concave rank sequences proposed in [65, 72] to which we already alluded.
We note that Corollary 6.6 in §6 was subsequently proved by different methods in
[48], where it was additionally shown that almost exchangeable measures satisfy
the (strong) Feder-Mihail property.

The aforementioned connections between the negative dependence properties
satisfied by strongly Rayleigh measures and the geometry of zero sets of their gen-
erating polynomials are much in the same spirit as e.g. Rota's philosophy about
the ubiquity of zeros of polynomials in combinatorics [67]. In the present context,
these connections actually prove to have fruitful consequences for both the theory
of negative dependence and the theory of real stable and hyperbolic polynomials, as
they give new insight and further potential applications in these and related areas.

2. A PLETHORA OF NEGATIVE DEPENDENCE PROPERTIES AND CONJECTURES

In this section we explain in terms of generating polynomials the probabilistic
notions and negative dependence properties studied in e.g. [55, 65, 72], and we
clarify the connections between these as well as new classes of probability measures
that we introduce below. We then formulate the problems and conjectures made
in [55, 65, 72] that we solve in the next sections.

2.1. **Negative dependence concepts.** Recall that a (real or complex) multivari-
ate polynomial is said to be *multi-affine* if it has degree at most one in each variable.
Denote by $\mathfrak{P}_n$, $n \in \mathbb{N}$, the set of all probability measures on the Boolean algebra
$2^{[n]}$. (In the finite case we will always assume that the σ-algebra is the full alge-
bra.) For $i \in [n]$ the i-th coordinate function on $2^{[n]}$ is an atomic (binary) random
variable given by $X_i(S) = 1$ if $i \in S$ and 0 otherwise, where $S \subseteq [n]$, while the char-
acteristic function χ_S of S is defined by $\chi_S(T) = 1$ if $T = S$ and 0 otherwise. Using
the inclusion-exclusion principle one can show that any such characteristic function
may be written as a multi-affine polynomial in $X_1, \ldots, X_n$. Any scalar function
$F = \sum_{S \subseteq [n]} F(S)\chi_S$ on $2^{[n]}$ may therefore be viewed as a multi-affine polynomial
in $X_1, \ldots, X_n$. By abuse of notation the latter is written as $F(X_1, \ldots, X_n)$, and

then $\mu(F) = \int F d\mu$. However, in what follows the X_i's always stand for (probabilistic) random variables, and we will use other variables (such as x, y, z, w) when working with arbitrary functions/polynomials.

Let $\mathcal{P}_n$ be the set of all multi-affine polynomials in n variables $f(z_1, \ldots, z_n)$ with non-negative coefficients such that $f(\mathbf{1}) = 1$, where $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$ denotes the "all ones vector". There is a 1-1 correspondence between $\mathfrak{P}_n$ and $\mathcal{P}_n$: if $\mu \in \mathfrak{P}_n$ we may form its generating polynomial, namely

$$g_\mu(z) = \int z^S d\mu(S) = \sum_{S \subseteq [n]} \mu(S) z^S, \quad z = (z_1, \ldots, z_n), \quad z^S := \prod_{i \in S} z_i,$$

and if

(1) $$f(z) = \sum_{S \subseteq [n]} a_S z^S \in \mathcal{P}_n$$

we define a measure μ_f on $2^{[n]}$ by setting $\mu_f(S) = a_S$, $S \subseteq [n]$. It is clear that $g_\nu \in \mathcal{P}_n$, $\mu_f \in \mathfrak{P}_n$, $g_{\mu_f} = f$, and $\mu_{g_\nu} = \nu$ for any $\nu \in \mathfrak{P}_n$, $f \in \mathcal{P}_n$. For convenience, we will sometimes use the symbol ∂_i for $\partial/\partial z_i$ $1 \le i \le n$, and the multi-index notation $\partial^S = \prod_{i \in S} \partial_i$, $S \subseteq [n]$.

The NLC property for measures in $\mathfrak{P}_n$ translates into the corresponding property – again called the *negative lattice condition* and denoted by NLC – for polynomials in $\mathcal{P}_n$: a polynomial $f \in \mathcal{P}_n$ satisfies NLC if and only if its corresponding measure $\mu_f \in \mathfrak{P}_n$ does. This amounts to saying that

(NLC) $$\partial^S f(0, \ldots, 0) \, \partial^T f(0, \ldots, 0) \ge \partial^{S \cup T} f(0, \ldots, 0) \, \partial^{S \cap T} f(0, \ldots, 0),$$

which in notation (1) becomes $a_S a_T \ge a_{S \cup T} a_{S \cap T}$ for all $S, T \subseteq [n]$.

The so-called *"closure" operations* discussed in e.g. [65, II.A] may also be reformulated in terms of generating polynomials:

(i) *Products.* If $\mu_1 \in \mathfrak{P}_{n_1}$ has generating polynomial $g_{\mu_1}(z_1, \ldots, z_{n_1}) \in \mathcal{P}_{n_1}$ and $\mu_2 \in \mathfrak{P}_{n_2}$ has generating polynomial $g_{\mu_2}(z_1, \ldots, z_{n_2}) \in \mathcal{P}_{n_2}$, then the product $\mu_1 \times \mu_2$ is the measure $\mu \in \mathfrak{P}_{[n_1 + n_2]}$ with generating polynomial $g_\mu(z_1, \ldots, z_{n_1 + n_2}) = g_{\mu_1}(z_1, \ldots, z_{n_1}) g_{\mu_2}(z_{n_1 + 1}, \ldots, z_{n_1 + n_2}) \in \mathcal{P}_{n_1 + n_2}$.

(ii) *Projections.* Given $S \subseteq [n]$ and $\mu \in \mathfrak{P}_n$ the projection of μ onto 2^S is the measure $\mu' \in \mathfrak{P}_{|S|}$ with generating polynomial

$$g_\mu(z_1, \ldots, z_n)\big|_{z_i = 1, \, i \in [n] \setminus S} \in \mathcal{P}_{|S|}.$$

Note that in order for this definition to make sense we should also relabel the variables using indices in $[|S|]$. However, we allow ourselves this abuse of notation here and in what follows.

(iii) *Conditioning.* Let $\mu \in \mathfrak{P}_n$ with generating polynomial $g_\mu(z_1, \ldots, z_n) \in \mathcal{P}_n$ and fix some $i \in [n]$. The measure obtained from μ by conditioning the i-th random variable X_i to be 0 is the measure on $2^{[n] \setminus \{i\}}$ with generating polynomial

$$\frac{g_\mu(z_1, \ldots, z_n)\big|_{z_i = 0}}{g_\mu(z_1, \ldots, z_n)\big|_{z_i = 0, \, z_j = 1, \, j \ne i}} \in \mathcal{P}_{n-1},$$

while the measure obtained from μ by conditioning X_i to be 1 is the measure on $2^{[n]\setminus\{i\}}$ with generating polynomial

$$\lim_{z_i \to \infty} \frac{g_\mu(z_1, \ldots, z_n)}{g_\mu(z_1, \ldots, z_n)\big|_{z_j=1, \, j\neq i}} = \frac{\partial_i g_\mu(z_1, \ldots, z_n)}{\partial_i g_\mu(1, \ldots, 1)} \in \mathcal{P}_{n-1}.$$

Note that these conditional probability measures are well defined provided that the denominators appearing in the above expressions are non-zero.

(iv) *External fields* (as pointed out in [65, II.A], this name is borrowed from the Ising model). If $\mu \in \mathfrak{P}_n$ has generating polynomial $g_\mu(z_1, \ldots, z_n) \in \mathcal{P}_n$ and a_i, $1 \leq i \leq n$, are non-negative real numbers, then the measure obtained from μ by imposing the external field $(a_1, \ldots, a_n)$ is the measure in $\mathfrak{P}_n$ with generating polynomial

$$\frac{g_\mu(a_1 z_1, \ldots, a_n z_n)}{g_\mu(a_1, \ldots, a_n)} \in \mathcal{P}_n,$$

which is well defined provided that $g_\mu(a_1, \ldots, a_n) \neq 0$.

(v) *Symmetrization.* Denote the symmetric group on n elements by $\mathfrak{S}_n$. For $\sigma \in \mathfrak{S}_n$ and $S \subseteq [n]$ let $\sigma(S) = \{\sigma(s) : s \in S\}$. Given $\mu \in \mathfrak{P}_n$ define a measure $\sigma(\mu) \in \mathfrak{P}_n$ by setting $\sigma(\mu)(S) = \mu(\sigma(S))$, $S \subseteq [n]$. The (full or complete) symmetrization of μ is the measure

$$\mu_s = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sigma(\mu) \in \mathfrak{P}_n.$$

Note that the generating polynomial of μ_s is

(2)
$$g_{\mu_s}(z) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} g_{\sigma(\mu)}(z) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sum_{S \subseteq [n]} a_{\sigma(S)} z^S$$
$$= \sum_{k=0}^{n} \frac{\sum_{|S|=k} a_S}{\binom{n}{k}} e_k(z_1, \ldots, z_n) \in \mathcal{P}_n,$$

where $\sum_{S \subseteq [n]} a_S z^S = g_\mu(z)$ and e_k, $0 \leq k \leq n$, is the k-th elementary symmetric function (see [72, §3.4] for a proof of the last identity).

(vi) *Partial symmetrization.* Let $\mu \in \mathfrak{P}_n$, $0 \leq \theta \leq 1$, $1 \leq i < j \leq n$, and $\tau = (ij) \in \mathfrak{S}_n$ be the transposition that permutes i and j. The partial symmetrization of μ with respect to τ and θ is the measure

(3)
$$\mu^{\tau,\theta} = \theta\mu + (1-\theta)\tau(\mu) \in \mathfrak{P}_n$$

whose generating polynomial is obviously $g_{\mu^{\tau,\theta}} = \theta g_\mu + (1-\theta)g_{\tau(\mu)} \in \mathcal{P}_n$.

Remark 2.1. All these "closure" operations are in fact well defined for arbitrary complex measures on $2^{[n]}$ if one drops the normalization factors appearing in the denominators of the above expressions (these were used just to make sure that the resulting measures are again probability measures). In §4 we will actually use the partial symmetrization procedure (3) for complex measures on $2^{[n]}$.

The corresponding "closure" operations for generating polynomials – i.e., *products*, *projections*, *conditioning*, *external fields*, *symmetrization*, *partial symmetrization* and *truncations* (cf. Definition 2.15 in §2.5)– are defined simply by considering the resulting polynomials for each of the operations discussed above for measures.

Definition 2.1. A complex measure μ on $2^{[n]}$ is called *symmetric* or *exchangeable* if $\sigma(\mu) = \mu$ for any $\sigma \in \mathfrak{S}_n$, or equivalently, its generating polynomial g_μ is symmetric in all n variables. We say that μ is *almost symmetric* or *almost exchangeable* if g_μ is symmetric in all but possibly one variable.

Definition 2.2. If $\mathcal{A}$ is a collection of subsets of $[n]$, we let $\chi_\mathcal{A} : 2^{[n]} \to \mathbb{R}$ be the characteristic function of $\mathcal{A}$ defined by $\chi_\mathcal{A}(S) = 1$ if $S \in \mathcal{A}$ and $\chi_\mathcal{A}(S) = 0$ if $S \notin \mathcal{A}$. An *increasing event* $\mathcal{A}$ on $2^{[n]}$ is a collection of subsets of $[n]$ that is closed upwards under containment, i.e., if $A \in \mathcal{A}$ and $A \subseteq B \subseteq [n]$, then $B \in \mathcal{A}$. Such an event *depends* only on the set $\mathcal{I}(\mathcal{A}) := I_1 \cup \cdots \cup I_m$ (that is, on the variables/coordinate functions X_i with $i \in \mathcal{I}(\mathcal{A})$), where $I_1, \ldots, I_m$ are the minimal sets of $\mathcal{A}$ with respect to inclusion. Any (non-identically zero) non-negative increasing function f on $2^{[n]}$ may be written as $f = \sum_{i=1}^k a_i \chi_{\mathcal{A}_i}$ for some increasing events $\mathcal{A}_i$ on $2^{[n]}$ and $a_i > 0$, $1 \leq i \leq k$. Clearly, f depends on the set $\mathcal{I}(f) := \mathcal{I}(\mathcal{A}_1) \cup \cdots \cup \mathcal{I}(\mathcal{A}_k)$.

The weakest negative dependence property is the following (see, e.g., [65]).

Definition 2.3. A measure $\in \mathfrak{P}_n$ is said to be *pairwise negatively correlated*, or p-NC for short, if $\mu(X_i)\mu(X_j) \geq \mu(X_iX_j)$ whenever $1 \leq i \neq j \leq n$. In terms of the generating polynomial g_μ this translates into the inequalities

$$\partial_i g_\mu(\mathbf{1}) \partial_j g_\mu(\mathbf{1}) \geq \partial_i \partial_j g_\mu(\mathbf{1}), \quad 1 \leq i \neq j \leq n,$$

where $\mathbf{1} \in \mathbb{R}^n$ is as before the "all ones vector". We say that a polynomial $f \in \mathcal{P}_n$ is p-NC (pairwise negatively correlated) if it satisfies the above inequalities.

As shown by the following example, NLC does not necessarily imply p-NC.

Example 2.1. The measure $\mu \in \mathfrak{P}_4$ with generating polynomial

$$g_\mu(z_1, z_2, z_3, z_4) = \frac{1}{2}(z_1z_2 + z_3z_4)$$

trivially satisfies NLC but $\mu(X_1X_2) = \frac{1}{2}$ while $\mu(X_1)\mu(X_2) = \frac{1}{4}$, so μ is not p-NC.

Definition 2.4 (Definition 2.4 in [65]). A measure $\mu \in \mathfrak{P}_n$ or a polynomial $f \in \mathcal{P}_n$ satisfies the *hereditary negative lattice condition* or h-NLC if every projection satisfies NLC. One further says that μ (respectively, f) satisfies the *strong hereditary negative lattice condition* or h-NLC+ if every measure (respectively, polynomial) obtained from μ (respectively, f) by imposing an external field satisfies h-NLC.

Definition 2.5. A polynomial $f \in \mathcal{P}_n$ is called a *Rayleigh polynomial* if

$$(4) \qquad \frac{\partial f}{\partial z_i}(x)\frac{\partial f}{\partial z_j}(x) \geq \frac{\partial^2 f}{\partial z_i \partial z_j}(x)f(x)$$

for all $x = (x_1, \ldots, x_n) \in \mathbb{R}_+^n$ and $1 \leq i, j \leq n$, where as usual $\mathbb{R}_+ = (0, \infty)$. More generally, a multi-affine polynomial in $\mathbb{R}[z_1, \ldots, z_n]$ with non-negative coefficients is called a *Rayleigh polynomial* if it satisfies the above condition. A measure $\mu \in \mathfrak{P}_n$ is said to be a *Rayleigh measure* if its generating polynomial g_μ is Rayleigh.

The notion of a Rayleigh polynomial was introduced in [72, §3.1], the terminology being motivated by its similarity with the Rayleigh monotonicity property of the (Kirchhoff) effective conductance of linear resistive electrical networks; see [20, 73]. It was first considered for uniform measures on the set of bases of a matroid [20] as a strengthening of weaker notions studied in [29, 69]. The following properties of Rayleigh polynomials are consequences of Definition 2.5.

Proposition 2.1. *If $f(z_1, \ldots, z_n)$ is Rayleigh, then so are the following polynomials:*

    (1) $\partial^S f(z_1, \ldots, z_n)$ *for any $S \subseteq [n]$;*
    (2) $f(z_1 + \alpha_1, \ldots, z_n + \alpha_n)$ *whenever $\alpha_i \geq 0$, $i \in [n]$;*
    (3) $f(z_1, \ldots, z_n)|_{z_i = \alpha_i}$ *for any $\alpha_i \geq 0$, $1 \leq i \leq n$;*
    (4) $f(a_1 z_1, \ldots, a_n z_n)$ *for all $a_i \geq 0$, $i \in [n]$;*
    (5) *the "inversion" of f, i.e., the polynomial $z_1 \cdots z_n f(z_1^{-1}, \ldots, z_n^{-1})$;*
    (6) *the polynomial in nk variables $f\left(k^{-1} \sum_{i=1}^{k} z_{1i}, \ldots, k^{-1} \sum_{i=1}^{k} z_{ni}\right)$.*

To prove (1) note that by induction it is enough to do it for $S = \{i\}$ and arbitrary $i \in [n]$. This follows by first writing

$$f(z_1, \ldots, z_n) = z_i \partial_i f(z_1, \ldots, z_n) + f(z_1, \ldots, z_n)|_{z_i = 0}$$

and then letting $z_i \to \infty$ in the Rayleigh inequalities (4) for f corresponding to pairs of variables indexed by distinct $j, k \in [n] \setminus \{i\}$.

Proposition 2.2. *A measure in $\mathfrak{P}_n$ or a polynomial in $\mathcal{P}_n$ is Rayleigh if and only if it is h-NLC+.*

Proof. In [72, Theorem 4.4] it was proved that if $f \in \mathcal{P}_n$ is Rayleigh, then f satisfies NLC, which combined with Proposition 2.1 shows that f is h-NLC+. To prove the converse statement, let $i, j \in [n]$ with $i \neq j$ and set

$$g(z_i, z_j) = f(z_1, \ldots, z_n)|_{z_k = 1, \, k \in [n] \setminus \{i,j\}}.$$

If f is h-NLC+, then g satisfies NLC (which in this case implies that $g(z_i, z_j)$ is in fact a real stable polynomial; see Definition 2.9 in §2.2 and Theorem 4.1 in §4). In particular, we deduce that $\partial_i \partial_j f(\mathbf{1}) \leq \partial_i f(\mathbf{1}) \partial_j f(\mathbf{1})$, where $\mathbf{1}$ stands for the "all ones vector" as in Definition 2.3. Using the external field condition given in Definition 2.4 we see that the aforementioned inequality also holds for the polynomial $f(x_1 z_1, \ldots, x_n z_n)$ whenever $x_i \geq 0$, $i \in [n]$, which by Definition 2.5 amounts to saying that f is Rayleigh. $\qquad\square$

Remark 2.2. It is interesting to note that one actually has the following analog of Proposition 2.2 for the h-NLC property: a measure $\mu \in \mathfrak{P}_n$ – or its generating polynomial $g_\mu \in \mathcal{P}_n$ – is h-NLC if and only if g_μ satisfies the Rayleigh inequalities (4) for all $x = (x_1, \ldots, x_n)$ with $x_i \in \{0, 1, \infty\}$, $i \in [n]$.

Definition 2.6. A *homogeneous Rayleigh measure* is one whose generating polynomial is homogeneous and Rayleigh. The set of all measures that are projections of homogeneous Rayleigh measures is denoted by PHR.

Let us now recall the key concept of negative association. For probability measures this notion is usually defined as follows (cf., e.g., [55, 65, 72]).

Definition 2.7. A measure $\mu \in \mathfrak{P}_n$ is called *negatively associated* or NA if

$$\int F d\mu \int G d\mu \geq \int FG d\mu$$

for any increasing functions F, G on $2^{[n]}$ that depend on disjoint sets of coordinates (cf. Definition 2.2). One says that μ is *conditionally negatively associated* or CNA if each measure obtained from μ by conditioning on some (or none) of the values of the variables is NA. Finally, μ is called *strongly conditionally negatively associated* or

CNA+ if each measure obtained from μ by imposing external fields and projections is CNA.

Remark 2.3. It is clear from the definitions that each of the five properties h-NLC, Rayleigh/h-NLC+, NA, CNA, and CNA+ implies p-NC.

Remark 2.4. As explained in [65], one has the following subordination relations: CNA $\Rightarrow$ h-NLC and CNA+ $\Rightarrow$ Rayleigh/h-NLC+. In §4.2 and §7 below we show that PHR $\Rightarrow$ CNA+ and that this implication is strict.

We note that NA can fail to imply NLC even for symmetric measures:

Example 2.2. Let as before e_k, $0 \le k \le n$, be the k-th elementary symmetric function in n variables and consider the measure $\mu \in \mathfrak{P}_3$ with generating polynomial

$$g_\mu(z_1, z_2, z_3) = \frac{1}{15} \left[3 + 2e_1(z_1, z_2, z_3) + 2e_2(z_1, z_2, z_3) \right].$$

It is not difficult to check that μ is NA but not NLC.

Definition 2.8. The *diagonal specialization* of a polynomial $f \in \mathbb{C}[z_1, \ldots, z_n]$ is the univariate polynomial

$$t \mapsto \Delta(f)(t) := f(t, \ldots, t).$$

A real sequence $\{a_k\}_{k=0}^n$ is *log-concave* if $a_k^2 \ge a_{k-1}a_{k+1}$, $1 \le k \le n-1$, and it is said to have *no internal zeros* if the indices of its non-zero terms form an interval (of non-negative integers). We say that a *non-negative* sequence $\{a_k\}_{k=0}^n$ is

- LC if it is log-concave with no internal zeros;
- SLC (strongly log-concave) if the sequence $\{k!a_k\}_{k=0}^n$ is LC;
- ULC (ultra log-concave) if the sequence $\left\{a_k/\binom{n}{k}\right\}_{k=0}^n$ is LC.

Clearly, ULC $\Rightarrow$ SLC $\Rightarrow$ LC. If $\mu \in \mathfrak{P}_n$ the sequence

$$\left\{ \mu\left(\sum_{i=1}^n X_i = k \right) \right\}_{k=0}^n = \left\{ \frac{\Delta(g_\mu)^{(k)}(0)}{k!} \right\}_{k=0}^n$$

is called the *rank sequence* of μ as well as of its generating polynomial $g_\mu \in \mathcal{P}_n$. A measure in $\mathfrak{P}_n$ is then said to be ULC, SLC or LC if its rank sequence is ULC, SLC or LC, respectively.

2.2. **Strongly Rayleigh measures.** Motivated by similar notions in the theory of multivariate entire functions studied by Levin [50], the following definition was made in [7, 8, 9] (see also [10, 11]).

Definition 2.9. A polynomial $f \in \mathbb{C}[z_1, \ldots, z_n]$ is called *stable* if $f(z_1, \ldots, z_n) \ne 0$ whenever $\mathfrak{Im}(z_j) > 0$ for $1 \le j \le n$. A stable polynomial with all real coefficients is called *real stable*.

Multivariate stable polynomials are intimately connected with other fundamental objects such as hyperbolic polynomials, Lee-Yang polynomials and polynomials with the half-plane property. Polynomials of the aforementioned types are widely encountered in both mathematics and physics [3, 12, 7, 8, 9, 10, 15, 19, 36, 37, 51] (see also §4.1).

Let us define a class of measures induced by stable polynomials.

Definition 2.10. A measure $\mu \in \mathfrak{P}_n$ is called *strongly Rayleigh* if its generating polynomial $g_\mu \in \mathcal{P}_n$ is (real) stable. For convenience, we will sometimes refer to a real stable polynomial $f \in \mathcal{P}_n$ as a *strongly Rayleigh polynomial*.

Remark 2.5. As we explain in §3.1 and §4, a strongly Rayleigh measure is necessarily Rayleigh. The terminology adopted is motivated by the fact that the equivalent condition for real stability for multi-affine polynomials in Theorem 4.1 is a natural strengthening of the definition of a Rayleigh polynomial/measure (Definition 2.5). The term is taken from [20], where a matroid was defined to be strongly Rayleigh if (5) holds for its bases-generating polynomial. However, it was not known then that condition (5) is equivalent to stability.

A simple albeit important example of strongly Rayleigh measure is the following.

Definition 2.11. A *product measure* on $2^{[n]}$ is a measure $\mu \in \mathfrak{P}_n$ with generating polynomial g_μ of the form

$$g_\mu(z_1, \ldots, z_n) = \prod_{i=1}^{n}(q_i z_i + 1 - q_i), \quad 0 \le q_i \le 1, \, i \in [n].$$

One says that μ is *deterministic* or *non-random* if it is a point mass on $2^{[n]}$, i.e., if there exists $S \subseteq [n]$ such that $q_i = 1$ if $i \in S$ and $q_j = 0$ for $j \in [n] \setminus S$.

In §3 we show that the class of strongly Rayleigh measures contains numerous other important examples of measures appearing in matrix theory, combinatorics, probability theory, and particle systems, and in §4 we study this class in detail.

2.3. **Symmetric homogenization.** We will now define a symmetric homogenization procedure for arbitrary measures in $\mathfrak{P}_n$. This natural construction proves to be quite useful for studying strongly Rayleigh measures (§4.1).

Definition 2.12. Given a measure $\mu \in \mathfrak{P}_n$ define a new measure $\mu_{sh} \in \mathfrak{P}_{2n}$ called the *symmetric homogenization* of μ by

$$\mu_{sh}(S) = \begin{cases} \mu(S \cap [n]) \binom{n}{|S \cap [n]|}^{-1} & \text{if } |S| = n, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the generating polynomial of μ_{sh} may be written as

$$g_{\mu_{sh}}(z_1, \ldots, z_n, z_{n+1}, \ldots, z_{2n}) = \sum_{S \subseteq [n]} \mu(S) \binom{n}{|S|}^{-1} z^S e_{n-|S|}(z_{n+1}, \ldots, z_{2n}),$$

where $z = (z_1, \ldots, z_n)$ and where e_k, $0 \le k \le n$, denotes the k-th elementary symmetric function. Clearly, $g_{\mu_{sh}}$ is symmetric in the variables $z_{n+1}, \ldots, z_{2n}$ and μ_{sh} projects to μ, that is, $g_{\mu_{sh}}(z_1, \ldots, z_n, 1, \ldots, 1) = g_\mu(z_1, \ldots, z_n)$. It is not difficult to show that $g_{\mu_{sh}}$ is actually the unique polynomial in $\mathcal{P}_{2n}$ of total degree n that is symmetric in $z_{n+1}, \ldots, z_{2n}$ and satisfies the latter identity. Therefore, μ_{sh} is canonically determined by the aforementioned conditions.

2.4. **Measures on countably infinite sets.** In §5 we will need to extend some of the notions of negative dependence to probability measures on 2^E, where E is a countably infinite set. This is canonically done as follows.

Definition 2.13. Let μ be a probability measure on 2^E, where E is a countably infinite set satisfying

(A)   for all $x \in E$ the set $\left\{T \in 2^E : x \in T\right\}$ is measurable.

Let $\mathcal{P}$ be a property defined for measures on power-sets of finite sets. We say that μ has property $\mathcal{P}$ provided that the projection of μ to all finite subsets of E has property $\mathcal{P}$. (Note that condition (A) is necessary for projections on finite sets to be well-defined.)

2.5. **Towards a theory of negative dependence: Problems and solutions.** A series of problems and conjectures was proposed by several authors as natural steps toward building as general a theory of negative dependence as possible. In [65, §III] Pemantle pointed out that although they are desirable, the many negative dependence properties listed in §2.1 might not be mutually satisfiable, and he suggested a problem that may be formulated as follows.

Problem 2.1. Find a natural and useful class of "negatively dependent measures".

Note that there may well exist several such classes, if indeed any, depending on the applications that one has in mind. We show that the class of strongly Rayleigh measures introduced in §2.2 enjoys all the negative dependence properties defined in §2.1 as well as other properties, which provides an answer to Problem 2.1.

The problem of describing natural negative dependence properties that are preserved by symmetric exclusion evolutions has attracted some attention in the theory of interacting particle systems and Markov processes [52, 54, 55]; see Problem 3.5 in §3.5 and the discussion therein. In particular, the following conjecture was implicitly made in [65, p. 1374] and [55, p. 551].

Conjecture 2.2. *If the initial configuration of a symmetric exclusion process is non-random/deterministic, then the distribution at time t is NA for all $t > 0$.*

In §5 we prove a stronger form of this conjecture using the results on strongly Rayleigh measures obtained in §4.

There are several conjectures, first published in [65] but apparently of obscure folklore origin (cf. [72]), stating that various negative dependent properties give rise to ULC rank sequences. The strongest form of these is contained in the first part of Conjecture 4 in Pemantle's paper [65] (see also [72, Conjecture 3.9]) and reads as follows.

Conjecture 2.3. *If $\mu \in \mathfrak{P}_n$ is NA, then it is ULC.*

Inspired by Conjecture 2.3 and motivated by Mason's long-standing conjecture in enumerative graph/matroid theory [62], Wagner made the following conjecture in [72, Conjecture 3.4] and baptized it "The Big Conjecture" since – if true – it would prove the validity of Mason's conjecture for a large class of matroids.

Conjecture 2.4. *Any Rayleigh measure $\mu \in \mathfrak{P}_n$ is ULC.*

Since we now know that the Rayleigh and h-NLC+ properties are equivalent (Proposition 2.2), Conjecture 2.4 is actually already contained in the second part of Conjecture 4 in Pemantle's paper [65]. The weakest form of the latter conjecture is the one stated below.

Conjecture 2.5. *If $\mu \in \mathfrak{P}_n$ is CNA+, then it is ULC.*

In §7 we construct examples of measures in $\mathfrak{P}_n$ for any $n \geq 20$ that disprove all these three conjectures (for the strongest of these, namely Conjecture 2.3, we can actually find counterexamples already in $\mathfrak{P}_3$ and $\mathfrak{P}_4$; see Remark 7.2). In fact, our examples show that the strongest negative association property (CNA+) need not even give rise to an SLC rank sequence (Definition 2.8), thus invalidating a still weaker version of Conjecture 2.5. Moreover, these same examples show that neither the PHR property (Definition 2.6) nor indeed any of the properties weaker than strongly Rayleigh implies the ULC property, and also that Conjecture 2.5 fails even under extra assumptions like those suggested in [12, Problem 6]. This confirms once again the delicate nature of negative dependence.

The FKG theorem [30] is a powerful tool that allows one to establish (global) positive correlation inequalities and limit theorems from a local (usually easier to check) condition, namely the positive lattice condition PLC. By contrast, there is as yet no general and practical local-to-global device for negative dependence/association. The NLC property is not closed under projections and implies neither the NA nor the (weaker) p-NC property, so the negative version of the FKG theorem fails. As a potential remedy to this situation, in [65, Conjecture 2] Pemantle conjectured that the implications CNA $\Rightarrow$ h-NLC and CNA+ $\Rightarrow$ Rayleigh/h-NLC+ (established in [65]; cf. Remark 2.4) are actually equivalences:

Conjecture 2.6. *h-NLC $\Rightarrow$ CNA and Rayleigh/h-NLC+ $\Rightarrow$ CNA+.*

We prove Conjecture 2.6 for almost symmetric measures (cf. Definition 2.1). Since PHR measures are CNA+ (Theorem 4.10), we are naturally led to consider the following related problem.

Problem 2.7. Is PHR = Rayleigh/h-NLC+?

Note that if true, this would prove the second part of Conjecture 2.6. In §7 we show that the answer to Problem 2.7 is negative, and that in fact the class of PHR measures is strictly contained in the class of CNA+ measures. The question whether Conjecture 2.6 is true in full generality therefore remains open.

We summarize the implications between the various notions of negative dependence discussed so far in Figure 1. Compared with the one in [65, §II.B], this diagram contains five extra classes – three "old" (ULC, p-NC, NLC) and two "new" (strongly Rayleigh, PHR) – and disregards the classes called JNRD, JNRD+ in [65].

As we already noted, the above examples and those in the next sections show that the left horizontal, down left, and all vertical implications in Figure 1 are strict, while the question as to whether the right horizontal implications are also strict is open (cf. Conjecture 2.6). Note also that there is no arrow pointing towards ULC from anything "below" strongly Rayleigh.

In [65, §I.B & §III.C] it was pointed out that the notions of stochastic domination and truncation are also useful when studying (positive or) negative dependence.

Definition 2.14. If $\mu, \nu \in \mathfrak{P}_n$ are such that $\mu(\mathcal{A}) \geq \nu(\mathcal{A})$ for any increasing event $\mathcal{A}$ on $2^{[n]}$, one says that μ *stochastically dominates* ν, written $\mu \succcurlyeq \nu$ or $\nu \preccurlyeq \mu$.

Definition 2.15. Let $\mu \in \mathfrak{P}_n$ and $1 \leq p \leq q \leq n$. The *truncation* of μ to $[p, q]$ is the conditional measure

$$\mu_{p,q} := \left(\mu \,\middle|\, p \leq \sum_{i=1}^n X_i \leq q \right) \in \mathfrak{P}_n,$$
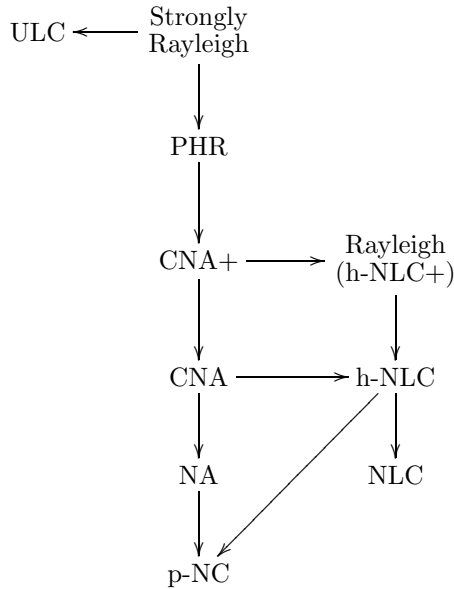
FIGURE 1. Subordination relations between negative dependence properties.

which is well defined provided that $\mu\left(\{S \in 2^{[n]} : p \leq |S| \leq q\}\right) \neq 0$. The generating polynomial of $\mu_{p,q}$ is thus given by

$$g_{\mu_{p,q}}(z_1, \ldots, z_n) = \frac{g_{p,q}(z_1, \ldots, z_n)}{g_{p,q}(1, \ldots, 1)} \in \mathcal{P}_n,$$

where $g_{p,q}(z) = \sum_{p \leq |S| \leq q} a_S z^S$ and $\sum_{S \subseteq [n]} a_S z^S = g_\mu(z)$.

For simplicity, we let $\mu_k = \mu_{k,k}$, $0 \leq k \leq n$. In [45, 65] and the references therein it has been discussed whether truncations of measures are operations that preserve negative dependence properties. Question 10 of [65] asks the following.

Problem 2.8. Under what hypotheses on μ can one prove that $\mu_k \preccurlyeq \mu_{k+1}$?

Moreover, the next conjecture was made in [65, Conjecture 8].

Conjecture 2.9. *Suppose that μ is CNA+ and that*

$$\mu(\{S \in 2^{[n]} : |S| = k\})\mu(\{S \in 2^{[n]} : |S| = k + 1\}) > 0.$$

Then $\mu_k \preccurlyeq \mu_{k+1}$.

To underline the relevance of these questions, in [65, §III.C] Pemantle argued that a combination of positive answers to Problem 2.8 and Conjecture 2.9 (for Rayleigh measures) would imply Conjecture 2.6. In §4.3 we give an answer to Problem 2.8 and prove that its conclusion (hence also that of Conjecture 2.9) holds for strongly Rayleigh measures, while in §7 we disprove Conjecture 2.9 in its full generality.

3. NEGATIVE DEPENDENCE IN MATHEMATICS AND PHYSICS

3.1. **Real stable polynomials and transcendental entire functions.** In [15] it was proved that a multi-affine polynomial $f \in \mathbb{R}[z_1, \ldots, z_n]$ is stable (in the sense

of Definition 2.9) if and only if

$$(5) \qquad \frac{\partial f}{\partial z_i}(x)\frac{\partial f}{\partial z_j}(x) \geq \frac{\partial^2 f}{\partial z_i \partial z_j}(x)f(x)$$

for all $x \in \mathbb{R}^n$ and $1 \leq i < j \leq n$ (see also Theorem 4.1 in §4 below). Therefore, a stable multi-affine polynomial with non-negative coefficients is Rayleigh/h-NLC+. As we will prove in §4, such polynomials are in fact CNA+.

Clearly, by setting all variables equal in a stable polynomial one gets a univariate stable polynomial. Hence, if $f(z_1, \ldots, z_n)$ is stable, then so is its diagonal specialization (see Definition 2.8)

$$\Delta(f)(t) = \sum_{k=0}^{n} \binom{n}{k} c_k t^k.$$

A univariate polynomial with real coefficients is stable if and only if it has all real zeros, so the familiar Newton inequalities (cf., e.g., [22, 38]) imply that f satisfies the conclusion of Conjecture 2.4 if f is stable with non-negative coefficients.

Multivariate generalizations of Newton's inequalities for homogeneous real stable polynomials with non-negative coefficients were obtained in [8, Theorem 3]. Corollaries 2 and 3 of *op. cit.* show that the coefficients of (symmetric) such polynomials also satisfy unimodality properties in the sense of majorization/stochastic domination [61]. In §4.3 we establish natural analogs of these results for truncations of (non-homogeneous) multi-affine stable polynomials with non-negative coefficients.

The well-known Laguerre-Turán inequalities for univariate polynomials and transcendental entire functions in the Laguerre-Pólya class [22] amount to saying that the sequence of Taylor coefficients of such a function is SLC (Definition 2.8). One can argue that both Newton's inequalities and the Laguerre-Turán inequalities are natural manifestations of negative dependence properties encoded in the *geometry* of the zero sets of these functions. Indeed, as we will see in §4, the geometric features of real stability (Definition 2.9) are a key tool in our study of negative dependence properties for strongly Rayleigh measures and allow us to establish many such properties for this class of measures, including:

- the strongest form of negative association (CNA+);
- closure under operations (i)–(vi) defined in §2.1;
- closure under symmetric homogenization;
- stochastic domination properties for truncations;
- the ULC property for rank sequences.

Rayleigh polynomials that are not stable do not seem to have a geometric description by means of their zero sets similar to that of stable polynomials. This may explain why several negative dependence properties fail for arbitrary (non-strongly) Rayleigh polynomials/measures (see §7).

We summarize below some of the closure properties of real stable polynomials (see [8, 9] for their proofs) that we need for later use.

Proposition 3.1. *Let $\mathcal{H}_n(\mathbb{R})$ be the set of all real stable polynomials in n variables.*
1. *$f \in \mathcal{H}_n(\mathbb{R})$ if and only if for any $\lambda \in \mathbb{R}_+^n$ and $\mu \in \mathbb{R}^n$ the univariate polynomial $t \mapsto f(\lambda t + \mu)$ has all real zeros.*
2. *If $f \in \mathcal{H}_n(\mathbb{R})$ has degree d_j in variable z_j, $j \in [n]$, then*
   - *$\partial_j f \in \mathcal{H}_n(\mathbb{R}) \cup \{0\}$ for $1 \leq j \leq n$;*
   - *$f(z_1, \ldots, z_{j-1}, \alpha z_j, z_{j+1}, \ldots, z_n) \in \mathcal{H}_n(\mathbb{R})$ for $j \in [n]$, $\alpha > 0$;*

- $f(z_1, \ldots, z_{j-1}, \beta, z_{j+1}, \ldots, z_n) \in \mathcal{H}_{n-1}(\mathbb{R}) \cup \{0\}$ *for* $j \in [n]$, $\beta \in \mathbb{R}$;
- $z_1^{d_1} \cdots z_n^{d_n} f(\lambda_1 z_1^{-1}, \ldots, \lambda_n z_n^{-1}) \in \mathcal{H}_n(\mathbb{R})$ *if* $\pm(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}_+^n$;
- $f(z_1, \ldots, z_{i-1}, z_j, z_{i+1}, \ldots, z_n) \in \mathcal{H}_{n-1}(\mathbb{R})$ *for* $i \neq j \in [n]$.

(3) *If* $\{f_j\}_{j=1}^\infty \subset \mathcal{H}_n(\mathbb{R})$ *and* $f \in \mathbb{R}[z_1, \ldots, z_n] \setminus \{0\}$ *is the limit, uniformly on compact subsets of* $\mathbb{C}^n$, *of the sequence* $\{f_j\}_{j=1}^\infty$, *then* $f \in \mathcal{H}_n(\mathbb{R})$.

Important examples of real stable polynomials are given by the following construction (see, e.g., [8, 9]).

Proposition 3.2. *Let* $A_1, \ldots, A_m$ *be (complex) positive semi-definite matrices and let* B *be a (complex) Hermitian matrix, all matrices being of the same size.*

(1) *The polynomial*

(6) $$z = (z_1, \ldots, z_n) \mapsto f(z) = \det(z_1 A_1 + \cdots + z_m A_m + B)$$

   is either identically zero or real stable.

(2) *If* B *is also positive semi-definite, then* f *has all non-negative coefficients.*

In particular, if $Z = \mathrm{diag}(z_1, \ldots, z_m)$ *and* A *is a positive semi-definite* $m \times m$ *matrix, then* $\det(A+Z)$ *is a multi-affine real stable polynomial with all non-negative coefficients, hence a (positive) constant multiple of a strongly Rayleigh polynomial (cf. Definition 2.10).*

Proof. By the second part of Proposition 3.1 and a density argument, it suffices to prove (1) in the case when all matrices A_j, $1 \leq j \leq m$, are positive definite. Set $z(t) = \lambda t + \mu$ with $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}_+^n$, $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$, and $t \in \mathbb{C}$. Note that $P := \sum_{j=1}^m \lambda_j A_j$ is positive definite and thus it is invertible and has a square root Q (recall that $\mathbb{R}_+ = (0, \infty)$). Then

$$f(z(t)) = \det(P) \det(tI + QHQ^*),$$

where $H := B + \sum_{j=1}^m \mu_j A_j$ is a Hermitian matrix. Therefore, $f(z(t))$ is a polynomial in t that is a constant multiple of the characteristic polynomial of a Hermitian matrix, and so it must have all real zeros. By the first part of Proposition 3.1, this proves (1), and then (2) follows by noticing that the coefficients of the polynomial in (6) are products of non-negative numbers, namely principal minors of the positive semi-definite matrices $A_1, \ldots, A_m, B$. $\qquad\square$

Further properties of (not necessarily multi-affine) real stable polynomials, their linear preservers and applications to entire function theory in one or several variables, matrix theory, combinatorics, convex optimization and statistical mechanics may be found in [7, 8, 9, 15, 19, 36, 51].

3.2. **Hadamard-Fischer inequalities and GKK-matrices.** Given an $n \times n$ matrix A over $\mathbb{C}$ and $S \subseteq [n]$, denote by $A[S]$ the principal submatrix of A with rows and columns indexed by S and let $A\langle S \rangle = \det(A[S'])$ be the principal minor of A with rows and columns indexed by $S' = [n] \setminus S$, where $A\langle [n] \rangle = \det(A[\emptyset]) := 1$. The matrix A is called a *P-matrix* if all its principal minors are positive. A much studied class of P-matrices that unifies several important types of matrices is the class of GKK-matrices; see [28, 40] and the references therein. A P-matrix A is a *GKK-matrix* (after Gantmacher-Krein and Kotelyansky) if it satisfies the *Hadamard-Fischer-Kotelyansky inequalities*, that is,

(7) $$A\langle S \rangle A\langle T \rangle \geq A\langle S \cup T \rangle A\langle S \cap T \rangle, \quad S, T \subseteq [n].$$

In other words, an $n \times n$ P-matrix A is GKK if and only if the probability measure μ_A on $2^{[n]}$ defined by

$$\mu_A(S) = A\langle S \rangle \det(A + I)^{-1} \tag{8}$$

satisfies the negative lattice condition (NLC). It is not hard to see that the generating polynomial of such a measure is

$$g_{\mu_A}(z) = \det(A+I)^{-1} \sum_{S \subseteq [n]} A\langle S \rangle z^S = \det(A+I)^{-1} \det(A+Z), \quad z^S := \prod_{i \in S} z_i,$$

where $Z = \mathrm{diag}(z_1, \ldots, z_n)$. Examples of GKK-matrices are:

   (I) Positive definite matrices.
  (II) Totally positive matrices, i.e., matrices for which *all* minors are positive.
 (III) Non-singular M-matrices; recall that an *M-matrix* (named after Minkowski) is a matrix with all non-negative principal minors and all non-positive off-diagonal entries.

An ambitious project would be to classify GKK-matrices according to the various negative dependence properties (introduced in §2) that these matrices satisfy. This would take us beyond the scope of this paper. In what follows we will only briefly discuss some of these aspects.

Definition 3.1. We say that a P-matrix A is a *Rayleigh matrix* if the associated probability measure μ_A defined in (8) is Rayleigh (in the sense of Definition 2.5).

By [72, Theorem 4.4], Rayleigh matrices are GKK. In fact, the following holds.

Theorem 3.3. *Let A be an $n \times n$ matrix over $\mathbb{C}$. The following are equivalent:*

  (1) *A is a Rayleigh matrix;*
  (2) *$A + X$ is a GKK-matrix for all $X = \mathrm{diag}(x_1, \ldots, x_n)$, where $x_i \geq 0$ for all $1 \leq i \leq n$.*

Proof. (1) $\Rightarrow$ (2): As we already noted, by [72, Theorem 4.4] Rayleigh matrices are GKK, and by definition the set of all Rayleigh matrices is obviously closed under adding positive diagonal matrices.

(2) $\Rightarrow$ (1): Let $f = \det(A + Z)$. Then

$$\frac{\partial f}{\partial z_i}(x) \frac{\partial f}{\partial z_j}(x) - \frac{\partial^2 f}{\partial z_i \partial z_j}(x) f(x)$$
$$= (A+X)\langle i \rangle \cdot (A+X)\langle j \rangle - (A+X)\langle i, j \rangle \cdot (A+X)\langle \emptyset \rangle.$$

Using the Hadamard-Fischer-Kotelyansky inequalities (7) for $A + X$ with $S = \{i\}$ and $T = \{j\}$ one gets the desired conclusion. $\qquad\square$

It follows immediately from Theorem 3.3 that non-singular M-matrices and positive definite matrices are Rayleigh. Now for any $s, t \in (0, 1)$ one can easily check that the matrix

$$B = \begin{bmatrix} 1 & 1/2 & s/4 \\ 1/2 & 1 & 1/2 \\ t/4 & 1/2 & 1 \end{bmatrix} \tag{9}$$

is totally positive. Given an $n \times n$ matrix A and two subsets $S, T \subseteq [n]$ of the same size, we let $A(S, T)$ denote the minor of A lying in rows indexed by S and columns

indexed by T. It was proved by Gantmacher-Krein [31] and Carlson [17] that a necessary and sufficient condition for a P-matrix to be GKK is that

$$A(S,T)A(T,S) \geq 0 \text{ for any } S, T \subseteq [n] \text{ with } |S| = |T| = |S \cup T| - 1.$$

Since

$$(B+I)(12, 23) = (1 - 2s)/4 \text{ and } (B+I)(23, 12) = (1 - 2t)/4$$

we may choose $s, t \in (0, 1)$ so that $(B+I)(12, 23)(B+I)(23, 12) < 0$. We conclude that for such values of s, t the matrix $B + I$ fails to be GKK.

It is easy to see that the measures associated with positive definite and totally positive matrices have ULC rank sequences. This follows from Newton's inequalities and the fact that Hermitian matrices and totally positive matrices have real-rooted characteristic polynomials (in the case of totally positive matrices this goes back to Gantmacher and Krein [31]). For M-matrices it is not as obvious but nonetheless true. In [40] Holtz and Schneider conjectured that M-matrices satisfy the conclusion of Conjecture 2.4. This was subsequently confirmed by Holtz in [39] using a certain monotonicity property for so-called symmetrized Fischer products of M-matrices due to James-Johnson-Pierce [44] (see [8] for further results on Fischer products).

We end this section with some open questions.

Question 3.1. Is μ_A an ULC measure whenever A is a Rayleigh matrix?

Question 3.2. For which GKK-matrices A is μ_A negatively associated?

Again, negative association fails for measures associated with totally positive matrices. Indeed, let B be the totally positive 3×3 matrix defined in (9) and fix $s, t \in (0, 1)$ such that $A := B + I$ is not GKK (see the above discussion). It is known [17, 31] that an $n \times n$ P-matrix C is GKK if and only if

$$(10) \qquad C\langle S \cup \{i\} \rangle \cdot C\langle S \cup \{j\} \rangle \geq C\langle S \cup \{i, j\} \rangle \cdot C\langle S \rangle$$

for all $S \subseteq [n]$ and $i, j \in [n] \setminus S$ with $i \neq j$. Hence, since A fails to be GKK it follows that (10) fails for A and some S, i, j. Unraveling the definitions, this means that for some principal submatrix B' of B the corresponding measure $\mu_{B'}$ will fail to have pairwise negatively correlated variables (cf. Definition 2.3). In particular, since NA $\Rightarrow$ p-NC, the measure $\mu_{B'}$ cannot be negatively associated.

Question 3.3. Describe all $n \times n$ Rayleigh matrices.

Question 3.4. Characterize the class of all real stable $n \times n$ matrices, that is, $n \times n$ matrices A such that $\det(A + Z)$ is real stable.

3.3. **Determinantal probability measures.** In [57] Lyons defined a *determinantal probability measure* on $2^{[n]}$ to be a measure $\mu \in \mathfrak{P}_n$ such that there is an $n \times n$ matrix A so that for any subset S of $[n]$ one has

$$(11) \qquad \mu(\{T : S \subseteq T\}) = \det(A[S]),$$

where as before $A[S]$ denotes the principal submatrix of A whose rows and columns are indexed by S. As discussed in e.g. [5, 57], such measures have important applications ranging from fermionic processes/continuous scaling limits of discrete point processes [23, 59] to the distribution of eigenvalues of random matrices and zeros of the Riemann zeta function [21], transfer current matrices [16], non-intersecting random walks [47], Poissonized versions of the Plancherel measure on partitions [13] and Young diagrams [64], discrete orthogonal polynomial ensembles [46], etc.

Recall that a square matrix A is said to be a *contraction* if $||A|| \leq 1$, where $|| \cdot ||$ denotes the supremum (operator) norm. Clearly, a positive semi-definite matrix is a contraction if and only if all its eigenvalues are in the interval $[0, 1]$. Such matrices were called *positive contractions* in [57].

Theorem 3.4. *Suppose that μ is a determinantal measure on $2^{[n]}$ whose corresponding matrix is a positive contraction. Then μ is CNA+.*

This is one of the main results of [57] (cf. [57, Theorem 8.1]), where exterior (Grassmann) algebra methods were used to investigate determinantal probability measures and, in particular, to prove the above theorem. As noted in *loc. cit.*, Theorem 3.4 provides a powerful tool for studying determinantal probability measures, comparable to the FKG theorem in the theory of positive association, and is crucial to most of the results in e.g. [59].

Let us show that Theorem 3.4 actually follows from Theorem 4.9 of §4.2 and the fact that determinantal measures induced by positive contractions are strongly Rayleigh:

Proposition 3.5. *Suppose that μ is a determinantal measure on $2^{[n]}$ whose corresponding matrix is a positive contraction. Then μ is strongly Rayleigh.*

Proof. Since positive definite matrices are dense in the set of all positive semi-definite matrices, by Proposition 3.1 (3) it is enough to prove the statement for invertible positive contractions. Elementary computations show that

$$g_\mu(z_1, \ldots, z_n) = \det(I - A + AZ) = \det(A) \cdot \det(A^{-1} - I + Z),$$

where $Z = \mathrm{diag}(z_1, \ldots, z_n)$. Since A is an invertible positive contraction, $A^{-1} - I$ is positive semi-definite and $\det(A) > 0$, so by (6) the polynomial g_μ is real stable. $\square$

Another main result of [57] is that if A and B are commuting positive contractions and $A \leq B$, then the measure corresponding to B stochastically dominates the one corresponding to A (Definition 2.14). In §4.3 we extend this result using the theory of stability and strongly Rayleigh measures and we prove that the condition $[A, B] = 0$ (where $[\cdot, \cdot]$ denotes the Lie bracket) can be dropped from the hypothesis.

3.4. **Graphs, Laplacians, spanning trees and the random cluster model.** Let $G = (V, E)$ be a graph with vertex set $V = [n]$ and edge set E. Associate to each edge $e \in E$ a variable w_e. If e connects i and j let A_e be the $n \times n$ positive semi-definite matrix with the ii-entry and the jj-entry equal to 1, the ij-entry and ji-entry equal to -1, and all other entries equal to 0. The *Laplacian*, $L(G)$, of G is defined by

$$L(G) = \sum_{e \in E} w_e A_e.$$

Let

$$f_G(z, w) = \det(L(G) + Z).$$

Thus, by (6), f_G is a multi-affine real stable polynomial with non-negative coefficients. The Principal Minors Matrix-Tree Theorem (see, e.g., [18]) says that

$$f_G(z, w) = \sum_F z^{\mathrm{roots}(F)} w^{\mathrm{edges}(F)},$$

where the sum is over all rooted spanning forests F in G, $\mathrm{roots}(F) \subseteq V$ is the set of roots of F and $\mathrm{edges}(F) \subseteq E$ is the set of edges used in F. Alternatively, we may write f_G as

$$f_G(z, w) = \sum_F w^F \prod_{C \in \mathcal{C}(F)} (\sum_{j \in C} z_j),$$

where $\mathcal{C}(F)$ is the partition of V whose parts are the vertices in the maximal trees in F and the sum is over all forests in G. Since the class of stable polynomials is closed under differentiation and specialization of variables at real values (see, e.g., [9]) we have that the *spanning tree polynomial* $T_G(w) = \sum_T w^T$, where the sum is over all spanning trees, is real stable (which is widely known). The corresponding probability measure is usually called the uniform random spanning tree measure and is consequently strongly Rayleigh. Now if $\mathbf{1}$ denotes the vector of all ones as in §2, then

$$f_G(\mathbf{1}, w) = \sum_F \omega_G(F) w^F,$$

where the sum is over all spanning forests (independent sets) in G and $\omega_G(F)$ is the product of the sizes of the connected components of F. In [72, Conjecture 5.10] Wagner conjectured that $f_G(\mathbf{1}, w)$ is a Rayleigh polynomial. By the above discussion we have shown that more is actually true, namely:

Theorem 3.6. *For any graph G the polynomial $f_G(\mathbf{1}, w)$ is real stable.*

Let $G = (V, E)$ be a graph and let $q > 0$ and $0 \le p \le 1$ be parameters. The *random cluster (RC) model* (see [33]) is the measure μ on 2^E defined by

$$\mu(F) = \frac{1}{Z_{\mathrm{RC}}} q^{k(F)} p^{|F|} (1 - p)^{|E \setminus F|}, \quad F \subseteq E,$$

where $k(F)$ is the number of connected components in the graph (V, F) and Z_{RC} is the appropriate normalizing factor. The uniform spanning tree measure can be obtained from this model by letting $p \to 0$ and $q/p \to 0$; see [33, Section 1.5]. If $q \ge 1$, the RC model satisfies PLC and is therefore positively associated. If $q \le 1$, NLC is satisfied, but other negative correlation properties are largely a matter of conjecture. We are interested in properties closed under external fields, so we assume that $p = 1/2$. The generating polynomial of μ is then a constant multiple of the multivariate Tutte polynomial (see [70]):

$$Z_G(z, q) = \sum_{F \subseteq E} q^{k(F)} z^F, \quad z = (z_e)_{e \in E}.$$

If $G = C_n, n \ge 3$, is a cycle and $E = [n]$, then

$$Z_G(z, q) = \prod_{j=1}^n (q + z_j) + (q - 1) z_1 \cdots z_n$$

and

$$\frac{\partial Z_G}{\partial z_1} \frac{\partial Z_G}{\partial z_2} - Z_G \frac{\partial^2 Z_G}{\partial z_1 \partial z_2} = q^2 (1 - q) \prod_{j=3}^n z_j (q + z_j),$$

so in this case the RC measure is Rayleigh, but not strongly Rayleigh.

If G is a tree, then

$$Z_G(z, q) = q \prod_{e \in E} (q + z_e),$$

so the RC model is strongly Rayleigh. We conclude that on a general graph G, the RC model is strongly Rayleigh if and only if G is acyclic.

With the help of $Mathematica^{\copyright}$, the third author has verified that the RC model is Rayleigh for all simple graphs with at most five vertices.

3.5. **Interacting particle systems and exclusion evolutions.** The exclusion process is one of the main models considered in the area of Probability Theory known as Interacting Particle Systems. The idea is that particles move in continuous time on a countable set S, in such a way that there is always at most one particle per site. The motion of each particle would be a continuous time Markov chain on S, except that transitions to occupied sites are forbidden. This process has been used to model many situations, including biopolymers and traffic flow. We refer to [52, Chap. VIII] and [54, Part 3] for detailed treatments of this process.

The (finite) *symmetric* exclusion process generates a continuous time evolution on the set $\mathfrak{P}_n$. The limiting measure under the symmetric exclusion evolution with initial measure μ is the symmetrization μ_s of μ as defined in §2.1. A problem that has attracted some attention is the following (see, e.g., [52, 55]).

Problem 3.5. Find a natural negative dependence property that is preserved by symmetric exclusion evolutions.

Since the limiting distribution of the evolution as $t \to \infty$ is the symmetrization of the initial distribution defined in (2), any such property would have to be preserved by this symmetrization procedure. In order for a solution to Problem 3.5 to be useful in settling Conjecture 2.2, the negative dependence property should be satisfied by non-random/deterministic distributions and should imply NA. Problem 3.5 is also motivated by a theorem of Harris [52, Theorem 2.14] that provides a general criterion for the preservation of positive dependence for Markov processes. Theorems 3.1 and 3.3 of [55] give negative dependence properties that are preserved by the symmetric exclusion evolution. However, one can make an argument that neither of these is ideal – the first is too weak, while the second is too strong. In [55, Theorem 3.2] it was proved that NA is not preserved by the symmetric exclusion process. The examples that we construct in §7 further show that in fact none of the properties NLC, h-NLC, Rayleigh/h-NLC+, CNA, CNA+ defined in §2.1 is preserved by such evolutions. This is a consequence of the following criterion established by Pemantle in [65, Theorem 2.7] (see also [72, Proposition 3.6]) for the ultra log-concavity of rank sequences of symmetric polynomials in $\mathcal{P}_n$:

Theorem 3.7 ([65])**.** *For* $0 \le k \le n$ *let* $e_k(z_1, \ldots, z_n)$ *be the k-th elementary symmetric function on* $\{z_1, \ldots, z_n\}$ *and consider the polynomial*

$$f(z_1, \ldots, z_n) = \sum_{k=0}^{n} a_k e_k(z_1, \ldots, z_n),$$

where $a_k \ge 0$, $0 \le k \le n$. *The following are equivalent:*
  (1) *f satisfies the conclusion of Conjecture* 2.4*; that is, it has an ULC rank sequence (cf. Definition* 2.8*);*
  (2) *f has either (and then all) of the following five properties: NLC, h-NLC, Rayleigh/h-NLC+, CNA, and CNA+.*

In §5 we show that the strongly Rayleigh property is preserved by (finite) symmetric exclusion evolutions and thus provide an answer to Problem 3.5. We then

prove that the distribution at time $t > 0$ for the (not necessarily finite) symmetric exclusion process is CNA+ whenever the initial distribution is a product measure, which confirms Conjecture 2.2.

Symmetric strongly Rayleigh measures are described in the following theorem. Recall that a (possibly infinite) matrix A is said to be totally non-negative of order p, or TP_p for short, if for any $k \leq p$ all its $k \times k$ minors are non-negative [49]. If A has all non-negative minors, then A is called totally non-negative, or TP for short.

Theorem 3.8 (Exchangeable Strongly Rayleigh Case)**.** *Suppose $\mu \in \mathcal{P}_n$ has a symmetric generating polynomial and rank sequence $\{r_k\}_{k=0}^n$. The following are equivalent:*

   (a) *The generating polynomial of μ is stable, i.e., μ is strongly Rayleigh.*
   (b) *All zeros of the univariate polynomial $\sum_{k=0}^n r_k z^k$ are real.*
   (c) *The infinite Toeplitz matrix $(r_{i-j})_{i,j=0}^\infty$ is TP, where $r_k = 0$ if $k \notin \{0\} \cup [n]$.*

Proof. The equivalence (b) $\Leftrightarrow$ (c) is a classical result due to Aissen, Schoenberg and Whitney [1], while the equivalence (a) $\Leftrightarrow$ (b) is a consequence of the Grace-Walsh-Szegö coincidence theorem (see Theorem 4.6 below). $\qquad\square$

4. NEGATIVE DEPENDENCE THEORY FOR STRONGLY RAYLEIGH MEASURES

The importance of strongly Rayleigh measures – that is, measures with stable generating polynomials (cf. Definition 2.10) – in negative dependence stems from the theorem below proved by one of the authors in [15] (cf. §3.1).

Theorem 4.1. *Let $g \in \mathbb{R}[z_1, \ldots, z_n]$ be a multi-affine polynomial. Then g is stable if and only if*

$$\frac{\partial g}{\partial z_i}(x)\frac{\partial g}{\partial z_j}(x) \geq \frac{\partial^2 g}{\partial z_i \partial z_j}(x)g(x)$$

for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $i, j \in [n]$.

Hence, perhaps somewhat surprisingly, if the Rayleigh condition in Definition 2.5 is enforced to hold for all real vectors, then this stronger correlation condition is equivalent to real stability. Clearly, a strongly Rayleigh measure is automatically Rayleigh (cf. Remark 2.5). As we will see in this section, the geometric properties encoded in the concept of real stability allow us to show that strongly Rayleigh measures enjoy all virtues of negative dependence, including the strongest form of negative association CNA+.

4.1. **Symmetric homogenization via Gårding hyperbolic polynomials.** The following theorem shows that the symmetric homogenization (cf. §2.3) of any strongly Rayleigh measure is again strongly Rayleigh, so that strongly Rayleigh measures belong to the class PHR (Definition 2.6).

Theorem 4.2. *If $\mu \in \mathfrak{P}_n$ is strongly Rayleigh, then so is its symmetric homogenization $\mu_{sh} \in \mathfrak{P}_{2n}$.*

The proof of Theorem 4.2 makes use of the theory of hyperbolic polynomials that has its origins in partial differential equations and was developed by Petrovsky, Gårding, Hörmander, Atiyah and Bott [3, 37, 43]. Recently, hyperbolic polynomials have proved to be quite useful in other areas of mathematics such as convex optimization [4, 36], complex analysis [7], matrix theory and combinatorics [34, 35].

Let $e \in \mathbb{R}^n$. A homogeneous polynomial $p \in \mathbb{R}[z_1, \ldots, z_n]$ is said to be *(Gårding) hyperbolic with respect to e* if $p(e) \neq 0$ and for all $x \in \mathbb{R}^n$ the univariate polynomial $t \mapsto p(x + te)$ has only real zeros. Recall that the *homogenization* of a polynomial $f(z_1, \ldots, z_n) = \sum_{\alpha \in \mathbb{N}^n} a(\alpha) z^\alpha \in \mathbb{C}[z_1, \ldots, z_n]$ of degree d is given by

$$f_H(z_1, \ldots, z_{n+1}) = z_{n+1}^d f(z_1/z_{n+1}, \ldots, z_n/z_{n+1}).$$

Recall the characterization of real stable polynomials given in Proposition 3.1 (1). The relationship between real stable polynomials and hyperbolic polynomials is made explicit by the following result; see [9, Proposition 1].

Proposition 4.3. *Let $f \in \mathbb{R}[z_1, \ldots, z_n]$. Then f is real stable if and only if f_H is hyperbolic with respect to all vectors $e \in \mathbb{R}^{n+1}$ of the form $e = (e_1, \ldots, e_n, 0)$, where $e_i > 0$ for $1 \leq i \leq n$.*

Let p be hyperbolic with respect to $e \in \mathbb{R}^n$. The cone

$$C_e(p) = \{x \in \mathbb{R}^n : p(x + te) \neq 0, t \geq 0\}$$

is called the *hyperbolicity cone* of p. The following fundamental properties of the hyperbolicity cone are due to Gårding [37]; see also [3, 36, 43].

Proposition 4.4. *Let $p \in \mathbb{R}[z_1, \ldots, z_n]$ be hyperbolic with respect to $e \in \mathbb{R}^n$. Then*

  (a) $C_e(p)$ *is convex;*
  (b) $C_e(p)$ *is equal to the connected component of the set $\{x \in \mathbb{R}^n : p(x) \neq 0\}$ that contains e;*
  (c) $(s, t) \mapsto p(x + su + te)$ *is real stable for any $x \in \mathbb{R}^n$ and $u \in C_e(p)$;*
  (d) *p is hyperbolic with respect to any $u \in C_e(p)$ and $C_u(p) = C_e(p)$.*

Below are examples of hyperbolic polynomials and their hyperbolicity cones:

  (I) The polynomial $p(z_1, \ldots, z_n) = z_1 z_2 \cdots z_n$ is hyperbolic with respect to the direction $e = (1, \ldots, 1) \in \mathbb{R}^n$, and the hyperbolicity cone is the positive orthant $C_e(p) = \mathbb{R}^n_+$.
 (II) The (principal) symbol of the hyperbolic wave equation, i.e., the polynomial

$$p(z_1, \ldots, z_n) = z_1^2 - \sum_{j=2}^n z_j^2,$$

   is hyperbolic with respect to the direction $e = (1, 0, \ldots, 0) \in \mathbb{R}^n$, and the hyperbolicity cone is the Lorentz cone

$$C_e(p) = \left\{ z \in \mathbb{R}^n : \sqrt{z_2^2 + \cdots + z_n^2} \leq z_1 \right\}.$$

(III) Let z_{ij}, $1 \leq i \leq j \leq n$, be $n(n+1)/2$ different variables and let $X = (x_{ij})_{i,j=1}^n$ be the matrix with entries

$$x_{ij} = \begin{cases} z_{ij} & \text{if } i \leq j, \\ z_{ji} & \text{if } i > j. \end{cases}$$

   The polynomial (in $n(n+1)/2$ variables) $p(z_{11}, \ldots, z_{nn}) = \det(X)$ is hyperbolic with respect to the $n \times n$ identity matrix I, and the hyperbolicity cone is the cone of positive definite $n \times n$ matrices.

(IV) Let $A_1, \ldots, A_m$ be symmetric $n \times n$ matrices and let $e = (e_1, \ldots, e_m) \in \mathbb{R}^m$ be such that $e_1 A_1 + \cdots + e_m A_m$ is positive definite. Then the polynomial

$$p(z_1, \ldots, z_m) = \det\left(\sum_{j=1}^{m} z_j A_j \right)$$

is hyperbolic with respect to e, and the hyperbolicity cone is given by

$$C_e(p) = \left\{ x \in \mathbb{R}^m : \sum_{j=1}^{n} x_j A_j \text{ is positive definite} \right\}.$$

Theorem 4.5. *Suppose that all the coefficients of $f \in \mathbb{R}[z_1, \ldots, z_n]$ are non-negative. The following are equivalent:*

(1) *f_H is stable;*
(2) *f is stable;*
(3) *f_H is hyperbolic with respect to some vector e with $e_i \geq 0$; $1 \leq i \leq n+1$;*
(4) *f_H is hyperbolic with respect to any vector e with $e_i > 0$, $1 \leq i \leq n+1$.*

Proof. By the fact that the stability property is closed under setting variables equal to real numbers (Proposition 3.1) and Proposition 4.3, we have that $(1) \Rightarrow (2) \Rightarrow (3)$. Now if (3) holds, then since f has all non-negative coefficients it follows from Proposition 4.4 (b) that $C_e(f_H)$ contains the cone $\mathbb{R}_+^{n+1}$, which by Proposition 4.4 (d) proves (4). Finally, $(4) \Rightarrow (1)$ by Proposition 3.1. $\qquad\square$

Let $f \in \mathbb{C}[z_1, \ldots, z_n]$ be a polynomial of degree d_i in z_i, $1 \leq i \leq n$. The *polarization* $\pi(f)$ of f is the unique polynomial in $\sum_{i=1}^{n} d_i$ variables z_{ij}, $1 \leq i \leq n$, $1 \leq j \leq d_i$, satisfying:

(i) $\pi(f)$ is multi-affine;
(ii) $\pi(f)$ is symmetric in the variables $z_{i1}, \ldots z_{id_i}$ for $1 \leq i \leq n$;
(iii) if we let $z_{ij} = z_i$ for all i, j in $\pi(f)$, we recover f.

Recall that a *circular region* in $\mathbb{C}$ is either an open or closed affine half-plane or the open or closed interior or exterior of a circle. The famous Grace-Walsh-Szegö coincidence theorem [32, 66, 74] is stated next. For a proof of the following version we refer to [19, Theorem 2.12].

Theorem 4.6 (Grace-Walsh-Szegö). *Let $f \in \mathbb{C}[z_1, \ldots, z_n]$ be symmetric and multi-affine and let C be a circular region containing the points $\zeta_1, \ldots, \zeta_n$. Suppose that either f has total degree n or C is convex (or both). Then there exists at least one point $\zeta \in C$ such that $f(\zeta_1, \ldots, \zeta_n) = f(\zeta, \ldots, \zeta)$.*

From the Grace-Walsh-Szegö coincidence theorem we deduce:

Corollary 4.7. *Let $f \in \mathbb{C}[z_1, \ldots, z_n]$. Then f is stable if and only if $\pi(f)$ is stable.*

Proof. If $\pi(f)$ is stable, then so is f since the latter is recovered from $\pi(f)$ by setting variables equal as in (iii) above. Now suppose that $\pi(f)$ is not stable. Then there are numbers $\zeta_{ij} \in \{z \in \mathbb{C} : \mathfrak{Im}(z) > 0\}$, $1 \leq i \leq n$, $1 \leq j \leq d_i$, such that $\pi(f)(\zeta) = 0$, where $\zeta = (\zeta_{11}, \ldots, \zeta_{1d_1}, \ldots, \zeta_{n1}, \ldots, \zeta_{nd_n})$. By successively applying the Grace-Walsh-Szegö coincidence theorem we see that there is a vector $\zeta' = (\zeta'_{11}, \ldots, \zeta'_{1d_1}, \ldots, \zeta'_{n1}, \ldots, \zeta'_{nd_n})$ with coordinates in $\{z \in \mathbb{C} : \mathfrak{Im}(z) > 0\}$ for which $\zeta'_{ij} = \zeta'_{k\ell}$ whenever $i = k$ and $\pi(f)(\zeta') = \pi(f)(\zeta) = 0$. Hence $f(\zeta'_{11}, \zeta'_{21}, \ldots, \zeta'_{n1}) = \pi(f)(\zeta') = 0$, so f is not stable either. $\qquad\square$

We now have all the tools to prove Theorem 4.2.

Proof of Theorem 4.2. Let g be the generating polynomial of $\mu \in \mathfrak{P}_n$ and suppose that g is stable of degree d. Then so is the homogenization g_H of g by Proposition 4.3. Now straightforward computations show that the generating polynomial of the symmetric homogenization $\mu_{sh} \in \mathfrak{P}_{2n}$ is given by $\pi(z_{n+1}^{n-d} g_H)$, so the theorem follows from Corollary 4.7.   □

4.2. **Stability implies negative association.** The main argument in the proof of the following theorem goes back to Feder-Mihail [29], although the proof of the full theorem is scattered over the literature [24, 57, 58, 72].

Theorem 4.8. *Let $\mathcal{S}$ be a class of probability measures satisfying:*
1. *Each $\mu \in \mathcal{S}$ is a measure on 2^E, where E is a finite subset of $\{1, 2, \ldots\}$ depending on μ.*
2. *$\mathcal{S}$ is closed under conditioning.*
3. *For each $\mu \in \mathcal{S}$ the variables X_e, $e \in E$, are pairwise negatively correlated.*
4. *Each $\mu \in \mathcal{S}$ has a homogeneous generating polynomial.*

Then all measures in $\mathcal{S}$ are CNA (conditionally negatively associated).

The homogeneity condition in (4) above is essential and ensures the existence of a variable of so-called "positive influence" [24] (various notions of "influence" have been studied by Bourgain *et al.* in [14]).

A sketch of the proof of Theorem 4.8 is as follows. Recall from §2.1 the definition of an increasing event. Since any increasing function depending on a set E_0 can be written as a positive linear combination of characteristic functions, $\chi_{\mathcal{A}}$, of increasing events $\mathcal{A}$ depending on E_0, it is enough to prove the theorem for such characteristic functions. The proof of negative association in the case when $F = \chi_{\mathcal{A}}$ and $G = X_e$ in Definition 2.7 are such that $\mathcal{A}$ does not depend on e is recovered from Feder-Mihail's proof [29, 58]. The general case – when $F = \chi_{\mathcal{A}_1}$ and $G = \chi_{\mathcal{A}_2}$ in Definition 2.7 are such that $\mathcal{A}_1$ and $\mathcal{A}_2$ depend on disjoint sets of variables – then follows from the proof of [57, Theorem 6.5].

Using Theorem 4.8 we can establish the following result.

Theorem 4.9. *If $\mu \in \mathfrak{P}_n$ is strongly Rayleigh, then it is CNA+.*

Proof. Let $\mathcal{S}$ be the class of all probability measures μ such that
- each $\mu \in \mathcal{S}$ is a measure on 2^E, where E is a finite subset of $\{1, 2, \ldots\}$ depending on μ, and
- μ has a stable homogeneous generating polynomial.

Since stable homogeneous polynomials are pairwise negatively correlated and are also closed under conditioning (cf. §2.1), the class $\mathcal{S}$ satisfies all the hypotheses required in Theorem 4.8. Therefore, all the measures in $\mathcal{S}$ are negatively associated. Further, let $\widehat{\mathcal{S}}$ be the class of all probability measures μ such that
- each $\mu \in \widehat{\mathcal{S}}$ is a measure on 2^E, where E is a finite subset of $\{1, 2, \ldots\}$ depending on μ, and
- μ has a stable generating polynomial.

Now by Theorem 4.2 every measure in $\widehat{\mathcal{S}}$ is the projection of a measure in $\mathcal{S}$. Since negative association is closed under projections and the strongly Rayleigh property is closed under conditioning and external fields, the theorem follows.   □

Note that since the Rayleigh/h-NLC+ property is also closed under conditioning, projections, and external fields, Theorem 4.8 and slight modifications of the above arguments actually yield the following stronger version of Theorem 4.9.

Theorem 4.10. *If $\mu \in \mathfrak{P}_n$ is PHR, then it is CNA+.*

Since homogeneous Rayleigh polynomials need not be stable, strongly Rayleigh $\Rightarrow$ PHR strictly. As we will see in §7, one also has PHR $\Rightarrow$ CNA+ strictly.

4.3. **Stochastic domination and truncations.** Recall the notion of stochastic domination from Definition 2.14 in §2.5. It turns out that for strongly Rayleigh measures, stochastic domination is intimately connected with central notions in the theory of stable and real-rooted polynomials, namely the notions of interlacing zeros and proper position [7, 8, 9, 15, 50, 66]. In particular, this connection allows us to extend one of Lyons' main results [57] and to answer Pemantle's questions on stochastic domination for truncations of "negatively dependent" measures [65] (cf. §2.5).

4.3.1. *Partial orders for strongly Rayleigh measures and polynomials.* Let $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$ and $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_m$ be the zeros of two real-rooted polynomials $p, q \in \mathbb{R}[z]$. These zeros are *interlacing* if they can be ordered so that either $\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \beta_2 \leq \cdots$ or $\beta_1 \leq \alpha_1 \leq \beta_2 \leq \alpha_2 \leq \cdots$. If the zeros of p, q interlace, then the *Wronskian* $W[p, q] := p'q - pq'$ is either non-negative or non-positive on the whole real axis $\mathbb{R}$; see, e.g., [50, 66]. We say that p and q are in *proper position*, denoted $p \ll q$, if the zeros of p and q interlace and $W[p, q] \leq 0$. For technical reasons, we also say that the zeros of the polynomial 0 interlace the zeros of any (non-zero) real-rooted polynomial and write $0 \ll p$ and $p \ll 0$.

The following theorem proved in [9] provides a notion of proper position for multivariate polynomials.

Theorem 4.11. *Let $f, g \in \mathbb{R}[z_1, \ldots, z_n]$. The following are equivalent:*
   (1) *The polynomial $g + if \in \mathbb{C}[z_1, \ldots, z_n]$ is stable.*
   (2) *The polynomial $g + z_{n+1}f \in \mathbb{R}[z_1, \ldots, z_n, z_{n+1}]$ is real stable.*
   (3) *For all $\lambda \in \mathbb{R}_+^n$ and $\mu \in \mathbb{R}^n$ we have*

$$f(\lambda t + \mu) \ll g(\lambda t + \mu),$$

   where $f(\lambda t + \mu) = f(\lambda_1 t + \mu_1, \ldots, \lambda_n t + \mu_n)$, $t \in \mathbb{C}$.

We say that the polynomials $f, g \in \mathbb{R}[z_1, \ldots, z_n]$ are in *proper position*, written $f \ll g$, if any of the equivalent conditions in Theorem 4.11 are satisfied.

Remark 4.1. The typical proper position is between a real stable polynomial and any of its partial derivatives: if $f \in \mathbb{R}[z_1, \ldots, z_n]$ is real stable, then $\partial_j f \ll f$ for all $j \in [n]$. If $f \ll g$ and α is a real number, then either $f(\alpha, z_2, \ldots, z_n) \ll g(\alpha, z_2, \ldots, z_n)$ or $f(\alpha, z_2, \ldots, z_n) = g(\alpha, z_2, \ldots, z_n) \equiv 0$; see [9]. Note also that by Theorem 4.11 (3) one has $f \ll g \Leftrightarrow \alpha f \ll \beta g$ for any $\alpha, \beta \in \mathbb{R}_+$.

We define a partial order $\trianglelefteq$ on the set of strongly Rayleigh probability measures on $2^{[n]}$ as follows.

Definition 4.1. Let $\mathcal{SR}_n$ be the set of strongly Rayleigh probability measures on $2^{[n]}$. First, define a (preliminary) partial order $\trianglelefteq'$ on $\mathcal{SR}_n$ by setting $\mu \trianglelefteq' \nu$ if there exists a sequence of strongly Rayleigh probability measures $\mu = \mu_0, \mu_1, \ldots, \mu_\ell = \nu$

such that $g_{\mu_0} \ll g_{\mu_1} \ll \cdots \ll g_{\mu_\ell}$. Then define the partial order $\trianglelefteq$ on $\mathcal{SR}_n$ as the closure of $\trianglelefteq'$; i.e., set $\mu \trianglelefteq \nu$ if there are two sequences $\{\mu_j\}_{j=0}^\infty, \{\nu_j\}_{j=0}^\infty \subset \mathcal{SR}_n$ such that

$$\lim_{j\to\infty} \mu_j = \mu, \ \lim_{j\to\infty} \nu_j = \nu, \ \text{and } \mu_j \trianglelefteq' \nu_j \text{ for all } j \in \mathbb{N}$$

(as the measures are defined on $2^{[n]}$ it does not matter what notion of limit we use). We will also consider $\trianglelefteq$ to be a partial order on the set of generating polynomials of strongly Rayleigh measures and write $g_\mu \trianglelefteq g_\nu$ whenever $\mu \trianglelefteq \nu$.

Remark 4.2. It seems likely that the partial orders $\trianglelefteq'$ and $\trianglelefteq$ on $\mathcal{SR}_n$ defined above are actually the same. However, since it is irrelevant for the present purposes, we do not further pursue this issue here.

As we noted in §2.5, the stochastic domination relation $\preccurlyeq$ defines a partial order on $\mathfrak{P}_n$ (cf. [61]). The anti-symmetry of $\trianglelefteq$ follows from our next result.

Proposition 4.12. *Let μ and ν be strongly Rayleigh probability measures on $2^{[n]}$. Then $\mu \trianglelefteq \nu \implies \mu \preccurlyeq \nu$.*

Proof. Clearly, we may assume that $g_\mu \ll g_\nu$, so that by Theorem 4.11 the polynomial in $n+1$ variables $G = (g_\nu + z_{n+1} g_\mu)/2$ is stable. Let $\gamma \in \mathfrak{P}_{n+1}$ be the measure with generating polynomial G. By Theorem 4.9, γ is NA. Let $\mathcal{A}$ be an increasing event not depending on the last coordinate and set $\mathcal{B} = \{S \in 2^{[n+1]} : n+1 \in S\}$. Then

$$\frac{1}{2}\mu(\mathcal{A}) = \gamma(\mathcal{A} \cap \mathcal{B}) \leq \gamma(\mathcal{A})\gamma(\mathcal{B}) = \left(\frac{1}{2}\mu(\mathcal{A}) + \frac{1}{2}\nu(\mathcal{A})\right)\frac{1}{2},$$

which gives $\mu(\mathcal{A}) \leq \nu(\mathcal{A})$. $\qquad\square$

Recall that the standard partial order on the set of all Hermitian $n \times n$ matrices – often referred to as the *Loewner order* [61] – is induced by the cone of all positive semi-definite $n \times n$ matrices: $A \leq B$ (respectively, $A < B$) means that $B - A$ is positive semi-definite (respectively, positive definite). Using our discussion so far and the arguments further below, we can prove the following extension of one of the main results in [57]:

Theorem 4.13. *Let A and B be positive contractions such that $A \leq B$ and let μ and ν be their corresponding determinantal measures. Then $\mu \trianglelefteq \nu$ and thus $\mu \preccurlyeq \nu$.*

In [57, Theorem 8.1] Lyons proved the last statement ($\mu \preccurlyeq \nu$) in Theorem 4.13 under the additional hypothesis $[A, B] = 0$.

We will need the following elementary lemma.

Lemma 4.14. *If A, B are positive semi-definite $n \times n$ matrices such that $A \leq B < I$, then $A(I - A)^{-1} \leq B(I - B)^{-1}$.*

Proof. By a standard density argument it is enough to prove the lemma for positive definite matrices. Note first that if C is a positive definite $n \times n$ matrix such that $C \leq I$, then $C^{-1} \geq I$. Moreover, if D is a positive definite $n \times n$ matrix such that $C \leq D$, then $M^*CM \leq M^*DM$ for any $n \times n$ matrix M since

$$\langle M^*CMu, u \rangle = \langle CMu, Mu \rangle \leq \langle DMu, Mu \rangle = \langle M^*DMu, u \rangle$$

for any $u \in \mathbb{C}^n$. The latter property with $M = D^{-1/2}$ yields $D^{-1/2}CD^{-1/2} \leq I$. Then by taking inverses and using the first property we get $D^{1/2}C^{-1}D^{1/2} \geq I$, and finally by again using the second property with $M = D^{-1/2}$ we obtain $C^{-1} \geq D^{-1}$.

Therefore, if A, B are positive definite matrices as in the lemma, it follows that $I - A^{-1} \leq I - B^{-1}$, hence

$$A(I - A)^{-1} = -\left(I - A^{-1}\right)^{-1} \leq -\left(I - B^{-1}\right)^{-1} = B(I - B)^{-1},$$

which is the desired conclusion. $\qquad\square$

Proposition 4.15. *Let A and B be positive semi-definite $n \times n$ matrices such that $A \leq B$ and set $Z = \mathrm{diag}(z_1, \ldots, z_n)$. Then*

$$\frac{\det(I + AZ)}{\det(I + A)} \trianglelefteq \frac{\det(I + BZ)}{\det(I + B)} \quad and \quad \frac{\det(B + Z)}{\det(B + I)} \trianglelefteq \frac{\det(A + Z)}{\det(A + I)}.$$

Moreover, if A and B are positive contractions, then

$$\det(I - A + AZ) \trianglelefteq \det(I - B + BZ).$$

Proof. We just prove the first inequality. The second inequality then follows upon considering the operation

$$\det(I + MZ) := f_M(z_1, \ldots, z_n) \mapsto z_1 \cdots z_n f_M(z_1^{-1}, \ldots, z_n^{-1})$$

with $M = A$ and $M = B$, respectively, which preserves the set of stable polynomials with non-negative coefficients (Proposition 3.1) and obviously reverses the partial order $\trianglelefteq$ (Theorem 4.11). Similarly, for the third inequality note that if $A \leq B$ and A, B are positive contractions, then by continuity we may assume that $B < I$. But then $A' = A(I - A)^{-1} \leq B(I - B)^{-1} = B'$ by Lemma 4.14, hence

$$\det(I - A + AZ) = \frac{\det(I + A'Z)}{\det(I + A')} \trianglelefteq \frac{\det(I + B'Z)}{\det(I + B')} = \det(I - B + BZ).$$

We claim that the polynomial $f \in \mathbb{R}[z_1, \ldots, z_n, y]$ defined by

$$f(z, y) = \det(y(I + AZ) + I + BZ) := y^n P_0(z) + y^{n-1} P_1(z) + \cdots + P_n(z),$$

where $z = (z_1, \ldots, z_n)$ is real stable with non-negative coefficients. Assuming this claim we see that since $P_k(z) = (n - k)!^{-1} \partial^{n-k} f/\partial y^{n-k}\big|_{y=0}$, by Remark 4.1 we have that

$$\det(I + AZ) = P_0 \ll P_1 \ll \cdots \ll P_n = \det(I + BZ),$$

which then settles the lemma (note that $P_k(0) = \binom{n}{k}$, so none of the P_k's is identically zero). To prove the remaining claim, note that the polynomial $G(z, y) = \det(A + y(B - A) + Z)$ is stable with non-negative coefficients by Proposition 3.2. Hence, so are the polynomials

$$F(z, y) = z_1 \cdots z_n y^n G(z_1^{-1}, \ldots, z_n^{-1}, y^{-1}) = \det(y(I + AZ) + (B - A)Z)$$

and $F(z, y + 1) = \det(y(I + AZ) + I + BZ) = f(z, y)$. $\qquad\square$

Proof of Theorem 4.13. The generating polynomials of μ and ν are $\det(I - A + AZ)$ and $\det(I - B + BZ)$, respectively. Combine Propositions 4.12 and 4.15. $\qquad\square$

We end this section with an open question, namely whether the following converse to Proposition 4.15 holds.

Question 4.1. Let A and B be positive semi-definite $n \times n$ matrices such that

$$\frac{\det(I + AZ)}{\det(I + A)} \trianglelefteq \frac{\det(I + BZ)}{\det(I + B)}, \quad \text{or equivalently,} \quad \frac{\det(B + Z)}{\det(B + I)} \trianglelefteq \frac{\det(A + Z)}{\det(A + I)},$$

where $Z = \mathrm{diag}(z_1, \ldots, z_n)$. Is it true that $A \leq B$?

Note that an affirmative answer to Question 4.1 would give a characterization of the Loewner order on the cone of all positive semi-definite $n \times n$ matrices in terms of determinantal stable polynomials.

Remark 4.3. Stochastic domination is closely related to the notion of majorization (or spectral order) for real vectors that was axiomatized by Hardy-Littlewood-Pólya and Schur [38, 61]. The latter also induces a partial order (though not the same as the proper position) on univariate real-rooted polynomials; see, e.g., [6].

4.3.2. *Truncations of strongly Rayleigh measures and polynomials.* We will now discuss certain truncations of stable polynomials and strongly Rayleigh measures.

Lemma 4.16. *Suppose that $f(z) = \sum_{\alpha \in \mathbb{N}^n} a(\alpha) z^\alpha \in \mathbb{R}[z_1, \ldots, z_n]$ is stable of total degree at most d and has non-negative coefficients. For $0 \leq k \leq d$ let*

$$E_k(z) = \sum_{|\alpha|=k} a(\alpha) z^\alpha.$$

If $0 \leq p \leq q \leq d$, then

$$\text{(12)} \qquad \sum_{k=p}^{q} \binom{q-p}{k-p} \frac{E_k(z)}{\binom{d}{k}} y^{q-k}$$

is a stable polynomial in the variables $z_1, \ldots, z_n, y$.

Proof. Define a sequence of polynomials in $\mathbb{R}[z_1, \ldots, z_n, y]$ as follows:

$$
\begin{aligned}
f_1(z, y) &= \sum a(\alpha) z^\alpha y^{d-|\alpha|} = \sum_{k=0}^{d} E_k(z) y^{d-k}, \\
f_2(z, y) &= \partial^{d-q} f_1 / \partial y^{d-q}, \\
f_3(z, y) &= y^q z_1^d \cdots z_n^d f_2(z_1^{-1}, \ldots, z_n^{-1}, y^{-1}), \\
f_4(z, y) &= \partial^p f_3 / \partial y^p, \\
f_5(z, y) &= y^{q-p} z_1^d \cdots z_n^d f_4(z_1^{-1}, \ldots, z_n^{-1}, y^{-1}).
\end{aligned}
$$

By Theorem 4.5, f_1 is a stable polynomial, and by the closure properties of stable polynomials (Proposition 3.1) we also have that $f_2, \ldots, f_5$ are stable. The polynomial f_5 is a constant multiple of (12). $\qquad \square$

The above lemma gives the following generalization of Newton's inequalities, which correspond to the special case when $f(z) = (1 + z_1) \cdots (1 + z_n)$.

Corollary 4.17. *Let $f(z)$ and $E_k(z)$, $k = 0, \ldots, d$, be as in the statement of Lemma 4.16. Then*

$$\text{(13)} \qquad \frac{E_k^2(z)}{\binom{d}{k}^2} \geq \frac{E_{k-1}(z)}{\binom{d}{k-1}} \frac{E_{k+1}(z)}{\binom{d}{k+1}}, \quad z \in \mathbb{R}^n, 1 \leq k \leq d-1.$$

Proof. By Lemma 4.16, the polynomial

$$\text{(14)} \qquad (z, y) \mapsto \frac{E_{k-1}(z)}{\binom{d}{k-1}} y^2 + 2 \frac{E_k(z)}{\binom{d}{k}} y + \frac{E_{k+1}(z)}{\binom{d}{k+1}}$$

is stable. Hence, for any fixed $z \in \mathbb{R}^n$ the resulting (univariate) polynomial in y is real-rooted. The corollary follows upon taking the discriminant in (14). $\qquad \square$

Recall from Definition 2.15 the truncation $\mu_{p,q}$ of a measure μ to $[p, q]$. It has been discussed whether truncations of measures preserve negative dependence properties; cf. [45, 65] and the references therein. As we show in Corollary 4.18 below, this is the case for strongly Rayleigh measures and truncations of length $q - p \leq 1$. Truncations of length $q - p \geq 2$ do not preserve the strong Rayleigh property. However, the weighted truncation suggested by (12) obviously does.

Corollary 4.18. *Suppose that μ is a strongly Rayleigh probability measure on $2^{[n]}$ and that $0 \leq p \leq q \leq n$ with $q - p \leq 1$. Then $\mu_{p,q}$ is strongly Rayleigh.*

Proof. We prove the corollary for $q - p = 1$. Let f be the generating polynomial of μ and let E_k be as in the statement of Lemma 4.16. Then the polynomial

$$(15) \qquad g(z, y) = \frac{E_p(z)}{\binom{n}{p}} y + \frac{E_q(z)}{\binom{n}{q}}$$

is stable. Clearly, the generating polynomial of $\mu_{p,q}$ is a (positive) constant multiple of $g(z, q/(n - p))$. The corollary follows since stable polynomials are closed under setting variables equal to real numbers; see Proposition 3.1. $\qquad\square$

In §7 we show that Pemantle's Conjecture 2.9 (see §2.5) fails in its full generality. However, it is true for strongly Rayleigh measures:

Theorem 4.19. *Let μ be a strongly Rayleigh probability measure on $2^{[n]}$, and let $1 \leq k \leq n$. If $\mu(\{S : |S| = k - 1\})\mu(\{S : |S| = k\}) \neq 0$, then $\mu_{k-1} \preccurlyeq \mu_k$.*

Proof. Let f be the generating polynomial of μ and let E_k be as in the statement of Lemma 4.16. By assumption one has $E_{k-1}(\mathbf{1})E_k(\mathbf{1}) \neq 0$, where $\mathbf{1} \in \mathbb{R}_+^n$ is as before the "all ones vector". From Lemma 4.16 and the last part of Remark 4.1 we deduce that the polynomial

$$F(z_1, \ldots, z_{n+1}) = z_{n+1}\frac{E_{k-1}(z_1, \ldots, z_n)}{E_{k-1}(\mathbf{1})} + \frac{E_k(z_1, \ldots, z_n)}{E_k(\mathbf{1})}$$

is stable. Since the above quotients are the generating polynomials of μ_{k-1} and μ_k, respectively, we have $\mu_{k-1} \trianglelefteq \mu_k$ by Theorem 4.11 and Definition 4.1. The desired conclusion follows from Proposition 4.12. $\qquad\square$

4.4. **The partial symmetrization procedure.** Recall Remark 2.1, the partial symmetrization of (complex) measures on $2^{[n]}$ defined in (3), and Definition 2.9. The main result of this section is the following theorem.

Theorem 4.20. *Let μ be a complex measure on $2^{[n]}$ with a (complex) stable generating polynomial. Then for any transposition $\tau \in \mathfrak{S}_n$ and $\theta \in [0, 1]$ the generating polynomial of the partial symmetrization $\mu^{\tau,\theta} = \theta\mu + (1-\theta)\tau(\mu)$ is (complex) stable.*

In particular, Theorem 4.20 implies that the class of strongly Rayleigh measures on $2^{[n]}$ is invariant under the partial symmetrization procedure. Clearly, the latter result is real stable in nature. However, in order to establish it we have to consider the (wider) complex stable context (cf. Definition 2.9), as in the above theorem.

For the proof of Theorem 4.20 we need a multivariate generalization of the classical Obreschkoff theorem that was obtained in [9].

Theorem 4.21 (Multivariate Obreschkoff theorem)**.** *Let $f, g \in \mathbb{R}[z_1, \ldots z_n]$. Then all non-zero polynomials in the space $\{\alpha f + \beta g : \alpha, \beta \in \mathbb{R}\}$ are stable if and only if either $f + ig$ or $f - ig$ is stable or $f = g \equiv 0$.*

Remark 4.4. There are now elementary proofs of Theorem 4.20 that avoid the multivariate Obresckoff theorem; see [10, 56]. Also, the referee sketched an elementary proof similar to those given in [10, 56].

Proof of Theorem 4.20. Let $f \in \mathbb{C}[z_1, \ldots, z_n]$ be the generating polynomial of μ. Hence f is stable and multi-affine. Assuming, as we may, that $i = 1$ and $j = 2$, we need to prove that

$$\theta f(\xi_1, \xi_2, \ldots, \xi_n) + (1 - \theta)f(\xi_2, \xi_1, \ldots, \xi_n) \neq 0$$

whenever $\xi_1, \ldots, \xi_n \in \{z : \mathfrak{Im}(z) > 0\}$ (cf. Definition 2.9). By fixing $\xi_3, \ldots, \xi_n$ arbitrarily in the open upper half-plane and considering the multi-affine polynomial in variables z_1, z_2 given by

$$(z_1, z_2) \mapsto g(z_1, z_2) := f(z_1, z_2, \xi_3, \ldots, \xi_n),$$

we see that the problem reduces to proving that the linear operator T_θ on multi-affine polynomials in two variables defined by

$$T_\theta(h)(z_1, z_2) = \theta h(z_1, z_2) + (1 - \theta)h(z_2, z_1)$$

preserves stability. Let us first show that T_θ preserves real stability. For this let

$$f(z_1, z_2) = a_{00} + a_{01}z_2 + a_{10}z_1 + a_{11}z_1z_2 \in \mathbb{R}[z_1, z_2] \setminus \{0\}$$

be real stable and set $D_f = a_{00}a_{11} - a_{01}a_{10}$. By Theorem 4.1 we know that f is real stable if and only if $D_f \leq 0$. Since

$$D_{T_\theta(f)} = D_f - \theta(1 - \theta)(a_{01} - a_{10})^2 \leq D_f,$$

it follows that $T_\theta(f)$ is also real stable, as required. To deal with the (non-real) stable case first let $h = g + if \in \mathbb{C}[z_1, z_2]$ be *strictly stable* and multi-affine. This means that $h(z_1, z_2) \neq 0$ whenever $\mathfrak{Im}(z_1) \geq 0$ and $\mathfrak{Im}(z_2) \geq 0$. By the multivariate Obreschkoff theorem we have that $\alpha g + \beta f$ is real stable or zero for all $\alpha, \beta \in \mathbb{R}$. Hence, by the above, we also have that $\alpha T_\theta(f) + \beta T_\theta(g)$ is real stable or zero for all $\alpha, \beta \in \mathbb{R}$. By the multivariate Obreschkoff theorem again this gives that $T_\theta(g + if)$ or $T_\theta(g - if)$ is stable. However, one has

$$T_\theta(h)(z, z) = h(z, z) = g(z, z) + if(z, z),$$

which is strictly stable by assumption. If $T_\theta(g - if)$ is stable, then $g(z, z) - if(z, z)$ is stable and $g(z, z) + if(z, z)$ is strictly stable, which is impossible since it would imply that all zeros of (the strictly stable polynomial) $g(z, z) + if(z, z)$ were real. Hence $T_\theta(g + if)$ is stable.

Now if $\epsilon > 0$ and $p \in \mathbb{C}[z_1, z_2]$ is stable, then clearly $p(z_1 + i\epsilon, z_2 + i\epsilon)$ is strictly stable (in the sense defined above), and thus any stable polynomial is the uniform limit on compact sets of strictly stable polynomials. By the multivariate version of Hurwitz' classical theorem on the "continuity of zeros" (see, e.g., [8, Theorem 2.3] and [19, Footnote 3, p. 96]) this proves that T_θ preserves stability. $\square$

Remark 4.5. From Theorem 4.20 it follows that the symmetrization μ_s of a strongly Rayleigh measure μ is also strongly Rayleigh. A different proof is as follows. If μ is a strongly Rayleigh measure on $2^{[n]}$, then (2) implies that the diagonal specialization $\Delta(g_{\mu_s})(t) = \Delta(g_\mu)(t)$ (cf. Definition 2.8) is a real-rooted univariate polynomial and by the Grace-Walsh-Szegö theorem we conclude that the symmetrization μ_s of μ is also strongly Rayleigh.

5. Preservation of negative dependence
by symmetric exclusion processes

In this section we study some consequences of the results in §4 for symmetric exclusion processes and provide an answer to the problem of finding a natural negative dependence property that is preserved by such evolutions (Problem 3.5).

Let us first describe these processes in more detail. A configuration of particles on a countable set S is a point η on the Boolean lattice 2^S, which we identify with $\{0,1\}^S$, so $\eta(i) = 1$ means that site i is occupied, while $\eta(i) = 0$ means that it is vacant. One is given a non-negative symmetric kernel $Q = (q_{i,j})_{i,j \in S}$ satisfying

$$(16) \qquad \sup_i \sum_j q_{i,j} < \infty.$$

The corresponding exclusion process η_t is the Markov process on $\{0,1\}^S$ in which $\eta \to \tau_{i,j}(\eta)$ at rate $q_{i,j}$, where $\tau_{i,j}$ is the transposition that interchanges the coordinates $\eta(i)$ and $\eta(j)$. Note that if $\eta(i) = \eta(j)$ (i.e., the two sites are both occupied or both empty), this transposition has no effect, while if $\eta(i) \neq \eta(j)$, the transposition has the effect of moving the particle from the occupied site to the vacant site. In the language of semigroups and generators, the semigroup $T(t)$ on $C(\{0,1\}^S)$ defined by $T(t)F(\eta) = E^\eta F(\eta_t)$ has a generator

$$\mathcal{L}F(\eta) = \frac{1}{2} \sum_{i,j} q_{i,j} [F(\tau_{i,j}(\eta)) - F(\eta)]$$

for functions F that depend on finitely many coordinates. Condition (16) guarantees that the process is well defined and uniquely determined by Q.

One can give a more probabilistic description of the symmetric exclusion process in terms of Poisson processes. Recall that a Poisson process $N(t)$ with rate $\lambda > 0$ is a random increasing step function with jumps of size one, in which the times between jumps are independent exponentially distributed random variables τ_k with parameter λ: $P(\tau_k > t) = e^{-\lambda t}$. Now take a collection $N_{i,j}(t)$ of independent Poisson processes indexed by unordered pairs i, j ($N_{i,j}(t)$ with rate $q_{i,j}$) and apply the transposition $\tau_{i,j}$ to η at the jump times of $N_{i,j}(t)$. We will denote probabilities for the process with initial configuration η by P^η.

Negative correlation inequalities have played an important role in the development of the theory of symmetric exclusion processes. For example, Proposition 1.7 of [52, Chap. VIII] implies that

$$(17) \qquad P^\eta(\eta_t \equiv 1 \text{ on } A) \leq \prod_{i \in A} P^\eta(\eta_t(i) = 1), \quad A \subseteq S.$$

This inequality was used in an essential way in the characterization of stationary distributions of the process; see Theorem 1.44 in *op. cit.* and the end of this section. Inequality (17) was generalized by Andjel in [2] to

$$(18) \qquad \begin{aligned} &P^\eta(\eta_t \equiv 1 \text{ on } A \cup B) \leq P^\eta(\eta_t \equiv 1 \text{ on } A)P^\eta(\eta_t \equiv 1 \text{ on } B), \\ &A, B \subseteq S, \ A \cap B = \emptyset, \end{aligned}$$

and used to prove an ergodic theorem. For further applications of correlation inequalities in this setting; see the references and discussion in [55].

In that paper, one of us conjectured that the distribution of η_t at time t with deterministic initial configuration η is negatively associated. Even the special case

(19)
$$P^\eta(\eta_t \equiv 1 \text{ on } A, \eta_t \equiv 0 \text{ on } B) \geq P^\eta(\eta_t \equiv 1 \text{ on } A)P^\eta(\eta_t \equiv 0 \text{ on } B),$$
$$A, B \subseteq S, \ A \cap B = \emptyset$$

remained open (note that in (19), one of the events is increasing while the other is decreasing). Surprisingly, despite the similarity between (18) and (19), neither the proof of (18) in [2], nor a somewhat different one given in [55], extends to prove (19). We are now able to prove this conjecture; see Theorem 5.2 below. The latter is actually a consequence of a yet stronger property for finite symmetric exclusion processes that we will now establish.

Proposition 5.1. *Suppose that S is finite and the initial distribution η_0 of a symmetric exclusion process on $\{0,1\}^S$ is strongly Rayleigh. Then so is the distribution of η_t for all $t > 0$.*

Proof. First, we observe that it suffices to prove the statement in case $q_{i,j} \neq 0$ for only one pair i, j. To see this, suppose $T_1(t)$ and $T_2(t)$ are semigroups for finite state Markov chains on the same state space with generators $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively, and let $T(t)$ be the semigroup with generator $\mathcal{L}_1 + \mathcal{L}_2$. The Trotter product formula (see [27, p. 33]) gives

$$T(t) = \lim_{n \to \infty} \big[T_1(t/n)T_2(t/n)\big]^n.$$

It follows that if a closed set of probability measures $\mathcal{M}$ on the state space has the property that $\mu \in \mathcal{M}$ implies $\mu T_i(t) \in \mathcal{M}$ for $i = 1, 2$ and $t > 0$, then $\mu \in \mathcal{M}$ implies $\mu T(t) \in \mathcal{M}$ for $t > 0$. Now write the generator of the symmetric exclusion process as the following sum, corresponding to transitions at individual pairs of sites:

$$\mathcal{L} = \sum_{i,j} \mathcal{L}_{i,j},$$

where

$$\mathcal{L}_{i,j}F(\eta) = \frac{1}{2}q_{i,j}[F(\tau_{i,j}(\eta)) - F(\eta)].$$

Repeated application of the preceding observation with $\mathcal{M} =$ the set of probability measures on $\{0,1\}^S$ with stable generating polynomials, implies that the statement of the proposition is correct for the process with generator $\mathcal{L}$ if it is correct for the process with generator $\mathcal{L}_{i,j}$ for each i, j.

In the case of a single non-zero $q_{i,j}$, the generating polynomial of the distribution of η_t is given by $\alpha f(z) + (1 - \alpha)f(\tau_{i,j}(z))$, where $\alpha = P(N_{i,j}(t) \text{ is even})$ and f is the generating polynomial for the initial distribution. Therefore, the result follows from Theorem 4.20. $\square$

Recall from §2.4 that if S is a countably infinite set, a probability measure on $\{0,1\}^S$ is strongly Rayleigh if every projection onto a finite subset of S is strongly Rayleigh. Note that this property is preserved by weak convergence of measures.

Theorem 5.2. *Suppose that S is countable and the initial distribution η_0 of a symmetric exclusion process on $\{0,1\}^S$ is strongly Rayleigh. Then the distribution of η_t is strongly Rayleigh, and hence CNA+, for all $t > 0$.*

Proof. Since stability (the strongly Rayleigh property) implies negative association by Theorem 4.9, the result follows from Proposition 5.1 when S is finite. For general S, approximate the exclusion process on S by processes on finite subsets S_n of S that increase to S, by taking the transposition rates for the process on S_n to be $q_{i,j}$ if $i,j \in S_n$ and zero otherwise. By the Trotter-Kurtz semigroup convergence theorem (see [27, p. 28]), the processes on S_n converge to the process on S, so the result holds on S as well. (See [52, §I.3] for details.) □

Remark 5.1. In particular, Theorem 5.2 proves Conjecture 2.2 and shows that its conclusion remains valid if the initial configuration is a product measure (since any such measure is obviously strongly Rayleigh; cf. Definition 2.11 and §2.4).

We conclude this section with an application of Theorem 5.2 to the extremal stationary distributions for an irreducible symmetric exclusion process. Theorem 1.44 of [52, Chap. VIII] states that these are exactly $\{\mu_\alpha : \alpha \in \mathcal{H}\}$, where

(a) $\mathcal{H} = \left\{ \alpha : S \to [0,1] : \sum_j q_{i,j}[\alpha(j) - \alpha(i)] = 0 \text{ for each } i \right\}$ are the harmonic functions for Q with values in $[0,1]$, and

(b) $\mu_\alpha = \lim_{t \to \infty} \nu_\alpha T(t)$, where ν_α is the product measure on $\{0,1\}^S$ satisfying

$$\nu_\alpha\{\eta : \eta \equiv 1 \text{ on } A\} = \prod_{i \in A} \alpha(i), \quad A \subseteq S.$$

Very little is known about μ_α other than

(I) $\mu_\alpha = \nu_\alpha$ if and only if α is constant, and

(II) $\mu_\alpha\{\eta : \eta(i) = 1\} = \alpha(i)$ for each $i \in S$.

As a consequence of Theorem 5.2, we can now add

(III) μ_α is negatively associated for each $\alpha \in \mathcal{H}$.

Remark 5.2. If S is finite and $q_{i,j} > 0$ for all i,j, then as we noted in §3.5 the limiting distribution of η_t as $t \to \infty$ is the symmetrization of the initial distribution. In §7 we construct examples of measures in $\mathfrak{P}_n$, $n \geq 20$, that are CNA+ but do not have ULC rank sequences. By Theorem 3.7 in §3.5, these counterexamples to Conjectures 2.3–2.5 show that none of the properties NLC, h-NLC, Rayleigh/h-NLC+, CNA, or CNA+ is preserved by the symmetrization procedure (2), and by [55, Theorem 3.2], the same can be said about the NA property. Therefore, the analog of Proposition 5.1 fails for all the aforementioned properties.

Remark 5.3. The analog of Theorem 5.2 for asymmetric exclusion processes is false. A simple example is obtained by taking $S = \{1,2\}$ in which only transitions from state 1 to state 2 are allowed. If the initial distribution gives probability $\frac{1}{4}$ to each of the configurations in $\{0,1\}^S$, then the limiting distribution as $t \to \infty$ is given by $\mu(\{11\}) = \mu(\{00\}) = \frac{1}{4}$, $\mu(\{10\}) = 0$, and $\mu(\{01\}) = \frac{1}{2}$. This measure is not NA.

6. Almost exchangeable measures

By Pemantle's result (Theorem 3.7 in §3.5), for symmetric measures in $\mathfrak{P}_n$ all five conditions NLC, h-NLC, Rayleigh/h-NLC+, CNA, and CNA+ are equivalent to the ULC condition for the rank sequence. Therefore, the relations between all these negative dependence properties are completely understood in the exchangeable case. In this section we study the almost exchangeable case, that is, measures whose generating polynomials are symmetric in all but possibly one variable (Definition 2.1). We prove that for such measures one has both h-NLC $\Leftrightarrow$ CNA and

Rayleigh/h-NLC+ $\Leftrightarrow$ CNA+, which confirms Pemantle's Conjecture 2.6 in this case. In the next section we will use this result to show that the counterexamples in 20 or more variables that we construct there satisfy all negative dependence properties other than strongly Rayleigh.

Let C be an $m \times n$ matrix (where m, n are allowed to be infinite), and let $\mathbf{i}$ and $\mathbf{j}$ be finite equi-numerous subsets of $[m]$ and $[n]$, respectively. Denote by $C(\mathbf{i}, \mathbf{j})$ the minor of C with rows indexed by $\mathbf{i}$ and columns indexed by $\mathbf{j}$. The Cauchy-Binet theorem, see, e.g., [49], asserts that

$$(AB)(\mathbf{i}, \mathbf{j}) = \sum_{\mathbf{k},\, |\mathbf{k}| = |\mathbf{i}| = |\mathbf{j}|} A(\mathbf{i}, \mathbf{k}) B(\mathbf{k}, \mathbf{j})$$

for any $m \times n$ matrix A and $n \times q$ matrix B, where $1 \leq m, n, q \leq \infty$, provided that the sum in the right-hand side converges absolutely.

Recall the notion of (non-negative) LC sequence from Definition 2.8. For convenience, we first prove a technical result that will be frequently used later on.

Lemma 6.1. *Let $\{a_j\}_{j=0}^n$ and $\{b_j\}_{j=0}^n$ be two LC sequences that satisfy*
   (i) $\{a_j + b_j\}_{j=0}^n$ *is LC, and*
   (ii) $a_j b_{j+1} \geq a_{j+1} b_j$ *for all $0 \leq j \leq n - 1$.*
If $\ell < k$ and $b_\ell a_k > 0$, then $b_r a_r > 0$ for all $\ell \leq r \leq k$.

Proof. Suppose that $a_r = 0$ for some $\ell \leq r < k$ and let s be the largest such index. Assumption (ii) for $j = s$ gives $b_s = 0$, hence $a_s + b_s = 0$. This contradicts the "no internal zeros" part of assumption (i) (cf. Definition 2.8). Therefore $a_r > 0$, and a similar argument shows that $b_r > 0$ as well. $\qquad\square$

The following proposition shows that for almost exchangeable measures the Rayleigh property is the same as the negative lattice condition (NLC) closed under conditioning and external fields on the last variable. Indeed, condition (ii) follows directly from NLC while (i) follows from NLC after having imposed an external field on the last coordinate and then projected onto the first n coordinates.

Proposition 6.2. *Let $\mu \in \mathfrak{P}_{n+1}$ be such that its generating polynomial $g = g_\mu$ is symmetric in its first n variables, so that*

$$g(z_1, \ldots, z_{n+1}) = z_{n+1} \sum_{k=0}^n a_k e_k(z_1, \ldots, z_n) + \sum_{k=0}^n b_k e_k(z_1, \ldots, z_n).$$

Then μ is Rayleigh if and only if
   (i) $\{\theta a_k + (1 - \theta) b_k\}_{k=0}^n$ *is LC for all $0 \leq \theta \leq 1$, and*
   (ii) $a_k b_{k+1} \geq a_{k+1} b_k$ *for all $0 \leq k \leq n - 1$.*

Proof. Suppose first that μ is Rayleigh. By Proposition 2.1 we get that if $0 \leq \theta < 1$, then $g(z_1, \ldots, z_n, \theta/(1 - \theta))$ is a Rayleigh polynomial that is symmetric in all its n variables. Condition (i) now follows from the exchangeable case (Theorem 3.7). Let $S \subseteq [n - 1]$ be of cardinality i. The NLC condition gives

$$a_i b_{i+1} = \mu(S \cup \{n+1\}) \mu(S \cup \{n\}) \geq \mu(S \cup \{n, n+1\}) \mu(S) = a_{i+1} b_i,$$

which verifies (ii).

Assume now that $\{a_k\}_{k=0}^n$ and $\{b_k\}_{k=0}^n$ satisfy (i) and (ii). The inequality

$$\frac{\partial g}{\partial z_i}(x) \frac{\partial g}{\partial z_j}(x) \geq \frac{\partial^2 g}{\partial z_i \partial z_j}(x) g(x), \quad x \in \mathbb{R}_+^{n+1}, \ i, j \in [n],$$

follows from the exchangeable case and (i). What is left to prove is the inequality

$$\frac{\partial g}{\partial z_i}(x)\frac{\partial g}{\partial z_{n+1}}(x) \geq \frac{\partial^2 g}{\partial z_i \partial z_{n+1}}(x)g(x), \quad x \in \mathbb{R}_+^{n+1}, \, i \in [n],$$

and by symmetry (in the first n variables) we only need to check it for, e.g. $i = n$. We may write g as

$$g(z_1, \ldots, z_{n+1}) = \sum_{k=0}^{n} b_k e_k(z_1, \ldots, z_{n-1}) + z_n \sum_{k=0}^{n} b_{k+1} e_k(z_1, \ldots, z_{n-1})$$

$$+ z_{n+1} \sum_{k=0}^{n} a_k e_k(z_1, \ldots, z_{n-1}) + z_n z_{n+1} \sum_{k=0}^{n} a_{k+1} e_k(z_1, \ldots, z_{n-1}),$$

so the desired inequality is equivalent to

$$(20) \qquad \left(\sum_j a_j e_j\right)\left(\sum_j b_{j+1} e_j\right) - \left(\sum_j a_{j+1} e_j\right)\left(\sum_j b_j e_j\right) \geq 0$$

for all $x \in \mathbb{R}_+^{n-1}$, where $e_j = e_j(x_1, \ldots, x_{n-1})$. Now let

$$C = \left(\begin{array}{ccccc} a_0 & a_1 & a_2 & \ldots & a_n \\ b_0 & b_1 & b_2 & \ldots & b_n \end{array}\right)$$

and recall the definition of a TP_p matrix from §3.5. We claim that C is TP_2. Let $1 \leq i < j \leq n$. We need to prove that $a_i b_j \geq b_i a_j$. Clearly, we may assume that $b_i a_j \neq 0$ and that $j - i \geq 2$. By Lemma 6.1 we have $b_r a_r > 0$ for $i \leq r \leq j$ and thus

$$\frac{a_i}{b_i} \geq \frac{a_{i+1}}{b_{i+1}} \geq \cdots \geq \frac{a_j}{b_j}$$

by (ii), which shows that C is indeed TP_2. Let $E = (e_{i-j})_{i,j=0}^{n}$. By Theorem 3.8 we have that E is TP_2, and by the Cauchy-Binet formula so is CE. Now

$$CE = \left(\begin{array}{ccc} \sum a_j e_j & \sum a_{j+1} e_j & \cdots \\ \sum b_j e_j & \sum b_{j+1} e_j & \cdots \end{array}\right),$$

from which (20) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The next result is the analog of Proposition 6.2 for the h-NLC property.

Proposition 6.3. *Let $\mu \in \mathfrak{P}_{n+1}$ be such that its generating polynomial $g = g_\mu$ is symmetric in its first n variables, so that*

$$g(z_1, \ldots, z_{n+1}) = z_{n+1} \sum_{k=0}^{n} a_k e_k(z_1, \ldots, z_n) + \sum_{k=0}^{n} b_k e_k(z_1, \ldots, z_n).$$

Then μ is h-NLC if and only if

(i) $\{a_k\}_{k=0}^{n}$, $\{b_k\}_{k=0}^{n}$, *and* $\{a_k + b_k\}_{k=0}^{n}$ *are LC, and*
(ii) $a_k b_{k+1} \geq a_{k+1} b_k$ *for all* $0 \leq k \leq n - 1$.

Proof. Let $S, T \subset [n]$. The NLC condition applied to each of the three pairs (S, T), $(S, T \cup \{n + 1\})$, and $(S \cup \{n + 1\}, T \cup \{n + 1\})$ gives, respectively,

$$b_{|S|} b_{|T|} \geq b_{|S \cup T|} b_{|S \cap T|}, \quad b_{|S|} a_{|T|} \geq a_{|S \cup T|} b_{|S \cap T|}, \quad a_{|S|} a_{|T|} \geq a_{|S \cup T|} a_{|S \cap T|}.$$

Therefore, setting $\ell = |S|$, $k = |T|$, $m = |S \cap T|$, we see that NLC is equivalent to

(21) $b_\ell b_k \geq b_{k+\ell-m} b_m, \quad b_\ell a_k \geq a_{k+\ell-m} b_m, \quad a_\ell a_k \geq a_{k+\ell-m} a_m$

for $m \leq k, \ell \leq n$ and $k + \ell \leq n + m$.

If $S \subseteq [n]$ has cardinality $n - j$, the measure obtained by projecting μ onto $2^{S \cup \{n+1\}}$ is of the same form as μ with coefficients

$$a_k(j) = \sum_{i=0}^{j} \binom{j}{i} a_{k+i} \quad \text{and} \quad b_k(j) = \sum_{i=0}^{j} \binom{j}{i} b_{k+i}, \quad 0 \leq k \leq n - j,$$

while the one obtained by projecting μ onto 2^S is exchangeable with coefficients $a_k(j) + b_k(j)$, $0 \leq k \leq n - j$. Therefore, h-NLC is equivalent to the statement that for each $0 \leq j \leq n$ one has

$$b_\ell(j) b_k(j) \geq b_{k+\ell-m}(j) b_m(j),$$

(22) $b_\ell(j) a_k(j) \geq a_{k+\ell-m}(j) b_m(j), \ a_\ell(j) a_k(j) \geq a_{k+\ell-m}(j) a_m(j),$

$$(a_\ell(j) + b_\ell(j))(a_k(j) + b_k(j)) \geq (a_{k+\ell-m}(j) + b_{k+\ell-m}(j))(a_m(j) + b_m(j))$$

for $m \leq k, \ell \leq n - j$ and $k + \ell \leq n - j + m$.

It follows from this that the h-NLC property implies that the sequences in (i) have no internal zeros. For example, if $k \leq \ell$, $b_{k-1} > 0, b_k = \cdots = b_\ell = 0$, $b_{\ell+1} > 0$ contradicts $b_k b_\ell \geq b_{k-1} b_{\ell+1}$. Given this, the fact that the three sequences are LC and satisfy (ii) is a consequence of the first and last inequalities in (21) with $\ell = k$, $m = k - 1$, the middle inequality in (21) with $\ell = k + 1$, $m = k$, and the last inequality in (22) with $j = 0$, $\ell = k$, $m = k - 1$.

For the converse, assume that (i) and (ii) hold. Note that for each j, $\{a_k(j)\}_{k=0}^{n-j}$ is a convolution of two LC sequences, and hence is LC by [49, Theorem 1.2 on p. 394], and the same holds for the sequences $b_k(j)$ and $a_k(j) + b_k(j)$. This gives three of the four inequalities in (22). For the other one, note that since evaluating the elementary symmetric polynomials at the "all ones vector" produces binomial coefficients, it follows from (20) that

(23) $a_k(j) b_{k+1}(j) \geq a_{k+1}(j) b_k(j).$

We need to check that

$$a_k(j) b_\ell(j) \geq a_{k+\ell-m}(j) b_m(j)$$

for $m \leq k, \ell$, and in doing so, can assume that the right-hand side is not zero. By Lemma 6.1, it follows that $a_i(j) b_i(j) > 0$ for $m \leq i \leq k + \ell - m$. To conclude the proof, write

$$\frac{a_k(j)}{a_{k+\ell-m}(j)} \geq \frac{a_m(j)}{a_\ell(j)} \geq \frac{b_m(j)}{b_\ell(j)}.$$

The first inequality comes from the already proved part of (22), while the second is obtained by repeated application of (23).   □

Theorem 6.4. *Let $\mu \in \mathfrak{P}_{n+1}$ be such that its generating polynomial $g = g_\mu$ is symmetric in its first n variables, so that*

$$g(z_1, \ldots, z_n) = z_{n+1} \sum_{k=0}^{n} a_k e_k(z_1, \ldots, z_n) + \sum_{k=0}^{n} b_k e_k(z_1, \ldots, z_n).$$

Suppose that $\{a_k\}_{k=0}^n$ and $\{b_k\}_{k=0}^n$ are LC sequences that satisfy

(i) $\{a_k + b_k\}_{k=0}^n$ *is LC, and*

(ii) $a_k b_{k+1} \geq a_{k+1} b_k$ *for all $0 \leq k \leq n - 1$.*

Then μ is CNA.

Proof. As in §2.1, we denote by X_i, $i \in [n+1]$, the i-th coordinate function/random (binary) variable on $2^{[n+1]}$. First, observe that it is enough to prove that μ is NA. To see this, note that if μ is conditioned on $X_i = 1$ for k values of $i \leq n$, on $X_i = 0$ for l values of $i \leq n$, and possibly on the value of X_{n+1}, the resulting measure is of the same form with a new value n' of n and new a_i's and b_i's that again satisfy the assumptions of the theorem. Using primes to denote the new values, we have $n' = n - k - l$, and if the value of X_{n+1} is not conditioned on, then, except for a constant factor, $a_i' = a_{i+k}$ and $b_i' = b_{i+k}$. If the value of X_{n+1} is conditioned on, then $a_i' = b_i' = a_{i+k}$ if the conditioning is on $X_{n+1} = 1$ and $a_i' = b_i' = b_{i+k}$ if the conditioning is on $X_{n+1} = 0$. Now take two increasing functions F and G on $2^{[n+1]}$ that depend on disjoint sets of coordinates (cf. Definition 2.2). Without loss of generality, we may assume that there is an m satisfying $1 \leq m \leq n$ such that F depends on the coordinates $\{X_1 \ldots, X_m\}$, and G depends on the coordinates $\{X_{m+1}, \ldots, X_{n+1}\}$. We need to prove

$$(24) \qquad \int FG d\mu \leq \int F d\mu \int G d\mu.$$

Since μ is symmetric in the first n coordinates, these three integrals are not changed if F and G are replaced by the functions obtained by symmetrizing them with respect to the coordinates $\{X_1, \ldots, X_m\}$ and $\{X_{m+1}, ..., X_n\}$, respectively. These symmetrized functions are also increasing. Therefore, we may assume that F and G are of the following form:

$$F(X_1, \ldots, X_m) = f_k \text{ if } \sum_{i=1}^m X_i = k,$$

$$G(X_{m+1}, \ldots, X_n, X_{n+1} = 0) = g_k \text{ if } \sum_{i=m+1}^n X_i = k, \text{ and}$$

$$G(X_{m+1}, \ldots, X_n, X_{n+1} = 1) = h_k \text{ if } \sum_{i=m+1}^n X_i = k,$$

where f_k, g_k and h_k are increasing, and $g_k \leq h_k$ for each k. Using the fact that the covariance of F and G with respect to the measure μ can be written as

$$\frac{1}{2} \sum_{S,T \in 2^{[n+1]}} [F(S) - F(T)][G(S) - G(T)]\mu(S)\mu(T),$$

(24) becomes

$$\sum_{i,j,k,\ell} \binom{m}{i}\binom{m}{j}\binom{n-m}{k}\binom{n-m}{\ell}(f_j - f_i)\left[(h_k - h_\ell)a_{k+i}a_{\ell+j}\right.$$

$$\left. + (g_k - g_\ell)b_{k+i}b_{\ell+j} + (h_k - g_\ell)a_{k+i}b_{\ell+j} + (g_k - h_\ell)a_{\ell+j}b_{k+i}\right] \geq 0.$$

Note that we are using the usual convention that $\binom{n}{m} = 0$ unless $0 \le m \le n$. Since the summand above is not changed if the roles of i and j are interchanged and the roles of k and ℓ are interchanged, this can be written as

$$(25) \quad \sum_{i<j} \sum_{k,\ell} \binom{m}{i}\binom{m}{j}\binom{n-m}{k}\binom{n-m}{\ell}(f_j - f_i)\Big[(h_k - h_\ell)a_{k+i}a_{\ell+j}$$

$$+ (g_k - g_\ell)b_{k+i}b_{\ell+j} + (h_k - g_\ell)a_{k+i}b_{\ell+j} + (g_k - h_\ell)a_{\ell+j}b_{k+i}\Big] \ge 0.$$

So, it is enough to show that for fixed $i < j$ and fixed p,

$$\sum_{k+\ell=p} \binom{n-m}{k}\binom{n-m}{\ell}\Big[(h_k - h_\ell)a_{k+i}a_{\ell+j}$$

$$+ (g_k - g_\ell)b_{k+i}b_{\ell+j} + (h_k - g_\ell)a_{k+i}b_{\ell+j} + (g_k - h_\ell)a_{\ell+j}b_{k+i}\Big] \ge 0.$$

Furthermore, it is enough to prove this in the case when the g_k's and h_k's take only the values 0 and 1, since any increasing function on a partially ordered set (in this case, $\{0,\ldots,n\} \times \{0,1\}$) can be written as a positive linear combination of increasing functions that take only the values 0 and 1.

We will in fact prove that for such g_k's and h_k's,

$$S_q = \sum_{k+\ell=p,\,|k-\ell|\le 2q} \binom{n-m}{k}\binom{n-m}{\ell}\Big[(h_k - h_\ell)a_{k+i}a_{\ell+j}$$

$$+ (g_k - g_\ell)b_{k+i}b_{\ell+j} + (h_k - g_\ell)a_{k+i}b_{\ell+j} + (g_k - h_\ell)a_{\ell+j}b_{k+i}\Big] \ge 0$$

by induction on q. For the basis step, note that $S_0 = 0$ if p is odd, while if $p = 2r$ is even, then

$$S_0 = \binom{n-m}{r}^2 (h_r - g_r)(a_{r+i}b_{r+j} - a_{r+j}b_{r+i}),$$

which is non-negative by (ii), since $i < j$.

We will now prove that $S_q \ge 0$, assuming that $S_{q-1} \ge 0$. The difference $S_q - S_{q-1}$ consists of two summands, which when combined, become

(26)
$$\binom{n-m}{k}\binom{n-m}{\ell}\Big[(h_k-h_\ell)(a_{k+i}a_{\ell+j}-a_{\ell+i}a_{k+j})+(g_k-g_\ell)(b_{k+i}b_{\ell+j}-b_{\ell+i}b_{k+j})$$

$$+ (h_k - g_\ell)(a_{k+i}b_{\ell+j} - a_{k+j}b_{\ell+i}) + (g_k - h_\ell)(a_{\ell+j}b_{k+i} - a_{\ell+i}b_{k+j})\Big],$$

where $k + \ell = p$, while $k - \ell = 2q$ if p is even and $k - \ell = 2q - 1$ if p is odd. Now we consider the possible values of g_ℓ, h_k for this pair k, ℓ. Recall that $g_\ell \le h_\ell, g_k \le h_k$. If $g_\ell = h_k$, then $g_t = h_t =$ this common value for all $\ell \le t \le k$, so that (26) is zero. So, we may assume that $g_\ell = 0, h_k = 1$. Now there are four possible values for the pair (h_ℓ, g_k) (note that the use of the fractions that appear below is justified by Lemma 6.1):

  (a) $h_\ell = g_k = 1$. In this case, (26) becomes

$$\binom{n-m}{k}\binom{n-m}{\ell}\Big[(b_{k+i}b_{\ell+j} - b_{\ell+i}b_{k+j}) + (a_{k+i}b_{\ell+j} - a_{k+j}b_{\ell+i})\Big].$$

This is non-negative by the log-concavity of the b's and (ii):

$$\frac{a_{k+i}}{a_{k+j}} \geq \frac{b_{k+i}}{b_{k+j}} \geq \frac{b_{\ell+i}}{b_{\ell+j}}.$$

(b) $h_\ell = 0, g_k = 1$. Now (26) is

$$\binom{n-m}{k}\binom{n-m}{\ell}\Big[(a_{k+i}+b_{k+i})(a_{\ell+j}+b_{\ell+j}) - (a_{\ell+i}+b_{\ell+i})(a_{k+j}+b_{k+j})\Big],$$

which is non-negative by (i).

(c) $h_\ell = g_k = 0$. This time (26) is

$$\binom{n-m}{k}\binom{n-m}{\ell}\Big[(a_{k+i}a_{\ell+j} - a_{\ell+i}a_{k+j}) + (a_{k+i}b_{\ell+j} - a_{k+j}b_{\ell+i})\Big],$$

which is non-negative by the log-concavity of the a's and (ii).

In these three cases, (26) is non-negative, and therefore $S_q \geq 0$ by the induction hypothesis. The fourth case is different, since (26) need not be non-negative:

(d) $h_\ell = 1, g_k = 0$. In this case, $g_t \equiv 0$ and $h_t \equiv 1$ for $\ell \leq t \leq k$, so

$$S_q = \sum_{r+s=p,\, |r-s|\leq 2q} \binom{n-m}{r}\binom{n-m}{s}(a_{r+i}b_{s+j} - a_{s+j}b_{r+i}).$$

Now use summation by parts, together with the fact that

$$u(r) = \binom{n-m}{r}\binom{n-m}{p-r}$$

is decreasing in r for $r \geq p/2$, to argue that it suffices to show that

$$(27) \qquad T_q = \sum_{r+s=p,\, r,s\geq 0,\, |r-s|\leq 2q} (a_{r+i}b_{s+j} - a_{s+j}b_{r+i}) \geq 0$$

for each q. Here is the argument: noting that

$$S_q - S_{q-1} = (T_q - T_{q-1})u\big(q + [p/2]\big)$$

and summing, we have

$$S_q = \sum_{r\leq q}(S_r - S_{r-1}) = \sum_{r\leq q}(T_r - T_{r-1})u\big(r + [p/2]\big)$$
$$= T_q u\big(q + [p/2]\big) + \sum_{r<q} T_r\big[u\big(r + [p/2]\big) - u\big(r+1+[p/2]\big)\big].$$

Now to check (27), we rearrange the sum as follows, noting that there may be some cancellation of terms, leading to the truncation of the upper limit of the sum:

$$\sum_{r=\ell+i}^{(k+i)\wedge(\ell+j-1)} (a_r b_{p-r+i+j} - a_{p-r+i+j}b_r).$$

Here again $k+\ell = p$, and $k-\ell = 2q$ if p is even and $k-\ell = 2q-1$ if p is odd. The summands in this sum are non-negative by (ii) since for all r appearing in the sum, $r \leq p - r + i + j$. To check this, write $2r = r + r \leq (k+i) + (\ell+j) = p+i+j$. $\square$

Corollary 6.5. *Suppose $\{a_k\}_{k=0}^n$ and $\{b_k\}_{k=0}^n$ are LC sequences that satisfy*

(i) *$\{\lambda a_k + b_k\}_{k=0}^n$ is LC for all $\lambda > 0$, and*

(ii) *$a_k b_{k+1} \geq a_{k+1}b_k$ for $0 \leq k < n$.*

Then μ is CNA+.

Proof. Let $\{w_i > 0\}_{i=1}^{n+1}$ be the external field that is applied to μ, and μ' be the resulting measure:

$$\mu'(S) = \text{constant} \times \prod_{i=1}^{n+1} w_i^{X_i(S)} \mu(S), \quad S \subseteq [n+1].$$

We need to show that (24) holds with μ replaced by μ'. First, note that the effect of w_{n+1} is to replace a_k by a constant multiple of a_k, so since we have introduced a general positive λ in assumption (i), we may assume that $w_{n+1} = 1$. The analog of (25) that we must check is

$$(28) \quad \sum_{i<j}\sum_{k,\ell}(f_j - f_i)c_ic_jd_kd_\ell\bigg[(h_k - h_\ell)a_{k+i}a_{\ell+j}$$

$$+ (g_k - g_\ell)b_{k+i}b_{\ell+j} + (h_k - g_\ell)a_{k+i}b_{\ell+j} + (g_k - h_\ell)a_{\ell+j}b_{k+i}\bigg] \geq 0,$$

where now the f_k's, g_k's and h_k's are obtained by symmetrizing F and G with respect to the field on the set of coordinates $\{X_1, \ldots, X_m\}$ in the case of F and $\{X_{m+1}, \ldots, X_n\}$ in the case of G. For example,

$$f_k = c_k^{-1} \cdot \sum F(S) \prod_{i=1}^{m} w_i^{X_i(S)},$$

where

$$c_k = \sum \prod_{i=1}^{m} w_i^{X_i(S)},$$

and the above sums are taken over all $S \subseteq [m]$ such that $\sum_{i=1}^{m} X_i(S) = k$. The formulas for g_k and h_k are exactly analogous, and the corresponding normalization factor (instead of c_k) is

$$d_k = \sum \prod_{i=m+1}^{n} w_i^{X_i(S)},$$

the sum being taken over all $S \subseteq \{m+1, \ldots, n\}$ such that $\sum_{i=m+1}^{n} X_i(S) = k$. Note that if $w_i \equiv 1$, then $c_k = \binom{m}{k}$ and $d_k = \binom{n-m}{k}$, so that f_k reduces to the symmetrization used in the proof of Theorem 6.4, and (28) reduces to (25).

   In order to deduce (28) from the arguments used to prove (25), we then need to check the following statements:
   (a) $f_k \leq f_{k+1}$, $g_k \leq g_{k+1}$, $h_k \leq h_{k+1}$ and $g_k \leq h_k$ for all k, and
   (b) $\{d_k\}_{k=0}^{n-m}$ is an LC sequence.
The latter is needed to check that $d_r d_{p-r}$ is decreasing in r for $r \geq p/2$, which is then used in the analog of the treatment of case (d) in the proof of Theorem 6.4. Note that statement (b) follows from e.g. Theorem 3.8, since d_k is just the k-th elementary symmetric function on $\{w_{m+1}, \ldots, w_n\}$.

   Now the fact that $g_k \leq h_k$ is immediate from the monotonicity of G in its last argument. The verification of the first three statements in (a) is the same for all of them, so we check (a) only for f_k. Letting as usual $|S|$ denote the cardinality of a set S, write

$$c_k c_{k+1}(f_{k+1} - f_k) = \sum_{|S|=k,\,|T|=k+1} [F(T) - F(S)] \prod_{i=1}^{m} w_i^{X_i(S)+X_i(T)}.$$

Consider the sum of the terms corresponding to S's and T's that satisfy $X_i(S) + X_i(T) = z_i$ for all i and an arbitrarily fixed $z \in \{0, 1, 2\}^m$. If A, B, C are the index sets corresponding to those coordinates X_i for which $z_i = 0, 1, 2$, respectively, we see that we need to check

$$\text{(29)} \qquad \sum_{\substack{|T|=k+1,\, X_i(T)=0,\, i \in A, \\ X_i(T)=1,\, i \in C}} F(T) \geq \sum_{\substack{|S|=k,\, X_i(S)=0,\, i \in A, \\ X_i(S)=1,\, i \in C}} F(S).$$

Noting that the sums above have the same number of summands, namely

$$\binom{2k - 2|C| + 1}{k - |C| + 1} \quad \text{and} \quad \binom{2k - 2|C| + 1}{k - |C|},$$

respectively, we see that (29) is a consequence of the monotonicity of F. $\square$

Recall from Remark 2.4 that CNA+ $\Rightarrow$ Rayleigh/h-NLC+ and CNA $\Rightarrow$ h-NLC. By combining Proposition 6.2 and Corollary 6.5 on the one hand, and Proposition 6.3 and Theorem 6.4 on the other hand, we deduce the following result.

Corollary 6.6. *If μ is an almost symmetric measure (i.e., its generating polynomial is symmetric in all but possibly one variable; cf. Definition 2.1), then*

(1) *μ is Rayleigh/h-NLC+ $\Longleftrightarrow$ μ is CNA+;*
(2) *μ is h-NLC $\Longleftrightarrow$ μ is CNA.*

This shows that Pemantle's Conjecture 2.6 is true for almost symmetric measures (recall that by the results in §4.2, Conjecture 2.6 also holds for strongly Rayleigh measures and PHR measures).

7. Negative results on negative dependence

We will now use the results from the previous sections, in particular those of §6, to construct counterexamples to the general cases of Pemantle's and Wagner's conjectures on ULC rank sequences (Conjectures 2.3–2.5) and Pemantle's conjecture on stochastic domination and truncations (Conjecture 2.9). We also answer in the negative some related problems as well as Problem 2.7 and [12, Problem 6].

Counterexample 1. Suppose that μ is a probability measure on $2^{[n+1]}$ whose generating polynomial

$$f(z_1, \ldots, z_{n+1}) = z_{n+1} \sum_{k=0}^{n} a_k e_k(z_1, \ldots, z_n) + \sum_{k=0}^{n} b_k e_k(z_1, \ldots, z_n)$$

is symmetric in its first n variables. Clearly,

$$\Delta(f)(t) = \sum_{k=0}^{n+1} \binom{n+1}{k} c_k t^k = \sum_{k=0}^{n+1} \left[\binom{n}{k-1} a_{k-1} + \binom{n}{k} b_k \right] t^k,$$

where

$$c_k = \frac{k a_{k-1} + (n + 1 - k) b_k}{n + 1}.$$

In light of Proposition 6.2 and Corollary 6.6, Conjectures 2.4 and 2.5 reduce to the following problem in the almost exchangeable case.

Problem 7.1. Let $\{a_k\}_{k=0}^m$, $\{b_k\}_{k=0}^m$ be two non-negative sequences satisfying

(1) $\{\theta a_k + (1 - \theta) b_k\}_{k=0}^m$ is LC for all $0 \leq \theta \leq 1$, and
(2) $a_k b_{k+1} \geq a_{k+1} b_k$ for $0 \leq k \leq m - 1$.

Is it true that the sequence

$$\{ka_{k-1} + (m + 1 - k)b_k\}_{k=0}^{m+1}, \quad \text{where } a_{-1} = b_{m+1} = 0,$$

is LC?

A related problem is the following.

Problem 7.2. Let $\{a_k\}_{k=0}^n$, $\{b_k\}_{k=0}^n$ be two non-negative sequences satisfying
  (1) $\{\theta a_k + (1 - \theta)b_k\}_{k=0}^n$ is LC for all $0 \le \theta \le 1$, and
  (2) $a_k b_{k+1} \ge a_{k+1} b_k$ for $0 \le k \le n - 1$.
Is it true that the sequence

$$\{ka_{k-1} + b_k\}_{k=0}^{n+1}, \quad \text{where } a_{-1} = b_{n+1} = 0,$$

is LC?

Note that a counterexample $\{a_k\}_{k=0}^n$, $\{b_k\}_{k=0}^n$ to the latter problem would actually give a counterexample to the former as well. This is because for some large integer $m \ge n$ the sequence $\{d_k\}_{k=0}^m$ defined by

$$(30) \qquad d_k = ka_{k-1} + \frac{m + 1 - k}{m}b_k, \quad \text{where } d_k = 0 \text{ if } k > n + 1,$$

will fail to be LC, and the sequences $\{a_k\}_{k=0}^m$ and $\{b_k/m\}_{k=0}^m$ will therefore constitute a counterexample to Problem 7.1.

It remains to find negative solutions to Problems 7.1 and 7.2. Set

$$a_0 = 2t^2, \quad a_1 = 2t, \quad a_2 = \frac{4}{9} - t, \quad a_3 = \frac{2}{3}, \quad a_4 = 1, \quad a_5 = \frac{2}{3}t,$$

$$b_0 = 9t^3, \quad b_1 = 9t^2, \quad b_2 = 3t, \quad b_3 = 1, \quad b_4 = 3, \quad b_5 = 9 - t.$$

For t sufficiently small and positive this is seen to be a negative solution to Problem 7.2. Indeed,

$$\lambda a_0 + b_0 = t^2(2\lambda + 9t), \quad \lambda a_1 + b_1 = t(2\lambda + 9t), \quad \lambda a_2 + b_2 = \left(\frac{4}{9} - t\right)\lambda + 3t,$$

$$\lambda a_3 + b_3 = \frac{2}{3}\lambda + 1, \quad \lambda a_4 + b_4 = \lambda + 3, \quad \lambda a_5 + b_5 = \frac{2}{3}t\lambda + 9 - t,$$

hence

$$(\lambda a_1 + b_1)^2 - (\lambda a_0 + b_0)(\lambda a_2 + b_2) = t^2(2\lambda + 9t)\left[\left(\frac{14}{9} + t\right)\lambda + 6t\right] \ge 0,$$

$$(\lambda a_2 + b_2)^2 - (\lambda a_1 + b_1)(\lambda a_3 + b_3) = \lambda^2\left[\left(\frac{4}{9} - t\right)^2 - \frac{4}{9}t\right] + 2t\lambda\left(\frac{1}{3} - 6t\right) \ge 0,$$

$$(\lambda a_3 + b_3)^2 - (\lambda a_2 + b_2)(\lambda a_4 + b_4) = t\lambda^2 + (1 - 9t) \ge 0,$$

$$(\lambda a_4 + b_4)^2 - (\lambda a_3 + b_3)(\lambda a_5 + b_5) = \left(1 - \frac{4}{9}t\right)\lambda^2 + t \ge 0$$

for any $\lambda \ge 0$ if e.g. $0 < t \le \frac{1}{20}$. Moreover,

$$a_0 b_1 - a_1 b_0 = 0, \quad a_1 b_2 - a_2 b_1 = t^2(2 + 9t) \ge 0, \quad a_2 b_3 - a_3 b_2 = \frac{4}{9} - 3t \ge 0,$$

$$a_3 b_4 - a_4 b_3 = 1, \quad a_4 b_5 - a_5 b_4 = 9 - 3t \ge 0$$

for $0 < t \le \frac{4}{27}$. However,

$$(b_4 + 4a_3)^2 - (b_3 + 3a_2)(b_5 + 5a_4) = -\frac{5}{9} + \frac{133}{3}t - 3t^2 < 0$$

for $0 < t \le \frac{1}{80}$. Thus, if $0 < t \le \frac{1}{80}$, then both sequences $\{a_k\}_{k=0}^5$ and $\{b_k\}_{k=0}^5$ are positive and satisfy conditions (1) and (2) in Problems 7.1 and 7.2. By taking e.g. $t = 10^{-4}$ one can then check that the corresponding sequences $\{a_k\}_{k=0}^m$ and $\{b_k/m\}_{k=0}^m$ constructed in (30) produce a counterexample to Problem 7.1 for any integer $m \ge 19$. Therefore, both Conjecture 2.4 and Conjecture 2.5 fail whenever the (total) number of variables satisfies $n + 1 \ge 20$.

We note that our construction disproves slightly more than Conjecture 2.4 and Conjecture 2.5. Indeed, the rank sequence of the measures obtained in this fashion (i.e., measures on $2^{[m+1]}$ with a generating polynomial of the form

$$z_{m+1} \sum_{k=0}^m a_k e_k(z_1, \ldots, z_m) + \sum_{k=0}^m \frac{b_k}{m} e_k(z_1, \ldots, z_m),$$

where $m \ge 19$ and $\{a_k\}_{k=0}^m$, $\{b_k\}_{k=0}^m$ are the sequences constructed above with $t = 10^{-4}$) is not even SLC, which is weaker than ULC (Definition 2.8).

In [12, Problem 6] it was asked whether Conjecture 2.4 would hold under the additional assumption that the Rayleigh measure μ is decreasing on $2^{[n+1]}$, i.e., the coefficients of its generating polynomial $\sum a_S z^S$ satisfy the monotonicity condition $a_S \ge a_T$ whenever $S \subseteq T \subseteq [n + 1]$. This fails as well, since if f is as in the above construction with $n + 1 \ge 20$, then the polynomial $f(\lambda z_1, \ldots, \lambda z_{n+1})$ is still Rayleigh and satisfies the additional monotonicity assumption stated above for all sufficiently small $\lambda \in \mathbb{R}_+$, but its rank sequence $\{\lambda^k r_k\}_{k=0}^{n+1}$ is not ULC. $\qquad\square$

Remark 7.1. In [72] it was shown that Conjectures 3.11 and 3.4 in *loc. cit.* are equivalent and would follow from either Conjecture 3.13 or Conjecture 3.14 in that same paper. Counterexample 1 above shows that all these conjectures fail.

Next we show that the above construction actually yields a counterexample to Conjecture 2.9 as well.

Counterexample 2. Let

$$f(z_1, \ldots, z_{n+1}) = z_{n+1} \sum_{j=0}^n a_j e_j(z_1, \ldots, z_n) + \sum_{j=0}^n b_j e_j(z_1, \ldots, z_n)$$

be the $n + 1 = 20$ variable polynomial in the previous counterexample and let μ be the corresponding probability measure in $\mathfrak{P}_{n+1}$, so that μ is CNA+. For $\mathcal{A} = \{S \in 2^{[n+1]} : n + 1 \in S\}$ we have

$$\mu_k(\mathcal{A}) = \frac{a_{k-1}\binom{n}{k-1}}{a_{k-1}\binom{n}{k-1} + b_k\binom{n}{k}}, \qquad \mu_{k+1}(\mathcal{A}) = \frac{a_k\binom{n}{k}}{a_k\binom{n}{k} + b_{k+1}\binom{n}{k+1}}.$$

Hence,

$$\mu_k(\mathcal{A}) \le \mu_{k+1}(\mathcal{A}) \iff a_{k-1}\binom{n}{k-1}b_{k+1}\binom{n}{k+1} \le a_k\binom{n}{k}b_k\binom{n}{k}.$$

This fails for $k = 3$ and $0 < t < \frac{1}{18}$. $\qquad\square$

As we noted in §2.5, a positive answer to Problem 2.7 would imply Conjecture 2.6. However, the answer to Problem 2.7 is negative, which we will now prove.

Counterexample 3. We claim that if $\mu \in \mathfrak{P}_m$ satisfies

 (1) μ is Rayleigh,
 (2) $\mu = \theta\mu_k + (1-\theta)\mu_{k+1}$ for some $0 < \theta < 1$,
 (3) $\mu_k \not\preccurlyeq \mu_{k+1}$,

where μ_k, μ_{k+1} are the truncations of μ as in Definition 2.15, then μ is not PHR (projection of a Rayleigh measure with homogeneous generating polynomial; cf. Definition 2.6). To prove this claim, let g_k and g_{k+1} be the generating polynomials for μ_k and μ_{k+1}, respectively, and suppose that

$$f(z_1, \ldots, z_\ell) = h_1(z_{m+1}, \ldots, z_\ell)g_k(z_1, \ldots, z_m) + h_2(z_{m+1}, \ldots, z_\ell)g_{k+1}(z_1, \ldots, z_m)$$

is a homogeneous Rayleigh polynomial that projects to $\theta g_k + (1-\theta)g_{k+1}$ and has a minimal number of variables $\ell \geq m+1$. If the same variable, say z_j, appears in both $h_1(z_{m+1}, \ldots, z_\ell)$ and $h_2(z_{m+1}, \ldots, z_\ell)$, then we can differentiate f with respect to z_j and get a new homogeneous polynomial in fewer variables that still projects to $\theta g_k + (1-\theta)g_{k+1}$. Moreover, if either $h_1(z_{m+1}, \ldots, z_\ell)$ or $h_2(z_{m+1}, \ldots, z_\ell)$ is the positive sum of several monomials, we could set a variable equal to zero and again get a polynomial in fewer variables with the desired properties. Hence, both $h_1(z_{m+1}, \ldots, z_\ell)$ and $h_2(z_{m+1}, \ldots, z_\ell)$ are monomials, and by the minimality of ℓ we have that $\ell = m + 1$ and

$$f(z_1, \ldots, z_{m+1}) = \theta z_{m+1} g_k(z_1, \ldots, z_m) + (1-\theta)g_{k+1}(z_1, \ldots, z_m).$$

Since f is a homogeneous Rayleigh polynomial, by Theorem 4.10 the corresponding measure $\mu_f \in \mathfrak{P}_{m+1}$ is (strongly) negatively associated. As in the proof of Proposition 4.12 we then deduce that $\mu_k \preccurlyeq \mu_{k+1}$, which contradicts assumption (3) above.

It remains to find a measure μ satisfying properties (1)–(3) above. If we let μ have as generating polynomial a constant multiple of

$$z_{n+1}(a_2 e_2(z_1, \ldots, z_n) + a_3 e_3(z_1, \ldots, z_n)) + b_3 e_3(z_1, \ldots, z_n) + b_4 e_4(z_1, \ldots, z_n),$$

where $m = n + 1$ and a_2, a_3, b_3, b_4 are as in Counterexample 1, we see that μ satisfies all of these three properties. Clearly, this counterexample fails to have a ULC rank sequence. Therefore, we conclude that not even the PHR property implies ULC. $\square$

Remark 7.2. We note that to disprove just the strongest version of Pemantle's conjecture on ULC rank sequences (Conjecture 2.3) we can simply invoke Example 2.2 in §2.1, which combined with Theorem 3.7 shows that Conjecture 2.3 actually fails even for symmetric measures. To give yet another counterexample to Conjecture 2.3 we can also employ a construction used in [55]. Indeed, the measure in $\mathfrak{P}_4$ considered in [55, Theorem 3.5] is seen to be NA but not ULC.

Note though that both examples mentioned above fail to be NLC (hence also CNA+, CNA, h-NLC+, and h-NLC), so they do not provide counterexamples to the weaker conjectures (Conjectures 2.4 and 2.5).

Acknowledgements

References

[1] M. Aissen, I. J. Schoenberg, A. Whitney, *On the generating functions of totally positive sequences I*, J. Analyse Math. **2** (1952), 93–103. MR0053174 (14:732d)

[2] E. D. Andjel, *A correlation inequality for the symmetric exclusion process*, Ann. Probab. **16** (1988), 717–721. MR929073 (89d:60192)

[3] M. F. Atiyah, R. Bott, L. Gårding, *Lacunas for hyperbolic differential operators with constant coefficients I*, Acta Math. **124** (1970), 109–189. MR0470499 (57:10252a)

[4] H. H. Bauschke, O. Güler, A. S. Lewis, H. S. Sendov, *Hyperbolic polynomials and convex analysis*, Canad. J. Math. **53** (2001), 470–488. MR1827817 (2002c:90099)

[5] J. Ben Hough, M. Krishnapur, Y. Peres, B. Virág, *Determinantal processes and independence*, Probab. Surv. **3** (2006), 206–229. MR2216966 (2006m:60068)

[6] J. Borcea, *Spectral order and isotonic differential operators of Laguerre-Pólya type*, Ark. Mat. **44** (2006), 211–240. MR2292718 (2008c:30010)

[7] J. Borcea, P. Brändén, *Pólya-Schur master theorems for circular domains and their boundaries*, to appear in Ann. of Math., http://annals.math.princeton.edu.

[8] J. Borcea, P. Brändén, *Applications of stable polynomials to mixed determinants: Johnson's conjectures, unimodality, and symmetrized Fischer products*, Duke Math. J. **143** (2008), 205–223. MR2420507

[9] J. Borcea, P. Brändén, *Multivariate Pólya-Schur classification problems in the Weyl algebra*, arXiv:math/0606360.

[10] J. Borcea, P. Brändén, *The Lee-Yang and Pólya-Schur programs I. Linear operators preserving stability*, arXiv:0809.0401.

[11] J. Borcea, P. Brändén, *The Lee-Yang and Pólya-Schur programs II. Theory of stable polynomials and applications*, preprint (2008).

[12] J. Borcea, P. Brändén, G. Csordas, V. Vinnikov, *Pólya-Schur-Lax problems: hyperbolicity and stability preservers*, http://www.aimath.org/pastworkshops/polyaschurlax.html.

[13] A. Borodin, A. Okounkov, G. Olshanski, *Asymptotics of Plancherel measures for symmetric groups*, J. Amer. Math. Soc. **13** (2000), 481–515. MR1758751 (2001g:05103)

[14] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, N. Linial, *The influence of variables in product spaces*, Israel J. Math. **77** (1992), 55–64. MR1194785 (94g:05091)

[15] P. Brändén, *Polynomials with the half-plane property and matroid theory*, Adv. Math. **216** (2007), 302–320. MR2353258 (2008h:05022)

[16] R. Burton, R. Pemantle, *Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances*, Ann. Probab. **21** (1993), 1329–1371. MR1235419 (94m:60019)

[17] D. Carlson, *Weakly sign-symmetric matrices and some determinantal inequalities*, Colloq. Math. **17** (1967), 123–129. MR0212041 (35:2916)

[18] S. Chaiken, *A combinatorial proof of the all minors matrix tree theorem*, SIAM J. Alg. Disc. Meth. **3** (1982), 319–329. MR666857 (83h:05062)

[19] Y. Choe, J. Oxley, A. Sokal, D. G. Wagner, *Homogeneous multivariate polynomials with the half-plane property*, Adv. Appl. Math. **32** (2004), 88–187. MR2037144 (2005d:05043)

[20] Y. Choe, D. G. Wagner, *Rayleigh matroids*, Combin. Prob. Comput. **15** (2006), 765–781. MR2248326 (2007f:05032)

[21] J. B. Conrey, *The Riemann hypothesis*, Notices Amer. Math. Soc. **50** (2003), 341–353. MR1954010 (2003j:11094)

[22] T. Craven, G. Csordas, *Jensen polynomials and the Túran and Laguerre inequalities*, Pacific J. Math **136** (1989), 241–260. MR978613 (90a:26035)

[23] D. J. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes. Springer, New York, 1988. MR950166 (90e:60060)

[24] D. Dubhashi, J. Jonasson, D. Ranjan, *Positive influence and negative dependence*, Combin. Probab. Comput. **16** (2007), 29–41. MR2286510 (2008h:62035)

[25] D. Dubhashi, V. Priebe, D. Ranjan, *Negative Dependence Through the FKG Inequality*, Technical Report RS-96-27, BRICS Report Series, Basic Research In Computer Science, Århus, Denmark, 1996; web version available at http://citeseer.ist.psu.edu/352490.html.

[26] D. Dubhashi, D. Ranjan, *Balls and bins: A study in negative dependence*, Random Struct. Algorithms **13** (1998), 99–124. MR1642566 (99k:60048)

[27] S. N. Ethier, T. G. Kurtz, Markov Processes: Characterization and Convergence. Wiley, 1986. MR838085 (88a:60130)

[28] S. M. Fallat, C. R. Johnson, *Determinantal Inequalities: Ancient History and Recent Advances*, Contemp. Math. **259** (2000), 199–212. MR1778502 (2001m:15019)

[29] T. Feder, M. Mihail, *Balanced matroids*, in "Proceedings of the 24th Annual ACM (STOC)", ACM Press, New York, 1992.

[30] C. M. Fortuin, P. W. Kasteleyn, J. Ginibre, *Correlation inequalities on some partially ordered sets*, Commun. Math. Phys. **22** (1971), 89–103. MR0309498 (46:8607)

[31] F. R. Gantmacher, M. G. Krein, *Oscillation matrices and kernels and small vibrations of mechanical systems*, Gostechizdat, 1950.

[32] J. H. Grace, *The zeros of a polynomial*, Proc. Cambridge Philos. Soc. **11** (1902), 352–357.

[33] G. Grimmett, The Random Cluster Model. Springer, 2006. MR2243761 (2007m:60295)

[34] L. Gurvits, *A proof of hyperbolic van der Waerden conjecture: the right generalization is the ultimate simplification*, arXiv:math/0504397.

[35] L. Gurvits, *Hyperbolic polynomials approach to van der Waerden/Schrijver-Valiant like Conjectures: sharper bounds, simpler proofs and algorithmic applications*, Proceedings of the 38th Annual ACM Symposium on Theory of Computing, ACM, New York, 2006, pp.417–426. MR2277167

[36] O. Güler, *Hyperbolic polynomials and interior point methods for convex programming*, Math. Oper. Res. **22** (1997), 350–377. MR1450796 (98d:90084)

[37] L. Gårding, *An inequality for hyperbolic polynomials*, J. Math. Mech. **8** (1959), 957–965. MR0113978 (22:4809)

[38] G. H. Hardy, J. E. Littlewood, G. Pólya, Inequalities. 2nd ed., Cambridge Univ. Press, Cambridge, United Kingdom, 1988. MR944909 (89d:26016)

[39] O. Holtz, *M-matrices satisfy Newton's inequalities*, Proc. Amer. Math. Soc. **133** (2005), 711–717. MR2113919 (2005k:15042)

[40] O. Holtz, H. Schneider, *Open problems on GKK τ-matrices*, Linear Algebra Appl. **345** (2002), 263–267. MR1883278

[41] O. Holtz, B. Sturmfels, *Hyperdeterminantal relations among symmetric principal minors*, J. Algebra **316** (2007), 634–648. MR2358606

[42] O. Häggström, *Random-cluster measures and uniform spanning trees*, Stoch. Proc. Appl. **59** (1995), 267–275. MR1357655 (97b:60170)

[43] L. Hörmander, Notions of Convexity. Progr. Math. **127**. Birkhäuser, Boston, MA, 1994. MR1301332 (95k:00002)

[44] G. James, C. Johnson, S. Pierce, *Generalized matrix function inequalities on M-matrices*, J. London Math. Soc (2) **57** (1998), 562–582. MR1659833 (2000b:15005)

[45] K. Joag-Dev, F. Proschan, *Negative association of random variables with applications*, Ann. Stat. **11** (1983), 286–295. MR684886 (85d:62058)

[46] K. Johansson, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, Ann. of Math. (2) **153** (2001), 259–296. MR1826414 (2002g:05188)

[47] K. Johansson, *Determinantal processes with number variance saturation*, Commun. Math. Phys. **252** (2004), 111-148. MR2103906 (2006b:82110)

[48] J. Kahn, M. Neiman, *Negative correlation and log-concavity*, arXiv:math/0712.3507.

[49] S. Karlin, Total Positivity. Vol. I. Stanford Univ. Press, Stanford, CA, 1968. MR0230102 (37:5667)

[50] B. Ja. Levin, Distribution of Zeros of Entire Functions. Transl. Math. Monogr. **5**, Amer. Math. Soc., Providence, R.I., 1980. MR589888 (81k:30011)

[51] E. H. Lieb, A. D. Sokal, *A general Lee-Yang theorem for one-component and multicomponent ferromagnets*, Commun. Math. Phys. **80** (1981), 153–179. MR623156 (83c:82008)

[52] T. M. Liggett, Interacting Particle Systems. Springer, 1985. MR776231 (86e:60089)

[53] T. M. Liggett, *Ultra log-concave sequences and negative dependence*, J. Combin. Theory Ser. A **79** (1997), 315–325. MR1462561 (98j:60018)

[54] T. M. Liggett, Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Springer, 1999. MR1717346 (2001g:60247)

[55] T. M. Liggett, *Negative correlations and particle systems*, Markov Proc. Rel. Fields **8** (2002), 547–564. MR1957219 (2004f:60207)

[56] T. M. Liggett, *Distributional limits for the symmetric exclusion process*, to appear in Stoch. Proc. Appl, arXiv:math/0710.3606.

[57] R. Lyons, *Determinantal probability measures*, Publ. Math. Inst. Hautes Études Sci. **98** (2003), 167–212. MR2031202 (2005b:60024)

[58] R. Lyons, Y. Peres, Probability on Trees and Networks. Book in progress, web version available at http://mypage.iu.edu/˜rdlyons/prbtree/prbtree.html.

[59] R. Lyons, J. Steif, *Stationary determinantal processes: phase multiplicity, Bernoullicity, entropy, and domination*, Duke Math. J. **120** (2003), 515–575. MR2030095 (2004k:60100)

[60] K. Markström, *Negative association does not imply log-concavity of the rank sequence*, J. Appl. Probab. **44** (2007), 1119–1121; updated version: `http://abel.math.umu.se/~klasm/`. MR2382951

[61] A. Marshall, I. Olkin, Inequalities: Theory of Majorization and Its Applications. Academic Press, New York, 1979. MR552278 (81b:00002)

[62] J. H. Mason, *Matroids: unimodal conjectures and Motzkin's theorem,* Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972), pp. 207–220. Inst. Math. Appl., Southend-on-Sea, 1972. MR0349445 (50:1939)

[63] C. Newman, *Normal fluctuations and the FKG inequalities*, Commun. Math. Phys. **74** (1980), 119–128. MR576267 (81i:82070)

[64] A. Okounkov, N. Reshetikhin, *Correlation function of Schur process with applications to local geometry of a random 3-dimensional Young diagram*, J. Amer. Math. Soc. **16** (2003), 581–603. MR1969205 (2004b:60033)

[65] R. Pemantle, *Towards a theory of negative dependence*, J. Math. Phys. **41** (2000), 1371–1390. MR1757964 (2001g:62039)

[66] Q. I. Rahman, G. Schmeisser, Analytic Theory of Polynomials. London Math. Soc. Monogr. (N. S.) **26**, Oxford Univ. Press, New York, NY, 2002. MR1954841 (2004b:30015)

[67] G.-C. Rota, D. Sharp, *Mathematics, Philosophy, and Artificial Intelligence: a Dialogue with Gian-Carlo Rota and David Sharp*, Los Alamos Science, Spring/Summer 1985.

[68] C. Semple, D. J. A. Welsh, *Negative correlation in graphs and matroids*, Combin. Prob. Comput. **15** (2006), 765–781.

[69] P. D. Seymour, D. J. A. Welsh, *Combinatorial applications of an inequality from statistical mechanics*, Math. Proc. Camb. Phil. Soc. **77** (1975), 485–495. MR0376378 (51:12554)

[70] A. D. Sokal, *The multivariate Tutte polynomial (alias Potts model) for graphs and matroids*, in "Surveys in Combinatorics, 2005" (B. S. Webb, ed.), Cambridge Univ. Press, Cambridge, United Kingdom, 2005. MR2187739 (2006k:05052)

[71] G. Szegö, *Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln algebraischer Gleichungen*, Math. Z. **13** (1922), 28–55. MR1544526

[72] D. G. Wagner, *Negatively correlated random variables and Mason's conjecture for independent sets in matroids*, Ann. Combin. **12** (2008), 211–239.

[73] D. G. Wagner, *Matroid inequalities from electrical network theory*, Electron. J. Combin. **11** (2005), A1 (17pp). MR2195432 (2006j:05041)

[74] J. L. Walsh, *On the location of the roots of certain types of polynomials*, Trans. Amer. Math. Soc. **24** (1922), 163–180. MR1501220

DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-106 91 STOCKHOLM, SWEDEN
E-mail address: `julius@math.su.se`

DEPARTMENT OF MATHEMATICS, ROYAL INSTITUTE OF TECHNOLOGY, SE-100 44 STOCKHOLM, SWEDEN
E-mail address: `pbranden@math.kth.se`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CALIFORNIA 90095-1555
E-mail address: `tml@math.ucla.edu`