COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **28** (2) (2008) 209–227
DOI: 10.1007/s00493-008-2249-5

HIGHER CORRELATION INEQUALITIES

SIDDHARTHA SAHI

Received September 17, 2005

We prove a correlation inequality for n increasing functions on a distributive lattice, which for $n=2$ reduces to a special case of the FKG inequality. The key new idea is to reformulate the inequalities for all n into a single positivity statement in the ring of formal power series. We also conjecture that our results hold in greater generality.

1. Introduction

The purpose of this paper is to prove a correlation inequality for n increasing functions on a distributive lattice. For finite lattices, it is possible to combine the entire family of inequalities (for all n) into a *single* statement involving the ring of formal power series. This is the form in which we first formulate and prove our result.

Thus, let X be a finite set and regard the power set 2^X as a partially ordered set with respect to inclusion. Let $\mathcal{R} := \mathbb{R}[[t]]$ be the space of formal power series in the variable t with real coefficients. The set

$$\mathcal{P} := \left\{a_1 t + a_2 t^2 + \cdots \in \mathcal{R} \mid a_i \geq 0 \text{ for all } i\right\}$$

is a convex cone in $\mathcal{R}$, and we define

$$\mathcal{R}[X] := \left\{F \mid F : 2^X \to \mathcal{R}\right\} \text{ and } \mathcal{P}[X] := \left\{F \mid F : 2^X \to \mathcal{P}\right\}.$$

Now $\mathcal{P}[X]$ is a convex cone in $\mathcal{R}[X]$ and we will refer to its elements as *positive* functions on 2^X. We define the subcone of *increasing* positive functions

Mathematics Subject Classification (2000): 05A20, 26D07, 60E15, 82B20 to be

$$\mathcal{I}[X] := \{F \in \mathcal{P}[X] : F(T) - F(S) \in \mathcal{P} \text{ for all } S \subseteq T \subseteq X\}.$$

To state our main result we need to consider a further subcone which we now introduce. For a function F in $\mathcal{R}[X]$ we define its "cumulation" to be the function F^+ in $\mathcal{R}[X]$ given by

$$F^+(T) := \sum_{S \subseteq T} F(S) \text{ for } T \subseteq X. \tag{1}$$

We write $\mathcal{C}[X]$ for the class of cumulations of *positive* functions; thus

$$\mathcal{C}[X] := \left\{F^+ \mid F \in \mathcal{P}[X]\right\}.$$

It is easy to see that $\mathcal{C}[X]$ is a subcone of $\mathcal{I}[X]$.

For each $x \in X$ we fix a real number m_x in $[0,1]$ and equip the power set 2^X with the product probability measure. Thus

$$\mu(S) := \prod_{x \in S} m_x \prod_{y \notin S} (1 - m_y) \text{ for } S \subseteq X. \tag{2}$$

Our main result is as follows:

Theorem 1. *For any F in $\mathcal{C}[X]$, and any measure μ as above, we have*

$$1 - \prod_{S \subseteq X} (1 - F(S))^{\mu(S)} \in \mathcal{P}. \tag{3}$$

Using this result, we can obtain an infinite family of correlation inequalities for ordinary (real valued) functions on 2^X. For this we define

$$R[X] := \left\{f \mid f : 2^X \to \mathbb{R}\right\} \text{ and } P[X] := \left\{f \mid f : 2^X \to \mathbb{R}_+\right\}.$$

Similarly, we define

$$I[X] := \{f \in P[X] \mid f(T) - f(S) \in \mathbb{R}_+ \text{ for all } S \subseteq T \subseteq X\},$$

and

$$C[X] := \left\{f^+ \mid f \in P[X]\right\}.$$

Note that $R[X]$ is a real vector space of dimension $2^{|X|}$, and the sets

$$C[X] \subseteq I[X] \subseteq P[X]$$

are polyhedral convex cones of dimension $2^{|X|}$.

The statement of our correlation inequalities involves a small amount of combinatorial notation. First of all, given functions $f_1,\dots,f_n$ in $R[X]$, and a subset τ of $\{1,\dots,n\}$, we define the corresponding correlation to be

$$E_\tau = E_\tau(f_1,\dots,f_n) = \mathbb{E}\left(\prod_{i\in\tau} f_i\right) := \sum_{S\subseteq X}\mu(S)\prod_{i\in\tau} f_i(S)\,. \tag{4}$$

Next, given a partition π of $\{1,\dots,n\}$ into disjoint subsets

$$\pi = \pi_1 \cup \cdots \cup \pi_l$$

we define the correlation product

$$E_\pi = \prod_{i=1}^{l} E_{\pi_i}.$$

Clearly if π is a partition of $\{1,\dots,n\}$, the cardinalities of the subsets constitute a partition $\lambda(\pi)$ of the integer n

$$\lambda(\pi) = (|\pi_1|,\dots,|\pi_l|)\,.$$

(We assume the indexing is chosen so that the cardinalities of the various subsets satisfy $|\pi_1|\geq\ldots\geq|\pi_l|$.) Now for a given partition λ of n, we define

$$E_\lambda = \sum_{\pi:\lambda(\pi)=\lambda} E_\pi. \tag{5}$$

For each partition λ of n, we introduce the coefficients

$$c_\lambda = (-1)^{l(\lambda)-1}\prod_{i=1}^{l(\lambda)}(\lambda_i - 1)! \tag{6}$$

where, as usual, $l(\lambda)$ denotes the number of parts of λ. Finally, we define

$$E_n = E_n(f_1,\dots,f_n) = \sum_{\lambda\vdash n} c_\lambda E_\lambda. \tag{7}$$

By construction, $E_n(f_1,\dots,f_n)$ is a linear function of each f_i and is a signed sum of products of correlations of the functions f_i with respect to the measure μ. Our main theorem implies the following result:

Theorem 2. *For any n-tuple of functions $f_1,\dots,f_n$ in $C[X]$, and for any product measure μ as in (2) we have*

$$E_n(f_1,\dots,f_n) \geq 0.$$

For $n=2$ we have two partitions, (2) and $(1,1)$, with coefficients

$$c_{(2)} = 1 \text{ and } c_{(1,1)} = -1.$$

Thus

$$E_2(f_1, f_2) = E_{(2)} - E_{(1,1)} = \mathbb{E}(f_1 f_2) - \mathbb{E}(f_1)\,\mathbb{E}(f_2)$$

and we get the following:

Corollary 3. *For f_1, f_2 in $C[X]$ we have*

$$\mathbb{E}(f_1 f_2) - \mathbb{E}(f_1)\,\mathbb{E}(f_2) \geq 0.$$

This is of course a special case of the Harris inequality which is known to hold for all increasing functions f_1, f_2 in $I[X]$. In turn the Harris inequality itself is a special case of the FKG inequality [3], which holds for all probability measures μ on 2^X satisfying

$$\mu(S \cup T)\,\mu(S \cap T) \geq \mu(S)\,\mu(T) \quad \text{for all } S, T \subseteq X. \tag{8}$$

(Of course if μ is as in (2) then it satisfies (8) with equality.)

The FKG inequality has proved extremely useful in several different areas of mathematics including mathematical physics, especially lattice models and percolation [7,5]; combinatorics, especially graph theory [4,2]; and statistics [10]. We refer the reader to [8] for a discussion of some of these applications, as well as an extensive bibliography.

It seems to be the case that our main result also holds in this generality. Thus we formulate the following conjecture:

Conjecture 4. For any F in $\mathcal{I}[X]$, and any μ satisfying (8), we have

$$1 - \prod_{S \subseteq X} (1 - F(S))^{\mu(S)} \in \mathcal{P}.$$

As above, this implies the following:

Conjecture 5. For any n-tuple of functions $f_1, \ldots, f_n$ in $I[X]$, and any μ satisfying (8) we have

$$E_n(f_1, \ldots, f_n) \geq 0.$$

We have been unable to establish these conjectures in general. The problem is that none of the usual inductive proofs of the FKG inequality seem to apply in this setting, and a slight natural strengthening of the conjecture turns out to be false. However we *have* managed to prove the conjectures for $|X| \leq 2$. For $|X| \leq 1$, Conjecture 4 coincides with Theorem 1; however the case $|X|=2$ already presents some novel features. We postpone further discussion of these matters to the appendix.

For $n=3,4,5$ we have

$$E_3 = 2E_{(3)} - E_{(2,1)} + E_{(1,1,1)},$$
$$E_4 = 6E_{(4)} - 2E_{(3,1)} - E_{(2,2)} + E_{(2,1,1)} - E_{(1,1,1,1)},$$
$$E_5 = 24E_{(5)} - 6E_{(4,1)} - 2E_{(3,2)} + 2E_{(3,1,1)} + E_{(2,2,1)} - E_{(2,1,1,1)} + E_{(1,1,1,1,1)}.$$

The positivity of $E_n(f_1,\ldots,f_n)$ for $n = 3,4,5$ was first observed by Richards in [8], who in the general case considered slightly different expressions of the form

$$\kappa'_n(f_1,\ldots,f_n) = \sum_{\lambda\vdash n} c'_\lambda E_\lambda$$

where

$$c'_\lambda = (-1)^{l(\lambda)-1}\left(l\left(\lambda'\right)-1\right)! \tag{9}$$

and λ' is the conjugate partition of λ.

Richards referred to the expressions κ'_n as "conjugate cumulants" since they are closely related to the usual cumulants, which are given by the formula

$$\kappa_n(f_1,\ldots,f_n) = \sum_{\lambda\vdash n} (-1)^{l(\lambda)-1}(l(\lambda)-1)!E_\lambda.$$

Indeed the κ'_n may be obtained from κ_n by formally switching the numerical coeffcient of E_λ with that of $E_{\lambda'}$, but keeping the same signs.

Now one has the identity

$$l\left(\lambda'\right) = \lambda_1$$

and moreover if λ is a partition of $n \leq 5$, then the second and subsequent parts of λ are either 2 or 1, and so

$$(\lambda_i - 1)! = 1 \quad \text{for all } i \geq 2.$$

Comparing formulas (6) and (9) we observe that we have

$$c_\lambda = c'_\lambda \quad \text{for all } \lambda \vdash n \leq 5.$$

Thus, by an amazing coincidence, one get

$$\kappa'_n = E_n \quad \text{for } n \leq 5.$$

However for $n=6$ we have

$$\kappa_6' \neq E_6$$

since the two functions differ in the coefficient of $E_{(3,3)}$. Indeed Richards observed that the inequality

$$\kappa_6' \geq 0$$

is not sharp, since one has

$$\kappa_6' (1,1,1,1,1,1) > 0,$$

and he asked about the possibility of modifying the definition of κ_n' to obtain a sharp inequality.

More precisely, Richards conjectured the existence of functionals of the form

$$P_n (f_1, \dots, f_n) = \sum_{\lambda \vdash n} c_\lambda E_\lambda$$

with some unknown coefficients c_λ, such that:

1. $P_n (f_1, \dots, f_n) \geq 0$ for all f_i in $I[X]$ and for all probability measures μ satisfying (8).
2. For each $n \geq 3$ there is a constant d_n such that
$$P_n (f_1, \dots f_{n-1}, 1) = d_n P_{n-1} (f_1, \dots f_{n-1}).$$
3. For all $n \geq 2$ one has
$$P_n (1, \dots, 1, 1) = 0.$$

It is now easy to prove the following result:

Theorem 6. *The functionals $E_n (f_1, \dots, f_n)$ satisfy conditions 2 and 3 with*

$$d_n = n - 2.$$

Moreover E_n satisfies condition 1 for all f_i in $C[X]$ and all μ of the form (2).

Of course, Conjecture 5 asserts precisely that E_n satisfies condition 1 in general.

Finally, we mention that there have been earlier generalizations of the FKG inequality. One of the most influential of these is the Ahlswede–Daykin 4-function theorem [1], which has been generalized in [9]. It would be interesting to find a generalization of this theorem along the lines of the present paper.

2. Power series

2.1. Difference operators

If X is a finite set, then for each x in X we define operators

$$\rho_x^+, \rho_x^-, \delta_x : \mathcal{R}[X] \to \mathcal{R}[X \setminus \{x\}].$$

Here ρ_x^+, ρ_x^- are the "restriction" operators defined by

$$\rho_x^+ F(U) = F(U \cup \{x\}), \quad \rho_x^- F(U) = F(U) \quad \text{for all } U \subset X \setminus \{x\}$$

and $\delta_x = \rho_x^+ - \rho_x^-$ is the "difference" operator given by

$$\delta_x F(U) = F(U \cup \{x\}) - F(U) \quad \text{for all } U \subset X \setminus \{x\}.$$

Lemma 7. *The operators $\rho_x^+, \rho_x^-, \delta_x$ map the cone $\mathcal{C}[X]$ to $\mathcal{C}[X \setminus \{x\}]$.*

Proof. If $G \in \mathcal{C}[X]$, then there exists F in $\mathcal{P}[X]$ such that

$$G(T) = F^+(T) = \sum_{S \subseteq T} F(S).$$

It now follows from the definition of the operators $\rho_x^+, \rho_x^-, \delta_x$ that

$$\rho_x^+ G = \left(\rho_x^+ F + \rho_x^- F\right)^+, \quad \rho_x^- G = \left(\rho_x^- F\right)^+ \quad \text{and} \quad \delta_x G = \left(\rho_x^+ F\right)^+.$$

Since $\rho_x^+ F$, $\rho_x^- F$ and $\rho_x^+ F + \rho_x^- F$ belong to $\mathcal{P}[X \setminus \{x\}]$, the result follows. ∎

Now observe that for $x \neq y$ the restriction operators $\rho_x^\pm$ commute with $\rho_y^\pm$. Hence we have

$$\delta_x \delta_y = \delta_y \delta_x.$$

Thus if T any subset of X, we can define the iterated difference operator

$$\delta_T := \prod_{x \in T} \delta_x : \mathcal{R}[X] \to \mathcal{R}[X \setminus S]$$

which is given by the explicit formula

$$\delta_T F(V) = \sum_{S \subseteq T} (-1)^{|T|-|S|} F(V \cup S) \quad \text{for all } V \subset X \setminus T.$$

The next key fact is that the iterated difference operators can be used to invert the cumulation operator $F \mapsto F^+$, defined in formula (1).

Lemma 8. *For $F \in \mathcal{R}[X]$, define $F^- \in \mathcal{R}[X]$ by the formula*

$$F^-(U) = \delta_U F(\emptyset).$$

Then we have

$$\left(F^-\right)^+ = \left(F^+\right)^- = F.$$

Proof. For $U \subseteq X$ we have

$$\left(F^-\right)^+(U) = \sum_{T\subseteq U} F^-(T) = \sum_{T\subseteq U} \delta_T F(\emptyset) = \sum_{T\subseteq U} \sum_{S\subseteq T} (-1)^{|T|-|S|} F(S).$$

Interchanging the order of summation, this becomes

$$\sum_{S\subseteq U} \sum_{\substack{T \\ S\subseteq T\subseteq U}} (-1)^{|T|-|S|} F(S) = \sum_{S\subseteq U} \left[\sum_{V\subseteq U\setminus S} (-1)^{|V|} \right] F(S). \tag{10}$$

Now for any finite set Z, the binomial formula implies the identity

$$\sum_{V\subseteq Z} a^{|V|} = (1+a)^{|Z|}.$$

Thus, for $S \subsetneq U$, the inner sum on the right in (10) is

$$(1-1)^{|U\setminus S|} = 0;$$

while for $S = U$, the inner sum is 1. Therefore we get

$$\left(F^-\right)^+(U) = F(U).$$

A similar calculation yields

$$\left(F^+\right)^-(U) = F(U),$$

and the result follows. ∎

As an immediate consequence we obtain the following:

Corollary 9. *For $F \in \mathcal{R}[X]$ we have*

$$F \in \mathcal{C}[X] \iff F^- \in \mathcal{P}[X].$$

Combining the previous results, we obtain an inductive characterization of the cone $\mathcal{C}[X]$.

Lemma 10. *For $F \in \mathcal{P}[X]$ we have*

$$F \in \mathcal{C}[X] \iff \delta_x F \in \mathcal{C}[X \setminus \{x\}] \quad \textit{for all } x \textit{ in } X. \tag{11}$$

Proof. First suppose F is in $\mathcal{C}[X]$. Then Lemma 7 implies $\delta_x F \in \mathcal{C}[X\setminus\{x\}]$.

Conversely, suppose $F\in\mathcal{P}[X]$ satisfies $\delta_x F\in\mathcal{C}[X\setminus\{x\}]$ for all x in X. To prove that $F\in\mathcal{C}[X]$, by the previous corollary, it suffices to show that

$$F^{-}(U)\in\mathcal{P}\quad\text{for all } U\subseteq X.$$

If U is the empty set, then since F is in $\mathcal{P}[X]$, we have

$$F^{-}(\emptyset)=[\delta_{\emptyset}F](\emptyset)=F(\emptyset)\in\mathcal{P}.$$

If U is not empty, we fix $x\in U$. Then since $\delta_x F\in\mathcal{C}[X\setminus\{x\}]$, we get

$$F^{-}(U)=[\delta_U F](\emptyset)=\left[\delta_{U\setminus\{x\}}[\delta_x F]\right](\emptyset)=[\delta_x F]^{-}(U\setminus\{x\})\in\mathcal{P}. \qquad \blacksquare$$

2.2. Positive functions

In this subsection we prove Theorem 1. The key ingredients are various properties of the cone $\mathcal{C}[X]$ of cumulative functions.

Since $\mathcal{C}[X]$ is a cone, it is closed under multiplication by positive scalars, and under (pointwise) addition. More generally, if we have an infinite sequence of functions $\{F_i\}_{i=1}^{\infty}\subseteq\mathcal{C}[X]$, such that for each $U\subseteq X$,

$$F(U)=\sum_{i=1}^{\infty}F_i(U)$$

is a well-defined element of $\mathcal{R}$, then F belongs to $\mathcal{C}[X]$.

Now since $\mathcal{R}$ is a ring, $\mathcal{R}[X]$ is also a ring under pointwise multiplication, and it is easy to see that the cones $\mathcal{P}[X]$ and $\mathcal{I}[X]$ are closed under multiplication. We now show that the same is true for the cone $\mathcal{C}[X]$.

Lemma 11. *For F,G in $\mathcal{C}[X]$ the product FG belongs to $\mathcal{C}[X]$.*

Proof. Clearly FG belongs to $\mathcal{P}[X]$, and so by the previous lemma, it suffices to prove that

$$\delta_x[FG]\in\mathcal{C}[X\setminus\{x\}]\quad\text{for all } x \text{ in } X.$$

For this we compute

$$\delta_x[FG]=\left[\rho_x^{+}F\right]\left[\rho_x^{+}G\right]-\left[\rho_x^{-}F\right]\left[\rho_x^{-}G\right]=\left[\rho_x^{+}F\right][\delta_x G]+[\delta_x F]\left[\rho_x^{-}G\right].$$

By induction on $|X|$ and Lemma 7, the RHS is in $\mathcal{C}[X\setminus\{x\}]$. $\blacksquare$

We can now prove Theorem 1.

Proof of Theorem 1. Let F be in $\mathcal{C}[X]$ and let μ be a product measure of the form (2). Then we have to prove

$$1-\prod_{S\subseteq X}(1-F(S))^{\mu(S)}\in\mathcal{P}.$$

We will proceed by induction on $|X|$. For $X=\emptyset$ the expression becomes

$$1-[1-F(\emptyset)]^{1}=F(\emptyset)$$

which belongs to $\mathcal{P}$ since F is in $\mathcal{P}[X]$.

Now suppose X is a non-empty set. Fix $x\in X$ and write $Y=X\setminus\{x\}$. Then we can rewrite the expression of the theorem in the form

$$1-\prod_{S\subseteq X}(1-F(S))^{\mu(S)}=1-\prod_{T\subseteq Y}(1-F_x(T))^{\nu(T)}$$

where ν is the product measure on Y given by

$$\nu(T)=\prod_{y\in T}m_y\prod_{w\in Y\setminus T}(1-m_w)$$

and F_x is the function in $\mathcal{R}[Y]$ defined by

$$1-F_x(T)=\left[1-\rho_x^{-}F(T)\right]^{1-m_x}\left[1-\rho_x^{+}F(T)\right]^{m_x}. \tag{12}$$

By the inductive hypothesis it suffices to prove that F_x is in $\mathcal{C}(Y)$. For this we rewrite the right side above in the form

$$\left[1-\rho_x^{-}F(T)\right]\left[\frac{1-\rho_x^{+}F(T)}{1-\rho_x^{-}F(T)}\right]^{m_x}=\left[1-\rho_x^{-}F(T)\right]\left[1-\frac{\delta_xF(T)}{1-\rho_x^{-}F(T)}\right]^{m_x}$$

and expand using the binomial theorem to get

$$F_x(T)=\rho_x^{-}F(T)+\sum_{i\geq 1}(-1)^{i-1}\binom{m_x}{i}\left[\delta_xF(T)\right]^{i}\left[1-\rho_x^{-}F(T)\right]^{1-i}.$$

Expanding further we obtain

$$F_x=\rho_x^{-}F+\sum_{i\geq 1}\sum_{j\geq 0}(-1)^{i-1}\binom{m_x}{i}(-1)^{j}\binom{1-i}{j}\left[\delta_xF\right]^{i}\left[\rho_x^{-}F\right]^{j}.$$

Now since $0\leq m_x\leq 1$ we have

$$(-1)^{i-1}\binom{m_x}{i}=\frac{m_x(1-m_x)(2-m_x)\cdots(i-1-m_x)}{i!}\geq 0.$$

Also since $i \geq 1$ we have

$$(-1)^j \binom{1-i}{j} = \frac{(i-1)(i)\cdots(i+j-1)}{j!} \geq 0.$$

Thus all the coefficients in the expression for F_x are positive, and by Lemmas 7 and 11 we see that F_x belongs to $\mathcal{C}(Y)$. ∎

3. Correlation inequalities

In this section we show how Theorem 1 can be used to derive the correlation inequalities of Theorem 2 for real valued functions on 2^X.

3.1. Polarization

We first show how the correlations $E_n(f_1, \ldots, f_n)$ arise from a "polarization" of the expression (3). For this we introduce the ring of formal power series in n variables

$$\mathcal{R}_n = \mathbb{R}[[t_1, \ldots, t_n]]$$

and for a finite set X, define the corresponding function ring on 2^X

$$\mathcal{R}_n[X] := \left\{F \mid F : 2^X \to \mathcal{R}_n\right\}.$$

Now if $f_1, \ldots, f_n$ are n (real-valued) functions in $\mathbb{R}[X]$, then we define

$$F_n := t_1 f_1 + \cdots + t_n f_n \in \mathcal{R}_n[X]. \tag{13}$$

Proposition 12. *The correlation $E_n(f_1, \ldots, f_n)$ of formula (7) is coefficient of $t_1 t_2 \cdots t_n$ in the expansion of*

$$1 - \prod_{S \subseteq X} (1 - F_n(S))^{\mu(S)}. \tag{14}$$

Proof. It suffices to show that $-E_n$ is the coefficient of $t_1 t_2 \cdots t_n$ in

$$\Pi = \prod_{S \subseteq X} (1 - F_n(S))^{\mu(S)}.$$

Taking log of both sides, we get

$$\log(\Pi) = \sum_{S \subseteq X} \mu(S) \log(1 - F_n(S)) = \mathbb{E}(\log(1 - F_n)) = -\sum_{k \geq 1} \frac{1}{k} \mathbb{E}\left(F_n^k\right).$$

Expanding using (13) and collecting terms of order ≤ 1 in each t_i we get

$$\log(\Pi) = -\sum_{k\geq 1}\frac{1}{k}\sum_{\{i_1,\ldots,i_k\}} k! t_{i_1}\cdots t_{i_k}\mathbb{E}(f_{i_1}\cdots f_{i_k}) + \cdots = \sum_{\tau} a_{|\tau|}E_\tau t^\tau + \cdots$$

where τ ranges over non-empty subsets of $\{1,\ldots,n\}$, $t^\tau$is the monomial $\prod_{i\in\tau} t_i$, E_τ is the correlation defined in (4) and

$$a_k := -(k-1)!\,.$$

Exponentiating, we get

$$\Pi = \exp\Big(\sum_\tau a_{|\tau|}E_\tau t^\tau + \cdots\Big) = \sum_{l\geq 0}\frac{1}{l!}\Big[\sum_\tau a_{|\tau|}E_\tau t^\tau\Big]^l + \cdots$$

where the ignored terms have order ≥ 2 in some t_i.

Therefore the coefficient of $t_1 t_2\cdots t_n$ in Π is

$$\sum_l \frac{1}{l!}\sum_{\pi\vdash[n],l(\pi)=l}\left[l! a_{|\pi_1|}\cdots a_{|\pi_l|}E_\pi\right] = \sum_{\pi\vdash[n]} a_{|\pi_1|}\cdots a_{|\pi_{l(\pi)}|}E_\pi$$

where the last sum ranges over all set partitions $\pi = \left(\pi_1,\cdots,\pi_{l(\pi)}\right)$ of $[n] := \{1,\ldots,n\}$ and $E_\pi = E_{\pi_1}\cdots E_{\pi_{l(\pi)}}$ as before. Since the coefficient of E_π depends only on $\lambda(\pi)$ we can rewrite this as

$$\sum_{\lambda\vdash n}(-1)^{l(\lambda)}\prod_{i=1}^{l(\lambda)}(\lambda_i - 1)!\sum_{\pi\vdash[n],\lambda(\pi)=\lambda} E_\pi = -\sum_{\lambda\vdash n} c_\lambda E_\lambda = -E_n,$$

recalling the definitions of c_λ (6) and E_λ (5). ∎

3.2. Specialization

We can now prove Theorem 2 by a "specialization" argument.

Proof of Theorem 2. Let us write $c(i_1,\ldots,i_n)$ for the coefficient of the monomial $t_1^{i_1}t_2^{i_2}\cdots t_n^{i_n}$ in the expansion of

$$1 - \prod_{S\subseteq X}\left(1 - F_n(S)\right)^{\mu(S)}.$$

In view of the previous proposition it suffices to prove that if $f_1,\ldots,f_n$ are in $C[X]$ then $c(1,\ldots,1)$ is non-negative.

We deduce this from Theorem 1 by a simple specialization "trick".

First, note that if we specialize

$$t_i = t^{k_i}, \quad i = 1, \ldots, n$$

for some positive integers $k_1, \ldots, k_n$, then F_n specializes to an element of $\mathcal{C}[X]$ and so by Theorem 1, the expression (14) specializes to an element of $\mathcal{P}$. Thus *after* the specialization, all coefficients of (14) are non-negative.

Now the coefficient of t^k in the specialized expression is

$$\sum_{(i_1,\ldots,i_n)} c(i_1, \ldots, i_n)$$

where the sum ranges over all n-tuples $(i_1, \ldots, i_n)$ of non-negative integers satisfying

$$i_1 k_1 + \cdots + i_n k_n = k.$$

Thus it suffices to show that we can choose $k_1, \ldots, k_n$ such that

$$i_1 k_1 + \cdots + i_n k_n = k_1 + \cdots + k_n \quad \Rightarrow \quad i_1 = \cdots = i_n = 1,$$

since then the corresponding sum, which is positive, reduces to the single term $c(1, \ldots, 1)$.

For this we choose distinct primes $p_1, \ldots, p_n$ and put

$$k_i = (p_1 \cdots p_n) / p_i.$$

Now consider the equation

$$i_1 k_1 + \cdots + i_n k_n = k_1 + \cdots + k_n.$$

Clearly no i_j could be 0, since otherwise the prime p_j would divide the left side, but not the right side. However if all the i_j's are ≥ 1, then they must all be equal to 1, since otherwise the left side would be strictly larger than the right. ∎

3.3. Appendix

The key to the inductive proof of Theorem 1 is the result that

$$F \in \mathcal{C}(X) \Rightarrow F_x \in \mathcal{C}(X \setminus \{x\}) \quad \text{for } x \in X,$$

where F_x is as defined in (12). We show by an example that this is false for $F \in \mathcal{I}(X)$.

Example 13. Let $X=\{1,2\}$, let μ be the uniform product measure so that $\mu(S)=1/4$ for all $S\subseteq X$, and let $F\in\mathcal{R}[\{1,2\}]$ be defined by

$$F(\emptyset)=0,\quad F(\{1\})=t^2,\quad F(\{2\})=t^2+t^3,\quad F(\{1,2\})=t^2+t^3.$$

Clearly we have $F\in\mathcal{I}[\{1,2\}]$. However using (12) we get

$$F_2(\{1\})-F_2(\emptyset)=(1-0)^{1/2}\left(1-t^2-t^3\right)^{1/2}-\left(1-t^2\right)^{1/2}\left(1-t^2-t^3\right)^{1/2}.$$

Computing the coefficient of t^5 in this expression we get

$$\left(1-\left(1-t^2\right)^{1/2}\right)\left(1-t^2-t^3\right)^{1/2}\sim\left(\frac{1}{2}t^2\right)\left(-\frac{1}{2}t^3\right)=-\frac{1}{4}t^5.$$

The negative sign shows that $F_2\notin\mathcal{I}[\{1\}]$.

Nevertheless we can prove the main result in this case. We start with the following result:

Lemma 14. *Let $X=\{1,2\}$ and let $F\in\mathcal{I}(X)$. Then there exist a,b,c,d,e in $\mathcal{P}$ such that*

$$F(\emptyset)=a,\;\; F(\{1\})=a+b+c,\;\; F(\{2\})=a+b+d,\;\; F(X)=a+b+c+d+e.$$

Proof. Arguing coefficient by coefficient, it suffices to prove the same result for real valued functions $F\in I(X)$, with $a,b,c,d,e\in\mathbb{R}_+$. In this case it is easy to verify that the following choices work:

$$a=F(\emptyset),\quad b=\min(F(\{1\})-a,F(\{2\})-a)$$
$$c=[F(\{1\})-a]-b,\quad d=[F(\{2\})-a]-b$$

and finally

$$e=F(X)-\max\left(F(\{1\}),F(\{2\})\right)=F(X)-a-b-c-d.$$ ∎

Proposition 15. *Conjecture 4 holds for $|X|\leq 2$.*

Proof. We need to show that for $F\in\mathcal{I}(X)$ and μ satisfying the FKG condition (8) we have

$$1-\prod_{S\subseteq X}(1-F(S))^{\mu(S)}\in\mathcal{P}.$$

For $|X|\leq 1$ we have $\mathcal{I}(X)=\mathcal{C}(X)$, and condition (8) holds for every measure μ. Thus in this case Conjecture 4 reduces to Theorem 1.

Now suppose $|X|=2$, say $X=\{1,2\}$. Let μ be a probability measure on X and write

$$\mu(\emptyset)=\delta, \quad \mu(\{1\})=\gamma, \quad \mu(\{2\})=\beta, \quad \mu(X)=\alpha.$$

Then we have

$$\alpha+\beta+\gamma+\delta=1 \quad \text{and} \quad \alpha\delta \geq \beta\gamma.$$

Next, let $F \in \mathcal{I}(X)$. By the previous lemma there exist a,b,c,d,e in $\mathcal{P}$ such that

$$F(\emptyset)=a, \;\; F(\{1\})=a+b+c, \;\; F(\{2\})=a+b+d, \;\; F(X)=a+b+c+d+e.$$

Thus we need to show that

$$1-\Pi \in \mathcal{P}$$

where

$$\Pi=(1-a)^{\delta}(1-a-b-c)^{\gamma}(1-a-b-d)^{\beta}(1-a-b-c-d-e)^{\alpha}.$$

It the next lemma we show that in the expansion of $1-\Pi$ as a formal power series in $\{a,b,c,d,e\}$, all coefficients are ≥ 0. Specializing a,b,c,d,e to elements of $\mathcal{P}$, the result follows. ∎

To complete the proof of the previous proposition we need to prove the following result:

Lemma 16. *Suppose $\alpha,\beta,\gamma,\delta$ are non-negative real numbers satisfying*

$$\alpha+\beta+\gamma+\delta=1 \quad \text{and} \quad \alpha\delta \geq \beta\gamma. \tag{15}$$

Write Π for the product

$$\Pi=(1-a)^{\delta}(1-a-b-c)^{\gamma}(1-a-b-d)^{\beta}(1-a-b-c-d-e)^{\alpha}$$

and consider its formal power series expansion in $\{a,b,c,d,e\}$

$$\Pi=1-\sum_{i+j+k+l+m\geq 1} c_{ijklm}a^i b^j c^k d^l e^m.$$

Then all the coefficients c_{ijklm} are ≥ 0.

Proof. We first show that we can reduce to the case where $a = 0$, i.e. to prove non-negativity of the coefficients c_{0jklm}. For this we rewrite the product above as

$$\Pi = (1-a)^{\alpha+\beta+\gamma+\delta} \left(1 - \frac{b+c}{1-a}\right)^{\gamma} \left(1 - \frac{b+d}{1-a}\right)^{\beta} \left(1 - \frac{b+c+d+e}{1-a}\right)^{\alpha}.$$

Since $\alpha+\beta+\gamma+\delta=1$, setting

$$b' = \frac{b}{1-a}, \quad \ldots, \quad e' = \frac{e}{1-a}$$

we get

$$\Pi = (1-a)\Pi'$$

where

$$\Pi' = \left(1 - b' - c'\right)^{\gamma} \left(1 - b' - d'\right)^{\beta} \left(1 - b' - c' - d' - e'\right)^{\alpha}.$$

Now Π' has the expansion

$$1 - \sum_{j+k+l+m\geq 1} c_{0jklm} \left(b'\right)^{j} \left(c'\right)^{k} \left(d'\right)^{l} \left(e'\right)^{m} = 1 - \sum_{j+k+l+m\geq 1} \frac{c_{0jklm} b^j c^k d^l e^m}{(1-a)^{j+k+l+m}}.$$

Therefore Π has the expansion

$$\Pi = (1-a)\Pi' = 1 - \left(a + \sum_{j+k+l+m\geq 1} c_{0jklm} \frac{b^j c^k d^l e^m}{(1-a)^{j+k+l+m-1}} \right).$$

Suppose that all c_{0jklm} are non-negative. Since

$$j+k+l+m-1 \geq 0$$

by the binomial theorem the expansion of $\frac{1}{(1-a)^{j+k+l+m-1}}$ has only positive terms. Thus further expanding the terms on the right, we see that all coefficients in the parethesis are ≥ 0.

Therefore, it suffice to prove the lemma with $a=0$, i.e to consider

$$\Pi' = (1-b-c)^{\gamma} (1-b-d)^{\beta} (1-b-c-d-e)^{\alpha}.$$

Using the previous strategy we now show how to reduce to the case where $b=0$. For this we rewrite

$$\Pi' = (1-b)^{\alpha+\beta+\gamma} \, \Pi''$$

where

$$\Pi'' = \left(1-c''\right)^{\gamma}\left(1-d''\right)^{\beta}\left(1-c''-d''-e''\right)^{\alpha} \quad \text{and} \quad c'' = \frac{c}{1-b}, \text{ etc.}$$

Since

$$\Pi'' = 1 - \sum_{k+l+m\geq 1} c_{00klm}\left(c''\right)^{k}\left(d''\right)^{l}\left(e''\right)^{m} = 1 - \sum_{k+l+m\geq 1} c_{00klm}\frac{c^{k}d^{l}e^{m}}{(1-b)^{k+l+m}}$$

we get

$$\Pi' = (1-b)^{\alpha+\beta+\gamma} - \left(\sum_{k+l+m\geq 1} c_{00klm}\frac{c^{k}d^{l}e^{m}}{(1-b)^{k+l+m-(\alpha+\beta+\gamma)}}\right).$$

Suppose all c_{00klm} are non-negative. Then since $\alpha+\beta+\gamma\leq 1$ we have

$$k+l+m-(\alpha+\beta+\gamma)\geq 0$$

and as before we conclude that, after further expansion in b, all the terms in the parenthetical sum are non-negative.

Also since $0\leq\alpha+\beta+\gamma\leq 1$, by the binomial theorem we get

$$(1-b)^{\alpha+\beta+\gamma} = 1 - (\text{non-negative terms})\,.$$

Thus we deduce that

$$\Pi' = 1 - (\text{non-negative terms})\,.$$

So finally it suffices to prove the lemma with $a=b=0$, i.e., to consider

$$\Pi'' = (1-c)^{\gamma}(1-d)^{\beta}(1-c-d-e)^{\alpha}\,.$$

We rewrite this as

$$\Pi'' = (1-c)^{\alpha+\gamma}(1-d)^{\alpha+\beta}\left(1-\frac{cd+e}{(1-c)(1-d)}\right)^{\alpha}.$$

Expanding the last expression by the binomial theorem we get

$$\Pi'' = (1-c)^{\alpha+\gamma}(1-d)^{\alpha+\beta} - \sum_{k\geq 1}(-1)^{k-1}\binom{\alpha}{k}\frac{(cd+e)^{k}}{(1-c)^{k-\alpha-\gamma}(1-d)^{k-\alpha-\beta}}.$$

Since we have $0 \le \alpha \le 1$ the coefficient $(-1)^{k-1}\binom{\alpha}{k}$ is non-negative. After further expansion in c and d we see that all the terms in the sum are non-negative. Thus, dropping the terms for $k \ge 2$, as well as the terms involving e for $k = 1$, it suffices to prove that the expression

$$\Pi''' = (1-c)^{\alpha+\gamma}(1-d)^{\alpha+\beta} - \alpha \frac{cd}{(1-c)^{1-\alpha-\gamma}(1-d)^{1-\alpha-\beta}}$$

is of the form

$$\Pi''' = 1 - (\text{non-negative terms})\,.$$

By the binomial expansion we can rewrite the first term in Π''' as

$$1 - (\text{non-negative terms}) + \sum_{k,l \ge 1} \left| \binom{\alpha+\gamma}{k} \binom{\alpha+\beta}{l} \right| c^k d^l$$

while the second term in Π'''can be rewritten as

$$-\alpha \sum_{k,l \ge 1} \left| \binom{\alpha+\gamma-1}{k-1} \binom{\alpha+\beta-1}{l-1} \right| c^k d^l.$$

Thus it suffices to prove the following inequality for binomial coefficients

$$\alpha \left| \binom{\alpha+\gamma-1}{k-1} \binom{\alpha+\beta-1}{l-1} \right| \ge \left| \binom{\alpha+\gamma}{k} \binom{\alpha+\beta}{l} \right| \quad \text{for } k,l \ge 1.$$

Dividing out by common factors, we are left with needing to show

$$\alpha \ge \frac{\alpha+\gamma}{k} \frac{\alpha+\beta}{l} \quad \text{for } k,l \ge 1.$$

It suffices to prove this for $k = l = 1$, and in this case using (15) we get

$$(\alpha+\gamma)(\alpha+\beta) = \alpha(\alpha+\beta+\gamma) + \beta\gamma \le \alpha(\alpha+\beta+\gamma) + \alpha\delta = \alpha. \qquad \blacksquare$$

References

[1] R. Ahlswede and D. E. Daykin: An inequality for the weights of two families of sets, their unions and intersections; *Z. Wahrsch. Verw. Gebiete* **43** (1978), 183–185.

[2] N. Alon and J. H. Spencer: *The Probabilistic Method*, John Wiley & Sons, Inc., New York, 1992.

[3] C. Fortuin, P. Kasteleyn and J. Ginibre: Correlation inequalities on some partially ordered sets, *Comm. Math. Phys.* **22** (1971), 89–103.

[4] R. L. Graham: Applications of the FKG inequality and its relatives, in: *Mathematical Programming: The State of the Art*, Springer, Berlin, 1983, pp. 115–131.

[5] J. Glimm and A. Jaffe: *Quantum Physics: A Functional Integral Point of View*, 2nd ed., Springer, New York, 1987.

[6] S. Karlin and Y. Rinott: A generalized Cauchy–Binet formula and applications to total positivity and majorization, *J. Multivar. Anal.* **27** (1988), 284–299.

[7] J. L. Lebowitz: Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems, *Comm. Math. Phys.* **28** (1972), 313–321.

[8] D. Richards: Algebraic method toward higher-order probability inequalities II, *Ann. of Prob.* **32** (2004), 1509–1544.

[9] Y. Rinott and M. Saks: Correlation inequalities and a conjecture for permanents, *Combinatorica* **13(3)** (1993), 269–277.

[10] L. A. Shepp: The XYZ conjecture and the FKG inequality, *Ann. of Prob.* **10** (1982), 824–827.

Siddhartha Sahi

Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
USA
sahi@math.rutgers.edu