Journal of Mathematical Physics

On the extension of the FKG inequality to *n* functions

Cite as: J. Math. Phys. **63**, 043301 (2022); https://doi.org/10.1063/5.0065285
Submitted: 31 July 2021 • Accepted: 10 March 2022 • Published Online: 04 April 2022

Elliott H. Lieb and Siddhartha Sahi

Journal of Mathematical Physics | ARTICLE | scitation.org/journal/jmp

On the extension of the FKG inequality to *n* functions

Cite as: J. Math. Phys. 63, 043301 (2022); doi: 10.1063/5.0065285
Submitted: 31 July 2021 • Accepted: 10 March 2022 •
Published Online: 4 April 2022

Elliott H. Lieb[1,a) and **Siddhartha Sahi**[2,b)]

AFFILIATIONS

[1] Departments of Mathematics and Physics, Jadwin Hall, Princeton University, Princeton, New Jersey 08544, USA
[2] Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road, Piscataway, New Jersey 08854, USA

Note: This paper is part of the Special Collection in Honor of Freeman Dyson.
[a)] **Email:** lieb@princeton.edu
[b)] **Author to whom correspondence should be addressed:** sahi@math.rutgers.edu

ABSTRACT

The 1971 Fortuin–Kasteleyn–Ginibre inequality for two monotone functions on a distributive lattice is well known and has seen many applications in statistical mechanics and other fields of mathematics. In 2008, one of us (Sahi) conjectured an extended version of this inequality for all $n > 2$ monotone functions on a distributive lattice. Here, we prove the conjecture for two special cases: for monotone functions on the unit square in $\mathbb{R}^k$ whose upper level sets are k-dimensional rectangles and, more significantly, for *arbitrary* monotone functions on the unit square in $\mathbb{R}^2$. The general case for $\mathbb{R}^k, k > 2$, remains open.

I. INTRODUCTION

For functions f, g on a probability space (L, μ), their expectation and correlation are defined by

$$E_1(f) = \mathscr{E}(f) := \int_L f d\mu \quad \text{and} \quad E_2(f,g) = \mathscr{E}(fg) - \mathscr{E}(f)\mathscr{E}(g). \tag{1.1}$$

Now suppose further that L is a distributive lattice[8] and that the probability measure μ satisfies

$$\mu(a \vee b)\mu(a \wedge b) \geq \mu(a)\mu(b). \tag{1.2}$$

In this situation, if f, g are positive monotone (decreasing) functions[9] on L, then one has

$$E_1(f) \geq 0 \quad \text{and} \quad E_2(f,g) \geq 0. \tag{1.3}$$

The first inequality is obvious, while the second is the celebrated FKG inequality of Fortuin–Kasteleyn–Ginibre[1] that plays an important role in several areas of mathematics/physics. We will refer to a distributive lattice L with probability measure μ satisfying (1.2) as an *FKG poset*.

In formulating (1.2), we have tacitly assumed that the poset L is a discrete set. However, the FKG inequality also has important continuous versions, which can be proved by a discrete approximation. For example, if $Q_k = [0,1]^k$ is the unit hypercube in $\mathbb{R}^k$ equipped with the partial order: $x \geq y$ if and only if $x_i \geq y_i$ for all i, then the FKG inequality holds for the Lebesgue measure and, more generally, for any absolutely continuous measure whose density function satisfies (1.2).

In Ref. 5, Sahi introduced a sequence of multilinear functionals $E_n(f_1, \ldots, f_n), n = 1, 2, 3, \ldots$, generalizing E_1 and E_2 (see Definition 3.1) and proposed the following conjecture:

Conjecture 1.1 (Ref. 5, Conjecture 5). If $f_1, \ldots, f_n$ are positive monotone functions on an FKG poset, then

$$E_n(f_1, \ldots, f_n) \geq 0. \tag{1.4}$$

Sahi[5] proved the conjecture for the lattice $\{0,1\} \times \{0,1\}$ and for a certain *subclass* of positive monotone functions on the general power set lattice $\{0,1\}^k$ equipped with a product measure. Since the functionals E_n satisfy the following "branching" property (Ref. 5, Theorem 6)

$$E_n(f_1, \ldots, f_{n-1}, 1) = (n-2)E_{n-1}(f_1, \ldots, f_{n-1}), \tag{1.5}$$

the inequalities (1.4) form a *hierarchy* in the following sense: if C_n denotes the n-function positivity conjecture, then C_n implies C_{n-1} for $n > 2$.

The work of Sahi was inspired by that of Richards[4] who first had the idea of generalizing the FKG inequality to more than two functions. A natural first candidate for such an inequality is the cumulant (Ursell function) κ_n, but an easy example shows that the inequality already fails for κ_3. Nevertheless, Richards [Ref. 4, Conjecture 2.5] conjectured the *existence* of such a hierarchy of inequalities, although without an explicit formula for E_n.

Indeed for $n = 3, 4, 5$, Sahi's functional E_n coincides with the "conjugate" cumulant κ'_n introduced by Richards [Ref. 4, formula (2.2)], although for $n \geq 6$ one has $E_n \neq \kappa'_n$. We note also that Ref. 4 contains two "proofs" of the positivity of $\kappa'_3, \kappa'_4, \kappa'_5$—one for a discrete lattice and the other for a continuous analog. However, it seems to us that both proofs have essential gaps. Thus, beyond the special cases treated in Ref. 5, Conjecture 1.1 remains a conjecture, even for $n = 3, 4, 5$.

In this paper, we provide further evidence in support of Conjecture 1.1. We consider the continuous case of the Lebesgue measure on the unit hypercube $Q_k = [0,1]^k$ in $\mathbb{R}^k$, and we prove the inequalities (1.4) for the following two additional cases:

- for arbitrary positive monotone functions on the unit square in $\mathbb{R}^2$ and
- for monotone characteristic functions of k-dimensional rectangles in $[0,1]^k$ and, by multlinearity of E_n, for functions whose level sets are (not necessarily homothetic) rectangles.

First, we treat the case of three functions on $\mathbb{R}^2$ in Sec. II. This introduces several key ideas, including a reduction to a non-linear inequality involving decreasing sequences. In Sec. III, we define E_n for arbitrary n and prove Conjecture 1.1, first for characteristic functions of k-dimensional rectangles and then for general monotone functions on $\mathbb{R}^2$, that is, we extend Sec. II to all $n > 3$. This requires additional ideas involving the symmetric group S_n and an intricate induction on n. Subsections III A and III B are written in complete generality, and we hope these ideas will help in the eventual resolution of Conjecture 1.1.

Since the FKG inequality has many applications in probability, combinatorics, statistics, and physics, it reasonable to suppose that the generalized inequality will likewise prove to be useful in one or more of these areas. Although we do not have a compelling application in mind, we feel that it is important to find such an application. Indeed, the right application might provide additional insight into Conjecture 1.1 and perhaps even suggest a line of attack.

To end this section, we tantalize the reader with an interesting reformulation of the inequalities $E_n \geq 0$ in terms of a formal power series from Ref. 5. First, if $F(x)$ is a positive function on a probability space L, then it is natural to define the geometric mean of F by

$$G(F) = \exp(\mathscr{E}(\log F)). \tag{1.6}$$

Now, suppose $F(x,t)$ is a power series of the form

$$F(x,t) = 1 - f_1(x)t - f_2(x)t^2 - \cdots. \tag{1.7}$$

Then, $\log(F(x,t))$ is a well defined power series, and formula (1.6) gives

$$G(F) = \exp(\mathscr{E}(\log(F(x,t))) = 1 - c_1 t - c_2 t^2 - \cdots, \tag{1.8}$$

where the constants c_j are certain algebraic expressions in various $\mathscr{E}(f_{i_1} f_{i_2} \cdots f_{i_p})$.

Conjecture 1.2 (Ref. 5, Conjecture 4). If the $f_1(x), f_2(x), \ldots$ is a sequence of positive monotone functions on an FKG poset, then $c_n \geq 0$ for all n.

It turns out that Conjectures 1.1 and 1.2 are *equivalent*. One implication has already been established in Sec. III of Ref. 5, and we prove the other direction in the Appendix. We also refer the reader to Refs. 6 and 7 for related inequalities in an algebraic setting.

II. THE INEQUALITY FOR THREE FUNCTIONS

For three functions, the multilinear functional E_n introduced in Ref. 5 is given by

$$E_3(f,g,h) = 2\mathscr{E}(fgh) + \mathscr{E}(f)\mathscr{E}(g)\mathscr{E}(h) - \mathscr{E}(f)\mathscr{E}(gh) - \mathscr{E}(g)\mathscr{E}(fh) - \mathscr{E}(h)\mathscr{E}(fg). \tag{2.1}$$

We note that E_3 is *different* from the cumulant (Ursell function), which is given by

$$\kappa_3(f,g,h) = \mathscr{E}(fgh) + 2\mathscr{E}(f)\mathscr{E}(g)\mathscr{E}(h) - \mathscr{E}(f)\mathscr{E}(gh) - \mathscr{E}(g)\mathscr{E}(fh) - \mathscr{E}(h)\mathscr{E}(fg). \tag{2.2}$$

We will consider the functional E_3 for functions on the unit hypercube,

$$Q_k = [0,1]^k = \{x = (x_1, \ldots, x_k) | \ 0 \leq x_i \leq 1\} \tag{2.3}$$

equipped with the Lebesgue measure and the usual partial order: $x \leq x'$ if and only if $x_i \leq x_i'$ for all i. We say that a real valued function f on Q_k is **monotone** (decreasing) if $x \leq x'$ implies $f(x) \geq f(x')$. We note that the FKG inequality is usually stated for monotonically increasing functions, but this is a somewhat arbitrary choice. Indeed, FKG and our theorems for decreasing functions are equivalent to the corresponding results for increasing functions. For a general FKG poset, this follows by reversing the partial order, and for Q_k, this follows by the change of variables $x_i \mapsto 1 - x_i$. We also note that monotonicity for Q_1 has the usual 1-variable meaning of a decreasing function.

Theorem 2.1. *If f, g, h are positive monotone functions on $[0,1]^2$, then $E_3(f,g,h) \geq 0$.*

The generalization of Theorem 2.1 to n functions is given in Theorem 3.6. We now reduce Theorem 2.1 to characteristic functions $\chi_S, S \subset Q_k$. These are defined by $\chi_S(x) = 1$ if $x \in S$ and $\chi_S(x) = 0$ if $x \notin S$. We will say S is **monotone** if χ_S is monotone.

Lemma 2.2. It suffices to prove Theorem 2.1 for χ_S, χ_T, χ_U for all monotone S, T, U.

Proof. Any positive f can be written as an integral over the characteristic functions of its upper level sets. Thus, $f(x) = \int_0^\infty \xi_s(x) \mathrm{d}s$, with $\xi_s(x) = 1$ if $f(x) > s$ and 0 otherwise (see the "layer cake principle" in Ref. 2). If f is monotone, then ξ_s is monotone for every s. Since E_3 is multi-linear in f, g, h, this reduces Theorem 2.1 to the case of monotone characteristic functions. ■

We now describe a further reduction of Theorem 2.1 to a discrete family of characteristic functions. Let $\mathscr{A} = \mathscr{A}(m)$ be the set of decreasing m-tuples of integers, each between 0 and m,

$$\mathscr{A}(m) := \{a \in \mathbb{Z}^m \mid m \geq a_1 \geq \cdots \geq a_m \geq 0\}. \tag{2.4}$$

For each $a \in \mathscr{A}$, we define a monotone subset S_a of $Q_2 = [0,1]^2$ as follows. Divide Q_2 uniformly into m^2 little squares, write $D_{i,j}$ for the square with top right vertex $(i/m, j/m)$, and set

$$S_a = \bigcup\nolimits_{j \leq a_i} D_{i,j}, \quad \chi_a = \chi_{S_a}. \tag{2.5}$$

Then, S_a is a monotone subset of Q_2, and, conversely, *any* monotone union of $D_{i,j}$ is of this form.

Lemma 2.3. It suffices to prove Theorem 2.1 for χ_a, χ_b, χ_c ; $a, b, c \in \mathscr{A}(m)$; for all m.

Proof. By Lemma 2.2, it suffices to consider monotone characteristic functions χ_S, χ_T, χ_U. Divide Q_2 uniformly into m^2 little squares $D_{i,j}$ as before, and let S^m, T^m, U^m be the unions of the $D_{i,j}$ contained in S, T, U, respectively; then, these are monotone subsets of Q_2 of the form (2.5). Moreover, $\chi_{S^m}, \chi_{S^m}\chi_{T^m}$, etc., converge to $\chi_S, \chi_S\chi_T$, etc., in L^1 as $m \to \infty$. Thus, if $E_3(\chi_{S^m}, \chi_{T^m}, \chi_{U^m}) \ge 0$, then we get $E_3(\chi_S, \chi_T, \chi_U) = \lim_{m\to\infty} E_3(\chi_{S^m}, \chi_{T^m}, \chi_{U^m}) \ge 0$. ∎

A. Proof of the three function inequality in two dimensions

We now prove Theorem 2.1 for χ_a, χ_b, χ_c, which suffices by Lemma 2.3. To simplify the notation, we work directly with a, b, c and we define the product ab, expectation $\mathscr{E}(a)$, etc., as follows:

$$(ab)_i = \min\{a_i, b_i\}, \tag{2.6}$$

$$E_1(a) = \mathscr{E}(a) = (a_1 + \cdots + a_m)/m^2, \tag{2.7}$$

$$E_2(a,b) = \mathscr{E}(ab) - \mathscr{E}(a)\mathscr{E}(b), \tag{2.8}$$

$$E_3(a,b,c) = 2\mathscr{E}(abc) + \mathscr{E}(a)\mathscr{E}(b)\mathscr{E}(c) - \mathscr{E}(a)\mathscr{E}(bc) - \mathscr{E}(b)\mathscr{E}(ac) - \mathscr{E}(c)\mathscr{E}(ab). \tag{2.9}$$

Then, we have $\chi_{ab} = \chi_a\chi_b$, $\mathscr{E}(a) = \mathscr{E}(\chi_a)$, $E_2(a,b) = E_2(\chi_a, \chi_b)$, $E_3(a,b,c) = E_3(\chi_a, \chi_b, \chi_c)$.

In particular, by the FKG inequality, we obtain the following lemma:

Lemma 2.4. For all a, b in $\mathscr{A}$, we have $E_2(a,b) \ge 0$. ∎

To study $E_3(a,b,c)$, we consider certain perturbations of a. We say that $a \in \mathscr{A}$ has a **descent** at i if $a_i > a_{i+1}$, and in this case, we can define three new sequences $a^- = a^{-,i}$, $a^+ = a^{+,i}$, $a^\star = a^{\star,i}$, *also* in $\mathscr{A}$, in which the following changes, and *only these*, are made to a:

$$a_i^- = a_{i+1}, \quad a_{i+1}^+ = a_i, \quad a_{i+1}^\star = a_{i+1} + 1. \tag{2.10}$$

Lemma 2.5. If a has a descent at i, but b does not, then we have $\mathscr{E}(a^+b) + \mathscr{E}(a^-b) = 2\mathscr{E}(ab)$.

Proof. Let $b_i = b_{i+1} = \beta$, say, then we have

$$(a^+b)_i = (a^+b)_{i+1} = \min\{a_i, \beta\} = (ab)_i, \tag{2.11}$$

$$(a^-b)_i = (a^-b)_{i+1} = \min\{a_{i+1}, \beta\} = (ab)_{i+1}. \tag{2.12}$$

Since the three sequences a^+b, a^-b, and ab coincide except at $i, i+1$, the result follows. ∎

Proposition 2.6. If a has a descent at i, but b and c do not, then

$$E_3(a^+,b,c) + E_3(a^-,b,c) = 2E_3(a,b,c). \tag{2.13}$$

Proof. Each term of (2.9) has a unique factor involving a, which is of the form $\mathscr{E}(ad)$, where $d = 1, b, c, bc$ is a sequence in $\mathscr{A}$ that does not have a descent at i. By Lemma 2.5, we get

$$\mathscr{E}(a^+d) + \mathscr{E}(a^-d) = 2\mathscr{E}(ad). \tag{2.14}$$

The result now follows from formula (2.9). ∎

Lemma 2.7. If a, b have a descent at i and $b_{i+1} \le a_{i+1}$, then $a^\star b = ab$.

Proof. Evidently, $(a^\star b)_j = (ab)_j$ for $j \ne i+1$, and since $b_{i+1} \le a_{i+1}$, we also have

$$(a^\star b)_{i+1} = b_{i+1} = (ab)_{i+1}. \tag{2.15}$$

Thus, we get $a^\star b = ab$, as claimed. ∎

Proposition 2.8. If a and b have a descent at i and $b_{i+1} \le a_{i+1}$, then we have $a^\star b = ab$ and

$$E_3(a^\star,b,c) \le E_3(a,b,c) \qquad \text{for all } c. \tag{2.16}$$

Proof. By Lemma 2.7, we get $\mathscr{E}(a^\star b) = \mathscr{E}(ab)$, $\mathscr{E}(a^\star bc) = \mathscr{E}(abc)$, and it follows that

$$E_3(a,b,c) - E_3(a^\star,b,c) = E_2(b,c)\big[\mathcal{E}(a^\star) - \mathcal{E}(a)\big] + \mathcal{E}(b)\big[\mathcal{E}(a^\star c) - \mathcal{E}(ac)\big]. \tag{2.17}$$

Evidently, we have $\mathcal{E}(a^\star) \geq \mathcal{E}(a)$ and $\mathcal{E}(a^\star c) \geq \mathcal{E}(ac)$, and by the FKG inequality, we also have $E_2(b,c) \geq 0$. Thus, all terms on the right-hand side of (2.17) are positive, which proves the result. ∎

Theorem 2.9. *For all a, b, c in $\mathcal{A}$, we have $E_3(a,b,c) \geq 0$.*

Proof. Let $\mathcal{U}$ be the set of triples (a,b,c) in $\mathcal{A}$ for which $E_3(a,b,c)$ attains its *minimum*, and let $\mathcal{V}$ be the subset of $\mathcal{U}$ for which the quantity $\mathcal{E}(a) + \mathcal{E}(b) + \mathcal{E}(c)$ attains its *maximum*.

We claim that if $(a,b,c) \in \mathcal{V}$, then a, b, c are constant sequences. If this is not the case, then a, say, has a descent at some i. If b, c do not have a descent at i, then by Proposition 2.6 we get

$$E_3(a,b,c) = \big(E_3(a^+,b,c) + E_3(a^-,b,c)\big)/2.$$

By minimality, $E_3(a,b,c) \leq E_3(a^\pm,b,c)$, which forces $E_3(a,b,c) = E_3(a^\pm,b,c)$. Replacing a by a^+, we reach a contradiction since $\mathcal{E}(a^+) > \mathcal{E}(a)$.

If b, say, also has a descent at i, then by symmetry we may assume $b_{i+1} \leq a_{i+1}$. Then, by Proposition 2.8, $E_3(a^\star,b,c) \leq E_3(a,b,c)$, and we again reach a contradiction since $\mathcal{E}(a^\star) > \mathcal{E}(a)$.

Now, we may assume a, b, c are constant sequences and, by symmetry, further assume that

$$a \equiv m\alpha, b \equiv m\beta, c \equiv m\gamma, \quad 0 \leq \alpha \leq \beta \leq \gamma \leq 1,$$

and it follows that $E_3(a,b,c) = 2\alpha + \alpha\beta\gamma - (\alpha\beta + \alpha\beta + \alpha\gamma) = \alpha(1-\beta)(2-\gamma) \geq 0$. ∎

This proves Theorem 2.1 for χ_a, χ_b, χ_c and, thus, by Lemma 2.3, in general.

III. THE INEQUALITY FOR *n* FUNCTIONS

A. The definition of E_n

In this subsection and Subsection III B, we work with arbitrary functions on a probability space. We start by recalling the definition of the multilinear functional $E_n(f_1,\ldots,f_n)$ from Ref. 5. This involves the decomposition of a permutation σ in the symmetric group S_n as a product of disjoint cycles,

$$\sigma = (i_1,\ldots,i_p)(j_1,\ldots,j_q)\cdots. \tag{3.1}$$

For σ as in (3.1), we write C_σ for the number of cycles in σ and we set

$$E_\sigma(f^1,\ldots,f^n) = \mathcal{E}(f^{i_1}\cdots f^{i_p})\mathcal{E}(f^{j_1}\cdots f^{j_q})\cdots. \tag{3.2}$$

Then, the following definition is due to Sahi.[5]

Definition 3.1. For functions $f^1,\ldots,f^n$ on a probability space X, we define

$$E_n(f^1,\ldots,f^n) = \textstyle\sum_{\sigma\in S_n}(-1)^{C_\sigma - 1}E_\sigma(f^1,\ldots,f^n). \tag{3.3}$$

Using (3.3), one can easily verify that E_1, E_2, E_3 coincide with their earlier definitions. We note that the factor of 2 in the term $2E(f^1f^2f^3)$ in formula (2.1) comes from the two 3-cycles (123) and (213). More generally, E_n will have repeated terms because E_σ is unchanged if we rearrange the indices within a cycle. For example, for $n = 4$, we have

$$\begin{aligned} E_4(f^1,f^2,f^3,f^4) = {} & 6\mathcal{E}(f^1f^2f^3f^4) - 2\big[\mathcal{E}(f^1)\mathcal{E}(f^2f^3f^4) + \mathcal{E}(f^2)\mathcal{E}(f^1f^3f^4) + \cdots\big] \\ & + \big[\mathcal{E}(f^1)\mathcal{E}(f^2)\mathcal{E}(f^3f^4) + \mathcal{E}(f^1)\mathcal{E}(f^3)\mathcal{E}(f^2f^4) + \cdots\big] \\ & - \big[\mathcal{E}(f^1f^2)\mathcal{E}(f^3f^4) + \mathcal{E}(f^1f^3)\mathcal{E}(f^2f^4) + \cdots\big] - \mathcal{E}(f^1)\mathcal{E}(f^2)\mathcal{E}(f^3)\mathcal{E}(f^4). \end{aligned}$$

We now give an explicit formula for E_n in a special case.

Lemma 3.2. Let $X = [0,1]$ be the unit interval equipped with Lebesgue measure, and let f^i be the characteristic function $\chi_{[0,a_i]}$, $0 \le a_i \le 1$, with $0 \le a_1 \le \cdots \le a_n \le 1$. Then, we have

$$E_n(f^1,\ldots,f^n) = a_1(1-a_2)\cdots(n-1-a_n).$$

We note that the above formula implies that E_n is positive, i.e., Conjecture 1.1 holds for the Lebesgue measure on $[0,1]$. While it is easy enough to give a direct proof the lemma, we prefer to postpone the proof to Subsection III B where we will derive it as a consequence of a more general result.

B. Algebraic properties of E_n

We first prove a recursive formula relating E_n to E_{n-1}.

Proposition 3.3. We have $E_n(f^1,\ldots,f^{n-1},f) = e_1 + \cdots + e_{n-1} - e_n$, where

$$e_i = \begin{cases} E_{n-1}(f^1,\ldots,f^i f,\ldots,f^{n-1}) & \text{if } 1 \le i \le n-1, \\ E_{n-1}(f^1,\ldots,f^{n-1})\,\mathcal{E}(f) & \text{if } i = n. \end{cases} \tag{3.4}$$

Proof. We write $f = f^n$ and consider expression (3.3) for $E_n(f^1,\ldots,f^n)$ as a sum over the symmetric group S_n. We decompose S_n as a disjoint union,

$$S_n = S^{(1)} \cup \cdots \cup S^{(n)}, \quad S^{(i)} = \{\sigma \in S_n \mid \sigma(i) = n\}. \tag{3.5}$$

Then, $S^{(n)}$ is a subgroup of S_n, naturally isomorphic to S_{n-1}. By (3.3), we have

$$E_n(f^1,\ldots,f^n) = \Sigma^1 + \cdots + \Sigma^n, \quad \Sigma^i = \textstyle\sum_{\sigma\in S^{(i)}} (-1)^{C_\sigma - 1} E_\sigma(f^1,\ldots,f^n). \tag{3.6}$$

To study Σ^i, we consider the map $\sigma \mapsto \overline{\sigma}$ defined by dropping n from the cycle decomposition of σ. Thus, for $n = 5$, we have $(13)(245) \mapsto (13)(24)$, $(12)(34)(5) \mapsto (12)(34)$, etc. Then, $\sigma \mapsto \overline{\sigma}$ defines a bijection from *each* $S^{(i)}$ to S_{n-1}. If σ is in $S^{(i)}$ and $i \neq n$, then i and n occur in the same cycle of σ, and dropping n does not change the cycle count. This gives

$$C_\sigma = C_{\overline{\sigma}}, \quad E_\sigma(f^1,\ldots,f^{n-1},f) = E_{\overline{\sigma}}(f^1,\ldots,f^i f,\ldots,f^{n-1}),$$

which implies $\Sigma^i = e_i$. If σ is in $S^{(n)}$, then (n) occurs as a separate cycle in σ and we get

$$C_\sigma = C_{\overline{\sigma}} + 1, \quad E_\sigma(f^1,\ldots,f^{n-1},f) = E_{\overline{\sigma}}(f^1,\ldots,f^{n-1})\,\mathcal{E}(f),$$

which gives $\Sigma^n = -e_n$. This proves the proposition. ■

Lemma 3.2 is now an easy consequence.

Proof of Lemma 3.2. Let $f_i = \chi_{[0,a_i]}$. Since $a_i \le a_n$ for all i, we get

$$f_i f_n = \chi_{[0,a_i]}\chi_{[0,a_n]} = \chi_{[0,a_n]} = f_i.$$

Now applying Proposition 3.3 with $f = f_n$, we deduce that

$$E_n(f^1,\ldots,f^n) = \big((n-1) - \mathcal{E}(f^n)\big)E_{n-1}(f^1,\ldots,f^{n-1}) = (n-1-a_n)E_{n-1}(f^1,\ldots,f^{n-1}).$$

The result follows by a straightforward induction on n. ■

We next establish a useful formula for the partial sum P_c of E_n over the set of permutations containing a fixed cycle c.

Proposition 3.4. Let S^c denote the set of permutations $\sigma \in S_n$ that contain a fixed cycle $c = (i_1, \ldots, i_p)$ and let $J_c = \{j_1, j_2, \ldots\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_p\}$, then we have

$$P_c := \sum_{\sigma \in S^c} (-1)^{C_\sigma - 1} E_\sigma(f^1, \ldots, f^n) = \begin{cases} -\mathscr{E}(f^{i_1} \cdots f^{i_p}) E_{n-p}(f^{j_1}, f^{j_2}, \ldots) & \text{if } p < n, \\ \mathscr{E}(f^1 \cdots f^n) & \text{if } p = n. \end{cases} \tag{3.7}$$

Proof. The set S^c consists of a single permutation if $p = n$. Otherwise, it consists of permutations of the form $\sigma = c \cdot \tau$, where τ is a permutation of J^c. Evidently, the number of cycles in σ and τ are related by $C_\tau = C_\sigma - 1$. Thus, in this case, we have

$$(-1)^{C_\sigma - 1} E_\sigma(f^1 \cdots f^n) = -\mathscr{E}(f^{i_1} \cdots f^{i_p})(-1)^{C_\tau - 1} E_\tau(f^{j_1}, f^{j_2}, \ldots). \tag{3.8}$$

Now, the result follows by summing (3.8) over τ. ■

C. Proof of the n function inequality for rectangles in any dimension

By a rectangle in dimension k or a k-rectangle, we mean a subset of $[0,1]^k$ of the form

$$[0, r_1] \times \cdots \times [0, r_k], \quad 0 \le r_1, \ldots, r_k \le 1.$$

Theorem 3.5. *If f^i are characteristic functions of k-rectangles, then $E_n(f^1, \ldots, f^n) \ge 0$.*

Proof. We proceed by induction on $k \ge 1$ and for a given k by induction on $n \ge 1$. The base cases $k = 1$ and $n = 1$ are straightforward and the former by Lemma 3.2. Thus, we may assume $k > 1$ and $n > 1$, and we can write

$$f^i = g^i \times \chi_{[0, a_i]},$$

where g^i is the characteristic function of a $(k-1)$-rectangle. By symmetry of E_n, we may assume

$$0 \le a_1 \le \cdots \le a_n \le 1. \tag{3.9}$$

We note that the assumption (3.9) on a_i means that we have

$$\mathscr{E}(f^{i_1} \cdots f^{i_p}) = a_l \mathscr{E}(g^{i_1} \cdots g^{i_p}), \quad l = \min\{i_1, \ldots, i_p\}. \tag{3.10}$$

Moreover, it follows from (3.3) and (3.10) that if $a_2 = \cdots = a_n = 1$, then we have

$$E_n(f^1, \ldots, f^n) = a_1 E_n(g^1, \ldots, g^n). \tag{3.11}$$

We now fix an index $i > 1$ and let $C(i)$ denote all set of all cycles containing i, then we have

$$E_n(f^1, \ldots, f^n) = \sum_{c \in C(i)} P_c,$$

where P_c is as in Proposition 3.4. If i is not minimal in c, then P_c is independent of a_i by (3.10). If i is minimal in c, then $1 \notin c$; hence, c has length $p < n$, and by (3.7) and (3.10), we get

$$P_c = -a_i b_c, \quad b_c = \mathscr{E}(g^{i_1} \cdots g^{i_p}) E_{n-p}(f^{j_1}, f^{j_2}, \ldots).$$

By induction on n, we have $b_c \geq 0$ for such c. This means that $E_n(f^1, \ldots, f^n)$ decreases as we increase $a_2, \ldots, a_n$ subject, of course, to condition (3.9). In particular, E_n decreases as we successively increase

$$a_n \nearrow 1, \quad a_{n-1} \nearrow 1, \quad \ldots, \quad a_2 \nearrow 1.$$

By (3.11), we get $E_n(f^1, \ldots, f^n) \geq a_1 E_n(g_1, \ldots, g_n)$, which is positive by induction on k. ■

If f is the characteristic function of a rectangle, then any level set of f is either the same rectangle or empty. However, using the layer-cake principle[2] and multilinearity as in the Proof of Lemma 2.2, we obtain the following immediate extension of the previous result.

Corollary 3.6. If $f^1, \ldots, f^n$ are positive, monotone functions whose level sets are (not necessarily homothetic) rectangles, then $E_n(f^1, \ldots, f^n) \geq 0$. ■

D. Proof of the *n* function inequality in two dimensions

Our main result is as follows:

Theorem 3.6. *If $f^1, \ldots, f^n$ are positive and monotone on $[0,1]^2$, then $E_n(f^1, \ldots, f^n) \geq 0$.*

As before, we can deduce this from the special case of χ_a as in (2.5).

Lemma 3.7. It suffices to prove Theorem 3.6 for $\chi_{a^1}, \ldots, \chi_{a^n}$, $a^i \in \mathscr{A}(m)$, for all m.

Proof. This is proved along the same lines as Lemmas 2.2 and 2.3. ■

In this section, we work with $\mathscr{A} = \mathscr{A}(m)$ and to simplify the notation for $a^1, \ldots, a^n$ in $\mathscr{A}$, we set

$$E_\sigma(a^1, \ldots, a^n) = \mathscr{E}(\chi_{a^{i_1}}, \ldots, \chi_{a^{i_p}}) \mathscr{E}(\chi_{a^{j_1}}, \ldots, \chi_{a^{j_q}}) \cdots, \tag{3.12}$$

$$E_n(a^1, \ldots, a^n) = \textstyle\sum_{\sigma \in S_n} (-1)^{C_\sigma - 1} E_\sigma(a^1, \ldots, a^n). \tag{3.13}$$

Then, we have $E_n(\chi_{a^1}, \ldots, \chi_{a^n}) = E_n(a^1, \ldots, a^n)$.

To study the positivity of E_n, we first consider a special case.

Proposition 3.8. If $a^i \equiv m\alpha_i$ are constant sequences with $1 \leq \alpha_1 \leq \cdots \leq \alpha_n \leq 0$, then

$$E_n(a^1, \ldots, a^n) = \alpha_1(1 - \alpha_2) \cdots (n - 1 - \alpha_n). \tag{3.14}$$

Proof. Let $L_n = E_n(a^1, \ldots, a^n)$. Since $\alpha_i \leq \alpha_n$, we have $a^i a^n = a^i$ for all i. Thus, we get

$$L_n = \textstyle\sum_{i=1}^{n-1} L_{n-1} - L_{n-1} \mathscr{E}(a^n) = (n - 1 - \alpha_n) L_{n-1}$$

by Proposition 3.3. Now, (3.14) follows by induction on n, the case $n = 1$ being obvious. ■

We now prove the generalization of Proposition 2.6.

Proposition 3.9. If a has a descent at i, but $a^1, \ldots, a^{n-1}$ do not, then we have

$$2E_n(a^1, \ldots, a^{n-1}, a) = E_n(a^1, \ldots, a^{n-1}, a^+) + E_n(a^1, \ldots, a^{n-1}, a^-). \tag{3.15}$$

Proof. This is proved for each term E_σ in (3.13) in exactly the same way as Proposition 2.6 by applying Lemma 2.5 to the unique factor of E_σ involving $a = a^n$ in (3.12). ■

We shall prove the next three theorems *together* by induction on n.

Theorem 3.10. *If $a^1, \ldots, a^{n-2}, b$ are in $\mathscr{A}$; S is a subset of Q_2; and $\chi_b \chi_S = 0$, then*

$$E_n(\chi_{a^1}, \ldots, \chi_{a^{n-2}}, \chi_b, \chi_S) \le 0. \tag{3.16}$$

Theorem 3.11. *If $a^1, \ldots, a^{n-2}, b, c$ are in $\mathscr{A}$; b, c have a descent at i; and $b_{i+1} \le c_{i+1}$, then*

$$E_n(a^1, \ldots, a^{n-2}, b, c^\star) \le E_n(a^1, \ldots, a^{n-2}, b, c). \tag{3.17}$$

Theorem 3.12. *For all $a^1, \ldots, a^n$ in $\mathscr{A}$, we have*

$$E_n(a^1, \ldots, a^n) \ge 0. \tag{3.18}$$

Proof. Let us write $A(n)$, $B(n)$, and $C(n)$ for the assertions of Theorems 3.10, 3.11, and 3.12. Then, $A(1), B(1)$ are vacuously true, while $C(1)$ is evident. Therefore, it suffices to prove the implications $A(n-1) \wedge C(n-1) \Longrightarrow A(n)$ and $A(n) \Longrightarrow B(n) \Longrightarrow C(n)$ for all $n \ge 2$.

$A(n-1) \wedge C(n-1) \Longrightarrow A(n)$: By assumption, we have $\chi_b \chi_S = 0$, and we also have $\chi_{a^i} \chi_S = \chi_{S^i}$, where $S^i = S \cap S_{a^i}$. Thus, by Proposition 3.3, we get

$$\begin{aligned} &E_n(\chi_{a^1}, \ldots, \chi_{a^{n-2}}, \chi_b, \chi_S) = e_1 + \cdots + e_{n-2} + e_{n-1} - e_n, \\ &\text{where } \; e_i := E_{n-1}(\chi_{a^1}, \ldots, \chi_{S^i}, \ldots, \chi_{a^{n-2}}, \chi_b), \quad i \le n-2, \\ &\qquad e_{n-1} := E_{n-1}(\chi_{a^1}, \ldots, \chi_{a^{n-2}}, 0), \\ &\qquad e_n := E_{n-1}(\chi_{a^1}, \ldots, \chi_{a^{n-2}}, \chi_b)\mathscr{E}(\chi_S). \end{aligned}$$

Now, $e_n \ge 0$ by $C(n-1)$ and $e_{n-1} = 0$ by (3.12) and (3.13). In addition, $\chi_b \chi_{S^i} = (\chi_b \chi_S)\chi_{a^i} = 0$, and so by symmetry, we can apply $A(n-1)$ to conclude $e_i \le 0$ for $i \le n-2$. This implies $A(n)$, (3.16).

$A(n) \Longrightarrow B(n)$: Define $S_c, S_{c^\star}$ as in (2.5) and put $S = S_{c^\star} \backslash S_c$, then by Lemma 2.7 we have

$$\chi_S \chi_b = (\chi_{c^\star} - \chi_c)\chi_b = \chi_{c^\star b} - \chi_{cb} = 0.$$

Thus, by $A(n)$, (3.16), we get $E_n(\chi_{a^1}, \ldots, \chi_{a^{n-2}}, \chi_b, \chi_{c^\star} - \chi_c) \le 0$, which implies $B(n)$, (3.17).

$B(n) \Longrightarrow C(n)$: This argument is similar to the Proof of Theorem 2.9. Let $\mathscr{M}$ be the set of n-tuples $\mathbf{a} = (a^1, \ldots, a^n)$ in $\mathscr{A}$ for which $E_n(\mathbf{a})$ achieves its *minimum*, and let $\mathscr{N}$ be the subset of $\mathscr{M}$ for which $\lambda(\boldsymbol{a}) = \mathscr{E}(a^1) + \cdots + \mathscr{E}(a^n)$ achieves its *maximum* on $\mathscr{M}$. We claim that for $\mathbf{a}$ in $\mathscr{N}$ each a^i is a constant sequence; by Proposition 3.8, this clearly implies $C(n)$, $E_n(\mathbf{a}) \ge 0$.

If the claim is not true, then one of the sequences has a descent at some i. First suppose that only one sequence, by symmetry $a^n = a$, has a descent at i. By Proposition 3.9 and minimality of $E_n(\mathbf{a})$, we deduce $E_n(\mathbf{a}) = E_n(a^1, \ldots, a^{n-1}, a^\pm)$. Thus, replacing a by a^+ preserves $E_n(\mathbf{a})$ but increases $\lambda(\mathbf{a})$, which is a contradiction. If two sequences have a descent at i, then by symmetry we may assume these are $a^{n-1} = b$, $a^n = c$ with $b_{i+1} \le c_{i+1}$. Now, $B(n)$, (3.17), implies that replacing c by $c^\star$ does not increase $E_n(\mathbf{a})$, but it does increase $\lambda(\mathbf{a})$, which is a contradiction. ■

This proves Theorem 3.6 for χ_{a^i} and, thus, by Lemma 3.7, in general.

ACKNOWLEDGMENTS

This work was partially supported by the NSF (Grant Nos. DMS-1939600 and DMS-2001537) and the Simons Foundation (Grant No. 509766). The hospitality of the Institute for Advanced Study is gratefully acknowledged.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX: THE EQUIVALENCE OF CONJECTURES 1.1 AND 1.2

We start by recalling some basic facts about partitions and permutations. For more background and details involving these ideas, we refer the reader to Ref. 3.

A partition λ of n, of length l, is a weakly decreasing sequence of positive integers,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l > 0 \quad \text{such that} \quad \lambda_1 + \cdots + \lambda_l = n,$$

we say that the λ_j are the *parts* of λ, and we write $l(\lambda) = l$ and $|\lambda| = n$.

The conjugation action of S_n permutes the indices in the cycle decomposition (3.1) of an element σ. Thus, the class of σ is uniquely determined by its "cycle type," i.e., the partition λ whose parts are the cycle lengths of σ, arranged in decreasing order. Moreover, if $m_i = m_i(\lambda)$ denotes the number of parts of size i, then the conjugacy class of cycle type λ contains $n!/z_\lambda$ elements, where

$$z_\lambda = \prod\nolimits_{i\geq 1} i^{m_i}(m_i). \tag{A1}$$

For a function f on a probability space, we define its *moments* by the formula

$$p_d(f) = \mathscr{E}(f^d) \quad \text{and} \quad p_\lambda(f) = p_{\lambda_1}(f)\cdots p_{\lambda_l}(f). \tag{A2}$$

Lemma A.1. We have $E_n(f,\ldots,f) = n!\sum_{|\lambda|=n}(-1)^{l(\lambda)-1} z_\lambda^{-1} p_\lambda(f)$.

Proof. If σ is of class λ, then the number of disjoint cycles in σ is $l(\lambda)$, and by (3.2), we have $E_\sigma(f,\ldots,f) = p_\lambda(f)$. Thus, the sum (3.3) for $E_n(f,\ldots,f)$ is constant over conjugacy classes, with class λ contributing $n!/z_\lambda$ identical terms. This implies the result. ∎

If f is as above and u is a parameter, then we can define the formal logarithm

$$\log(1-uf) = -\sum\nolimits_{i\geq 1} u^i f^i/i. \tag{A3}$$

Proposition A.2 We have $\exp(\mathscr{E}(\log(1-uf))) = 1 - \sum_{n\geq 1} u^n E_n(f,\ldots,f)/n!$

Proof. Let $Z = \mathscr{E}(\log(1-uf))$, then by (A3) we have

$$Z = -\sum\nolimits_{i\geq 1} u^i p_i(f)/i. \tag{A4}$$

Writing $p_k = p_k(f)$ and $p_\lambda = p_\lambda(f)$ for simplicity, we get

$$\exp(Z) = \prod\nolimits_{i\geq 1}\sum\nolimits_{m_i\geq 0}(-1)^{m_i}(u^i p_i)^{m_i}/i^{m_i} m_i! = \sum\nolimits_\lambda (-1)^{l(\lambda)} z_\lambda^{-1} p_\lambda u^{|\lambda|}. \tag{A5}$$

Now, the result follows from Lemma A.1. ∎

Proposition A.3. If $f_1, f_2, \ldots$ *are functions on a probability space, then we have*

$$1 - \exp\left(\mathscr{E}\left(\log\left(1 - \sum\nolimits_i f_i t^i\right)\right)\right) = \sum\nolimits_{n\geq 1}\sum\nolimits_{i_1,\ldots,i_n} E_n(f_{i_1},\ldots,f_{i_n})t^{i_1+\cdots+i_n}/n!.$$

Proof. Let us write $A = f_1 t + f_2 t^2 + \cdots$, then by Proposition A.2, we get

$$1 - \exp(\mathscr{E}(\log(1-A))) = \sum\nolimits_{n\geq 1} E_n(A,\ldots,A)/n!,$$

and by multilinearity of E_n, we have $E_n(A,\ldots,A) = \sum_{i_1,\ldots,i_n} E_n(f_{i_1},\ldots,f_{i_n})t^{i_1+\cdots i_n}$. ∎

Theorem A.4. *For a set of functions $\mathscr{I}$ on a probability space, the following are equivalent:*

1. *For all n, we have $E_n(f_1, \ldots, f_n) \geq 0$ if $f_1, \ldots, f_n \in \mathscr{I}$.*
2. *The power series $1 - \exp\big(\mathscr{E}(\log(1 - \sum_i f_i t^i))\big)$ has positive coefficients if $f_1, f_2, \ldots \in \mathscr{I}$.*

Proof. The first statement implies the second by Proposition A.3. The converse was proved in Ref. 5, but we recall it here for completeness. Let $p_1, p_2, \ldots, p_n$ be the first n primes; define

$$k = p_1 p_2 \cdots p_n, \quad k_j = k/p_j, \quad N = k_1 + \cdots + k_n,$$

and consider possible solutions of the equation $s_1 k_1 + \cdots + s_n k_n = N$, where $s_1, \ldots, s_n$ are integers ≥ 0. If some s_j were 0, then p_j would divide the left side but not the right; thus, we must have all $s_j > 0$ and, hence, that $s_1 = \cdots = s_n = 1$. Now, it follows from Proposition A.3 that the coefficient of t^N in the power series $1 - \exp\big(\mathscr{E}(\log(1 - \sum_{j=1}^n f_j t^{k_j}))\big)$ is precisely $E_n(f_1, \ldots, f_n)$. Thus, the second statement implies the first. ∎

The previous theorem proves the equivalence of Conjectures 1.1 and 1.2. In particular, our Theorem 3.6 implies Conjecture 1.2 for the Lebesgue measure on the unit square in $\mathbb{R}^2$.

REFERENCES

[1] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre, "Correlation inequalitiess on some partially ordered sets," Commun. Math. Phys. **22**(2), 89–103 (1971).

[2] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics Vol. 14 (American Mathematical Society, 2014), ISBN: 10: 0-8218-2783-9.

[3] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford Mathematical Monographs (Oxford Science Publications; The Clarendon Press; Oxford University Press, New York, 1995).

[4] D. St. P. Richards, "Algebraic methods toward higher order probability inequalities, II," Ann. Probab. **32**(2), 1509–1544 (2004).

[5] S. Sahi, "Higher correlation inequalities," Combinatorica **28**(2), 209–227 (2008).

[6] S. Sahi, "The FKG inequality for partially ordered algebras," J. Theor. Probab. **21**(2), 449–458 (2008).

[7] S. Sahi, "Correlation inequalities for partially ordered algebras," in *The Mathematics of Preference, Choice and Order*, Studies in Choice and Welfare (Springer, Berlin, 2009), pp. 361–369.

[8] A distributive lattice is a partially ordered set, closed under join (supremum) $\vee$ and meet (infimum) $\wedge$ such that each operation distributes over the other. A key example is the power set of a set, partially ordered by inclusion.

[9] In this paper, we use *positive* as a synonym for *non-negative* and *monotone* for *monotone decreasing*. By reversing the partial order, our results and conjectures hold equally for monotone increasing functions. We note further that the positivity requirement on functions is redundant for the second inequality but essential for the first.