Applications of the FKG Inequality and Its Relatives

R. L. Graham
Bell Laboratories, Department of Discrete Mathematics, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

Introduction

In 1971, C. M. Fortuin, P. W. Kasteleyn and J. Ginibre [FKG] published a remarkable inequality relating certain real functions defined on a finite distributive lattice. This inequality, now generally known as the FKG inequality, arose in connection with these authors' investigations into correlation properties of Ising ferromagnet spin systems and generalized earlier results of Griffiths [Gri] and Harris [Har] (who was studying percolation models). The FKG inequality in turn has stimulated further research in a number of directions, including a variety of interesting generalizations and applications, particularly to statistics, computer science and the theory of partially ordered sets. It turns out that special cases of the FKG inequality can be found in the literature of at least a half dozen different fields, and in some sense can be traced all the way back to work of Chebyshev.

In this paper, I will survey some of this history as well as the more recent extensions and applications. I will also discuss various open problems along the way which I hope will convince the reader that this exciting area is still fertile ground for further research.

Background

We begin with an old result of Chebyshev (see [HLP]) which asserts that if f and g are both increasing (or both decreasing) functions on [0, 1] then the average value of the product fg is at least as large as the product of the average values of f and g, where the average is taken with respect to some measure μ on [0, 1].

In symbols, this is just

$$(1)\qquad \int_0^1 fg\,d\mu \geqslant \int_0^1 f\,d\mu \int_0^1 g\,d\mu$$

In the case that μ is a discrete measure we can restate (1) as follows: If $f(k)$ and $g(k)$ are both increasing (or both decreasing) and $\mu(k) \geqslant 0$ for $k = 1, 2, 3, \ldots$, then

$$\frac{\sum_k f(k)g(k)\mu(k)}{\sum_k \mu(k)} \geqslant \frac{\sum_k f(k)\mu(k)}{\sum_k \mu(k)} \cdot \frac{\sum_k g(k)\mu(k)}{\sum_k \mu(k)}$$

i. e.,

(2) $$\sum_k f(k)g(k)\mu(k) \sum_k \mu(k) \geqslant \sum_k f(k)\mu(k) \sum_k g(k)\mu(k).$$

The proofs of (1) and (2) follow immediately by expanding the inequality

$$\sum_{i,j} (f(i)-f(j))(g(i)-g(j))\mu(i)\mu(j) \geqslant 0.$$

Basically, the FKG inequality represents a way of extending (2) to the situation in which the underlying index set is only *partially* ordered, as opposed to the *totally* ordered index set of integers occurring in (2). The setting is as follows. Let $(\Gamma, <)$ be a finite distributive lattice, i. e., Γ is a finite set, partially ordered by $<$, for which the two functions $\wedge$ (meet or greatest lower bound) and $\vee$ (join or least upper bound) defined by:

$$x \wedge y := \max\{z \in \Gamma : z \leqslant x, z \leqslant y\},$$
$$x \vee y := \min\{z \in \Gamma : z \geqslant x, z \geqslant y\}$$

are well-defined and satisfy the distributive laws:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

for all $x, y, z \in \Gamma$.

It is well known that any such lattice can be realized as a sublattice of the lattice of all subsets of some finite set partially ordered by inclusion and with $x \wedge y = x \cap y$ and $x \vee y = x \cup y$.

We now suppose $\mu : \Gamma \to \mathbb{R}_0$, the nonnegative reals, satisfies

(*) $$\mu(x \wedge y)\mu(x \vee y) \geqslant \mu(x)\mu(y) \quad \text{for all } x, y \in \Gamma.$$

For reasons we shall mention later, a function μ satisfying (*) is often called *log supermodular*. Finally, a function $f: \Gamma \to \mathbb{R}$ is called *increasing* if

$$x \leqslant y \Rightarrow f(x) \leqslant f(g) \quad \text{for } x, y \in \Gamma$$

(with *decreasing* defined similarly).

The FKG inequality states:

If f and g are both increasing (or both decreasing) real functions on a finite distributive lattice Γ and $\mu : \Gamma \to \mathbb{R}_0$ is log supermodular then

(3) $$\sum_{x \in \Gamma} f(x)g(x)\mu(x) \sum_{x \in \Gamma} \mu(x) \geqslant \sum_{x \in \Gamma} f(x)\mu(x) \sum_{x \in \Gamma} g(x)\mu(x).$$

The original proof of (3) was somewhat complicated [FKG]. Several years after (3) appeared, Holley found the following beautiful generalization:

Theorem (Holley [Hol])

Suppose $\alpha, \beta: \Gamma \to \mathbb{R}_0$ satisfy

$$\alpha(x \wedge y)\beta(x \vee y) \geqslant \alpha(x)\beta(y) \quad \text{for all } x, y \in \Gamma.$$

Then for any increasing function $\theta: \Gamma \to \mathbb{R}_0$

(4) $$\sum_{x \in \Gamma} \alpha(x)\beta(x) \geqslant \sum_{x \in \Gamma} \beta(x)\theta(x).$$

However, this result was itself soon superseded by a striking result of Ahlswede and Daykin [AD 1] which we now describe.

The Ahlswede-Daykin Inequality

To state the inequality of Ahlswede and Daykin, we need the following simplifying notation. For subsets X and Y of Γ, define

$$X \wedge Y = \{x \wedge y : x \in X, y \in Y\},$$
$$X \vee Y = \{x \vee y : x \in X, y \in Y\}$$

and, for a function of $f: \Gamma \to \mathbb{R}$, define

$$f(X) = \sum_{x \in X} f(x).$$

As before, Γ denotes a finite distributive lattice.

Theorem [AD]

Suppose $\alpha, \beta, \gamma, \delta: \Gamma \to \mathbb{R}_0$ satisfy

(5) $$\alpha(x)\beta(y) \leqslant \gamma(x \vee y)\delta(x \wedge y) \quad \text{for all } x, y \in \Gamma.$$

Then

(5′) $$\alpha(X)\beta(Y) \leqslant \gamma(X \vee Y)\delta(X \wedge Y) \quad \text{for all } X, Y \subseteq \Gamma.$$

Note the attractive similarity between the hypothesis (5) and the conclusion (5′). This certainly contributes to the relative simplicity of the proof of the theorem, which we now give (also, see [Kle 2]).

Proof: It follows from our previous remarks that it suffices to prove the theorem for the case that $\Gamma = 2^{[N]}$ is the lattice of subsets of $[N] = \{1, 2, \dots, N\}$ partially ordered by inclusion. In this case, the hypothesis is

(6) $$\alpha(x)\beta(y) \leqslant \gamma(x \cup y)\delta(x \cap y) \quad \text{for all } x, y \in 2^{[N]}$$

and the desired conclusion is

(6′) $$\alpha(X)\beta(Y) \leqslant \gamma(X \vee Y)\delta(X \wedge Y) \quad \text{for all } X, Y \subseteq 2^{[N]}.$$

The proof proceeds by induction on N. We first consider the case $N=1$, in which case $2^{[N]}=2^{[1]}=\{\phi, \{1\}\}$. For $\sigma=\alpha, \beta, \gamma$ or δ, let σ_0 denote $\sigma(\phi)$ and σ_1 denote $\sigma(\{1\})$. Then (6) becomes

$$\begin{aligned} &\alpha_0\beta_0 \leqslant \gamma_0\delta_0,\\ &\alpha_1\beta_0 \leqslant \gamma_1\delta_0,\\ &\alpha_0\beta_1 \leqslant \gamma_1\delta_0,\\ &\alpha_1\beta_1 \leqslant \gamma_1\delta_1. \end{aligned} \tag{7}$$

It is easy to check that (6′) holds if either X or Y consists of a single element. This leaves only the case $X=\{\phi, \{1\}\}=Y$ to deal with. In this case, the inequality we must prove is

$$(\alpha_0+\alpha_1)(\beta_0+\beta_1) \leqslant (\gamma_0+\gamma_1)(\delta_0+\delta_1). \tag{7′}$$

Note that (7′) would follow at once from (7) if one of the occurrences of $\gamma_1\delta_0$ in (7′) were $\gamma_0\delta_1$ instead (by summing). As it is we have to work slightly harder. If any of $\alpha_0, \beta_0, \gamma_0, \delta_0$ is zero then (7′) follows at once. It follows from this (and a little computation) that it is enough to consider the special case $\alpha_0=\beta_0=\gamma_0=\delta_0=1$. Now (7) becomes

$$\alpha_1 \leqslant \gamma_1, \quad \beta_1 \leqslant \gamma_1, \quad \alpha_1\beta_1 \leqslant \gamma_1\delta_1 \tag{8}$$

and (7′) becomes

$$(1+\alpha_1)(1+\beta_1) \leqslant (1+\gamma_1)(1+\delta_1). \tag{8′}$$

Again, (8′) is immediate if $\gamma_1=0$ so we may assume $\gamma_1>0$. Since (8′) becomes harder to satisfy as δ_1 decreases, it suffices to prove (8′) when δ_1 is as small as possible, i.e., (by (8)) when $\delta_1=\alpha_1\beta_1/\gamma_1$. In this case (8′) becomes

$$(1+\alpha_1)(1+\beta_1) \leqslant (1+\gamma_1)\left(1+\frac{\alpha_1\beta_1}{\gamma_1}\right)$$

i.e.,

$$\alpha_1+\beta_1 \leqslant \gamma_1 + \frac{\alpha_1\beta_1}{\gamma_1}.$$

However, this last inequality is an immediate consequence of

$$(\gamma_1-\alpha_1)(\gamma_1-\beta_1) \geqslant 0$$

which is implied by (8). This proves the result for the case $N=1$.

Assume now that the assertion holds for $N=n-1$ for some $n \geqslant 2$. Suppose $\alpha, \beta, \gamma, \delta$: $2^{[n]} \to \mathbb{R}_0$ satisfy the hypothesis (6) with $N=n$ and let $X, Y \subseteq 2^{[n]}$ be given. We will define new functions $\alpha', \beta', \gamma', \delta'$ mapping $2^{[n-1]} := T'$ into $\mathbb{R}_0$ as follows:

$$\alpha'(x') = \sum_{\substack{x \in X \\ x' = x \setminus \{n\}}} \alpha(x),$$

$$\beta'(y') = \sum_{\substack{y \in Y \\ y' = y \setminus \{n\}}} \beta(y),$$

$$\gamma'(z') = \sum_{\substack{z \in X \cup Y \\ z' = z \setminus \{n\}}} \gamma(z),$$

$$\delta'(w') = \sum_{\substack{w \in X \cap Y \\ w' = w \setminus \{n\}}} \delta(w).$$

Thus, for $x' \in T'$,

$$\alpha'(x') = \begin{cases} \alpha(x') + \alpha(x' \cup \{n\}) & \text{if } x' \in X,\ x' \cup \{n\} \in X, \\ \alpha(x') & \text{if } x' \in X,\ x' \cup \{n\} \notin X, \\ \alpha(x' \cup \{n\}) & \text{if } x' \notin X,\ x' \cup \{n\} \in X, \\ 0 & \text{if } x' \notin X,\ x' \cup \{n\} \notin X. \end{cases}$$

Observe that with these definitions

$$\alpha(X) = \sum_{x \in X} \alpha(x') = \sum_{x' \in T'} \alpha'(x') = \alpha'(T')$$

and, similarly,

$$\beta(Y) = \beta'(T'), \qquad \gamma(X \vee Y) = \gamma'(T'), \qquad \delta(X \wedge Y) = \delta'(T').$$

Therefore, *if*

(10) $$\alpha'(x')\beta'(y') \leqslant \gamma'(x' \cup y')\delta'(x' \cap y') \quad \text{for all } x', y' \in T'$$

holds, then by the induction hypotheses we would have

$$\alpha(X)\beta(Y) = \alpha'(T')\beta'(T') \leqslant \gamma'(T')\delta'(T') = \gamma(X \vee Y)\delta(X \wedge Y)$$

since $T' \vee T' = T'$, $T' \wedge T' = T'$, which is just (6′), the desired conclusion.

So, it remains to prove (10). However, note that by (9) this is exactly the same computation as that performed for the case $N=1$ with $x' \leftrightarrow \phi$ and $x' \cup \{n\} \leftrightarrow \{1\}$. Since we have already treated this case then the proof of the induction step is completed. This proves the theorem of Ahlswede and Daykin. □

We next indicate how the FKG inequality follows from the AD (= Ahlswede-Daykin) inequality. As usual it suffices to prove this in the case that $\Gamma = 2^{[N]} := T$ partially ordered by inclusion. Note that if A and B are upper ideals in T (i.e., $x, y \in A \Rightarrow x \cup y \in A$) then the indicator functions $f = I_A$ and $g = I_B$ (where $I_A(x) = 1$ if and only if $x \in A$) are increasing. Taking $\alpha = \beta = \gamma = \delta = \mu$ in (5) and $X = A$, $Y = B$ in (5′) we have

(11) $$\mu(A)\mu(B) \leqslant \mu(A \vee B)\mu(A \wedge B).$$

But

$$\mu(A) = \sum_{x \in A} \mu(x) = \sum_{x \in T} f(x)\mu(x),$$

$$\mu(B) = \sum_{x \in T} g(x)\mu(x),$$

$$\mu(A \wedge B) = \sum_{z \in A \wedge B} \mu(z) = \sum_{z \in T} f(z)g(z)\mu(z),$$

$$\mu(A \vee B) = \sum_{z \in A \vee B} \mu(z) \leqslant \sum_{z \in T} \mu(z).$$

Thus, (11) implies

$$\sum_z f(z)g(z)\mu(z) \sum_z \mu(z) \geqslant \sum_z f(z)\mu(z) \sum_z g(z)\mu(z)$$

which is just the FKG inequality for this case. The general FKG inequality is proved in just this way by first writing an arbitrary increasing function f on T as $f = \sum_i \lambda_i I_{A_i}$ where $\lambda_i \geqslant 0$ and the A_i are suitable upper ideals in T. That is, for

$$f = \sum_i \lambda_i I_{A_i}, \qquad \lambda_i \geqslant 0,$$

we have

$$\sum_{z \in T} f(z)\mu(z) = \sum_{z \in T} \sum_i \lambda_i I_{A_i}(z)\mu(z)$$
$$= \sum_i \lambda_i \sum_{z \in T} I_{A_i}(z)\mu(z), \qquad \text{etc.,}$$

and we can now apply the preceding inequality.

Some Consequences of the AD Inequality

By specializing the choices of α, β, γ and δ in (5), many results which have appeared in various places and times in the literature can be obtained. We now describe some of these.

To begin with, setting $\alpha = \beta = \gamma = \delta = 1$ we have $\alpha(T) = |T|$ for $T \subseteq [N]$ and we obtain

(12) $$|X|\,|Y| \leqslant |X \vee Y|\,|X \wedge Y| \quad \text{for all } X, Y \subseteq 2^{[N]}$$

This was first proved by Daykin [Day 1] (who also showed that this implies a lattice is distributive) and has as immediate corrollaries:

(a) (Seymour [Sey]). For any two upper ideals U, U' of $2^{[N]}$,

$$|U|\,|U'| \leqslant |U \cap U'| \cdot 2^N.$$

(b) (Kleitman [Kle 1]). For any upper ideal U and lower ideal L of $2^{[N]}$,

$$|U|\,|L| \geqslant |U \cap L| \cdot 2^N.$$

(c) (Marica-Schönheim [MS])

$$|A| \leqslant |A \backslash A| \qquad \text{for all } A \subseteq 2^{[N]}.$$

Kleitman's result first appeared in 1966 and, in fact, directly implies (a). The 1969 result (c) of Marica and Schönheim arose in connection with the following (still unresolved) number-theoretic conjecture of the author.

Conjecture: If $0 < a_1 < a_2 < \ldots < a_n$ are integers then

$$\max_{i,j} \frac{a_i}{gcd(a_i, a_j)} \geqslant n.$$

The Marica-Schönheim inequality is equivalent to the validity of the conjecture when all the a_k are squarefree. To see this, suppose

$$a_i = \prod_k p_k^{\varepsilon_{ik}}, \qquad \varepsilon_{ik} = 0 \text{ or } 1,$$

where p_k denotes the k^{th} prime. To each a_i associate the set $S_i \subseteq \mathbb{Z}^+$ defined by

$$S_k = \{k : \varepsilon_{ik} = 1\}.$$

Then

$$a_i = \prod_{k \in S_i} p_k$$

and

$$\frac{a_i}{gcd(a_i, a_j)} = \prod_k p_k^{\varepsilon_{ik} - \min(\varepsilon_{ik}, \varepsilon_{jk})} = \prod_{k \in S_i \setminus S_j} p_k.$$

By (c) there must be at least n *different* sets $S_i \setminus S_j$ and so, at least n different integers of the form $\frac{a_i}{gcd(a_i, a_j)}$. Thus, the largest such value must be at least as large as n, which is the desired conclusion. (Further generalizations can be found in [Mar] and [DL].)

The implication (12) $\Rightarrow$ (c) is short and sweet – simply note that for $A, B \subseteq 2^{[N]}$,

$$\begin{aligned} |A|\,|B| &= |A|\,|\bar{B}| \leqslant |A \vee \bar{B}|\,|A \wedge \bar{B}| \\ &= |\overline{A \vee \bar{B}}|\,|A \wedge \bar{B}| \\ &= |\bar{A} \wedge B|\,|A \wedge \bar{B}| \\ &= |B \setminus A|\,|A \setminus B| \end{aligned}$$

and set $A = B$.

Partial results currently available for the conjecture can be found in the surveys [Won] and [EG].

Linear Extensions of Partial Orders

An important area in which the FKG and related inequalities have had important applications has been in the theory of partially ordered sets and the analysis of related sorting algorithms in computer science. Many algorithms for sorting n numbers $\{a_1, a_2, \ldots, a_n\}$ proceed by using binary comparisons $a_i : a_j$ to construct successively stronger partial orders P until a linear order finally emerges (e.g., see Knuth [Knu]). A fundamental quantity in analyzing the expected behavior of such algorithms is $Pr(a_i < a_j | P)$, the probability that the result of $a_i : a_j$ is $a_i < a_j$ when all linear orders consistent with P are considered equally likely. In this section we describe a number of recent results of this type.

First, we need some notation. Let $(P, <)$ be a finite partially ordered set and for $n=|P|$, let Λ denote the set of all $1-1$ mappings of P onto $[n]$. We will assign a uniform probability distribution on Λ so that each $\lambda\in\Lambda$ has the probability $\frac{1}{n!}$ of occurring. A map $\lambda\in\Lambda$ is said to be a *linear extension* of P if

$$x<y \text{ in } P \Rightarrow \lambda(x)<\lambda(y).$$

We denote the set of linear extensions of P by $\Lambda(P)$. We are going to consider the probabilities of certain "events" occuring where we will think of an event as some other partial order on the elements of P which is preserved by various $\lambda\in\Lambda$. Thus, $Pr(Q)=\frac{|\Lambda(Q)|}{n!}$ is just the fraction of $\lambda\in\Lambda$ which are linear extensions of Q, i.e., $u<v$ in Q implies $\lambda(u)<\lambda(v)$. Similarly, $Pr(P$ and $Q)$ is the fraction of $\lambda\in\Lambda$ which are linear extensions of both P and Q, while $Pr(P|Q)$ is defined as usual to mean $Pr(P$ and $Q)/Pr(Q)$, provided the denominator does not vanish.

To get into the spirit of this topic we first give a relatively simple result. Suppose $(P, <)$ consists of two *chains* $A=\{a_1<\ldots<a_m\}$ and $B=\{b_1<\ldots<b_n\}$. Of course, relations of the form $a_i<b_j$ and $b_r<a_s$ are also allowed. Suppose Q and Q' are two events both of which are unions of sets of the form $\{a_{i_1}<b_{j_1}, a_{i_2}<b_{j_2}, \ldots\}$, i.e., such that a's are less than b's. It is quite natural to believe that the events Q and Q' are *positively correlated* when considering linear extensions of P, i.e.,

$$Pr(Q|P \text{ and } Q')\geq Pr(Q|P)$$

or, more symmetrically,

$$Pr(Q \text{ and } Q'|P)\geq Pr(Q|P)\,Pr(Q'|P). \tag{13}$$

In fact, the inequality in (13) is a theorem of Graham, Yao and Yao [GYY]. The first proof of (13) was a rather complicated combinatorial proof which used, among other things, the Marica-Schönheim inequality. Shortly after [GYY] appeared, Shepp [She 1] (and independently Kleitman and Shearer [KS]) provided rather short proofs based on the FRG inequality. Before sketching Shepp's proof, we give an example which shows that if the hypothesis that P consists of two chains is weakened even slightly then (13) no longer remains valid.

Example: Let $(P, <)$ consist of the set $\{a_1, a_2, b_1, b_2, b_3\}$ with the following partial order:

$$\{a_1<a_2, a_1<b_2, a_2<b_1, b_1<b_3, b_2<b_3\}.$$

Let Q denote the event $\{a_1<b_1\}$ and let Q' denote the event $\{a_2<b_3\}$. A straightforward calculation shows that

$$Pr(Q|P \text{ and } Q')=\tfrac{3}{5}<\tfrac{5}{8}=Pr(Q|P),$$

which violates (13).

Proof of (13). (Shepp [She 1]). Define a lattice Γ with elements of the form $\bar{x}=(x_1,\dots,x_m)$ where $1\leqslant x_1<x_2<\dots<x_m\leqslant m+n$. Let us define $\bar{x}\leqslant\bar{x}'$ to mean $x_i\leqslant x_i'$ for $1\leqslant i\leqslant m$. Thus, we have

$$\bar{x}\vee\bar{x}'=(\dots,\max(x_i,x_i'),\dots)$$

$$\bar{x}\wedge\bar{x}'=(\dots,\min(x_i,x_i'),\dots).$$

It is easily checked that with these definitions Γ is a distributive lattice. For each $\bar{x}\in\Gamma$ we can associate a unique mapping $\lambda_{\bar{x}}\in\Lambda$ by setting:

$$\lambda_{\bar{x}}(a_i)=x_i, \qquad \lambda_{\bar{x}}(b_j)=y_j$$

where $[m+n]\backslash\{x_1<\dots<x_m\}=\{y_1<\dots<y_n\}$.

Finally, define

$$\mu(\bar{x})=\begin{cases}1 & \text{if } \lambda_{\bar{x}}\in\Lambda(P)\\ 0 & \text{otherwise}\end{cases}$$

$$f(\bar{x})=\begin{cases}1 & \text{if } \lambda_{\bar{x}}\in\Lambda(Q)\\ 0 & \text{otherwise}\end{cases}$$

$$f'(\bar{x})=\begin{cases}1 & \text{if } \lambda_{\bar{x}}\in\Lambda(Q')\\ 0 & \text{otherwise}\end{cases}$$

To apply the FKG inequality, we must first verify the log supermodularity of μ:

(14) $$\mu(\bar{x})\mu(\bar{x}')\leqslant\mu(\bar{x}\vee\bar{x}')\mu(\bar{x}\wedge\bar{x}') \quad \text{for all } \bar{x},\bar{x}'\in\Gamma.$$

Suppose $\mu(\bar{x})\mu(\bar{x}')=1$. Then $\lambda_{\bar{x}}\in\Lambda(P)$, $\lambda_{\bar{x}'}\in\Lambda(P)$. If $a_i<a_j$ in P then

$$\lambda_{\bar{x}}(a_i)=x_i<x_j=\lambda_{\bar{x}}(a_j)$$

$$\lambda_{\bar{x}'}(a_i)=x_i'<x_j'=\lambda_{\bar{x}'}(a_j)$$

and so,

$$\lambda_{\bar{x}\vee\bar{x}'}(a_i)=\max(x_i,x_i')<\max(x_j,x_j')=\lambda_{\bar{x}\vee\bar{x}'}(a_j).$$

Similarly, if $b_i<b_j$ in P then

$$\lambda_{\bar{x}\vee\bar{x}'}(b_i)<\lambda_{\bar{x}\vee\bar{x}'}(b_j).$$

More interesting is the case that $a_i<b_j$ in P. Then

$$\lambda_{\bar{x}}(a_i)=x_i<y_j=\lambda_{\bar{x}}(b_j),$$

$$\lambda_{\bar{x}'}(a_i)=x_i'<y_j'=\lambda_{\bar{x}'}(b_j).$$

But this implies

$$x_i\leqslant i+j-1, \qquad x_i'\leqslant i+j-1$$

and consequently,

$$\lambda_{\bar{x}\vee\bar{x}'}(a_i)=\max(x_i,x_i')\leqslant i+j-1,$$

i.e.,

$$\lambda_{\bar{x}\vee\bar{x}'}(a_i)<\lambda_{\bar{x}\vee\bar{x}'}(b_j).$$

The argument for $b_i < a_j$ is similar. This shows that $\lambda_{\bar{x} \vee \bar{x}'} \in \Lambda(P)$, i.e., $\mu(\bar{x} \vee \bar{x}') = 1$. In basically the same way it follows that $\mu(\bar{x} \wedge \bar{x}) = 1$. Therefore, we have shown that

$$\mu(\bar{x})\mu(\bar{x}') = 1 \Rightarrow \mu(\bar{x} \vee \bar{x}')\mu(\bar{x} \wedge \bar{x}') = 1$$

and so (14) always holds.

The final hypothesis to check before we can apply the FKG inequality is that f and f' are both decreasing. To see this, suppose $\bar{x} \leqslant \bar{x}'$ and $f(\bar{x}') = 1$. Then, by definition,

$$\lambda_{\bar{x}'} \in \Lambda(Q) = \bigcup_k \Lambda(Q_k)$$

where

$$Q_k = \{a_{i_1} < b_{j_1}, a_{i_2} < b_{j_2}, \ldots\}.$$

Thus, for some k, the elements x'_i of $\bar{x}'$ satisfy all the constraints $x'_{i_1} \leqslant i_1 + j_1 - 1$, $x'_{i_2} \leqslant i_2 + j_2 - 1$, ..., imposed by $\Lambda(Q_k)$. However, since $\bar{x} \leqslant \bar{x}'$ then $x_i \leqslant x'_i$ for all i, and in particular, $x_{i_1} \leqslant x'_{i_1} \leqslant i_1 + j_1 - 1$, etc. This implies that $\lambda_{\bar{x}} \in \Lambda(Q_k) \subseteq \Lambda(Q)$ and $f(\bar{x}) = 1$, and consequently, f is decreasing.

The FKG inequality can now be applied to the functions we have defined, yielding

$$\sum_{\bar{x} \in \Gamma} f(\bar{x}) f'(\bar{x}) \mu(\bar{x}) \sum_{\bar{x} \in \Gamma} \mu(\bar{x}) \geqslant \sum_{\bar{x} \in \Gamma} f(\bar{x}) \mu(\bar{x}) \sum_{\bar{x} \in \Gamma} f'(\bar{x}) \mu(\bar{x}).$$

Interpreting this in terms of P, Q and Q', we obtain

$$|Q \cap Q' \cap P| \, |Q| \geqslant |Q \cap P| \, |Q' \cap P|$$

i.e.,

$$Pr(Q \text{ and } Q'|P) \geqslant Pr(Q|P) Pr(Q'|P)$$

which is just (13). □

It is natural to try and extend (13) to more general partial orders P than just those which can be covered by two chains. However, examples such as the one previously given show that *some* additional hypotheses must be assumed in order for the desired positive correlation to hold. One such extension was provided by Shepp in the following result.

Theorem (Shepp [She 1]). Suppose $(P, <)$ is a union of two *disjoint* partial orders, i.e., $P = A \cup B$ where each pair a, b with $a \in A$ and $b \in B$ are incomparable. Then any two events Q and Q', each being the union of sets of the form $\{a_{i_1} < b_{j_1}, a_{i_2} < b_{j_2}, \ldots\}$ are positively correlated, i.e.,

$$Pr(Q \text{ and } Q'|P) \geqslant Pr(Q|P) Pr(Q'|P).$$

The only proof known for this result uses the FKG inequality in a somewhat more subtle way than in the preceding theorem (a similar use of FKG will be given in the next section).

Universal correlation. Suppose Q and Q' are a pair of partial orders on a common underlying set S. Following Winkler [Win 1], let us call Q and Q' *univer-*

sally correlated if for all possible partial orders P on S,

(15) $$Pr(Q \text{ and } Q'|P) \geqslant Pr(Q|P)\, Pr(Q'|P)$$

As an example, it had been a long-standing conjecture until quite recently that the pair $Q=\{x<y\}$ and $Q'=\{x<z\}$ were universally correlated. This conjecture, due to I. Rival and W. Sands and known as the XYZ conjecture, also has finally been settled (affirmatively) by Shepp, not surprisingly (by now) using the FKG inequality. An interesting application of this result can be found in [Win 2]. We give an outline of Shepp's snappy proof.

Theorem (Shepp [She 2]). For any partial order P on the set $\{x_1, x_2, \ldots, x_n\}$, the sets $Q=\{x_1<x_2\}$ and $Q'=\{x_1<x_3\}$ are positively correlated, i.e.,

$$Pr(Q \text{ and } Q'|P) \geqslant Pr(Q|P)\, Pr(Q'|P).$$

Proof (outline). Choose a large $N \gg n$ and let $\Omega=[N]^n=\{\bar{x}=(x_1, \ldots, x_m): x_i \in [N]\}$. Define a partial order on Ω by:

$$\bar{x} \leqslant \bar{y} \text{ if and only if } x_1 \geqslant y_1,\ x_i - x_1 \leqslant y_i - y_1, \quad i=2,3,\ldots,N.$$

It is easy to verify that this indeed does define a partial order on Ω and further that

$$(\bar{x} \wedge \bar{y})_i = \min(x_i - x_1, y_i - y_1) + \max(x_1, y_1), \quad i \in [n]$$

$$(\bar{x} \vee \bar{y})_i = \max(x_i - x_1, y_i - y_1) + \min(x_1, y_1), \quad i \in [n].$$

Using these expressions together with the fact that the reals partially ordered by magnitude form a distributive lattice (so that $\min(a, \max(b,c)) = \max(\min(a,b), \min(a,c))$, etc.), it follows (after a little computation) that Γ also forms a distributive lattice.

Let f be the indicator function of the event $\{x_1 \leqslant x_2\}$, i.e.,

$$f(\bar{x}) = \begin{cases} 1 & \text{if } x_1 \leqslant x_2, \\ 0 & \text{otherwise} \end{cases}$$

and let g be the indicator function of the event $\{x_1 \leqslant x_3\}$. Again, an easy calculation shows that f and g are both increasing.

Finally, take μ to be the indicator function of P, i.e.,

$$\mu(\bar{x}) = \begin{cases} 1 & \text{if } \bar{x} \text{ satisfies the inequalities in } P, \\ 0 & \text{otherwise} \end{cases}$$

In order to apply the FKG inequality it remains to verify that $\mu(\bar{x}) \geqslant 0$ (obvious) and that μ is log supermodular, i.e.,

(16) $$\mu(\bar{x} \wedge \bar{y})\mu(\bar{x} \vee \bar{y}) \geqslant \mu(\bar{x})\mu(\bar{y}) \quad \text{for all } \bar{x}, \bar{y} \in \Omega.$$

Suppose $\mu(\bar{x}) = \mu(\bar{y}) = 1$. If $x_i < x_j$ in P then $x_i < x_j$ and $y_i < y_j$ and consequently

$$(\bar{x} \wedge \bar{y})_i \leqslant \min(x_j - x_1, y_j - y_1) + \max(x_1, y_1) = (\bar{x} \wedge \bar{y})_j$$

and, in a similar way,

$$(\bar{x} \vee \bar{y})_i \leqslant (\bar{x} \vee \bar{y})_j.$$

Thus, since $\bar{x} \wedge \bar{y}$ and $\bar{x} \vee \bar{y}$ both satisfy each inequality in P, $\mu(\bar{x} \wedge \bar{y}) = \mu(\bar{x} \vee \bar{y}) = 1$ and so, (16) holds.

Having verified the hypotheses of the FKG inequality, we can apply its conclusion. This yields

$$(17) \qquad Pr(x_1 \leqslant x_2, x_1 \leqslant x_3, P)\, Pr(P) \geqslant Pr(x_1 \leqslant x_2, P)\, Pr(x_1 \leqslant x_3, P).$$

Letting $N \to \infty$, the probability that $x_i = x_j$ tends to zero and so, it follows that (17) also holds for the permutations (i.e., $1-1$ maps) induced by $x_1, \ldots, x_n$. Finally, dividing by $Pr(P)\, Pr(x_1 < x_3, P)$ we obtain the desired inequality and the theorem is proved. □

It is natural to ask for other examples of universally correlated partial orders, or even more ambitiously, to ask for a characterization of all such pairs. The apparent difficulty of the second task is increased when one notes that these are easy examples which show that such "reasonable" sets as $\{x_1 < x_2 < x_4\}$ and $\{x_1 < x_3 < x_4\}$ are *not* universally correlated. It is therefore rather surprising that there is in fact a striking characterization for such pairs.

In order to state this result (due to Winkler [Win 1]) we first need a definition. For a partial order P, let us call an inequality $x < y$ in P *minimal* if there is no z such that $x < z < y$ in P. We denote the set of minimal inequalities in P by $\Delta(P)$.

Theorem (Winkler [Win 1]). Two partial orders Q and Q' (on the same underlying set) are universally correlated if and only if

(i) Q and Q' are consistent (i.e., for no x and y do we have $x < y$ in Q and $y < x$ in Q')

(ii) $u < v$ in $\Delta(Q \cup Q') - \Delta(Q)$ and
$x < y$ in $\Delta(Q \cup Q') - \Delta(Q')$
$\Rightarrow$
$u = x$ or $v = y$.

In particular, Winkler shows that whenever Q and Q' are universally correlated, it can in fact be proved by repeated applications of the XYZ theorem and so, ultimately rests on the FKG inequality. As pointed out in [Win 1], it follows from the theorem that the pairs

$$Q = \{x < y < z\}, \quad Q' = \{x < z\}$$
$$Q = \{x < y, x < z\}, \quad Q' = \{x < u, x < v\}$$
$$Q = \{u < v, x < y, x < z\}, \quad Q' = \{x < v\}$$
$$Q = \{u < v, x < y\}, \quad Q' = \{x < v, u < y\}$$

are universally correlated, whereas the (nearly) equally plausible pairs

$$Q = \{x < y\}, \quad Q' = \{x < u < v\}$$
$$Q = \{x < y < z, u < w\}, \quad Q' = \{x < z, u < v < w\}$$
$$Q = \{x < u, y < v\}, \quad Q' = \{x < v, y < v\}$$

are *not* universally correlated.

In fact, Winkler shows that if Q and Q' are partial orders on a common n-element set which are not universally correlated then there is a partially ordered set on at most $n+1$ elements on which Q and Q' are negatively correlated.

More Applications of FKG

Consider the set of all $2^{\binom{n}{2}}$ labelled graphs on the vertex set $[n]$ with the uniform probability of $2^{-\binom{n}{2}}$ assigned to each graph. Let us call a property P of a graph *increasing* if any graph formed by adjoining additional edges to a graph with property P also has property P. Examples of increasing properties are:
(a) G is Hamiltonian;
(b) G has chromatic number at least k;
(c) G has independence number at most k;
(d) G is connected (or, more generally, k-connected);
(e) G has girth at most k;
(f) G contains a clique of size k.
Using the FKG inequality, it is almost trivial to show that any pair of increasing (or any pair of decreasing) properties are positively correlated. To see this, simply define:

$\Gamma^n :=$ set of all graphs with vertex set $[n]$ partially ordered by inclusion of the edge sets.
$f = I_P$ – the indicator function of property P,
$f' = I_{P'}$ – the indicator function of property P',
$\mu = 1$.

Thus, μ is automatically log supermodular and $\mu(X) = |X|$ for $X \subseteq \Gamma_n$. The FKG inequality therefore applies to this situation and implies at once that P and P' are positively correlated.

Another area in which the FKG inequality has been applied effectively is that of unimodality and log convexity of sequences. Recall that a real sequence $(a_0, a_1, \ldots, a_n)$ is *log convex* if

$$a_k^2 \leqslant a_{k-1} a_{k+1} \quad \text{for } 1 \leqslant k \leqslant n-1.$$

An important property of log convex sequences is that they are unimodal. A typical result of this type (due to Seymour and Welsh [SW]) is the following.

Theorem [SW]. If $(a_0, a_1, \ldots, a_n)$ is log convex and positive and $(b_0, b_1, \ldots, b_n)$ and $(c_0, c_1, \ldots, c_n)$ are both increasing (or both decreasing) then

$$\sum_{k=0}^{n} a_k b_k c_k \sum_{k=0}^{n} a_k \geqslant \sum_{k=0}^{n} a_k b_k \sum_{k=0}^{n} a_k c_k \tag{18}$$

Note the similarity in form of (18) to the FKG inequality (which is used in [SW] to prove (18)). If all a_k are set equal to 1 then (18) reduces to the previously mentioned result (2) of Chebyshev.

Order Preserving Maps of Partial Orders

Given a finite partially ordered set $(P, <)$ one can weaken the notion of a linear extension and require that a map $\rho: P \to [n]$ only satisfy

$$x < y \text{ in } P \Rightarrow \rho(x) < \rho(y)$$

We will call such a ρ *order-preserving*. Note that in particular ρ need not be $1-1$. It is only natural to expect that many of the results which hold for linear extensions also hold for order-preserving maps. While this in fact may well be true, there are still relatively few theorems available for this class of maps (no doubt, due in part to the fact that they have not been studied as much).

As an example of such an analogue, we mention the following (due to J. W. Daykin and the author).

For a partially ordered set P, define $R(P, n)$ to be the set of all order-preserving maps of P into $[n]$, and for $x \in P$, let

$$\text{range}(x) := \{\rho(x) : \rho \in R(P, n)\}.$$

Theorem: Suppose P is covered by two disjoint *chains* A and B. Let Q and Q' be partial orders on $A \cup B$ both being unions of sets of the form $\{a_{i_1} < b_{i_1}, a_{i_2} < b_{i_2}, \ldots\}$. Furthermore, assume for all pairs $a \in A$, $b \in B$ which are comparable in P that

$$\text{range}(a) \cap \text{range}(b) = \phi. \tag{19}$$

Then Q and Q' are positively correlated in P, i.e.,

$$Pr(Q \text{ and } Q' | P) \geqslant Pr(Q | P)\, Pr(Q' | P)$$

where we assume that all maps of P into $[n]$ are equally likely.

The proof is a modification of that used for (13), complicated by the fact that ρ need not be $1-1$. The key new ingredient needed is the result (pointed out by J. W. Daykin [JWD]) that range(a) is always convex, i.e., an interval in $[n]$.

A related result for linear extensions which is not yet known to hold for order-preserving maps is the following beautiful result of Stanley.

For a finite partially ordered set P, an arbitrary element $x \in P$, and an arbitrary positive integer n, let $N_i(P, x, n)$ denote the number of linear extensions $\lambda: P \to [n]$ with $\lambda(x) = i$.

Theorem (Stanley [Sta]). For any P, $x \in P$ and $n \in \mathbb{Z}^+$, the sequence $N_i(P, x, n)$, $1 \leqslant i \leqslant n$, is log concave.

This was conjectured by Chung, Fishburn and Graham (strengthening an earlier conjecture of Rivest [Riv] that the $N_i(p, x, n)$ were always unimodal) who proved it when P could be covered by two disjoint chains (see [CFG]). Stanley's proof uses the Aleksandrov-Fenchel inequalities from the theory of mixed volumes (see [Bus], [Fen]).

So far, no one has been able to establish the corresponding result for order-preserving maps of P into $[n]$ although it must certainly be true.

It would seem that Stanley's theorem (and the analogue for order-preserving maps) should have a proof based on the FKG or AD inequalities. However, such a proof has up to now successfully eluded all attempts to find it.

Concluding Remarks

It is not possible, of course, because of space limitations to explore fully all the recent developments concerning the FKG inequality and its various generalizations and applications. We mention here several sources where the interested reader can find additional material on these topics.

To begin with, a wide variety of FKG-like inequalities have been investigated by Ahlswede and Daykin [AD 1], [AD 2], [AD 3] and Daykin [Day 2], [Day 4]. Kemperman [Kem 1] has given some very pretty extensions of work of Holley [Hol], Preston [Pre] and others [Car], [Bru], [KS] which consider the FKG inequality for measures on partially ordered measure spaces. A number of inequalities related to the FKG inequality have been developed in connection with certain concepts in the statistical theory of reliability (going back at least to Esary, Proschan and Walkup [EPW] in 1967 and Sarkar [Sar] in 1969). The interested reader will find many of these in the book of Barlow and Proschan [BP]. In fact, the FKG inequality is just one among a large class of statistical multivariate inequalities about which a number of survey papers and books have recently appeared (e.g., see [Eat], [Jog], [Ton], [MO], [Kem 2]). An interesting connection between the FKG inequality and majorization on partially ordered sets is given in [Lih]. Also, recent applications of various forms of the FKG inequality to modern theoretical physics can be found in [BR], which in particular contains the following (perhaps unfamiliar) version of FKG:

Theorem: Let $d\nu = e^{w} d^n q$ be a probability measure on $\mathbb{R}^n$ with $w \in C^2(\mathbb{R}^n)$. Suppose

$$\partial^2 w / \partial q_i \partial q_j \geqslant 0, \quad i \neq j.$$

Then

$$\int fg\, d\nu \geqslant \int f\, d\nu \int g\, d\nu$$

for all increasing functions of $\mathbb{R}^n$ (such that f, g and fg are $d\nu$-integrable).

We remark in closing that because of the intimate relation between log supermodular functions and ordinary sub- and supermodular functions (f is supermodular iff $\exp f$ is log supermodular), one suspects that there are in fact deeper connections between inequalities such as FKG and the many other striking properties enjoyed by such functions than are currently known. There is every reason to believe that we have yet to realize the full potential such a more complete understanding could provide.

References

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

[AD 2] –, –, Inequalities for a pair of maps $S \times S \rightarrow S$ with S a finite set. Math. Z. 165 (1979), 267–289.

[AD 3] –, –, Integral inequalities for increasing functions, Math. Proc. Camb. Phil. Soc. 86 (1979), 391–394.

[BP] R. E. Barlow and F. Proschan, Statistical Theory of Reliability and Life Testing, Holt, New York, 1975.

[BR] G. A. Battle and L. Rosen, The FKG inequality for the Yukawa_2 quantum field theory, J. Statist. Physics 2 (1980), 123–193.

[Bru] Brunel, A., Une preuve elementaire du lemme de Holley. Théorie Ergodique, Lecture Notes in Mathematics, no. 532, Springer-Verlag (1976), 15–18.

[Bus] H. Busemann, Convex Surfaces, Interscience, New York, 1958.

[Car] Cartier, P., Inégalités de corrélation en mécanique statistique. Seminaire Bourbaki 25 (1973), 431-(01–23).

[CFG] F. R. K. Chung, P. C. Fishburn and R. L. Graham, On unimodality for linear extensions of partial orders, SIAM J. Algebraic and Discrete Methods 1 (1980), 405–410.

[Day 1] D. E. Daykin, A lattice is distributive iff $|A||B| \leq |A \vee B||A \wedge B|$, Nanta Math., 10 (1977), 58–60.

[Day 2] –, Poset functions commuting with the product and yielding Chebychev type inequalities. In: Problèmes Combinatoires et Théorie des Graphs (Paris 1976), 93–98. Colloques Internationaux du C.N.R.S., No. 260, Paris: C.N.R.S. 1978.

[Day 3] –, Inequalities among the subsets of a set, Nanta Math., 12 (1980), 137–145.

[Day 4] –, A hierarchy of inequalities, Studies in Appl. Math., 63 (1980), 263–274.

[DL] – and L. Lovász, The number of values of Boolean functions, J. London Math. Soc. ser. 2, 12 (1976), 225–230.

[JWD] J. W. Daykin (personal communication).

[Eat] M. L. Eaton, A review of selected topics in multivariate probability inequalities, Ann. Statist. 10 (1982), 11–43.

[EG] P. Erdös and R. L. Graham, Old and new problems and results in combinatorial number theory, Monographie no. 28, l'Enseignement Math., Genève, 1980.

[EPW] J. D. Esary, F. Proschan and D. W. Walkup, Association of random variables with applications, Ann. Math. Statist., 38 (1967), 1466–1474.

[Fen] W. Fenchel, Inégalités quadratique entre les volumes mixtes des corps convexes, C. R. Acad. Sci., Paris, 203 (1936), 647–650.

[FKG] C. M. Fortuin, J. Ginibre and P. N. Kasteleyn, Correlation inequalities on some partially ordered sets, Commun. Math. Phys. 22 (1971), 89–103.

[Gra] R. L. Graham, Linear extensions of partial orders and the FKG inequality, in Ordered Sets, I. Rival, ed., 213–236, D. Reidel, Dordrecht, Holland, 1982.

[GYY] -, A. C. Yao and F. F. Yao, Some monotonicity properties of partial orders, SIAM J. Algebraic and Discrete Methods 1 (1980), 251-258.

[Gri] R. B. Griffiths, Correlations in Ising Ferromagnets, J. Math. Phys., 8 (1967), 478-483.

[HLP] G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, 2nd ed., Cambridge University Press, New York, 1959.

[Har] T. E. Harris, A lower bound for the critical probability in a certain percolation process, Proc. Camb. Phil. Soc., 56 (1960), 13-20.

[Hol] R. Holley, Remarks on the FKG inequalities, Comm. Math. Phys., 36 (1974), 227-231.

[Jog] K. Jogdeo, Association and probability inequalities, Ann. Statist. 5 (1977), 495-504.

[KS] D. G. Kelley and S. Sherman, General Griffith's inequalities on correlations in Ising ferromagnets, J. Math. Phys. 9 (1968), 466-484.

[Kem 1] J. H. B. Kemperman, On the FKG-inequality for measures on a partially ordered space, Indag. Math. 39 (1977), 313-331.

[Kem 2] -, Book Review, Bull. Amer. Math. Soc. (New Series) 5 (1981), 319-324.

[Kle 1] D. J. Kleitman, Families of non-disjoint sets, J. Combinatorial Th., 1 (1966), 153-155.

[Kle 2] -, Extremal hypergraph problems, Surveys in Combinatorics, B. Bollobás, ed., London Math. Soc., Lecture Note Series 38, Cambridge Univ. Press (1979), 44-65.

[KS] - and J. B. Shearer, Some monotonicity properties of partial orders, Studies in Appl. Math. 65 (1981), 81-83.

[Knu] D. E. Knuth, The Art of Computer Programming, Vol. 3, Sorting and Searching, Addison-Wesley, Reading, Mass., 1973.

[Lith] K.-W. Lih, Majorization on finite partially ordered sets. (To appear)

[Mar] J. Marica, Orthogonal families of sets, Canad. Math. Bull. 14 (1971), 573.

[MS] - and J. Schönheim, Differences of sets and a problem of Graham, Canad. Math. Bull., 12 (1969), 635-637.

[MO] A. W. Marshall and I. Olkin, Inequalities: Theory of Majorization and Its Applications, Academic Press, New York, 1979.

[Pre] C. J. Preston, A generalization of the FKG inequalities, Comm. Math. Phys. 36 (1974), 233-241.

[Riv] R. Rivest (personal communication).

[Sar] T. K. Sarkar, Some lower bounds of reliability, Tech. Rep. No. 124, Depts. of Op. Res. and Stat., Stanford Univ., 1969.

[Sey] P. D. Seymour, On incomparable collections of sets, Mathematika, 20 (1973), 208-209.

[SW] - and D. J. A. Welsh, Combinatorial applications of an inequality from statistical mechanics, Math. Proc. Camb. Phil. Soc. 77 (1975), 485-495.

[She 1] L. A. Shepp, The FKG inequality and some monotonicity properties of partial orders, SIAM J. Algebraic and Discrete Methods 1 (1980), 295-299.

[She 2] -, The XYZ conjecture and the FKG inequality, Ann. of Prob. 10 (1982), 824-827.

[Sta] R. P. Stanley, Two combinatorial applications of the Aleksandrov-Fenchel inequalities, J. Combinatorial Theory (Ser. A) 31 (1981), 56-65.

[Ton] Y. L. Tong, Probability Inequalities in Multivariate Distributions, Academic Press, New York, 1980.

[Win 1] P. Winkler, Correlation among partial orders. (To appear)

[Win 2] -, Average height in a partially ordered set, Dis. Math. 39 (1982), 337-341.

[Won] W. Wong, On a number theoretic conjecture of Graham, Ph. D. Dissertation, Dept. of Math., Columbia Univ., 1981.