Negative Dependence via the FKG Inequality

Devdatt Dubhashi Department of Computer Science and Engineering
Chalmers University
SE 412 96, Göteborg, SWEDEN
dubhashi@cs.chalmers.se

Abstract

We show how the FKG Inequality can be used to prove negative dependence properties.

1 Introduction

Pemantle[21] points out that there is a need for a theory of negatively dependent events analogous to the existing natural and useful theory of positively dependent events. Informally speaking, random variables are said to be *negatively dependent*, if they have the following property: if any one subset of the variables is "high", then other (disjoint) subsets of the variables are "low". He gives a survey of various notions of negative dependence, two of which are reviewed below, and points out several applications of consequences of one of these concepts: negative association. These examples include natural stochastic processes and structures such as the uniform random spanning tree[16], the simple exclusion process[15], the random cluster model and occupancy numbers of competing bins[6]. Such variables arise frequently in the analysis of algorithms where a stream of random bits influences either the input or the execution of the algorithm. Negative dependence may also be used to obtain information on the distribution of functionals such as the sum of the variables involved (or more generally, a non-decreasing function). Newman[19] shows that under negative dependence, one obtains a Central Limit Theorem (CLT) for stationary sequences of random variables. Dubhashi and Ranjan[6] show that under negative dependence conditions, one can employ classical tools of concentration of measure such as the Chernoff-Hoeffding (CH) bounds and related martingale inequalities (see also [20, 23, 4, 14]).

A classical case of negative dependence occurs in *occupancy problems*, where m balls are randomly allocated into n bins. Typical random variables of interest are the occupancy numbers B_i, $i \in [n]$, that is, the number of balls that are contained in bin i. The B_i's are dependent, since $\sum_i B_i = m$. The intuitive argument from above—if one of the B_i's is "large", the other variables are less likely to be "large" as well—suggests that they are negatively dependent. Occupancy problems arise in the analysis of algorithms from areas as diverse as dynamic load balancing [1], simulation of parallel computer models on realistic parallel machines [4], and distributed graph coloring [20]. There are also certain non-linear and non-independent generalizations of the basic balls and bins model that have been studied, in particular, in the literature in applied Economics [3, 11, 24]. Recently [5] such models have also been shown to have relevance in Computer Science in random graph models of the structure of the internet.

Another paradigm example of negative dependence are the so-called *permutation distributions* that arise in many settings corresponding to resource sharing by random reordering strategies.

Establishing negative dependence in these examples is useful in order to carry out probabilistic analysis by employing standard probabilistic tools such as central limit theorems, large deviation estimates and martingale inequalities. However, this can be a hard task, and it is mostly accomplished by *ad-hoc* techniques.

In this paper, we establish negative dependence in the examples mentioned above in two strong forms: negative association and negative dependence, two concepts that are defined in § 3 below. We use the FKG inequality, a celebrated tool in the well developed theory of positive dependence [8, 9, 2]. We show that it can be adapted also to give proofs of negative association involving occupancy numbers and permutation distributions.

1.1 Organization

The paper is organized as follows. A detailed description of the probabilistic experiments is given in Section 2. We review the notions of negative association and negative regression in § 3 and the FKG inequality in § 4. In § 5, we show that permutation distributions are negatively associated by an application of the FKG inequality in an interesting lattice. In § 6, we obtain certain inequalities involving sums of occupancy numbers that are reminiscent of majorization inequalities. Curiously, some of these inequalities involve non-disjoint sets of variables, so they do not follow directly from other results on negative association of occupancy numbers such as in [6].

2 Examples

For a positive integer n, let $[n] := \{1, \ldots, n\}$; for $I \subseteq [n]$, let $\bar{I} := [n] - I$. We investigate probabilistic experiments where m balls are randomly distributed among n bins. Let B_i, $i \in [n]$, denote the *occupancy number* of bin i, that is, the number of balls that are contained in bin i at the end of the experiment.

2.1 Independent Balls

Balls are thrown independently into bins with $\Pr(\text{ball } j \text{ goes into bin } i) = p_{i,j}$, $i \in [n], j \in [m]$, and for each ball j, $\sum_i p_{i,j} = 1$. In the uniform case where $p_{i,j} = p_i$ for each $j \in [m]$, $(B_1, \ldots, B_n)$ have the usual multinomial distribution with

$$\Pr(B_1 = m_1, \ldots, B_n = m_n) = \frac{n!}{m_1! \cdots m_n!} \cdot p_1^{m_1} \cdots p_n^{m_n} \ ,$$

when $\sum_i m_i = m$. This is sometimes called the *Maxwell–Boltzmann model*.

2.2 Permutation Distributions

In the so-called *Fermi–Dirac model*, bins contain at most one ball, and each distribution of balls among the bins is equally likely to occur. (This requires $m < n$.) The B_i's are indicator variables in this case, and for $m_i \in \{0, 1\}$, $i \in [n]$, with $\sum_i m_i = m$,

$$\Pr(B_1 = m_1, \ldots, B_n = m_n) = \binom{n}{m}^{-1} \ .$$

The joint distribution of $(B_1, \ldots, B_n)$ in the Fermi–Dirac model is a special case of a *permutation distribution* for n random variables.

Definition 1 *Let n be a positive integer.*

1. *The random variables $J_1, \ldots, J_n$ have the* permutation distribution on $[n]$, *if they take values in $[n]$ and, for every permutation $\sigma : [n] \to [n]$,*

$$\Pr(J_1 = \sigma(1), \ldots, J_n = \sigma(n)) = \frac{1}{n!} \ .$$

2. *Let $x_1, \ldots, x_n$ be arbitrary real numbers. The random variables $X_1, \ldots, X_n$ are said to have a* permutation distribution on $(x_1, \ldots, x_n)$, *if there is a set of random variables $J_1, \ldots, J_n$ with the permutation distribution on $[n]$ and $X_i = x_{J_i}$ for each $i \in [n]$.*

We shall refer to either situation as a permutation distribution.

Remark 2 If $x_1, \ldots, x_n$ are all distinct, then this definition is equivalent to stating that

$$\Pr(X_1 = x_{\sigma(1)}, \ldots, X_n = x_{\sigma(n)}) = \frac{1}{n!}$$

for every permutation $\sigma : [n] \to [n]$. This is apparently the definition of Joag-Dev and Proschan [12].[1] However, this is *not equivalent* if the x_i's are not all distinct, which is the case needed in our application to the Fermi–Dirac model.

3 Negative Dependence of Random Variables

We consider only discrete random variables. $\mathbf{X} = (X_1, \ldots, X_n)$ denotes a tuple of random variables $X_1, \ldots, X_n$; we will assume that all expectations $E[h(\mathbf{X})]$ exist.

Two random variables X, Y are called *negatively correlated*, if $\mathrm{cov}(X, Y) := E[XY] - E[X]E[Y] \leq 0$. The following definition from [12] is a natural generalization of negative correlation (and other notions of negative dependence) to the case of n random variables.

Definition 3 (–A) *The random variables $\mathbf{X} = (X_1, \ldots, X_n)$ are* negatively associated *if for every index set $I \subseteq [n]$, $\mathrm{cov}(f(X_i, i \in I), g(X_i, i \in \bar{I})) \leq 0$, that is,*

$$E[f(X_i, i \in I)g(X_i, i \in \bar{I})] \leq E[f(X_i, i \in I)]E[g(X_i, i \in \bar{I})] \ ,$$

for all non-decreasing functions $f : \mathbb{R}^{|I|} \to \mathbb{R}$ and $g : \mathbb{R}^{n-|I|} \to \mathbb{R}$. (A function $h : \mathbb{R}^k \to \mathbb{R}$ is said to be non-decreasing, if $h(\mathbf{x}) \leq h(\mathbf{y})$ whenever $\mathbf{x} \leq \mathbf{y}$ in the component-wise ordering on $\mathbb{R}^k$.)

Note that the same inequality will hold if f and g are both non-*increasing* functions.

Negative association of random variables is preserved under taking subsets, forming unions of independent sets, and forming sets of non-decreasing functions that are defined on disjoint subsets of the random variables. The following proposition makes some of these very useful properties more precise, see [12].

Proposition 4 1. *If $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ both satisfy $(-A)$ and are mutually independent, then the augmented vector $(\mathbf{X}, \mathbf{Y}) = (X_1, \ldots, X_n, Y_1, \ldots, Y_m)$ satisfies $(-A)$.*

[1] They use the term "expermutation" which we were not able to locate in the literature.

2. *Let $\mathbf{X} := (X_1, \ldots, X_n)$ satisfy $(-A)$. Let $I_1, \ldots, I_k \subseteq [n]$ be disjoint index sets, for some positive integer k. For $j \in [k]$, let $h_j : \mathbb{R}^{|I_k|} \to \mathbb{R}$ be non-decreasing functions, and define $Y_j := h_j(X_i, i \in I_j)$. Then the vector $\mathbf{Y} := (Y_1, \ldots, Y_k)$ also satisfies $(-A)$. That is, non-decreasing functions of disjoint subsets of negatively associated variables are also negatively associated. The same is true if each h_j is a non-increasing function.*

Remark 5 It is obvious from the definition that two negatively associated random variables are negatively correlated. In general, the notion of negative association is much stronger than the notion of negative correlation.

A different (and probably incomparable) notion of negative dependence is:

Definition 6 (–R) *The random variables $\mathbf{X} = (X_1, \ldots, X_n)$ satisfy* negative regression *if for every two disjoint index sets $I, J \subseteq [n]$, and all non-decreasing functions $f : \mathbb{R}^{|I|} \to \mathbb{R}$,*

$$E[f(X_i, i \in I) \mid X_j = a_j, j \in J]$$

is non-increasing in each $a_j, j \in J$.

4 The FKG Inequality

We recall some concepts from the theory of partial orders. A (finite) *lattice* $(L, \leq_L)$ is a (finite) set L, partially ordered by $\leq_L$, in which every two elements x, y have a least upper bound, denoted $x \vee y$ and called the *join* of x and y, and a greatest lower bound, denoted $x \wedge y$ and called the *meet* of x and y. A lattice L is called *distributive*, if, for all $x, y, z \in L$, we have the following two *distributive laws*:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \quad \text{or, equivalently,} \quad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \ .$$

A function $f : L \to \mathbb{R}$ on a lattice $(L, \leq_L)$ is said to be *non-decreasing* (*non-increasing*) with respect to $\leq_L$, if $x \leq_L y$ implies $f(x) \leq f(y)$ (respectively, $x \leq_L y$ implies $f(x) \geq f(y)$). A function $\mu : L \to \mathbb{R}^+$ is called *log-supermodular*, if, for all $x, y \in L$,

$$\mu(x)\mu(y) \leq \mu(x \vee y)\mu(x \wedge y) \ . \tag{4.1}$$

We give two examples of lattices that we will use in later sections.

Example 7 For positive integers n, m, define $L := [n]^m$ and $\leq_L$ to be the component-wise order, that is, for $\mathbf{a} = (a_1, \ldots, a_m), \mathbf{b} = (b_1, \ldots, b_m) \in L$,

$$\mathbf{a} \leq_L \mathbf{b} \iff a_k \leq b_k \text{ for each } k \in [m] \ .$$

Join and meet are given by the following equations on the components,

$$(\mathbf{a} \vee \mathbf{b})_k := \max\{a_k, b_k\} \quad \text{and} \quad (\mathbf{a} \wedge \mathbf{b})_k := \min\{a_k, b_k\} \ ;$$

and it turns out that $(L, \leq_L)$ is a distributive lattice because of the following property of integers,

$$\begin{aligned} \min\{u, \max\{v, w\}\} &= \max\{\min\{u, v\}, \min\{u, w\}\} \ , \\ \max\{u, \min\{v, w\}\} &= \min\{\max\{u, v\}, \max\{u, w\}\} \ . \end{aligned}$$

Example 8 For $m < n$, let L_m be the set of (ordered) m-element subsets $S = \{s_1, \ldots, s_m\}$ of $[n]$, that is, $s_1 < \cdots < s_m$. For $S, S' \in L_m$, we define $S \preceq S'$ if $s_k \leq s'_k$ for all $k \in [m]$. If we identify S with $(s_1, \ldots, s_m) \in [n]^m$, we can view $(L_m, \preceq)$ as a sublattice of the lattice $(L, \leq_L)$ from the previous example. (Note that L_m is closed under $\vee$ and $\wedge$. For m-element subsets S, S' of $[n]$ and any $k \in [m-1]$, $(S \vee S')_{[k+1]} := \max\{s_{[k+1]}, s'_{[k+1]}\} > \max\{s_{[k]}, s'_{[k]}\} = (S \vee S')_{[k]}$, since $s_{[k+1]} > s_{[k]}$ and $s'_{[k+1]} > s'_{[k]}$. Therefore, $(S \vee S') \in L_m$, and $(S \wedge S') \in L_m$ is proved similarly.) $(L_m, \preceq)$ is distributive, since it is a sublattice of the distributive lattice $(L, \leq_L)$. (The lattice $(L_m, \preceq)$ has also been considered in [25].

There is an interesting relationship between $(L_m, \preceq)$ and $(L_{n-m}, \preceq)$. For m-element subsets S, S' of $[n]$, $S \preceq S'$ if and only if $\overline{S'} \preceq \overline{S}$. For $i \in [m]$, let $\overline{S'}_i =: \ell + i$, $\ell \geq 0$. This means $\ell + i - 1 \geq S'_\ell \geq S_\ell$, since $S \preceq S'$, and, in turn, $\ell + i - 1 \leq \overline{S}_{i-1}$. This implies $\overline{S}_i \geq \ell + i = \overline{S'}_i$, that is, $\overline{S'} \preceq \overline{S}$.

The FKG inequality extends the correlation of monotone functions on the real line to the situation in which functions are defined on a lattice.

Theorem 9 (FKG Inequality [8, 22, 2]) *Let L be a finite, distributive lattice and let $\mu : L \to \mathbb{R}^+$ be a log-supermodular function. Then, if $f, g : L \to \mathbb{R}$ are both non-decreasing or both non-increasing with respect to $\leq_L$, we have*

$$\sum_{x \in L} f(x)\mu(x) \cdot \sum_{x \in L} g(x)\mu(x) \leq \sum_{x \in L} f(x)g(x)\mu(x) \cdot \sum_{x \in L} \mu(x) \ .$$

If one of the functions is non-decreasing and the other is non-increasing, then the reverse inequality holds.

It is helpful to view μ as a measure on L. Assuming that μ is not identically zero, we can define, for any $f : L \to \mathbb{R}$, its expectation $E[f] := (\sum_{x \in L} f(x)\mu(x))/(\sum_{x \in L} \mu(x))$. In this notation, the FKG inequality asserts, for example, that for any log-supermodular μ and functions $f, g : L \to \mathbb{R}$,

$$E[f] \cdot E[g] \geq E[f \cdot g] \ ,$$

if one of the functions is non-decreasing and the other one is non-increasing. This should be taken not only as a formal similarity with Definition 3 but as an indication why the FKG inequality is at the core of many proofs of negative association among random variables.

5 Permutation Distributions are Negatively Associated

We will prove that random variables having a permutation distribution are negatively associated. Basically, this result already appears in [12, Theorem 2.11]. Here we give a new short proof of this result via the FKG inequality.

Theorem 10 *Random variables having the permutation distribution are negatively associated.*

Proof. We shall first show that for any positive integer n, the permutation distribution on $[n]$ is negatively associated. Let $J_1, \ldots, J_n$ have the permutation distribution on $[n]$. Let $I \subseteq [n]$ be an arbitrary index set with $|I| = k \leq n$. For a k-element subset $S = \{S_1, \ldots, S_k\} \subseteq [n]$ and a permutation τ on S, we shall write $\tau(S)$ for the vector $(\tau(S_1), \ldots, \tau(S_k))$.

Let $(L_k, \preceq)$ be the lattice on the k-element subsets of $[n]$ as defined in Example 8. For non-decreasing functions $f : \mathbb{R}^k \to \mathbb{R}$, $g : \mathbb{R}^{n-k} \to \mathbb{R}$, we define real-valued functions f', g' on $(L_k, \preceq)$ by setting

$$f'(S) := \frac{1}{k!} \sum_{\tau} f(\tau(S)) \ , \quad g'(S) := \frac{1}{(n-k)!} \sum_{\rho} g(\rho(\overline{S})) \ ,$$

where τ ranges over all permutations of S and ρ ranges over all permutations of $\overline{S}$. Then f' is non-decreasing and g' is non-increasing on the lattice. To see that f' is non-decreasing, that is, $f(S) \leq f(S')$ if $S \preceq S'$, merely do a term-wise comparison of the two summations. To see that g' is non-increasing, observe in addition that $S \preceq S'$ if and only if $\overline{S} \succeq \overline{S'}$, see Example 8. Set $\mu(S) := \binom{n}{k}^{-1}$ to get a trivially log-supermodular measure. Observe now that (with σ varying over all permutations of $[n]$)

$$\sum_S f'(S)\mu(S) = \sum_S \sum_\tau f(\tau(S))(n-k)!/n! = \sum_\sigma f(\sigma(i), i \in I)/n! = E[f(J_i, i \in I)] \ .$$

Similarly,

$$\sum_S g'(S)\mu(S) = E[g(J_i, i \in \bar{I})]$$

and

$$\begin{aligned} \sum_S f'(S)g'(S)\mu(S) &= \sum_S \sum_\tau f(\tau(S)) \sum_\rho g(\rho(\overline{S}))/n! \\ &= \sum_\sigma f(\sigma(i), i \in I) g(\sigma(i), i \in \bar{I})/n! \\ &= E[f(J_i, i \in I) g(J_i, i \in \bar{I})] \ . \end{aligned}$$

Applying the FKG inequality, we conclude that $J_1, \ldots, J_n$ are negatively associated.

We deduce that for any reals $x_1, \ldots, x_n$, random variables $X_1, \ldots, X_n$ having the permutation distribution on $(x_1, \ldots, x_n)$ are negatively associated. Indeed, $X_i = h_i(J_i) := x_{J_i}$ are non-decreasing functions of distinct variables; hence, by Proposition 4(2), we conclude that any permutation distribution is negatively associated. □

An immediate corollary is negative association of Fermi-Dirac distributions:

Corollary 11 *The indicator variables in the Fermi–Dirac model satisfy the negative association condition* $(-A)$*.*

5.1 Application

This corollary a simple analysis of the following probabilistic experiment from [17]. Consider a $k \times n$ matrix A that is defined as follows. Row entries $A_{i\cdot}$, $i \in [k]$, are independent random variables, and for each row i, the entries A_{ij}, $j \in [n]$, are indicator variables distributed according to the Fermi–Dirac model, that is, each row of A is a random 0-1 vector of length n with exactly m ones.

Let $f(A)$ be the number of all-zero columns in A. By Corollary 11 and Proposition 4(1), the random variables A_{ij}, $i \in [k], j \in [n]$, are negatively associated and so are the random variables $C_j := 1 - \operatorname{sgn} \sum_{i \in [k]} A_{ij}$, $j \in [n]$, by Proposition 4(2) ($\operatorname{sgn} 0 := 0, \operatorname{sgn} x := 1$ for $x > 0$). Note that $f(A) = \sum_{j \in [n]} C_j$, and so one can apply Chernoff–Hoeffding bounds on $f(A)$.

In [17], Mehlhorn and Priebe consider shortest path problems on complete digraphs (with loops) with respect to *simple* weight functions. On a graph with n vertices, for every vertex v

and every integer $j \in [n]$, there is exactly one edge of length j leaving v. Among other facts, Mehlhorn and Priebe use large deviation estimates for $f(A)$ to deduce that on random simple weight functions, any algorithm for the single source shortest path problem has complexity $\Omega(n \log n)$ with high probability.

6 Correlation Inequalitites for Sums of Occupancy Numbers

Correlations of the occupancy numbers $B_1, \ldots, B_n$ in our first experiment are extensively studied in [6]; it turns out that $(B_1, \ldots, B_n)$ satisfy a number of negative dependence conditions, including negative association. By Proposition 4, this implies general correlation inequalities for non-decreasing functions of disjoint subsets of the occupancy numbers. We now show that correlation inequalities involving sums of these occupancy numbers can be obtained in a more direct way via the FKG inequality.

A possible configuration of the experiment can be represented by a vector $\mathbf{a} := (a_1, \ldots, a_m)$, with $a_k \in [n]$ for each $k \in [m]$. This is the configuration where ball k goes into bin a_k for each $k \in [m]$. Define the lattice $(L, \leq_L)$ on all such configurations as in Example 7 and define $\mu : L \to \mathbb{R}^+$ by $\mu(\mathbf{a}) := \prod_k p_{a_k,k}$ for each $\mathbf{a} \in L$. For any $\mathbf{a}, \mathbf{b} \in L$, we have $\mu(\mathbf{a})\mu(\mathbf{b}) = \mu(\mathbf{a} \vee \mathbf{b})\mu(\mathbf{a} \wedge \mathbf{b})$, and so μ is log-supermodular.

Let $I, J \subseteq [n]$ be two index sets such that either $I \cap J = \varnothing$ or $I \cup J = [n]$; without loss of generality, we can arrange it by renumbering that $J = \{1, \ldots, |J|\}$ and $I = \{n-|I|+1, \ldots, n\}$. Let t_I, t_J be arbitrary non-negative integers and define $f, g : L \to \{0, 1\}$ to be the indicator functions of the events $(\sum_{i \in I} B_i \geq t_I)$ and $(\sum_{j \in J} B_j \geq t_J)$, respectively, where B_i, $i \in [n]$, are the (random) occupancy numbers. (The occupancy number of bin i on configuration $\mathbf{a}$ is given by $B_i(\mathbf{a}) := |\{j \mid a_j = i\}|$.) The definition of the lattice order $\leq_L$ ensures that f is non-decreasing, while g is non-increasing on L for any fixed integers t_I, t_J. Applying the FKG inequality, we get the following correlation inequality on the random variables B_i, $i \in [n]$.

Theorem 12 *Let $I, J \subseteq [n]$ be index sets such that either $I \cap J = \varnothing$ or $I \cup J = [n]$, and let t_I, t_J be arbitrary non-negative integers. Then*

$$\Pr\left(\sum_{i \in I} B_i \geq t_I, \sum_{j \in J} B_j \geq t_J\right) \leq \Pr\left(\sum_{i \in I} B_i \geq t_I\right) \cdot \Pr\left(\sum_{j \in J} B_j \geq t_J\right) \quad . \tag{6.1}$$

Remark 13 (6.1) is referred to as the *negative quadrant dependence* condition for $X := \sum_{i \in I} B_i$ and $Y := \sum_{j \in J} B_j$. It is known to be equivalent to the negative association condition $(-A)$ for X, Y, [12]. This can also be easily seen by replacing f, g in the proof of Theorem 12 by arbitrary non-decreasing functions. In fact, even more general correlation inequalities reminiscent of majorization inequalities [10] follow along the same lines. For example, if we define a partial order on tuples of occupancy numbers (for a fixed number of balls) by

$$(B_1, \ldots, B_n) \preceq (B'_1, \ldots, B'_n) \iff \sum_{k \leq i \leq n} B_i \leq \sum_{k \leq i \leq n} B'_i \text{ for all } k \in [n-1] \ ,$$

then $\mathbf{a} \leq_L \mathbf{b}$ implies $(B_1(\mathbf{a}), \ldots, B_n(\mathbf{a})) \preceq (B_1(\mathbf{b}), \ldots, B_n(\mathbf{b}))$ and, hence, the FKG inequality on L can be applied to functions on $(B_1, \ldots, B_n)$ that are non-decreasing or non-increasing with respect to $\preceq$. Such inequalities involve non-disjoint sets of random variables and thus do not follow directly from the negative assoctaion of the occupancy numbers (Theorem 14 in [6]).

7 In Praise of Dexter's Style

Those of us fortunate to have known Dexter at close quarters are unanimous in admiration of the hallmarks of Dexter's style and work: the elegance and power of his mathematical techniques. The deliberations above are offered as a humble effort to imitate some of that style.

References

[1] Adler, M., Chakrabarti, S., Mitzenmacher, M. and Rasmussen, L. (1995) Parallel randomized load balancing. In *Proc. 27th Annual ACM Symposium on the Theory of Computing* 238–247

[2] Alon, N., Spencer, J. H. and Erdős, P. (2000) *The Probabilistic Method*, second edition, John Wiley & Sons, New York.

[3] B. Arthur, *Increasing Returns and path Dependence in the Economy*, The Univ. of Michigan Press, 1994.

[4] Dietzfelbinger, M. and Meyer auf der Heide, F. (1993) Simple, efficient shared memory simulations. In *Proc. 5th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '93)* 110–119

[5] E. Drinea, A. Frieze and M. Mitzenmacher, "Balls and Bins Models with Feedback", *Symposium on Discrete Algorithms* (SODA) 2002.

[6] D. Dubhashi and D. Ranjan, "Balls and Bins: A Study in Negative Dependence", *Random Structures and Algorithms*, 13 pp. 99 -124 1998

[7] Ebrahimi, N. and Ghosh, M. (1981) Multivariate negative dependence. *Comm. Statist. Theory Methods* **A10** 307–337

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

[9] Graham, R. L. (1983) Applications of the FKG Inequality and Its Relatives. In A. Bachem, M. Grötschel and B. Korte (Eds.), *Mathematical Programming: The State of the Art.* Springer-Verlag, Berlin, Heidelberg, New York: Springer-Verlag, 1983.

[10] G.H. Hardy, J.E. Littlewood and G. Polya, *Inequalities*, cambridge University Press, 1952.

[11] N. Johnson and S. Kotz, *Urn Models and their applications: An Approcah to Modern Discrete Probability Theory*, Wiley, New York, 1977.

[12] Joag-Dev, K. and Proschan, F. (1983) Negative association of random variables, with applications. *Ann. Statist.* **11** 286–295

[13] J. Jonasson and O. Nerman (1997) "On maximum entropy πps-sampling with fixed sample size", Report, Chalmers University of Technology, 199

[14] Kamath, A., Motwani, R., Palem, K. and Spirakis, P. (1995) Tail bounds for occupancy and the satisfiability threshold conjecture. *Random Structures Algorithms* **7** 59–80

[15] T. Liggett, "The stochastic evolution of infinite systems of interacting particles", *lecture Notes in Mathematics* vol 598, Springer-Verlag, 1977

[16] R, Lyons and Y. Peres, *Probability on Networks and Trees*, Book manuscript to appear in Cambridge University Press 2003, `http://www.math.gatech.edu/ rdlyons/prbtree/prbtree.html`

[17] Mehlhorn, K. and Priebe, V. (1995) On the all-pairs shortest path algorithm of Moffat and Takaoka. In P. Spirakis (Ed.), *Algorithms — ESA '95* (Proc. Third Annual European Symposium on Algorithms) [Lecture Notes in Computer Science **979**] 185–198. Springer-Verlag, Berlin.

[18] Motwani, R. and Raghavan, P. (1995) *Randomized Algorithms.* Cambridge University Press, Cambridge.

[19] C. Newman, "Asymptotic independence and limit theorems for positively and negatively associated random variables", in *Inequalities in Statistics and Probability*, edited by Y.L. Tong, I.M.S. Lecture notes-monograph series, vol 5, pp. 127–140.

[20] Panconesi, A. and Srinivasan, A. (to appear) Randomized distributed edge coloring via an extension of the Chernoff–Hoeffding bounds. *SIAM J. Comput.* Extended abstract appeared as A. Panconesi and A. Srinivasan (1992) Fast randomized algorithms for distributed edge coloring. In *Proc. 11th Annual ACM Symposium on Principles of Distributed Computing* 251–262.

[21] R. Pemantle, "Towards a Theory of Negative Dependence", *Journal of Mathematical Physics*, 41:3, pp. 1371 – 1390, 2000

[22] Sarkar, T. K. (1969) *Some Lower Bounds of Reliability.* Tech. Report 124, Dept. of Operations Research and Statistics, Stanford University.

[23] Schmidt, J. P., Siegel, A. and Srinivasan, A. (1995) Chernoff–Hoeffding bounds for applications with limited independence. *SIAM J. Discrete Math.* **8** 223–250

[24] C. Shapiro and H. Varian, *Information Rules*, Harvard Business School Press, Boston, MA 1999.

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

[26] Shwartz, A. and Weiss, A. (1995) *Large Deviations for Performance Analysis.* Chapman & Hall, London. Springer-Verlag, 2000.