Basic Research in Computer Science

BRICS RS-96-27 Dubhashi et al.: Negative Dependence Through the FKG Inequality

Negative Dependence Through the FKG Inequality

Devdatt Dubhashi
Volker Priebe
Desh Ranjan

BRICS Report Series RS-96-27

ISSN 0909-0878 July 1996

See back inner page for a list of recent publications in the BRICS Report Series. Copies may be obtained by contacting:

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK - 8000 Aarhus C
Denmark

Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

BRICS publications are in general accessible through WWW and anonymous FTP:

```
http://www.brics.dk/
ftp ftp.brics.dk (cd pub/BRICS)
```

Negative Dependence Through the FKG Inequality*

Devdatt Dubhashi[†]
BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade
DK-8000 Aarhus C, Denmark
dubhashi@daimi.aau.dk

Volker Priebe
Max-Planck-Institut für Informatik
Im Stadtwald
66123 Saarbrücken, Germany
priebe@mpi-sb.mpg.de

Desh Ranjan[†]
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003, USA
dranjan@cs.nmsu.edu

Abstract

We investigate random variables arising in occupancy problems, and show the variables to be negatively associated, that is, negatively dependent in a strong sense. Our proofs are based on the FKG correlation inequality, and they suggest a useful, general technique for proving negative dependence among random variables. We also show that in the special case of two binary random variables, the notions of negative correlation and negative association coincide.

*This work was supported by the ESPRIT II Basic Research Actions Program of the EC under contract no. 7141 (project ALCOM II) and by the Danish National Research Foundation through the Centre for Basic Research in Computer Science (BRICS).

[†]Work partly done when the authors were visiting Max-Planck-Institut für Informatik, Saarbrücken. Desh Ranjan also acknowledges the hospitality of BRICS.

1 Introduction

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". Such variables arise frequently in the analysis of algorithms, for which a stream of random bits influences either the input or the execution of the algorithm. To give a more specific example, we consider *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 [3], and distributed graph coloring [11].

Dependence among random variables makes the analysis of an algorithm more difficult, since independent random variables obey many simple laws that do not hold for dependent random variables. The well-known Chernoff–Hoeffding bounds from the theory of large deviations are an excellent example in this respect. Much effort has been made to salvage these sharp bounds in the more general situation under consideration; see, for example, [11, 13, 3, 8]. It turns out that one can apply the Chernoff–Hoeffding bounds to sums of "strongly" negatively dependent random variables just as one would apply them to independent random variables; see Section 6. Hence, it is useful to establish negative dependence among random variables. However, this can be a hard task, and it is mostly accomplished by ad-hoc techniques. In this paper, we show that the FKG correlation inequality from the theory of partial orders can be a useful, general tool. We give simple proofs based on the FKG inequality, establishing negative dependence among random variables in three different settings. The results are not new, but we give new proofs that are more elegant than those appearing in the literature.

Our results involve a strong notion of negative dependence called *negative association*, which, in general, is much stronger than the better known negative correlation. In Section 5.1, we give a short proof that for the special case of two binary random variables, the two notions coincide. (This result can also be obtained by combining results in [5, 7].)

We further deal with two different types of occupancy experiments. In the first one, balls are thrown independently into bins—this gives rise to a (generalized) multinomial distribution for $(B_1, \ldots, B_n)$. The B_i's are known to be negatively associated, [4]. In Section 5.3, we give a more direct proof of negative association for sums of the B_i's. In the second experiment, we assume that $m < n$, that bins contain at most one ball, and that each distribution of balls among the bins is equally likely to occur. This is the so-called Fermi–Dirac model, which can be viewed as a special case of the more general *permutation distribution*. We prove in Section 5.2 that random variables with a permutation distribution satisfy the negative association condition, a result already mentioned in [7]. In particular, it follows that the occupancy numbers in the Fermi–Dirac model are negatively associated.

The paper is organized as follows. A detailed description of the probabilistic experiments is given in Section 2. We review the notion of negative association and the FKG inequality in Sections 3 and 4, respectively. Our results are proved in Section 5, and we give an application of our results to a probabilistic analysis in Section 6.

2 Experiments

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.

We consider two types of experiments. In the first one, 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*.

In the second experiment, 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 [7].[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.

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

Two random variables X, Y are called *negatively correlated*, if $\text{cov}(X,Y) := E[XY] - E[X]E[Y] \leq 0$. The following definition from [7] 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* $I \subseteq [n]$, $\text{cov}(f(X_i, i \in I), g(X_i, i \in \bar{I})) \leq 0$, *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 [7].

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* $(-A)$.
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; however, see Theorem 9.

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 6 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 7 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 [14]. Figure 4.1 shows $(L_2, \preceq)$ for $n = 5$.)

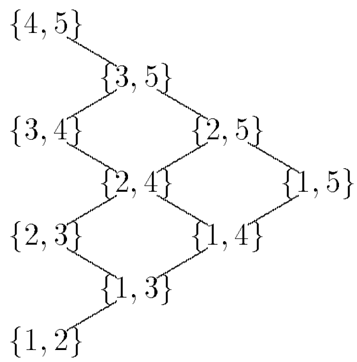

Figure 4.1: The lattice $(L_2, \preceq)$ of ordered 2-element subsets of $\{1, \ldots, 5\}$.

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 8 (FKG Inequality [6, 12, 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 Results on Negative Dependence

5.1 Negatively Correlated Coins Satisfy $(-A)$

As already mentioned in Remark 5, two random variables are negatively correlated if they are negatively associated. We now show that the converse is true if both variables are binary (that is, are "coins"); cf. [5, Theorem 2] and [7, p. 287].

Theorem 9 *Binary random variables are negatively associated if and only if they are negatively correlated.*

Proof. In view of Remark 5, it remains to be proved that two binary random variables satisfy the negative association condition $(-A)$ if they are negatively correlated. Let X, Y be negatively correlated binary random variables, that is, $\mathrm{cov}(X, Y) = E[XY] - E[X]E[Y] \leq 0$. For $i, j \in \{0, 1\}$, define $\mu_{i,j} := \Pr(X = i, Y = j)$. We have $E[X] = \Pr(X = 1) = \mu_{1,0} + \mu_{1,1}$, and $\mathrm{cov}(X, Y) \leq 0$ reads $\mu_{1,1} \leq (\mu_{1,0} + \mu_{1,1}) \cdot (\mu_{0,1} + \mu_{1,1})$. Since $\mu_{0,0} + \mu_{0,1} + \mu_{1,0} + \mu_{1,1} = 1$, this is equivalent to

$$\mu_{1,1} \cdot \mu_{0,0} = \mu_{1,1} \cdot (1 - \mu_{0,1} - \mu_{1,0} - \mu_{1,1}) \leq \mu_{1,0} \cdot \mu_{0,1} \ . \tag{5.1}$$

Let $L := \{0, 1\}^2$ and for $x = (x_1, x_2), y = (y_1, y_2) \in L$, $x \leq_L y$ if and only if $x_1 \leq y_1$ and $x_2 \geq y_2$; see Figure 5.1. It is easily seen that $(L, \leq_L)$ defines a distributive lattice.

$$\begin{tikzcd}[column sep=small] & (1,0) & \\ (0,0) \arrow[ur, dash] & & (1,1) \arrow[ul, dash] \\ & (0,1) \arrow[ul, dash] \arrow[ur, dash] & \end{tikzcd}$$

Figure 5.1: The lattice $(L = \{0,1\}^2, \leq_L)$ defined in the proof of Theorem 9.

To establish the negative association condition for (X,Y), we have to prove that $E[f(X)g(Y)] \leq E[f(X)]E[g(Y)]$ for non-decreasing functions $f, g : \mathbb{R} \to \mathbb{R}$. (All other cases are trivial.) Define real-valued functions f', g' on $(L, \leq_L)$ by setting

$$f'(x_1, x_2) := f(x_1) \ , \quad g'(x_1, x_2) := g(x_2) \ .$$

By definition of $(L, \leq_L)$, f' is non-decreasing and g' is non-increasing on the lattice. For $x = (x_1, x_2) \in L$, define $\mu(x) := \mu_{x_1, x_2}$. Note that

$$\sum_{x \in L} f'(x)\mu(x) = E[f(X)] \ , \quad \sum_{x \in L} g'(x)\mu(x) = E[g(Y)] \ , \qquad \text{and}$$
$$\sum_{x \in L} f'(x)g'(x)\mu(x) = E[f(X)g(Y)] \ .$$

Therefore, the desired result follows from the FKG inequality as soon as we have established log-supermodularity of μ. Again, there is only one non-trivial case to check, namely, $x = (1,1)$ and $y = (0,0)$ with $x \vee y = (1,0)$ and $x \wedge y = (0,1)$. However, for these elements, the log-supermodularity condition (4.1) is nothing else but (5.1), that is, log-supermodularity of μ is implied by the assumption that X, Y are negatively correlated. $\square$

5.2 Permutation Distribution Satisfies $(-A)$

In this paragraph, we prove that the indicator variables $B_1, \ldots, B_n$ in the Fermi–Dirac model are negatively associated. We will prove the stronger result that random variables having the permutation distribution are negatively associated. Basically, this result already appears in [7, 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 7. For non-decreasing functions $f : \mathbb{R}^k \rightarrow \mathbb{R}$, $g : \mathbb{R}^{n-k} \rightarrow \mathbb{R}$, 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 termwise 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 7. 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. □

The desired result is now an immediate corollary.

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

5.3 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 [4]; 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 6 and define $\mu : L \rightarrow \mathbb{R}^+$ 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 $f, g : L \rightarrow \{0, 1\}$ 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

$$\Pr\left(\sum_{i\in I} B_i \geq t_I, \sum_{j\in J} B_j \geq t_J\right) \leq \Pr\left(\sum_{i\in I} B_i \geq t_I\right) \cdot \Pr\left(\sum_{j\in J} B_j \geq t_J\right) \ . \tag{5.2}$$

Remark 13 (5.2) 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, [7]. 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 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 $\preceq$.

6 Applications

Chernoff–Hoeffding bounds are large deviation estimates for sums $S = \sum_{i=1}^n X_i$ of independent, identically distributed random variables X_i, that is, they provide bounds of the form

$$\Pr(S > an) \leq \inf_{t>0} e^{-ant}(E[\exp(tX_1)])^n \quad \text{for } a > E[X_1] = E[S]/n \ ; \tag{6.1}$$

see, for example, [15, 10]. The independence assumption can be replaced by the requirement that

$$E[\exp(t \sum_i X_i)] \leq \prod_i E[\exp(tX_i)] \ .$$

Because of Proposition 4, this is easily seen to be fulfilled if the X_i's are negatively associated.

Theorem 14 *Let $X_1, \ldots, X_n$ be identically distributed random variables whose joint distribution satisfies the negative association condition $(-A)$. Then the Chernoff–Hoeffding bounds (6.1) apply for $S := \sum_i X_i$.*

Analogues of this result are true for other notions of "strong" negative dependence among random variables, but the argument gets slightly more involved; see [4] for a more detailed account.

Theorem 14 allows, for example, a simple analysis of the following probabilistic experiment from [9]. 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 Theorem 14 allows to apply Chernoff–Hoeffding bounds on $f(A)$.

In [9], 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.

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. (1992) *The Probabilistic Method.* John Wiley & Sons, New York.

[3] 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

[4] Dubhashi, D. and Ranjan, D. (1996) Balls and Bins: A Study in Negative Dependence. Manuscript

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

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

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

[8] 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

[9] 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.

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

[11] 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.

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

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

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

[15] Shwartz, A. and Weiss, A. (1995) *Large Deviations for Performance Analysis*. Chapman & Hall, London.

Recent Publications in the BRICS Report Series

RS-96-27 Devdatt Dubhashi, Volker Priebe, and Desh Ranjan. ***Negative Dependence Through the FKG Inequality*****. July 1996. 15 pp.**

RS-96-26 Nils Klarlund and Theis Rauhe. ***BDD Algortihms and Cache Misses*****. July 1996. 15 pp.**

RS-96-25 Devdatt Dubhashi and Desh Ranjan. ***Balls and Bins: A Study in Negative Dependence*****. July 1996. 27 pp.**

RS-96-24 Henrik Ejersbo Jensen, Kim G. Larsen, and Arne Skou. ***Modelling and Analysis of a Collision Avoidance Protocol using SPIN and UPPAAL*****. July 1996. 20 pp. Presented at** ***DIMACS Workshop SPIN96 – 2nd International SPIN Verification Workshop on Algorithms, Applications, Tool Use, Theory*** **(Rutgers University, New Jersey, USA, August 5, 1996).**

RS-96-23 Luca Aceto, Wan J. Fokkink, and Anna Ingólfsdóttir. ***A Menagerie of Non-Finitely Based Process Semantics over BPA*: From Ready Simulation Semantics to Completed Tracs*****. July 1996. 38 pp.**

RS-96-22 Luca Aceto and Wan J. Fokkink. ***An Equational Axiomatization for Multi-Exit Iteration*****. June 1996. 30 pp.**

RS-96-21 Dany Breslauer, Tao Jiang, and Zhigen Jiang. ***Rotation of Periodic Strings and Short Superstrings*****. June 1996. 14 pp.**

RS-96-20 Olivier Danvy and Julia L. Lawall. ***Back to Direct Style II: First-Class Continuations*****. June 1996. 36 pp. A preliminary version of this paper appeared in the proceedings of the 1992 ACM Conference on Lisp and Functional Programming, William Clinger, editor, LISP Pointers, Vol. V, No. 1, pages 299–310, San Francisco, California, June 1992. ACM Press.**