Technical Report
MSR-TR-98-42

THE VAN DEN BERG-KESTEN-REIMER INEQUALITY: A REVIEW

C. Borgs[1,a], J. T. Chayes[1,b] and D. Randall[2,c]

[1] *Microsoft Research*
One Microsoft Way, Redmond, WA 98052

[2] *School of Mathematics & College of Computing*
Georgia Institute of Technology, Atlanta GA 30332-0160

November 1998

Abstract. We present a variant of Reimer's proof of the van den Berg-Kesten conjecture.

1. Introduction.

In this note, in honor of Harry Kesten's $66\frac{2}{3}$th birthday, we give an expository treatment of a result that is near and dear to his heart, namely the van den Berg - Kesten (BK) inequality. Specifically, we give a variant of Reimer's proof of the BK inequality for percolation.

The BK inequality provides an upper bound on the probability of the disjoint occurrence of two events in terms of the product of their probabilities. The inequality was first proved by van den Berg and Kesten in 1985 [BK] for the case in which the two events in question are both increasing or both decreasing in the sense of Fortuin, Kasteleyn and Ginibre [FKG]; the BK proof holds for a class of measures called strongly new better than used (SNBU) which includes the product measure.

[a] On leave from Universität Leipzig, Germany.
[b] On leave from University of California, Los Angeles.
[c] Work partially supported by by NSF Career Award No. CCR-9703206.

Van den Berg and Kesten conjectured that the inequality holds for all events in percolation.

During the next decade, there were many attempts to prove the BK conjecture. Van den Berg and Fiebig [BF] refined the conjecture and showed that it holds whenever each of the events in question is the intersection of an FKG increasing and an FKG decreasing event. Finally, in 1994, a general proof by Reimer [Re] confirmed the belief that the inequality holds for all events in percolation.

A version of the Reimer proof was presented by one of us in some lectures given at the Institute for Advanced Study in 1996 [CPS]. The proof given there was based on a copy of a preliminary manuscript and some notes by D. Reimer [Re], as well as on a lecture by J. Kahn, and on some comments by C. Borgs, H. Kesten and P. Deligne. The proof in [CPS] modified some of Reimer's notation and added a few details to the proofs previously seen, but the main proof presented there was very similar to that given by Reimer. In particular, the proof of the main lemma in [CPS] used the notion of butterflies, introduced by Reimer, as the principle construct.

Here we give another treatment of Reimer's proof. Although our treatment follows quite closely that of [CPS], it differs in a number of important respects: First, we review the proofs of both van den Berg and Fiebig [BF] and Fishburn and Shepp [FS], on which Reimer's proof depends. The proof presented here is therefore entirely self-contained. Second, in contrast to [CPS], we use the more familiar notions of cylinders and subcubes, rather than butterflies. We hope that the reader will find it easier to understand the proof in this more familiar language. Finally, we have streamlined and generalized aspects of the proof of Reimer's main lemma, see Lemmas 5.2 and 5.3 and the remarks in Section 5.

2. Formal Statement of the BKR Inequality.

We consider a finite probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is a product of finite sets Ω_1, Ω_2, $\dots$, Ω_n,

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \,,$$

$\mathcal{F} = 2^\Omega$ is the set of all subsets of Ω, and μ is a product of n probability measures μ_1, μ_2, $\dots$, μ_n,

$$\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n \,.$$

As usual, elements $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \Omega$ are called configurations, and two configurations $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \Omega$ and $\tilde\omega = (\tilde\omega_1, \tilde\omega_2, \dots, \tilde\omega_n) \in \Omega$ are are said to be *equal on* $S \subset [n] := \{1, 2, \dots, n\}$ if $\omega_i = \tilde\omega_i$ for all $i \in S$.

Definition: Let $S \subset [n]$, and let $S^c = [n] \setminus S$. An event $A \in \mathcal{F}$ is said to *occur on the set* S in the configuration ω if A occurs using only the random variables over

S, i.e., if A occurs independent of the values of $\{\omega_i\}_{i\in S^c}$. We denote the collection of all such ω by $A|_S$:

$$A|_S = \{\omega : \forall\tilde{\omega},\, \tilde{\omega} = \omega \text{ on } S \Rightarrow \tilde{\omega} \in A\}\,. \tag{2.1}$$

Two events $A_1, A_2 \in \Omega$ are said to *occur disjointly*, denoted by $A_1 \circ A_2$, if there are two disjoint sets on which they occur:

$$A_1 \circ A_2 = \{\omega : \exists S_1, S_2 \subset [n],\ S_1 \cap S_2 = \emptyset,\ \omega \in A_1|_{S_1} \cap A_2|_{S_2}\}\,. \tag{2.2}$$

BK Conjecture. *Let $n \in \mathbb{N}$, let Ω_i be finite sets and μ_i be probability measures on Ω_i, $i \in [n]$. Let $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n$ and let $\mathcal{F} = 2^\Omega$. Then*

$$\mu(A \circ B) \le \mu(A)\mu(B) \tag{2.3}$$

for all $A, B \in \mathcal{F}$.

Theorem 2.1 (Reimer [Re]**).** *The BK conjecture holds.*

3. Equivalent Forms of the Inequality.

First we reformulate the disjoint occurrence event $A \circ B$ in terms of cylinders. Given a configuration $\omega \in \Omega$ and a set of points $S \subset [n]$, we define the *cylinder* $[\omega]_S$ by

$$[\omega]_S = \{\tilde{\omega} : \tilde{\omega}_i = \omega_i \quad \forall i \in S\}. \tag{3.1}$$

We say that a set $A \subset \Omega$ is a cylinder set if $\exists\omega \in \Omega$ and $S \subset [n]$ such that $A = [\omega]_S$. With this definition, we may rewrite $A \circ B$ as

$$A \circ B = \{\omega : \exists S = S(\omega) \subset [n],\ [\omega]_S \subset A,\ [\omega]_{S^c} \subset B\}. \tag{3.2}$$

The first simplification of the BK conjecture (2.3) was due to van den Berg and Fiebig [BF] who showed that it is sufficient to prove the inequality for $\Omega = \{0,1\}^n$ and the uniform measure on Ω, i.e., for the pure percolation problem at density 1/2. This was a significant simplification because it turned the inequality into a purely combinatorial one.

Proposition 3.1 (van den Berg-Fiebig [BF]**).** *The BK conjecture holds if for all $n \in \mathbb{N}$ it holds for the uniform measure on $\{0,1\}^n$, i.e. if for all $n \in \mathbb{N}$ and for all $A, B \subset \{0,1\}^n$*

$$|A \circ B|\, 2^n \le |A|\, |B|\,. \tag{3.3}$$

Proof. Assume that (3.3) holds for all $n \in \mathbb{N}$ and for all $A, B \subset \{0,1\}^n$. Let $\tilde{n} \in \mathbb{N}$, let $\tilde{\Omega}_i = \{\omega_{i1}, \omega_{i2}, \ldots, \omega_{im_i}\}$ and let μ_i be probability measures on $\tilde{\Omega}_i$, $i \in [\tilde{n}]$. Let

$\tilde\Omega = \tilde\Omega_1 \times \tilde\Omega_2 \times \cdots \times \tilde\Omega_{\tilde n}$, $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_{\tilde n}$ and let $\tilde A, \tilde B \subset \tilde{\mathcal F} = 2^{\tilde\Omega}$. We then have to show that

$$\mu(\tilde A \circ \tilde B) \le \mu(\tilde A)\mu(\tilde B)\,. \tag{3.4}$$

In order to prove (3.4), we will approximate the measure $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_{\tilde n}$ by a measure $\tilde\mu = \tilde\mu_1 \times \tilde\mu_2 \times \cdots \times \tilde\mu_{\tilde n}$ for which (3.4) can be reduced to (3.3). To this end, we approximate the probability measures μ_i on $\tilde\Omega_i$ by probability measures $\tilde\mu_i = \tilde\mu_i^{(K)}$ on $\tilde\Omega_i$ with the property that

$$k_{ij} := 2^K \tilde\mu_i(\omega_{ij}) \in \mathbb{N} \cup \{0\} \tag{3.5}$$

for all $i \in [\tilde n]$ and all $j \in [m_i]$. Obviously, the sequence $\tilde\mu^{(K)}$ can be chosen in such a way that $\tilde\mu^{(K)}$ converges weakly to μ. As a consequence, it is enough to show (3.4) for measures $\tilde\mu = \tilde\mu_1 \times \tilde\mu_2 \times \cdots \times \tilde\mu_{\tilde n}$ which obey the condition (3.5).

To prove (3.4) for measures $\tilde\mu$ which obey the condition (3.5), let $n = K\tilde n$,

$$\Omega_i = \{0,1\}^K, \qquad \text{and} \qquad \Omega = \Omega_1 \times \cdots \times \Omega_{\tilde n} = \{0,1\}^n\,. \tag{3.6}$$

Let

$$\Omega_i = \bigcup_{j=1}^{m_i} \Omega_{ij} \tag{3.7}$$

be an arbitrary decomposition of Ω_i into disjoint sets Ω_{ij} with

$$|\Omega_{ij}| = k_{ij}\,, \tag{3.8}$$

and let $f : \Omega \to \tilde\Omega$ be defined by

$$f = f_1 \times \cdots \times f_{\tilde n} \quad \text{with} \quad f_i : \Omega_i \to \tilde\Omega_i \quad \text{given by} \quad f_i(\omega) = \omega_{ij} \quad \text{if} \quad \omega \in \Omega_{ij}\,. \tag{3.9}$$

Then

$$\tilde\mu(\cdot) = \mu_0(f^{-1}(\cdot)) \tag{3.10}$$

where μ_0 is the uniform measure on $\Omega = \{0,1\}^n$. Defining

$$A = f^{-1}(\tilde A) \qquad \text{and} \qquad B = f^{-1}(\tilde B)\,, \tag{3.11}$$

we then may use (3.3) to conclude that

$$\tilde\mu(\tilde A)\tilde\mu(\tilde B) = \mu_0(A)\mu_0(B) \ge \mu_0(A \circ B)\,. \tag{3.12}$$

Next, we claim that

$$f^{-1}(\tilde{A} \circ \tilde{B}) \subset f^{-1}(\tilde{A}) \circ f^{-1}(\tilde{B})\,. \tag{3.13}$$

Indeed, let $x \in f^{-1}(\tilde{A} \circ \tilde{B})$, i.e. let $f(x) \in \tilde{A} \circ \tilde{B}$. By the definition (3.2) of disjoint occurrence, this implies that there exists a set $\tilde{S} \subset [\tilde{n}]$ such that $[f(x)]_{\tilde{S}} \subset \tilde{A}$ and $[f(x)]_{\tilde{S}^c} \subset \tilde{B}$. Therefore

$$f^{-1}([f(x)]_{\tilde{S}}) \subset f^{-1}(\tilde{A}) \qquad \text{and} \qquad f^{-1}([f(x)]_{\tilde{S}^c}) \subset f^{-1}(\tilde{B})\,.$$

Defining the set $S \subset [n]$ as $S = \{i \in [n] : \lceil i/K \rceil \in \tilde{S}\}$, we then have

$$[x]_S \subset \bigcup_{y \in f^{-1}(f(x))} [y]_S = f^{-1}([f(x)]_{\tilde{S}}) \subset f^{-1}(\tilde{A}) \tag{3.14}$$

and

$$[x]_{S^c} \subset \bigcup_{y \in f^{-1}(f(x))} [y]_{S^c} = f^{-1}([f(x)]_{\tilde{S}^c}) \subset f^{-1}(\tilde{B})\,. \tag{3.15}$$

The only subtle step in (3.14) and (3.15) are the equalities, which follow from the fact that $f : \Omega \to \tilde{\Omega}$ and the set S are defined to respect the product structure (3.6) of Ω. By the definition of disjoint occurrence, (3.14) and (3.15) imply that $x \in f^{-1}(\tilde{A}) \circ f^{-1}(\tilde{B})$ which in turn implies (3.13).

Combining (3.10) with (3.13), (3.11) and (3.12), we get that

$$\begin{aligned}
\tilde{\mu}(\tilde{A} \circ \tilde{B}) &= \mu_0(f^{-1}(\tilde{A} \circ \tilde{B})) \\
&\leq \mu_0(f^{-1}(\tilde{A}) \circ f^{-1}(\tilde{B})) \\
&= \mu_0(A \circ B) \\
&\leq \tilde{\mu}(\tilde{A})\tilde{\mu}(\tilde{B})\,.
\end{aligned} \tag{3.16}$$

This gives (3.4) for all product measures obeying the condition (3.5). Choosing a sequence of measures $\tilde{\mu} = \tilde{\mu}^{(K)}$ that converges weakly to μ, we obtain the BK inequality (3.4) for general product measures μ. $\square$

Fishburn and Shepp [FS] derived yet another way of expressing the BK inequality, and it was their form that was ultimately proved by Reimer [Re]. While Fishburn and Shepp stated their inequality in the special case of the uniform measure on $\{0,1\}^n$, we will state it here in the general context of the full BK conjecture.

We need some notation: Let $X \subset \Omega$, and let $S : X \to 2^{[n]} : x \mapsto S(x) \subset [n]$ be an arbitrary map from X into $2^{[n]}$. We then define

$$[X]_S = \bigcup_{x \in X} [x]_{S(x)} \tag{3.17}$$

and

$$[X]_{S^c} = \bigcup_{x \in X} [x]_{S(x)^c}\,, \tag{3.18}$$

where, as before, $S(x)^c = [n] \setminus S(x)$.

Proposition 3.2 (Fishburn-Shepp [FS]**).** *Let $n \in \mathbb{N}$, let Ω_i be finite sets and μ_i be probability measures on Ω_i, $i \in [n]$. Let $\Omega = \Omega_1 \times \Omega_2 \cdots \Omega_n$, $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n$ and let $\mathcal{F} = 2^{\Omega}$. Then the following two statements are equivalent:*
i) For all $A, B \in \mathcal{F}$,

$$\mu(A \circ B) \leq \mu(A)\mu(B). \tag{3.19}$$

ii) For all $X \subset \Omega$ and all $S : X \to 2^{[n]}$,

$$\mu(X) \leq \mu\left([X]_S\right)\, \mu\left([X]_{S^c}\right). \tag{3.20}$$

Proof.
i) $\implies$ ii): Let $X \subset \Omega$, and let $S : X \to 2^{[n]}$. Consider

$$A = [X]_S = \bigcup_{x \in X} [x]_{S(x)} \quad \text{and} \quad B = [X]_{S^c} = \bigcup_{x \in X} [x]_{S(x)^c}.$$

Then

$$X \subset A \circ B\,.$$

Hence, by i),

$$\mu(X) \leq \mu(A \circ B) \leq \mu(A)\,\mu(B) = \mu\left([X]_S\right)\, \mu\left([X]_{S^c}\right),$$

which shows that i) implies ii).

ii) $\implies$ i): Let $A, B \subset \Omega$. Let $X = A \circ B$. By the definition (3.2) of $A \circ B$, for each $x \in X$ there is an $S(x)$ such that $[x]_{S(x)} \subset A$ and $[x]_{S(x)^c} \subset B$. We therefore have

$$[X]_S = \bigcup_{x \in X} [x]_{S(x)} \subset A$$

and

$$[X]_{S^c} = \bigcup_{x \in X} [x]_{S(x)^c} \subset B.$$

So, by ii),

$$\mu(A \circ B) = \mu(X) \leq \mu\left([X]_S\right)\, \mu\left([X]_{S^c}\right) \leq \mu(A)\,\mu(B),$$

which shows that ii) implies i). $\square$

4. Reduction of the BKR Inequality to Reimer's Main Lemma.

We now come to the main lemma in the proof of the BKR inequality. As noted before, it is enough to show the BKR inequality (2.3), and hence the Fishburn-Shepp inequality (3.20), in the special case in which μ is the uniform measure on $\Omega = \{0,1\}^n$. From now on, we will restrict ourselves to this case.

We will use the notation $\bar{x}$ to denote the bitwise complement of a configuration $x \in \Omega$, i.e. $\bar{x}_i = 1 - x_i$. For a cylinder $A = [y]_\Lambda$ and $x \in \Omega$, we define $\bar{x}^{(A)}$ to be the complement of x in A, i.e.

$$\bar{x}_i^{(A)} = \begin{cases} x_i & \text{if} \quad i \in \Lambda \\ \bar{x}_i & \text{if} \quad i \notin \Lambda \ . \end{cases} \tag{4.1}$$

For a set T, we write $\bar{T} = \bigcup_{x \in T} \bar{x}$, and $\bar{T}^{(A)} = \bigcup_{x \in T} \bar{x}^{(A)}$.

Lemma 4.1 (Reimer's Main Lemma). *Let $n \in \mathbb{N}$, $X \subset \Omega = \{0,1\}^n$ and $S : X \to 2^{[n]} : x \mapsto S(x)$. Let*

$$U = [X]_S \qquad \textit{and} \qquad V = [X]_{S^c} \ . \tag{4.2}$$

Then

$$\left|U \cap \bar{V}\right| \ = \ \left|\bar{U} \cap V\right| \ \geq \ |X|. \tag{4.3}$$

In this section, we will show that the above lemma implies the BKR inequality. It turns out that the sufficiency of the main lemma was already known independently to van den Berg [Be] and Talagrand [Ta]; however, they did not have a proof of the lemma. As noted earlier, it suffices to show the Fishburn-Shepp form (3.20) of the BK inequality for the uniform measure on $\Omega = \{0,1\}^n$. For $x, y \in \Omega$, let $\langle x, y\rangle$ be the cylinder $\langle x, y\rangle = \{z \in \Omega : z_i = x_i \text{ whenever } x_i = y_i\}$. Then

$$\begin{aligned} |U|\,|V| &= |\{(u,v) \in U \times V\}| \\ &= \sum_A |\{(u,v) \in U \times V : \langle u,v\rangle = A\}| \\ &= \sum_A |\{(u,v) \in (U\cap A) \times (V \cap A) : \langle u,v\rangle = A\}| \end{aligned} \tag{4.4}$$

where the sum runs over all cylinder sets $A \subset \Omega$. Defining

$$U_A = U \cap A \qquad \text{and} \qquad V_A = V \cap A \tag{4.5}$$

and observing that that $\langle u,v\rangle = A$ if and only if $u \in A$ and $v = \bar{u}^{(A)}$, we get that

$$\begin{aligned} |U|\,|V| &= \sum_A |\{(u,v) \in U_A \times V_A : v = \bar{u}^{(A)}\}| \\ &= \sum_A |U_A \cap \bar{V}_A^{(A)}|\,. \end{aligned} \tag{4.6}$$

We claim that Reimer's Main Lemma can be used to show that for each cylinder set $A \subset \Omega$

$$|U_A \cap \bar{V}_A^{(A)}| \geq |X_A|\,, \qquad \text{where} \qquad X_A = X \cap A\,. \tag{4.7}$$

Indeed, let $A = [z]_\Lambda$ for some $\Lambda \subset [n]$ and some $z \in \Omega$. Let $\Omega_A = \{0,1\}^{\Lambda^c}$, let $f : \Omega_A \to A : \omega \mapsto f(\omega)$ be the bijection

$$f(\omega)_i := \begin{cases} \omega_i & \text{if} \quad i \in \Lambda^c \\ z_i & \text{if} \quad i \in \Lambda\,. \end{cases} \tag{4.8}$$

and let

$$\tilde{X}_A = f^{-1}(X_A)\,. \tag{4.9}$$

Introducing

$$\tilde{S} : \tilde{X}_A \to 2^{\Lambda^c} : \tilde{x} \mapsto S(f(\tilde{x})) \cap \Lambda^c \tag{4.10}$$

and it's complement in Λ^c,

$$\tilde{S}^c : \tilde{X}_A \to 2^{\Lambda^c} : \tilde{x} \mapsto \Lambda^c \setminus (S(f(\tilde{x})) \cap \Lambda^c), \tag{4.11}$$

Reimer's Main Lemma now implies that

$$|X_A| = |\tilde{X}_A| \leq |[\tilde{X}_A]_{\tilde{S}} \cap [\overline{\tilde{X}}_A]_{\tilde{S}^c}| = |[X_A]_{S\cup\Lambda} \cap [\bar{X}_A^{(A)}]_{S^c\cup\Lambda}|\,, \tag{4.12}$$

where

$$[X_A]_{S\cup\Lambda} = \bigcup_{x\in X_A} [x]_{S(x)\cup\Lambda} \qquad \text{and} \qquad [\bar{X}_A^{(A)}]_{S^c\cup\Lambda} = \bigcup_{x\in \bar{X}_A^{(A)}} [x]_{S(x)^c\cup\Lambda}\,, \tag{4.13}$$

with, as before, $S^c(x) = [n] \setminus S(x)$.

Next we claim that

$$[X_A]_{S\cup\Lambda} \cap [\bar{X}_A^{(A)}]_{S^c\cup\Lambda} \subset U_A \cap \bar{V}_A^{(A)}\,. \tag{4.14}$$

Indeed, since $x \in A$ implies that $A = [x]_\Lambda$ which in turn implies that $[x]_{S(x)\cup\Lambda} = A \cap [x]_{S(x)}$, we have

$$\begin{aligned}[X_A]_{S\cup\Lambda} &= \bigcup_{x\in X\cap A} [x]_{S(x)\cup\Lambda} = \bigcup_{x\in X\cap A} A\cap [x]_{S(x)} \\ &\subset \bigcup_{x\in X} A\cap [x]_{S(x)} = U_A \end{aligned} \tag{4.15}$$

In a similar way, one gets

$$[\bar{X}_A^{(A)}]_{S^c\cup\Lambda} \subset \bar{V}_A^{(A)}\,. \tag{4.16}$$

The relations (4.15) and (4.16) imply (4.14). Together with (4.14), the bound (4.12) now gives the bound (4.7).

Combining (4.6) and (4.7), we get

$$|U|\,|V| \;\geq\; \sum_A |X\cap A|\,. \tag{4.17}$$

An easy counting argument gives that the right hand side of (4.17) is equal to $|X||\Omega|$. Indeed,

$$\sum_A |X\cap A| \,=\, \sum_A \sum_{x\in X\cap A} 1 \,=\, \sum_{x\in X}\sum_{A\ni x} 1 \,=\, |X||\Omega|\,,$$

which, together with (4.16), implies that

$$|U|\,|V| \;\geq\; |X||\Omega|\,, \tag{4.18}$$

the Fishburn-Shepp inequality (3.20) for the uniform measure on $\Omega = \{0,1\}^n$.

5. Proof of Reimer's Main Lemma.

The first half of the statement of the main lemma just follows from the simple observation that $x \in U\cap\bar{V} \iff \bar{x} \in \bar{U}\cap V$. We therefore have to show that $|U\cap\bar{V}| \geq |X|$. Using de Morgan's laws, this is equivalent to showing that

$$|U^c \cup \bar{V}^c| \leq |\Omega| - |X|, \tag{5.1}$$

or

$$|U^c| + |U\cap\bar{V}^c| + |X| \leq |\Omega| = 2^n\,. \tag{5.2}$$

Since $|\bar{U}^c| = |U^c|$, this is equivalent to

$$|\bar{U}^c| + |U \cap \bar{V}^c| + |X| \leq |\Omega| = 2^n. \tag{5.3}$$

To obtain (5.3), we will construct injective maps α, β and γ from $\bar{U}^c$, $U \cap \bar{V}^c$ and X into $\mathbb{R}^{2^n}$. We will show that the images of these maps are disjoint and that the union of the images is a set of linearly independent vectors in $\mathbb{R}^{2^n}$. This immediately implies that the number of elements in the union, and hence on the left hand side of (5.3), is bounded above by 2^n.

We begin by defining the maps α, β and γ. While these three maps were defined separately in Reimer's original proof, our treatment allows for a unified definition in terms of one function $\Phi : \Omega \times 2^{[n]} \to \mathbb{R}^{2^n} : (x, S) \mapsto \Phi(x, S)$. In particular, the oberservation that (5.2) is equivalent to (5.3) allows the definition of a single function. In terms of the (still to be defined) function Φ, the maps α, β and γ are defined as

$$\begin{aligned}
\alpha : \bar{U}^c \to \mathbb{R}^{2^n} : &\quad x \mapsto \Phi(x, \emptyset) \\
\beta : U \cap \bar{V}^c \to \mathbb{R}^{2^n} : &\quad x \mapsto \Phi(x, [n]) \\
\gamma : X \to \mathbb{R}^{2^n} : &\quad x \mapsto \Phi(x, S(x)).
\end{aligned} \tag{5.4}$$

To define $\Phi(\cdot, S)$, we first define functions $\varphi_i(\cdot, S)$ on a single bit x_i:

$$\varphi_i(x_i, S) = \begin{cases} (x_i, -1) & \text{if } i \notin S \\ (1, x_i) & \text{if } i \in S. \end{cases} \tag{5.5}$$

To define Φ on $\Omega = \{0,1\}^{[n]}$, we must set some notation. Let $\oplus$ denote concatenation given by, $(a, b) \oplus (c, d) = (a, b, c, d)$. Let $\otimes$ be the tensor product given by $(a, b) \otimes v = av \oplus bv$ for $a, b \in \mathbb{R}$ and $v \in \mathbb{R}^m$. Equipping $\mathbb{R}^{2^n}$ with the standard inner product: $\langle v \,|\, w \rangle = \sum_{i=1}^{2^n} v_i w_i$, notice that an easy inductive proof yields

$$\langle \bigotimes_{i=1}^n v_i \,|\, \bigotimes_{i=1}^n w_i \rangle = \prod_{i=1}^n \langle v_i \,|\, w_i \rangle \tag{5.6}$$

for $v_i, w_i \in \mathbb{R}^2, 1 \leq i \leq n$. With this notation in hand, let

$$\Phi(x, S) = \bigotimes_{i=1}^n \varphi_i(x_i, S) \tag{5.7}$$

for each $x \in \Omega$.

It suffices to verify the following six statements to show linear independence:

(1) $\Phi(y,\emptyset) \perp \Phi(z,[n])$ for all $y \in \bar{U}^c$ and all $z \in U \cap \bar{V}^c$.
(2) $\Phi(y,\emptyset) \perp \Phi(x,S(x))$ for all $y \in \bar{U}^c$ and all $x \in X$.
(3) $\Phi(z,[n]) \perp \Phi(x,S(x))$, for all $z \in U \cap \bar{V}^c$ and all $x \in X$.
(4) $\{\Phi(x,S(x)) : x \in X\}$ is linearly independent.
(5) $\Phi(\bar{U}^c,\emptyset)$ is linearly independent.
(6) $\Phi(U \cap \bar{V}^c,[n])$ is linearly independent.

The function Φ has been defined so that most of this will be routine.

(1) $\Phi(y,\emptyset) \perp \Phi(z,[n])$ for all $y \in \bar{U}^c$ and all $z \in U \cap \bar{V}^c$.
If $y \in \bar{U}^c$ and $z \in U \cap \bar{V}^c$, then $\bar{y} \notin U$ and $z \in U$, so in particular $\bar{y} \neq z$. Then $y_i = z_i$ for some i and

$$\langle \varphi_i(y_i,\emptyset) \,|\, \varphi_i(z_i,[n])\rangle = \langle (y_i,-1) \,|\, (1,z_i)\rangle = 0.$$

Recalling (5.6), we have that

$$\langle \Phi(y,\emptyset) \,|\, \Phi(z,[n])\rangle = 0.$$

Since it is easy to see that neither $\Phi(y,\emptyset)$ nor $\Phi(z,[n])$ can be the zero vector, it follows that $\Phi(y,\emptyset) \perp \Phi(z,[n])$.

(2) $\Phi(y,\emptyset) \perp \Phi(x,S(x))$ for all $y \in \bar{U}^c$ and all $x \in X$.
If $y \in \bar{U}^c$ and $x \in X$, then $\bar{y} \notin U$ which implies there exists $i \in S(x)$ such that $y_i = x_i$. Thus, it follows that

$$\langle \varphi_i(y_i,\emptyset) \,|\, \varphi_i(x_i,S(x))\rangle = \langle (y_i,-1) \,|\, (1,x_i)\rangle = 0.$$

Hence, $\Phi(y,\emptyset) \perp \Phi(x,S(x))$.

(3) $\Phi(z,[n]) \perp \Phi(x,S(x))$, for all $z \in U \cap \bar{V}^c$ and all $x \in X$.
If $z \in U \cap \bar{V}^c$ and $x \in X$, then $\bar{z} \notin V$ which implies there exists $i \in S(x)^c$ such that $z_i = x_i$. It follows that

$$\langle \varphi_i(z_i,[n]) \,|\, \varphi_i(x_i,S(x))\rangle = \langle (1,z_i) \,|\, (x_i,-1)\rangle = 0.$$

Hence $\Phi(z,[n]) \perp \Phi(x,S(x))$.

(4) $\{\Phi(x,S(x)) : x \in X\}$ is a set of linearly independent vectors.
This statement is the core of Reimer's proof. For this argument, it is sufficient to prove the independence on $\mathbb{Z}_2^{2^n}$ rather than $\mathbb{R}^{2^n}$, and, as will become clear, it turns out to be much simpler for $\mathbb{Z}_2^{2^n}$. For the moment, simply note that, in $\mathbb{Z}_2^2$,

if $x_i = 1$ then $\varphi_i(x_i, S) = (1,1)$ whether or not $i \in S$. Notice that since $X \subseteq \Omega$, we can think of $S : x \mapsto S(x)$ as a function from $X \to 2^{[n]}$. We can extend this by defining $S(x) \in \Omega$ for all $x \in \Omega \setminus X$ arbitrarily. This in turn induces a function $x \mapsto \Phi(x, S(x)) : \Omega \to \mathbb{R}^{2^n}$ (or $\mathbb{Z}_2^{2^n}$) which coincides with γ when $x \in X$. In order to prove (4), it is therefore enough to prove that for all $S : \Omega \to 2^{[n]}$, the set $\{\Phi(x, S(x)) : x \in \Omega\}$ is a set of linearly independent vectors in $\mathbb{Z}_2^{2^n}$. This is the content of Lemma 5.2 below.

(5) $\Phi(\bar{U}^c, \emptyset)$ is linearly independent, and
(6) $\Phi(U \cap \bar{V}^c, [n])$ is linearly independent.
Although these can both be argued by completely elementary methods, both of these statements follow as a special case of the statement that for all $S : \Omega \to 2^{[n]} : x \mapsto S(x)$, the set $\{\Phi(x, S(x)) : x \in \Omega\}$ is a set of linearly independent vectors in $\mathbb{R}^{2^n}$ (choose the constant functions $S(x) \equiv \emptyset$ and $S(x) \equiv [n]$, respectively).

The proof of Reimer's Main Lemma is therefore reduced to the proof of the following:

Lemma 5.2. *Let $n \in \mathbb{N}$, and let $\Phi : \{0,1\}^n \times 2^{[n]} \to \mathbb{R}^{2^n}$ be defined by (5.5) and (5.7). Let $S : x \mapsto S(x) \subset [n]$ be an arbitrary function from $\{0,1\}^n$ into $2^{[n]}$. Then the vectors $\Phi(x, S(x))$, $x \in \{0,1\}^n$, are linearly independent in $\mathbb{Z}_2^{2^n}$, and hence in $\mathbb{R}^{2^n}$.*

Proof. For $0 < k \leq 2^n$, let y^k be the configuration in Ω given by the binary representation of $k-1$ so that $\Omega = \{y^k : 0 < k \leq 2^n\}$, with $k = 1$ corresponding to $y_i^k \equiv 0$, $k = 2$ corresponding to $y_n^k = 1$ and $y_i^k = 0$ for all $i \leq n-1$, etc. For the configuration y^k in $\{0,1\}^n$, we let $0y^k$ be the configuration corresponding to the binary representation of $k-1$ in $\{0,1\}^{n+1}$, and $1y^k$ be the configuration corresponding to the binary representation of $2^n + k - 1$ in $\{0,1\}^{n+1}$.

If we let $A_S^{(n)}$ be the $2^n \times 2^n$ matrix formed by letting row k be the vector $\Phi(y^k, S(y^k))$,

$$A_S^{(n)}(k, \cdot) = \Phi(y^k, S(y^k)) \tag{5.8}$$

then it suffices to show that for all functions $S : \Omega \to 2^{[n]}$, the matrix $A_S^{(n)}$ satisfies

$$\det A_S^{(n)} = 1. \tag{5.9}$$

We will prove this using induction on n. The base case $n = 1$ is trivial to check. So suppose that for all $S : \{0,1\} \to 2^{[n]}$ we have $\det A_S^{(n)} = 1$ by induction. Analyzing the case $n+1$, let now $\Omega = \{0,1\}^{n+1}$, and let S be a function from $\{0,1\}^{n+1}$ into $2^{[n+1]}$. Note that the binary representation of each of the first 2^n

configurations begins with 0. So $\varphi_1(y_1^k, S(y^k)) = (1,0)$ or $(0,-1)$ (which equals $(0,1)$ in $\mathbb{Z}_2$), depending on whether $1 \in S(y^k)$ or not. Therefore, defining $S^0 : \{0,1\}^n \to 2^{[n]}$ by $S^0(y^k) = \{i \in [n] : i+1 \in S(0y^k)\}$, we get that for each $0 \le k < 2^n$, either $1 \in S(y^k)$ and

$$\begin{aligned} A_S^{(n+1)}(k,\cdot) &= (1,0) \otimes \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \\ &= \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \oplus \bigoplus_{j=1}^{2^n} 0 \\ &= A_{S^0}^{(n)}(k,\cdot) \oplus \bigoplus_{j=1}^{2^n} 0 \end{aligned}$$

or $1 \notin S(y^k)$ and

$$\begin{aligned} A_S^{(n+1)}(k,\cdot) &= (0,1) \otimes \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \\ &= \bigoplus_{j=1}^{2^n} 0 \oplus \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)). \\ &= \bigoplus_{j=1}^{2^n} 0 \oplus A_{S^0}^{(n)}(k,\cdot) \ . \end{aligned}$$

Defining $\varepsilon_k = 1$ if $1 \in S(y^k)$ and $\varepsilon_k = 0$ if $1 \notin S(y^k)$, we therefore have that

$$A_S^{(n+1)}(k,\cdot) = \varepsilon_k A_{S^0}^{(n)}(k,\cdot) \oplus (1-\varepsilon_k) A_{S^0}^{(n)}(k,\cdot).$$

Meanwhile, note that $(1,-1) = (1,1)$ in $\mathbb{Z}_2^2$, so that $\varphi_1(y_1^k, S(y^k)) = (1,1)$ if the binary representation of k starts with 1. Therefore, defining $S^1 : \{0,1\}^n \to 2^{[n]}$ by $S^1(y^k) = \{i \in [n] : i+1 \in S(1y^k)\}$, we get that for each $2^n < k \le 2^{n+1}$

$$\begin{aligned} A_S^{(n+1)}(k,\cdot) &= (1,1) \otimes \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \\ &= \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \oplus \bigotimes_{i=2}^{n+1} \varphi_i(y_i^k, S(y^k)) \\ &= A_{S^1}^{(n)}(k,\cdot) \oplus A_{S^1}^{(n)}(k,\cdot) \ . \end{aligned}$$

Hence

$$A_S^{(n+1)} = \begin{pmatrix} \varepsilon_k A_{S^0}^{(n)}(k,\cdot) & (1-\varepsilon_k) A_{S^0}^{(n)}(k,\cdot) \\ \\ A_{S^1}^{(n)} & A_{S^1}^{(n)} \end{pmatrix}.$$

Although this matrix looks messy, a few column operations—actually 2^n of them—will improve things, without changing the determinant, of course. By adding column $k+1$ to column $k+1+2^n$ (for each $0 \le k < 2^n$) which, in $\mathbb{Z}_2$, is the same as subtracting column $k+1$ from column $k+1+2^n$, we can conclude that

$$\begin{aligned} \det A_S^{(n+1)} &= \det \begin{pmatrix} \varepsilon_k A_{S^0}^{(n)}(k,\cdot) & A_{S^0}^{(n)}(k,\cdot) \\ \\ A_{S^1}^{(n)} & 0 \end{pmatrix} \\ &= \det A_{S^0}^{(n)} \det A_{S^1}^{(n)} \\ &= 1, \end{aligned}$$

where the final step follows by induction. □

This completes the proof of Reimer's Main Lemma, and hence the proof of the BK conjecture, Theorem 2.1.

Remarks.

i) The proof given above for independence of $\{\Phi(x, S(x)) : x \in \Omega\}$ follows quite closely that presented in [CPS], which was in turn based on the matrix proof given by Reimer. Alternatively, there is a more algebraic proof, similar to one presented by J. Kahn and also to one suggested to us by P. Deligne. At the end of this section we give a version of such a proof, originally presented by one of us (J.T.C.) in the Kac Seminars of 1995, and reviewed in [CPS].

ii) A close analysis of the above proof of Lemma 5.2 shows that it holds in the more general case in which (5.5) is replaced by the definition

$$\varphi_i(x_i, S) = \begin{cases} \psi_0(x_i) & \text{if } i \notin S \\ \psi_1(x_i) & \text{if } i \in S, \end{cases}$$

provided the four vectors $\psi_0(0)$, $\psi_0(1)$, $\psi_1(0)$ and $\psi_1(1) \in \mathbb{Z}_2^2$ are chosen in such a way that for each pair of vectors $\psi(0) \in \{\psi_0(0), \psi_1(0)\}$ and $\psi(1) \in \{\psi_0(1), \psi_1(1)\}$, the set $\{\psi(0), \psi(1)\}$ is a basis for $\mathbb{Z}_2^2$. This is easy to see once it is realized that all possible cases can be reduced to the following three cases:

(a) $\psi_0(0) = (0, 1)$, $\psi_1(0) = (1, 0)$, and $\psi_0(1) = \psi_1(1) = (1, 1)$, the case studied in the proof of Lemma 5.2;

(b) $\psi_0(0) = (0, 1)$, $\psi_1(0) = (1, 1)$, and $\psi_0(1) = \psi_1(1) = (1, 0)$, and

(c) $\psi_0(0) = (1,0)$, $\psi_1(0) = (1,1)$, and $\psi_0(1) = \psi_1(1) = (0,1)$.
Revisiting the proof of Lemma 5.2, it can be easily seen that the inductive proof works for cases (b) and (c) in a similar way to case (a). We therefore obtain the following generalization of Lemma 5.2:

Lemma 5.3. *Let $\psi_0(0)$, $\psi_0(1)$, $\psi_1(0)$ and $\psi_1(1) \in \mathbb{Z}_2^2$ be chosen in such a way that for each pair of vectors $\psi(0) \in \{\psi_0(0), \psi_1(0)\}$ and $\psi(1) \in \{\psi_0(1), \psi_1(1)\}$, the set $\{\psi(0), \psi(1)\}$ is a basis for $\mathbb{Z}_2^2$. Let $n \in \mathbb{N}$, let S be an arbitrary function from $\{0,1\}^n$ into $2^{[n]}$, and let*

$$\varphi_i(x_i, S(x)) = \begin{cases} \psi_0(x_i) & \text{if } i \notin S(x) \\ \psi_1(x_i) & \text{if } i \in S(x). \end{cases} \tag{5.10}$$

and

$$\Phi_S(x) := \bigotimes_{i=1}^{n} \varphi_i(x_i, S(x)) \,. \tag{5.11}$$

Then the set $\{\Phi_S(x) : x \in \{0,1\}^n\}$ is a set of linearly independent vectors in $Z_2^{2^n}$.

Remarks. iii) If the set $S(x) \subset [n]$ is independent of $x \in \{0,1\}^n$, the statement of the lemma just follows from the well known fact that whenever $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ are bases for two vector spaces V and W, then $\{v_i \otimes w_j\}_{(i,j)\in[n]\times[m]}$ is a basis for $V \otimes W$.

iv) The generalization of Lemma 5.3 to $\mathbb{Z}^2$ or $\mathbb{R}^2$ is false. Indeed, taking $n = 2$ and choosing $\psi_0(0) = (0,1)$, $\psi_1(0) = (1,1)$, $\psi_0(1) = (1,0)$, $\psi_1(1) = (-1,1)$, $S(00) = S(11) = \{2\}$ and $S(01) = S(10) = \{1\}$, we find that the four vectors

$$\begin{aligned}
\Phi_S(00) &= \psi_0(0) \otimes \psi_1(0) = \quad (0,1) \otimes (1,1) \quad = (0,0,1,1)\,, \\
\Phi_S(01) &= \psi_1(0) \otimes \psi_0(1) = \quad (1,1) \otimes (1,0) \quad = (1,0,1,0)\,, \\
\Phi_S(10) &= \psi_1(1) \otimes \psi_0(0) = (1,-1) \otimes (0,1) \quad = \ (0,1,0,-1), \ \text{ and} \\
\Phi_S(11) &= \psi_0(1) \otimes \psi_1(1) = \quad (1,0) \otimes (1,-1) = (1,-1,0,0)
\end{aligned}$$

are linearly dependent, since $(0,0,1,1) - (1,0,1,0) + (0,1,0,-1) + (1,-1,0,0) = 0$.

We close this paper with the alternative proof of Lemma 5.2 mentioned in Remark (i) above.

Alternative Proof of Lemma 5.2. As in the above proof, we will establish independence on $\mathbb{Z}_2^{2^n}$ rather than $\mathbb{R}^{2^n}$. For convenience, we define $\mathbf{w} = (1,0)$ and

$\mathbf{v} = (1,1)$. Then $\varphi_i(x_i,S) = \mathbf{w}$ if $x_i = 0$ and $i \in S$, $\varphi_i(x_i,S) = \mathbf{v}$ if $x_i = 1$, and $\varphi_i(x_i,S) = \mathbf{w} + \mathbf{v}$ if $x_i = 0$ and $i \notin S$.

To show that the vectors in $\{\Phi(x,S(x)) : x \in \Omega\}$ are linearly independent, it suffices to expand them in a basis in $\otimes_{i=1}^n \mathbb{Z}_2^2$ and show that the coefficient matrix has nonzero determinant. To this end, let I be the set $I \subset [n]$. Our basis in $\otimes_{i=1}^n \mathbb{Z}_2^2$ will be $\{u_I \mid I \subset [n]\}$ where

$$u_I = \bigotimes_{j \notin I} \mathbf{v} \otimes \bigotimes_{j \in I} \mathbf{w}\,.$$

In order to expand $\Phi(x,S(x))$ in the $\{u_I\}$, we let $I_y = \{i \mid y_i = 0\}$. Then

$$\begin{aligned}
\Phi(x,S(x)) &= \bigotimes_{i=1}^{n} \varphi_i(x_i,S(x)) \\
&= \bigotimes_{i \notin I_y} \mathbf{v} \otimes \bigotimes_{i \in I} \begin{cases} \mathbf{w} & \text{if } x_i = 0 \\ \mathbf{w}+\mathbf{v} & \text{if } x_i = 1 \end{cases} \\
&= \bigotimes_{i \notin I_y} \mathbf{v} \otimes \bigotimes_{i \in I_y} (\mathbf{w} + x_i \mathbf{v}) \\
&= \sum_{J \subseteq I_y} \bigotimes_{i \notin I_y} \mathbf{v} \otimes \bigotimes_{i \in J} \mathbf{w} \otimes \bigotimes_{i \in I_y \setminus J} x_i \mathbf{v}\,.
\end{aligned}$$

Noting that $I_y^c \cup I_y \setminus J = J^c$, and defining $\epsilon(J) =: \prod_{I_y \setminus J} x_i \in \{0,1\}$, we have

$$\begin{aligned}
\Phi(x,S(x)) &= \sum_{J \subseteq I_y} \epsilon(J) \bigotimes_{i \notin J} \mathbf{v} \otimes \bigotimes_{i \in J} \mathbf{w} \\
&= \sum_{J \subseteq I_y} \epsilon(J)\, u_J \\
&= u_{I_y} \; + \sum_{J \subset I_y} \epsilon(J)\, u_J\,,
\end{aligned}$$

where J is a proper subset of I_y in the final sum.

Now the above matrix is an upper triangular matrix with 1's along the diagonal. If the index set I_y were a totally ordered set, this would immediately imply the determinant of this matrix is one, and hence that the vectors $\Phi(x,S(x))$, $x \in \Omega$ are linearly independent. Since the index set is only partially ordered, this requires a little additional argument, which we leave to the reader. It is easy to verify using e.g. the expansion of the determinant in minors. $\square$

References

[Be] J. van den Berg, *private communication.*

[BF] J. van den Berg and U. Fiebig, *On a combinatorial conjecture concerning disjoint occurrences of events*, Ann. Probab. **15** (1987), 354–374.

[BK] J. van den Berg and H. Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. **22** (1985), 556–569.

[CPS] J. T. Chayes, A. Puha and T. Sweet, *Independent and dependent percolation*, IAS/Park City Mathematics Series, Vol. 6, AMS, Providence, 1998.

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

[FS] P.C. Fishburn and L.A. Shepp, *On the FKB conjecture for disjoint intersections*, Discrete Math. **98** (1991), 105–122.

[Re] D. Reimer, *preliminary manuscript* (1994).

[Ta] M. Talagrand, *Some remarks on the Berg-Kesten inequality, Probab. in Banach Spaces, 9 (Sandjberg, 1993)*, Prog. Probab. **35** (1994), Birkhäuser, Boston, 293–297.

CHRISTIAN BORGS
1 MICROSOFT WAY,
REDMOND, WA 98052
E-mail address: borgs@microsoft.com

JENNIFER T. CHAYES
1 MICROSOFT WAY,
REDMOND, WA 98052
E-mail address: jchayes@microsoft.com

DANA RANDALL
SCHOOL OF MATHEMATICS AND COLLEGE OF COMPUTING
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA GA 30332-0160
E-mail address: randall@math.gatech.edu