Bernoulli **24**(1), 2018, 433–448
DOI: 10.3150/16-BEJ883

The van den Berg–Kesten–Reimer operator and inequality for infinite spaces

RICHARD ARRATIA[1], SKIP GARIBALDI[2,3] and ALFRED W. HALES[3]

[1]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA. E-mail: rarratia@usc.edu*
[2]*UCLA Institute for Pure and Applied Mathematics, Box 957121, Los Angeles, CA 90095-7121, USA. E-mail: skip@member.ams.org*
[3]*Center for Communications Research, San Diego, CA 92121, USA*

We remove the hypothesis "S is finite" from the BKR inequality for product measures on S^d, which raises some issues related to descriptive set theory. We also discuss the extension of the BKR operator and inequality, from 2 events to 2 or more events, and we remove, in one sense, the hypothesis that d be finite.

Keywords: BKR inequality; projective set; van den Berg–Kesten–Reimer

1. The classic BKR inequality

The BKR inequality, named for van den Berg–Kesten–Reimer, was conjectured in [19] and proved in [18] and [15]; see [5] or [4] for a clear exposition. The setup involves a probability space of the form S^d, with S finite, and $\mathbb{P}$ a product measure, and the inequality takes the form: for two events $A, B \subset S^d$, with $A \mathbin{\Box} B$ for the event that, informally, "A and B occur for disjoint reasons,"

$$\mathbb{P}(A \mathbin{\Box} B) \le \mathbb{P}(A)\mathbb{P}(B).$$

The somewhat convoluted history is summarized as follows: Kesten and van den Berg [19] defined the operation $A \mathbin{\Box} B$ on subsets of S^d, and proved the (BK) inequality for the special case where A and B are assumed to be increasing events. Then van den Berg and Fiebig [18] proved a conditional implication, not involving increasing events: "If the inequality holds for the cases $S^d = \{0, 1\}^d$ and $\mathbb{P}$ is the uniform distribution, with all 2^d points of S^d equally likely, then the inequality holds for any finite S and any product measure on S^d." Finally, Reimer [15] proved the inequality in $\{0, 1\}^d$, a purely combinatorial fact, so that combined with the earlier conditional implication from [18], the general inequality was established.

In [2], still in the context of S finite and $\mathbb{P}$ a product measure on S^d, we had a Florida-lottery-crimefighting reason to need an extension of the BKR inequality, from $r = 2$ events, to the more general case $r = 2, 3, \ldots$. An easy example shows that sometimes $(A \mathbin{\Box} B) \mathbin{\Box} C \ne A \mathbin{\Box} (B \mathbin{\Box} C)$, so we gave a natural definition for the r-fold operator $\bigsqcap_1^r A_i$, proved that $\bigsqcap_1^r A_i \subset (\cdots((A_1 \mathbin{\Box} A_2) \mathbin{\Box} A_3) \cdots \mathbin{\Box} A_r)$, and gave the easy induction, from the classic BKR inequality, to conclude that

$$\mathbb{P}\left(\square_1^r A_i\right) \le \prod_1^r \mathbb{P}(A_i).$$

Although the case S finite was sufficient for our application, it seemed strange to have to quote the hypothesis "S is finite," before invoking the inequality. Indeed, the first draft of [2] made the mistake of omitting this hypothesis – but thankfully was called to the carpet by a referee.

In this paper, we remove the restriction that S be finite, allowing $S = \mathbb{N}$ or $S = \mathbb{R}$, along with an arbitrary product probability measure on S^d, for our main result, Theorem 7. This raises issues related to descriptive set theory; the BKR combination of Borel sets need not be a Borel set, and the BKR combination of Lebesgue measurable sets need not be Lebesgue measurable, see Example 3. We will also, in Section 7, remove the restriction that d be finite, for one of the two natural ways of generalizing the BKR operator to spaces of the form $S^{\mathbb{N}}$.

Other extensions and complements to the BKR inequality are given in [1,9,11]. In greater detail, [9] gives a generalization of the BKR operator and inequality which applies to spaces such as $\mathbb{R}^d$; however, the combination of sets which [9], formula (5), identifies as "A and B occur for disjoint reasons" is somewhat different from the original BKR combination $A \square B$, and depends on the choice of measure and notions of essential infimum. It is easy to see the BKR combination of events from [9] is a superset of the standard $A \square B$, hence the result from [9], with the corrections and improvements provided in [10], proves the outer measure assertion in our Theorems 4 and 7, via a method which finesses all issues of projective sets by an appeal to Tonelli's theorem. In contrast to their approach, ours extends the BKR operator and inequality to infinite spaces in a way that closely follows the original definitions, meaning as a combination of events, rather than a combination of events and measures.

2. Definition of the BKR operators

The formal definition of $A \square B$, copied from [19], begins with the notation $\omega = (\omega_1, \ldots, \omega_d)$ or $\overline{\omega} = (\overline{\omega}_1, \ldots, \overline{\omega}_d)$ for elements of S^d. For $\omega \in S^d$ and $K \subset [d] := \{1, \ldots, d\}$, consider the *thin cylinder* $\mathrm{Cyl}(K, \omega) := \{\overline{\omega} : \overline{\omega}_i = \omega_i, i \in K\}$. For $A, B \subset S^d$ define $A \square B$ as the set of ω for which there exists a $K \subset [d]$ such that $\mathrm{Cyl}(K, \omega) \subset A$ and $\mathrm{Cyl}(K^c, \omega) \subset B$, where $K^c := [d] \setminus K$ is the complement of K relative to the universe of coordinate indices.

A small paraphrase of this definition is based on $[A]_K$ defined to be the largest cylinder set[1] contained in A and free in the directions indexed by K^c:

$$\text{for } A \subset S^d, \qquad [A]_K := \{\omega : \mathrm{Cyl}(K, \omega) \subset A\}. \tag{1}$$

[1]Both $\mathrm{Cyl}(K, \omega)$ and $[A]_K$ are defined relative to S^d. We have several occasions in this paper to work simultaneously with two different sets in the role of S, and it should be understood that the definition of the BKR operator for sets $A, B \subset S^d$ also involves the choice of S and d. Apart from Section 8, we use the same symbol $\square$ for every operator of this form, and leave it to the reader to understand the appropriate context.

With this notation,

$$\text{for } A, B \subset S^d, \qquad A \,\square\, B := \bigcup_{K \subset [d]} [A]_K \cap [B]_{K^c}. \tag{2}$$

An obvious relation, that $J \subset K \subset [d]$ implies $[A]_J \subset [A]_K$, shows that (2) is equivalent to the following:

$$\text{for } A, B \subset S^d, \qquad A \,\square\, B := \bigcup_{\text{disjoint } J, K \subset [d]} [A]_J \cap [B]_K. \tag{3}$$

The definition of the simultaneous r-fold BKR operator given in [2] is, for $A_1, \ldots, A_r \subset S^d$,

$$\bigsqcup_{1 \le i \le r} A_i \equiv A_1 \,\square\, A_2 \,\square \cdots \square\, A_r := \bigcup_{J_1, \ldots, J_r} [A_1]_{J_1} \cap [A_2]_{J_2} \cap \cdots \cap [A_r]_{J_r}, \tag{4}$$

where the union is taken over disjoint subsets $J_1, \ldots, J_r$ of $\{1, \ldots, d\}$. It is clear that for the case $r = 2$, definition (4) agrees with (3), and hence with (2).

2.1. Careful notation for cylinders, projections, extensions

We follow the strict convention that, for any sets U, V, the set U^V is the set of all functions from V to U, and an element $f \in U^V$ carries the information: what is the domain of f, and what is the range of f. For the case $V = \varnothing$, there is one point exactly in U^V. Since we use the notational convention, common in combinatorics, that for $d = 0, 1, 2, \ldots,$ $[d] := \{1, 2, \ldots, d\}$, the d-fold Cartesian product of a set S with itself, S^d, is exactly equal to $S^{[d]}$. But for $0 \le k \le d$, there are $\binom{d}{k}$ subsets $K \subset [d]$, with $|K| = k$, and there are $\binom{d}{k}$ different sets S^K; only one of these is equal to S^k, namely, the one with $K = [k]$.

It will be convenient to work first with the case $S = [0, 1]$, allowing us to specialize to the uniform distribution.

For $K \subset [d]$, the projection

$$\text{proj}_K : [0, 1]^{[d]} \to [0, 1]^K$$

is, naturally, the function $f \mapsto f|_K$ which restricts a function $f \in [0, 1]^{[d]}$ to have domain K. There is a single one-to-many *relation* ext_d, with domain $\bigcup_{K \subset [d]} [0, 1]^K$, which serves as the inverse for *all* of maps proj_K, namely, $(g, f) \in \text{ext}_d$ if and only if, for some $K \subset [d]$, $g \in [0, 1]^K$, $f \in [0, 1]^{[d]}$, and $g = f|_K$.

For any set D, we write 2^D for the power set of D, that is, the set of all subsets of D. We will be fussy, to distinguish a function from D to D', and its inverse relation, written with lowercase, from the induced functions, mapping 2^D to $2^{D'}$ and back, written with uppercase.

Thus, we have 2^d projection functions

$$\text{Proj}_K : 2^{[0,1]^{[d]}} \to 2^{[0,1]^K},$$

and a single extension function,

$$\mathrm{Ext}_d : 2^{\bigcup_K [0,1]^K} \to 2^{[0,1]^{[d]}}.$$

In particular, for $K \subset [d]$,

$$\text{for } C \subset [0,1]^K, \qquad \mathrm{Ext}_d(C) := \{ f \in [0,1]^{[d]} : f|_K \in C \} = \mathrm{Proj}_K^{-1}(C).$$

3. Measurability considerations

3.1. Introductory motivation

In 1905, Lebesgue stated, incorrectly, that projections of Borel sets are Borel sets, and Souslin showed otherwise. Superficially, this is an obstacle to extending the BKR inequality from S^d with S countable to the case with $S = [0,1]$, since in $[0,1]^d$, even starting with Borel sets A, B, we cannot assert that $A \square B$ is also a Borel set. In more detail,

$$A \square B := \bigcup_{K \subset [d]} [A]_K \cap [B]_{K^c},$$

where $[A]_K$ is the maximal cylinder subset of A free in the directions in $[d] \setminus K$, equivalently, using notation from Section 2.1,

$$[A]_K := \mathrm{Ext}_d\big([0,1]^K \setminus \mathrm{Proj}_K\big(A^c\big)\big). \tag{5}$$

However, Lusin and Sierpiński showed that projections of Borel sets are nice, in the concrete sense of having equal inner and outer measure, that is, being measurable in the *completion* of the Borel sigma-algebra with respect to Lebesgue measure [6], Theorem 8.4.1. For history, see [7], page 500, [14], page 232, [12,13].

We write λ_d for the usual Lebesgue measure on $[0,1]^d$. The $\binom{d}{k}$ different spaces $[0,1]^K$, for $K \subset [d]$ with $|K| = k$, are all naturally measure isomorphic to $[0,1]^k$. Rather than writing the explicit isomorphism, or naming the corresponding copies of Lebesgue measure λ_K, we simply write λ_k.

With $k = |K|$, if $C \subset [0,1]^K$ is Lebesgue measurable then so is $\mathrm{Ext}_d(C)$, and $\lambda_k(C) = \lambda_d(\mathrm{Ext}_d(C))$.

3.2. Details for measurability

Lemma 1. *Assume that A is a Borel subset of $[0,1]^d$, and that $K \subset [d]$. Then, the cylinder $[A]_K$ is Lebesgue measurable. If A, B are Borel subsets of $[0,1]^d$, then $A \square B$ is Lebesgue measurable, and if $A_1, \dots, A_r$ are Borel subsets of $[0,1]^d$, then $\square_1^r A_i$ is Lebesgue measurable.*

Proof. To start, A is a Borel subset of $[0,1]^d$ so $A^c := [0,1]^d \setminus A$ is also Borel, and the projection $C := \mathrm{Proj}_K(A^c)$ is an analytic subset of $[0,1]^K$. In particular, C^c is universally measurable, and the pullback $[A]_K$ of C^c is measurable with respect to the product measure λ_d.

The Lebesgue measurability claims for $A \,\square\, B$ and $\square_1^r A_i$ now follow immediately from the definitions (2) and (4). $\square$

Lemma 1 remains true if "Borel" is replaced by "co-analytic," as is clear from the proof.

Remark 2. Similarly, if A is open, then the continuous image C of the compact set A^c is closed, and therefore $[A]_K$ is open. And if A, B are both open, then $A \,\square\, B$ is open.

The following example shows why, in Corollary 5, with the hypothesis that A and B are Lebesgue measurable, we could not simply state that $\lambda_d(A \,\square\, B) \le \lambda_d(A)\lambda_d(B)$.

Example 3. The BKR combination of Lebesgue measurable sets need not be Lebesgue measurable, as shown by this example with $d = 2$. Take a set $C \subset [0,1]$ which is *not* a Lebesgue measurable subset of $[0,1]$. Then $C^2 \subset [0,1]^2$ is *not* a Lebesgue measurable subset of $[0,1]^2$. The diagonal in $[0,1]^2$ is

$$D := \big\{(x,x) : x \in [0,1]\big\} \subset [0,1]^2,$$

and this is a Borel subset of $[0,1]^2$, with $\lambda_2(D) = 0$. Hence the set

$$E := \big\{(x,x) : x \in \big([0,1] \setminus C\big)\big\} \subset D \subset [0,1]^2$$

is Lebesgue measurable, with $\lambda_2(E) = 0$. Now, taking complement relative to $[0,1]^2$, let

$$A := [0,1]^2 \setminus E,$$

so that A is Lebesgue measurable, with $\lambda_2(A) = 1$. We have

$$[A]_{\{1\}} = C \times [0,1], \qquad [A]_{\{2\}} = [0,1] \times C,$$

and with $B := A$ we have

$$A \,\square\, B = C^2.$$

4. Approximation, from [0, 1] to a finite set

Theorem 4. *For Borel subsets A, B in $[0,1]^d$,*

$$\lambda_d(A \,\square\, B) \le \lambda_d(A)\lambda_d(B).$$

4.1. Overview of the argument

We want to prove that, for Borel $A, B \subset [0,1]^d$, we have $\lambda_d(A \,\square\, B) \leq \lambda_d(A)\lambda_d(B)$, and we proceed by contradiction. Thus, we assume that we have A, B with

$$4\varepsilon := \lambda_d(A \,\square\, B) - \lambda_d(A)\lambda_d(B) > 0, \tag{6}$$

and we work to provide an example, with finite S, in which the classic BKR inequality on S^d is violated.

In this example, for some large but finite n, we have $|S| = 2^n$, $|S^d| = 2^{nd}$, corresponding to the number of atoms in the "observe the first n bits" sigma-algebra $\mathcal{F}_n^{(d)}$ on $[0,1]^d$. The product measure $\mathbb{P}$ on S^d will be the uniform distribution, with mass 2^{-nd} at each point of S^d. We will produce subsets $A'', B'' \subset S^d$ for which

$$\mathbb{P}\big(A'' \,\square\, B''\big) \geq \lambda_d(A \,\square\, B) - \varepsilon \tag{7}$$

and

$$\mathbb{P}\big(A''\big) \leq \lambda_d(A) + \varepsilon, \qquad \mathbb{P}\big(B''\big) \leq \lambda_d(B) + \varepsilon, \tag{8}$$

so that A'', B'' violate the classic BKR inequality.

4.2. Set approximation, in 1 dimension

To lighten the notational burden, we start with dimension 1, and review a familiar martingale, from for example [3], Examples 35.3, 35.10. The probability space is $[0,1]$, with the Borel sigma-algebra, and the probability measure is λ_1. For $n = 0, 1, 2, \ldots$, define $\mathcal{F}_n$ to be the sigma-algebra generated by the 2^n disjoint intervals, $[0, 1/2^n)$, $[1/2^n, 2/2^n), \ldots,$ $[(n-2)/2^n, (n-1)/2^n)$, $[1 - 1/2^n, 1]$. Note that the last of these intervals is exceptional, in that it is closed at both ends, but all 2^n intervals I have length $\lambda_1(I) = 1/2^n$. The sigma-algebra $\mathcal{F}_n$ has 2^n atoms, and is a family of 2^{2^n} subsets of $[0,1]$. These sigma-algebras are nested, and $\sigma(\bigcup_{n\geq 0}\mathcal{F}_n)$ is the usual Borel sigma-algebra on $[0,1]$.

Hence for any Borel measurable $h : [0,1] \to [0,1]$, $M_n := \mathbb{E}(h|\mathcal{F}_n)$ is a martingale. Explicitly, on an atom I of $\mathcal{F}_n$, $M_n = 2^n\mathbb{E}(h; I) = 2^n \int_I h(x)\,dx$. The martingale convergence theorem implies that M_n converges to h, almost surely and in L_1, with the L_1 convergence meaning that $\mathbb{E}|M_n - h| \to 0$ as $n \to \infty$.

In particular, given a Borel measurable $C \subset [0,1]$, we take h to be the indicator function $h = 1_C$. Explicitly, on an atom I of $\mathcal{F}_n$, $M_n = 2^n\lambda_1(C \cap I)$. From this martingale, we round values in $[0, 1/2]$ down to 0, and values in $(1/2, 1]$ up to 1, to get a deterministic set $C_n \in \mathcal{F}_n$. Explicitly,

$$C_n := \big\{\omega \in [0,1] : M_n(\omega) > 1/2\big\}.$$

For a point x to be in the symmetric difference set, $C\Delta C_n$, the rounding error is at least one half. This implies that $\lambda_1(C\Delta C_n) \leq 2\mathbb{E}|M_n - 1_C|$.

4.3. Set approximation, in k dimensions

The above extends to dimension k, for $k = 1, 2, \ldots,$ with no difficulties, only extra notation. The probability space is $[0, 1]^k$, with λ_k serving as the probability measure. We define, for $n = 0, 1, 2, \ldots,$ the analogous sigma-algebra $\mathcal{F}_n^{(k)}$ with 2^{nk} atoms, and for any Borel measurable set $C \subset [0, 1]^k$, the martingale argument gives us determinstic sets C_n, with

$$\lambda_k(C \Delta C_n) \to 0, \tag{9}$$

and C_n is $\mathcal{F}_n^{(k)}$ measurable.

4.4. Approximation in $[0, 1]^d$ to control BKR ingredients

Recall that for $A \subset [0, 1]^d$ and $K \subset [d]$, $[A]_K \subset [0, 1]^d$ is the (maximal) cylinder subset of A, in the directions not restricted by K.

We write

$$[[A]]_K := \mathrm{Proj}_K\big([A]_K\big) = [0, 1]^K \setminus \mathrm{Proj}_K\big(\big[A^c\big]_K\big) \subset [0, 1]^K \tag{10}$$

for the base of this cylinder. From the proof of Lemma 1, $[[A]]_K$ is Lebesgue measurable, and there is a Borel subset $C \subset [0, 1]^K$ with

$$C \subset [[A]]_K, \qquad \lambda_k(C) = \lambda_k\big([[A]]_K\big). \tag{11}$$

Observe that, with $1 \le k = |K| < d$,

$$[A]_K = \mathrm{Ext}_d\big([[A]]_K\big) \supset \mathrm{Ext}_d(C)$$

and

$$\lambda_d\big([A]_K\big) = \lambda_k\big([[A]]_K\big) = \lambda_k(C) = \lambda_d\big(\mathrm{Ext}_d(C)\big).$$

Taking A or B, and $K \subset [d]$, we have 2^{d+1} instances of a set $C \subset [[A]]_K$ or $C \subset [[B]]_K$, with $0 \le k := |K| \le d$, to serve as the target for an approximation as given by the martingale argument, summarized by (9). Since

$$A \,\square\, B = \bigcup_K [A]_K \cap [B]_{K^c}, \tag{12}$$

has 2^{d+1} ingredients, we take

$$\delta := \varepsilon / 2^{d+1},$$

and pick a single value of n so that for each of the instances of C,

$$\lambda_k(C \Delta C_n) < \delta. \tag{13}$$

When $C \subset [[A]]_K$, the dyadic approximation C_n is a subset of $[0,1]^K$, and we write

$$A_{n,K} := \mathrm{Ext}_d(C_n) \subset [0,1]^d$$

for the cylinder set whose base is C_n. Thus, with similar notation for B and approximations $B_{n,K}$ to $[B]_K$, we have, from (11) and (13), that

$$\lambda_d\big([A]_K \setminus A_{n,K}\big) < \delta, \qquad \lambda_d\big([B]_K \setminus B_{n,K}\big) < \delta, \tag{14}$$

and since $[A]_K \subset A$, $[B]_K \subset B$,

$$\lambda_d(A_{n,K} \setminus A) < \delta, \qquad \lambda_d(B_{n,K} \setminus B) < \delta. \tag{15}$$

Note also that $A_{n,K}, B_{n,K} \in \mathcal{F}_n^{(d)}$. We take

$$A' := \bigcup_K A_{n,K}, \qquad B' := \bigcup_K B_{n,K}, \tag{16}$$

so that

$$A', B' \in \mathcal{F}_n^{(d)},$$

and for every K, $[A']_K \supset A_{n,K}$, similarly for B, so that by (14),

$$\lambda_d\big([A]_K \setminus [A']_K\big) < \delta, \qquad \lambda_d\big([B]_K \setminus [B']_K\big) < \delta. \tag{17}$$

Using (17),

$$\lambda_d\big([A']_K \cap [B']_{K^c}\big) > \lambda_d\big([A]_K \cap [B]_{K^c}\big) - 2\delta,$$

and hence for the unions, with 2^d values for K, using $2^{d+1}\delta = \varepsilon$,

$$\lambda_d\big(A' \square B'\big) > \lambda_d(A \square B) - \varepsilon.$$

To get an inequality in the opposite direction, combining (15) with (16),

$$\lambda_d\big(A' \setminus A\big) < 2^d\delta < \varepsilon \qquad \text{hence } \lambda_d\big(A'\big) < \lambda_d(A) + \varepsilon,$$

and similarly $\lambda_d(B') < \lambda_d(B) + \varepsilon$.

Finally, since $A', B' \in \mathcal{F}_n^{(d)}$, we take equivalence classes modulo the atoms of $\mathcal{F}_n^{(d)}$, to produce our sets $A'', B'' \in S^d$ for S with $|S| = 2^n$, to get the example satisfying (7) and (8). This completes a proof of Theorem 4.

Corollary 5. *For Lebesgue measurable $A, B \subset [0,1]^d$, there exists a Borel set C, with*

$$(A \square B) \subset C \quad \text{and} \quad \lambda_d(C) \le \lambda_d(A)\lambda_d(B).$$

Proof. Take Borel sets $A_1, B_1 \subset [0,1]^d$ with $A \subset A_1$, $B \subset B_1$, and $\lambda_d(A_1) = \lambda_d(A)$, $\lambda_d(B_1) = \lambda_d(B)$. Obviously $A \square B \subset A_1 \square B_1$, and Theorem 4 implies the existence of a Borel set C with $A_1 \square B_1 \subset C$ and $\lambda_d(C) \le \lambda_d(A_1)\lambda_d(B_1)$. $\square$

5. Extension to 3 or more events

Theorem 6. *For Borel subsets $A_1, \ldots, A_r$ in $[0,1]^d$,*

$$\lambda_d\left(\mathop{\Box}_1^r A_i\right) \le \prod_1^r \lambda_d(A_i). \tag{18}$$

For Lebesgue measurable $A_1, \ldots, A_r$ in $[0,1]^d$, there exists a Borel set D with $\Box_1^r A_i \subset D$ and $\lambda_d(D) \le \prod \lambda_d(A_i)$.

Proof. Define sets $B_1, B_2, \ldots, B_r \subset [0,1]^d$, Lebesgue measurable sets $C_1, C_2, \ldots, C_r \subset [0,1]^d$, and Borel sets $D_1, D_2, \ldots, D_r$ recursively, with

$$A_1 = B_1 = C_1 = D_1$$

and for $i = 2$ to r, using Lemma 1,

$$\begin{aligned} B_i &= B_{i-1} \Box A_i, \\ C_i &= D_{i-1} \Box A_i, \\ D_i \text{ is a Borel set with } C_i \subset D_i, \qquad \lambda_d(C_i) &= \lambda_d(D_i). \end{aligned}$$

The BKR monotonicity relation that $B \subset D$ implies $B \Box A \subset D \Box A$, and induction, shows that for all i, $B_i \subset C_i \subset D_i$. We check that C_i is Lebesgue measurable by noting the it is the BKR combination of two Borel sets, namely D_{i-1} and A_i.

Theorem 4 implies that $\lambda_d(C_i) \le \lambda_d(D_{i-1})\lambda_d(A_i)$, and together with the defining property of D_i this yields

$$\lambda_d(C_i) \le \lambda_d(C_{i-1})\lambda_d(A_i)$$

and it follows by induction that $\lambda_d(C_r) \le \prod_1^r \lambda_d(A_i)$.

It is shown in [2] that $\Box_1^r A_i \subset B_r$.

Combined with $B_r \subset C_r$, we have $\Box_1^r A_i \subset C_r$. Lemma 1 shows that $\Box_1^r A_i$ is Lebesgue measurable, so we have proved (18).

The case with Lebesgue measurable inputs $A_1, \ldots, A_r$ now follows from the Borel case, by the same reasing used to derive Corollary 5 from Theorem 4. □

6. Extension of the BKR inequalities to $\mathbb{R}^d$

Say we are given a product probability measure $\mathbb{P}$ on $\mathbb{R}^d$. This is equivalent to saying that $\mathbb{P}$ is the law, with the Borel sigma-algebra on $\mathbb{R}^d$, of $\mathbf{X} = (X_1, \ldots, X_d)$, with $X_1, X_2, \ldots, X_d$ mutually independent, *and* with *some* given marginal distributions – given by, say, the cumulative distribution functions F_i, where $F_i(t) := \mathbb{P}(X_i \le t)$ for $-\infty < t < \infty$. Let G_i be what is commonly called "F_i^{-1}, the inverse cumulative distribution function for X_i," or "the quantile function for the distribution of X_i." Specifically, we take the domain of G_i to be (0, 1), and for $0 < u < 1$,

$$G_i(u) := \sup\bigl\{x : \mathbb{P}(X_i \le x) \le u\bigr\},$$

this being a choice that makes $G_i(\cdot)$ right-continuous. It is standard to use this in a coupling: with U uniformly distributed in (0, 1), $G_i(U)$ is equal in distribution to X_i.

The net effect of this is to reassure the reader we have no claim to originality, if we *define*

$$g : (0,1)^d \to \mathbb{R}^d, \qquad \mathbf{u} = (u_1, \ldots, u_d) \mapsto \mathbf{x} := \bigl(G_1(u_1), \ldots, G_d(u_d)\bigr). \tag{19}$$

Also, it is *obvious* that under the uniform distribution on $(0, 1)^d$, $g(\omega)$ is equal in distribution to $\mathbf{X}$, that is, for every Borel set A in $\mathbb{R}^d$, $\lambda_d(g^{-1}(A)) = \mathbb{P}(A)$.

Theorem 7. *For Borel subsets A, B of $\mathbb{R}^d$, under any complete product probability measure $\mathbb{P}$ on $\mathbb{R}^d$,*

$$\mathbb{P}(A \,\square\, B) \le \mathbb{P}(A)\mathbb{P}(B). \tag{20}$$

For Borel subsets $A_1, \ldots, A_r$ of $\mathbb{R}^d$, under any complete product probability measure $\mathbb{P}$ on $\mathbb{R}^d$,

$$\mathbb{P}\left(\mathop{\Box}_{1}^{r} A_i\right) \le \prod_{1}^{r} \mathbb{P}(A_i). \tag{21}$$

Proof. The map g defined by (19) is Borel measurable. Since the ith coordinate of $g(\mathbf{u})$ depends only on u_i, the BKR operators respect g, that is,

$$\text{for } a := g^{-1}(A), \qquad b := g^{-1}(B) \subset [0,1]^d, \qquad a \,\square\, b = g^{-1}(A \,\square\, B). \tag{22}$$

Of course, the BKR operator $\square$ appearing in $a \,\square\, b$ in (22) is defined for $[0, 1]^d$ by (2) and (5), while the BKR operator $\square$ appearing in $A \,\square\, B$ in (22) is defined for $\mathbb{R}^d$ by the appropriate analog of (5); these *are* different operators.

Now apply Theorem 4 to get (20). For the r-fold BKR operator, the same g, combined with Theorem 6, implies (21). $\square$

Corollary 8. *Suppose $S \subset \mathbb{R}$ is Borel measurable. For Borel subsets A, B and $A_1, \ldots, A_r$ of S^d, under any complete product probability measure $\mathbb{P}$ on S^d, (20) and (21) hold.*

Proof. Extend $A \subset S^d$ to $\hat{A} \subset \mathbb{R}^d$ given by $\hat{A} := A \cup (\mathbb{R}^d \setminus S^d)$, likewise extend B or $A_1, \ldots, A_d$, and apply Theorem 7. $\square$

7. Infinite products

How should the BKR operator be extended from S^d to $S^\infty \equiv S^{\mathbb{N}}$? For $A \subset S^{\mathbb{N}}$, and $K \subset \mathbb{N}$, the definition of $[A]_K$ extends in the obvious way from (1): $[A]_K$ is the maximal cylinder subset of A, free in all coordinates indexed by $\mathbb{N} \setminus K$.

Definition (3) for the BKR operator $\square$ on spaces of the form S^d, if modified to apply to $S^{\mathbb{N}}$ merely by replacing $[d]$ by $\mathbb{N}$, yields an operator we shall call $\square_{=\infty}$:

$$A \square_{=\infty} B := \bigcup_{\text{disjoint } J,K \subset \mathbb{N}} [A]_J \cap [B]_K. \tag{23}$$

One problem with this operator is that it involves an uncountable union, so in the measurability argument from Lemma 1, the cylinders such as $[A]_J$ are Lebesgue measurable, but this fails to imply that for Borel set A, B, the result $A \square_{=\infty} B$ is Lebesgue measurable. A more severe problem with definition (23) is that it does not seem to yield to any approximation scheme down to a known version of the BKR inequality, as in the heart of this paper, Section 4.1.

Hence, for spaces of the form $S^{\mathbb{N}}$, we adopt the following definitions:

$$\text{for } A, B \subset S^{\mathbb{N}}, \qquad A \square B := \bigcup_{\text{finite disjoint } J,K \subset \mathbb{N}} [A]_J \cap [B]_K \tag{24}$$

and for $A_1, \ldots, A_r \subset S^{\mathbb{N}}$,

$$\mathop{\Large\square}_{1 \le i \le r} A_i \equiv A_1 \square \cdots \square A_r := \bigcup_{\text{finite disjoint } J_1,\ldots,J_r \subset \mathbb{N}} \bigcap_1^r [A_i]_{J_i}. \tag{25}$$

It *may* have been nice to use the customary BKR symbol $\square$ in the above definitions, rather than contrive new notation, perhaps $\square_{\text{finite}}$ or $\square_{\infty}$. It *is* valid, and would allow a single universal definition, to replace all of (3), (4), (24), and (25): for countable index set I (such as $I = [d]$ or $I = \mathbb{N}$), for $r \ge 2$ and for $A_1, \ldots, A_r \subset S^I$, we define the event *that* $A_1, \ldots, A_r$ *occur for* finite *disjoint sets of reasons*,

$$\mathop{\Large\square}_1^r A_i := A_1 \square \cdots \square A_r := \bigcup_{\text{finite disjoint } J_1,\ldots,J_r \subset I} \bigcap_1^r [A_i]_{J_i}. \tag{26}$$

However, in light of the natural alternate extension given by (23), users of the symbol $\square$ in the context of infinite products spaces *should* attach warning prose, as we do in Theorems 10 and 11 below.

Example 9. Consider $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = [0,1]^{\mathbb{N}}$, $\mathcal{F} =$ the Borel sets, and $\mathbb{P} = \lambda$, Lebesgue measure; as usual let $X_i :=$ the ith coordinate, $S_n := X_1 + \cdots + X_n$. Let $A = \{\limsup S_n/n \ge 0.2\}$. Then $A \square B = \varnothing$ for *every* event B, but $\mathbb{P}(A \square_{=\infty} A) = 1$, which can be seen by taking $J =$ the odd positive integers, $K =$ the even positive integers. Consider the r-fold $\square_{=\infty}$ operator defined in the natural way. Take $A_1 = \cdots = A_r = A$, $B_r := A_1 \square_{=\infty} A_2 \square_{=\infty} \cdots \square_{=\infty} A_r$, and $C_r := A_1 \square A_2 \square \cdots \square A_r \equiv \square_1^r A_i$. We have $B_r = \varnothing$ if and only if $r > 5$, and $\mathbb{P}(B_1) = \mathbb{P}(B_2) = 1$, $\mathbb{P}(B_3) = \mathbb{P}(B_4) = \mathbb{P}(B_5) = 0$. We have $C_r = \varnothing$ for $r = 1, 2, 3, \ldots$ with the case $r = 1$ serving to highlight a difference between the two forms of notation, $\square_1^r A_i$ and $A_1 \square \cdots \square A_r$ – does the latter reduce to A_1 when $r = 1$? Of course not.

The extension of Theorems 4 and 6 from $[0,1]^d$ to $[0,1]^{\mathbb{N}}$ is relatively easy. We write λ for Lebesgue measure on the Lebesgue measurable subsets of $[0,1]^{\mathbb{N}}$.

Theorem 10. *Consider the BKR combination of events, that they occur for* finite *disjoint sets of reasons, as specified by* (24) *and* (25). *For Borel subsets A, B in $[0,1]^{\mathbb{N}}$, $A \square B$ is Lebesgue measurable, and*

$$\lambda(A \square B) \le \lambda(A)\lambda(B).$$

For Borel subsets $A_1, \ldots, A_r$ in $[0,1]^{\mathbb{N}}$, $A_1 \square \cdots \square A_r$ is Lebesgue measurable, and

$$\lambda\left(\mathop{\square}_{1}^{r} A_i\right) \le \prod_{1}^{r} \lambda(A_i). \tag{27}$$

Proof. The Lebesgue measurability of the BKR products is clear from the sentence following (23). Define the level-d BKR operator on $[0,1]^{\mathbb{N}}$ by

$$A \square_d B := \bigcup_{\text{disjoint } J, K \subset [d]} [A]_J \cap [B]_K. \tag{28}$$

It is obvious that $A \square B$ is the countable, nested union of these, hence

$$A \square B = \bigcup_{d \ge 0} A \square_d B \quad \text{and} \quad \lim_{d \to \infty} \lambda(A \square_d B) = \lambda(A \square B).$$

Therefore, it suffices to show that for $d < \infty$, $\lambda(A \square_d B) \le \lambda(A)\lambda(B)$.

Fix d and let $C = A \square_d B$. Extend the notation $[[A]]_K$ for the base of the cylinder $[A]_K$, from (10) to the situation with $A \subset [0,1]^{\mathbb{N}}$, and apply it with $K = [d]$. Take $A' := [[A]]_{[d]} \subset [0,1]^d$, so $[A]_{[d]} \subset A$, and $\lambda_d(A') = \lambda([A]_{[d]}) \le \lambda(A)$. Similarly take $B' := [[B]]_{[d]}$ and $C' := [[C]]_{[d]}$. Note that C is a cylinder, free in the coordinates of index greater than d, so $C = [C]_{[d]}$ and $\lambda(C) = \lambda_d(C')$. It is "obvious" (and we supply details in the next paragraph) that with the usual BKR operator on $[0,1]^d$, $A' \square B' = C'$, so Corollary 5 applies, showing that $\lambda_d(C') \le \lambda_d(A')\lambda_d(B')$, and chaining together inequalities completes the proof that $\lambda(A \square B) \le \lambda(A)\lambda(B)$.

Details for $A' \square B' = C'$: We start with $C := A \square_d B$ as defined by (28), and apply $\mathrm{Proj}_{[d]}$. The relation $([A]_J)_K = [A]_{J \cap K}$ in $[0,1]^{\mathbb{N}}$, used with $K = [d]$, shows that for $J \subset [d]$,

$$\big([A]_{[d]}\big)_J = [A]_J \quad \text{and hence, in } [0,1]^d \quad \big[A'\big]_J = \mathrm{Proj}_{[d]}\big([A]_J\big).$$

The function $\mathrm{Proj} = \mathrm{Proj}_{[d]}$, which is the set-to-set function induced by $\mathrm{proj}_{[d]} : [0,1]^{\mathbb{N}} \to [0,1]^d$, distributes over unions. For $J, K \subset [d]$, $[A']_J = \mathrm{Proj}([A]_J)$ and $[B']_K = \mathrm{Proj}([B]_K)$, also, both $[A]_J$ and $[B]_K$ are cylinders free in all coordinates of index greater than d, so that $\mathrm{Proj}([A]_J) \cap \mathrm{Proj}([B]_K) = \mathrm{Proj}([A]_J \cap [B]_K)$. Hence, with all unions taken over disjoint $J, K \subset [d]$,

$$\begin{aligned} A' \,\square\, B' &= \bigcup [A']_J \cap [B']_K \\ &= \bigcup \mathrm{Proj}\big([A]_J\big) \cap \mathrm{Proj}\big([B]_K\big) \\ &= \mathrm{Proj}\Big(\bigcup [A]_J \cap [B]_K\Big) \\ &= \mathrm{Proj}(A \,\square_d\, B) = \mathrm{Proj}(C) = \mathrm{Proj}\big([C]_{[d]}\big) = C'. \end{aligned}$$

Finally, the result for the simultaneous r-fold BKR operator follows by a similar argument, starting with an extension of (28) to define a level-d r-fold BKR operator. □

Theorem 11. *Consider the BKR combination of events, that they occur for* finite *disjoint sets of reasons, as specified by* (24) *and* (25). *For Borel subsets A, B of $\mathbb{R}^{\mathbb{N}}$, under any complete product probability measure $\mathbb{P}$ on $\mathbb{R}^{\mathbb{N}}$,*

$$\mathbb{P}(A \,\square\, B) \le \mathbb{P}(A)\mathbb{P}(B). \tag{29}$$

For Borel subsets $A_1, \ldots, A_r$ of $\mathbb{R}^{\mathbb{N}}$, under any complete product probability measure $\mathbb{P}$ on $\mathbb{R}^{\mathbb{N}}$,

$$\mathbb{P}\left(\mathop{\square}_{1}^{r} A_i\right) \le \prod_1^r \mathbb{P}(A_i). \tag{30}$$

Proof. The result follows immediately from Theorem 10, by adapting (19) and the argument used to prove Theorem 7, from the context of $\mathbb{R}^d$, to the context of $\mathbb{R}^{\mathbb{N}}$. □

8. Relaxing the sample space

In this paper, we consider a sample space S^I for I countable and $S = [0, 1]$ – with Lebesgue measure on S^I, or $S = \mathbb{R}$, with arbitrary complete product probability measure on S^I. However, all results can be carried over to the superficially more general case $\Omega := \prod_{i \in I} S_i$ for S_i a Polish subspace (equivalently, G_δ subset) of $\mathbb{R}$, each S_i is endowed with a probability measure $\mathbb{P}_i$ defined on the Borel subsets, and Ω has the product measure $\mathbb{P} = \prod \mathbb{P}_i$.

Extend $\mathbb{P}_i, \mathbb{P}$ to measures $\hat{\mathbb{P}}_i, \hat{\mathbb{P}}$ on $\mathbb{R}, \mathbb{R}^I$ respectively, by taking them to be 0 on the complement. The definition of the BKR operation from (4) or (25) rephrases in a natural way to Ω. One finds, for Borel sets $A_j \subset \Omega$: (a) $\square_j A_j$ is $\mathbb{P}$-measurable by the argument of Lemma 1, and (b) writing $\hat{\square}$ for the BKR operation computed with respect to $\mathbb{R}^I$ and $\hat{A}_j := A_j \cup (\mathbb{R}^I \setminus \Omega)$, that

$$\square_j A_j = (\hat{\square}_j \hat{A}_j) \cap \Omega.$$

Therefore, $\mathbb{P}(\square_j A_j) = \hat{\mathbb{P}}(\hat{\square}_j \hat{A}_j)$ and $\prod_j \hat{\mathbb{P}}(\hat{A}_j) = \prod_j \mathbb{P}(A_j)$, and it is clear that Theorems 7 and 11 for $\mathbb{R}^I$ imply the BKR inequality for $\prod_{i \in I} S_i$.

9. From Ω to Ω

It is tempting to attempt to extend our results to get something symmetric, where we assume that the inputs A, B are in a larger family of sets than the Borel sets, and the output $A \square B$, satisfying $\lambda_d(A \square B) \le \lambda_d(A)\lambda_d(B)$, is in the same family. Since defining the BKR product requires only complement, countable union, and projection, the "larger family" should be the class of projective sets, the smallest extension of the class of Borel sets closed under projection, complement and countable union, see [12,13]. Then the version of Lemma 1, *If A, B are projective, then the cylinders $[A]_K$ and the BKR product $A \square B$ are also projective*, is immediately true.

Probabilists may be familiar with the construction of the family of Borel sets, starting from the family of open sets, take complements and countable unions, to get a larger family, then iterate – see [3], pages 30–32. The construction of projective sets is similar; start with the Borel sets, take projections, countable unions, and complements, to get a larger family, then iterate. But there is a difference: the construction of Borel sets requires iteration out to the first uncountable ordinal, usually denoted Ω, while the construction of projective sets is finished at the first infinite ordinal ω.

In view of Corollary 5, to get BKR inequalities, we need only show that Lebesgue measure extends to projective sets. Here the situation is somewhat complex. It is *consistent* with ZFC to assume that such extension is false, in fact that there are nonmeasurable projective sets only one level in the projective hierarchy above analytic sets [8]. On the other hand, the existence of an inaccessible cardinal would imply that all projective sets are measurable [17]. Though such existence cannot be proved to be consistent with ZFC, it is widely assumed that this (consistency) is true – and often such existence is accepted as a useful extra axiom.

10. Open problems

Problem 12. For the BKR operator $\square_{=\infty}$ defined by (23), prove or give a counterexample: *For Borel subsets A, B in $[0,1]^{\mathbb{N}}$, there exists a Borel set C, with $A \square B \subset C$ and $\lambda(C) \le \lambda(A)\lambda(B)$.*

It is not hard to determine, for the special case $d = 2$, when the BKR inequality holds with equality: for Borel sets $A, B \subset [0,1]^2$, $\lambda_2(A \square B) = \lambda_2(A)\lambda_2(B)$ if and only if if and only if (0) $\lambda_1(A)\lambda_2(B) = 0$, or (1) A or B is all of $[0,1]^2$, or (2) A and B are each unions of a "cylinder" and a measure zero set, with the two cylinders being "orthogonal," that is, in different directions.

Problem 13. Give a simple necessary and sufficient condition for $A, B \subset [0,1]^d$, to satisfy $\lambda_d(A \square B) = \lambda_d(A)\lambda_d(B)$.

As background for Problems 14 and 15: in [2], Proposition 3, for arbitrary S and $A_1, \ldots, A_r \subset S^d$, we showed that $\square_1^r A_i \subset ((\cdots((A_1 \square A_2) \square A_3) \cdots \square A_{r-1}) \square A_r)$. For brevity, we omit the symbol for binary BKR operator, and write simply $\square_1^r A_i \subset ((\cdots((A_1 A_2) A_3) \cdots A_{r-1}) A_r)$. For a binary operator, the number of ways to associate a product with r factors is given by the Catalan number C_{r-1}, and the same argument shows that the simultaneous r-fold BKR product, $\square_1^r A_i$, is a subset of each of the binary-associated products.

Problem 14. Prove or disprove: for $r = 3, 4, \dots,$ there exist S and d, and $A_1, \dots, A_r \subset S^d$, such that the C_{r-1} binary-associated products for $A_1 A_2 \cdots A_r$ are all distinct.

Problem 15. For $r = 3, 4, \dots,$ for any S and d, and for any $A_1, \dots, A_r \subset S^d$, we already know that $\square_1^r A_i$ is a subset of the *intersection* of the C_{r-1} binary-associated products for $A_1 A_2 \cdots A_r$. Prove or disprove: for $r = 3, 4, \dots,$ there exists an example where the containment of $\square_1^r A_i$ is *strict*.

Now consider cases where all r factors are the same set A. Commutativity of the binary BKR product implies that $(A \,\square\, A) \,\square\, A = A \,\square\, (A \,\square\, A)$. The next example resolves the situation for $r = 4$.

Example 16 ($((AA)A)A \neq (AA)(AA)$ ***can occur).*** In $\{0, 1\}^6$, let A be the union of the following 2-cylinders, each of which is a set of size 16:

$$\begin{array}{cccc} 11{*}{*}{*}{*}, & {*}{*}11{*}{*}, & 1{*}{*}0{*}{*}, & {*}11{*}{*}{*}, \\ {*}{*}00{*}{*}, & {*}{*}{*}{*}00, & {*}{*}1{*}{*}0, & {*}{*}{*}00{*}. \end{array}$$

Note that the first two 2-cylinders combine to show that $1111{*}{*} \subset AA$, the next two show that $1110{*}{*} \subset AA$. Hence the first four 2-cylinders show that $111{*}{*}{*} \subset AA$. Similarly, the last four 2-cylinders show that ${*}{*}{*}000 \subset AA$. Combining, we see that $111000 \in (AA)(AA)$. Computer-exhaustive checking shows that $((AA)A)A = \varnothing$, hence $((AA)A)A \neq (AA)(AA)$.

In honor of Wedderburn [20], [16], Sequence A001190, write W_n for the number of ways to binary-associate a product of the form A^n, up to equivalence modulo the commutative property of the binary relation; for example, $W_2, W_3, \dots, W_7 = 1, 1, 1, 2, 3, 6, 11$.

Problem 17. (a) For $r = 5, 6, \dots,$ does there exist an example with a single set A, such that all W_r equivalence classes of association yield different results?

(b) As above, with the additional restriction that $A \subset \{0, 1\}^d$ for some d depending on r.

(c) If, for a given r, there is an example with $A \subset \{0, 1\}^d$ such that all W_r equivalence classes of association yield different results, write D_r for the smallest such d, following the notation Ramsey numbers. Example 16 shows that $D_4 \leq 6$. Can you prove that $D_4 > 5$? Can you determine D_5? Or give nontrivial upper or lower bounds for D_r for general r?

Acknowledgments

We thank Yiannis Moschovakis for a helpful conversation and the referee for their helpful remarks. Garibaldi's research was partially supported by NSF Grant DMS-12-01542.

References

[1] Alexander, K.S. (1993). A note on some rates of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 81–90. MR1202516

[2] Arratia, R., Garibaldi, S., Mower, L. and Stark, P.B. (2015). Some people have all the luck. *Math. Mag.* **88** 196–211. MR3383910

[3] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1324786

[4] Borgs, C., Chayes, J.T. and Randall, D. (1999). The van den Berg–Kesten–Reimer inequality: A review. In *Perplexing Problems in Probability*. *Progress in Probability* **44** 159–173. Boston, MA: Birkhäuser. MR1703130

[5] Chayes, J.T., Puha, A.L. and Sweet, T. (1999). Independent and dependent percolation. In *Probability Theory and Applications* (*Princeton*, *NJ*, 1996). *IAS/Park City Math. Ser.* **6** 49–166. Providence, RI: Amer. Math. Soc. MR1678308

[6] Cohn, D.L. (2013). *Measure Theory*, 2nd ed. *Birkhäuser Advanced Texts*: *Basler Lehrbücher*. [*Birkhäuser Advanced Texts*: *Basel Textbooks*]. New York: Birkhäuser/Springer. MR3098996

[7] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358

[8] Gödel, K. (1940). *The Consistency of the Continuum Hypothesis*. *Annals of Mathematics Studies*, *No*. 3. Princeton, NJ: Princeton Univ. Press. MR0002514

[9] Goldstein, L. and Rinott, Y. (2007). Functional BKR inequalities, and their duals, with applications. *J. Theoret. Probab.* **20** 275–293. MR2324531

[10] Goldstein, L. and Rinott, Y. (2015). Functional van den Berg–Kesten–Reimer inequalities, and their duals, with applications. Available at arXiv:1508.07267.

[11] Kahn, J., Saks, M. and Smyth, C. (2011). The dual BKR inequality and Rudich's conjecture. *Combin. Probab. Comput.* **20** 257–266. MR2769191

[12] Kechris, A.S. (1995). *Classical Descriptive Set Theory*. *Graduate Texts in Mathematics* **156**. New York: Springer. MR1321597

[13] Moschovakis, Y.N. (2009). *Descriptive Set Theory*, 2nd ed. *Mathematical Surveys and Monographs* **155**. Providence, RI: Amer. Math. Soc. MR2526093

[14] Potter, M. (2004). *Set Theory and Its Philosophy*: *A Critical Introduction*. New York: Oxford Univ. Press. MR2050123

[15] Reimer, D. (2000). Proof of the van den Berg–Kesten conjecture. *Combin. Probab. Comput.* **9** 27–32. MR1751301

[16] Sloane, N.J.A. The On-Line Encyclopedia of Integer Sequences. Available at oeis.org.

[17] Solovay, R.M. (1970). A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. of Math.* (2) **92** 1–56. MR0265151

[18] van den Berg, J. and Fiebig, U. (1987). On a combinatorial conjecture concerning disjoint occurrences of events. *Ann. Probab.* **15** 354–374. MR0877608

[19] van den Berg, J. and Kesten, H. (1985). Inequalities with applications to percolation and reliability. *J. Appl. Probab.* **22** 556–569. MR0799280

[20] Wedderburn, J.H.M. (1922). The functional equation $g(x^2) = 2\alpha x + [g(x)]^2$. *Ann. of Math.* (2) **24** 121–140. MR1502633

Received December 2015 and revised June 2016