



On the Borel–Cantelli Lemmas, the Erdős–Rényi Theorem, and the Kochen–Stone Theorem

ROB ARTHAN
PAULO OLIVA

Abstract: In this paper we present a quantitative analysis of the first and second Borel–Cantelli Lemmas and of two of their generalisations: the Erdős–Rényi Theorem, and the Kochen–Stone Theorem. We will see that the first three results have *direct* quantitative formulations, giving an explicit relationship between quantitative formulations of the assumptions and the conclusion. For the Kochen–Stone Theorem, however, we can show that the numerical bounds of a direct quantitative formulation are not computable in general. Nonetheless, we obtain a quantitative formulation of the Kochen–Stone Theorem using Tao’s notion of *metastability*.

2010 Mathematics Subject Classification 60A10 (primary); 03F60 (secondary)

Keywords: Quantitative analysis, metastability, Borel–Cantelli Lemma, Kochen–Stone Theorem, Erdős–Rényi Theorem

1 Introduction

Let $(A_i)_{i=1}^\infty$ be an infinite sequence of events in a probability space $(\mathcal{S}, \mathcal{E}, \mathbb{P})$. The Borel–Cantelli Lemma is a classical result in probability theory, relating the convergence or divergence of the sum $\sum_{i=1}^\infty \mathbb{P}[A_i]$ with the probability of the event “ A_i infinitely often”, which is defined as follows

$$A_i \text{ i.o.} = \bigcap_{n=1}^\infty \bigcup_{i \geq n} A_i$$

ie $\omega \in \mathcal{S}$ happens infinitely often in $(A_i)_{i=1}^\infty$ if for all n there exists an $i \geq n$ such that $\omega \in A_i$. Note that, as $\bigcup_{i \geq 1} A_i \supseteq \bigcup_{i \geq 2} A_i \supseteq \dots$, we have:

$$\mathbb{P}[A_i \text{ i.o.}] = \lim_{n \rightarrow \infty} \mathbb{P} \left[\bigcup_{i \geq n} A_i \right]$$

The Borel–Cantelli Lemma (see, for example, Feller [5]) is normally presented in two parts. The first part says that when the sum $\sum_{i=1}^{\infty} \mathbb{P}[A_i]$ converges, then the event A_i i.o. has probability zero:

Theorem 1.1 (First Borel–Cantelli Lemma) *Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] < \infty$. Then $\mathbb{P}[A_i \text{ i.o.}] = 0$.*

The second part says that when $\sum_{i=1}^{\infty} \mathbb{P}[A_i]$ diverges, and when the A_i are mutually independent, then the event A_i i.o. has probability one:

Theorem 1.2 (Second Borel–Cantelli Lemma) *Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of mutually independent events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$. Then $\mathbb{P}[A_i \text{ i.o.}] = 1$.*

In [7], Kochen and Stone presented a result that generalises the Second Borel–Cantelli Lemma in two directions: (i) it gives a lower bound on $\mathbb{P}[A_i \text{ i.o.}]$ when the A_i are not mutually independent; and (ii) it can be used to show that the assumption of mutual independence in the original lemma can be weakened to pairwise independence. We formulate this generalisation following Yan [16]:

Theorem 1.3 (Kochen–Stone) *Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$. Then:*

$$(1) \quad \mathbb{P}[A_i \text{ i.o.}] \geq \limsup_{n \rightarrow \infty} \frac{(\sum_{k=1}^n \mathbb{P}[A_k])^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}$$

Erdős and Rényi [4] gave a result that is intermediate between the second Borel–Cantelli Lemma and the Kochen–Stone Theorem. Like the Kochen–Stone Theorem it implies that the assumption of mutual independence in the second Borel–Cantelli Lemma can be weakened to pairwise independence. Erdős and Rényi applied their theorem to the study of generalised Cantor expansions for real numbers.

Theorem 1.4 (Erdős–Rényi) *Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$ and:*

$$(2) \quad \liminf_{n \rightarrow \infty} \frac{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}{(\sum_{k=1}^n \mathbb{P}[A_k])^2} = 1$$

Then $\mathbb{P}[A_i \text{ i.o.}] = 1$.

The aim of the present note is to investigate “quantitative” versions of each of these “qualitative” results. The methods we use come from the *proof mining* programme (see Kohlenbach [8])—where numerical information is obtained from (often non-constructive) proofs via logical methods. For some noteworthy applications of these methods, see the work of Avigad and collaborators on Ergodic Theory [1] and Kohlenbach and collaborators on Fixed-Point Theory [9, 10].

Terence Tao’s programme of bridging “soft” and “hard” analysis [13] was an independent rediscovery of some of these ideas. The results as presented above are results of “soft analysis”: they relate statements about convergence or divergence, without giving any numeric information about the corresponding rates of convergence or divergence. For instance, regarding the First Borel–Cantelli Lemma, it is natural to ask how the rate of convergence of the sequence of partial sums $(\sum_{i=1}^n \mathbb{P}[A_i])_{n=1}^{\infty}$ relates to the rate of convergence of the sequence of probabilities $(\mathbb{P}[\bigcup_{i \geq n} A_i])_{n=1}^{\infty}$, whose limit is $\mathbb{P}[A_i \text{ i.o.}]$. Similar questions arise regarding the other three results.

We provide here answers to these four questions. We will find in Section 2 that the answer is almost trivial for the First Borel–Cantelli Lemma, as it has a very direct (constructive) proof. It turns out that the sequence $(\mathbb{P}[\bigcup_{i \geq n} A_i])_{n=1}^{\infty}$ converges with the same rate as the sequence $(\sum_{i=1}^n \mathbb{P}[A_i])_{n=1}^{\infty}$. The answers are found to be less trivial in Section 2 for the Second Borel–Cantelli Lemma and in Section 3 for the Erdős–Rényi Theorem, but still, the quantitative versions of these follow the standard proofs of the qualitative versions quite closely. As we will see in Section 4, in the case of the Kochen–Stone Theorem the situation is more complicated. In Section 4.2, we prove that a direct (computable) rate of convergence does not exist: we can find a concrete sequence of events (with computable probabilities $\mathbb{P}[A_i]$) such that the rate of convergence for the quantitative version of the theorem is not computable. To allow for this, we use Tao’s notion of “rate of metastability” (see Section 1.1), a concept which is logically equivalent to convergence, but is computationally weaker. We give a quantitative version of the Kochen–Stone Theorem that provides a rate of metastability for inequality (1) as a computable function of the rate of divergence of the sequence of partial sums $(\sum_{i=1}^n \mathbb{P}[A_i])_{n=1}^{\infty}$.

In Section 5 we also consider the optimality of the bounds we obtain. For the First and Second Borel–Cantelli Lemmas, we can argue that the bounds obtained are in some sense best possible. For the Erdős–Rényi Theorem we conjecture that the bounds we present are optimal, but do not have a proof yet. For the Kochen–Stone Theorem, however, the “metastable” reformulation makes it much less clear what the right notion of optimality should be, and we leave this to future work.

The work presented in this paper was motivated by our on-going work in the area of metric Diophantine approximation, more specifically, on quantitative analyses of generalisations of the Khintchine–Groshev Theorem on approximability of real numbers by rationals. To introduce these generalisations, let \mathbb{I}^{nm} denote the unit cube $[0, 1]^{nm}$ in \mathbb{R}^{nm} , and let $\psi: \mathbb{N} \rightarrow \mathbb{R}^+$ be given. A point $\mathbf{X} \in \mathbb{I}^{nm}$, viewed as an $n \times m$ matrix, is said to be ψ -approximable if there are infinitely many $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n$ such that $\|\mathbf{q}\mathbf{X} + \mathbf{p}\| < \psi(\|\mathbf{q}\|)$ (where $\|\cdot\|$ is the supremum norm). Let $\mathcal{A}_{n,m}$ denote the set of ψ -approximable points $\mathbf{X} \in \mathbb{I}^{nm}$. The generalised Khintchine–Groshev Theorems are 0–1 laws for the Lebesgue measure of the sets $\mathcal{A}_{n,m}$ governed by assumptions on the divergence (and possibly monotonicity) of sequences defined in terms of ψ . The most general theorem of this form is given in Beresnevich and Velani [2], which improves on earlier work of Gallagher [6] dealing with the case $n = 1$. In both of these works, the proofs break into two parts: A proof of a 0–1 law and a proof that a certain set has positive measure, and hence measure 1 by the 0–1 law. The Borel–Cantelli Lemmas and their generalisations are important tools in some of these proofs. In particular, the Kochen–Stone Theorem is a key step¹ in Beresnevich and Velani’s work [2]. A quantitative analysis of these tools seemed to us to be a worthwhile investigation in its own right.

1.1 Rate of convergence vs rate of metastability

As mentioned above, in our quantitative analysis of the Kochen–Stone Theorem we will make use of Terence Tao’s notion of *metastability* [13]. As an example, consider the statement that a sequence of reals $(x_n)_{n=1}^\infty$ is Cauchy convergent:

$$(3) \quad \forall \ell \exists k \forall m, n > k \left(|x_m - x_n| < \frac{1}{2^\ell} \right)$$

A *rate of Cauchy convergence* (or just *rate of convergence*) for the sequence is a function $\phi: \mathbb{N} \rightarrow \mathbb{N}^+$ (where \mathbb{N}^+ denotes the positive integers) such that:

$$(4) \quad \forall \ell \forall m, n > \phi(\ell) \left(|x_m - x_n| < \frac{1}{2^\ell} \right)$$

While the mere existence of ϕ in (4) is equivalent to (3), if one has an explicit ϕ for which (4) holds, one has a quantitative rather than merely qualitative understanding of the convergence of the sequence $(x_n)_{n=1}^\infty$. However, for a given sequence that is known

¹Beresnevich and Velani in fact refer to Sprindzuk [12, Lemma 5] for the result and not to Kochen and Stone [7]. Presumably Sprindzuk obtained the result independently.

to be convergent, it may not be possible to provide an explicit rate of convergence: in many interesting cases, it can be shown that no computable function ϕ satisfying (4) exists. In such cases, it is often worth considering the equivalent “metastable” version of (3), namely:

$$(5) \quad \forall \ell, f: \mathbb{N}^+ \rightarrow \mathbb{N}^+ \exists k \forall m, n \in [k, f(k)] \left(|x_m - x_n| < \frac{1}{2^\ell} \right)$$

Clearly (3) directly implies (5). But (5) also implies (3). Assume (5) and suppose (3) does not hold for some ℓ , ie

$$\forall k \exists m, n > k \left(|x_m - x_n| \geq \frac{1}{2^\ell} \right)$$

and let $f: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ be any function which provides an upper bound for m and n for each given k , ie:

$$\forall k \exists m, n \in [k, f(k)] \left(|x_m - x_n| \geq \frac{1}{2^\ell} \right)$$

Taking this f in (5) leads to a contradiction.

In cases where there is no computable rate of convergence (as will be the case with the Kochen–Stone Theorem), one can still attempt to find a computable *rate of metastability* instead. In the convergence example above, the rate of metastability would be a function $\phi: \mathbb{N} \times (\mathbb{N}^+ \rightarrow \mathbb{N}^+) \rightarrow \mathbb{N}^+$ such that:

$$(6) \quad \forall \ell, f: \mathbb{N}^+ \rightarrow \mathbb{N}^+ \exists k \leq \phi(\ell, f) \forall m, n \in [k, f(k)] \left(|x_m - x_n| < \frac{1}{2^\ell} \right)$$

One should think of the function $f: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ as potentially producing longer and longer intervals $[k, f(k)]$, and the rate of metastability ϕ as trying to find, for arbitrarily large ℓ , an interval in which the sequence is $\frac{1}{2^\ell}$ -stable, ie $\forall m, n \in [k, f(k)] (|x_m - x_n| < \frac{1}{2^\ell})$.

1.2 Rate of divergence

Let $(x_i)_{i=1}^\infty$ be a non-decreasing sequence of real numbers. The sequence is said to *diverge* if

$$\forall N \exists i (x_i \geq N).$$

We say that a function $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ gives a *rate of divergence* for $(x_i)_{i=1}^\infty$ if

$$\forall N (x_{\omega(N)} \geq N).$$

In the sequel, the sequence x_i will typically comprise the partial sums of a series of terms in the interval $[0, 1]$, for which we have the following lemma.

Lemma 1.5 Let $x_n = \sum_{i=1}^n a_i$, where $0 \leq a_i \leq 1$ for all i , and assume $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ is a rate of divergence for $(x_n)_{n=1}^\infty$. Then, for all n, N we have:

$$(7) \quad \sum_{i=n}^{\omega(n+N-1)} a_i \geq N$$

Proof Since ω is a rate of divergence and each $a_i \leq 1$, we have that:

$$(8) \quad \sum_{i=n}^{\omega(n+N-1)} a_i = \sum_{i=1}^{\omega(n+N-1)} a_i - \sum_{i=1}^{n-1} a_i \geq (n+N-1) - (n-1) = N \quad \square$$

2 Quantitative Borel–Cantelli Lemmas

Given an infinite sequence of events $(A_i)_{i=1}^\infty$, the first Borel–Cantelli Lemma says that the probability of A_i i.o. is 0 when $\sum_{i=1}^\infty \mathbb{P}[A_i]$ is finite. Here the assumption “ $\sum_{i=1}^\infty \mathbb{P}[A_i]$ is finite” is equivalent to the convergence of the sequence of partial sums $s_k = \sum_{i=1}^k \mathbb{P}[A_i]$. In quantitative terms, that implies the existence of a rate of Cauchy convergence $\psi(\ell)$ for s_k , ie

$$(9) \quad \forall \ell, m, n > \psi(\ell) \left(|s_m - s_n| < \frac{1}{2^\ell} \right)$$

or, equivalently, the existence of a function $\phi(\ell)$ such that:

$$(10) \quad \forall \ell \forall m > \phi(\ell) \left(\sum_{i=\phi(\ell)}^m \mathbb{P}[A_i] < \frac{1}{2^\ell} \right)$$

Theorem 2.1 (First Borel–Cantelli Lemma: Quantitative Version) Let $(A_i)_{i=1}^\infty$ be an infinite sequence of events. Assume that $(\sum_{i=1}^m \mathbb{P}[A_i])_{m=1}^\infty$ converges with a rate of convergence $\phi: \mathbb{N} \rightarrow \mathbb{N}^+$, ie that for all $\ell \geq 0$ and $m > \phi(\ell)$:

$$\sum_{i=\phi(\ell)}^m \mathbb{P}[A_i] \leq \frac{1}{2^\ell}$$

Then the sequence $(\mathbb{P}[\bigcup_{i=1}^m A_i])_{m=1}^\infty$ converges to 0 with the same rate, ie for all $\ell \geq 0$ and $m > \phi(\ell)$:

$$\mathbb{P} \left[\bigcup_{i=\phi(\ell)}^m A_i \right] \leq \frac{1}{2^\ell}$$

Proof By subadditivity we have

$$(11) \quad \mathbb{P} \left[\bigcup_{i=\phi(\ell)}^m A_i \right] \leq \sum_{i=\phi(\ell)}^m \mathbb{P}[A_i] \leq \frac{1}{2^\ell}$$

for all $\ell > 0$ and $m > \phi(\ell)$. □

The second Borel–Cantelli Lemma says that, under the extra assumption that the events are mutually independent, the probability of A_i i.o. is 1 whenever the sum $\sum_{i=1}^\infty \mathbb{P}[A_i]$ diverges. In our quantitative version of this lemma we will estimate, for each n , how fast the sequence

$$\left(\mathbb{P} \left[\bigcup_{i=n}^m A_i \right] \right)_{m=1}^\infty$$

converges to 1, given a rate of divergence for the sequence $(\sum_{i=n}^m \mathbb{P}[A_i])_{m=1}^\infty$.

Theorem 2.2 (Second Borel–Cantelli Lemma: Quantitative Version) *Let $(A_i)_{i=1}^\infty$ be an infinite sequence of events which are mutually independent. Assume that the sequence $(\sum_{i=1}^n \mathbb{P}[A_i])_{n=1}^\infty$ diverges with rate $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$, ie for all N*

$$\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \geq N$$

then, for all n and N :

$$\mathbb{P} \left[\bigcup_{i=n}^{\omega(n+N-1)} A_i \right] \geq 1 - e^{-N}$$

Proof Fix n and N . Let us write \bar{A}_i for the complement of the event A_i . The independence of the events implies:

$$(12) \quad \mathbb{P} \left[\bigcap_{i=n}^{\omega(n+N-1)} \bar{A}_i \right] = \prod_{i=n}^{\omega(n+N-1)} \mathbb{P}[\bar{A}_i]$$

$$(13) \quad = \prod_{i=n}^{\omega(n+N-1)} (1 - \mathbb{P}[A_i])$$

Taking the natural logarithm on both sides we have

$$(14) \quad \ln \left(\mathbb{P} \left[\bigcap_{i=n}^{\omega(n+N-1)} \bar{A}_i \right] \right) = \ln \left(\prod_{i=n}^{\omega(n+N-1)} (1 - \mathbb{P}[A_i]) \right)$$

$$(15) \quad = \sum_{i=n}^{\omega(n+N-1)} \ln(1 - \mathbb{P}[A_i])$$

$$(16) \quad \leq - \sum_{i=n}^{\omega(n+N-1)} \mathbb{P}[A_i]$$

$$(17) \quad \leq -N$$

where inequality (16) follows from the fact that $\ln(1+x) \leq x$, for all $x \in (-1, \infty)$ and inequality (17) follows from Lemma 1.5. Hence

$$(18) \quad \mathbb{P} \left[\bigcap_{i=n}^{\omega(n+N-1)} \bar{A}_i \right] \leq e^{-N}$$

and so:

$$(19) \quad \mathbb{P} \left[\bigcup_{i=n}^{\omega(n+N-1)} A_i \right] \geq 1 - e^{-N} \quad \square$$

2.1 Proving qualitative version from quantitative one

That Theorem 1.1 (First Borel–Cantelli Lemma) follows from Theorem 2.1 (its quantitative version) is clear. Let us show that Theorem 1.2 (Second Borel–Cantelli Lemma) also follows from Theorem 2.2 (its quantitative version).

Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of mutually independent events. Assuming $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$, there exists $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$, such that for all N :

$$\left(\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \right) \geq N$$

By Theorem 2.2, for all n and N

$$\mathbb{P} \left[\bigcup_{i=n}^{\omega(n+N-1)} A_i \right] \geq 1 - e^{-N}$$

which implies that

$$\mathbb{P} \left[\bigcup_{i=n}^{\infty} A_i \right] \geq 1 - e^{-N}$$

and hence $\mathbb{P}[A_i \text{ i.o.}] = 1$.

3 Quantitative Erdős–Rényi Theorem

In this section, we present a quantitative version of the Erdős–Rényi Theorem. Our proof follows that of Erdős and Rényi [4], but uses more modern notation: X, Y for random variables, $\mathbb{E}(X)$ for the expectation of X (or mean value in Erdős and Rényi’s terminology), and $\sigma(X)$ for the standard deviation of X .

A simplistic logical formalisation of equation (2), involving \liminf , would give a formula with quantifier prefix $\forall\exists\forall\exists$. We can simplify this to a $\forall\exists$ using the following lemmas.

Lemma 3.1 For any sequence of events $(A_i)_{i=1}^\infty$ we have that for all $n \geq 1$:

$$(20) \quad \frac{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}{(\sum_{k=1}^n \mathbb{P}[A_k])^2} \geq 1$$

Proof Let X_i be the random variable given by the indicator function of the event A_i . Then $\mathbb{E}(X_i) = \mathbb{P}[A_i]$ and $\mathbb{E}(X_i X_k) = \mathbb{P}[A_i A_k]$. Define $Y_n = \sum_{i=1}^n X_i$ so that:

$$(21) \quad \frac{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}{(\sum_{i=1}^n \mathbb{P}[A_i])^2} = \frac{\mathbb{E}(Y_n^2)}{\mathbb{E}^2(Y_n)}$$

Since the variance $\sigma^2(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) \geq 0$ for any X , the result follows. \square

Lemma 3.2 For any sequence of reals $d_n \geq 1$, the following are equivalent:

$$(22) \quad \liminf_{n \rightarrow \infty} d_n = 1$$

$$(23) \quad \forall \ell, n \exists k \geq n \left(d_k \leq 1 + 2^{-\ell} \right)$$

Proof Because $d_n \geq 1$, we have that $\liminf_{n \rightarrow \infty} d_n = 1$ is equivalent to

$$\forall \ell \exists m \forall i \geq m \left(\inf_{k \geq i} d_k \leq 1 + 2^{-\ell} \right)$$

which, by the definition of \inf , is equivalent to:

$$\forall \ell \exists m \forall i \geq m \exists k \geq i \left(d_k \leq 1 + 2^{-\ell} \right)$$

This is easily seen to be equivalent to $\forall \ell, n \exists k \geq n \left(d_k \leq 1 + 2^{-\ell} \right)$ (for the right-to-left direction take $i = \max(m, n)$). \square

In the quantitative version of the Erdős–Rényi Theorem, we will assume that we are given a rate of divergence for the sequence $(\sum_{i=1}^m \mathbb{P}[A_i])_{m=1}^\infty$, and a function ϕ witnessing

$$\forall \ell, n \exists m \geq n \left(\frac{\sum_{i,k=1}^m \mathbb{P}[A_i A_k]}{(\sum_{k=1}^m \mathbb{P}[A_k])^2} \leq 1 + \frac{1}{2^\ell} \right)$$

which, by Lemmas 3.1 and 3.2, is equivalent to the assumption:

$$\liminf_{n \rightarrow \infty} \frac{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}{(\sum_{k=1}^n \mathbb{P}[A_k])^2} = 1$$

Theorem 3.3 (Erdős–Rényi: Quantitative Version) *Let $(A_i)_{i=1}^\infty$ be an infinite sequence of events. Let $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ be such that for all N*

$$(24) \quad \left(\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \right) \geq N$$

and let $\phi: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be such that:

$$(25) \quad \forall \ell, n \left(\phi(\ell, n) \geq n \wedge \frac{\sum_{i,k=1}^{\phi(\ell, n)} \mathbb{P}[A_i A_k]}{(\sum_{i=1}^{\phi(\ell, n)} \mathbb{P}[A_i])^2} \leq 1 + \frac{1}{2^\ell} \right)$$

Define $n_1 = \phi(1, 1)$ and, for $k > 1$, $n_k = \phi(k, \max(n_{k-1}, k))$. Then, for all n and ℓ

$$(26) \quad \mathbb{P} \left[\bigcup_{i=n}^{n_m} A_i \right] \geq 1 - \frac{1}{2^\ell}$$

where $m = \max(\omega(2n), \ell + 3)$.

Proof Let X_i and Y_n be as in the proof of Lemma 3.1. Assumption (25) gives us:

$$(27) \quad \forall \ell, n \left(\phi(\ell, n) \geq n \wedge \frac{\mathbb{E}(Y_{\phi(\ell, n)}^2)}{\mathbb{E}^2(Y_{\phi(\ell, n)})} \leq 1 + \frac{1}{2^\ell} \right)$$

Since $\mathbb{E}(Y_n^2) = \sigma^2(Y_n) + \mathbb{E}^2(Y_n)$, we have:

$$(28) \quad \frac{\mathbb{E}(Y_n^2)}{\mathbb{E}^2(Y_n)} = \frac{\sigma^2(Y_n)}{\mathbb{E}^2(Y_n)} + 1$$

By (27) and (28), we get:

$$(29) \quad \forall \ell, n \left(\phi(\ell, n) \geq n \wedge \frac{\sigma^2(Y_{\phi(\ell, n)})}{\mathbb{E}^2(Y_{\phi(\ell, n)})} \leq \frac{1}{2^\ell} \right)$$

Let $n_1 = \phi(1, 1)$ and, for $k > 1$, $n_k = \phi(k, \max(n_{k-1}, k))$. The above implies (taking $\ell = k$ and $n = \max(n_{k-1}, k)$):

$$(30) \quad \forall k \left(\frac{\sigma^2(Y_{n_k})}{\mathbb{E}^2(Y_{n_k})} \leq \frac{1}{2^k} \right)$$

The Chebyshev inequality tells us that:

$$(31) \quad \mathbb{P}[|Y_n - \mathbb{E}(Y_n)| \geq \lambda \sigma(Y_n)] \leq \frac{1}{\lambda^2}$$

Taking $\lambda = \frac{\varepsilon \mathbb{E}(Y_n)}{\sigma(Y_n)}$, we have that, for any given $\varepsilon \in (0, 1)$:

$$(32) \quad \mathbb{P}[Y_n \leq (1 - \varepsilon)\mathbb{E}(Y_n)] \leq \frac{\sigma^2(Y_n)}{\varepsilon^2 \mathbb{E}^2(Y_n)}$$

From (30) and (32) (taking $n = n_k$ and $\varepsilon = 1/2$), we find that:

$$(33) \quad \forall k \left(\mathbb{P} \left[Y_{n_k} \leq \frac{\mathbb{E}(Y_{n_k})}{2} \right] \leq \frac{1}{2^{k-2}} \right)$$

Let B_k be the event $Y_{n_k} \leq \frac{\mathbb{E}(Y_{n_k})}{2}$, so that (33) together with the formula for the partial sums of a geometric series implies:

$$(34) \quad \forall \ell \forall m \geq \ell + 3 \left(\sum_{k=\ell+3}^m \mathbb{P}[B_k] \leq \frac{1}{2^\ell} \right)$$

By the quantitative version of the First Borel–Cantelli Lemma (Theorem 2.1):

$$(35) \quad \forall \ell \forall m \geq \ell + 3 \left(\mathbb{P} \left[\bigcup_{k=\ell+3}^m B_k \right] \leq \frac{1}{2^\ell} \right)$$

Hence

$$(36) \quad \forall \ell \forall m \geq \ell + 3 \left(\mathbb{P} \left[\bigcap_{k=\ell+3}^m \bar{B}_k \right] \geq 1 - \frac{1}{2^\ell} \right)$$

where \bar{B}_k is the complement of B_k , ie the event $Y_{n_k} > \frac{\mathbb{E}(Y_{n_k})}{2}$. Hence, an outcome $x \in \bar{B}_k$ if $Y_{n_k}(x) = \sum_{i=1}^{n_k} X_i(x) > \frac{\mathbb{E}(Y_{n_k})}{2}$. Therefore, $x \in \bar{B}_k$ iff $x \in A_i$ for at least $\frac{\mathbb{E}(Y_{n_k})}{2}$ values of $i \in [1, n_k]$.

Fix n and ℓ . We will show that

$$(37) \quad \mathbb{P} \left[\bigcup_{i=n}^{n_m} A_i \right] \geq 1 - \frac{1}{2^\ell}$$

for $m = \max(\omega(2n), \ell + 3)$. Let $C_\ell^m = \bigcap_{k=\ell+3}^m \bar{B}_k$. By (36), inequality (37) will follow if we can show that:

$$(38) \quad C_\ell^m \subseteq \bigcup_{i=n}^{n_m} A_i$$

So, let us assume that $x \in C_\ell^m$. Since, $x \in C_\ell^m$ iff $\forall k \in [\ell + 3, m](x \in \bar{B}_k)$, we have, taking $k = m$:

$$(39) \quad Y_{n_m}(x) > \frac{\mathbb{E}(Y_{n_m})}{2}$$

By the definition of n_k and the assumption (25), which implies $\phi(\ell, n) \geq n$, we have $n_{\omega(2n)} \geq \omega(2n)$, and by the definition of m we also have $m \geq \omega(2n)$. Hence

$$(40) \quad Y_{n_m}(x) > \frac{\mathbb{E}(Y_{n_m})}{2} \geq \frac{\mathbb{E}(Y_{\omega(2n)})}{2} \geq n$$

using, for the last inequality, that assumption (24) implies $\forall N (\mathbb{E}(Y_{\omega(N)}) \geq N)$. Since, $Y_{n_m}(x) = \sum_{i=1}^{n_m} X_i(x)$, (40) and the pigeon-hole principle imply that for at least one $i \in [n, n_m]$ we have $X_i(x) = 1$, ie $x \in A_i \subseteq \bigcup_{j=n}^{n_m} A_j$. Since x was an arbitrary element of C_ℓ^m , this gives us (38) and hence (37). \square

Remark To obtain a quantitative version of the Erdős–Rényi Theorem, we have had to make two choices about points that are left open in the qualitative proof of [4]. The first choice is essentially forced upon us by our decision to use $\frac{1}{2^\ell}$ in the formulation of assumption (25). This means that where Erdős and Rényi take the n_k to be any sequence such that $\sum_{k=1}^{\infty} \frac{\sigma^2(Y_{n_k})}{\mathbb{E}^2(Y_{n_k})}$ converges, we have had to choose a sequence such that the series is dominated by a geometric series (see formula (30)). The second choice is that in formula (33), we have taken ε to be $\frac{1}{2}$, where Erdős and Rényi leave it unspecified. As pointed out by one of the referees, both the statements and the proofs above (Lemma 3.2 and Theorem 3.3) could be made more general, but rather more complicated, by introducing a convergent series in place of $\frac{1}{2}$ and a constant $\varepsilon \in (0, 1)$ in place of $\frac{1}{2}$ as parameters.

3.1 Proving qualitative version from quantitative one

Let us show that the Erdős–Rényi Theorem (Theorem 1.4) follows directly from our quantitative version (Theorem 3.3).

Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$ and assume:

$$(41) \quad \liminf_{n \rightarrow \infty} \frac{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}{(\sum_{k=1}^n \mathbb{P}[A_k])^2} = 1$$

The above assumptions imply that there exists an $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ such that

$$(42) \quad \forall N \left(\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \geq N \right)$$

and a function $\phi: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ such that:

$$(43) \quad \forall \ell, n \left(\phi(\ell, n) \geq n \wedge \frac{\sum_{i,k=1}^{\phi(\ell, n)} \mathbb{P}[A_i A_k]}{\left(\sum_{i=1}^{\phi(\ell, n)} \mathbb{P}[A_i] \right)^2} \leq 1 + 2^{-\ell} \right)$$

From these and Theorem 3.3 we have (26), which implies $\mathbb{P}[A_i \text{ i.o.}] = 1$.

4 Quantitative Kochen–Stone Theorem

As with the quantitative version of the second Borel–Cantelli Lemma, we will also assume that we are given a rate of divergence for the sequence $(\sum_{i=n}^m \mathbb{P}[A_i])_{m \in \mathbb{N}^+}$. Our quantitative version will follow closely the very concise proof of the Kochen–Stone Theorem discovered by Yan [16].

After expressing $\mathbb{P}[A_i \text{ i.o.}]$ as a limit, the Kochen–Stone inequality (1) has the form $\lim_{n \rightarrow \infty} p_n \geq \limsup_{n \rightarrow \infty} q_n$. Much as in Lemma 3.2, we can be more economical with the quantifiers using the following lemma.

Lemma 4.1 *For any sequence of events $(A_i)_{i=1}^\infty$ and sequence of reals $(x_i)_{i=1}^\infty$, the following are equivalent:*

$$(44) \quad \mathbb{P}[A_i \text{ i.o.}] \geq \limsup_{i \rightarrow \infty} x_i$$

$$(45) \quad \forall m, \ell \exists n > m \forall j > n \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq x_j \right)$$

Proof By the definition of $A_i \text{ i.o.}$, (44) is equivalent to:

$$(46) \quad \forall m \left(\mathbb{P} \left[\bigcup_{i=m+1}^\infty A_i \right] \geq \limsup_{i \rightarrow \infty} x_i \right)$$

Let us first show that the above is equivalent to:

$$(47) \quad \forall m, \ell \exists n > m \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \sup_{i > n} x_i \right)$$

Assume (46), and fix m and ℓ . Pick $n_1 > m$ such that

$$(48) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{n_1} A_i \right] + \frac{1}{2^{\ell+1}} > \mathbb{P} \left[\bigcup_{i=m+1}^{\infty} A_i \right]$$

and $n_2 > m$ such that:

$$(49) \quad \limsup_{i \rightarrow \infty} x_i + \frac{1}{2^{\ell+1}} > \sup_{i > n_2} x_i$$

Then, taking $n = \max(n_1, n_2)$, by (46), (48) and (49) we get:

$$(50) \quad \mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} > \mathbb{P} \left[\bigcup_{i=m+1}^{\infty} A_i \right] + \frac{1}{2^{\ell+1}} \geq \limsup_{i \rightarrow \infty} x_i + \frac{1}{2^{\ell+1}} > \sup_{i > n} x_i$$

Thus (46) implies (47). Now, suppose (47) holds but (46) does not. Then, for some m and ℓ , we have that:

$$(51) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{\infty} A_i \right] + \frac{1}{2^\ell} < \limsup_{i \rightarrow \infty} x_i$$

But by (47) we have an $n > m$ such that

$$(52) \quad \mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \sup_{i > n} x_i$$

and hence

$$(53) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{\infty} A_i \right] + \frac{1}{2^\ell} \geq \mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \sup_{i > n} x_i \geq \limsup_{i \rightarrow \infty} x_i$$

contradicting (51). Thus (47) implies, and hence is equivalent to (46). Finally, (47) is equivalent to (45), by the definition of sup. \square

Using the above lemma we can show that the Kochen–Stone inequality (1) can be equivalently written as a $\forall\exists\forall$ -statement:

Lemma 4.2 *The Kochen–Stone inequality (1) is equivalent to:*

$$(54) \quad \forall m, \ell \exists n > m \forall j > n \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \frac{(\sum_{k=1}^j \mathbb{P}[A_k])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \right)$$

Proof This follows from Lemma 4.1 using the sequence $x_j = \frac{(\sum_{k=1}^j \mathbb{P}[A_k])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]}$. \square

As we will see in Section 4.2, there can be no computable function of m and ℓ that bounds n in (54). Therefore, we will consider its metastable counterpart

$$(55) \quad \forall m, \ell, g \exists n > m \forall j \in [n, g(n)] \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \frac{(\sum_{k=1}^j \mathbb{P}[A_k])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \right)$$

and will produce an explicit computable bound on n as a function of m, ℓ and the function g .

Theorem 4.3 (Kochen–Stone: Quantitative Version) *Let $(A_i)_{i=1}^\infty$ be an infinite sequence of events. Let $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ be such that for all N :*

$$\left(\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \right) \geq N$$

Then, for all m and ℓ and $g: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ such that $g(i) > i$, for all i , there exists an $n > m$ such that:

- $n \leq g^{(2^{\ell+1})}(\max(\omega(2^{\ell+2} \sum_{i=1}^m \mathbb{P}[A_i]), m))$, and
- for all $j \in [n, g(n)]$:

$$(56) \quad \mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \frac{(\sum_{i=1}^j \mathbb{P}[A_i])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]}$$

Remark Since $\mathbb{P}[A_i] \leq 1$, we have that $\sum_{i=1}^m \mathbb{P}[A_i] \leq m$. Hence, we can obtain a bound $g^{(2^{\ell+1})}(\omega(2^{\ell+2}m))$ on n which is completely independent of the actual events A_i , but only depends on the parameters ω, g, m and ℓ .

Before we embark on the proof of Theorem 4.3, we need three further lemmas.

Lemma 4.4 (Chung–Erdős Inequality) *For every n and sequence of events A_1, \dots, A_n :*

$$\mathbb{P} \left[\bigcup_{k=1}^n A_k \right] \geq \frac{(\sum_{k=1}^n \mathbb{P}[A_k])^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]}$$

Proof See Chung and Erdős [3] or Yan [16]. □

Lemma 4.5 *Let a and b be such that $0 < a \leq b$. Assume $0 \leq x < a, 0 \leq y < b$ and $0 < \varepsilon$. If $b \geq 4x/\varepsilon^2$ then:*

$$\frac{(\sqrt{a} - \sqrt{x})^2}{b - y} + \varepsilon \geq \frac{a}{b}$$

Proof Since

$$\frac{(\sqrt{a} - \sqrt{x})^2}{b - y} = \frac{a - 2\sqrt{xa} + x}{b - y} \geq \frac{a - 2\sqrt{xa}}{b}$$

it is enough to show that:

$$\frac{2\sqrt{xa}}{b} \leq \varepsilon$$

But since $a \leq b$, that follows from

$$\frac{2\sqrt{xb}}{b} = 2\frac{\sqrt{x}}{\sqrt{b}} \leq \varepsilon$$

which follows from $b \geq 4x/\varepsilon^2$. \square

Lemma 4.6 Let $(A_i)_{i=1}^{\infty}$ and $\omega: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ be as in the statement of Theorem 4.3. Then for all m and $\varepsilon > 0$ and all $j > \max(\omega(\lceil \frac{2}{\varepsilon} \sum_{i=1}^m \mathbb{P}[A_i] \rceil), m)$:

$$\mathbb{P} \left[\bigcup_{i=m+1}^j A_i \right] + \varepsilon \geq \frac{(\sum_{i=1}^j \mathbb{P}[A_i])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]}$$

Proof Let m and $\varepsilon > 0$ be fixed. Let² $a_n = (\sum_{i=1}^n \mathbb{P}[A_i])^2$ and $b_n = \sum_{i,k=1}^n \mathbb{P}[A_i A_k]$. The assumption in the statement of Theorem 4.3 says that $\sqrt{a_n}$ diverges with rate ω . By Lemma 4.4, $b_n \geq a_n$. Hence, for all N :

$$b_{\omega(N)} \geq a_{\omega(N)} \geq N^2$$

Since $\sum_{i,k=m+1}^n \mathbb{P}[A_i A_k] \leq b_n - b_m$, by Lemma 4.4 we have for all $j > m$:

$$(57) \quad \mathbb{P} \left[\bigcup_{i=m+1}^j A_i \right] + \varepsilon \geq \frac{(\sqrt{a_j} - \sqrt{a_m})^2}{b_j - b_m} + \varepsilon$$

Let $M = \max(\omega(\lceil \frac{2}{\varepsilon} \sum_{i=1}^m \mathbb{P}[A_i] \rceil), m)$ and let $j > M$. By assumption we have that:

$$b_j \geq b_{\omega(\lceil \frac{2}{\varepsilon} \sum_{i=1}^m \mathbb{P}[A_i] \rceil)} \geq \left(\frac{2}{\varepsilon} \sum_{i=1}^m \mathbb{P}[A_i] \right)^2 \geq \frac{4a_m}{\varepsilon^2}$$

By Lemma 4.5, we have that:

$$(58) \quad \frac{(\sqrt{a_j} - \sqrt{a_m})^2}{b_j - b_m} + \varepsilon \geq \frac{a_j}{b_j} = \frac{(\sum_{i=1}^j \mathbb{P}[A_i])^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]}$$

The result follows from (57) and (58). \square

²In the notation of the proof of Lemma 3.1, $a_n = \mathbb{E}^2(Y_n)$ and $b_n = \mathbb{E}(Y_n^2)$.

Proof of Theorem 4.3 Let $n_0 = \max(\omega(\lceil 2^{\ell+2} \sum_{i=1}^m \mathbb{P}[A_i] \rceil), m)$ and $n_{r+1} = g(n_r)$. We claim that with $n = n_r$ for some $r \leq 2^{\ell+1}$, the conclusion holds. Assume this is not the case. Then for each r there is a $j_r \in [n_r, n_{r+1}]$ such that:

$$(59) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{n_r} A_i \right] + \frac{1}{2^\ell} < \frac{\left(\sum_{i=1}^{j_r} \mathbb{P}[A_i] \right)^2}{\sum_{i,k=1}^{j_r} \mathbb{P}[A_i A_k]}$$

But by Lemma 4.6, with $\varepsilon = \frac{1}{2^{\ell+1}}$

$$(60) \quad \frac{\left(\sum_{i=1}^{j_r} \mathbb{P}[A_i] \right)^2}{\sum_{i,k=1}^{j_r} \mathbb{P}[A_i A_k]} \leq \mathbb{P} \left[\bigcup_{i=m+1}^{j_r} A_i \right] + \frac{1}{2^{\ell+1}}$$

and hence, combining (59) and (60) and subtracting $\frac{1}{2^{\ell+1}}$, we have:

$$(61) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{n_r} A_i \right] + \frac{1}{2^{\ell+1}} < \mathbb{P} \left[\bigcup_{i=m+1}^{j_r} A_i \right] \leq \mathbb{P} \left[\bigcup_{i=m+1}^{n_{r+1}} A_i \right]$$

Chaining together the inequalities given by (61) for $r = 0$ to $2^{\ell+1}$, we have

$$(62) \quad \mathbb{P} \left[\bigcup_{i=m+1}^{n_0} A_i \right] + 1 < \mathbb{P} \left[\bigcup_{i=m+1}^{n_{2^{\ell+1}}} A_i \right] \leq 1$$

which is a contradiction. □

4.1 Proving qualitative version from quantitative one

Let us argue that the qualitative version of the Kochen–Stone Theorem (Theorem 4.3) directly implies the original qualitative version. Theorem 4.3 implies that for all m, ℓ and $g: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ (with $g(i) > i$) there exists an n such that:

$$\forall j \in [n, g(n)] \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \right)$$

But (see Section 1.1), the above is equivalent to: for all m, ℓ there exists an n such that:

$$\forall j \geq n \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \right)$$

Then, this is equivalent to: For all m, ℓ there exists an n such that:

$$\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \sup_{j \geq n} \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]}$$

Hence, for all m and ℓ there exists an n such that:

$$\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \sup_{j \geq n} \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \geq \limsup_{n \rightarrow \infty} \left(\frac{\left(\sum_{k=1}^n \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]} \right)$$

Therefore, for all m

$$\mathbb{P} \left[\bigcup_{i=m+1}^{\infty} A_i \right] \geq \limsup_{n \rightarrow \infty} \left(\frac{\left(\sum_{k=1}^n \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]} \right)$$

and hence:

$$\mathbb{P}[A_i \text{ i.o.}] \geq \limsup_{n \rightarrow \infty} \left(\frac{\left(\sum_{k=1}^n \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]} \right)$$

4.2 Necessity for use of metastability

We wish to show that there is no effective bound on the witness n in (54), so that the approach via metastability is necessary. To this end, we will need examples where the Kochen–Stone inequality (1) is actually an equality with $\mathbb{P}[A_i \text{ i.o.}] < 1$. To do this we will use the following result of Yan which shows that the diagonal terms in the sums on the right-hand side of the inequality are negligible. Yan’s sketch of the proof in [16] is very terse, so we give more detail here.

Theorem 4.7 *Let $(A_i)_{i=1}^{\infty}$ be an infinite sequence of events such that $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$. Then:*

$$(63) \quad \limsup_{n \rightarrow \infty} \frac{\left(\sum_{k=1}^n \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^n \mathbb{P}[A_i A_k]} = \limsup_{n \rightarrow \infty} \frac{\sum_{1 \leq i < k \leq n} \mathbb{P}[A_i] \mathbb{P}[A_k]}{\sum_{1 \leq i < k \leq n} \mathbb{P}[A_i A_k]}$$

Proof Define sequences s_n, t_n, b_n and c_n as follows:

$$\begin{aligned} s_n &= \sum_{k=1}^n \mathbb{P}[A_k] & t_n &= \sum_{1 \leq i < k \leq n} \mathbb{P}[A_i] \mathbb{P}[A_k] \\ b_n &= \sum_{i,k=1}^n \mathbb{P}[A_i A_k] & c_n &= \sum_{1 \leq i < k \leq n} \mathbb{P}[A_i A_k] \end{aligned}$$

Since $\sum_{i=1}^{\infty} \mathbb{P}[A_i] = \infty$, we have $s_n = o(s_n^2)$. Hence, as $2t_n \leq s_n^2 \leq 2t_n + s_n$, $\lim_{n \rightarrow \infty} s_n^2/2t_n = 1$. By inequality (1), $s_n^2 \leq (1 + o(1))b_n$. Hence, as $2c_n \leq b_n = 2c_n + s_n$, $\lim_{n \rightarrow \infty} b_n/2c_n = 1$. It follows that $\limsup_{n \rightarrow \infty} s_n^2/b_n = \limsup_{n \rightarrow \infty} t_n/c_n$, which is what we wish to prove. \square

Let $(q_n)_{n=1}^{\infty}$ be any non-decreasing sequence of elements of the open unit interval $(0, 1)$, let $q = \lim_{n \rightarrow \infty} q_n$ and let A_i be the event that a uniformly random member of the unit interval $[0, 1]$ lies in $[0, q_i]$. Then $\mathbb{P}[A_i] = q_i$ and $\mathbb{P}[A_i \text{ i.o.}] = q$. Moreover, $\mathbb{P}[A_i A_k] = \mathbb{P}[A_i]$ for $i < k$. Let us define u_n, v_n and w_n as follows:

$$(64) \quad u_n = q_1 \sum_{k=2}^n q_k + q_2 \sum_{k=3}^n q_k + \dots + q_{n-1} q_n$$

$$(65) \quad v_n = (n-1)q_1 + (n-2)q_2 + \dots + q_{n-1}$$

$$(66) \quad w_n = \frac{u_n}{v_n}$$

Then, rearranging the terms in the sums on the right-hand side of equation (63), we find that, given equation (63), the inequality (1) is equivalent to $q \geq \limsup_{n \rightarrow \infty} w_n$. The following lemma implies that equality holds in the Kochen–Stone inequality for any sequence of events $(A_i)_{i=1}^{\infty}$ constructed in this way.

Lemma 4.8 *Let w_n and q be as above. Then $w_n \rightarrow q$ as $n \rightarrow \infty$.*

Proof We have:

$$(67) \quad \sum_{k=i}^n q_k = q \left(n - i + 1 - \frac{1}{q} \sum_{k=i}^n (q - q_k) \right)$$

Let us write σ_i^j for $\sum_{k=i}^j (q - q_k)$. From the above, multiplying by q_{i-1} and summing for i from 2 to n , we have:

$$(68) \quad u_n = qv_n - \sum_{i=2}^n q_{i-1} \sigma_i^n$$

Define the sequence r_n by:

$$(69) \quad r_n = \frac{\sum_{i=2}^n q_{i-1} \sigma_i^n}{v_n}$$

We claim that $r_n \rightarrow 0$ as $n \rightarrow \infty$ so that $w_n = u_n/v_n = (q - r_n) \rightarrow q$, which is what we have to prove. So given $\varepsilon > 0$, let $\varepsilon_0 = \frac{q_1}{4} \varepsilon$ and choose N such that for all $n > N$, we have $q - q_n < \varepsilon_0$. Define C by:

$$(70) \quad C = q_1 \sigma_2^N + q_2 \sigma_3^N + \dots + q_{N-1} \sigma_N^N$$

Then, for $n > N$

$$(71) \quad r_n = \frac{C}{v_n} + \frac{q_1 \sigma_{N+1}^n + \dots + q_N \sigma_{N+1}^n + q_{N+1} \sigma_{N+2}^n + \dots + q_{n-1} \sigma_n^n}{(n-1)q_1 + (n-2)q_2 + \dots + q_{n-1}}$$

$$(72) \quad \leq \frac{C}{v_n} + \frac{(n-1)\sigma_{N+1}^n}{q_1((n-1) + (n-2) + \dots + 1)}$$

$$(73) \quad \leq \frac{C}{v_n} + \frac{1}{q_1} \cdot \frac{(n-1)(n-N)\varepsilon_0}{\frac{1}{2}(n-1)n}$$

$$(74) \quad \rightarrow \frac{2}{q_1} \varepsilon_0 = \frac{\varepsilon}{2} \text{ as } n \rightarrow \infty$$

where in equation (71) we have expanded the denominator of the second fraction using the definition of v_n and where the bounds (72) and (73) are obtained using the facts that $0 < q_1 \leq q_i < 1$ and $\sigma_{N+m}^n \leq \sigma_{N+1}^n \leq (n-N)\varepsilon_0$. Hence we can choose $M > N$, such that for $n > M$, we have $|r_n - \frac{\varepsilon}{2}| < \frac{\varepsilon}{2}$, giving $r_n < \varepsilon$. Hence $r_n \rightarrow 0$ as $n \rightarrow \infty$. \square

Recall that a *Specker sequence* (see Specker [11] or Toelstra and van Dalen [14]) is a computable, monotone increasing, bounded sequence of rationals whose limit is not a computable real number.³

Theorem 4.9 *There is a sequence of events $(A_i)_{i=1}^\infty$ and a computable function $\omega : \mathbb{N}^+ \rightarrow \mathbb{N}^+$ such that, for any N*

$$\left(\sum_{i=1}^{\omega(N)} \mathbb{P}[A_i] \right) \geq N$$

but for which there is no computable function $\phi : \mathbb{N} \times \mathbb{N}^+ \rightarrow \mathbb{N}^+$ satisfying:

$$(75) \quad \forall m, \ell \exists n \in [m, \phi(m, \ell)] \left(\mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] + \frac{1}{2^\ell} \geq \limsup_{j \rightarrow \infty} \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k] \right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} \right)$$

Proof Take A_i to be the event that a uniformly random element of $[0, 1]$ lies in the interval $[0, q_i]$ where the $q_i \in (0, 1)$ form a Specker sequence with limit q (note that $q < 1$, since q is not a computable real). Also $\sum_{i=1}^\infty q_i$ diverges with the rate of divergence given by the computable function $\omega(N) = \lceil \frac{N}{q_1} \rceil$. Then, for $n > m$, we have

$$(76) \quad \mathbb{P} \left[\bigcup_{i=m+1}^n A_i \right] = \mathbb{P}[A_n] = q_n$$

³A real number is said to be *computable* if it is the limit of a computable sequence of rationals, with a computable rate of convergence (see Weihrauch [15]).

and, by Lemma 4.8 and the discussion preceding it, we have:

$$(77) \quad \limsup_{j \rightarrow \infty} \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k]\right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} = \lim_{j \rightarrow \infty} \frac{\left(\sum_{k=1}^j \mathbb{P}[A_k]\right)^2}{\sum_{i,k=1}^j \mathbb{P}[A_i A_k]} = q$$

Assume ϕ satisfies (75), ie:

$$(78) \quad \forall m, \ell \exists n \in (m, \phi(m, \ell)] \left(q_n + \frac{1}{2^\ell} \geq q \right)$$

Hence

$$(79) \quad \forall \ell \exists n \leq \phi(0, \ell) \left(|q - q_n| \leq \frac{1}{2^\ell} \right)$$

implying that, $g(\ell) = \phi(0, \ell)$ is a rate of convergence for the Specker sequence q_n . But since $q = \lim_{n \rightarrow \infty} q_n$ is not a computable real, it cannot be approximated by a sequence with a computable rate of convergence. It follows that g , and hence also ϕ , is not a computable function. \square

5 Optimality of the estimates

It is easy to argue that the numeric bounds in Theorem 2.1 are best possible. Indeed, let $(A_i)_{i=1}^\infty$ be a sequence of mutually exclusive events. In this case, the first inequality in (11) is actually an equality and hence the given estimate is optimal.

The estimate given by Theorem 2.2 is also the best possible amongst estimates that do not depend on n . To see this, consider the probability space whose outcomes are functions $\alpha: \mathbb{N}^+ \rightarrow \{1, \dots, k\}$ representing an infinite sequence of throws of a fair k -sided die. Let A_i be the event $\alpha(i) = k$, so that $\mathbb{P}[A_i] = 1/k$. Clearly $\sum_{i=1}^\infty \mathbb{P}[A_i]$ diverges with (optimal) rate $\omega(N) = kN$. We have that:

$$(80) \quad \mathbb{P} \left[\bigcap_{i=n}^{k(n+N-1)} \bar{A}_i \right] = \left(1 - \frac{1}{k} \right)^{k(n+N-1)-n+1}$$

This is a decreasing function of n , so the worst case for our estimate is when $n = 1$, but in that case we have

$$(81) \quad \mathbb{P} \left[\bigcap_{i=1}^{kN} \bar{A}_i \right] = \left(1 - \frac{1}{k} \right)^{kN} \rightarrow e^{-N} \quad \text{as } k \rightarrow \infty$$

showing that the bounds in (18) and (19) are tight.

Unlike the quantitative proofs of the First and Second Borel–Cantelli Lemmas, where we have argued above that the numerical bounds obtained have optimal rate of growth, in the case of the quantitative Erdős–Rényi Theorem (Theorem 3.3), it is plausible that the bounds we have achieved are optimal for the given formulation of assumption (25) (see remark after Theorem 3.3), but we have been unable to prove that.

In the case of the quantitative Kochen–Stone Theorem (Theorem 4.3), we have had to reformulate the original theorem using Tao’s notion of metastability. This means that the numerical bound given also depends on a new functional input $g: \mathbb{N} \rightarrow \mathbb{N}$. This makes an asymptotic analysis of the growth rate of the bound much more difficult. It is again unclear whether the bound we have obtained has optimal dependency on any of the parameters ω, g, m and ℓ . We leave this for future work. It is worth emphasising, however, that all our numerical bounds are uniform in (ie independent of) the actual sequence of events $(A_i)_{i=1}^{\infty}$.

Acknowledgements

We would like to thank Iosif Pinelis for the proof of Theorem 4.7, and the anonymous referees for their constructive feedback, corrections and suggestions.

References

- [1] **J Avigad**, *The metamathematics of ergodic theory*, Annals of Pure and Applied Logic 157 (2009) 64–76; <https://doi.org/10.1016/j.apal.2008.09.001>
- [2] **V Beresnevich, S Velani**, *Classical metric Diophantine approximation revisited: the Khintchine–Groshev theorem*, Int. Math. Res. Not. 2010 (2010) 69–86; <https://doi.org/10.1093/imrn/rnp119>
- [3] **KL Chung, P Erdős**, *On the application of the Borel–Cantelli lemma*, Trans. Am. Math. Soc. 72 (1952) 179–186; <https://doi.org/10.2307/1990661>
- [4] **P Erdős, A Rényi**, *On Cantor’s series with convergent $\sum 1/q_n$* , Ann. Univ. Sci. Budap. Rolando Eötvös, Sect. Math. 2 (1959) 93–109
- [5] **W Feller**, *An Introduction to Probability Theory and Its Applications. I. Third Edition*, John Wiley and Sons, Inc. (1968)
- [6] **PX Gallagher**, *Metric simultaneous diophantine approximation. II*, Mathematika 12 (1965) 123–127; <https://doi.org/10.1112/jlms/s1-37.1.387>
- [7] **S Kochen, C Stone**, *A note on the Borel–Cantelli lemma*, Ill. J. Math. 8 (1964) 248–251

- [8] **U Kohlenbach**, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Monographs in Mathematics, Springer (2008); <https://doi.org/10.1007/978-3-540-77533-1>
- [9] **U Kohlenbach, L Leustean**, *A quantitative Mean Ergodic Theorem for uniformly convex Banach spaces*, *Ergodic Theory and Dynamical Systems* 29 (2009) 1907–1915; <https://doi.org/10.1017/S0143385708001004>
- [10] **K Schade, U Kohlenbach**, *Effective metastability for modified Halpern iterations in $CAT(0)$ spaces*, *Fixed Point Theory Appl.* 191 (2012) 19pp.; <https://doi.org/10.1186/1687-1812-2012-191>
- [11] **E Specker**, *Nicht konstruktiv beweisbare Sätze der Analysis*, *J. Symb. Log.* 14 (1949) 145–158; <https://doi.org/10.2307/2267043>
- [12] **V G Sprindzuk**, *Metric Theory of Diophantine Approximations*, John Wiley & Sons, New York (1979)
- [13] **T Tao**, *Structure and randomness. Pages from year one of a mathematical blog*, American Mathematical Society (2008)
- [14] **A S Troelstra, D van Dalen**, *Constructivism in Mathematics. An Introduction. Volume I*, volume 121 of *Stud. Logic Found. Math.*, North-Holland (1988)
- [15] **K Weihrauch**, *Computable Analysis*, Springer, Berlin (2000); <https://doi.org/10.1007/978-3-642-56999-9>
- [16] **J Yan**, *A simple proof of two generalized Borel–Cantelli lemmas*, from: “In memoriam Paul-André Meyer. Séminaire de probabilités XXXIX”, (M Émery, M Yor, editors), *Lecture Notes in Mathematics 1874*, Springer (2006) 77–79; <https://doi.org/10.1007/b128398>

School of Electronic Engineering and Computer Science, Queen Mary University of London
r.arthan@qmul.ac.uk, p.oliva@qmul.ac.uk

Received: 12 January 2019 Revised: 15 December 2021