Contextuality in composite systems: the role of entanglement in the Kochen–Specker theorem

Victoria J Wright[1] and Ravi Kunjwal[2]

[1]ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels, Spain
[2]Centre for Quantum Information and Communication, Ecole polytechnique de Bruxelles, CP 165, Université libre de Bruxelles, 1050 Brussels, Belgium

The Kochen–Specker (KS) theorem reveals the nonclassicality of single quantum systems. In contrast, Bell's theorem and entanglement concern the nonclassicality of composite quantum systems. Accordingly, unlike incompatibility, entanglement and Bell nonlocality are not necessary to demonstrate KS-contextuality. However, here we find that for multiqubit systems, entanglement and non-locality are *both* essential to proofs of the Kochen–Specker theorem. Firstly, we show that unentangled measurements (a strict superset of local measurements) can never yield a logical (state-independent) proof of the KS theorem for multiqubit systems. In particular, unentangled but nonlocal measurements—whose eigenstates exhibit "nonlocality without entanglement"—are insufficient for such proofs. This also implies that proving Gleason's theorem on a multiqubit system necessarily requires entangled projections, as shown by Wallach [Contemp Math, 305: 291-298 (2002)]. Secondly, we show that a multiqubit state admits a statistical (state-dependent) proof of the KS theorem if and only if it can violate a Bell inequality with projective measurements. We also establish the relationship between entanglement and the theorems of Kochen–Specker and Gleason more generally in multiqudit systems by constructing new examples of KS sets. Finally, we discuss how our results shed new light on the role of multiqubit contextuality as a resource within the paradigm of quantum computation with state injection.

Victoria J Wright: victoria.wright@icfo.eu
Ravi Kunjwal: ravi.kunjwal@ulb.be

1 Introduction

Quantum theory's 'departure from classical lines of thought' [1] is today a driving force behind the promise of quantum technologies. Quantum theory abandons assumptions implicit in classical physical theories, such as the assumptions that physics must be fundamentally deterministic or that all observables are jointly measurable. This allows for an array of typically quantum phenomena such as entanglement, uncertainty relations, Bell nonlocality, contextuality, etc. We generically refer to the possibility of these previously forbidden properties as the *nonclassicality* of quantum theory. Such nonclassical properties of quantum theory are key to the many advantages quantum information processing and quantum computation hold over their classical counterparts. However, the exact relationship between different notions of nonclassicality and such advantages is often unclear and remains an active area of research [2, 3, 4, 5, 6, 7, 8, 9].

The case of multiqubit systems is of particular importance given their ubiquity throughout quantum technologies, particularly noisy intermediate-scale quantum (NISQ) technologies [10, 11]. However, the nonclassicality of qubit systems remains an anomalous case. In-

dividually, for example, a qubit cannot display Kochen–Specker (KS) contextuality [12] while, collectively, multiqubit systems (which exhibit KS-contextuality) derail the neat narrative of such contextuality powering quantum computational advantage [5, 13]. Furthermore, entanglement is often considered a key indicator of nonclassicality in these systems but the sense in which it relates to fundamental notions of nonclassicality witnessed by the theorems of Bell [14, 15], Kochen–Specker [12], and Gleason [16] needs clarification. Since the question of nonclassicality is essentially a foundational one [17], we approach the study of multiqubit systems through this lens.

Entanglement is an intrinsically compositional property and is, therefore, only relevant to the study of nonclassicality in composite (i.e. multipartite) systems. Schrödinger claimed the entanglement of quantum states to be 'the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought' [1]. Bell's theorem [14, 15] exemplifies this point by revealing the nonclassicality of correlations arising from local measurements on composite quantum systems in entangled states, the simplest case being a two-qubit system.

The Kochen–Specker theorem [12], on the other hand, can reveal the nonclassicality of correlations between measurements implemented on an *indivisible* quantum system, the simplest case being a qutrit. For quantum systems of dimension at least three, the theorem states that there cannot exist an underlying ontological model—known as a *KS-noncontextual* ontological model—that reproduces the predictions of quantum theory. In such a model the outcomes of projective measurements are fully determined by the ontic state of the system. Furthermore, the outcomes are independent of *context*[1] and respect the functional relationships between commut-

ing measurements.[2]

The key insight of Kochen and Specker was that one can witness the impossibility of such models with a *finite* set of rank-1 projections (a *KS set*) on a three-dimensional Hilbert space. A KS set thus constitutes a *logical* proof of the KS theorem [18], exemplified by the original result of Kochen and Specker [12]. It is also possible to demonstrate the inadequacy (rather than impossibility) of KS-noncontextual ontological models in a (weaker) *statistical* sense. Such statistical proofs of the KS theorem [19] are exemplified by the proof due to Klyachko *et al.* [20].

Prior to the Bell and Kochen–Specker theorems, Gleason's theorem [16] demonstrated that, for any quantum system of dimension at least three, the unique way to assign probabilities to the outcomes of projective measurements is via the Born rule. In particular, Gleason's theorem excludes any deterministic probability rule given by a $\{0, 1\}$-valued assignment of probabilities to all the self-adjoint projections on the system's Hilbert space. This exclusion thus implies the KS theorem, but, unlike the proof of Kochen and Specker [12], it requires an uncountably infinite KS set.

A single qubit does not support any of the three theorems (Bell, KS, or Gleason) and is, by that token, rather "classical".[3] A single qutrit can support the Gleason and KS theorems but not Bell's theorem. Hence, the smallest quantum system on which one can meaningfully study the interplay of Gleason, Bell, and KS theorems is a two-qubit sys-

---

[1]The context of a projective measurement here refers to other projective measurements with which it is jointly performed.

[2]For example, the outcome of measuring an observable with operator A^2 should be the square of the outcome of measuring A.

[3]This assumes a restriction to pure states and projective measurements on a qubit, as is traditionally the case in these theorems. Such a 'pure' qubit can exhibit nonclassicality in other (weaker) respects like the existence of incompatible measurements, something reproducible in classical theories with an epistemic restriction [17]. However, once mixed states and generalised measurements are included, a single qubit can support proofs of generalised contextuality [21, 22, 23]. Allowing generalised measurements also permits Gleason-type theorems for the qubit case [24, 25, 26].

tem. The nonclassicality witnessed by Bell's theorem in qubit pairs (and more generally) clearly depends on entanglement, since Bell inequality violations can only be observed when the quantum systems used are described by an entangled state. Is entanglement, however, also necessary for the theorems of KS and Gleason in a two-qubit system, and more generally, in multiqubit systems? Since both theorems appeal to the structure of quantum measurements, we need to go beyond states and consider the role of entanglement in measurements as well.

The question of entanglement and Gleason's theorem is already resolved. A result by Wallach [27] showed that the set of unentangled multiqudit projections yields a proof of Gleason's theorem if and only if each qudit has a Hilbert space dimension three or more. In particular, the set of unentangled multiqubit projections cannot yield Gleason's theorem.

This work will largely consider rank-one projective measurements that are *unentangled*, i.e. measurements in bases comprising only product vectors. Unlike quantum states, measurements can be "nonlocal" without being entangled, that is, there exist measurements which cannot be implemented via local operations and classical communication (LOCC) but nevertheless only involve projections onto unentangled subspaces.[4] For example, one cannot perform a measurement in the unentangled three qubit basis

$$\begin{aligned} \{ & |000\rangle, |+10\rangle, |0+1\rangle, |10+\rangle, \\ & |111\rangle, |-10\rangle, |0-1\rangle, |10-\rangle \}, \end{aligned} \quad (1)$$

via LOCC. It could in principle be that "nonlocality" of unentangled measurements is sufficient to recover a logical proof of the KS theorem [12, 18]. However, in the first main result of this work, Theorem 3, we show that this is not the case: the unentangled rays of a multiqubit system are KS-colourable and,

---

[4]The eigenstates for such a measurement exhibit a phenomenon called 'nonlocality without entanglement', i.e. they form a set of mutually orthogonal states that cannot be distinguished perfectly via LOCC [28].

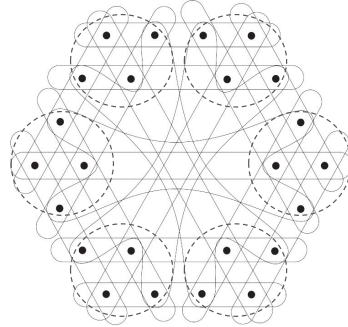| $X \otimes I$ | $I \otimes X$ | $X \otimes X$ |
| $I \otimes Y$ | $Y \otimes I$ | $Y \otimes Y$ |
| $X \otimes Y$ | $Y \otimes X$ | $Z \otimes Z$ |

Figure 1: The Peres-Mermin square [29] consisting of two-qubit Pauli matrices together with the contextuality scenario defined by the orthogonality relations between Peres's 24 rays [30]. Each row or column of the Peres-Mermin square is associated to a two-qubit orthonormal basis in which all the operators in that row or column are diagonal. In the contextuality scenario, the six dashed hyperedges denote the six orthonormal bases corresponding to the rows and columns of the Peres-Mermin square. In particular, the basis that diagonalises the third column, $\{XX, YY, ZZ\}$, is the Bell basis.

thus, no unentangled form of nonclassicality suffices for a logical proof of KS-contextuality in this setting.

The well-known proof of the KS theorem via the Peres-Mermin square [31, 30, 29] at first appears to not involve entanglement (see Fig. 1). However, the joint measurement of the Pauli observables $\{XX, YY, ZZ\}$, for example, necessitates a measurement in the Bell basis. In this sense, our result shows that this entanglement is not accidental but, rather, unavoidable; any logical (hence, state-independent) [18] proof of the KS theorem for multiqubit systems necessarily requires entangled measurements.

We also construct a KS-noncontextual ontological model for the fragment of multiqubit quantum theory containing unentangled measurements and product states. The model can be viewed as a generalization of the single-qubit model of Kochen and Specker [12], but the proof of its validity relies on our result—Lemma 2—to show that the on-

tic states are indeed valid KS-noncontextual ontic states. The model admits a simple extension to the case of separable states which renders it preparation contextual [21], but still KS-noncontextual.

The existence of this model implies that for a proof of the KS theorem on a multiqubit system, one requires either (i) entanglement in the measurements, in which case one can provide a logical (and *state-independent*) proof, or (ii) entanglement in the state, in which case the violation of Bell inequalities, for example, provides a *state-dependent* proof without any entangled measurements [32]. In the second main result of this paper, we demonstrate that such Bell inequality violations are the *only* way to prove the KS theorem in this setting. We show that an entangled state enables a (finite) statistical proof of the KS theorem with unentangled measurements if and only if it also violates a Bell inequality with local projective measurements.

Thus, we must conclude that, just as in the case of Bell's and Gleason's theorems, the nonclassicality of multiqubit systems is underpinned by entanglement in the case of the Kochen–Specker theorem. This discovery is in surprising contrast to the usual intuition that takes the KS theorem as witnessing nonclassicality that is independent of entanglement because it applies, for example, to a single qutrit [12]. Hence, in the simplest case where *all* three theorems apply, i.e. a two-qubit system, entanglement is necessary for all of them.

Exploring the relationship between the KS theorem and entanglement in multiqudit systems further, we provide two new constructions of KS sets on multiqudit systems that allow us to obtain an overall picture of this relationship (Fig. 2). As displayed in Fig. 2, our results recover the result of Wallach [27] in the multiqubit case as a corollary, namely, that entangled projections are necessary to obtain Gleason's theorem.

Finally, we discuss implications of our results for the role of contextuality in multiqubit quantum computation with state injection. In particular, we highlight how the choice of measurements in the schemes of Ref. [13] ap-

when $\dim(\mathcal{H}_j)$:	≥ 3 for all j	2 for some j ≥ 3 for some j	2 for all j
Gleason (Unentangled)	yes $\Downarrow$	no	no $\Uparrow$
K-S (Unentangled)	yes +**direct**	**yes** +**direct**	**no**

Figure 2: The existence of logical proofs of the KS theorem and proofs of Gleason's theorem without entanglement on a Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, where $1 \leq j \leq n$. The implication arrows show which results follow from each other. The addition of +direct signifies the result also holds for the subset of unentangled measurements given by direct products bases, see Sec. 7. The results in bold and purple are introduced in the present work.

pear natural in view of our results.

The structure of the paper is as follows. In Sec. 2, we provide preliminary notions that we will need in the rest of the paper. In Sec. 3, we prove our first main result, i.e. the necessity of entanglement for multiqubit KS sets. Sec. 4 presents a KS-noncontextual ontological model for product states and unentangled measurements of multiqubit systems. Sec. 5 establishes our second main result: an entangled state yields a statistical proof of the KS theorem if and only if it yields a proof of Bell's theorem. In Sec. 6, we construct a two-qubit KS set without any fully entangled bases. Sec. 7 proves the existence of unentangled KS sets on any multiqudit system that contains at least one qudit of Hilbert space dimension three or more. Sec. 8 discusses implications of our results for the role of contextuality in multiqubit schemes for quantum computation with state injection (QCSI). Sec. 9 concludes with a discussion of our results.

2 Preliminaries

Given a separable Hilbert space, $\mathcal{H}$, a self-adjoint projection on $\mathcal{H}$ is a linear operator Π satisfying $\Pi^2 = \Pi^\dagger = \Pi$, where $\Pi^\dagger$ denotes the adjoint of Π. We will denote the set of self-adjoint projections on $\mathcal{H}$ by $\mathcal{P}(\mathcal{H})$. In this work, we will consider *projective measurements* on quantum systems, which we de-

scribe by sets $\{\Pi_1, \Pi_2, \ldots\}$ of self-adjoint, mutually orthogonal projections $\Pi_j \in \mathcal{P}(\mathcal{H})$ on a separable Hilbert space, $\mathcal{H}$, that sum to the identity operator.

Rank-one projections are generally sufficient for our purposes so we equivalently consider rays in the projective Hilbert space $\mathcal{R}(\mathcal{H})$. The projective Hilbert space $\mathcal{R}(\mathcal{H})$ is given by the set of equivalence classes, or *rays*, of non-zero vectors in $\mathcal{H}$, under the equivalence relation $\psi \sim \chi$ if and only if $\psi = \alpha\chi$ for some non-zero $\alpha \in \mathbb{C}$. We represent each ray with one of its unit vectors, which we denote by a "ket", such as $|\psi\rangle$. For brevity, we will often refer to this vector as the ray it represents.

There is a bijection between the sets of rays and rank-one projections on a Hilbert space. A rank-one projective measurement is therefore uniquely defined by a *complete set* of rays, i.e. a set of d mutually orthogonal rays, where d is the (possibly infinite) dimension of the Hilbert space. A complete set is represented by an orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots\}$ of $\mathcal{H}$, and thus we say we are performing a measurement *in a basis* and often refer to a complete set of rays as simply a basis.

The KS theorem states the impossibility of an outcome-deterministic and measurement noncontextual ontological model [21] (briefly, a *KS-noncontextual* ontological model) for quantum theory. We will refer to this fact as the KS-contextuality of quantum theory. The original result due to Kochen and Specker [12] provides a *logical* proof of the KS theorem [18] and we state it below in a formulation most relevant to the present work:

Theorem 1 (Kochen–Specker [12]). *Given a separable Hilbert space $\mathcal{H}$ of dimension at least three, there does not exist a map*

$$c : \mathcal{R} \to \{0, 1\} \tag{2}$$

such that for any complete set of rays $\{|\psi_1\rangle, |\psi_2\rangle, \ldots\}$, we have $c(|\psi_j\rangle) = 1$ for exactly one value of $j \in \{1, 2, \ldots\}$.

This formulation of the result implies the traditional statement of the Kochen–Specker theorem in terms of *valuation* functions on

self-adjoint operators (see Appendix A). The KS theorem also admits *statistical* proofs, whereby it is shown that the statistics produced by a collection of measurements performed on a fixed quantum state cannot be derived from a KS-noncontextual ontological model [20, 19].[5] We will discuss these proofs of the KS theorem, and how they differ from logical proofs, more extensively in Sec. 5.

A set of rays in a Hilbert space can be represented by a hypergraph in which there is a vertex for each ray and each complete set of rays constitutes a hyperedge. Such a hypergraph is an example of a *contextuality scenario* [32]. The map c in Theorem 1 then defines a special type of 2-colouring of this hypergraph, which we will call a *KS-colouring*.[6] A set of rays for which there does not exist a KS-colouring is known as a *KS set*.

A KS-colouring of a set of rays is more than simply a mathematical tool: KS-colourings define the ontic states in a KS-noncontextual ontological model. The ontic state determines the outcome of any measurement with certainty, and this choice of one deterministic outcome from each measurement is exactly a KS-colouring. The existence of a KS set in any Hilbert space of dimension greater than two then proves the KS theorem (via Theorem 1)

The fact that all the rays in a Hilbert space, $\mathcal{H}$, of dimension three or more form a KS set (i.e., a proof of the KS theorem) follows from Gleason's theorem. Kochen and Specker [12],

---

[5]The notion of contextuality was extended beyond the Kochen–Specker notion to a generalised notion of contextuality in Ref. [21]. KS-noncontextuality, within this generalised framework, is recovered as a conjunction of two assumptions on ontological models of any operational theory: measurement noncontextuality and outcome determinism for projective measurements. As we will not use much of the machinery of generalised contextuality in this paper (and since we are focusing on quantum theory rather than operational theories in general), we refer the interested reader to Refs. [21, 18, 23, 19, 33, 34] for discussions of generalised contextuality and its connection with KS-contextuality.

[6]Here "0" and "1" stand in for two possible "colours" that could be assigned to vertices according to the rule specified in Theorem 1.

however, gave an explicit construction of a *finite* KS set in a three-dimensional Hilbert space, arguably providing a much simpler proof of the KS theorem than relying on Gleason's theorem.[7]

Theorem 2 (Gleason [16])**.** *Let $\mathcal{H}$ be a separable Hilbert space of dimension at least three. Any map $f : \mathcal{P}(\mathcal{H}) \to [0,1]$ satisfying*

$$f(\Pi_1) + f(\Pi_2) + \cdots = f(\Pi_1 + \Pi_2 + \cdots) , \quad (3)$$

for any set of mutually orthogonal projections $\{\Pi_1, \Pi_2, \ldots\}$, and $f(\mathrm{I}_{\mathcal{H}}) = 1$ where $\mathrm{I}_{\mathcal{H}}$ is the identity operator on $\mathcal{H}$, admits an expression

$$f(\Pi) = \mathrm{Tr}(\Pi \rho) , \quad (4)$$

for some density operator ρ on $\mathcal{H}$.

The maps f in Gleason's theorem are known as *frame functions*. Any KS-colouring, c, on the rays (and equivalently the rank-one projections) of a Hilbert space $\mathcal{H}$ would extend to a $\{0,1\}$-valued frame function on $\mathcal{P}(\mathcal{H})$ via $c(\Pi_1 + \Pi_2 + \cdots) \equiv c(\Pi_1) + c(\Pi_2) + \cdots$ for mutually orthogonal sets of rank-one projections $\{\Pi_1, \Pi_2, \ldots\}$. In dimensions greater than two, however, Gleason's theorem shows that such a frame function does not exist, hence the set of all rays is not KS-colourable.

Neither the KS theorem nor Gleason's theorem hold for a single two-level system, i.e. a qubit. The theorems do, however, hold for systems of multiple qubits. We will examine whether the onset of the applicability of these theorems is due to the presence entanglement in the rays of the measurements or if it is due to a weaker notion of nonlocality (without entanglement) [28]. Specifically, if we describe a projective measurement on a composite Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ by a sequence of projections $\Pi_1, \Pi_2, \ldots$, we say the measurement is *unentangled* if the support of Π_k admits a basis of *product* vectors, i.e. vectors $\psi = \psi_1 \otimes \cdots \otimes \psi_n \in \mathcal{H}$, where $\psi_j \in \mathcal{H}_j$ for

---

[7]Hrushovski and Pitowsky [35], however, showed that Gleason's theorem, when combined with the compactness theorem of first-order logic, implies the *existence* of a finite KS set. Hence, Gleason's theorem implies not only the KS theorem but something stronger, i.e., the existence of finite KS sets.

all j. Rank-one unentangled measurements are sufficient for our argument. Each such measurement can be described by a complete set of *product rays*, that is, rays consisting of product vectors.

In the case of Gleason's theorem, Wallach [27] showed that for a multiqudit system in which each subsystem has dimension at least three, Gleason's theorem can be proved using only rank-one unentangled projections. Specifically, frame functions on rank-one unentangled measurements can always be described by the Born rule, as in Eq. (4). However, if even one subsystem has dimension two, the result fails to hold, i.e. there exist frame functions on unentangled projections which cannot be described as in Eq. (4).

To address the case of the KS theorem, we will be interested in rays in $\mathbb{C}^2$, or *qubit rays*, and rays in $\mathbb{C}^{2 \otimes n}$, or *n-qubit rays*. Given a pair of orthogonal unit vectors $\{|0\rangle, |1\rangle\}$ in $\mathbb{C}^2$ a generic qubit ray can be expressed as

$$|\psi\rangle = \cos(\theta/2) |0\rangle + e^{i\phi} \sin(\theta/2) |1\rangle , \quad (5)$$

for some $0 \le \theta \le \pi$ and $0 \le \phi < 2\pi$. The parameters θ and ϕ define a point on the Bloch sphere (see Fig. 3). The pairs of orthogonal rays in $\mathbb{C}^2$ are given exactly by the pairs of antipodal points of the Bloch sphere. The product rays in $\mathbb{C}^{2 \otimes n}$ are then rays admitting an expression $|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$ where $|\psi_j\rangle$ are qubit rays for all $1 \le j \le n$.

3 Kochen–Specker theorem for multiqubit systems

In this section, we will show that any logical proof of the KS theorem on a multiqubit system requires entangled measurements, i.e.

Theorem 3. *Any multiqubit Kochen–Specker set necessarily contains entangled rays.*

We consider a KS-colouring c_n on product rays of $\mathbb{C}^{2 \otimes n}$ defined in terms of a KS-colouring c_1 on the rays of $\mathbb{C}^2$. Specifically, $c_n(|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle) = \prod_{j=1}^{n} c_1(|\psi_j\rangle)$. We choose c_1 to be particularly easy to visualise KS-colouring but the argument follows for any
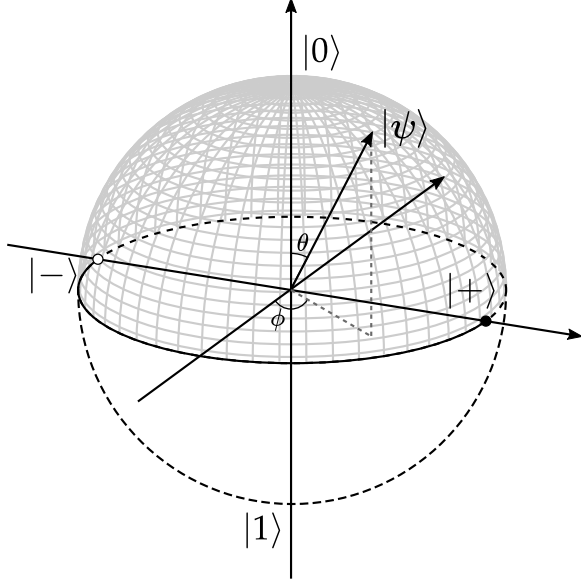
Figure 3: The Bloch sphere with the qubit north rays depicted by the hatched grey hemisphere including the solid point at $|+\rangle$ and the solid black half of the equator but not the open point at $|-\rangle$ or the dashed half of the equator.

choice. To specify our KS-colouring, we first define a *north* qubit ray:

Definition 1. A *north* qubit ray is of the form

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin\theta/2|1\rangle \qquad (6)$$

where $0 \le \theta < \pi/2$ and $0 \le \phi < 2\pi$ or $\theta = \pi/2$ and $\pi < \phi \le 2\pi$. We denote the set of north rays, depicted on the Bloch sphere in Fig. 3, by $\mathcal{N}$.

Note that, for any north qubit ray $|\psi\rangle$, the ray $|\psi^\perp\rangle$ satisfying $\langle\psi|\psi^\perp\rangle = 0$ is not a north ray, and vice versa, i.e. every pair of orthogonal qubit rays contains exactly one north ray. This definition naturally extends to the case of n-qubit rays:

Definition 2. An *all-north* n-qubit ray is a product ray $|\psi\rangle = |\psi_1\rangle \cdots |\psi_n\rangle$ such that $|\psi_j\rangle \in \mathcal{N}$ for all $j \in \{1, \dots, n\}$. We denote the set of all-north n-qubit rays by $\mathcal{N}^n$.

We need the following two lemmas in order to prove our first main result:

Lemma 1. *Any complete set of two-qubit product rays contains exactly one all-north ray.*

Proof. A generic two-qubit product basis takes the form

$$\left\{ |\psi_1\rangle|\psi_2\rangle, |\psi_1\rangle|\psi_2^\perp\rangle, |\psi_1^\perp\rangle|\psi_3\rangle, |\psi_1^\perp\rangle|\psi_3^\perp\rangle \right\}, \qquad (7)$$

for $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \in \mathbb{C}^2$, noting that the systems may be swapped. By inspection, one can see that exactly one of the rays in Eq. (7) is all-north. $\square$

The proof of the following lemma is based on that of Prop. 1 in [36].

Lemma 2. *Any complete set of n-qubit product rays contains exactly one all-north ray.*

Proof. We will prove this statement by induction on the number of qubits.

Firstly, an n-qubit complete set can contain at most one all-north ray since no two all-north rays are orthogonal.

Now, assume the result holds for m qubits and consider a complete set of $(m+1)$-qubit product rays

$$P = \{ |\Psi_1\rangle|\psi_1\rangle, \dots, |\Psi_{2^{m+1}}\rangle|\psi_{2^{m+1}}\rangle \}, \qquad (8)$$

where $|\Psi_j\rangle \in \mathbb{C}^{2^{\otimes m}}$ and $|\psi_j\rangle \in \mathbb{C}^2$ for $j \in \{1, \dots, 2^{m+1}\}$. Let B denote the set of distinct rays of the final qubit, i.e.

$$B = \Big\{ |\psi\rangle \in \mathbb{C}^2 \,\Big|\, |\Psi\rangle|\psi\rangle \in P$$
$$\text{for some } |\Psi\rangle \in \mathbb{C}^{2^{\otimes m}} \Big\}, \qquad (9)$$

noting that we may have $|\psi_j\rangle = |\psi_k\rangle$ for $j \ne k$, and let E be a maximal set of non-orthogonal rays from B, i.e. each ray in $B \backslash E$ is orthogonal to some ray in E. Denote by J the subset of $\{1, \dots, 2^{m+1}\}$ such that $|\psi_j\rangle \in E$ for all $j \in J$.

Define the map μ to count the number of rays in the $(m+1)$-qubit basis, P, for which the final qubit is associated to a given ray, i.e.

$$\mu(|\psi\rangle) = \Big| \Big\{ |\Psi\rangle|\psi\rangle \,\Big|\, |\Psi\rangle \in \mathbb{C}^{2^{\otimes m}}, |\Psi\rangle|\psi\rangle \in P \Big\} \Big|. \qquad (10)$$

We may assume (without loss of generality) that the set E is chosen such that

$$\mu(|\psi\rangle) \ge \mu(|\psi^\perp\rangle), \qquad (11)$$

7

for all $|\psi\rangle \in E$, since if $\mu(|\psi\rangle) < \mu(|\psi^\perp\rangle)$ then $|\psi\rangle$ could be replaced by $|\psi^\perp\rangle$ in E without violating the requirements on the set E.

The maximality of E implies that for every $|\psi\rangle \in B \backslash E$ we have $|\psi^\perp\rangle \in E$. Let $\mu(E) = \sum_{|\psi\rangle \in E} \mu(|\psi\rangle) = |J|$ and $\mu^\perp(E) = \sum_{|\psi\rangle \in E} \mu(|\psi^\perp\rangle)$. Note that

$$\mu(E) + \mu^\perp(E) = \sum_{|\psi\rangle \in B} \mu(|\psi\rangle) = 2^{m+1} \quad (12)$$

and $\mu(E) \geq \mu^\perp(E)$ by the assumption in Eq. (11).

If $|\psi_j\rangle, |\psi_k\rangle \in E$ for $j \neq k$ then $\langle \Psi_j | \Psi_k \rangle = 0$, since $\langle \psi_j | \psi_k \rangle \neq 0$. Hence the set

$$\{|\Psi_j\rangle \,|\, j \in J\}, \quad (13)$$

comprises $\mu(E)$ mutually orthogonal rays of $\mathbb{C}^{2 \otimes m}$. It follows that $\mu(E) \leq 2^m$ which gives

$$2\mu(E) \leq 2^{m+1} = \mu(E) + \mu^\perp(E), \quad (14)$$

and $\mu(E) \leq \mu^\perp(E)$. Therefore, we find

$$\mu(E) = \mu^\perp(E). \quad (15)$$

Furthermore, it then follows from Eq. (11) that

$$\mu(|\psi\rangle) = \mu(|\psi^\perp\rangle), \quad (16)$$

for all $|\psi\rangle \in E$ and hence for all $|\psi\rangle \in B$ (since, by the maximality of E, if $|\psi\rangle \in B$ is not contained in E then $|\psi^\perp\rangle \in E$). The definition of μ is independent of E, and we find that assumption (11) holds with equality for any choice of E. We may now choose E to consist entirely of north rays: given any maximal set E' of non-orthogonal rays from B we have by Eq. (16) that $|\psi^\perp\rangle$ is in B for any $|\psi\rangle \in E'$ and therefore all $|\psi\rangle \in E'$ that are not north rays can be exchanged with the north rays $|\psi^\perp\rangle$ to produce the all-north set E.

We also find from Eq. (15) that $\mu(E) = 2^m = |J|$ so that the set of Eq. (13) is a complete set of product rays in $\mathbb{C}^{2 \otimes m}$. By our assumption (for the proof by induction), this m-qubit complete set contains exactly one all-north ray, say $|\Psi_k\rangle$. The ray $|\Psi_k\rangle |\psi_k\rangle \in P$ is then an all-north ray since $k \in J$ implies that

$|\psi_k\rangle \in E$, which is an all-north set of rays. Since P can contain at most one all-north ray, it contains exactly one all-north ray.

By Lemma 1, since the statement of Lemma 2 holds for $n = 2$, it holds for all $n \in \mathbb{N}$ by induction. $\qquad \square$

We can now prove Theorem 3.

Theorem 3. *Any multiqubit Kochen–Specker set necessarily contains entangled rays.*

Proof. We will show the contrapositive: the contextuality scenario generated by all n-qubit product rays is KS-colourable. Define $c_n : \Sigma\left(\mathbb{C}^{2 \otimes n}\right) \to \{0, 1\}$ as follows.

$$c_n(|\psi\rangle) = \begin{cases} 1 \text{ if } |\psi\rangle \in \mathcal{N}^n, \\ 0 \text{ otherwise.} \end{cases} \quad (17)$$

By Lemma 2, every complete set of n-qubit product rays contains exactly one all-north ray, hence c_n defines a KS-colouring on the contextuality scenario generated by all n-qubit product rays. $\qquad \square$

4 A Kochen–Specker noncontextual ontological model for unentangled n-qubit systems

In the previous section we demonstrated the existence of KS-colourings of the unentangled rays of n-qubit systems. Such colourings define ontic states in a KS-noncontextual ontological model. However, the existence of such ontic states alone is insufficient to conclude that some quantum statistics can be reproduced by such a model.

In this section we will show how these ontic states can, indeed, yield a KS-noncontextual ontological model for the fragment of quantum theory consisting of n-qubit product states and n-qubit unentangled measurements. We do so via an extension of the single qubit model of Kochen and Specker [12] using the formulation of Leifer [37].

Before we proceed further, we note that if we were to restrict our analysis to measurements that can be performed locally on

8

each qubit—either only measurements in *direct* product bases (see Eq. (42) or Ref. [47]) or, more generally, those that can be implemented via LOCC—construction of a KS-noncontextual model would not require the results from the previous section. However, we are considering the strictly larger class of all unentangled measurements (see Sec. 2) including those that cannot be implemented via LOCC, for example, a measurement in the basis of Eq. (1). Thus, our KS-noncontextual ontological model requires KS-colourings of *all* unentangled multiqubit measurement bases and not only those implementable via LOCC. With these clarifications out of the way, we can now proceed to detail our construction.

Given a unit vector $|\psi\rangle$ in Hilbert space we will denote by $[\psi]$ the rank-one projection onto the subspace spanned by $|\psi\rangle$. The ontic state space for the model is given by $\Lambda = S_1^2 \times \cdots \times S_n^2$, where S_j^2 is a two-dimensional sphere representing the ontic state space of the jth qubit ($j = 1, 2, \ldots, n$) and $\times$ is the Cartesian product. We represent each n-qubit ontic state $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \Lambda$ by a single vector $\vec{\lambda}$ in a real vector space $\mathbb{R}^{3n}$:

$$
\vec{\lambda} = \vec{\lambda}_1 + \vec{\lambda}_2 + \cdots
$$
$$
= \begin{pmatrix} \sin\theta_1 \cos\varphi_1 \\ \sin\theta_1 \sin\varphi_1 \\ \cos\theta_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sin\theta_2 \cos\varphi_2 \\ \sin\theta_2 \sin\varphi_2 \\ \cos\theta_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots
$$
(18)

where $0 \leq \theta_j \leq \pi$ and $-\pi < \varphi_j \leq \pi$ are the coordinates of λ in the jth copy of S^2. Similarly, we may describe an unentangled rank-one projection $[\psi] = [\psi_1 \otimes \cdots \otimes \psi_n]$ by a vector

$$
\vec{\psi} = \vec{\psi}_1 + \vec{\psi}_2 + \cdots
$$
$$
= \begin{pmatrix} \sin\theta_1^\psi \cos\varphi_1^\psi \\ \sin\theta_1^\psi \sin\varphi_1^\psi \\ \cos\theta_1^\psi \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sin\theta_2^\psi \cos\varphi_2^\psi \\ \sin\theta_2^\psi \sin\varphi_2^\psi \\ \cos\theta_2^\psi \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots
$$
(19)

where $0 \leq \theta_j^\psi \leq \pi$ and $-\pi < \varphi_j^\psi \leq \pi$ are the spherical coordinates of $\vec{\psi}_j$ on the jth sphere S^2.

We will now define a deterministic response function whereby, essentially, an outcome $[\psi]$ occurs with probability one in a state λ exactly when $\vec{\psi}_j$ is in the hemisphere centred around $\vec{\lambda}_j$ for all $j \in \{1, 2, \ldots, n\}$. First, we make the following definitions. Let $|\lambda_j\rangle$ denote the Hilbert space vector with spherical coordinates (θ_j, φ_j) on the jth qubit, i.e. $|\lambda_j\rangle = \cos(\theta_j/2)|0\rangle + e^{i\varphi_j}\sin(\theta_j/2)|1\rangle$. Then we define a unitary that maps $|\lambda_j\rangle$ to the north pole $|0\rangle$, i.e. $U_{\lambda_j} = |0\rangle\langle\lambda_j| + |1\rangle\langle\lambda_j^\perp|$.

Now, explicitly, given an unentangled projection $[\psi] = [\psi_1 \otimes \cdots \otimes \psi_n]$, the probability of observing an outcome associated with $[\psi]$ of a measurement M on a system in ontic state λ is given by

$$
\Pr\left([\psi]|M, \lambda\right) = \begin{cases} 1 & \text{if } \bigotimes_{j=1}^n U_{\lambda_j}|\psi\rangle \in \mathcal{N}^n, \\ 0 & \text{otherwise} . \end{cases}
$$
(20)

Lemma 2 shows that Eq. (20) defines a valid response function for n-qubit product projections.

The following inequalities will be useful in proving that our ontological model reproduces quantum theory:

$$
\prod_{j=1}^n H^0(\vec{\psi}_j \cdot \vec{\lambda}) \leq \Pr\left([\psi]|M, \lambda\right) \leq \prod_{j=1}^n H^1(\vec{\psi}_j \cdot \vec{\lambda})
$$
(21)

for all unentangled n-qubit projections $[\psi]$ and $\lambda \in \Lambda$, where $H^{0,1} : \mathbb{R} \to \{0,1\}$ are two conventions for the Heaviside step function,

$$H^y(x) = \begin{cases} 1 & \text{if } x > 0 \\ y & \text{if } x = 0 \\ 0 & \text{if } x < 0 \,. \end{cases} \quad (22)$$

We now define the epistemic states of the model. Given that a product state $[\chi] = [\chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_n]$ (which can be described by a vector as in Eq. (19) with parameters $0 \leq \theta_j^\chi \leq \pi$ and $-\pi < \varphi_j^\chi \leq \pi$) is prepared, the probability measure over the ontic states of the system is given by

$$\mu_\chi(\Omega) = \int_\Omega \prod_{j=1}^n p_{\vec{\chi}_j}(\vec{\lambda}) \sin \theta_j d\theta_j d\varphi_j, \quad (23)$$

for $\Omega \subseteq \Lambda$, where

$$p_{\vec{\chi}_j}(\vec{\lambda}) = \frac{1}{\pi} H^0(\vec{\chi}_j \cdot \vec{\lambda}) \vec{\chi}_j \cdot \vec{\lambda}. \quad (24)$$

Thus the probability of observing outcome $[\psi]$ of a measurement M performed on a system prepared in quantum state $[\chi]$ is given by

$$\Pr([\psi]|M, [\chi])$$
$$= \int_\Lambda \Pr([\psi]|M, \lambda) \, d\mu_\chi(\lambda)$$
$$= \int_\Lambda \Pr([\psi]|M, \lambda) \prod_{j=1}^n p_{\vec{\chi}_j}(\vec{\lambda}) \sin \theta_j d\theta_j d\varphi_j$$
$$(25)$$

Since the terms of the integrand are all non-negative, it follows from Eq. (21) that

$$\int_\Lambda \prod_{j=1}^n H^0(\vec{\psi}_j \cdot \vec{\lambda}) \prod_{j=1}^n p_{\vec{\chi}_j}(\vec{\lambda}) \sin \theta_j d\theta_j d\varphi_j$$
$$(26)$$
$$\leq \Pr([\psi]|M, [\chi]) \quad (27)$$
$$\leq \int_\Lambda \prod_{j=1}^n H^1(\vec{\psi}_j \cdot \vec{\lambda}) \prod_{j=1}^n p_{\vec{\chi}_j}(\vec{\lambda}) \sin \theta_j d\theta_j d\varphi_j$$
$$(28)$$

For both of these bounds ($y \in \{0,1\}$) we find

$$\int_\Lambda \prod_{j=1}^n H^y(\vec{\psi}_j \cdot \vec{\lambda}) \prod_{j=1}^n p_{\vec{\chi}_j}(\vec{\lambda}) \sin \theta_j d\theta_j d\varphi_j$$
$$(29)$$
$$= \int_\Lambda \prod_{j=1}^n \frac{1}{\pi} H^y(\vec{\psi}_j \cdot \vec{\lambda}) H^0(\vec{\chi}_j \cdot \vec{\lambda}) \vec{\chi}_j \cdot \vec{\lambda}$$
$$\sin \theta_j d\theta_j d\varphi_j \quad (30)$$
$$= \prod_{j=1}^n \int_0^\pi \int_{-\pi}^\pi \frac{1}{\pi} H^y(\vec{\psi}_j \cdot \vec{\lambda}) H^0(\vec{\chi}_j \cdot \vec{\lambda}) \vec{\chi}_j \cdot \vec{\lambda}$$
$$\sin \theta_j d\theta_j d\varphi_j. \quad (31)$$

We can now evaluate each term in this product for both cases $y = 0$ and 1, following the method of Ref. [37] (where the $y = 0$ case is considered), to find

$$\int_0^\pi \int_{-\pi}^\pi \frac{1}{\pi} H^y(\vec{\psi}_j \cdot \vec{\lambda}) H^0(\vec{\chi}_j \cdot \vec{\lambda}) \vec{\chi}_j \cdot \vec{\lambda}$$
$$\sin \theta_j d\theta_j d\varphi_j$$
$$= |\langle \psi_j, \chi_j \rangle|^2, \quad (32)$$

for both $y = 0, 1$. The argument goes as follows. Firstly, we choose our coordinates θ_j and φ_j such that $\vec{\chi}_j = (1, 0, 0)$ and $\vec{\psi}_j = (\cos \phi, \sin \phi, 0)$ for some $-\pi < \phi \leq \pi$. Then since $\vec{\lambda}_j = (\sin \theta_j \cos \varphi_j, \sin \theta_j \sin \varphi_j, \cos \theta_j)$, we find

$$\vec{\chi}_j \cdot \vec{\lambda}_j = \sin \theta_j \cos \varphi_j, \quad (33)$$
$$\vec{\psi}_j \cdot \vec{\lambda}_j = \sin \theta_j \cos(\varphi_j - \phi). \quad (34)$$

The integrand of Eq. (32) is non-zero when $\vec{\chi}_j \cdot \vec{\lambda}_j$ is positive and when $\vec{\psi}_j \cdot \vec{\lambda}_j$ is positive (non-negative) for the case $y = 0$ ($y = 1$). These conditions are achieved for $y = 0$ when $-\pi/2 < \varphi_j < \pi/2$ and $-\pi/2 + \phi < \varphi_j < \pi/2 + \phi$ and similarly for $y = 1$ but when the latter inequalities are no longer strict, i.e. $-\pi/2 + \phi \leq \varphi_j \leq \pi/2 + \phi$. Since the integrals over a closed or open interval are equal, in both

cases $y = 0$, or 1, if ϕ is non-negative we find

$$\int_0^\pi \int_{-\pi}^\pi \frac{1}{\pi} H^y(\vec{\psi_j} \cdot \vec{\lambda}) H^0(\vec{\chi_j} \cdot \vec{\lambda})$$
$$\vec{\chi_j} \cdot \vec{\lambda} \sin \theta_j d\theta_j d\varphi_j$$
$$= \frac{1}{\pi} \int_0^\pi \sin^2 \theta_j d\theta_j \int_{-\frac{\pi}{2}+\phi}^{\frac{\pi}{2}} \cos \varphi_j d\varphi_j \quad (35)$$
$$= \frac{1}{2} \left(1 + \cos \phi\right) = |\langle \psi_j, \chi_j \rangle|^2.$$

We find the same value if ϕ is negative. Finally, we have shown

$$\Pr\left([\psi]|M, [\chi]\right) = \prod_{j=1}^n |\langle \psi_j, \chi_j \rangle|^2 \quad (36)$$
$$= |\langle \psi, \chi \rangle|^2.$$

The model is easily extended to higher rank projections via $\Pr\left(\sum_j [\psi^j]|M, \lambda\right) = \sum_j \Pr\left([\psi^j]|M, \lambda\right)$ giving $\Pr\left(\sum_j [\psi^j]|M, [\chi]\right) = \sum_j \Pr\left([\psi^j]|M, [\chi]\right)$ for any mutually orthogonal projections $[\psi^j]$. The fact that this extension is well-defined follows from Lemma 2. Thus the model reproduces the predictions of quantum theory for n-qubit product states and unentangled measurements.

Note that for the case $n = 1$, this ontological model is equivalent to the Kochen–Specker model for projective measurements and pure states of a single qubit [12, 37]. Our generalization, however, is nontrivial because it crucially relies on Lemma 2 to define the n-qubit response function of Eq. (20). For a single qubit, on the other hand, Lemma 2 is trivial because every qubit state appears in exactly one basis which already implies that each basis contains exactly one north state.

The model can be extended to include mixtures of product states, e.g. $\sum_{i=1}^N q_i [\chi^i]$, for product states $|\chi^i\rangle$ by simply defining the probability measure for such a mixture as the same mixture of probability measures, i.e. $\sum_{i=1}^N q_i \mu_{\chi^i}$. However, note that such an extension is *preparation contextual*, in the sense that there is no fixed probability measure for a given separable density operator since different decompositions of the same density operator as a mixture of product

states result in different probability measures. Indeed, there exists no preparation noncontextual extension to mixed states, as follows from the impossibility of a preparation noncontextual ontological model for single qubit mixed states [21].

An immediate consequence of the existence of this KS-noncontextual ontological model is that it allows us to obtain a tight statement on the relationship between entanglement and KS-contextuality in multiqubit systems. Namely, entanglement is necessary for any proof of the KS theorem, whether logical [12, 18] or statistical [20, 19]. Theorem 3 showed that multiqubit entangled measurements are necessary for logical proofs of the KS theorem. Our construction of the KS-noncontextual ontological model implies that even for statistical proofs of the KS theorem on multiqubit systems, one requires entanglement, either in the state or in the measurements. A well-known instance of such a statistical proof is the violation of Bell inequalities by implementing local (hence, unentangled) measurements on an entangled multiqubit state. One can also construct statistical proofs of the KS theorem (where there exist KS-colourings of the measurements involved) in which the quantum state is a product state and there is entanglement in the measurements. For example, given a multiqubit entangled state that violates a Bell inequality, one can apply a global unitary to the state making it a product state and then apply the same unitary to all the measurements involved in the violation. This transformation preserves the probabilities, the orthogonality relations, and hence, the statistical proof, whilst inevitably making some of the measurements entangled.

5 Bell and Kochen–Specker coincide for multiqubit systems with unentangled measurements

In the previous section we found a KS-noncontextual ontological model for unentangled projective measurements and prod-

uct pure states of multiqubit systems. The model straightforwardly extends to separable states, allbeit in a preparation contextual way. It follows that to obtain a multiqubit statistical proof of the Kochen–Specker theorem from unentangled measurements one must employ an entangled multiqubit state.[8] Is entanglement of the multiqubit state, however, sufficient to yield such a statistical proof? In this section, we find the answer to be negative, i.e. there must exist an KS-noncontextual ontological model that includes—besides separable states and unentangled projective measurements—certain entangled states. Specifically, we find that the multiqubit states that can demonstrate KS-contextuality with unentangled measurements (in a finite contextuality scenario) are exactly those that can violate a Bell inequality with local projective measurements.

The fact that there exist entangled multiqubit states yielding a statistical proof of the KS theorem with only unentangled measurements follows from the existence of entangled multiqubit states that violate Bell inequalities with only local projective measurements. This implication can be seen from both the more traditional *observable-based* perspective [39], as well as the *event-based* perspective [32, 40].

In the observable-based perspective we consider the local observables of each party in a Bell experiment. The set of local observables corresponding to one choice of setting for each party commute pairwise and, thus, form a context. We may then consider the KS-noncontextual ontological models respecting all such contexts. The mathematical constraints defining these models in this situation are equivalent to the constraints imposed

by Bell's assumption of local causality in the original Bell scenario.[9] Hence, a quantum violation of a Bell inequality using projective measurements also provides a set of quantum statistics deriving from local, and hence unentangled, measurements that are incompatible with a KS-noncontextual ontological model.

To prove our result the event-based perspective will be more convenient. In order the see the connection between Bell scenarios and KS-contextuality in this perspective one needs to go beyond the idea of KS-colourings of contextuality scenarios to that of *probabilistic models* on contextuality scenarios following the framework of Ref. [32].

In the following subsection, we define probabilistic models on contextuality scenarios and related notions, then show how a proof of Bell's theorem yields a statistical proof of the KS theorem (a known connection). In the next subsection, we then show that the converse relationship also holds for multiqubit systems and unentangled measurements. Thus, we arrive at the second main contribution of this work: a multiqubit entangled state can yield a statistical proof of the KS theorem with unentangled measurements if and only if it violates a Bell inequality with local projective measurements.

5.1 Bell implies KS

A probabilistic model is a probability assignment to the vertices of a contextuality scenario (i.e. a hypergraph) such that the probabilities assigned in each hyperedge sum to one. Explicitly, given a contextuality scenario H with vertices $V(H)$ and hyperedges $E(H)$, a probabilistic model on the contextuality scenario is a map $p : V(H) \to [0, 1]$ such that $\sum_{v \in e} p(v) = 1$ for all $e \in E(H)$. A KS-colouring is a probabilistic model that only takes values 0 and 1. A *classical model* is a probabilistic model that can be decomposed into a convex combination of KS-colourings, where a convex combination, q, of probabilis-

---

[8]Note the contrast between preparation contextuality and KS-contextuality for multiqubit systems: the former requires no entanglement but for the latter, entanglement is necessary. It is known that any preparation noncontextual ontological model of a fragment of quantum theory is also necessarily KS-noncontextual but not conversely [38, 23]. We can thus conclude that in any ontological model of multiqubit systems, KS-contextuality implies both, entanglement and preparation contextuality.

[9]Both sets of constraints can be viewed as instances of the marginal problem for classical probability distributions via Fine's theorem [41, 42, 44, 43, 39].

tic models p and p' is the probabilistic model $q(v) = \omega p(v) + (1 - \omega)p'(v)$ for some $\omega \in [0, 1]$ and all $v \in V(H)$.

A probabilistic model p is quantum if and only if for some separable Hilbert space $\mathcal{H}$ there exists (i) a projection Π_v for every $v \in V(H)$ such that $\sum_{v \in e} \Pi_v = I_{\mathcal{H}}$ for all $e \in E(H)$, and (ii) a density operator ρ on $\mathcal{H}$, such that $p(v) = \text{Tr}(\Pi_v \rho)$.

A quantum model that is not classical is said to provide a *statistical proof* of the KS theorem [20, 19]. Such nonclassicality can be witnessed by the violation of a Bell-KS inequality that bounds the polytope of classical models, e.g. the KCBS inequality [20]. On the other hand, if a contextuality scenario admits no classical models (hence no KS-colourings) but it admits a quantum model, we have a *logical proof* of the KS theorem [12, 18] which demonstrates a stronger form of KS-contextuality: not only does there exist a quantum model outside the set of classical models, the set of classical models is, in fact, empty, i.e. no KS inequalities exist and *every* probabilistic model (hence every quantum model) on the contextuality scenario fails to be classical.

A Bell scenario is an experiment in which n parties each perform a measurement x_r and observe an output a_r for $1 \le r \le n$ on some system such that the parties' individual experiments are space-like separated. We call a measurement on the entire system performed by all n parties a global measurement. Each global measurement is specified by a list of the local measurement settings $\mathbf{x} = x_1 x_2 \ldots x_n$ and has a set of possible outcomes $\mathbf{a}|\mathbf{x}$, where $\mathbf{a} = a_1 a_2 \ldots a_n$. We consider the case in which each local measurement has two possible outcomes, i.e. $a_r \in \{0, 1\}$ for all $1 \le r \le n$. This is the natural setting for a multiqubit Bell experiment where the parties implement local projective measurements.

A *behaviour* in a Bell scenario is a probability distribution, $p(\mathbf{a}|\mathbf{x})$, on the measurement outcomes. A local behaviour is a behaviour that can be decomposed into a convex combination of local *deterministic* behaviours, i.e. of behaviours $p(\mathbf{a}|\mathbf{x}) =$

$p_1(a_1|x_1) \cdots p_n(a_n|x_n)$, where $p_r(a_r|x_r) \in \{0, 1\}$ denotes the probability of party r observing a_r given they performed measurement x_r for all $1 \le r \le n$.

A behaviour that can be realised by each party performing a projective measurement on some subsystem of a quantum system is called a projective quantum behaviour. For a projective quantum behaviour, $p(\mathbf{a}|\mathbf{x})$, there exists a separable Hilbert space, $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, a projective measurement $\{\Pi_{a_r}^{r,x_r}\}_{a_r}$ on $\mathcal{H}_r$ for each measurement setting x_r of the r-th party for all $1 \le r \le n$, and a density operator ρ on $\mathcal{H}$ such that $p(\mathbf{a}|\mathbf{x}) = \text{Tr}(\bigotimes_{r=1}^n \Pi_{a_r}^{r,x_r} \rho)$.

A Bell scenario can be mapped to a contextuality scenario which has (i) a vertex for every possible global measurement outcome $\mathbf{a}|\mathbf{x}$ for all $\mathbf{a}$ and $\mathbf{x}$, (ii) a hyperedge consisting of all the possible outcomes $\mathbf{a}|\mathbf{x}$ of fixed global measurement $\mathbf{x}$, i.e. each set $\{(\mathbf{a}|\mathbf{x})|a_r \in \{0, 1\}$ for all $1 \le r \le n\}$ is a hyperedge, *and* (iii) a hyperedge deriving from each no-signalling condition. The non-signalling hyperedges consist of sets of possible outcomes of (hypothetical) *adaptive* measurements in which one party performs a measurement first and depending on their outcome a second party selects a measurement setting and so on [32]. For example, the hyperedge $\{00|00, 01|00, 10|01, 11|01\}$ features in the hypergraph of the two party Bell scenario with binary inputs and outputs. This hyperedge can be thought of as the outcomes of an adaptive measurement protocol in which the first party always chooses measurement setting 0, and the second party chooses the measurement setting 0 if the first party obtains outcome 0 and 1 if they obtain 1.[10]

There is a bijection between local behaviours in the Bell scenario and classical models in the corresponding contextuality

---

[10]Note that the global measurements from the Bell scenario are also examples of adaptive measurements in which each party's measurement setting does not actually change based on the outcome of any other party. Thus all the hyperedges in a contextuality scenario representing a Bell experiment can be thought of as adaptive measurements.

scenario. Furthermore, each quantum behaviour deriving from projective measurements in the Bell scenario defines a quantum model in the contextuality scenario. A collection of multiqubit projective measurements and a density operator that violate a Bell inequality then generate a quantum model on the corresponding contextuality scenario that is not a classical model and thus yields a statistical proof of the KS theorem. For a detailed account, see Ref. [32].

5.2 KS implies Bell

We have seen that states that can violate Bell inequalities also yield statistical proofs of the KS theorem with unentangled measurements. We now show that the converse relationship also holds. Namely, for any collection of product multiqubit rays and an entangled state which yield a statistical proof of the KS theorem, the entangled state necessarily violates a Bell inequality.

The proof will proceed along the following lines. Consider a hypergraph H with a non-classical quantum model given by some multiqubit product rays, $\mathcal{S}$, and a density operator ρ. We will expand the set of multiqubit rays $\mathcal{S}$ to a set of rays $\mathcal{S}'$ that correspond to all the outcomes of a quantum measurement strategy in some Bell scenario, $\mathcal{B}(\mathcal{S})$. For example, if $\mathcal{S}$ consisted of $\{|00\rangle, |+1\rangle, |0+\rangle\}$ it would be expanded to

$$\begin{aligned} \mathcal{S}' = \{ &|00\rangle, |01\rangle, |10\rangle, |11\rangle, \\ &|+0\rangle, |+1\rangle, |-0\rangle, |-1\rangle, \\ &|0+\rangle, |0-\rangle, |1+\rangle, |1-\rangle, \\ &|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\}, \end{aligned} \quad (37)$$

which contains all the rays corresponding to the outcomes of measurements in a two-party Bell scenario where each party has two possible measurement settings given by $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$. We can then extend the hypergraph, H, to the hypergraph $G \supseteq H$ generated by this extended set of rays and their orthogonality relations. We can also extend the quantum model on H given by $\mathcal{S}$ and ρ to a quantum model on G given by $\mathcal{S}'$ and ρ; this quantum model on G continues to be

non-classical. Let H' be the hypergraph corresponding to the Bell scenario $\mathcal{B}(\mathcal{S})$. The hypergraph G will contain all the vertices and hyperedges of the hypergraph H' but possibly with additional hyperedges deriving from non-local bases in $\mathcal{S}'$ such as the basis in Eq. (1). The set $\mathcal{S}'$ and state ρ also give a quantum model on the Bell hypergraph H', since every hyperedge of H' is contained in G. Finally, we will show, following [45], that the sets of classical models on G and H' are identical despite the difference in hyperedges. Thus, the non-classical model on G is also non-classical on H', meaning ρ yields a non-local, projective quantum behaviour in $\mathcal{B}(\mathcal{S})$. This leads us to the following theorem:

Theorem 4. *Any multiqubit density operator, ρ, that can yield a statistical proof of the KS theorem with a finite set of unentangled projective measurements can violate a Bell inequality with local projective measurements.*

A full proof of Theorem 4 is given in Appendix B.

It follows that it must be possible to extend our KS-noncontextual ontological model to include, besides separable states, multiqubit entangled states (e.g. Werner states) that cannot yield Bell violations when subjected to local projective measurements. We leave the construction of such a multiqubit KS-noncontextual ontological model as an open problem for future work.

6 Do we need fully entangled bases?

Our initial question concerning the necessity of entanglement in any multiqubit logical proof of the KS theorem [12] was motivated by the presence of entanglement in the Peres-Mermin magic square (Fig. 1). Another curious feature of the Peres-Mermin argument is the appearance of bases that are not merely entangled but, in fact, *fully entangled*, i.e. these bases contain no product projections. Is there, then, an even stronger requirement on the entanglement needed for a multiqubit KS theorem, i.e. just as the presence of entangled projections in a KS set is generic

and not particular to the Peres-Mermin case, must a multiqubit KS set always contain fully entangled bases? Here we rule out this possibility by explicitly constructing a two-qubit KS set that does not require fully entangled bases, i.e. every basis contains at least one product ray.

Our construction makes use of Peres' 33-ray proof of the KS theorem in $\mathbb{C}^3$ which, when represented as a contextuality scenario, actually requires 57 rays (since the proof makes use of the orthogonality of certain pairs of the 33 rays for which the final orthogonal ray is missing). We construct a KS set by embedding two copies of the 57 Peres rays in two different three dimensional subspaces of $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Denote by $|\phi_j\rangle$, for integers $1 \leq j \leq 57$, the rays in the Peres 57-ray contextuality scenario. The rays can be carved up into 40 distinct orthonormal bases.

We will now embed Peres' three-dimensional rays in the three-dimensional subspace of the two-qubit Hilbert space orthogonal to the ray $|00\rangle$. Given a ray $|v\rangle = \alpha |0\rangle + \beta |1\rangle + \gamma |2\rangle \in \mathcal{R}(\mathbb{C}^3)$, denote $|v^{00}\rangle = \alpha |01\rangle + \beta |10\rangle + \gamma |11\rangle \in \mathcal{R}(\mathbb{C}^2 \otimes \mathbb{C}^2)$. For each of the 40 bases $\{|\phi_a\rangle, |\phi_b\rangle, |\phi_c\rangle\}$ among the Peres rays, we add the basis

$$\left\{ |00\rangle, |\phi_a^{00}\rangle, |\phi_b^{00}\rangle, |\phi_c^{00}\rangle \right\}, \qquad (38)$$

to our hypergraph.

We now perform the analogous operation in the subspace orthogonal to $|01\rangle$, but we first transform Peres' rays by a unitary matrix

$$U = \frac{1}{3} \begin{pmatrix} 1 + \sqrt{2} & \tau_- & \tau_+ \\ \tau_+ & 1 + \sqrt{2} & \tau_- \\ \tau_- & \tau_+ & 1 + \sqrt{2} \end{pmatrix}, \quad (39)$$

where $\tau_\pm = \left(2 - \sqrt{2} \pm \sqrt{6}\right)/2$, and denote the new set of rays $|U\phi_j\rangle$. Note that the orthogonality relations between the rays are preserved under this transformation. Explicitly, we add the basis

$$\left\{ |01\rangle, |U\phi_a^{01}\rangle, |U\phi_b^{01}\rangle, |U\phi_c^{01}\rangle \right\}, \qquad (40)$$

to our hypergraph for every basis $\{|\phi_a\rangle, |\phi_b\rangle, |\phi_c\rangle\}$, where analogously

$|v^{01}\rangle = \alpha |00\rangle + \beta |10\rangle + \gamma |11\rangle \in \mathcal{R}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ for any $|v\rangle = \alpha |0\rangle + \beta |1\rangle + \gamma |2\rangle \in \mathcal{R}(\mathbb{C}^3)$.

The unitary U has been chosen such that $\langle \phi_j^{00} | U\phi_k^{01}\rangle \neq 0$ for all entangled $|\phi_j^{00}\rangle$ and $|U\phi_k^{01}\rangle$. It follows that the entangled rays of the hypergraph cannot be formed into a basis, i.e. there are no fully entangled bases in our construction. This fact can be verified by consulting the Mathematica notebook available at Ref. [46].

The contextuality scenario can now be seen to be KS-uncolourable as follows. Any KS-colouring, c, of our hypergraph should assign zero to at least one of the rays $|00\rangle$ and $|01\rangle$, since they appear in a hyperedge—the hyperedge consisting of $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ which is one of the bases given by Eq. (38). If $|00\rangle$ is assigned zero by the colouring, c, then the assignments to the rays $|\phi_j^{00}\rangle$ would constitute a KS-colouring of the Peres hypergraph, which does not exist. Similarly, if $|01\rangle$ is assigned zero in the colouring, c, then the assignments to the rays $|U\phi_j^{01}\rangle$ would also constitute a KS-colouring of the Peres hypergraph. Hence such a colouring, c, cannot exist.

7 Unentangled Kochen–Specker sets

Between the case of multiqubit systems (wherein each subsystem has dimension two) and the case of multiqudit systems, wherein each subsystem has dimension at least three, we have the possibility of multiqudit systems consisting of both qubits and higher-dimensional qudits. For these systems, Wallach showed that unentangled projections are still insufficient to yield Gleason's theorem [27]. Is it, however, possible to obtain the KS theorem with unentangled projections in this case, unlike the multiqubit case? In this section, we show that this is indeed possible. In fact, the presence of just one qutrit in an otherwise multiqubit system is enough to allow constructions of KS sets with unentangled projections. Our argument, not surprisingly, relies on the fact that a single qutrit admits

KS sets on its own [12].

Theorem 5. *There exists a KS set consisting entirely of product rays in any separable Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ where $\dim(\mathcal{H}_j) \geq 3$ for some $1 \leq j \leq n$.*

Here we give an idea of the proof with an example and give the general proof in Appendix C. We will show that if there were no KS set consisting of product rays in $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, i.e. if there existed a KS-colouring c on the product rays in $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, then we would be able to use c to define a KS-colouring c' on any set of bases of $\mathcal{H}_j$, contradicting the KS theorem. For example, assume c is a KS-colouring of the product rays in $\mathbb{C}^2 \otimes \mathbb{C}^3$. Now take any set of bases $\{B_1, \ldots, B_N\}$ in $\mathbb{C}^3$, for example the bases from the Peres 33-vector proof (57 vectors in 40 bases), and take the element-wise tensor product of each basis B_k with the basis $\{|0\rangle, |1\rangle\}$ of $\mathbb{C}^2$. The resulting set of bases, $\{\{|0\rangle, |1\rangle\} \otimes B_k\}_k$, of $\mathbb{C}^2 \otimes \mathbb{C}^3$ would be KS-colourable by the definition of c. Now given a basis B_k we define a map c' on B_k by

$$c'(|\psi\rangle) = \begin{cases} 1 & \text{if } c(|0\rangle |\psi\rangle) = 1 \text{ or } c(|1\rangle |\psi\rangle) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

(41)

For example, given the basis $\{|0\rangle, |1\rangle, |2\rangle\}$, if $c(|02\rangle) = 1$ then $c'(|2\rangle) = 1$ and $c'(|0\rangle) = c'(|1\rangle) = 0$. By this definition and the fact that c assigns one to exactly one element of $\{|0\rangle, |1\rangle\} \otimes B_k$ we find that, likewise, c' assigns one to exactly one element of B_k. Then we see that c' is well-defined across all the elements of the bases B_k since it is defined independently from the basis B_k. The map c' would therefore be a KS-colouring of the bases $\{B_1, \ldots, B_N\}$, contradicting the result of Peres.

Note that the proof of Theorem 5 in Appendix C actually shows the stronger result that the hypergraph of product rays of $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ with only the hyperedges deriving from *direct* product bases is also KS-uncolourable. A direct product basis [47] is a basis formed by taking the elementwise tensor product of a basis for each subsystem $\mathcal{H}_j$,

i.e. a basis

$$\{|\psi_1^1\rangle |\psi_1^2\rangle \cdots |\psi_1^n\rangle, |\psi_2^1\rangle |\psi_1^2\rangle \cdots |\psi_1^n\rangle,$$
$$\ldots, |\psi_{N_1}^1\rangle |\psi_{N_2}^2\rangle \cdots |\psi_{N_2}^n\rangle\} \quad (42)$$

where $\{|\psi_1^j\rangle, \ldots, |\psi_{N_j}^j\rangle\}$ is a basis for each $\mathcal{H}_j$ for $1 \leq j \leq n$.

The contrast between the *impossibility* of Gleason's theorem and the *possibility* of KS theorem with unentangled measurements in the case of composite systems that contain both qubits and higher-dimensional qudits may seem surprising at first glance. To gain some intuition for this contrast, consider two important facts: first, very simply, Gleason's theorem implies the KS theorem, but not conversely. Second, in more depth, the existence of a KS theorem but the lack of a corresponding Gleason's theorem reflects the fact that the former is a *no-go theorem* (ruling out certain types of probabilistic assignments) while the latter is a *"go theorem"* (specifying the allowed probabilistic assignments).

Consider, for example, a qubit-qutrit system. A single qubit does not permit a proof of Gleason's theorem because the structure of projective measurements on a qubit allows for many more probabilistic assignments than are dictated by the Born rule. Composing the qubit with a qutrit and considering unentangled measurements on the pair does not rule out these non-quantum assignments on the qubit, so Gleason's theorem continues to fail. On the other hand, a single qutrit admits a proof of the KS theorem and this no-go statement goes through even if we compose it with a qubit.

8 Implications for the role of contextuality in quantum computation

The role of contextuality within the paradigm of quantum computation with state injection (QCSI) has been a subject of active research in recent years [5, 13].[11] This approach to

---

[11] Although we focus on KS-contextuality in this paper, recent results [48] have shown connections be-

quantum computation relies on lifting stabiliser quantum circuits, which cannot implement universal quantum computation, to universality via the injection of non-stabiliser states called *magic states*. It has been shown that, in the case of odd-prime dimensional qudit (or, *quopit* [13]) circuits, contextuality is a necessary resource for quantum computation, i.e. the magic states needed for universality must exhibit contextuality with respect to stabiliser measurements [5]. Thus, if a state admits a KS-noncontextual ontological model with the stabiliser measurements, it cannot promote the circuit to universality. The converse claim, that (KS-)contextuality is sufficient for universal quantum computation, is conjectured but unproven [5].

The multiqubit case (i.e. the even-prime dimensional case) has, however, been a hurdle in interpreting contextuality of magic states as a resource for quantum computation [5]. The multiqubit stabiliser subtheory is classically efficiently simulable [49, 50] despite the presence of (state-independent) KS-contextuality in its observables (e.g. see Fig. 1). The fact that *any* state, stabiliser or not, exhibits contextuality with respect to such observables means that there is nothing special about the contextuality of magic states that renders contextuality a necessary resource for universal quantum computation. While magic states are still a necessary resource, their contextuality does not single them out (unlike the quopit case) since *all* states can exhibit contextuality.[12] To overcome this hurdle, *restricted* QCSI schemes have been proposed [13] which restore the status of contextuality as a resource for quantum computation.

In Ref. [13], such a restricted QCSI scheme $\mathcal{M}_\mathcal{O}$ is required to satisfy:

(C1) *Resource character.* There exists a quantum state that does not exhibit contextu-

tween QCSI and generalised contextuality [21], analogous to the conclusions of Ref. [5] with respect to KS-contextuality.

[12]This also means that, unlike the quopit case, contextuality cannot even be *conjectured* as a sufficient condition for universality in this case.

ality with respect to measurements available in $\mathcal{M}_\mathcal{O}$.

(C2) *Tomographic completeness.* For any state ρ, the expectation value of any Pauli observable can be inferred via the allowed operations of the scheme.

The requirement (C1) means that the measurements in the scheme cannot exhibit state-independent contextuality. It follows from Theorem 3 of the present manuscript that a sufficient condition for satisfying requirement (C1) is that every measurement in the scheme be unentangled. Specifically, the scheme $\mathcal{M}_\mathcal{O}$ prescribes the sets of observables in the scheme that are jointly measurable. If none of these joint measurements requires a measurement in an entangled basis then (C1) is satisfied.

The schemes proposed in Ref. [13] satisfy exactly this condition; they contain no entangled measurements. Thus, all the contextuality in these schemes derives from entanglement of the injected state. Furthermore, it follows from Theorem 4 that in order for the injected state to promote such a scheme to universality it must be capable of violating a Bell inequality with some local projective measurements.

9 Conclusions

In this work we have demonstrated the necessity of entanglement in proofs of the KS theorem for systems of multiple qubits, i.e., KS-contextuality necessitates not only incompatibility but also entanglement in multiqubit systems.

On the one hand, we showed KS sets for multiqubit systems must contain entangled rays/projections. It follows that any *logical* proof of the KS theorem for multiqubit systems relies upon entanglement in the measurements, just like in the case of the Peres-Mermin square (Fig. 1). However, unlike the Peres-Mermin square, multiqubit proofs of KS-contextuality do not, in general, require measurements in *fully entangled* bases.

On the other hand, the KS-noncontextual ontological model defined in Sec. 4 allows us to go further and also make statements about *statistical* (and state-dependent) proofs of the KS theorem, i.e. those proofs that are in the style of Klyachko *et al.* [20]. Specifically, this model can be easily extended to include unentangled projective measurements and separable states of a multiqubit system. Thus, any statistical proof of the KS theorem for such a system must employ entanglement, either in the state or in the measurements (or both). For example, proofs of Bell's theorem give rise to statistical proofs of the KS theorem in which the measurements are unentangled but the state is necessarily entangled [51, 33].

Our results also allow us to make the following comparisons with other forms of nonclassicality.

Firstly, it follows from our results that the nonclassicality present in unentangled measurements in the form of 'nonlocality without entanglement' [28] is insufficient to witness the (KS-)contextuality of multiqubit systems.

Secondly, Bell's theorem follows from unentangled measurements and requires an entangled state (although entanglement is not sufficient [52]). Similarly, it follows from the model in Sec. 4 that statistical proofs of the KS theorem employing unentangled measurements also require entangled states for multiqubit systems. Moreover, in Theorem 4, we have shown that the subset of entangled sets that violate a Bell inequality with local projective measurements are exactly those which can yield a finite statistical proof of the KS theorem, where by 'finite' we mean that the proof uses a finite set of measurements.

Finally, our results clarify the connection between Gleason's theorem and the KS theorem with respect to entanglement. While both theorems hold in dimensions greater than two and fail to hold in dimension two, their behaviour with respect to entanglement differs in multiqudit systems. In any multiqudit system that contains both, a subsystem of dimension two and another of dimension at least three, the two theorems diverge. Unentangled measurements are sufficient for a proof of the KS theorem but not Gleason's theorem for such multiqudit systems (Fig. 2).

We have also discussed the implications of our results for the program of understanding the role of contextuality in quantum computation. We find that the assumptions underlying some previously proposed QCSI schemes with qubits [13] become more intuitive in view of our results.

The questions we have raised and addressed in this paper have implications for both fundamental and applied aspects of quantum theory. Further development of the applied aspects, particularly with respect to quantum computation, will be taken up in future work.

Acknowledgements

A Traditional KS

A traditional proof of the KS theorem comprises a set of self-adjoint operators on a Hilbert space that lead to a contradiction

when one attempts to find a *valuation* (see Def. 3 below) on all such self-adjoint operators [12, 53]. A subset of these operators that commute pairwise is known as a context, since all the observables in the subset are jointly measurable. In finite dimensions, and in particular for multiqubit systems, the operators in a context have a simultaneous orthonormal eigenbasis, i.e. an orthonormal basis in which every vector is an eigenvector of every operator in the context. In this appendix we will describe the relationship between the existence of valuations and KS-colourings. Thus, we demonstrate how Theorem 3 implies that for multiqubit systems a set of operators leading to a traditional proof of the KS theorem must contain a context for which the simultaneous orthogonal eigenbasis contains entangled vectors.

Let $\mathcal{L}_{\mathrm{sa}}(\mathcal{H})$ denote the self adjoint operators of a separable Hilbert space $\mathcal{H}$.

Definition 3. A valuation is a map $v : \mathcal{L}_{\mathrm{sa}}(\mathcal{H}) \to \mathbb{R}$ satisfying:

(SPEC) $v(A) \in \sigma(A)$, where $\sigma(A)$ denotes the spectrum of A

(FUNC) $v(g(A)) = g(v(A))$ for any Borel function g.

The requirement (FUNC) implies, for example, that for commuting operators A and B we have $v(AB) = v(A)v(B)$ and $v(aA+bB) = av(A)+bv(B)$ for $a, b \in \mathbb{R}$. Therefore, if there did exist a valuation, v, on $\mathcal{L}_{\mathrm{sa}}(\mathcal{H})$ it would assign values zero or one to each rank-one projection and satisfy $v(\Pi_1) + v(\Pi_2) + \ldots = 1$ for sets of orthogonal projections $\{\Pi_1, \Pi_2, \ldots\}$ summing to the identity. In other words, the valuation would define a KS-colouring when restricted to the rank-one projections of the Hilbert space. The non-existence of such a colouring, stated in Theorem 1, therefore implies the non-existence of a valuation.

In the case of multiqubit product rays we have found that KS-colourings do, however, exist. We may use these colourings to define valuations on certain subsets of self-adjoint operators.

We define a traditional proof of the KS theorem as a set $\mathcal{A}$ of self-adjoint operators such that there does not exist a map v on $\mathcal{A}$ satisfying (SPEC) and (FUNC), where (FUNC) is limited to Borel functions g such that $g(A) \in \mathcal{A}$.[13] We find that such a set $\mathcal{A}$ must contain contexts for which the simultaneous orthogonal eigenbasis contains entangled vectors in multiqubit systems.

Lemma 3. *Consider a set of self-adjoint operators $\mathcal{A}$ on $\mathbb{C}^{2^{\otimes n}}$ such that the simultaneous eigenbases for each context in $\mathcal{A}$ consist entirely of product vectors. Then $\mathcal{A}$ does not form a traditional proof of the KS theorem.*

Proof. By Theorem 3, there exists a KS-colouring, c, on the hypergraph H generated by the simultaneous eigenbases of the contexts in $\mathcal{A}$. As in the treatment of Gleason's theorem in Sec. 2, this colouring can be equivalently defined on the rank-one projections corresponding to each ray and extended to higher rank projections via $c(\Pi_1 + \Pi_2 + \cdots) \equiv c(\Pi_1) + c(\Pi_2) + \cdots$ for mutually orthogonal sets of projections $\{\Pi_1, \Pi_2, \ldots\}$. Given $A \in \mathcal{A}$, let $A = \sum_j \lambda_j \Pi_j$ be its spectral decomposition, where Π_j are orthogonal projections and λ_j are the eigenvalues of A. We will show the map

$$v(A) = \sum_j \lambda_j c(\Pi_j), \qquad (43)$$

satisfies SPEC and FUNC for all $A \in \mathcal{A}$.

For a given operator $A \in \mathcal{A}$, the colouring c takes value one on exactly one of the projections Π_j and zero on the rest. It follows that v satisfies the (SPEC) principle. For any

---

[13]Note that under this definition the set of nine operators in the Peres-Mermin square does *not* produce a KS contradiction, however, if one adds the identity and negative identity to the set then the standard contradiction can be found.

Borel function g such that $g(A) \in \mathcal{A}$, we have

$$v\left(g(A)\right) = v\left(g\left(\sum_j \lambda_j \Pi_j\right)\right) \qquad (44)$$

$$= v\left(\sum_j g(\lambda_j)\Pi_j\right) \qquad (45)$$

$$= \sum_j g(\lambda_j)c(\Pi_j) = g(\lambda_k), \qquad (46)$$

for some $\lambda_k \in \sigma(A)$. And further

$$g(v(A)) = g\left(\sum_{j=1}^d \lambda_j c(\Pi_j)\right) = g(\lambda_k). \qquad (47)$$

Therefore, v also satisfies the (FUNC) principle. $\qquad\square$

B Proof of Theorem 4

Theorem 4. *Any multiqubit density operator, ρ, that can yield a statistical proof of the KS theorem with a finite set of unentangled projective measurements can violate a Bell inequality with local projective measurements.*

Proof. Let

$$\mathcal{S} = \left\{|\psi_j\rangle = |\psi_j^1\rangle |\psi_j^2\rangle \dots |\psi_j^n\rangle \,|\, 1 \le j \le m\right\}, \qquad (48)$$

be a set of m product rays in $\mathbb{C}^{2^{\otimes n}}$, and let ρ be a density operator on $\mathbb{C}^{2^{\otimes n}}$ that yields a statistical proof of the KS theorem with the hypergraph, H, generated by $\mathcal{S}$ and all the bases contained in $\mathcal{S}$. Explicitly, the hypergraph H has m vertices, v_j for $1 \le j \le m$, such that $\{v_j | j \in J\} \in E(H)$ if and only if $\{|\psi_j\rangle |j \in J\}$ is an orthonormal basis, where $J \subset \{1, \dots, m\}$ is some indexing set. The non-classical quantum model on H is given by $p_\rho^{\mathcal{S}}(v_j) = \langle\psi_j|\rho|\psi_j\rangle$. We will show that ρ necessarily violates a Bell inequality.

Consider the set $\mathcal{S}_r$ of distinct local rays of the r-th system that are pairwise non-orthogonal. Explicitly,

$$\mathcal{S}_r = \left\{|\psi_j^r\rangle |j \in J_r\right\}, \qquad (49)$$

where $j \in J_r \subseteq \{1, \dots, m\}$ if and only if $|\psi_j^r\rangle \ne |\psi_k^r\rangle$ and $\langle\psi_j^r|\psi_k^r\rangle \ne 0$ for all $k < j$.

Now we will relabel the states in $\mathcal{S}_r$ as follows

$$\mathcal{S}_r^{\text{re}} = \left\{|0_k^r\rangle |1 \le k \le |\mathcal{S}_r|\right\}.$$

Let $\mathcal{B}(\mathcal{S})$ be the n party Bell scenario in which the r-th party has $|\mathcal{S}_r|$ binary measurement settings. Consider the quantum strategy in this scenario in which the r-th party's k-th measurement is given by

$$x_r = k: \quad \{|0_k^r\rangle, |1_k^r\rangle\}, \qquad (50)$$

where $|1_k^r\rangle \in \mathbb{C}^2$ is the ray orthogonal to $|0_k^r\rangle$. For $a_r \in \{0,1\}$, we denote the n-qubit ray $\bigotimes_{r=1}^n |(a_r)_{x_r}^r\rangle$ by $|(\mathbf{a}|\mathbf{x})\rangle$—in correspondence with its outcome in the Bell scenario. Note that the global measurements, $\{|(\mathbf{a}|\mathbf{x})\rangle\}_{\mathbf{a}}$, in this Bell experiment contain all the rays in $\mathcal{S}$. Let $\mathcal{S}'$ denote the extended set (compared to $\mathcal{S}$) of rays $\{|(\mathbf{a}|\mathbf{x})\rangle\}_{\mathbf{a},\mathbf{x}}$.

Denote by H' the contextuality scenario corresponding to the Bell scenario $\mathcal{B}(\mathcal{S})$. Recall that each hyperedge of H' is given by the outcomes of an adaptive measurement. Equivalently, under the assignment of the rays from $\mathcal{S}'$ to each vertex, each hyperedge corresponds to a projective measurement measurement that can be performed via LOCC. In other words, the hypergraph H' is generated by $\mathcal{S}'$ and those orthonormal bases that can be implemented via LOCC rather than all possible orthonormal bases. It follows that, although the vertices of H' are a superset of those of the original contextuality scenario, H, there could be hyperedges in H which are not contained in $E(H')$ since they arise from "non-local" bases in H, such as the basis in Eq. (1). Denote by G the contextuality scenario generated by rays $\mathcal{S}'$ and *all* the bases, which gives $V(G) = V(H')$ and $E(G) \supseteq E(H') \cup E(H)$.

The entangled state ρ combined with the assignment of the ray $|(\mathbf{a}|\mathbf{x})\rangle$ to each vertex $\mathbf{a}|\mathbf{x}$ gives a quantum model p_ρ on both the hypergraphs, G and H'. Further, we have that p_ρ is a non-classical model on G since it is a non-classical model on a subset of G, namely, H[14]. Now, we will show that the classical

---

[14]The G can only add further constraints on the

models on G and H' coincide exactly, meaning that p_ρ is also a non-classical model on H'. Given the bijection between classical models on H' and local behaviours in the Bell scenario $\mathcal{B}(\mathcal{S})$, it follows that ρ violates a Bell inequality.

Clearly, a classical model on G is a classical model on H' since there are fewer constraints on the classical models on H' due to the reduced set of hyperedges. We now show the converse, using an argument from [45, Theorem 2]. Given an edge $e = \{(\mathbf{a}_j|\mathbf{x}_j) | 1 \leq j \leq 2^n\}$ of G we have that for any pair of distinct vertices, $\mathbf{a}_j|\mathbf{x}_j$ and $\mathbf{a}_k|\mathbf{x}_k$, in the edge the corresponding pair of rays, $|(\mathbf{a}_j|\mathbf{x}_j)\rangle$ and $|(\mathbf{a}_k|\mathbf{x}_k)\rangle$, are orthogonal by the definition of G. Since these rays are n-qubit product rays, we find that for at least one of the single-qubit subsystems, the qubit ray in $|\mathbf{a}_j|\mathbf{x}_j\rangle$ must the orthogonal to qubit ray of the same subsystem of $|\mathbf{a}_k|\mathbf{x}_k\rangle$. By construction, for each qubit ray that occurs in a given subsystem the orthogonal ray only occurs as the other outcome of the same measurement setting, cf. Eq. (50). If this orthogonality occurs in the r-th subsystem, explicitly, we find $(x_r)^j = (x_r)^k$ and $(a_r)^j \neq (a_r)^k$. Thus, we find in the two events $\mathbf{a}_j|\mathbf{x}_j$ and $\mathbf{a}_k|\mathbf{x}_k$ of the Bell scenario, party r has the same measurement setting $(x_r)^j = (x_r)^k$ but observes a different outcome, either $(a_r)^j$ or $(a_r)^k$, i.e. the two events are locally orthogonal in the terminology of Ref. [54].

It follows that for any local deterministic behaviour in the Bell scenario (i.e. classical model on H'), at most one of the events $\mathbf{a}_j|\mathbf{x}_j$ or $\mathbf{a}_k|\mathbf{x}_k$ occurs with probability one, whilst the other must occur with probability zero. Since this relationship holds between any pair of events in the edge $e \in E(G) \supseteq E(H')$, we find that a local deterministic behaviour on H' assigns one to at most one of these outcomes. Hence, for each $e \in E(G) \backslash E(H')$, we

classical models achievable on H, so any non-classical model on H will necessarily be non-classical on G. Or, to take the contrapositive, if a model is classical on G then it must be classical on H because every deterministic model on G will also be a deterministic model on H.

have the following Bell inequality for $\mathcal{B}(\mathcal{S})$ (or, equivalently, a KS-noncontextuality inequality on the contextuality scenario H')

$$\sum_{\mathbf{a}|\mathbf{x}\in e} p_L(\mathbf{a}|\mathbf{x}) \leq 1, \qquad (51)$$

where p_L is a local behaviour (or, equivalently, a classical model on H').

Finally, we show that the Bell inequality of Eq. (51) is saturated by all local behaviours in $\mathcal{B}(\mathcal{S})$. We do so by showing that an internal point of the polytope of local behaviours saturates the inequality and therefore the inequality is trivial, in the sense that it is exactly saturated by all models in the affine span of the local polytope. Observe that the uniform behaviour $p_U(v) = 1/2^n$ for all $v \in V(H')$ is an internal point of the local polytope and saturates the inequality (51).

Any internal behaviour $p_I(\mathbf{a}|\mathbf{x})$ of the local polytope may be expressed as a convex combination

$$p_I(\mathbf{a}|\mathbf{x}) = \omega p_D(\mathbf{a}|\mathbf{x}) + (1 - \omega)p_\delta(\mathbf{a}|\mathbf{x}), \quad (52)$$

of any local deterministic behaviour $p_D(\mathbf{a}|\mathbf{x})$ (a vertex of the local polytope) and some other behaviour on the boundary of the local polytope $p_\delta(\mathbf{a}|\mathbf{x})$, where $0 \leq \omega \leq 1$. Thus, we have

$$\sum_{j=1}^{2^n} p_I(\mathbf{a}^j|\mathbf{x}^j)$$
$$= \omega \left(\sum_{j=1}^{2^n} p_D(\mathbf{a}^j|\mathbf{x}^j) \right) + (1 - \omega) \left(\sum_{j=1}^{2^n} p_\delta(\mathbf{a}^j|\mathbf{x}^j) \right)$$
$$= 1,$$
$$(53)$$

and, therefore, by the inequality (51) we find

$$\sum_{j=1}^{2^n} p_D\left(\mathbf{a}^j|\mathbf{x}^j\right) = 1. \qquad (54)$$

Note that any affine combination of local deterministic behaviours, thus any non-signalling behaviour [55, Corollary 2], also obtains the value one for this Bell expression. In particular, we have

$$\sum_{j=1}^{2^n} p_L\left(\mathbf{a}^j|\mathbf{x}^j\right) = 1, \qquad (55)$$

for any local behaviour p_L in $\mathcal{B}(\mathcal{S})$.

We have shown that any local behaviour in $\mathcal{B}(\mathcal{S})$, and thus classical model, p_C, on H', satisfies $\sum_{v \in e} p_C(v) = 1$ for all hyperedges $e \in E(G)$. Thus, the classical models on H' are exactly the classical models on G. It follows that p_ρ is a non-classical quantum model on H' and therefore violates a Bell inequality. $\square$

C Proof of Theorem 5

Theorem 5. *There exists a KS set consisting entirely of product rays in any separable Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ where $\dim(\mathcal{H}_j) \geq 3$ for some $1 \leq j \leq n$.*

Proof. Let $\mathcal{H} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_{n-1}$ be some separable Hilbert space for which there exists a KS set of bases V^k for $1 \leq k \leq K$ consisting entirely of product vectors v_l for $1 \leq l \leq L$ (including the case $n = 2$ where we consider all vectors to be product). Let $W = \{w_j | 1 \leq j \leq J\}$ be a basis of a separable Hilbert space $\mathcal{H}_n$ and consider the product bases $W^k = \{w_j \otimes v_l | w_j \in W, v_l \in V^k\}$ of $\mathcal{H}_n \otimes \mathcal{H}$ for each $1 \leq k \leq K$. Now assume there is no unentangled KS theorem in $\mathcal{H}_n \otimes \mathcal{H}$ and, therefore, there exists a KS-colouring c of the bases W^k. Consider the map $c'(v_l) = \sum_j c(w_j \otimes v_l)$ on the elements of the bases V^k. We will show this map is a KS-colouring of the bases V^k and thus the assumption that there is no unentangled KS theorem in $\mathcal{H}_n \otimes \mathcal{H}$ must be false.

Firstly, we have that c assigns one to at most one of the vectors in $\{w_j \otimes v_l | 1 \leq j \leq J\}$ and assigns zero to the rest, since c is a KS-colouring and the vectors are mutually orthogonal. Therefore $c'(v_l) \in \{0, 1\}$ for all $1 \leq l \leq L$. Secondly, if $\langle v_l, v_{l'} \rangle = 0$ then the vectors $\{w_j \otimes v_l, w_j \otimes v_{l'} | 1 \leq j \leq J\}$ are mutually orthogonal and c assigns one to at most one of the vectors. It follows that $c'(v_l) + c'(v_{l'}) \leq 1$. Finally, for all W^k we have $c(w \otimes v) = 1$ for some $w \in W$ and $v \in V^k$. Therefore, $c'(v) = \sum_j c(w_j \otimes v) = 1$ for some $v \in V^k$ for all $1 \leq k \leq K$.

Since there exists a KS set in any separable

Hilbert space of dimension at least three, the desired result follows by induction (and the irrelevance of the order of the Hilbert spaces in the tensor product). $\square$

References

[1] Erwin Schrödinger. Discussion of probability relations between separated systems. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 31, pages 555–563. Cambridge University Press, 1935. `doi:10.1017/S0305004100013554`.

[2] Noah Linden and Sandu Popescu. Good dynamics versus bad kinematics: Is entanglement needed for quantum computation? *Phys. Rev. Lett.*, 87:047901, 2001. `doi:10.1103/PhysRevLett.87.047901`.

[3] Animesh Datta and Guifre Vidal. Role of entanglement and correlations in mixed-state quantum computation. *Phys. Rev. A*, 75:042310, 2007. `doi:10.1103/PhysRevA.75.042310`.

[4] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. Negative quasi-probability as a resource for quantum computation. *New J. Phys.*, 14(11):113011, 2012. `doi:10.1088/1367-2630/14/11/113011`.

[5] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, 2014. `doi:10.1038/nature13460`.

[6] Claudio Carmeli, Teiko Heinosaari, and Alessandro Toigo. Quantum random access codes and incompatibility of measurements. *EPL (Europhysics Letters)*, 130(5):50001, 2020. `doi:10.1209/0295-5075/130/50001`.

[7] Toby S Cubitt, Debbie Leung, William Matthews, and Andreas Winter. Improving zero-error classical communication with entanglement. *Phys. Rev.*

Lett., 104:230503, 2010. `doi:10.1103/PhysRevLett.104.230503`.

[8] Shiv Akshar Yadavalli and Ravi Kunjwal. Contextuality in entanglement-assisted one-shot classical communication. *arXiv:2006.00469*, 2020. `doi:10.48550/arXiv.2006.00469`.

[9] Máté Farkas, Maria Balanzó-Juandó, Karol Łukanowski, Jan Kołodyński, and Antonio Acín. Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols. *Phys. Rev. Lett.*, 127:050503, 2021. `doi:10.1103/PhysRevLett.127.050503`.

[10] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, 2018. `doi:10.22331/q-2018-08-06-79`.

[11] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. `doi:10.1038/s41586-019-1666-5`.

[12] Simon Kochen and Ernst P Specker. The problem of hidden variables in quantum mechanics. *J. Math. Mech.*, 17(1):59–87, 1967. `doi:10.1512/iumj.1968.17.17004`.

[13] Juan Bermejo-Vega, Nicolas Delfosse, Dan E Browne, Cihan Okay, and Robert Raussendorf. Contextuality as a resource for models of quantum computation with qubits. *Phys. Rev. Lett.*, 119:120505, 2017. `doi:10.1103/PhysRevLett.119.120505`.

[14] John Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(RX-1376):195–200, 1964. `doi:10.1103/PhysicsPhysiqueFizika.1.195`.

[15] John S Bell. On the Problem of Hidden Variables in Quantum Mechanics. *Rev. Mod. Phys.*, 38:447–452, 1966. `doi:10.1103/RevModPhys.38.447`.

[16] Andrew M Gleason. Measures on the closed subspaces of a Hilbert space. *Indiana Univ. Math. J*, 6:885, 1957. `doi:10.1512/iumj.1957.6.56050`.

[17] Robert W Spekkens. *Quasi-Quantization: Classical Statistical Theories with an Epistemic Restriction*, pages 83–135. Springer Netherlands, Dordrecht, 2016. `doi:10.1007/978-94-017-7303-4_4`.

[18] Ravi Kunjwal and Robert W Spekkens. From the Kochen-Specker theorem to noncontextuality inequalities without assuming determinism. *Phys. Rev. Lett.*, 115:110403, 2015. `doi:10.1103/PhysRevLett.115.110403`.

[19] Ravi Kunjwal and Robert W Spekkens. From statistical proofs of the Kochen-Specker theorem to noise-robust noncontextuality inequalities. *Phys. Rev. A*, 97:052110, 2018. `doi:10.1103/PhysRevA.97.052110`.

[20] Alexander A Klyachko, M Ali Can, Sinem Binicioğlu, and Alexander S Shumovsky. Simple Test for Hidden Variables in Spin-1 Systems. *Phys. Rev. Lett.*, 101:020403, 2008. `doi:10.1103/PhysRevLett.101.020403`.

[21] Robert W Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Phys. Rev. A*, 71:052108, 2005. `doi:10.1103/PhysRevA.71.052108`.

[22] Ravi Kunjwal and Sibasish Ghosh. Minimal state-dependent proof of measurement contextuality for a qubit. *Phys. Rev. A*, 89:042118, 2014. `doi:10.1103/PhysRevA.89.042118`.

[23] Ravi Kunjwal. Contextuality beyond the Kochen–Specker theorem. *arXiv:1612.07250*, 2016. `doi:10.48550/arXiv.1612.07250`.

[24] Paul Busch. Quantum states and generalized observables: a simple proof of Gleason's theorem. *Phys. Rev. Lett.*, 91:120403, 2003. `doi:10.1103/physrevlett.91.120403`.

[25] Carlton M Caves, Christopher A Fuchs, Kiran K Manne, and Joseph M Renes. Gleason-type derivations of the quantum probability rule for generalized measurements. *Found. Phys.*, 34:193–209, 2004. `doi:10.1023/b:foop.0000019581.00318.a5`.

[26] Victoria J Wright and Stefan Weigert. A Gleason-type theorem for qubits based on mixtures of projective measurements. *J. Phys. A*, 52:055301, 2019. `doi:10.1088/1751-8121/aaf93d`.

[27] Nolan R Wallach. An unentangled Gleason's theorem. *Contemp Math*, 305:291–298, 2002. `doi:10.1090/conm/305/05226`.

[28] Charles H Bennett, David P DiVincenzo, Christopher A Fuchs, Tal Mor, Eric Rains, Peter W Shor, John A Smolin, and William K Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, 1999. `doi:10.1103/PhysRevA.59.1070`.

[29] David N Mermin. Hidden variables and the two theorems of John Bell. *Rev. Mod. Phys.*, 65:803–815, 1993. `doi:10.1103/RevModPhys.65.803`.

[30] Asher Peres. Two simple proofs of the Kochen–Specker theorem. *J. Phys. A*, 24(4):L175, 1991. `doi:10.1088/0305-4470/24/4/003`.

[31] Asher Peres. Incompatible results of quantum measurements. *Phys. Lett. A*, 151(3-4):107–108, 1990. `doi:10.1016/0375-9601(90)90172-K`.

[32] Antonio Acín, Tobias Fritz, Anthony Leverrier, and Ana Belén Sainz. A combinatorial approach to nonlocality and contextuality. *Commun. Math. Phys.*, 334(2):533–628, 2015. `doi:10.1007/s00220-014-2260-1`.

[33] Ravi Kunjwal. Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements. *Quantum*, 3:184, 2019. `doi:10.22331/q-2019-09-09-184`.

[34] Ravi Kunjwal. Hypergraph framework for irreducible noncontextuality inequalities from logical proofs of the Kochen-Specker theorem. *Quantum*, 4:219, 2020. `doi:10.22331/q-2020-01-10-219`.

[35] Ehud Hrushovski and Itamar Pitowsky. Generalizations of Kochen and Specker's theorem and the effectiveness of Gleason's theorem. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 35(2):177–194, 2004. `doi:10.1016/j.shpsb.2003.10.002`.

[36] Lin Chen and Dragomir Z Djokovic. Orthogonal product bases of four qubits. *J. Phys. A*, 50(39):395301, 2017. `doi:10.1088/1751-8121/aa8546`.

[37] Matthew S Leifer. Is the quantum state real? An extended review of ψ-ontology theorems. *Quanta*, 3(1):67–155, 2014. `doi:10.12743/quanta.v3i1.22`.

[38] Matthew S Leifer and Owen J E Maroney. Maximally epistemic interpretations of the quantum state and contextuality. *Phys. Rev. Lett.*, 110:120401, 2013. `doi:10.1103/PhysRevLett.110.120401`.

[39] Ravi Kunjwal. Fine's theorem, noncontextuality, and correlations in Specker's scenario. *Phys. Rev. A*, 91:022108, 2015. `doi:10.1103/PhysRevA.91.022108`.

[40] Tomáš Gonda, Ravi Kunjwal, David Schmid, Elie Wolfe, and Ana Belén Sainz. Almost quantum correlations are inconsistent with Specker's principle. 2:87. `doi:10.22331/q-2018-08-27-87`.

[41] Arthur Fine. Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.*, 48:291–295, 1982. `doi:10.1103/physrevlett.48.291`.

[42] Arthur Fine. Joint distributions, quantum correlations, and commuting observables. *J. Math. Phys.*, 23(7):1306–1310, 1982. `doi:10.1063/1.525514`.

[43] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure

of non-locality and contextuality. *New J. Phys.*, 13(11):113036, 2011. `doi:10.1088/1367-2630/13/11/113036`.

[44] Rafael Chaves and Tobias Fritz. Entropic approach to local realism and noncontextuality. *Phys. Rev. A*, 85:032113, 2012. `doi:10.1103/PhysRevA.85.032113`.

[45] Remigiusz Augusiak, Tobias Fritz, Ma Kotowski, Mi Kotowski, Marcin Pawłowski, Maciej Lewenstein, and Antonio Acín. Tight Bell inequalities with no quantum violation from qubit unextendible product bases. *Phys. Rev. A*, 85(4):042113, 2012. `doi:10.1103/physreva.85.042113`.

[46] Victoria J Wright and Ravi Kunjwal. Embedding Peres. *GitHub repository*, 2021. URL: `https://github.com/vickyjwright/embeddingperes`.

[47] Daniel McNulty, Bogdan Pammer, and Stefan Weigert. Mutually unbiased product bases for multiple qudits. *J. Math. Phys.*, 57(3):032202, 2016. `doi:10.1063/1.4943301`.

[48] David Schmid, Haoxing Du, John H Selby, and Matthew F Pusey. The only noncontextual model of the stabilizer subtheory is Gross's. *Phys. Rev. Lett.*, 129:120403, 2021 `doi:10.1103/PhysRevLett.129.120403`.

[49] Daniel Gottesman. The Heisenberg representation of quantum computers. In *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, pages 32–43. Cambridge, MA, International Press, 1998. `doi:10.48550/arXiv.quant-ph/9807006`.

[50] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, 2004. `doi:10.1103/PhysRevA.70.052328`.

[51] Adán Cabello, Simone Severini, and Andreas Winter. Graph-theoretic approach to quantum correlations. *Phys. Rev. Lett.*, 112:040401, 2014. `doi:10.1103/PhysRevLett.112.040401`.

[52] Reinhard F Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989. `doi:10.1103/PhysRevA.40.4277`.

[53] Michael Redhead. *Incompleteness, nonlocality and realism: A prolegomenon to the philosophy of quantum mechanics.* Oxford University Press, 1987.

[54] Tobias Fritz, Ana Belén Sainz, Remigiusz Augusiak, J Bohr Brask, Rafael Chaves, Anthony Leverrier, and Antonio Acín. Local orthogonality as a multipartite principle for quantum correlations. *Nature communications*, 4(1):1–7, 2013. `doi:10.1038/ncomms3263`.

[55] Julien Degorre, Marc Kaplan, Sophie Laplante, and Jérémie Roland. The communication complexity of non-signaling distributions. In *Mathematical Foundations of Computer Science 2009*, pages 270–281, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. `doi:10.1007/978-3-642-03816-7_24`.