

**Minimal state-dependent proof of measurement contextuality for a qubit**

Ravi Kunjwal\* and Sibasish Ghosh†

*Optics & Quantum Information Group, The Institute of Mathematical Sciences, C.I.T Campus, Tharamani, Chennai 600 113, India*

(Received 6 June 2013; revised manuscript received 18 October 2013; published 25 April 2014)

We show that three unsharp binary qubit measurements are enough to violate a generalized noncontextuality inequality, the Liang-Spekkens-Wiseman inequality, in a state-dependent manner. For the case of trine spin axes we calculate the optimal quantum violation of this inequality. In addition, we show that unsharp qubit measurements do not allow a state-independent violation of this inequality. We thus provide a minimal state-dependent proof of measurement contextuality requiring one qubit and three unsharp measurements. Our result rules out generalized noncontextual models of these measurements which were previously conjectured to exist. More importantly, this class of generalized noncontextual models includes the traditional Kochen-Specker (KS) noncontextual models as a proper subset, so our result rules out a larger class of models than those ruled out by a violation of the corresponding KS inequality in this scenario.

DOI: [10.1103/PhysRevA.89.042118](https://doi.org/10.1103/PhysRevA.89.042118)

PACS number(s): 03.65.Ta, 03.65.Ud

**I. INTRODUCTION**

Quantum theory does not admit Bell-local or Kochen-Specker-noncontextual hidden variable (KS-NCHV) models. This is manifest in Bell nonlocality [1,2] and KS contextuality [3,4]. Both these features arise, at a mathematical level, from the lack of a global joint probability distribution over measurement outcomes that recovers the measurement statistics predicted by quantum theory. Traditionally, KS contextuality has been shown for KS-NCHV models of projective measurements for Hilbert spaces of a dimension of three or greater [4–12].

For projective measurements, KS noncontextuality assumes that the outcome of a measurement  $A$  is independent of whether it is performed together with a measurement  $B$ , where  $[A, B] = 0$ , or with measurement  $C$ , where  $[A, C] = 0$  and  $B$  and  $C$  are not compatible, i.e.,  $[B, C] \neq 0$ .  $B$  and  $C$  provide contexts for measurement of  $A$ . A qubit cannot yield a proof of KS contextuality because it does not admit such a triple of projective measurements. While a state-independent proof of KS contextuality holds for any state preparation, a state-dependent proof requires a special choice of the prepared state. The minimal state-independent proof of KS contextuality requires a qutrit and 13 projectors [12,13]. The minimal state-dependent proof [10,11], first given by Klyachko *et al.*, requires a qutrit and five projectors (Fig. 1). Thus a qutrit is the simplest quantum system that allows a proof of KS contextuality, both state-independent and state-dependent. However, we note that generalizations of KS noncontextuality for a qubit have been considered earlier [14–16] in a manner that is different from our approach. The precise difference, and the merits of our approach over earlier attempts, are discussed at length in Ref. [17]. Here we simply note that our approach builds upon the work of Spekkens [18] and Liang *et al.* [19], and we consider generalized noncontextuality proposed by Spekkens as the appropriate notion of noncontextuality for unsharp measurements [17,18]. Generalized noncontextuality allows

outcome-indeterministic response functions for unsharp measurements in the ontological model while KS-NCHV models insist on outcome-deterministic response functions. In particular, generalized noncontextuality insists on the noncontextuality of *probability* assignments to measurement outcomes rather than the stronger KS-noncontextual assumption of the noncontextuality of *value* assignments. A KS-NCHV model is necessarily generalized-noncontextual, but the converse is not true. Further, while the assumption of KS noncontextuality applies to projective measurements in quantum theory, the assumption of generalized noncontextuality applies to all experimental procedures (preparations, transformations, and measurements) in any operational theory.

We define a contextuality scenario as a collection of subsets, called “contexts,” of the set of all measurements. A context refers to measurements that can be jointly implemented. Conceptually, the simplest possible contextuality scenario, first considered by Specker [3] (Fig. 2), requires three two-valued measurements  $\{M_1, M_2, M_3\}$  to allow for three nontrivial contexts:  $\{\{M_1, M_2\}, \{M_1, M_3\}, \{M_2, M_3\}\}$ . Any other choice of contexts will be trivially KS-noncontextual, e.g.,  $\{\{M_1, M_2\}, \{M_1, M_3\}\}$  is KS-noncontextual because the joint probability distribution  $p(M_1, M_2, M_3) \equiv p(M_1, M_2)p(M_1, M_3)/p(M_1)$  reproduces the marginal statistics. By implication, it is also generalized-noncontextual. On assigning outcomes  $\{+1, -1\}$  noncontextually to the three measurements  $\{M_1, M_2, M_3\}$ , it becomes obvious that the maximum number of anticorrelated contexts possible in a single assignment is 2, e.g., for the assignment  $\{M_1 \rightarrow +1, M_2 \rightarrow -1, M_3 \rightarrow +1\}$ ,  $\{M_1, M_2\}$  and  $\{M_2, M_3\}$  are anticorrelated but  $\{M_1, M_3\}$  is not. This puts a KS-noncontextual upper bound of  $2/3$  on the probability of anticorrelation when a context is chosen uniformly at random. Specker’s scenario precludes projective measurements because a set of three pairwise commuting projective measurements is trivially jointly measurable and cannot show any contextuality. One may surmise that it represents a kind of contextuality that is not seen in quantum theory. However, as Liang *et al.* showed [19], this contextuality scenario can be realized using noisy spin-1/2 observables. If one does not assume outcome-determinism for unsharp measurements and models them stochastically but

\*rkunj@imsc.res.in

†sibasish@imsc.res.in

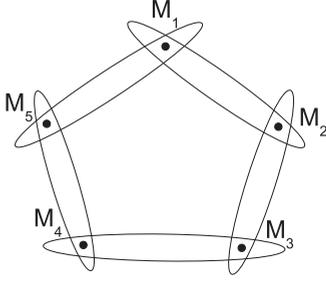


FIG. 1. The KCBS [10] contextuality scenario. The vertices represent the measurements and hyperedges represent jointly measurable contexts.

noncontextually, then this generalized-noncontextual model for noisy spin-1/2 observables will obey a bound of  $1 - \frac{\eta}{3}$ , where  $\eta \in [0, 1]$  is the sharpness associated with each observable. Formally,

$$R_3 \equiv \frac{1}{3} \sum_{(ij) \in \{(12), (23), (13)\}} \Pr(X_i \neq X_j | G_{ij}) \leq 1 - \frac{\eta}{3}, \quad (1)$$

where  $\Pr(X_i \neq X_j | G_{ij})$  is the probability of anticorrelation between the outcomes  $X_i, X_j \in \{+1, -1\}$  of measurements  $M_i$  and  $M_j$ , respectively.  $G_{ij}$  denotes the positive-operator valued measure (POVM) corresponding to the joint implementation of  $M_i$  and  $M_j$ . We will refer to this generalized noncontextuality inequality as the Liang-Spekkens-Wiseman (LSW) inequality. This is *not* a KS-noncontextual inequality, for which the bound would be  $2/3$ . A violation of the LSW inequality will rule out generalized noncontextuality and, by implication, KS noncontextuality. A discussion of this generalized noncontextual model and its ontological meaning, compared to the usual KS-NCHV model, is provided in Appendix A, where we also point out the merits of generalized noncontextuality over KS noncontextuality as a benchmark for nonclassicality. For a more detailed analysis of these issues we refer the interested reader to Refs. [17, 18].

After giving examples of orthogonal and trine spin axes that did not seem to show a violation of this inequality, Liang *et al.* left open the question of whether such a violation exists [19]. They conjectured that all such triples of POVMs will admit a generalized noncontextual model [18], i.e., the LSW inequality will not be violated.

Our main result is a proof that a state-dependent violation of the LSW inequality is possible. In Sect. II we set up the LSW inequality for three unsharp qubit POVMs, in Sec. III we

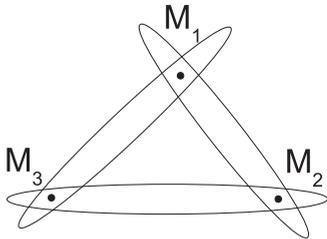


FIG. 2. Specker's [3] contextuality scenario. The vertices represent the measurements and hyperedges represent jointly measurable contexts.

obtain constraints on  $\eta$  from joint measurability, and Sec. IV provides construction of the joint measurement POVMs. In Sec. V we prove that noisy spin-1/2 observables do not allow a state-independent violation of the LSW inequality, followed by our main result in Sec. VI: a state-dependent violation of LSW inequality for the case of trine spin axes. We conclude with some discussion and open questions in Sec. VII.

## II. LSW INEQUALITY

The three POVMs considered,  $M_k = \{E_+^k, E_-^k\}$ ,  $k \in \{1, 2, 3\}$ , are noisy spin-1/2 observables of the form

$$E_{\pm}^k \equiv \frac{1}{2}I \pm \frac{\eta}{2}\vec{\sigma} \cdot \hat{n}_k, \quad 0 \leq \eta \leq 1. \quad (2)$$

That is,

$$E_{\pm}^k = \frac{1-\eta}{2}I + \eta\Pi_{\pm}^k, \quad (3)$$

where  $\Pi_{\pm}^k = \frac{1}{2}(I \pm \vec{\sigma} \cdot \hat{n}_k)$  are the corresponding projectors. So  $E_{\pm}^k$  are noisy versions of the projectors  $\Pi_{\pm}^k$ , and the observable  $\{E_+^k, E_-^k\}$  is therefore a noisy (or unsharp) version of the projective measurement  $P_k = \{\Pi_+^k, \Pi_-^k\}$  (for  $k \in \{1, 2, 3\}$ ). The LSW inequality concerns the following quantity:

$$R_3 \equiv \frac{1}{3} \sum_{(ij) \in \{(12), (23), (13)\}} p(X_i \neq X_j | G_{ij}), \quad (4)$$

where  $X_i, X_j \in \{+1, -1\}$  label measurement outcomes for  $M_i$  and  $M_j$ , respectively. The joint measurement POVM for the context  $\{M_i, M_j\}$  is denoted by  $G_{ij} \equiv \{G_{++}^{ij}, G_{+-}^{ij}, G_{-+}^{ij}, G_{--}^{ij}\}$ .  $G_{X_i, X_j}^{ij}$  is the joint measurement effect corresponding to the effects  $E_{X_i}^i$  and  $E_{X_j}^j$ , i.e.,  $\sum_{X_j} G_{X_i, X_j}^{ij} = E_{X_i}^i$ , and  $\sum_{X_i} G_{X_i, X_j}^{ij} = E_{X_j}^j$ .  $R_3$  is the average probability of anticorrelation when one of the three contexts is chosen uniformly at random.

Under a generalized noncontextual model for these noisy spin-1/2 observables, the following bound on  $R_3$  holds (cf. Ref. [19], Sec. 7.3):

$$R_3 \leq 1 - \frac{\eta}{3}. \quad (5)$$

The question is: Does there exist a triple of noisy spin-1/2 observables that will violate this inequality, perhaps for some specific state-preparation?

## III. JOINT MEASURABILITY CONSTRAINTS ON $\eta$

Testing the LSW inequality for a quantum mechanical violation requires a special kind of joint measurability, denoted by jointly measurable contexts  $\{\{M_1, M_2\}, \{M_2, M_3\}, \{M_1, M_3\}\}$ , i.e., pairwise joint measurability but no triplewise joint measurability. This can be achieved by adding noise to the projective measurements along three different axes. For a given choice of  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  in Eq. (2), denoting  $\hat{n}_i \cdot \hat{n}_j \equiv \cos \theta_{ij}$ , a sufficient condition for this kind of joint measurability is

$$\eta_l < \eta \leq \eta_u, \quad (6)$$

where

$$\eta_l = \frac{1}{3} \max_{X_1, X_2, X_3 \in \{\pm 1\}} \left\{ \sqrt{3 + 2 \sum_{k, j \in \{1, 2, 3\}, k < j} X_k X_j \cos \theta_{kj}} \right\} \quad (7)$$

and

$$\eta_u = \min_{(ij) \in \{(12), (23), (13)\}} \left\{ \frac{1}{\sqrt{1 + |\sin \theta_{ij}|}} \right\}. \quad (8)$$

These are obtained as special cases of the more general joint measurability conditions obtained in Appendix B, based on Refs. [19] and [20]. We note that this condition is necessary and sufficient for the special case of trine ( $\hat{n}_i \cdot \hat{n}_j = -1/2$ ) and orthogonal ( $\hat{n}_i \cdot \hat{n}_j = 0$ ) spin axes.

#### IV. JOINT MEASUREMENT EFFECTS

We construct the joint measurement POVM  $G_{ij} = \{G_{+++}^{ij}, G_{++-}^{ij}, G_{+-+}^{ij}, G_{--}^{ij}\}$ , such that the given POVMs,  $M_i = \{E_+^i, E_-^i\}$  and  $M_j = \{E_+^j, E_-^j\}$ , are recovered as marginals, i.e.,  $\sum_{X_j} G_{X_i, X_j}^{ij} = E_{X_i}^i$ ,  $\sum_{X_i} G_{X_i, X_j}^{ij} = E_{X_j}^j$ ,  $0 \leq G_{X_i, X_j}^{ij} \leq I$ , and  $\sum_{X_i, X_j} G_{X_i, X_j}^{ij} = I$ , where  $X_i, X_j \in \{+1, -1\}$ . The joint measurement POVM has the following general form:

$$G_{+++}^{ij} = \frac{1}{2} \left\{ \frac{\alpha_{ij}}{2} I + \vec{\sigma} \cdot \frac{1}{2} [\eta(\hat{n}_i + \hat{n}_j) - \vec{a}_{ij}] \right\}, \quad (9)$$

$$G_{++-}^{ij} = \frac{1}{2} \left\{ \left(1 - \frac{\alpha_{ij}}{2}\right) I + \vec{\sigma} \cdot \frac{1}{2} [\eta(\hat{n}_i - \hat{n}_j) + \vec{a}_{ij}] \right\}, \quad (10)$$

$$G_{+-+}^{ij} = \frac{1}{2} \left\{ \left(1 - \frac{\alpha_{ij}}{2}\right) I + \vec{\sigma} \cdot \frac{1}{2} [\eta(-\hat{n}_i + \hat{n}_j) + \vec{a}_{ij}] \right\}, \quad (11)$$

$$G_{--}^{ij} = \frac{1}{2} \left\{ \frac{\alpha_{ij}}{2} I + \vec{\sigma} \cdot \frac{1}{2} [\eta(-\hat{n}_i - \hat{n}_j) - \vec{a}_{ij}] \right\}, \quad (12)$$

where  $I$  is the  $2 \times 2$  identity matrix,  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  are the  $2 \times 2$  Pauli matrices,  $\alpha_{ij} \in \mathbb{R}$ , and  $\vec{a}_{ij} \in \mathbb{R}^3$ . The necessary and sufficient conditions for these to be valid qubit effects,  $0 \leq G_{X_i, X_j}^{ij} \leq I$ ,  $\forall X_i, X_j \in \{+1, -1\}$ , are equivalent to the following inequalities [21]:

$$\sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2 + 2\eta|(\hat{n}_i + \hat{n}_j) \cdot \vec{a}_{ij}|} \leq \alpha_{ij}, \quad (13)$$

$$\alpha_{ij} \leq 2 - \sqrt{2\eta^2(1 - \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2 + 2\eta|(\hat{n}_i - \hat{n}_j) \cdot \vec{a}_{ij}|}, \quad (14)$$

where  $\eta_l < \eta \leq \eta_u$ . The construction of the joint measurement POVM and derivation of the necessary and sufficient condition for its validity, Eqs. (13) and (14), are provided in Appendix C. The joint measurement effects corresponding to anticorrelation sum to

$$G_{+-}^{ij} + G_{-+}^{ij} = \left(1 - \frac{\alpha_{ij}}{2}\right) I + \frac{1}{2} \vec{\sigma} \cdot \vec{a}_{ij}. \quad (15)$$

#### V. NO STATE-INDEPENDENT VIOLATION

We will now show that no state-independent violation of the LSW inequality with qubit POVMs is possible.

*Theorem 1.* There exists no state-independent violation of the generalized-noncontextual inequality  $R_3 \leq 1 - \frac{\eta}{3}$  using a triple of qubit POVMs  $M_k \equiv \{E_{\pm}^k\}, k \in \{1, 2, 3\}$  that are pairwise jointly measurable but not triplewise jointly measurable.

*Proof.* In quantum theory, the probability  $R_3^Q$  for anticorrelation of measurement outcomes for pairwise joint measurements of  $M_k \equiv \{E_{\pm}^k, E_{\mp}^k\}$  (where  $k \in \{1, 2, 3\}$ ) has the following form for a qubit state  $\rho$ :

$$R_3^Q \equiv \frac{1}{3} \sum_{(ij) \in \{(12), (23), (13)\}} \text{Tr}[\rho(G_{+-}^{ij} + G_{-+}^{ij})]. \quad (16)$$

The condition for violation of noncontextuality inequality (5) is  $R_3^Q > 1 - \frac{\eta}{3}$ . Using Eq. (15), this reduces to

$$\text{Tr} \left( \rho \sum_{(ij)} (\alpha_{ij} I - \vec{\sigma} \cdot \vec{a}_{ij}) \right) < 2\eta. \quad (17)$$

Using the standard  $2 \times 2$  Pauli matrices and  $\rho$  parameterized by  $0 \leq q \leq 1$  and  $\hat{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ :

$$\rho = q |\psi\rangle\langle\psi| + (1 - q)(I - |\psi\rangle\langle\psi|), \quad (18)$$

$$|\psi\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix} = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle, \quad (19)$$

the condition for violation becomes

$$\sum_{(ij)} \alpha_{ij} + \lambda_{\rho} < 2\eta, \quad (20)$$

where

$$\lambda_{\rho} \equiv (1 - 2q) \vec{a} \cdot \hat{n} \in [-|\vec{a}|, |\vec{a}|] \quad (21)$$

denotes the state-dependent term in the condition and  $\vec{a} = (a_x, a_y, a_z)$  is given by

$$a_x = \sum_{(ij)} (\vec{a}_{ij})_x, \quad a_y = \sum_{(ij)} (\vec{a}_{ij})_y, \quad a_z = \sum_{(ij)} (\vec{a}_{ij})_z. \quad (22)$$

For a state-independent violation, either the state-dependent term in Eq. (20),  $\lambda_{\rho}$ , must vanish for all qubit states  $\rho$ , or  $\sum_{(ij)} \alpha_{ij} + \max_{\rho} \lambda_{\rho} < 2\eta$  should hold. The first case,  $\lambda_{\rho} = 0 \forall \rho$ , requires  $\vec{a} = 0$  since  $\vec{a}$  is the only term in  $\lambda_{\rho}$  that depends on the joint measurement POVM. This means  $a_x = a_y = a_z = 0$ , so that  $\lambda_{\rho} = 0$  for all  $\rho$ . The second case requires  $\sum_{(ij)} \alpha_{ij} + |\vec{a}| < 2\eta$ . In both cases, we have the following lower bound on  $\alpha_{ij}$ , from inequality (13):

$$\alpha_{ij} > \sqrt{2\eta\sqrt{1 + \hat{n}_i \cdot \hat{n}_j}}. \quad (23)$$

Taking the sum of  $\alpha_{ij}$ ,  $(ij) \in \{(12), (23), (13)\}$ , we have

$$\sum_{(ij)} \alpha_{ij} > \sqrt{2\eta} \sum_{(ij)} \sqrt{1 + \hat{n}_i \cdot \hat{n}_j}. \quad (24)$$

For the first case, the condition for state-independent violation is  $\sum_{(ij)} \alpha_{ij} < 2\eta$ , while for the second case the condition for such a violation is  $\sum_{(ij)} \alpha_{ij} + |\vec{a}| < 2\eta$ . Given the lower bound

on  $\sum_{(ij)} \alpha_{ij}$ , it follows that a necessary condition for state-independent violation of the LSW inequality is

$$\sum_{(ij)} \sqrt{1 + \hat{n}_i \cdot \hat{n}_j} < \sqrt{2}. \quad (25)$$

We will show that there exists no choice of measurement directions that will satisfy this necessary condition, thereby ruling out a state-independent violation of the LSW inequality. The particular cases of orthogonal axes ( $\hat{n}_i \cdot \hat{n}_j = 0$ ) or trine spin axes ( $\hat{n}_i \cdot \hat{n}_j = -1/2$ ), used by the authors of Ref. [19], are clearly ruled out by this necessary condition. Denoting  $\hat{n}_i \cdot \hat{n}_j \equiv \cos \theta_{ij}$ , the necessary condition for violation is

$$\left| \cos \frac{\theta_{12}}{2} \right| + \left| \cos \frac{\theta_{13}}{2} \right| + \left| \cos \frac{\theta_{23}}{2} \right| < 1. \quad (26)$$

Without loss of generality, the three directions can be parameterized as follows:

$$\hat{n}_1 \equiv (0, 0, 1), \quad (27)$$

$$\hat{n}_2 \equiv (\sin \theta_{12}, 0, \cos \theta_{12}), \quad (28)$$

$$\hat{n}_3 \equiv (\sin \theta_{13} \cos \phi_3, \sin \theta_{13} \sin \phi_3, \cos \theta_{13}). \quad (29)$$

where

$$0 < \frac{\theta_{ij}}{2} < \frac{\pi}{2} \quad \forall (ij) \in \{(12), (13), (23)\}, \quad 0 \leq \phi_3 < 2\pi,$$

and  $\cos \theta_{23} = \sin \theta_{12} \sin \theta_{13} \cos \phi_3 + \cos \theta_{12} \cos \theta_{13}$ . This implies

$$\cos(\theta_{12} + \theta_{13}) \leq \cos(\theta_{23}) \leq \cos(\theta_{12} - \theta_{13}). \quad (30)$$

Then

$$\begin{aligned} & \min_{\theta_{12}, \theta_{13}, \theta_{23}} \left\{ \left| \cos \frac{\theta_{12}}{2} \right| + \left| \cos \frac{\theta_{13}}{2} \right| + \left| \cos \frac{\theta_{23}}{2} \right| \right\} \\ & \geq \min_{\theta_{12}, \theta_{13}} \left\{ \left| \cos \frac{\theta_{12}}{2} \right| + \left| \cos \frac{\theta_{13}}{2} \right| + \sqrt{\frac{1 + \cos(\theta_{12} + \theta_{13})}{2}} \right\} \\ & > 1. \end{aligned}$$

This contradicts the necessary condition (26). Hence, there is no state-independent violation of the LSW inequality (5) allowed by noisy spin-1/2 observables. ■

## VI. STATE-DEPENDENT VIOLATION OF THE LSW INEQUALITY

Our main result is that the LSW inequality can be violated in a state-dependent manner. From the condition for violation (20), it follows that a necessary condition for state-dependent violation is  $\sum_{(ij)} \alpha_{ij} - |\vec{a}| < 2\eta$ . An optimal choice of  $\rho$  that yields  $\lambda_\rho = -|\vec{a}|$  corresponds to  $q = 1$  and  $\vec{a} \cdot \hat{n} = |\vec{a}|$ , i.e.,

$$\cos \theta = \frac{a_z}{|\vec{a}|}, \quad \tan \phi = \frac{a_y}{a_x}.$$

With this choice of  $\rho$  the question becomes: Does there exist a choice of  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}, \{\alpha_{ij}, \vec{a}_{ij}\}$  such that  $\sum_{(ij)} \alpha_{ij} - |\vec{a}| <$

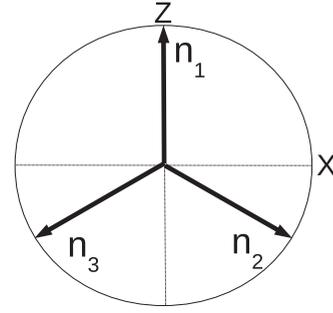


FIG. 3. Choice of measurement directions  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  along trine spin axes in the  $Z$ - $X$  plane.

$2\eta$ ? We show that this is indeed the case. We define

$$C \equiv 2\eta - \left( \sum_{(ij)} \alpha_{ij} - |\vec{a}| \right), \quad (31)$$

so that  $C > 0$  indicates a state-dependent violation. Note that violation of the LSW inequality  $R_3^Q \leq 1 - \frac{\eta}{3}$  is characterized by

$$S \equiv R_3^Q - \left( 1 - \frac{\eta}{3} \right) = \frac{C}{6}, \quad (32)$$

where  $S > 0$  for a state-dependent violation. Given a coplanar choice of  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$ , and  $\eta$  satisfying  $\eta_l < \eta \leq \eta_u$ , the optimal value of  $C$ , denoted as  $C_{\max}^{\{\hat{n}_i\}, \eta}$ , is given by

$$\begin{aligned} C_{\max}^{\{\hat{n}_i\}, \eta} &= 2\eta + \sum_{(ij)} \left( \sqrt{1 + \eta^4 (\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2} \right. \\ & \quad \left. - (1 + \eta^2 \hat{n}_i \cdot \hat{n}_j) \right), \end{aligned} \quad (33)$$

which is obtained in Appendix D. We obtain a state-dependent violation of the LSW inequality for trine axes (Fig. 3).

*Theorem 2.* The optimal violation of the LSW inequality for measurements along trine spin axes, i.e.,  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  such that  $\hat{n}_1 \cdot \hat{n}_2 = \hat{n}_2 \cdot \hat{n}_3 = \hat{n}_1 \cdot \hat{n}_3 = -1/2$ , occurs for  $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$  if the plane of the measurements is in the  $ZX$  plane. The lower and upper bounds on  $\eta$  are  $\eta_l = \frac{2}{3} \approx 0.667$  and  $\eta_u = \sqrt{3} - 1 \approx 0.732$ . The joint measurement POVM is given by  $\alpha_{ij} = 1 + \eta^2 \hat{n}_i \cdot \hat{n}_j$  and  $\vec{a}_{ij} = (0, \sqrt{1 + \eta^4 (\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2}, 0)$ . The optimal violation corresponds to  $\eta \rightarrow \eta_l$ , so that  $\alpha_{12} = \alpha_{13} = \alpha_{23} \rightarrow 1 - \frac{\eta_l^2}{2} = \frac{7}{9}$ ,  $|\vec{a}_{ij}| \rightarrow \frac{\sqrt{13}}{9} \quad \forall (ij)$ ,  $C_{\max}^{\text{trine}} \rightarrow \frac{\sqrt{13}}{3} - 1 \approx 0.20185$ , and  $S_{\max}^{\text{trine}} = \frac{C_{\max}^{\text{trine}}}{6} \rightarrow 0.03364$  or 3.36%.

Thus the quantum probability of anticorrelation can exceed the generalized-noncontextual bound by an amount arbitrarily close to 0.03364 or about 3.36% for trine spin axes. The quantum degree of anticorrelation for this violation is  $R_3^Q = S_{\max}^{\text{trine}} + (1 - \frac{\eta_l}{3}) \rightarrow 0.8114$  and the generalized-noncontextual bound is  $(1 - \frac{\eta_l}{3}) \rightarrow \frac{7}{9} \approx 0.7778$ . The proof of Theorem 2 follows from Appendix D.

## VII. CONCLUSION

A violation of the LSW inequality is interesting primarily because the benchmark for nonclassicality set by generalized noncontextuality is more stringent than the one set by the traditional notion of KS noncontextuality. The LSW inequality takes into account, for example, the possibility that the measurement apparatus could introduce anticorrelations that have nothing to do with hidden variable(s) one could associate with the system's preparation. This would allow violation of the KS-noncontextual bound of  $2/3$  when the measurement is unsharp ( $\eta < 1$ ) even though this violation could purely be a result of noise coming from elsewhere, such as the measurement apparatus, rather than a consequence of quantum theory. A violation of the LSW inequality, on the other hand, rules out this possibility and certifies genuine nonclassicality that cannot be attributed to hidden variables associated with either the preparation or the noise. As argued by Spekkens [18], the appropriate notion of noncontextuality for unsharp measurements is one that allows outcome-indeterministic response functions.

An interesting open question is whether such a violation is possible in higher-dimensional systems and whether the amount of violation could be higher for these than for a qubit. Whether a state-independent violation of the LSW inequality is possible in higher dimensions also remains an open question. Our result also hints at the fact that perhaps all contextuality scenarios may be realizable and contextuality demonstrated if we consider the possibilities that general quantum measurements allow. In particular, scenarios that involve pairwise compatibility between all measurements but no global compatibility may be realizable within quantum theory. Specker's scenario is the simplest such example we have considered. Indeed, as later shown by the authors of Ref. [22] recently, quantum theory does admit all contextuality scenarios since it allows one to realize any conceivable set of (in)compatibility relations between a set of observables.

In summary, the joint measurability allowed in a theory restricts the kind of contextuality scenarios that can arise in it. Quantum theory admits Specker's contextuality scenario with unsharp measurements [19]. Further, as we have shown, quantum theory allows violation of the LSW inequality in this scenario. Thus quantum theory is contextual even in the simplest contextuality scenario. Whether, and to what extent, this is the case with more complicated contextuality scenarios realizable, for example, via the construction in Ref. [22] remains to be explored.

*Note.* In Ref. [23], which appeared recently, the authors deal with the LSW inequality and make some remarks on the results of this paper. We refer the interested reader to Ref. [24] for a discussion of the claims in Ref. [23] compared to the results of this paper. See also Appendix E for a brief remark on the question of triplewise incompatibility.

## ACKNOWLEDGMENTS

R. K. would like to thank Rob Spekkens for comments on earlier drafts of this work; in particular, for asking whether a state-dependent violation of the LSW inequality was possible,

and for clarifying the operational meaning of the generalized-noncontextual bound in this inequality. Thanks are also due to Andreas Winter for asking some tough questions about benchmarks for nonclassicality. We also thank the anonymous referees for their comments which made us re-examine and revise some of our results.

## APPENDIX A: GENERALIZED-NONCONTEXTUAL MODELS VERSUS KS-NONCONTEXTUAL MODELS

The traditional assumption of KS noncontextuality entails two things: measurement noncontextuality and outcome-determinism for sharp measurements [18]. Given a set of measurements  $\{M_1, \dots, M_N\}$ , measurement noncontextuality is the assumption that the response function for each measurement is insensitive to contexts (jointly measurable subsets) that it may be a part of:  $\forall M_i, p(X_i|M_i; \lambda) \in [0, 1]$ . Here  $X_i$  is an outcome for measurement  $M_i$  and  $\lambda$  is the hidden variable associated with the system's preparation. Outcome-determinism is the further assumption that  $\forall M_i, p(X_i|M_i; \lambda) \in \{0, 1\}$ , i.e., response functions are outcome-deterministic. A KS-NCHV model is one that makes these two assumptions for sharp (projective) measurements. A KS inequality is a constraint on measurement statistics obtained under these two assumptions. A generalized-noncontextual model, on the other hand, derives outcome-determinism for sharp measurements as a consequence of preparation noncontextuality [18]. For unsharp measurements, however, outcome-determinism is not implied by generalized noncontextuality and one needs to model these measurements by outcome-indeterministic response functions [17, 18]. We refer the reader to Ref. [17] for a detailed critique of the assumption of outcome-determinism for unsharp measurements and for arguments on the reasonableness, and generality, of the notion of noncontextuality for unsharp measurements that is the basis of our work. Indeed, the qubit effects we need to write the response functions for are of the form  $E_{\pm}^k = \eta \Pi_{\pm}^k + (1 - \eta) \frac{1}{2}$ . We will relabel the outcomes according to  $\{+1 \rightarrow 0, -1 \rightarrow 1\}$  so that  $X_k \in \{0, 1\}$  in what follows. Liang, Spekkens, and Wiseman (LSW) argued [19] that the response function for these effects in a generalized-noncontextual model should be  $p(X_k|M_k; \lambda) = \eta[X_k(\lambda)] + (1 - \eta)(\frac{1}{2}[0] + \frac{1}{2}[1])$ , where  $p(X) = [x]$  denotes the point distribution given by the Kronecker delta function  $\delta_{X,x}$ . For  $\eta = 1$  (sharp measurements) this would be the traditional KS-noncontextual model. When  $\eta < 1$  (unsharp measurements), the second "coin flip" term in the response function  $(\frac{1}{2}[0] + \frac{1}{2}[1])$  begins to play a role. This term is not conditioned by  $\lambda$ , the hidden variable associated with the system's preparation, but is instead the response function for tossing a fair coin regardless of what measurement is being made. It characterizes the random noise introduced, for example, by the measuring apparatus. The important thing to note is that this noise is uncorrelated with the system's hidden variable  $\lambda$ .

Given these single-measurement response functions, one needs to figure out pairwise response functions for pairwise joint measurements of the three qubit POVMs. LSW [19] argued that the pairwise response functions maximizing the average anticorrelation  $R_3$  and consistent with the

single-measurement response functions are given by

$$p(X_i, X_j | G_{ij}; \lambda) = \eta [X_i(\lambda)] [X_j(\lambda)] + (1 - \eta) \left( \frac{1}{2} [0][1] + \frac{1}{2} [1][0] \right), \quad (\text{A1})$$

for all pairs of measurements  $(ij) \in \{(12), (13), (23)\}$ . This generalized-noncontextual model for these measurements turns out to be KS-contextual in the sense that the three pairwise response functions do not admit a joint probability distribution over the three measurement outcomes,  $p(X_1, X_2, X_3 | \lambda)$ , which is consistent with all three of them. Indeed this LSW model maximizes the average anticorrelation possible in Specker's scenario given the single-measurement response functions, thus allowing us to obtain the LSW inequality  $R_3 \leq 1 - \frac{\eta}{3}$ . Let us note the two bounds separately:

$$R_3 \leq R_3^{\text{KS}} \equiv \frac{2}{3}, \quad (\text{A2})$$

$$R_3 \leq R_3^{\text{LSW}} \equiv 1 - \frac{\eta}{3}. \quad (\text{A3})$$

To be clear, the assumptions leading to the LSW inequality are as follows:

- (1) Measurement noncontextuality;
- (2) Outcome-determinism for projective measurements; and
- (3) No outcome-determinism for nonprojective measurements.

On the other hand, the assumptions that lead to the corresponding KS inequality (upper bound of  $2/3$ ) are the following:

- (1) Measurement noncontextuality and
- (2) Outcome-determinism for *all* (projective as well as nonprojective) measurements.

The first set of assumptions is clearly weaker than the second set of assumptions and violation of the LSW inequality rules out even this weaker notion of noncontextuality. Our main result in this paper is that there exists a choice of measurement directions  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  and a choice of  $\eta$  for some state  $\rho$  such that the quantum probability of anticorrelation  $R_3^Q$  beats the generalized-noncontextual bound  $R_3^{\text{LSW}}$ . This rules out the possibility of being able to generate these correlations by classical means, as in the LSW model, for at least some values of sharpness parameter  $\eta$ . Of course, if  $\eta = 0$ , then the generalized-noncontextual bound becomes trivial and the question of violation does not arise: This situation corresponds to the case where for any of the three pairwise joint measurements, the measuring apparatus outputs one of the two anticorrelated outcomes by flipping a fair coin and there is no functional dependence of the response function on  $\lambda$ . In other words, one could generate perfect anticorrelation in a generalized-noncontextual model if  $\eta = 0$ . However, as long as one is performing a nontrivial measurement (where  $\eta > 0$ ) there is a constraint on the degree of anticorrelation imposed by generalized noncontextuality. What we establish is that generalized noncontextuality cannot account for the degree of anticorrelation observed in quantum theory. Clearly, quantum theory is nonclassical even given a more stringent benchmark than the one set by KS noncontextuality. A violation of the KS-noncontextual bound  $R_3^{\text{KS}}$  is possible in a generalized-noncontextual model, so such a violation is not

in itself a signature of nonclassicality. On the other hand, violation of the generalized-noncontextual bound  $R_3^{\text{LSW}}$  should be considered a signature of genuine nonclassicality in that it is not attributable either to the system's hidden variable or random noise (from the measuring apparatus or elsewhere).

## APPENDIX B: BOUNDS ON $\eta$ FROM JOINT MEASURABILITY

In Appendix F of Ref. [19], Theorem 13, the authors obtained the necessary and sufficient conditions for joint measurability of noisy spin-1/2 observables. We note, as pointed out by an anonymous referee, that the claimed necessary condition in the aforementioned theorem is incorrect, while the sufficient condition holds. Here we prove a necessary condition for joint measurability, one we use in the main text for triplewise joint measurability, by revising the argument for necessity made by Liang *et al.* [19].

*Theorem 3.* Given a set of qubit POVMs,  $\{\{E_{X_k}^k : X_k \in \{+1, -1\}\} | k \in \{1, \dots, N\}\}$ , of the form

$$E_{X_k}^k = \frac{1}{2} I + \frac{1}{2} \vec{\sigma} \cdot X_k \eta \hat{n}_k, \quad (\text{B1})$$

and defining  $2^N$  3-vectors

$$\vec{m}_{X_1, \dots, X_N} \equiv \sum_{k=1}^N X_k \hat{n}_k, \quad (\text{B2})$$

a necessary condition for all the POVMs to be jointly measurable is that

$$\eta \leq \frac{1}{N} \max_{X_1, \dots, X_N} \{|\vec{m}_{X_1, \dots, X_N}|\}, \quad (\text{B3})$$

and a sufficient condition is that

$$\eta \leq \frac{2^N}{\sum_{X_1, \dots, X_N} |\vec{m}_{X_1, \dots, X_N}|}. \quad (\text{B4})$$

*Proof.* We will only prove the necessary condition, which we use in the main text, and refer the reader to Ref. [19], Appendix F, for proof of the sufficient condition. Note that  $\eta = \text{Tr}[(\vec{\sigma} \cdot X_k \hat{n}_k) E_{X_k}^k]$ . Since this holds  $\forall X_k, k$ , we have

$$\eta = \frac{1}{2N} \sum_{k=1}^N \sum_{X_k} \text{Tr}[(\vec{\sigma} \cdot X_k \hat{n}_k) E_{X_k}^k]. \quad (\text{B5})$$

If all the POVMs are jointly measurable, then we must necessarily have a joint POVM  $\{E_{X_1, \dots, X_N}\}$  such that

$$E_{X_k}^k = \sum_{X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_N} E_{X_1, \dots, X_N}. \quad (\text{B6})$$

Then

$$\eta = \frac{1}{2N} \sum_{X_1, \dots, X_N} \text{Tr} \left[ \left( \vec{\sigma} \cdot \sum_{k=1}^N X_k \hat{n}_k \right) E_{X_1, \dots, X_N} \right], \quad (\text{B7})$$

and using  $\hat{m}_{X_1, \dots, X_N} \equiv \vec{m}_{X_1, \dots, X_N} / |\vec{m}_{X_1, \dots, X_N}|$ , we have

$$\eta = \frac{1}{2N} \sum_{X_1, \dots, X_N} |\vec{m}_{X_1, \dots, X_N}| \text{Tr}[(\vec{\sigma} \cdot \hat{m}_{X_1, \dots, X_N}) E_{X_1, \dots, X_N}]. \quad (\text{B8})$$

Further,

$$\text{Tr}[(\vec{\sigma} \cdot \hat{m}_{X_1, \dots, X_N}) E_{X_1, \dots, X_N}] \leq \text{Tr}[E_{X_1, \dots, X_N}], \quad (\text{B9})$$

which yields the inequality

$$\eta \leq \frac{1}{2N} \sum_{X_1, \dots, X_N} |\vec{m}_{X_1, \dots, X_N}| \text{Tr}[E_{X_1, \dots, X_N}]. \quad (\text{B10})$$

Now,  $\sum_{X_1, \dots, X_N} E_{X_1, \dots, X_N} = I$ , and therefore,

$$\sum_{X_1, \dots, X_N} \frac{1}{2} \text{Tr}[E_{X_1, \dots, X_N}] = 1. \quad (\text{B11})$$

Also,  $0 \leq \frac{1}{2} \text{Tr}[E_{X_1, \dots, X_N}] \leq 1$ , so  $\{\frac{1}{2} \text{Tr}[E_{X_1, \dots, X_N}]\}_{X_1, \dots, X_N}$  is a probability distribution and therefore

$$\eta \leq \frac{1}{N} \max_{X_1, \dots, X_N} \{|\vec{m}_{X_1, \dots, X_N}|\}, \quad (\text{B12})$$

which is a necessary condition for joint measurability. ■

For  $N = 3$  we obtain the necessary condition for triplewise joint measurability which is used in the main text for computing  $\eta_l$ . The necessary and sufficient condition for pairwise joint measurability is given by

$$1 + \eta^4 (\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2 \geq 0, \quad \forall (ij) \in \{(12), (13), (23)\}. \quad (\text{B13})$$

This is obtained as a special case, for the present problem, of the more general necessary and sufficient condition for joint measurability of unsharp qubit observables obtained by the authors of Ref. [20]. Using  $\hat{n}_i \cdot \hat{n}_j = \cos \theta_{ij}$ , this inequality becomes

$$\left( \eta^2 - \frac{1}{1 - |\sin \theta_{ij}|} \right) \left( \eta^2 - \frac{1}{1 + |\sin \theta_{ij}|} \right) \geq 0. \quad (\text{B14})$$

Since  $0 \leq \eta \leq 1$ , the necessary and sufficient condition for pairwise joint measurability becomes

$$\eta \leq \min_{(ij) \in \{(12), (23), (13)\}} \left\{ \frac{1}{\sqrt{1 + |\sin \theta_{ij}|}} \right\}, \quad (\text{B15})$$

which is used to compute  $\eta_u$  in the main text.

*Orthogonal spin axes:*  $\hat{n}_i \cdot \hat{n}_j = 0 \forall (ij) \in \{(12), (13), (23)\}$ . The necessary and sufficient joint measurability condition is

$$\frac{1}{\sqrt{3}} < \eta \leq \frac{1}{\sqrt{2}}. \quad (\text{B16})$$

*Trine spin axes:*  $\hat{n}_i \cdot \hat{n}_j = -1/2 \forall (ij) \in \{(12), (13), (23)\}$ . The necessary and sufficient joint measurability condition is

$$\frac{2}{3} < \eta \leq \sqrt{3} - 1. \quad (\text{B17})$$

### APPENDIX C: CONSTRUCTING THE JOINT MEASUREMENT POVM

The joint measurement POVM  $G_{ij}$  for  $\{M_i, M_j\}$  should satisfy the marginal condition

$$G_{++}^{ij} + G_{+-}^{ij} = E_+^i, \quad G_{-+}^{ij} + G_{--}^{ij} = E_-^i, \quad (\text{C1})$$

$$G_{++}^{ij} + G_{-+}^{ij} = E_+^j, \quad G_{+-}^{ij} + G_{--}^{ij} = E_-^j. \quad (\text{C2})$$

Also, the joint measurement should consist of valid effects

$$0 \leq G_{++}^{ij}, G_{+-}^{ij}, G_{-+}^{ij}, G_{--}^{ij} \leq I, \quad (\text{C3})$$

where  $I$  is the  $2 \times 2$  identity matrix. The general form of the joint measurement effects is

$$G_{++}^{ij} = \frac{1}{2} \left[ \frac{\alpha_{ij}}{2} I + \vec{\sigma} \cdot \vec{a}_{++}^{ij} \right], \quad (\text{C4})$$

$$G_{+-}^{ij} = \frac{1}{2} \left[ \left( 1 - \frac{\alpha_{ij}}{2} \right) I + \vec{\sigma} \cdot \vec{a}_{+-}^{ij} \right], \quad (\text{C5})$$

$$G_{-+}^{ij} = \frac{1}{2} \left[ \left( 1 - \frac{\alpha_{ij}}{2} \right) I + \vec{\sigma} \cdot \vec{a}_{-+}^{ij} \right], \quad (\text{C6})$$

$$G_{--}^{ij} = \frac{1}{2} \left[ \frac{\alpha_{ij}}{2} I + \vec{\sigma} \cdot \vec{a}_{--}^{ij} \right], \quad (\text{C7})$$

where each effect is parameterized by four real numbers. From the marginal condition, Eqs. (C1) and (C2), it follows that

$$\vec{a}_{++}^{ij} + \vec{a}_{+-}^{ij} = \eta \hat{n}_i, \quad \vec{a}_{-+}^{ij} + \vec{a}_{--}^{ij} = -\eta \hat{n}_i, \quad (\text{C8})$$

$$\vec{a}_{-+}^{ij} + \vec{a}_{++}^{ij} = \eta \hat{n}_j, \quad \vec{a}_{--}^{ij} + \vec{a}_{+-}^{ij} = -\eta \hat{n}_j. \quad (\text{C9})$$

These can be rewritten as

$$2\vec{a}_{++}^{ij} + \vec{a}_{+-}^{ij} + \vec{a}_{-+}^{ij} = \eta(\hat{n}_i + \hat{n}_j), \quad (\text{C10})$$

$$2\vec{a}_{+-}^{ij} + \vec{a}_{++}^{ij} + \vec{a}_{--}^{ij} = \eta(\hat{n}_i - \hat{n}_j), \quad (\text{C11})$$

$$2\vec{a}_{-+}^{ij} + \vec{a}_{++}^{ij} + \vec{a}_{--}^{ij} = \eta(-\hat{n}_i + \hat{n}_j), \quad (\text{C12})$$

$$2\vec{a}_{--}^{ij} + \vec{a}_{+-}^{ij} + \vec{a}_{-+}^{ij} = \eta(-\hat{n}_i - \hat{n}_j). \quad (\text{C13})$$

From Eqs. (C8) and (C9) it follows that

$$(\vec{a}_{++}^{ij} + \vec{a}_{--}^{ij}) + (\vec{a}_{-+}^{ij} + \vec{a}_{+-}^{ij}) = 0.$$

So one can define

$$\vec{a}_{ij} \equiv \vec{a}_{+-}^{ij} + \vec{a}_{-+}^{ij} \Rightarrow \vec{a}_{++}^{ij} + \vec{a}_{--}^{ij} = -\vec{a}_{ij}.$$

Now, from Eqs. (C10) to (C13) the following are obvious:

$$\vec{a}_{++}^{ij} = \frac{1}{2} [\eta(\hat{n}_i + \hat{n}_j) - \vec{a}_{ij}], \quad (\text{C14})$$

$$\vec{a}_{+-}^{ij} = \frac{1}{2} [\eta(\hat{n}_i - \hat{n}_j) + \vec{a}_{ij}], \quad (\text{C15})$$

$$\vec{a}_{-+}^{ij} = \frac{1}{2} [\eta(-\hat{n}_i + \hat{n}_j) + \vec{a}_{ij}], \quad (\text{C16})$$

$$\vec{a}_{--}^{ij} = \frac{1}{2} [\eta(-\hat{n}_i - \hat{n}_j) - \vec{a}_{ij}]. \quad (\text{C17})$$

This gives the general form of the joint measurement POVMs in the main text. For qubit effects,  $G_{X_i X_j}^{ij}$ , where  $X_i, X_j \in \{+1, -1\}$ , the valid effect condition (C3) is equivalent to the following [21]:

$$|\vec{a}_{++}^{ij}| \leq \frac{\alpha_{ij}}{2} \leq 2 - |\vec{a}_{++}^{ij}|, \quad (\text{C18})$$

$$|\vec{a}_{+-}^{ij}| \leq 1 - \frac{\alpha_{ij}}{2} \leq 2 - |\vec{a}_{+-}^{ij}|, \quad (\text{C19})$$

$$|\vec{a}_{-+}^{ij}| \leq 1 - \frac{\alpha_{ij}}{2} \leq 2 - |\vec{a}_{-+}^{ij}|, \quad (\text{C20})$$

$$|\vec{a}_{--}^{ij}| \leq \frac{\alpha_{ij}}{2} \leq 2 - |\vec{a}_{--}^{ij}|. \quad (\text{C21})$$

These inequalities can be combined and rewritten as

$$2 \max\{|\vec{a}_{++}^{ij}|, |\vec{a}_{--}^{ij}|\} \leq \alpha_{ij} \leq 2 - 2 \max\{|\vec{a}_{+-}^{ij}|, |\vec{a}_{-+}^{ij}|\}, \quad (\text{C22})$$

where

$$\begin{aligned} & \max\{|\vec{a}_{++}^{ij}|, |\vec{a}_{--}^{ij}|\} \\ &= \sqrt{\frac{\eta^2}{2}(1 + \hat{n}_i \cdot \hat{n}_j) + \frac{|\vec{a}_{ij}|^2}{4} + \frac{\eta}{2} |(\hat{n}_i + \hat{n}_j) \cdot \vec{a}_{ij}|} \end{aligned}$$

and

$$\begin{aligned} & \max\{|\vec{a}_{+-}^{ij}|, |\vec{a}_{-+}^{ij}|\} \\ &= \sqrt{\frac{\eta^2}{2}(1 - \hat{n}_i \cdot \hat{n}_j) + \frac{|\vec{a}_{ij}|^2}{4} + \frac{\eta}{2} |(\hat{n}_i - \hat{n}_j) \cdot \vec{a}_{ij}|}. \end{aligned}$$

This is the condition for a valid joint measurement used in inequalities (13) and (14) in the main text.

#### APPENDIX D: OPTIMAL STATE-DEPENDENT VIOLATION FOR MEASUREMENTS IN A PLANE

We need to maximize  $C \equiv 2\eta - (\sum_{(ij)} \alpha_{ij} - |\vec{a}|)$  to obtain the optimal violation of the LSW inequality. Subject to satisfaction of the joint measurability constraints (13) and (14) in the main text, we have

$$\begin{aligned} C_{\max} &= \max_{\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}, \{\vec{a}_{ij}\}, \eta} \left\{ 2\eta + |\vec{a}| - \sum_{(ij)} \alpha_{ij} \right\} \\ &\leq \max_{\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}, \{\vec{a}_{ij}\}, \eta} \left\{ 2\eta + \sum_{(ij)} |\vec{a}_{ij}| \right. \\ &\quad \left. - \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2} \right\}. \quad (\text{D1}) \end{aligned}$$

The inequality above follows from the fact that

$$|\vec{a}| = \sqrt{\sum_{(ij)} |\vec{a}_{ij}|^2 + 2(\vec{a}_{12} \cdot \vec{a}_{13} + \vec{a}_{12} \cdot \vec{a}_{23} + \vec{a}_{13} \cdot \vec{a}_{23})}, \quad (\text{D2})$$

so that  $|\vec{a}| \leq \sum_{(ij)} |\vec{a}_{ij}|$ , and

$$\begin{aligned} & \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2} \\ &\leq \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2 + 2\eta |(\hat{n}_i + \hat{n}_j) \cdot \vec{a}_{ij}|} \\ &\leq \sum_{(ij)} \alpha_{ij}. \end{aligned}$$

Also, we have

$$\begin{aligned} & \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2} \\ &\geq \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2}_{\text{coplanar}, \phi_3=\pi}. \end{aligned}$$

That is, for a fixed  $|\vec{a}_{ij}|$ ,  $\sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2}$  is smallest when the measurement directions  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  are coplanar and  $\phi_3 = \pi$ . From Eqs. (27) to (29) in the main text,  $\hat{n}_2 \cdot \hat{n}_3 = \cos \theta_{23} = \sin \theta_{12} \sin \theta_{13} \cos \phi_3 + \cos \theta_{12} \cos \theta_{13}$ . When  $\phi_3 = 0$  or  $\pi$ , the three measurements are coplanar and there are only two free angles,  $\hat{n}_1 \cdot \hat{n}_2 = \cos \theta_{12}$  and  $\hat{n}_1 \cdot \hat{n}_3 = \cos \theta_{13}$ , while the third angle is fixed by these two:  $\hat{n}_2 \cdot \hat{n}_3 = \cos \theta_{23} = \cos(\theta_{12} - \theta_{13})$  or  $\cos(\theta_{12} + \theta_{13})$ . Since  $\cos(\theta_{12} + \theta_{13}) \leq \cos(\theta_{23}) \leq \cos(\theta_{12} - \theta_{13})$ , for any given  $\theta_{12}$  and  $\theta_{13} \in (0, \pi)$ ,  $\cos \theta_{23}$  is smallest when  $\phi_3 = \pi$ . Hence, we choose the three measurements to be coplanar such that  $\phi_3 = \pi$  and  $\cos \theta_{23} = \cos(\theta_{12} + \theta_{13})$ . Any other choice of  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  will give a larger value of  $\cos \theta_{23}$ , hence also  $\sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2}$ . So,

$$\begin{aligned} C_{\max} &\leq \max_{\hat{n}_1, \hat{n}_2, \hat{n}_1, \hat{n}_3, \{\vec{a}_{ij}\}, \eta} \left\{ 2\eta + \sum_{(ij)} |\vec{a}_{ij}| \right. \\ &\quad \left. - \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2}_{\text{coplanar}, \phi_3=\pi} \right\}. \end{aligned}$$

We will now argue that this inequality for  $C_{\max}$  can be replaced by an equality. Let us take coplanar measurement directions  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  such that  $\phi_3 = \pi$ . We also take all the  $\vec{a}_{ij}$  parallel to each other, i.e.,  $\vec{a}_{12} \cdot \vec{a}_{13} = |\vec{a}_{12}| |\vec{a}_{13}|$ ,  $\vec{a}_{12} \cdot \vec{a}_{23} = |\vec{a}_{12}| |\vec{a}_{23}|$ , and  $\vec{a}_{13} \cdot \vec{a}_{23} = |\vec{a}_{13}| |\vec{a}_{23}|$ , so that  $|\vec{a}| = |\vec{a}_{12}| + |\vec{a}_{13}| + |\vec{a}_{23}|$ . Besides,  $|(\hat{n}_i + \hat{n}_j) \cdot \vec{a}_{ij}| = 0 \forall (ij) \in \{(12), (13), (23)\}$ . From these conditions it follows that each  $\vec{a}_{ij}$  is perpendicular to the plane and  $\forall (ij) \in \{(12), (13), (23)\}$ ,  $\vec{a}_{ij} \cdot \hat{n}_i = \vec{a}_{ij} \cdot \hat{n}_j = 0$ . This allows us to choose  $\alpha_{ij} = \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2}$ . So, in our optimal configuration, the measurement directions are coplanar while the  $\vec{a}_{ij}$ 's are parallel to each other and perpendicular to the plane of the measurements. Note that this also means  $\vec{a}$  will be parallel to  $\vec{a}_{ij}$  and therefore perpendicular to the plane of the measurements, and so will be the optimal state (which is parallel to  $\vec{a}$ ). With these optimality conditions satisfied, the optimal violation can now be written as

$$C_{\max} = \max_{\hat{n}_1, \hat{n}_2, \hat{n}_1, \hat{n}_3, \{\vec{a}_{ij}\}, \eta} \left\{ 2\eta + \sum_{(ij)} |\vec{a}_{ij}| \right\} \quad (\text{D3})$$

$$- \sum_{(ij)} \sqrt{2\eta^2(1 + \hat{n}_i \cdot \hat{n}_j) + |\vec{a}_{ij}|^2} \Big\}. \quad (\text{D4})$$

The constraints from joint measurability (13) and (14) become

$$|\vec{a}_{ij}| \leq \sqrt{1 + \eta^4(\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2}. \quad (\text{D5})$$

Now,

$$C_{\max} \leq \max_{\hat{n}_1, \hat{n}_2, \hat{n}_3, \{|\vec{a}_{ij}\rangle\}, \eta} \left\{ 2\eta + \sum_{(ij)} (\sqrt{1 + \eta^4(\hat{n}_i \cdot \hat{n}_j)^2} - 2\eta^2) - (1 + \eta^2 \hat{n}_i \cdot \hat{n}_j) \right\}.$$

The upper bound follows from the fact that  $f(x, y) = x - \sqrt{x^2 + 2\eta^2(1+y)}$ , where  $0 \leq x \leq \sqrt{1 + \eta^4 y^2 - 2\eta^2}$  and  $-1 < y < 1$ , is an increasing function of  $x$  for a fixed  $y$ , i.e.,  $(\frac{\partial f}{\partial x})_y > 0$ . Here  $x \equiv |\vec{a}_{ij}|$  and  $y \equiv \hat{n}_i \cdot \hat{n}_j$ . So, taking  $|\vec{a}_{ij}| = \sqrt{1 + \eta^4(\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2}$ , we have

$$C_{\max}^{(\hat{n}_i), \eta} \equiv 2\eta + \sum_{(ij)} [\sqrt{1 + \eta^4(\hat{n}_i \cdot \hat{n}_j)^2} - 2\eta^2 - (1 + \eta^2 \hat{n}_i \cdot \hat{n}_j)]. \quad (\text{D6})$$

Note that  $\alpha_{ij} = 1 + \eta^2 \hat{n}_i \cdot \hat{n}_j$  for  $|\vec{a}_{ij}| = \sqrt{1 + \eta^4(\hat{n}_i \cdot \hat{n}_j)^2 - 2\eta^2}$ .  $C_{\max}^{(\hat{n}_i), \eta}$  is the maximum value of  $C$  for a given coplanar choice of measurement directions  $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$  and sharpness parameter  $\eta$ .

#### APPENDIX E: NOTE ON TRIPLEWISE INCOMPATIBILITY

We have worked with the constraint of triplewise incompatibility originally employed by LSW [19]. The intuition behind

this constraint was to ensure that the binary qubit POVMs we consider do not trivially admit a joint distribution, in which case there should be no contextuality even with respect to the KS inequality bound of  $2/3$ . However, the situation in respect of POVMs turns out to be richer than expected: it is possible to construct pairwise joint measurements of these qubit POVMs such that they violate the LSW inequality, and it is also possible for these qubit POVMs to be triplewise jointly measurable. It is just that any triplewise joint measurement that one may construct for these qubit POVMs will not marginalize to pairwise joint measurements capable of violating the LSW inequality (or even the KS inequality for this scenario). However, it may still be possible to construct pairwise joint measurements for the qubit POVMs that do not arise as marginals of any triplewise joint measurement and can therefore violate the LSW inequality.

On relaxing the requirement of triplewise incompatibility in our optimal (trine axes) scenario,  $\eta > \eta_l = \frac{2}{3}$ , we find that the maximum violation of the LSW inequality occurs at  $\eta = 0.4566$ , with  $R_3^O = 0.9374$  and  $R_3^{\text{LSW}} = 0.8478$ , the violation being 0.0896 or about 8.96%. This is a straightforward consequence of Eq. (33) in the main text. This result has been quoted in Theorem 3 of Ref. [23], where the peculiar behavior of POVMs with respect to joint measurability has also been discussed. We note, however, that the claim in Ref. [23] that “*LSW’s inequality can be regarded as a genuine KS inequality*” is not true, which should be clear from our discussion of the LSW inequality and its comparison to the corresponding KS inequality in Appendix A. For further related discussion see Ref. [24].

- 
- [1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
  - [2] J. S. Bell, *Physics* **1**, 195 (1964). Reprinted in *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, New York, 1987), Chap. 2.
  - [3] E. P. Specker, *Dialectica* **14**, 239 (1960) (English translation: M. P. Seevinck, [arXiv:1103.4537v3](https://arxiv.org/abs/1103.4537v3)).
  - [4] S. Kochen and E. P. Specker, *J. Math. Mech.* **17**, 59 (1967).
  - [5] R. Clifton, *Am. J. Phys.* **61**, 443 (1993).
  - [6] A. Peres, *J. Phys. A: Math. Gen.* **24**, L175 (1991).
  - [7] N. D. Mermin, *Rev. Mod. Phys.* **65**, 803 (1993).
  - [8] A. Cabello, J. M. Estebaranz, and G. García-Alcaine, *Phys. Lett. A* **212**, 183 (1996).
  - [9] A. Cabello and G. García-Alcaine, *J. Phys. A* **29**, 1025 (1996).
  - [10] A. A. Klyachko, M. A. Can, S. Biniciolgu, and A. S. Shumovsky, *Phys. Rev. Lett.* **101**, 020403 (2008).
  - [11] P. Kurzyński, R. Ramanathan, and D. Kaszlikowski, *Phys. Rev. Lett.* **109**, 020404 (2012).
  - [12] S. Yu and C. H. Oh, *Phys. Rev. Lett.* **108**, 030402 (2012).
  - [13] A. Cabello, [arXiv:1112.5149v2](https://arxiv.org/abs/1112.5149v2).
  - [14] A. Cabello, *Phys. Rev. Lett.* **90**, 190401 (2003).
  - [15] P. Busch, *Phys. Rev. Lett.* **91**, 120403 (2003).
  - [16] C. M. Caves, C. A. Fuchs, K. K. Manne, and J. M. Renes, *Found. Phys.* **34**, 193 (2004).
  - [17] R. W. Spekkens, [arXiv:1312.3667](https://arxiv.org/abs/1312.3667).
  - [18] R. W. Spekkens, *Phys. Rev. A* **71**, 052108 (2005).
  - [19] Y. C. Liang, R. W. Spekkens, and H. M. Wiseman, *Phys. Rep.* **506**, 1 (2011).
  - [20] S. Yu, N. L. Liu, L. Li, and C. H. Oh, *Phys. Rev. A* **81**, 062116 (2010).
  - [21] T. Heinosaari, D. Reitzner, and P. Stano, *Found. Phys.* **38**, 1133 (2008).
  - [22] R. Kunjwal, C. Heunen, and T. Fritz, [arXiv:1311.5948](https://arxiv.org/abs/1311.5948).
  - [23] S. Yu and C. H. Oh, [arXiv:1312.6470](https://arxiv.org/abs/1312.6470).
  - [24] R. Kunjwal, [arXiv:1403.0470](https://arxiv.org/abs/1403.0470).