Supplemental Material: Probing the limits of correlations in an indivisible quantum system

M. Malinowski, C. Zhang, F. M. Leupold, J. Alonso,* and J. P. Home†

Institute for Quantum Electronics, ETH Zürich, Otto-Stern-Weg 1, 8093 Zürich, Switzerland

A. Cabello

Departamento de Física Aplicada II, Universidad de Sevilla, 41012 Sevilla, Spain

Appendix A: Qutrit transitions

Qutrit states are encoded into Zeeman sub-levels of a $^{40}\mathrm{Ca}^+$ ion as follows:

$$
\begin{aligned}
|0\rangle &= \left|S_{1/2},\, m_J = -1/2\right\rangle, \\
|1\rangle &= \left|D_{5/2},\, m_J = -3/2\right\rangle, \\
|2\rangle &= \left|D_{5/2},\, m_J = -1/2\right\rangle.
\end{aligned}
$$

An external magnetic field of $|B| \approx 3.73\,\mathrm{G}$ splits the $|0\rangle \leftrightarrow |1\rangle$ and $|0\rangle \leftrightarrow |2\rangle$ transitions by $\approx 6.27\,\mathrm{MHz}$. These transitions are driven by linearly polarized laser pulses at $\lambda \approx 729\,\mathrm{nm}$ propagating at an angle of 45° to the quantization axis and adjusted to be resonant with the transition of choice. We operate at low laser intensities to keep AC Stark shifts below 100 Hz. This allows us to, up to a good approximation, treat the laser pulses as inducing single-qubit rotations in a qutrit space:

$$
R^{(1)}(\theta,\phi) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\mathrm{i}e^{-\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right) & 0 \\ -\mathrm{i}e^{\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{A1a}
$$

$$
R^{(2)}(\theta,\phi) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & 0 & -\mathrm{i}e^{-\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right) \\ 0 & 1 & 0 \\ -\mathrm{i}e^{\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right) & 0 & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}. \tag{A1b}
$$

Appendix B: Qutrit coherence times

Using Ramsey techniques, we determine coherence times of $\sigma_t \approx 1.6\,\mathrm{ms}$ for the $|0\rangle \leftrightarrow |1\rangle$ and $|0\rangle \leftrightarrow |2\rangle$ transitions, and $\sigma_t \approx 7\,\mathrm{ms}$ for the $|1\rangle \leftrightarrow |2\rangle$ transition. Here we assumed for simplicity that the noise is Gaussian and slow compared to the timescale of a single experiment, causing a Ramsey experiment with wait time τ to lose contrast as $C = e^{-\tau^2/(2\sigma_t^2)}$ [1]. This reveals a noise component of $\approx$ 230 Hz (Full-Width Half-Maximum, FWHM) common to both the $|0\rangle \leftrightarrow |1\rangle$ and $|0\rangle \leftrightarrow |2\rangle$ transitions, which we believe originates due to Doppler shifts transmitted onto the ion from vibrations of the closed-cycle cryocooler. The differential noise has a width of $\approx$ 50 Hz (FWHM), which we associate with B-field fluctuations of $<5\,\mu\mathrm{G}$ (FWHM) and slow drifts. The latter result in changes of transition frequencies ($\sim$ 100 Hz on time scales of minutes), which we automatically re-calibrate every 30 s with 10 Hz resolution.

* alonso@phys.ethz.ch

† jhome@phys.ethz.ch

Appendix C: Ion cooling

At the start of every experiment the ion is optically pumped to $|0\rangle$ and cooled close to the motional ground state. The cooling is done in two steps. First, a Doppler-cooling sequence brings all three oscillation modes (one parallel, and two perpendicular to the trap axis) to a thermal state with mean phonon occupation $n_{\mathrm{th}} \approx 5$. Then we further cool the mode of motion parallel to the trap axis to $n_{\mathrm{th}} \approx 0.2$ using electromagnetically-induced transparency (EIT) cooling [2]. Qutrit transitions are driven with a laser beam propagating along the trap axis, deeming the influence of radial motion negligible.

The second instance of cooling occurs during state detection. Standard on-resonance fluorescence detection techniques cause motional heating due to photon recoil in case of a bright detection. To circumvent this problem, detection is conducted off-resonantly with settings close to those of Doppler cooling [3]. While this decreases the observed signal and hence lengthens detection pulses, it also brings the temperature of the bright ion close to the Doppler limit. Afterwards, we repeat the EIT cooling sequence described above to get back to $n_{\mathrm{th}} \approx 0.2$. In case of a dark detection, the state of the ion is unaffected by the detection or EIT beams. This leads to negligible dependence of the ion's motional state on the first measurement, since a dark ion is subject to motional heating at a rate of $\sim$ 200 quanta/s.

Appendix D: Qutrit detection

Fluorescence detection is implemented by shining a laser beam at 397 nm, which connects the $S_{1/2}$ states with a short-lived $P_{1/2}$ level, together with a repumping beam at 866 nm (Fig. 2, [3]). If the ion fluoresces ("bright detection") it is projected onto the $S_{1/2}$ manifold, while no fluorescence ("dark detection") projects it onto the $D_{5/2}$ manifold.

During a dark detection, the lack of physical interaction between ion and laser ensures that the initial qutrit state ρ_{in} is related to the post-measurement state $\rho_{\rm out}$ via

$$\rho_{\rm out} = \frac{P_{\rm D}\rho_{\rm in}P_{\rm D}}{{\rm tr}(P_{\rm D}\rho_{\rm in})}, \tag{D1}$$

where $P_{\rm D} = (|1\rangle\langle 1| + |2\rangle\langle 2|)$ is the projection operator onto the dark states. In other words, a dark detection implements a Lüders projective measurement [4]. A bright detection, on the other hand, projects the ion onto a mixture of $\{|S_{1/2}, m_J = -1/2\rangle, |S_{1/2}, m_J = +1/2\rangle\}$ states. We therefore complete the detection sequence with a spin-polarizing (SP) σ-polarized pulse at 397 nm, which pumps all the $S_{1/2}$ population into $|0\rangle = |S_{1/2}, m_J = -1/2\rangle$. This makes a bright measurement a Lüders measurement with projection operator $P_{\rm B} = |0\rangle\langle 0|$.

Bright states are discriminated from dark states by thresholding the number of photon counts collected within a detection window. In a typical experiment, we collect an average of ≈ 25 photons from a bright ion, with an average background of ≈ 1 photon, during a window of $\approx 200\,\mu$s. The optimal threshold is set close to the crossing point between the histograms resulting from bright and dark states [3]. We estimate a bright (dark) detection error of $\approx 2\times 10^{-5}$ ($\approx 1\times 10^{-4}$). Dark-detection errors are dominated by spontaneous decay of $D_{5/2}$ states into the $S_{1/2}$ ground state ($\tau_{\rm decay} \approx 1.2$ s).

Appendix E: Data collection and analysis

Consider an experiment with N observables (for KCBS, $N = 5$). For each opening angle setting $\theta_{\rm set}$ we perform a correlation measurement (Fig. 2) along a pair of directions $(|\psi_i\rangle, |\psi_j\rangle)$, with results $(A_i^{(1)} = \pm 1,\ A_j^{(2)} = \pm 1)$. Results +1 correspond to a bright ion and results -1 correspond to a dark ion, and the superscript in brackets denotes whether the observable is measured first or second. We also calculate the correlation $A_i^{(1)}A_j^{(2)}$. Each experiment is repeated 10,000 times, and the average results $(\langle A_i^{(1)}\rangle, \langle A_j^{(2)}\rangle, \langle A_i^{(1)}A_j^{(2)}\rangle)$ are extracted. We collect data for N observable pairs in one of two possible orders:

$$\text{normal order:}\quad (i,j) = (i, i+1),\quad i = 1,\ldots,N, \tag{E1}$$

$$\text{reverse order:}\quad (i,j) = (i, i-1),\quad i = 1,\ldots,N. \tag{E2}$$

We then evaluate the witness S_N starting from Eq. (8):

$$\text{normal order:}\quad S_N = \sum_{i=1}^{N}\langle A_i^{(1)}A_{i+1}^{(2)}\rangle, \tag{E3}$$

$$\text{reverse order:}\quad S_N = \sum_{i=1}^{N}\langle A_{i+1}^{(1)}A_i^{(2)}\rangle. \tag{E4}$$

We also estimate the pentagon opening angle θ using

$$\theta = \frac{1}{2}\arccos\left(\sum_{i=1}^{N}\frac{\langle A_i^{(1)}\rangle}{N}\right). \tag{E5}$$

The incompatibility term ϵ is evaluated according to

$$\epsilon = \sum_{i=1}^{N}\epsilon_i = \sum_{i=1}^{N}|\langle A_i^{(1)}\rangle - \langle A_i^{(2)}\rangle|, \tag{E6}$$

and the extended witness is given by $S_N^{(\rm ext)} = S_N + \epsilon$ (Eq. (4)). For each average $(\langle A_i^{(1)}\rangle, \langle A_j^{(2)}\rangle, \langle A_i^{(1)}A_j^{(2)}\rangle)$ we evaluate the sample standard deviation and use it to compute the standard error in the mean. We then propagate the standard errors of S_N, θ, ϵ and $S_N^{(\rm ext)}$ assuming independent errors. These standard errors are plotted as error bars in Figs. 3 and 4. We note that ϵ is defined as necessarily positive and its distribution is non-gaussian, hence the standard error of ϵ cannot always be treated as a confidence interval.

Appendix F: Measured observables

Consider operators $P_{\rm B} = |0\rangle\langle 0|$ and $P_{\rm D} = |1\rangle\langle 1| + |2\rangle\langle 2|$, which project the ion onto a bright and dark state respectively. We define the observable M_i of measurement along $|\psi_i\rangle = U_i|0\rangle$ as $M_i = P_{{\rm B},i} - P_{{\rm D},i}$, where

$$P_{{\rm B},i} = U_i P_{\rm B} U_i^\dagger = |\psi_i\rangle\langle\psi_i|, \tag{F1}$$

$$P_{{\rm D},i} = U_i P_{\rm D} U_i^\dagger. \tag{F2}$$

The outcome of measurement M_i is denoted as $A_i = \pm 1$. With these definitions, the outcome $A_i = +1$ corresponds to projection onto $P_{{\rm B},i}$, while $A_i = -1$ corresponds to projection onto $P_{{\rm D},i}$. Note that $[M_i, M_{i+1}] = 0$ when $\langle\psi_i|\psi_{i\pm1}\rangle = 0$ and that in QM two observables are compatible when their operators commute. If we consider states $|\psi_i\rangle$ on a real "qutrit sphere" (Fig. 1), this condition is equivalent to orthogonality between vectors.

Appendix G: Theoretical predictions for KCBS witnesses

Consider a sequence of measurements M_i, $M_{i\pm1}$ as defined in App. F, with U_i defined as in Eq. (1) and the initial state $\rho_{\rm in} = |0\rangle\langle 0|$. Their correlation can be evaluated to:

$$\begin{aligned}\langle A_i^{(1)}A_{\pm1}^{(2)}\rangle &= {\rm tr}\big(M_i M_{i\pm1}\rho_{\rm in}\big)\\ &= \frac{1}{8}\big(3 - \sqrt{5} + (5+\sqrt{5})\cos(4\theta)\big).\end{aligned} \tag{G1}$$

Then the KCBS witness can be simply calculated as $S_5(\theta) = 5\langle A_i^{(1)}A_{i\pm1}^{(2)}\rangle$ (Eq. (2)). This is plotted in solid blue in Fig. 1. Note in particular that the minimum value of $S_5(\theta)$ is obtained when $\theta = \frac{\pi}{2}$, where $S_5 = \frac{5}{4}(-\sqrt{5} - 1) \approx -4.045$. At the point of compatibility, where $[M_i, M_{i+1}] = 0$, which corresponds to $\theta = \theta_5 = \arccos\big(5^{-1/4}\big)$, we obtain $S_5 = 5 - 4\sqrt{5} \approx -3.944$.

In order to evaluate the extended KCBS witness $S_5^{(\rm ext)}(\theta)$ we first evaluate the expectation value of A_i as

$$\langle A_i^{(1)}\rangle = {\rm tr}(M_i\rho_{\rm in}) = \cos(2\theta), \tag{G2}$$

while the post-measurement state is given by

$$\rho_i = P_{\mathrm{B},i}\rho_{\mathrm{in}}P_{\mathrm{B},i} + P_{\mathrm{D},i}\rho_{\mathrm{in}}P_{\mathrm{D},i}. \tag{G3}$$

The average outcome of measurement $M_{i\pm1}$ is then given by

$$\langle A^{(2)}_{i\pm1}\rangle = \mathrm{tr}(M_{i\pm1}\rho_i). \tag{G4}$$

Due to the symmetry of the problem, $\langle A^{(2)}_{i+1}\rangle$ is expected to be the same for all i, and hence$\langle A^{(2)}_{i+1}\rangle = \langle A^{(2)}_{i}\rangle$. We use this to evaluate the extension term:

$$\begin{aligned}\epsilon_i &= \left|\langle A^{(1)}_i\rangle - \langle A^{(2)}_i\rangle\right| = \left|\langle A^{(1)}_i\rangle - \langle A^{(2)}_{i+1}\rangle\right| \\ &= \frac{1}{16}\left|(5-\sqrt{5}+5(3+\sqrt{5})\cos(2\theta))\sin^2(2\theta)\right|. \end{aligned} \tag{G5}$$

The extended KCBS witness $S^{(\mathrm{ext})}_5 = S_5 + \sum_{i=1}^{5}\epsilon_i$ is plotted in Fig. 1 in solid red. Note that at the point of compatibility $\epsilon_i = 0$ and $S^{(\mathrm{ext})}_5 = S_5 = 5-4\sqrt{5}$, while for all other values of θ we have $S^{(\mathrm{ext})}_5 > \bar{S}^{\mathrm{QM}}_5$.

Aside from penalizing systematic effects, ϵ_i also accounts for finite sample size. This is because the sample mean of ϵ_i is a biased estimator of the population mean of ϵ_i. Consider an experimental run with n measurement repetitions. The measured outcomes follow a normal distribution $Y_{\langle A_i\rangle} \propto \mathcal{N}(\langle A_i\rangle, \sigma^2_{A_i})$, where $\mathcal{N}(\mu,\sigma)$ denotes a normal distribution with mean μ and variance σ^2. Then $Y_{\langle A_i\rangle-\langle A_{i+1}\rangle} \propto \mathcal{N}(\langle A_i\rangle - \langle A_{i+1}\rangle, \sigma^2_{A_i} + \sigma^2_{A_{i+1}})$. Finally, $Y_{\epsilon_i} = \mathcal{F}(\langle A_i\rangle - \langle A_{i+1}\rangle, \sigma^2_{A_i} + \sigma^2_{A_{i+1}})$, where $\mathcal{F}(\mu,\sigma^2)$ is the so-called "folded normal distribution", obtained by taking the absolute value of a gaussian with mean μ and variance σ^2. The mean of a folded normal distribution is [5]

$$\mu_{\mathrm{F}}(\mu,\sigma) = \sigma\sqrt{\frac{2}{\pi}}\exp\left(\frac{-\mu^2}{2\sigma^2}\right) - \mu\,\mathrm{erf}\left(\frac{-\mu}{\sqrt{2\sigma^2}}\right), \tag{G6}$$

where erf is the error function. This mean represents an expectation value $\mathrm{E}(\epsilon_i)$ of ϵ_i in an experiment with n repetitions:

$$\mathrm{E}(\epsilon_i) = \mu_{\mathrm{F}}\left(\epsilon_i, \sigma^2_{A_i} + \sigma^2_{A_{i+1}}\right), \tag{G7}$$

$$\sigma^2_{A_i} = \frac{1-\langle A_i\rangle}{n}. \tag{G8}$$

The expression for the variance $\sigma^2_{A_i}$ represents the shot-noise contribution, hence we refer to this expectation value as "theoretical prediction with shot noise". $\mathrm{E}(S^{(\mathrm{ext})}_5) = S_5 + \sum_{i=1}^{5}\mathrm{E}(\epsilon_i)$ for $n = 10{,}000$ is plotted in Fig. 1 in dashed red. We see that the effect of shot noise is significant when $\epsilon_i \approx 0$, where it predicts a gap of size $\sqrt{2(1-\cos(2\theta_5))/(\pi n)}$ between the sample mean and the population mean of $S^{(\mathrm{ext})}_5$.

Appendix H: Relevance of KCBS and odd cycle NC inequalities

There are three different perspectives from which the KCBS and the N odd-cycle non-contextuality inequalities (with $N \geq 5$; the case $= 5$ is the KCBS inequality) are of fundamental importance for understanding the power of quantum systems. Each of these perspectives corresponds to a different approach for investigating the quantum vs classical advantage.

The first perspective focuses on *the quantum system that produces the quantum advantage*. In this respect, the KCBS and the other odd $N-$cycle NC inequalities are special because all of them achieve their maximal quantum violation using a *qutrit*, which is the simplest quantum system that produces contextuality [6, 7]. This is in contrast with the case of the CHSH and the rest of even $N-$cycle NC inequalities (called chained Bell inequalities when tested on pairs of systems [8, 9]), whose maximal quantum violation requires either a ququart, i.e. a four-dimensional quantum system, or a pair of qubits. Moreover, the KCBS inequality is the simplest (i.e. the one requiring the smallest number of observables) non-contextuality inequality violated by a qutrit.

The second perspective emphasizes *the scenario in which the quantum advantage occurs*. A scenario is defined by a set of observables, each with a certain number of possible outcomes, having certain relations of compatibility among them. These relations are typically represented by a graph called the *compatibility graph*, in which vertices represent observables and edges connect compatible observables. The name N-cycle inequality refers to the fact that the corresponding scenario has N observables whose compatibility relations are represented by a cycle of N vertices. For each scenario, the correlations between compatible measurements define a set whose points are vectors of probabilities. Geometrically, the classical set is a polytope whose facets define tight inequalities that are necessary and sufficient conditions for the existence of a classical model. For certain scenarios, the quantum set is larger than the classical set. However, only a few scenarios have been exhaustively explored [10–14] in the sense that we know all the classical inequalities and all the maximal quantum violations. Among these scenarios, the $N \geq 4$ cycle non-contextuality scenario (where there are N observables, each with 2 outcomes, such that observable j is compatible with observable $j+1$ with the sum modulo N) is the only one which is symmetric, i.e. with all tight inequalities and all maximal quantum violations being of the same type [13]. These inequalities are precisely the $N \geq 4$ cycle non-contextuality inequalities. As non-contextuality inequalities, all of them are tight. However, the case N even can be tested with pairs of particles and then the corresponding non-contextuality inequalities are also Bell inequalities, the so-called chained Bell inequalities [8, 9]. Notice, however, that the chained Bell inequalities are *not* tight Bell inequalities (i.e. they do not correspond to facets of the *local* polytope) [13].

The third perspective focuses on *the graph of exclusivity responsible for the quantum advantage*. In the graph-theoretic approach to quantum theory introduced in [15], every classical inequality is first converted into an inequality in which the quantity bounded is a sum of probabilities of events (i.e. state transformations produced by an ideal measurement), and then associated to a graph in which vertices represent the events in this quantity and edges connect events that are exclusive (i.e. that correspond to different outcomes of the same ideal measurement). This graph is called the exclusivity graph of the inequality. The exclusivity graph of (the events of) an inequality should not be confused with the compatibility graph of (the observables) of an scenario.

The first result of the graph-theoretic approach is that an exclusivity graph admits a quantum realization whose probabilities cannot be explained classically *if and only if* the graph has, as induced subgraphs, odd M-cycles with $M \geq 5$ or their complements [15]. Relevant to this study is the fact that the exclusivity graph of the M events needed to test the odd N-cycle NC inequalities is precisely an M-cycle with $M = N$.

The second result of the graph-theoretic approach is that, for any exclusivity graph, the maximum value allowed by QM is given by a characteristic number of the graph. The odd $N \geq 5$-cycle NC inequalities happen to saturate each of these maxima [15] and hence saturate also the strength of correlations allowed by QM, which is given by Eq. (9):

$$\bar{S}_M^{\text{QM}} = \frac{M - 3M\cos\left(\frac{\pi}{M}\right)}{1 + \cos\left(\frac{\pi}{M}\right)}. \tag{H1}$$

Together, these two results imply that, among all possible non-contextuality inequalities, the odd $N \geq 5$-cycle NC inequalities have fundamental importance for understanding the differences between quantum and classical resources. Moreover, the exclusivity graph with the smallest number of vertices featuring genuinely quantum probabilities is the pentagon [15], which is exactly the exclusivity graph of the five events needed to test the KCBS inequality.

Although the chained Bell inequalities (and other Bell inequalities) contain events associated with N-cycle exclusivity graphs, it has been proven [16, 17] that no Bell inequality contains events allowing the maximum quantum value for any N-cycle exclusivity graph. It has also been proven [17] that the quantum maximum of the odd $N \geq 5$-cycle exclusivity graph for Bell scenarios is achieved for the chained Bell inequalities for $(N-1)/2$ observables per party and is given by

$$\bar{S}_M^{\text{Bell}} = M - 4\left[\frac{1}{2} + \frac{M-1}{4}\left(1 + \cos\left(\frac{\pi}{M-1}\right)\right)\right]. \tag{H2}$$

For example, for the pentagon, the quantum maximum within Bell scenarios is achieved in the scenario with two dichotomic observables per party, i.e. in the CHSH scenario [18]. A direct comparison between the QM expectations for the CHSH and KCBS scenarios shows that correlations between observables are stronger in the latter: $\bar{S}_5^{\text{Bell}} \approx -3.828 > \bar{S}_5^{\text{KCBS}} \approx -3.944$. Although the quantum realization for the CHSH scenario takes place in a Hilbert space of dimension four while for the KCBS scenario takes place in dimension three, measurements in the former are restricted to be bipartite, which accounts for the weaker correlations.

Finally, the third important result of the graph-theoretic approach is that, for a given graph of exclusivity, the set of classical probabilities is the so-called stable set polytope of the graph, while the set of quantum probabilities is the so-called theta body of the graph [15]. Among all the basic graphs needed for genuinely quantum correlations, the pentagon is the one in which the difference between the quantum and the classical set is the biggest one [15]. This difference is precisely the one aimed by the initial state and the set of measurements used to test the violation of the KCBS inequality.

TABLE I. Experimental results for N-cycle witness measurements. The contextual fraction is calculated according to Eq. (10). Numbers in brackets are the standard errors in the means, calculated as described in App. E, with each pair of observables measured 10,000 times.

N	S_N	$S_N^{(\text{ext})}$	$\bar{S}_N^{\text{NC}}$	$\bar{S}_5^{\text{QM}}$	CF_N
5	-3.926(14)	-3.905(34)	-3	-3.944	0.463(7)
7	-6.208(12)	-6.124(39)	-6	-6.271	0.604(6)
11	-10.452(10)	-10.304(48)	-9	-10.545	0.726(5)
17	-16.538(10)	-16.538(59)	-15	-16.708	0.769(5)
23	-22.530(10)	-22.138(69)	-21	-22.785	0.765(5)
31	-30.599(9)	-30.172(79)	-29	-30.840	0.800(4)
41	-40.439(11)	-39.983(91)	-39	-40.879	0.719(5)
51	-50.422(11)	-49.740(102)	-49	-50.903	0.711(5)
61	-60.279(11)	-59.494(111)	-59	-60.919	0.640(6)
81	-79.972(14)	-79.058(128)	-79	-80.939	0.486(7)
101	-99.544(17)	-98.437(143)	-99	-100.951	0.272(8)
121	-117.686(25)	-116.443(157)	-119	-120.959	-0.657(12)

Appendix I: N-cycle NC inequalities and results

Our experimental results are summarized in Tab. I. In order to penalize for incompatibility we again follow the scheme introduced in Ref. [19]. Accordingly, the witness for an extended N-cycle NC inequality is given by

$$S_N^{(\text{ext})} = \sum_{i=1}^{N} \langle A_i^{(1)} A_{i+1}^{(2)} \rangle + \sum_{i=1}^{N} \epsilon_i \geq \bar{S}_N^{\text{NC}} = -N + 2, \tag{I1}$$

where $\epsilon_i = \left| \langle A_i^{(1)} \rangle - \langle A_i^{(2)} \rangle \right|$, as in Eq. (4). Like $S_5^{(\text{ext})}$, $S_N^{(\text{ext})}$ reduces to S_N when $\theta = \theta_N$, and miscalibration of θ always results in $S_N^{(\text{ext})} > \bar{S}_N^{\text{QM}}$.

Appendix J: Comparison with previous KCBS tests

We provide in Tab. II a collection of experimental KCBS tests to benchmark our data. Aside from being among the closest to QM predictions, our analysis is the only one that we are aware of that systematically characterizes signaling and is in agreement with theoretical expectations. We have calculated the signaling the different experiments would have incurred wherever the data was made available by the authors (second-to-last column). The meanings of the different comments in the last column are:

- *Detection loophole*. Experiments with photons suffer from significant detector inefficiencies and setup losses. Consequently, the results assume that the registered events form an unbiased sample of all the events. The loophole does not apply when measurements are conducted with high fidelity.

TABLE II. Experimental results of previous KCBS tests. Comments in the last column are discussed in the text. The results marked by an asterisk represent our own analysis of the results table from [20]. These differ significantly from the final results quoted in [20], which are currently revised due to errors in the original data analysis [21].

Reference	Platform	Saturation of QM limit $(S_5 - \bar{S}_5^{\mathrm{NC}})/(\bar{S}_5^{\mathrm{QM}} - \bar{S}_5^{\mathrm{NC}})$	Signaling $\sum_{i=1}^{5} \epsilon_i/(\bar{S}_5^{\mathrm{QM}} - \bar{S}_5^{\mathrm{NC}})$	Comments
Vienna, 2011 [22]	Photons	0.947(6)	0.08(3)	Detection loophole Simultaneous measurements Six observables
Stockholm, 2013 [23]	Photons	0.53(11) (normal order) 0.95(11) (reverse order)	No data	Detection loophole Order dependence
Beijing, 2013 [24]	Photons	0.977(11) 0.956(26)	0.267 0.291	Detection loophole Simultaneous measurements Large signaling
Beijing, 2013 [20]	Yb ion	0.589(24)*	0.119(24)*	Six observables Non-projective measurements
Brisbane, 2016 [25]	Superconducting circuits	0.520(1) (normal order) 0.541(1) (reverse order)	0.379(2) 0.379(2)	Large signaling
This work	Ca ion	0.969(14) (normal order) 0.992(14) (reverse order)	0.054(31) 0.050(31)	

- *Simultaneous measurements*. When pairs of measurements are conducted simultaneously, it is difficult to establish an operational definition of an individual measurement [25]. It has been argued that this so-called "individual existence loophole" can be addressed by performing measurements in a sequence [26].

- *Six observables*. Certain experimental arrangements may not allow for performing the same measurement in the same way in every context. In order to circumvent this problem, the KCBS inequality is extended to a six-observable inequality. However, the validity of this approach has been challenged by some authors [23, 27].

- *Order dependence*. When statistical errors dominate, the order of measurement of correlators should not influence the final outcome. Significant dependence on the order indicates the presence of uncontrolled systematic errors.

- *Large signaling*. When signaling is large compared to observed violations, the experiment is far from the ideal assumption of compatible sharp measurements. This makes is difficult to relate the strength of contextual correlations measured in such experiments to established theoretical results. An indicator of this effect is that the statistical uncertainty in the measured witness (third column) is much smaller than the signaling term (fourth column).

- *Non-projective measurements*. In Ref. [20] measurements are conducted in a sequence, but are not projective by design. Specifically, the post-measurement state, conditioned on photon detection, is a mixture of qutrit basis states. This precludes the implementation of sharp measurements.

Appendix K: Comparison with previous CHSH and chained Bell tests

As discussed in App. H, both the CHSH scenario for a Bell test and the KCBS scenario for a non-contextuality test feature pentagonal graphs in the exclusivity formalism. To the best of our knowledge, there are two previous experiments aiming at the limits of correlations in Bell scenarios: one carried out by the Kwiat group in Illinois [28], the other in the group of Kurtsiefer in Singapore [29]. Both were carried out with photons. The former came close to 99% of the QM prediction for the CHSH test (with statistical uncertainties well below 1%), and the latter came down to 99.97(2)% of the Tsirelson bound. Our result for KCBS brings us to 99.5(2)% of the QM limit and is the first experimental demonstration of stronger-than-Bell correlations in a non-contextuality test closing the detection loophole (Fig. 3).

In Ref. [28] they furthermore measured chained Bell inequalities with even numbers of observables all the way to $N = 90$ and found a maximal value of the contextual fraction $CF_{36} = 0.874(1)$. Our results complement these with odd numbers of observables up to $N = 121$ and we determine a maximal value of $CF_{31} = 0.800(4)$. Aside from closing the detection loophole, we have thoroughly characterized our signaling (or incompatibility) by scanning the relevant experimental parameters. This has allowed us to identify optimal working conditions and place bounds on the amount of signaling present in the experiment. This could be relevant if the presence of signaling indicated by preliminary data analysis of previous photon experiments is confirmed [30].

Appendix L: Data repository

The complete raw dataset is publicly available from an open repository on `http://www.tiqi.ethz.ch/`

publications-and-awards/public-datasets.html. We encourage readers who want to expand our work with further data analysis or representations to do so.

[1] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, J. Res. Natl. Inst. Stand. Technol. **103**, 259 (1998).
[2] J. Alonso, F. M. Leupold, Z. U. Solèr, M. Fadel, M. Marinelli, B. C. Keitch, V. Negnevitsky, and J. P. Home, Nature Communications **7**, 11243 (2016).
[3] F. M. Leupold, M. Malinowski, C. Zhang, A. Cabello, J. Alonso, and J. P. Home, arXiv:1706.07370 (2017).
[4] G. Lüders, Annalen der Physik **443**, 322 (1950).
[5] F. C. Leone, L. S. Nelson, and R. B. Nottingham, Technometrics **3**, 543 (1961).
[6] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[7] S. Kochen and E. P. Specker, Journal of Mathematics and Mechanics **17**, 59 (1967).
[8] P. M. Pearle, Phys. Rev. D **2**, 1418 (1970).
[9] S. L. Braunstein and C. M. Caves, Annals of Physics **202**, 22 (1990).
[10] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[11] C. Sliwa, Physics Letters A **317**, 165 (2003).
[12] D. Collins and N. Gisin, Journal of Physics A: Mathematical and General **37**, 1775 (2004).
[13] M. Araújo, M. T. Quintino, C. Budroni, M. T. Cunha, and A. Cabello, Phys. Rev. A **88**, 022118 (2013).
[14] S. López-Rosa, Z.-P. Xu, and A. Cabello, Phys. Rev. A **94**, 062121 (2016).
[15] A. Cabello, S. Severini, and A. Winter, Physical Review Letters **112**, 040401 (2014).
[16] M. Sadiq, P. Badziąg, M. Bourennane, and A. Cabello, Physical Review A **87**, 012128 (2013).
[17] R. Rabelo, M. T. Cunha, and A. Cabello, "In preparation," (2017).
[18] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Physical Review Letters **23**, 880 (1969).
[19] J. V. Kujala, E. N. Dzhafarov, and J.-A. Larsson, Phys. Rev. Lett. **115**, 150401 (2015).
[20] M. Um, X. Zhang, J. Zhang, Y. Wang, S. Yangchao, D.-L. Deng, L.-M. Duan, and K. Kim, Scientific Reports **3** (2013).
[21] K. Kim, "Private communication," (2017).
[22] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wieśniak, and A. Zeilinger, Nature **474**, 490 (2011).
[23] J. Ahrens, E. Amselem, A. Cabello, and M. Bourennane, Scientific Reports **3** (2013).
[24] D.-L. Deng, C. Zu, X.-Y. Chang, P.-Y. Hou, H.-X. Yang, Y.-X. Wang, and L.-M. Duan, arXiv (2013), 1301.5364.
[25] M. Jerger, Y. Reshitnyk, M. Oppliger, A. Potočnik, M. Mondal, A. Wallraff, K. Goodenough, S. Wehner, K. Juliusson, N. K. Langford, *et al.*, Nature communications **7** (2016).
[26] A. Cabello, in *FOUNDATIONS OF PROBABILITY AND PHYSICS 5*, Vol. 1101 (AIP Publishing, 2009) pp. 246–254.
[27] R. Lapkiewicz, P. Li, C. Schaeff, N. Langford, S. Ramelow, M. Wiesniak, and A. Zeilinger, arXiv (2013), 1305.5529.
[28] B. G. Christensen, Y.-C. Liang, N. Brunner, N. Gisin, and P. G. Kwiat, Phys. Rev. X **5**, 041052 (2015).
[29] H. S. Poh, S. K. Joshi, A. Cerè, A. Cabello, and C. Kurtsiefer, Physical Review Letters **115**, 180408 (2015).
[30] M. Smania, M. Kleinmann, A. Cabello, and M. Bourennane, arXiv (2018), 1801.05739.