Better local hidden variable models for two-qubit Werner states and an upper bound on the Grothendieck constant $K_G(3)$

Flavien Hirsch[1], Marco Túlio Quintino[1][2], Tamás Vértesi[3], Miguel Navascués[4], and Nicolas Brunner[1]

[1]Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

[2]Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

[3]Institute for Nuclear Research, Hungarian Academy of Sciences, H-4001 Debrecen, P.O. Box 51, Hungary

[4]Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmangasse 3, 1090 Vienna, Austria

17th April 2017

We consider the problem of reproducing the correlations obtained by arbitrary local projective measurements on the two-qubit Werner state $\rho = v\,|\psi_-\rangle\langle\psi_-| + (1-v)\mathbb{1}/4$ via a local hidden variable (LHV) model, where $|\psi_-\rangle$ denotes the singlet state. We show analytically that these correlations are local for $v = 999 \times 689 \times 10^{-6} \cos^4(\pi/50) \simeq 0.6829$. In turn, as this problem is closely related to a purely mathematical one formulated by Grothendieck, our result implies a new bound on the Grothendieck constant $K_G(3) \leq 1/v \simeq 1.4644$. We also present a LHV model for reproducing the statistics of arbitrary POVMs on the Werner state for $v \simeq 0.4553$. The techniques we develop can be adapted to construct LHV models for other entangled states, as well as bounding other Grothendieck constants.

A quantum Bell experiment consists of two (or more) distant observers performing local measurements on a shared entangled quantum state. Remarkably the predictions of quantum theory are here incompatible with a natural definition of locality formulated by Bell [1]. Specifically, the statistics of certain quantum Bell experiments are found to be nonlocal (in the sense of Bell), as witnessed via violation of Bell inequalities. This phenomenon, referred to as quantum nonlocality, represents a fundamental aspect of quantum theory as well as a central resource for quantum information processing [2].

Understanding the exact relation between entanglement and quantum nonlocality is a central problem in the foundations of quantum theory, with implications for quantum information processing. While the use of an entangled state is necessary for observing quantum nonlocal correlations, it is interesting to ask if the converse link also holds. That is, can any entangled state lead to a Bell inequality violation, when performing a set of (judiciously chosen, and possibly infinitely many) local measurements? For pure entangled states, the answer turns out to be positive [3]. For mixed entangled states, the situation is more complex, as first discovered by Werner [4], who presented a class of entangled quantum states (now referred to as Werner states) which admit a local hidden variable (LHV) model for any possible local projective measurements. Therefore such states, while being entangled—i.e. inseparable at the level of the Hilbert space—can never lead to nonlocal correlations. Notably, while Werner's original model focused on projective measurements, Barrett [5] presented a LHV model considering the most general non-sequential measurements, i.e. POVMs. These early results triggered much interest, and subsequent works presented various classes of entangled states admitting LHV models [6–13], including results for the multipartite case [14–16]; see [17] for a recent review. More sophisticated Bell scenarios have also been ex-

plored[1], but will not be discussed here.

In parallel, Acín, Gisin and Toner [18], based on previous work by Tsirelson [19], established a direct connection between these questions and a purely mathematical problem discussed by Grothendieck [20]. In particular, the problem of determining the range of visibilities $v \leq v_c$ (v_c denoting the critical visibility) for which the two-qubit Werner state

$$\rho(v) = v \left| \psi_- \right\rangle \left\langle \psi_- \right| + (1-v) \frac{\mathbb{1}}{4}, \qquad (1)$$

where $\left| \psi_- \right\rangle = (\left| 01 \right\rangle - \left| 10 \right\rangle)/\sqrt{2}$ is the singlet state and $\frac{\mathbb{1}}{4}$ is the maximally mixed state, admits a LHV model for arbitrary projective measurements, is directly related to the Grothendieck constant (of order 3) $K_G(3)$. Specifically, one has that $v_c = 1/K_G(3)$. While the exact value of the Grothendieck constants are generally unknown, existing (upper) bounds can be used for deriving lower bounds on v_c. Notably, a result of Krivine [21] for bounding $K_G(3)$ implies that $v_c \geq 0.6595$. Note also that upper bounds on v_c can be obtained by demonstrating explicit Bell inequality violations. From the well-known Clauser-Horne-Shimony-Holt Bell inequality, it follows that $v_c \leq 1/\sqrt{2} \simeq 0.7071$. This was later improved to $v_c \leq 0.7054$ [22, 23] and very recently to $v_c \leq 0.7012$ [24].

More generally, Grothendieck's problem has implications in many areas of mathematics. It had first major impact on Banach space theory and in $\mathcal{C}^*$-algebra theory. More recently, it impacted graph theory and computer science. For more details, we refer to the following review [25].

In this work, we present better LHV models for two-qubit Werner states. Let us set

$$v_1 = \frac{999 \cdot 689}{10^6} \cos^4 \left(\frac{\pi}{50} \right). \qquad (2)$$

We first consider the case of projective measurements and prove analytically that

$$v_c \geq v_1 \simeq 0.6829 \Leftrightarrow K_G(3) \leq \frac{1}{v_1} \simeq 1.4644 \quad (3)$$

which also leads to a better bound on the Grothendieck constant, as stated above. This result is derived by combining two recently introduced numerical methods. The first one is an algorithmic method for constructing LHV models [26, 27], which is applicable to arbitrary entangled states. The second is a numerical optimization algorithm for estimating the distance between a point and a convex set in $\mathbb{R}^d$ [24]. From the output of these numerical methods, we construct the LHV model analytically.

Our second result is a better LHV model for the two-qubit Werner state considering arbitrary POVMs. The model works for $v \leq 2v_1/3 \simeq 0.4553$. The proof relies on the fact that the statistics of arbitrary POVMs featuring a certain level of noise can be reproduced exactly via projective measurements only. By applying this observation to the LHV model we construct for projective measurements, the result follows.

More generally, we believe that the methods presented here open promising new possibilities for the construction of LHV models for quantum states, as well as for other convex set membership problems, such as steerability of quantum states (see e.g. references [7, 10]). We conclude by discussing several possible directions for future research.

1 Concepts and notations

Consider an experiment in which two distant parties share a two-qubit Werner state (1), perform some local measurements labelled by x, y and obtain outputs a, b, respectively. The statistics of the experiment is given by

$$p(ab|xy) = \text{Tr}[\rho(v) A_{a|x} \otimes B_{b|y}]. \qquad (4)$$

Here $A_{a|x}$ and $B_{b|y}$ are the operators representing the local measurements of Alice and Bob. They satisfy positivity and normalization, i.e. $A_{a|x} \geq 0$ and $\sum_a A_{a|x} = \mathbb{1}$, and similarly for $B_{b|y}$. These represent general POVMs, which we will consider in the second part of the paper. In the

first part, however, we focus on the case of local projective measurements, i.e. adding the constraints $a, b \in \{-1, +1\}$ and $A_{a|x} A_{a'|x} = \delta_{aa'} A_{a|x}$ for all x, a, a' and similarly for $B_{b|y}$.

Our goal is to determine the range of visibilities v for which the Werner state (1) admits a LHV model. That is, the measurement statistics (4) can be decomposed as

$$p(ab|xy) = \int q(\lambda) p_A(a|x\lambda) p_B(b|y\lambda) d\lambda \quad (5)$$

where λ represents the local variable, distributed according to the density $q(\lambda)$, and the distributions $p_A(a|x\lambda)$ and $p_B(b|y\lambda)$ are Alice and Bob's local response functions. If such a decomposition can be found for all local projective measurements, the state $\rho(v)$ is said to be local for projective measurements. Moreover, if this decomposition can be extended to all local POVMs, $\rho(v)$ is termed local for POVMs.

As mentioned above the case of projective measurements has a strong connection to the Grothendieck constant, a mathematical constant arising in the context of Banach space theory. Local dichotomic projective qubit measurements are conveniently described via observables of the form

$$O_{\hat{x}} = \hat{x} \cdot \vec{\sigma} , \ O_{\hat{y}} = \hat{y} \cdot \vec{\sigma} \quad (6)$$

where the measurement directions are given by unit vectors $\hat{x}$ and $\hat{y}$ on the Bloch sphere, i.e. $\hat{x}, \hat{y} \in \mathbb{R}^3$, and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices. The measurement statistics of the Werner state (1) are then simply characterized by the expectation values

$$\langle a \rangle = \text{Tr}[O_{\hat{x}} \otimes \mathbb{1} \rho(v)] = 0$$
$$\langle b \rangle = \text{Tr}[\mathbb{1} \otimes O_{\hat{y}} \rho(v)] = 0$$
$$\langle ab \rangle = \text{Tr}[O_{\hat{x}} \otimes O_{\hat{y}} \rho(v)] = -v \, (\hat{x} \cdot \hat{y}). \quad (7)$$

The problem is now to find the largest visibility, v_c, such that the above statistics admits a LHV model. For any visibility $v > v_c$, Bell inequality violation is then possible, even though the Bell test may require an infinite number of local measurements.

Interestingly this problem can be directly related to another purely abstract problem, discussed by Grothendieck. Consider any possible

$m \times m$ matrix M such that

$$\left| \sum_{i,j=1}^{m} M_{ij} \alpha_i \beta_j \right| \leq 1 \quad (8)$$

for any real numbers $\alpha_i, \beta_j \in [-1, +1]$. Then, $K_G(n)$ is the smallest number such that

$$\left| \sum_{i,j=1}^{m} M_{ij} \hat{\alpha}_i \cdot \hat{\beta}_j \right| \leq K_G(n) \quad (9)$$

for any unit vectors $\hat{\alpha}_i, \hat{\beta}_j \in \mathbb{R}^n$ and for any matrix M. This defines a set of numbers $K_G(n)$, called the Grothendieck constants of order n. The Grothendieck constant is then defined as $K_G = \lim_{n \to \infty} K_G(n)$. While the exact values of these constants is not known in general (except for $n = 2$ where $K_G(2) = \sqrt{2}$), lower and upper bounds were proven, see e.g. [28,29]. Of particular relevance to the present work is the constant of order 3, which relates to v_c. Indeed one has that

$$v_c = \frac{1}{K_G(3)}, \quad (10)$$

as shown in Ref. [18], see Theorem 1. This connection follows from early work by Tsirelson [19], who connected the Grothendieck's problem to Bell inequalities. Basically, the matrix M is associated to a Bell inequality, for which the local bound is 1, see equation (8). The largest possible violation of this Bell inequality requires, in general, a maximally entangled state of dimension $d \times d$, where $d = 2^{\lfloor \frac{n}{2} \rfloor}$ [19]. It then follows that $K_G(3)$ is the largest possible Bell violation for a maximally entangled two-qubit state, from which (10) follows. We refer the reader to Ref. [18] for more details.

2 LHV model for projective measurements

Our main result is the construction of a LHV model for projective measurements on Werner states $\rho(v)$ for a visibility v_1 in equation (2), whose approximate value is $v_1 \simeq 0.6829$. This implies the novel bounds on the critical visibility v_c and hence also on the Grothendieck constant $K_G(3)$ stated in equation (3).

We make use of a recently developed method for constructing LHV models for entangled quantum states [26, 27]. The method is algorithmic, in the sense that it is applicable to any entangled state in principle, and can be efficiently implemented on a standard computer (at least for low dimensions). For the case of interest to us, i.e. local projective qubit measurements on a Werner state, the method can be intuitively explained as follows.

Consider a finite set of qubit projective measurements, represented by a finite set of Bloch vectors $\hat{u}_i$, with $i = 1, ..., m$. These vectors form a polyhedron $\mathcal{P}$ contained in the Bloch sphere. Typically, the vectors will be chosen rather uniformly over the sphere, such that the radius of the largest sphere inscribed in $\mathcal{P}$ (and centered at the origin) is close to 1. We refer to this radius as the 'shrinking factor' η of the polyhedron $\mathcal{P}$.

Next we consider the measurement statistics obtained by performing the above set of local measurements (for both Alice and Bob) on the Werner state. Specifically, we get

$$p(ab|xy) = \text{Tr}[\rho(v) A_{a|x} \otimes B_{b|y}] \qquad (11)$$

with $A_{a|x} = \frac{1}{2}(\mathbb{1} + a\,\hat{u}_x \cdot \vec{\sigma})$ and $B_{b|y} = \frac{1}{2}(\mathbb{1} + b\,\hat{u}_y \cdot \vec{\sigma})$, where $a, b \in \{\pm 1\}$. Since we consider a finite set of m measurement settings (for both Alice and Bob), one can find the maximal visibility v^* for which the above measurement statistics admits a LHV model. In practice, this can be done efficiently (at least for $m \leq 10$) using linear programming, see e.g. [2].

As the Werner state $\rho(v^*)$ is local for the set of measurements given by Bloch vectors $\hat{u}_i$, it follows that $\rho(v^*)$ is also local for any noisy measurement of the form

$$A_{a|x}^{\eta} = \frac{1}{2}(\mathbb{1} + a\,\eta\hat{u}_x \cdot \vec{\sigma})$$
$$B_{b|y}^{\eta} = \frac{1}{2}(\mathbb{1} + b\,\eta\hat{u}_y \cdot \vec{\sigma}). \qquad (12)$$

Note that the above measurements form a continuous set, forming a 'shrunk' Block sphere, given by vectors $\eta\hat{u}$ where η is the shrinking factor of the polyhedron $\mathcal{P}$. As these shrunk vectors lie on the sphere inscribed in $\mathcal{P}$, they can be expressed as convex combinations of the finite set of vectors $\hat{u}_i$. One can then show that the distribution $p(ab|\hat{x}\hat{y}) = \text{Tr}[\rho(v^*) A_{a|\hat{x}}^{\eta} \otimes B_{b|\hat{y}}^{\eta}]$ is local, for any possible measurement directions $\hat{x}$ and $\hat{y}$. This follows from the linearity of the trace rule, and from the fact that the noisy measurement operators $A_{a|x}^{\eta}$, respectively $B_{b|y}^{\eta}$, can be expressed as convex combinations of the noiseless operators $A_{a|x}$, respectively $B_{b|y}$; see [26, 27] for more details.

Finally, note that the statistics of the noisy measurements (12) on $\rho(v^*)$ are in fact equivalent to the statistics of noiseless measurements on a slightly more noisy Werner state $\rho(\eta^2 v^*)$. Indeed, it is straightforward to verify that

$$\text{Tr}[\rho(v^*) A_{a|\hat{x}}^{\eta} \otimes B_{b|\hat{y}}^{\eta}] = \text{Tr}[\rho(\eta^2 v^*) A_{a|\hat{x}} \otimes B_{b|\hat{y}}]$$

which again holds for any possible measurement directions $\hat{x}$ and $\hat{y}$. We thus conclude that the Werner state $\rho(\eta^2 v^*)$ admits a LHV model for all local projective measurements.

As mentioned above, this method can be implemented in practice using standard tools when considering sets of relatively few measurements ($m \leq 10$). In Ref. [26], this was used to construct a LHV model for the Werner state for visibilities up to $v = 0.54$. While this construction improves on Werner's original model, which attained $v = 1/2$, it does not reach the best known value so far of $v = 0.6595$ obtained in Ref. [18] based on the connection to the Grothendieck constant. However, a remarkable feature of the above algorithm is the fact that it will converge to v_c when $m \rightarrow \infty$ [26]. Therefore by running the method for finite sets of measurements featuring a large (but nevertheless finite) number of vectors, one can expect to approach the optimal visibility v_c. In particular, one may expect to overcome the best known value of $v = 0.6595$ in case the latter is suboptimal. This is precisely what we implemented, using sets containing up to $m = 625$ vectors. This allows us to obtain the new bounds stated in equation (3).

It is however non-trivial to run the algorithm for sets containing so many vectors. Let us discuss why. Since, the local marginals vanish (see Eq. (7)), we can restrict ourselves to the set of joint correlation terms $\{\langle a_x b_y \rangle\}_{x,y}$. Therefore, in case of m binary measurements per party, the local polytope with completely random marginals, which we call correlation polytope and de-

note by $\mathcal{L}$, lives in dimension m^2. Each vertex of $\mathcal{L}$ corresponds to a local deterministic strategy $\lambda = (a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_m)$, where each a_x, b_y, $(x, y = 1, \ldots, m)$, may take the values ± 1, which amounts to 2^{2m} distinct strategies. A given λ strategy translates to a vertex $\vec{D}_\lambda$, which is a $m \times m$ matrix with entries $D_\lambda(x, y) = a_x b_y$. Hence, the polytope $\mathcal{L}$ features altogether 2^{2m} vertices.

Our goal is now to decompose a given quantum point $\vec{q}$, whose entries are $q(x, y) = \langle a_x b_y \rangle = -v \hat{x} \cdot \hat{y}$, as a convex combination of deterministic vertices: $\vec{q} = \sum_\lambda w_\lambda \vec{D}_\lambda$. This proves that $\vec{q}$ is local. In principle, this problem can be solved via linear programming. However, for $m = 625$ settings, even *inputting* all deterministic strategies to our linear programming solver is completely out of reach.

In order to circumvent these problems, we resort to a modified version of Gilbert's algorithm [30], a popular collision detection method used for instance by the video game industry. This algorithm can provide a $\vec{q}_\epsilon$ such that $\|\vec{q} - \vec{q}_\epsilon\| \le \epsilon$ without a full vertex characterization of the local polytope. The reader is referred to Ref. [24] for more details about this algorithm and its extension including convergence properties and further applications in quantum information. The algorithm is iterative and is given by:

1. Set $i = 0$ and pick an arbitrary point $\vec{q}_i$ inside the polytope $\mathcal{L}$.

2. Given the point $\vec{q}_i$ and the target point $\vec{q}$, run an oracle which maximizes the overlap $(\vec{q} - \vec{q}_i) \cdot \vec{l}$ over all $\vec{l} \in \mathcal{L}$. Let us denote the local point $\vec{l}$ returned by the oracle by $\vec{l}_i$.

3. Find the convex combination $\vec{q}_{i+1}$ of $\vec{q}_i$ and $\vec{l}_i$ that minimizes the distance $\|\vec{q} - \vec{q}_{i+1}\|$.

4. Let $i = i + 1$ and go to Step 2 until the distance $\|\vec{q} - \vec{q}_i\| \le \epsilon$. Return $\vec{q}_\epsilon \equiv q_i$.

Note that the distance $\|\vec{q} - \vec{q}_i\|$ is a decreasing function of i, and actually, when $\vec{q}$ happens to lie inside $\mathcal{L}$, the algorithm is guaranteed to stop after a number of steps $O(1/\epsilon^2)$ [30]. Since maximizing the overlap $(\vec{q} - \vec{q}_i) \cdot \vec{l}$ over all local vectors $\vec{l}$ is a NP-hard problem, in Step 2 we

must make use of a heuristic method, described in Appendix A.

Analytic lower bound for v_c. We now discuss explicitly the procedure we implemented in order to obtain the new bound (3) on $v_c = 1/K_G(3)$. It is important to note that, while our procedure is based on implementing on a computer the above methods (hence giving a numerical result), the final result is proven analytically. This is done as follows.

The finite set of measurement settings we use is based on a family of polyhedra parameterized by an integer n which results in $m = n^2$ ($m = n^2 - n + 1$) vertices in case of n odd (even). The shrinking factor of this polyhedron is given by $\eta = \cos^2(\pi/2n)$. Both the construction of the polyhedra in terms of unit vectors $\hat{u}(i_1, i_2)$ and the proof regarding the value of the shrinking factor can be found in Appendix B. We could implement the calculation up to $n = 25$ (i.e. $m = n^2 = 625$ settings) for which we find the lower bound (3).

We set the initial visibility of the Werner state to $v_0 = 689/1000$, which is guessed to be close to v_c. Combined with the above $m = 625$ measurement settings (for both Alice and Bob), we obtain the target quantum point $\vec{q}$.

After 23×10^6 iterations of the algorithm (*i.e.* we repeat 23×10^6 times the steps 2-4 of the algorithm), which was completed on a standard desktop PC within a week, we get numerically a point $\vec{q}_\epsilon$ such that $\|\vec{q} - \vec{q}_\epsilon\| \le 9.8484 \times 10^{-6}$. We then truncate $\vec{q}_\epsilon$ up to $k = 16$ digits, which results in the rational point $\vec{q}_r$. Note that $\vec{q}_r$ is now local by construction; see Appendix C. We now have that

$$\vec{q} = \vec{q}_r + \vec{q}_{\text{junk}}, \tag{13}$$

where $\vec{q}_{\text{junk}}$ takes care of the (small) difference between the analytical points $\vec{q}$ and $\vec{q}_r$. Let us now slightly shrink the point $\vec{q}$ towards the centre of the local correlation polytope (the origin) by rescaling $\vec{q}$ with a factor ν close to (but strictly smaller than) 1:

$$\nu \vec{q} = \nu \vec{q}_r + (1 - \nu) \frac{\nu \vec{q}_{\text{junk}}}{1 - \nu}. \tag{14}$$

Clearly, we see that $\nu \vec{q}$ is provenly local if the point $\vec{x} = \nu \vec{q}_{\text{junk}}/(1 - \nu)$ is local as well. By

rearranging (14), we have that

$$\vec{x} = \frac{\nu}{1 - \nu} (\vec{q} - \vec{q}_r), \qquad (15)$$

where each entry of $\vec{x}$ has an analytical form. Note that all the components of $\vec{x}$ are expected to be small (in norm), since the points $\vec{q}$ and $\vec{q}_r$ are very close (note that ν will be chosen such that the factor $\nu/(1 - \nu)$ is not too large). In fact, it can be proven that the point $\vec{x}$ is local, using the following result:

Lemma 1. A correlation point $\vec{x}$ is local if $\sum_{i,j} |x_{ij}| < 1$.

The proof, as well as more details on this analysis, can be found in Appendix C.

Next, setting $\nu = 999/1000$ we obtain the bound

$$\frac{\nu}{1 - \nu} \sum_{x,y=1}^{m} |q(x, y) - q_r(x, y)| < 1, \qquad (16)$$

implying that the point $\vec{q}$ is local via the above lemma.

To summarize, we obtain that the following lower bound for the critical visibility

$$v_c \geq \eta^2 \nu v_0 \simeq 0.6829 \qquad (17)$$

where $\eta = \cos^2(\pi/50)$ is the shrinking factor of the polyhedron. We provide a Mathematica file which gives the points $\vec{q}$ and $\vec{q}_r$ and checks the validity of condition (16), as well as a file containing all the data for the proof. This material is available online [31].

3 LHV model for POVMs

We also provide a better LHV model for Werner states considering arbitrary local measurements, i.e. moving from projective measurements to general POVMs. Specifically, we give a model for $v_2 = 2v_1/3 \simeq 0.4553$. This improves on a previous model of Barrett [5] that reached $v = 5/12 \simeq 0.4167$.

The construction of the model is based on the following argument. Essentially, any noisy qubit POVM can be expressed as a convex combination involving only projective qubit measurements, given the amount of white noise is above a certain threshold μ. Therefore, if the

statistics of certain Werner states $\rho(v)$ admit a LHV model, then so do the statistics of noisy POVMs (given the amount of noise is above the threshold μ). Again, this follows from the linearity of the trace rule. In turn this implies that the statistics of arbitrary (noiseless) POVMs on the slightly more noisy Werner state $\rho(\mu^2 v)$ is local.

More formally, we can make the following statement.

Lemma 2. Any noisy qubit POVM $M(\mu)$ with elements $\{M_i(\mu)\}_{i=1,...,4}$ proportional to rank-1 projectors for $\mu = \sqrt{\frac{2}{3}} - \epsilon$ can be written as a convex sum of rank-1 projectors, where ϵ may be arbitrary close to zero.

The proof is deferred to Appendix D. The above value $\mu = \sqrt{2/3}$ along with our lower bound v_1 for v_c implies the lower bound of the visibility

$$\mu^2 v_1 = \frac{2}{3} v_1 = \frac{2 \cdot 999 \cdot 689}{3 \cdot 10^6} \cos^4\left(\frac{\pi}{50}\right) \simeq 0.4553 \qquad (18)$$

for a LHV model for POVMs. We also refer to an independent related work [32], in particular for an alternative proof of Lemma 2.

4 Discussion

We have presented better LHV models for Werner states, as well as a new upper bound on the Grothendieck constant of order 3. The methods we develop provide analytical bounds, which will converge to the exact value of v_c (and $K_G(3)$) using increased computational power.

Clearly, these methods can be applied to construct LHV models for other classes of entangled states, in particular in higher Hilbert space dimensions. It would also be interesting to adapt the present technique to construct local hidden state models, a specific class of LHV models relevant in the context of quantum steering [7]. Finally, these methods could also be used to obtain bounds on other Grothendieck constants. While this looks computationally challenging at first sight, taking advantage of symmetry arguments could lead to progress.

Note added. In a related work, the authors of Ref. [32] also presented a better LHV model for

Werner states for POVMs, achieving a visibility similar to ours.

References

[1] J. S. Bell, "On the Einstein-Poldolsky-Rosen paradox," *Physics* **1**, 195–200 (1964). https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf.

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," *Reviews of Modern Physics* **86**, 419–478 (2014), arXiv:1303.2849 [quant-ph].

[3] N. Gisin, "Bell's inequality holds for all non-product states," *Physics Letters A* **154**, 201 – 202 (1991).

[4] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Phys. Rev. A* **40**, 4277–4281 (1989).

[5] J. Barrett, "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality," *Phys. Rev. A* **65**, 042302 (2002), quant-ph/0107045.

[6] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, "Noise Robustness of the Nonlocality of Entangled Quantum States," *Phys. Rev. Lett.* **99**, 040403 (2007), quant-ph/0703018.

[7] H. M. Wiseman, S. J. Jones, and A. C. Doherty, "Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox," *Phys. Rev. Lett.* **98**, 140402 (2007), quant-ph/0612147.

[8] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, "Genuine Hidden Quantum Nonlocality," *Phys. Rev. Lett.* **111**, 160402 (2013), arXiv:1307.4404 [quant-ph].

[9] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, "One-way Einstein-Podolsky-Rosen Steering," *Phys. Rev. Lett.* **112**, 200402 (2014), arXiv:1402.3607 [quant-ph].

[10] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz, and H. M. Wiseman, "Einstein-Podolsky-Rosen steering and the steering ellipsoid," *Journal of the Optical Society of America B Optical Physics* **32**, A40 (2015), arXiv:1411.1517 [quant-ph].

[11] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, "Local Hidden Variable Models for Entangled Quantum States Using Finite Shared Randomness," *Phys. Rev. Lett.* **114**, 120401 (2015), arXiv:1412.1416 [quant-ph].

[12] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, "Sufficient criterion for guaranteeing that a two-qubit state is unsteerable," *Phys. Rev. A* **93**, 022121 (2016), arXiv:1510.06721 [quant-ph].

[13] H. Chau Nguyen and T. Vu, "Necessary and sufficient condition for steerability of two-qubit states by the geometry of steering outcomes," *EPL (Europhysics Letters)* **115**, 10003 (2016), arXiv:1604.03815 [quant-ph].

[14] G. Tóth and A. Acín, "Genuine tripartite entangled states with a local hidden-variable model," *Phys. Rev. A* **74**, 030306 (2006), quant-ph/0512088.

[15] J. Bowles, J. Francfort, M. Fillettaz, F. Hirsch, and N. Brunner, "Genuinely Multipartite Entangled Quantum States with Fully Local Hidden Variable Models and Hidden Multipartite Nonlocality," *Phys. Rev. Lett.* **116**, 130401 (2016), arXiv:1511.08401 [quant-ph].

[16] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, "Entanglement and

Nonlocality are Inequivalent for Any Number of Parties," *Phys. Rev. Lett.* **115**, 030404 (2015), `arXiv:1407.3114 [quant-ph]`.

[17] R. Augusiak, M. Demianowicz, and A. Acín, "Local hidden-variable models for entangled quantum states," *Journal of Physics A Mathematical General* **47**, 424002 (2014), `arXiv:1405.7321 [quant-ph]`.

[18] A. Acín, N. Gisin, and B. Toner, "Grothendieck's constant and local models for noisy entangled quantum states," *Phys. Rev. A* **73**, 062105 (2006), `quant-ph/0606138`.

[19] B. S. Tsirelson, "Some results and problems on quantum Bell-type inequalities," *Hadronic Journal Supplement* **8**, 329–345 (1993). `http://www.tau.ac.il/~tsirel/download/hadron.pdf`.

[20] A. Grothendieck, "Résumé de la théorie métrique des produits tensoriels topologiques," *Bol. Soc. Mat. São Paulo* **8**, (1953). `https://www.ime.usp.br/acervovirtual/textos/estrangeiros/grothendieck/produits_tensoriels_topologiques/files/produits_tensoriels_topologiques.pdf`.

[21] J. Krivine, "Constantes de Grothendieck et fonctions de type positif sur les sphères," *Advances in Mathematics* **31**, 16 – 30 (1979).

[22] T. Vértesi, "More efficient Bell inequalities for Werner states," *Phys. Rev. A* **78**, 032112 (2008), `arXiv:0806.0096 [quant-ph]`.

[23] B. Hua, M. Li, T. Zhang, C. Zhou, X. Li-Jost, and S.-M. Fei, "Towards Grothendieck constants and LHV models in quantum mechanics," *Journal of Physics A Mathematical General* **48**, 065302 (2015), `arXiv:1501.05507 [quant-ph]`.

[24] S. Brierley, M. Navascues, and T. Vertesi, "Convex separation from convex optimization for large-scale problems," *ArXiv e-prints* (2016), `arXiv:1609.05011 [quant-ph]`.

[25] G. Pisier, "Grothendieck's Theorem, past and present," *Bull. Amer. Math. Soc.* **49**, 237–323 (2011), `arXiv:1101.4195 [math.FA]`.

[26] F. Hirsch, M. T. Quintino, T. Vértesi, M. F. Pusey, and N. Brunner, "Algorithmic Construction of Local Hidden Variable Models for Entangled Quantum States," *Phys. Rev. Lett.* **117**, 190402 (2016), `arXiv:1512.00262 [quant-ph]`.

[27] D. Cavalcanti, L. Guerini, R. Rabelo, and P. Skrzypczyk, "General Method for Constructing Local Hidden Variable Models for Entangled Quantum States," *Phys. Rev. Lett.* **117**, 190401 (2016), `arXiv:1512.00277 [quant-ph]`.

[28] S. R. Finch, *Mathematical constants*. Cambridge University Press, 2003. `http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521818052`.

[29] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor, "The Grothendieck constant is strictly smaller than Krivine's bound," *Forum of Mathematics, Pi* (2013), `arXiv:1103.6161 [math.FA]`.

[30] E. G. Gilbert, "An iterative procedure for computing the minimum of a quadratic form on a convex set," *SIAM Journal on Control* **4**, 61–80 (1966).

[31] `https://cloud.atomki.hu/s/KG3m625`.

[32] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, "Simulating positive-operator-valued measures with projective measurements," *ArXiv e-prints* (2016), `arXiv:1609.06139 [quant-ph]`.

[33] G. Mauro D'Ariano, P. Lo Presti, and P. Perinotti, "Classical randomness in quantum measurements," *Journal of Physics A Mathematical General* **38**, 5979–5991 (2005), `quant-ph/0408115`.

5 Appendix A. Description of the heuristic oracle

The oracle returning $\vec{l}_i$ in maximizing the overlap

$$S = (\vec{q} - \vec{q}_i) \cdot \vec{l} \qquad (19)$$

over all $\vec{l} \in \mathcal{L}$ is a heuristic one. This problem corresponds to Step 2 in the algorithm in Sec. 2 of the main text. It is first noted that it is enough to maximize over all vertices $\vec{D}_\lambda$ of the set $\mathcal{L}$, since $\mathcal{L}$ is a polytope. The objective S in Eq. (19) for a given strategy $\lambda = (a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_m)$ (and corresponding vertex $\vec{D}_\lambda$) is as follows:

$$S_\lambda = \sum_{x,y=1}^{m} (q(x,y) - q_i(x,y)) a_x b_y. \qquad (20)$$

The number of vertices (and different λ strategies) is 2^{2m}, hence evaluating S_λ for all λ and picking the biggest one is clearly not tractable in our range of $m > 600$. Therefore, instead of a brute force computation we have to resort to a heuristic method. Note that a heuristic method still suffices, since the intuition behind $\vec{l}_i$ is that it gives a direction for $\vec{q}_i$ to move towards a better point.

The iterative algorithm is as follows.

1. Choose randomly assignments $\{a_x \in \pm 1\}_x$ for the deterministic strategy.

2. Fixing $\{a_x\}_x$, maximize S in function of b_y. This amounts to setting $b_y = +1$ if $\sum_x (q(x,y) - q_i(x,y)) a_x > 0$, otherwise to setting $b_y = -1$ for all $y = 1, \ldots, m$.

3. Fixing $\{b_y\}_y$, maximize S in function of a_x. This amounts to setting $a_x = +1$ if $\sum_y (q(x,y) - q_i(x,y)) b_y > 0$, otherwise to setting $a_x = -1$ for all $x = 1, \ldots, m$.

4. Go back to step 2 until convergence of S is reached.

Note that in each iteration step the value of the objective S is guaranteed not to decrease. However, the algorithm may easily get stuck in a non-optimal S. To make the iterative procedure more efficient, we run it several times (in practice we chose to run it 100 times) starting from different random seeds, and pick the solution λ and corresponding vertex $\vec{D}_\lambda$ with the biggest value of S.

6   Appendix B. A family of polyhedra

The polyhedra are parameterized by an integer n. In case of n odd, the vertices $\hat{u}(i_1, i_2)$ are given by

$$\hat{u}(i_1, i_2) = \left(\cos\frac{i_1\pi}{n}\cos\frac{i_2\pi}{n}, \sin\frac{i_1\pi}{n}\cos\frac{i_2\pi}{n}, \sin\frac{i_2\pi}{n} \right),$$
$$(21)$$

where $i_1, i_2 \in \{0, 1, \ldots, n-1\}$ plus their antipodal points. This amounts to $m = n^2$ measurement settings. In case of n even, the number of vertices up to inversion (i.e. the number of settings) are $m = n^2 - n + 1$ due to possible redundancy of some of the vertices.

It is easy to derive an analytical expression for the shrinking factor associated with this polyhedron, which is given by $\eta = \cos^2(\pi/2n)$. Indeed, let us fix i_1, in which case the vertices $\pm\hat{u}(i_1, i_2)$, $i_2 = (0, \ldots, n-1)$ define a regular $2n$-gon in the two-dimensional plane with a planar shrinking factor of $\eta_{2d} = \cos(\pi/2n)$. Any point $\hat{u}$ on the unit sphere can be written as a convex combination of two neighboring planes defined by some $i_1 \in \{0, 1, \ldots, n-1\}$ and $i_1 + 1 \pmod{n}$. In the worst case, the point $\hat{u}$ lies just midway between these two planes, in which case we get the lower bound of $\eta = \eta_{2d}\cos(\pi/2n) = \cos^2(\pi/2n)$ on the 3-dimensional shrinking factor.

7   Appendix C. Going from numerical to exact precision

The algorithm described in the main text does not provide us precisely the point $\vec{q}$ but only a point $\vec{q}_\epsilon$, which is very close to $\vec{q}$, say,

$$\|\vec{q} - \vec{q}_\epsilon\| \le \epsilon. \qquad (22)$$

Note that $\vec{q}_\epsilon = \sum_\lambda w_\lambda \vec{D}_\lambda$, where the (positive) weights w_λ coming from the algorithm are given in double precision format, whereas $\vec{D}_\lambda$ are deterministic strategies with ± 1 entries.

In order to provide an analytical proof, we first transform the positive floating-point numbers w_λ to positive rationals w_λ^r. To this end, we use the truncation

$$w_\lambda^r = \frac{\lfloor 10^k w_\lambda \rfloor}{10^k}, \qquad (23)$$

where k denotes the number of figures kept behind the decimal point. Then we renormalize and obtain the rational point

$$\vec{q}_r = \frac{\sum_\lambda w_\lambda^r \vec{D}_\lambda}{\sum_\lambda w_\lambda^r} \qquad (24)$$

which is local by construction. Then we can write

$$\vec{q} = \vec{q}_r + \vec{q}_{\text{junk}}, \qquad (25)$$

where $\vec{q}_{\text{junk}}$ takes care of the (small) difference between the analytical points $\vec{q}$ and $\vec{q}_r$. Let us now slightly shrink the point $\vec{q}$ towards the centre of the local polytope by rescaling $\vec{q}$ with a factor ν smaller than 1:

$$\nu\vec{q} = \nu\vec{q}_r + \nu\vec{q}_{\text{junk}} = \nu\vec{q}_r + (1-\nu)\frac{\nu\vec{q}_{\text{junk}}}{1-\nu}. \quad (26)$$

From the right-hand side expression, it is clear that $\nu\vec{q}$ is provenly local if the point $\vec{x} = \nu\vec{q}_{\text{junk}}/(1-\nu)$ is local as well. By rearranging (26), we have

$$\vec{x} = \frac{\nu}{1-\nu}(\vec{q} - \vec{q}_r), \qquad (27)$$

where each entry of $\vec{x}$ has an analytical form. Moreover, due to equation (22), the entries of $\vec{x}$ are typically small in case of small ϵ, and ν not extremely close to 1. This suggests an easy test to decide if $\vec{x}$ is local. Namely,

Lemma 1. A correlation point $\vec{z}$ is local if $\sum_{i,j} |z_{ij}| < 1$.

Proof. The proof is based on an explicit decomposition

$$\vec{z} = \sum_{i,j} |z_{ij}| \operatorname{sgn}(z_{ij})\vec{E}_{i,j} + \left(1 - \sum_{i,j} |z_{ij}|\right) \vec{E}_0 \tag{28}$$

in terms of local points $\pm\vec{E}_{i,j}$ and $\vec{E}_0$, where all entries of the point $\vec{E}_{i,j}$ are zero but entry (i,j) where it takes up $+1$. On the other hand, $\vec{E}_0$ stands for the distribution with all entries zero. From the positivity of the weights in the above decomposition, it follows that $\vec{z}$ admits a LHV model if $\sum_{i,j} |z_{ij}| < 1$ as claimed in Lemma 2. Note also that $\sum_{i,j} |z_{ij}| < 1$ entails that all entries are bounded by ± 1, hence such a $\vec{z}$ is a valid correlation point by definition. $\qquad \square$

8 Appendix D. Decomposing noisy qubit POVMs in terms of projectors

An extremal POVM M for qubits can be characterized as follows. It has no more than four elements $\{M_i\}$, $i = 1, 2, 3, 4$ such that each element M_i is proportional to a rank 1 projector [33]. Let us define the vector of Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and write the POVM elements in the form

$$M_i = a_i \mathbb{1} + \vec{a}_i \cdot \vec{\sigma}, \qquad (29)$$

where $a_i = |\vec{a}_i|$, $\sum_i a_i = 1$, and $\sum_i \vec{a}_i = 0$. Similarly, we define the elements of a noisy POVM as follows

$$M_i(\mu) = a_i \mathbb{1} + \mu\vec{a}_i \cdot \vec{\sigma}, \qquad (30)$$

where $0 \le \mu \le 1$. Note that $\mu = 1$ corresponds to the noiseless case. Also an arbitrary qubit rank 1 projector P can be described with the following two elements

$$\begin{aligned} P_1 &= (1/2)\mathbb{1} + (1/2)\vec{b}\cdot\vec{\sigma}, \\ P_2 &= (1/2)\mathbb{1} - (1/2)\vec{b}\cdot\vec{\sigma}, \end{aligned} \qquad (31)$$

where $\vec{b}$ is a unit vector.

With these definitions, we state our Lemma 1 in the main text:

Any noisy qubit POVM $M(\mu)$ with elements $\{M_i(\mu)\}_{i=1,\dots,4}$ proportional with rank 1 projectors for $\mu = \sqrt{\frac{2}{3}} - \epsilon$ can be written as a convex sum of rank 1 projectors $P^{(k)}$, where ϵ may be arbitrary close to zero and k may run up to infinity.

In order to prove it, we start with the following lemma.

Lemma 3. Any noisy POVM $M(\mu)$ with elements $\{M_i(\mu)\}$, $i = 1, 2, 3, 4$ for $\mu = \sqrt{\frac{2}{3}} - \epsilon$ can be expressed in terms of a projector P and another noisy POVM $M'(\mu)$ with the same factor μ as follows

$$M(\mu) = pP + (1-p)M'(\mu), \qquad (32)$$

where p is strictly larger than zero if $\epsilon > 0$.

Note that the above lemma already provides us a constructive method to prove Lemma 1: We start with $M(\mu)$ and use Lemma 3 to decompose it as

$$M(\mu) = p_1 P^{(1)} + (1 - p_1)M^{(1)}(\mu) \qquad (33)$$

with $p_1 > 0$. Then we run the protocol again starting with $M^{(1)}(\mu)$

$$M^{(1)}(\mu) = p_2 P^{(2)} + (1-p_2)M^{(2)}(\mu) \qquad (34)$$

to get the decomposition

$$M(\mu) = p_1 P^{(1)} + (1-p_1)p_2 P^{(2)} + (1-p_1)(1-p_2)M^{(2)}(\mu). \qquad (35)$$

Running this iterative procedure up to n times, we get a decomposition of $M(\mu)$ in terms of n projectors $P^{(k)}$, $k = 1, 2, \ldots, n$ and a POVM $M^{(n)}(\mu)$ with an overall weight of $\prod_{k=1}^{n}(1-p_k)$ of this POVM. If we were demanding each p_k to be maximal at each step k, then, as n goes to infinity, this weight must go to zero. For, if it did not, then there would be a subsequence of $(M^{(k)}(\mu))_k$ converging to some POVM $\bar{M}(\mu)$ with the property that it cannot be decomposed further, a contradiction. Therefore one arrives at a decomposition of $M(\mu)$ only in terms of projectors.

We are now left with a proof of Lemma 3. To this end, we state another lemma.

Lemma 4. Given four nonzero vectors $\vec{a}_i$, $i = 1, 2, 3, 4$ in the three-dimensional Euclidean space such that they sum up to zero. Then the relation

$$\frac{\vec{a}_i \cdot \vec{a}_j}{a_i a_j} \le -\frac{1}{3} \qquad (36)$$

holds true for at least one pair (say, i and j), where we defined $a_i = |\vec{a}_i|$. In other words, one can always pick two vectors for which the angle θ_{ij} between them is at least $\theta_{ij} = \arccos(-1/3)$. Note the special case of the vertices of the regular tetrahedron for which each angle θ_{ij} between the vectors formed by the vertices is $\arccos(-1/3)$.

Proof. The proof is by contradiction. Suppose that the lemma is not true and we can pick vectors $\vec{a}_i$ such that

$$\vec{a}_1 \cdot \vec{a}_4 > -a_1 a_4/3$$
$$\vec{a}_2 \cdot \vec{a}_4 > -a_2 a_4/3$$
$$\vec{a}_3 \cdot \vec{a}_4 > -a_3 a_4/3. \qquad (37)$$

Then summing up the above three equations, and plugging $\vec{a}_4 = -\vec{a}_1 - \vec{a}_2 - \vec{a}_3$, we get $(a_1 + a_2 + a_3) > 3a_4$. If we choose for instance the ordering $a_4 \ge a_3 \ge a_2 \ge a_1$, the above relation is clearly not true. $\qquad \square$

With the above tools, we are ready to prove Lemma 3 which in turn proves Lemma 1. To this end we write out equation (32) for each POVM element and we assume w.l.o.g. that the (i, j) pair satisfying equation (36) is given by $(i, j) = (1, 2)$. Then we have

$$a_1 = (p/2) + (1-p)c_1$$
$$a_2 = (p/2) + (1-p)c_2$$
$$a_3 = (1-p)c_3$$
$$a_4 = (1-p)c_4 \qquad (38)$$

for the scalar terms and

$$\mu\vec{a}_1 = (p/2)\vec{b} + (1-p)\mu\vec{c}_1$$
$$\mu\vec{a}_2 = (-p/2)\vec{b} + (1-p)\mu\vec{c}_2$$
$$\mu\vec{a}_3 = (1-p)\mu\vec{c}_3$$
$$\mu\vec{a}_4 = (1-p)\mu\vec{c}_4 \qquad (39)$$

for the vectors, where the other POVM $M'(\mu)$ in (32) is defined by the elements $M'_i(\mu) = c_i + \mu\vec{c}_i \cdot \vec{\sigma}$. Let us denote the angle between $\vec{a}_1$ and $\vec{a}_2$ by θ_{12}, which we write in terms of two positive angles θ_1, θ_2 as follows $\theta_{12} = \theta_1 + \theta_2$ which are yet to be determined. After an appropriate rotation of the coordinate system, we can use the following parametrization

$$\vec{a}_1 = \sin\theta_1 e_x + \cos\theta_1 e_z$$
$$\vec{a}_2 = -\sin\theta_2 e_x + \cos\theta_2 e_z$$
$$\vec{b} = e_x \qquad (40)$$

and we can also assume w.l.o.g. that $a_1 \ge a_2$.

Using the first equation of (39), we separate the e_x and e_z terms which result in two equations

$$\mu a_1 \sin\theta_1 = (p/2) + (1-p)\eta c_{1x}$$
$$\mu a_1 \cos\theta_1 = (1-p)\mu c_{1z}, \qquad (41)$$

where we defined $\vec{c}_i = c_{ix}e_x + c_{iz}e_z$ for $i = 1, 2$. Combining (41) with the first equations of (38) and (39), p can be expressed as follows:

$$p = \frac{4a_1\mu}{1-\mu^2}(\sin\theta_1 - \mu). \qquad (42)$$

Similarly, from the second equations of (38) and (39), we arrive at

$$p = \frac{4a_2\mu}{1-\mu^2}(\sin\theta_2 - \mu). \qquad (43)$$

Given θ_{12}, $a_1 \geq a_2 > 0$ (which define $M(\mu)$), our goal is to find a p strictly greater than 0. Note that due to Lemma 4 we can also assume that $\theta_{12} \geq \arccos(-1/3)$. Note also that $a_1 = 0$ entails $a_2 = 0$, which means that $M(\mu)$ is already a projector and the proof can be finished. On the other hand, if $a_2 = 0$ we go back to the case of a three-outcome POVM which has to be treated similarly to the general four-outcome situation and will be discussed later.

Let us split the general case $a_1 \geq a_2$ into $a_1 = a_2$ and $a_1 > a_2$.

We start with $a = a_1 = a_2$. Here we take $\theta_1 = \theta_2 = \theta_{12}/2$ and the pair of equations (42,43) become a single equation for p:

$$p = \frac{4a\mu}{1-\mu^2}(\sin(\theta_{12}/2) - \mu). \qquad (44)$$

Since we discuss the noisy situation $0 < \mu < 1$, the factor $\frac{4a\mu}{1-\mu^2}$ is positive and we get the condition $\sin(\theta_{12}/2) > \mu$ to get an μ for which p is strictly larger than zero. The worst case scenario is defined by $\theta_{12} = \arccos(-1/3)$ which gives the critical value $\mu_{crit} = \sin(\arccos(-1/3)/2) = \sqrt{2/3}$ below which p is strictly larger than zero.

We now discuss the case $a_1 > a_2$. From the pair of equations (42,43) it is clear that if we want to satisfy both with a single value of p, it is required to have $\sin\theta_2 > \sin\theta_1$. Recalling that $\theta_1 + \theta_2 = \theta_{12} \geq \arccos(-1/3)$, where θ_1 and θ_2 are positive, we get the lower bound $\theta_2 \geq \arccos(-1/3)/2$ which entails the bound $\sin\theta_2 \geq \sqrt{2/3}$.

Altogether we obtain the result that in case of $a_1 \geq a_2 > 0$, p is strictly larger than zero whenever $\mu_{crit} = \sqrt{2/3} - \epsilon$, where ϵ can be arbitrary small.

Let us now come back to the situation when $a_2 = 0$. In that case, we get the very same equations for p as in Eqs. (42,43) with the only exception that instead of Lemma 4, we have

Lemma 5. Given three nonzero vectors $\vec{a}_i$, $i = 1, 2, 3$ in the three-dimensional Euclidean space such that they sum up to zero. Then the relation

$$\frac{\vec{a}_i \cdot \vec{a}_j}{a_i a_j} \leq -\frac{1}{2} \qquad (45)$$

holds true for at least one pair (say, i and j), where we defined $a_i = |\vec{a}_i|$.

The proof is analogous to the one of Lemma 4. However, the new condition in (45) implies the even larger bound of $\mu_{crit} = \sqrt{3}/2$. Hence, the case $a_2 = 0$ is solved as well. As a side result, we also have a theorem for 3-outcome POVMs:

Lemma 6. Any noisy three-outcome POVM $M(\mu)$ with elements $\{M_i(\mu)\}$, $i = 1, 2, 3$ for $\mu = \sqrt{3}/2 - \epsilon$ can be written as a convex sum of projectors $P^{(k)}$, where ϵ may go arbitrary close to zero and k may run up to infinity.