Geometric interpretation of the Clauser-Horne-Shimony-Holt inequality of nonmaximally entangled states

Johannes Seiler [1,*] Thomas Strohm [2] and Wolfgang P. Schleich [1,3,4]

[1]*Institut für Quantenphysik and Center for Integrated Quantum Science and Technology (IQST), Universität Ulm, D-89069 Ulm, Germany*
[2]*Corporate Research, Robert Bosch GmbH, D-71272 Renningen, Germany*
[3]*Institute of Quantum Technologies, German Aerospace Center (DLR), Söflinger Strasse 100, D-89077 Ulm, Germany*
[4]*Hagler Institute for Advanced Study, Institute for Quantum Science and Engineering (IQSE),*
and Texas A&M AgriLife Research, Texas A&M University, College Station, Texas 77843-4242, USA

We show that for pure and mixed states the problem of maximizing the correlation measure in the Clauser-Horne-Shimony-Holt inequality reduces to maximizing the perimeter of a parallelogram enclosed by an ellipse characterized by the entanglement contained in the bipartite system. Our geometrical description is valid for a nonmaximally entangled state. We also determine the corresponding optimal measurements.

I. INTRODUCTION

The Einstein-Podolsky-Rosen situation [1] represents a paradigm of quantum mechanics [2] and brings out the important role of entanglement [3]. This feature leads to a violation of the celebrated Bell inequality [4–6] or the Clauser-Horne-Shimony-Holt (CHSH) inequality [7]. In this article, we provide a geometric interpretation of the quantum mechanical correlation measure central to the CHSH inequality.

Over the years, Bell's inequality has been tested in many remarkable experiments [8–14], ever more strictly supporting quantum mechanics and its predictions. The key ingredient is the entanglement of the underlying systems. Indeed, in order to violate the CHSH inequality, the state has to be entangled [15], and every entangled two-qubit state leads to a violation [16], provided appropriate measurements are chosen.

In fact, the optimal measurement depends crucially on the entanglement of the system. For example, the measurement settings that lead to the maximal violation of the CHSH inequality for a given entangled state can fulfill the inequality for a state with different entanglement. In order to obtain a violation, one therefore has to adapt the measurement strategy to the state [15].

In recent years, the connection between Bell inequalities and entanglement has been exploited in many quantum informational tasks, such as quantum key distribution [17] and quantum random number generators [18,19]. Unfortunately, in real life experiments, the state is never maximally entangled or necessarily pure. Knowing which measurements lead to the maximal violation of the underlying Bell inequality for a given entanglement can therefore increase the efficiency of such tasks.

In this article, we derive for a fixed entanglement and for pure and mixed states all optimal measurement directions as to violate the CHSH inequality. We note that Refs. [16,19] have also addressed this question, however restricted to pure states and have only presented specific examples.

In particular, we consider the CHSH inequality for nonmaximally entangled states and introduce in Sec. II the state, the measurement operators, and the expectation value of interest. In Sec. III, we present a geometrical interpretation of the correlation function bounded by the CHSH inequality, where we connect this expectation value with the perimeter of a parallelogram caught by an ellipse whose semiaxes depend on the strength of the entanglement parametrized by the concurrence. This interpretation highlights the crucial role of the entanglement in violating the CHSH inequality. We then use this geometric argument in Sec. IV to analytically determine all optimal measurement strategies for a given entangled pure two-qubit state. We finally discuss the extension of our geometric interpretation to mixed states in Sec. V. We conclude in Sec. VI by summarizing and providing an outlook. In order to keep our article self-contained but focused on the main ideas, we have included additional material in two Appendixes. In Appendix A, we present the analytical optimization of the measurement strategies, while in Appendix B we prove that these strategies are found geometrically, when demanding that the tangents on the ellipse in the corners of the parallelogram are orthogonal.

II. FORMULATION OF THE PROBLEM

In this section, we briefly recall the essential ingredients of the CHSH inequality by first defining the corresponding quantum mechanical expectation value in terms of measurement operators. In order to lay the groundwork for our geometrical interpretation of the CHSH inequality, we then cast this expectation value in terms of the measurement directions and the correlation matrix. We conclude this section by presenting the Tsirelson bound.

*johannes.seiler@uni-ulm.de

Throughout this article, we consider the pure two-qubit state

$$|\Psi\rangle \equiv \sqrt{\frac{1 + \sqrt{1 - \mathcal{C}^2}}{2}} |0\rangle_A |1\rangle_B - \sqrt{\frac{1 - \sqrt{1 - \mathcal{C}^2}}{2}} |1\rangle_A |0\rangle_B, \tag{1}$$

in the Schmidt basis [20], where $|0\rangle$ and $|1\rangle$ are the eigenstates of the Pauli operator $\hat{\sigma}_z$ with the eigenvalues 1 and -1, respectively. Here, the concurrence $\mathcal{C}$ with $0 \leqslant \mathcal{C} \leqslant 1$ is a measure of the entanglement between the two subsystems. For $\mathcal{C} = 0$ the state is separable, while for $\mathcal{C} = 1$ it is a maximally entangled Bell state.

On each subsystem, we perform one of two measurements. In the subsystem A, we employ either the operator

$$\hat{Q} \equiv \boldsymbol{q} \cdot \hat{\boldsymbol{\sigma}}_A \quad \text{or} \quad \hat{R} \equiv \boldsymbol{r} \cdot \hat{\boldsymbol{\sigma}}_A, \tag{2}$$

with the three-dimensional unit vectors $\boldsymbol{q}$ and $\boldsymbol{r}$ describing the measurement directions on the Bloch sphere, and $\hat{\boldsymbol{\sigma}}_A$ denotes the vector formed by the Pauli matrices $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$.

Analogously, on the subsystem B we make either the measurement

$$\hat{S} \equiv \boldsymbol{s} \cdot \hat{\boldsymbol{\sigma}}_B \quad \text{or} \quad \hat{T} \equiv \boldsymbol{t} \cdot \hat{\boldsymbol{\sigma}}_B, \tag{3}$$

with the three-dimensional unit vectors $\boldsymbol{s}$ and $\boldsymbol{t}$, and $\hat{\boldsymbol{\sigma}}_B$ is again the vector of the Pauli matrices. The measurements on both subsystems are performed independently of each other.

We are interested in the expectation value

$$\mathcal{S} \equiv \langle \hat{Q} \otimes \hat{S} \rangle - \langle \hat{Q} \otimes \hat{T} \rangle + \langle \hat{R} \otimes \hat{S} \rangle + \langle \hat{R} \otimes \hat{T} \rangle, \tag{4}$$

which appears in the CHSH inequality and is a measure for the correlation between the two subsystems [7].

When we substitute the state $|\Psi\rangle$ defined by Eq. (1) into Eq. (4) and use the representations Eqs. (2) and (3) of the measurement operators, we get the representation

$$\mathcal{S} = \boldsymbol{q} K \cdot (\boldsymbol{s} - \boldsymbol{t}) + \boldsymbol{r} K \cdot (\boldsymbol{s} + \boldsymbol{t}) \tag{5}$$

in terms of multiplications between the Bloch vectors $\boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}, \boldsymbol{t}$ of the measurements and the 3×3 correlation matrix

$$K \equiv \langle \Psi | \hat{\boldsymbol{\sigma}}_A \otimes \hat{\boldsymbol{\sigma}}_B | \Psi \rangle. \tag{6}$$

This matrix depends only on the state $|\Psi\rangle$ of the complete system. For the state $|\Psi\rangle$ of Eq. (1), the correlation matrix is [21] the diagonal matrix

$$K = -\operatorname{diag}(\mathcal{C}, \mathcal{C}, 1) \tag{7}$$

and only depends on the concurrence $\mathcal{C}$ and thus on the entanglement of the state. The diagonal form of the correlation matrix, Eq. (7), is not a coincidence, but the result of defining the state $|\Psi\rangle$ in the Schmidt basis, in which the state is invariant under rotations around the z axis.

For a classical system, that is, one which obeys a local hidden variable theory, it is well known that the correlation measure is bounded by the CHSH inequality

$$|\mathcal{S}| \leqslant 2. \tag{8}$$

This bound can be violated by an entangled quantum mechanical state, and in fact, for any pure two-qubit quantum state the

correlation measure $\mathcal{S}$ is bounded [22,23] by

$$|\mathcal{S}| \leqslant 2\sqrt{2}, \tag{9}$$

which is usually referred to as the Tsirelson bound.

In order to reach this bound, we consider the maximally entangled state; that is, we set $\mathcal{C} = 1$ in Eq. (1), leading to the singlet state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}} (|0\rangle_A |1\rangle_B - |1\rangle_A |0\rangle_B), \tag{10}$$

and further choose the measurement vectors such that $\boldsymbol{q} \perp \boldsymbol{r}$ and $\boldsymbol{s} \equiv -(\boldsymbol{q} + \boldsymbol{r})/\sqrt{2}$ as well as $\boldsymbol{t} \equiv (\boldsymbol{q} - \boldsymbol{r})/\sqrt{2}$. Inserting these relations into Eq. (5), and using that for $\mathcal{C} = 1$ the correlation matrix $K = -\mathbb{1}$ is proportional to unity, we find

$$\mathcal{S} = \frac{2|\boldsymbol{q}|^2}{\sqrt{2}} + \frac{2|\boldsymbol{r}|^2}{\sqrt{2}} = 2\sqrt{2}, \tag{11}$$

that is, the Tsirelson bound, where we used that $\boldsymbol{q}$ and $\boldsymbol{r}$ are unit vectors. Note that by no means the conditions to maximally violate the CHSH inequality determine the measurement direction.

For a nonmaximally entangled state, that is, for $\mathcal{C} < 1$, the Tsirelson bound can no longer be obtained. Furthermore, the measurement vectors that maximize $|\mathcal{S}|$ in general are no longer the same as the ones that correspond to the Tsirelson bound. In the following, we will discuss the optimal measurement strategy to obtain the maximum of $|\mathcal{S}|$, for a nonmaximally entangled state, that is, a state of the form of Eq. (1) with $\mathcal{C} < 1$.

III. GEOMETRICAL INTERPRETATION

The expectation value $\mathcal{S}$ given by Eq. (5) has a geometrical interpretation shown in Fig. 1. The vectors $\boldsymbol{q}$ and $\boldsymbol{r}$ end on the Bloch sphere of the subsystem A, while the vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ terminate on the Bloch sphere of the subsystem B. The expectation value $\mathcal{S}$ contains all four possible pairwise combinations of one vector from subsystem A and one vector from the subsystem B. The correlation matrix K defines these pairwise combinations; that is, it defines how a vector from one subsystem is multiplied by a vector of the other subsystem. In a sense, it connects the two Bloch spheres with each other.
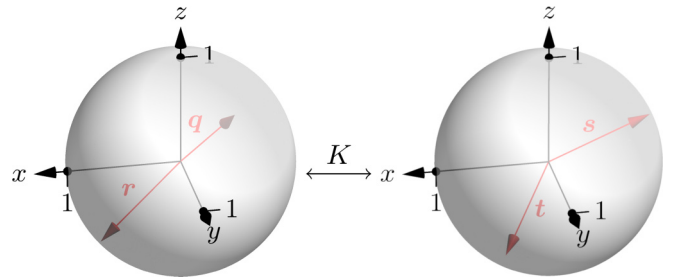

FIG. 1. Geometric interpretation of the expectation value $\mathcal{S}$ of the CHSH inequality in terms of Eq. (5). The two pairs of vectors $\boldsymbol{q}$ and $\boldsymbol{r}$ as well as $\boldsymbol{s}$ and $\boldsymbol{t}$ are unit vectors which end on the Bloch spheres of the subsystems A and B. The correlation matrix K establishes the connection between the vectors on the two Bloch spheres.
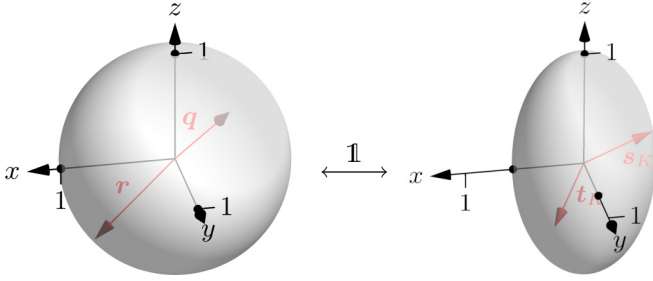
FIG. 2. Elimination of the correlation matrix K in favor of the new vectors s_K and t_K and geometrical interpretation of $\mathcal{S}$ according to Eq. (14). Now q and r are still unit vectors; that is, they still terminate on the Bloch sphere of the subsystem A. However, due to the matrix K the new vectors s_K and t_K are no longer unit vectors, but vectors pointing to the surface of a prolate spheroid with the semiminor axis given by the concurrence $\mathcal{C}$. The Bloch sphere of subsystem A and the prolate spheroid of subsystem B are now connected by the identity matrix $\mathbb{1}$.

Now we multiply the correlation matrix K by s and t of the subsystem B, leading to new vectors

$$s_K \equiv K s \qquad (12)$$

and

$$t_K \equiv K t, \qquad (13)$$

which are no longer unit vectors. They do not end on the unit sphere anymore, but rather on a prolate spheroid with semiaxes of length $\mathcal{C}$ and 1. This interpretation is depicted in Fig. 2.

When we insert the new vectors into Eq. (5) for $\mathcal{S}$ and regroup the four terms, the expectation value

$$\mathcal{S} = q \cdot (s_K - t_K) + r \cdot (s_K + t_K) \qquad (14)$$

is now only a combination of the scalar products between the vectors q and r from subsystem A, and $s_K + t_K$ and $s_K - t_K$ of subsystem B.

We can simplify Eq. (14) as

$$\mathcal{S} = |q| |s_K - t_K| \cos \alpha + |r| |s_K + t_K| \cos \beta, \qquad (15)$$

where α and β are the angles between q and $s_K - t_K$ and between r and $s_K + t_K$, respectively. Since both q and r are unit vectors, we find

$$\mathcal{S} = |s_K - t_K| \cos \alpha + |s_K + t_K| \cos \beta. \qquad (16)$$

The vectors q and r can be chosen independently of each other and of the vectors s_K and t_K. Therefore, the angles α and β are also independent of each other and we can easily maximize over the vectors q and r by choosing them such that either $\alpha = \beta = 0$ or $\alpha = \beta = \pi$.

Therefore, we maximize $\mathcal{S}$ by choosing q and r with s_K and t_K

$$q \equiv \pm \frac{s_K - t_K}{|s_K - t_K|} \qquad (17)$$

and

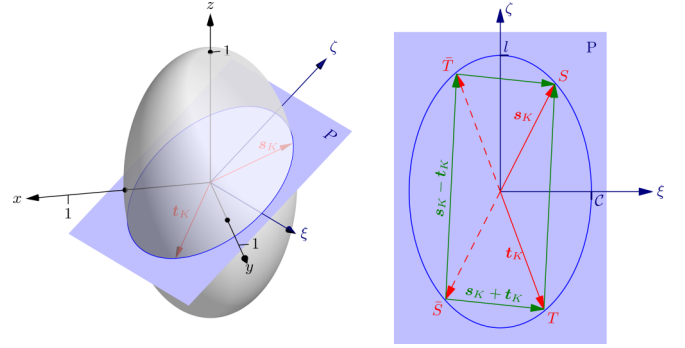$$r \equiv \pm \frac{s_K + t_K}{|s_K + t_K|}, \qquad (18)$$

FIG. 3. Geometrical interpretation of $\mathcal{S}$ as the circumference of a parallelogram enclosed by an ellipse. We only consider the subsystem B of Fig. 2 represented on the left by the prolate spheroid. The vectors s_K and t_K span a plane P which cuts the spheroid in an ellipse with a semiminor axis of length $\mathcal{C}$ and a semimajor axis of length l, with $\mathcal{C} \leqslant l \leqslant 1$, depending on the relative orientation of the plane to the ellipsoid. The vectors s_K and t_K, which terminate on this ellipse, together with the vectors $s_K + t_K$ and $s_K - t_K$, which are also in P, form a parallelogram whose corners $S, T, \bar{S}, \bar{T}$ lie on the ellipse. The task of maximizing the expectation value $\mathcal{S}_m(s_K, t_K)$ with respect to s_K and t_K is therefore equivalent to maximizing the perimeter of the parallelogram.

with which we find from Eq. (14) the expression

$$\mathcal{S}_m(s_K, t_K) \equiv \max_{q, r} |\mathcal{S}| = |s_K - t_K| + |s_K + t_K|, \qquad (19)$$

which only depends on vectors associated with the measurements performed on the subsystem B and the correlation matrix K. The optimization problem is thereby reduced to optimizing the two vectors s_K and t_K on a spheroid.

In general, s_K and t_K span a two-dimensional plane P, as shown in Fig. 3. This plane contains both the center of the spheroid as well as the linear combinations $s_K + t_K$ and $s_K - t_K$. The intersection between the plane and the spheroid therefore is an ellipse. Since the intersection contains the origin of the spheroid, the resulting ellipse has a semiminor axis of length $\mathcal{C}$ and a semimajor axis of length l, with $\mathcal{C} \leqslant l \leqslant 1$.

Indeed, the exact value of the semimajor axis l depends on the relative orientation of the plane to the spheroid. If, for example, P is the x-y plane, the intersection defines a circle with radius $\mathcal{C}$, and we have $l = \mathcal{C}$. On the other hand, if P contains the z axis, we have $l = 1$. In all other cases, l is in between these two extreme cases.

We now focus on the ellipse in the two-dimensional plane P as depicted on the right-hand side of Fig. 3. The vectors s_K and t_K touch the ellipse in the points S and T, respectively. The same is true for $-s_K$ and $-t_K$, and we call these points $\bar{S}$ and $\bar{T}$. Moreover, T and S, as well as $\bar{S}$ and $\bar{T}$, are connected by $s_K - t_K$. Similarly, S and $\bar{T}$, as well as $\bar{S}$ and T, are connected by $s_K + t_K$. Two sides of the parallelogram formed by the four points $S, T, \bar{S},$ and $\bar{T}$ have the length $|s_K - t_K|$ and two have $|s_K + t_K|$. The perimeter of the resulting parallelogram is thus $2\mathcal{S}_m$. Maximizing the expectation value $\mathcal{S}_m$ is therefore equivalent to finding the parallelogram enclosed by the ellipse with the maximal perimeter.

We note that the perimeter of the parallelogram increases when we increase the semiaxes of the ellipse. For a given

state with a fixed concurrence $\mathcal{C}$, the expectation value $\mathcal{S}_m$ is maximal, when the semimajor axis is as long as possible, that is, for $l = 1$. Throughout the remainder of our article, we set $l = 1$. As discussed before, this choice implies that the plane P contains the z axis.

IV. OPTIMIZATION OF THE MEASUREMENT STRATEGY

We now determine measurement strategies that maximize the expectation value $\mathcal{S}_m$. For this purpose, we first determine the angle between the vectors $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ that maximizes $\mathcal{S}_m$. We then address the orientations of the vectors $\boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}$, and $\boldsymbol{t}$ relative to each other.

A. A continuum of optimal angles

We start by expressing the three-dimensional unit vectors

$$\boldsymbol{s} = \begin{pmatrix} \cos\theta\cos\varphi_s \\ \sin\theta\cos\varphi_s \\ \sin\varphi_s \end{pmatrix} \tag{20}$$

and

$$\boldsymbol{t} = \begin{pmatrix} \cos\theta\cos\varphi_t \\ \sin\theta\cos\varphi_t \\ \sin\varphi_t \end{pmatrix} \tag{21}$$

in spherical coordinates with the polar angles φ_s and φ_t, as well as the azimuth θ, which is the same for both vectors, since they span the plane P which contains the z axis. Note that we have chosen the angles such that $0 \leqslant \varphi_s, \varphi_t < 2\pi$ and $0 \leqslant \theta < \pi$.

When we insert these representations into the definitions, Eqs. (12) and (13), of $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ as well as apply the explicit form, Eq. (7), of the correlation matrix K we find

$$\boldsymbol{s}_K = -\begin{pmatrix} \mathcal{C}\cos\theta\cos\varphi_s \\ \mathcal{C}\sin\theta\cos\varphi_s \\ \sin\varphi_s \end{pmatrix} \tag{22}$$

and

$$\boldsymbol{t}_K = -\begin{pmatrix} \mathcal{C}\cos\theta\cos\varphi_t \\ \mathcal{C}\sin\theta\cos\varphi_t \\ \sin\varphi_t \end{pmatrix}. \tag{23}$$

Throughout the remainder of this section, we choose a coordinate system in the two-dimensional plane P defined by the variables ξ and ζ, such that the ξ axis points along the semiminor axis, which lies in the x-y plane, and the ζ axis along the semimajor axis of the ellipse; that is, the ζ axis coincides with the z axis of the three-dimensional coordinate system.

In this two-dimensional system, we find the vector representations

$$\boldsymbol{s}_K = -\begin{pmatrix} \mathcal{C}\cos\varphi_s \\ \sin\varphi_s \end{pmatrix} \tag{24}$$

and

$$\boldsymbol{t}_K = -\begin{pmatrix} \mathcal{C}\cos\varphi_t \\ \sin\varphi_t, \end{pmatrix} \tag{25}$$

which only depend on the polar angles φ_s and φ_t of the vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ and the concurrence $\mathcal{C}$.

Inserting Eqs. (24) and (25) into Eq. (19), we find the expression

$$\begin{aligned} \mathcal{S}_m = &\sqrt{\mathcal{C}^2(\cos\varphi_s - \cos\varphi_t)^2 + (\sin\varphi_s - \sin\varphi_t)^2} \\ &+ \sqrt{\mathcal{C}^2(\cos\varphi_s + \cos\varphi_t)^2 + (\sin\varphi_s + \sin\varphi_t)^2} \end{aligned} \tag{26}$$

for the expectation value $\mathcal{S}_m$, as a function of φ_s and φ_t, which is already maximized over $\boldsymbol{q}$ and $\boldsymbol{r}$ by the choice made in Eqs. (17) and (18).

In Appendix A, we analytically maximize $\mathcal{S}_m$ and find the remarkable property that there exist not a *single* but *infinitely* many possible combinations of φ_s and φ_t, which can be parameterized by the average angle

$$\gamma \equiv (\varphi_s + \varphi_t)/2 \tag{27}$$

as

$$\varphi_s(\gamma) = \gamma + \delta \tag{28}$$

and

$$\varphi_t(\gamma) = \gamma - \delta \tag{29}$$

with the difference angle

$$\delta = \text{arccot}\sqrt{\frac{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}}. \tag{30}$$

This result implies that there exist *infinitely* many parallelograms which are inscribed by an ellipse with identical perimeters. This interesting property of ellipses and parallelograms has been shown before in Refs. [24,25], where also a geometric proof has been given.

In Fig. 4, we show the dependence of the optimal angles φ_s and φ_t on γ for three different values of $\mathcal{C}$, based on Eqs. (28) and (29).

For $\mathcal{C} = 1$, we get from Eqs. (28) and (29)

$$\varphi_s(\gamma) = \gamma + \frac{\pi}{4} \tag{31}$$

and

$$\varphi_t(\gamma) = \gamma - \frac{\pi}{4}. \tag{32}$$

Hence, the only requirement to maximize the expectation value $\mathcal{S}$ is that the two vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ are perpendicular, which is shown in Fig. 4. This result is in agreement with the literature [7].

For $\mathcal{C} < 1$, $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ are no longer perpendicular for all average angles γ. Instead, for $\gamma < \pi/4$ or $\gamma > 3\pi/4$, the angle between $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ is obtuse, increasing with decreasing concurrence, while for $\pi/4 < \gamma < 3\pi/4$ the angle between $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ is acute and decreases with decreasing concurrence.

In the limit of $\mathcal{C} = 0$, for $\gamma < \pi/2$ the angle $\varphi_s = \pi/2$ is constant. From the definition, Eq. (27), of the average angle γ it directly follows that $\varphi_t = 2\gamma - \pi/2$ depends linearly on γ. Since we have chosen $\varphi_s \geqslant \varphi_t$, for $\gamma > \pi/2$ the angle $\varphi_t = \pi/2$ is constant and $\varphi_s = 2\gamma - \pi/2$ depends linearly on γ.
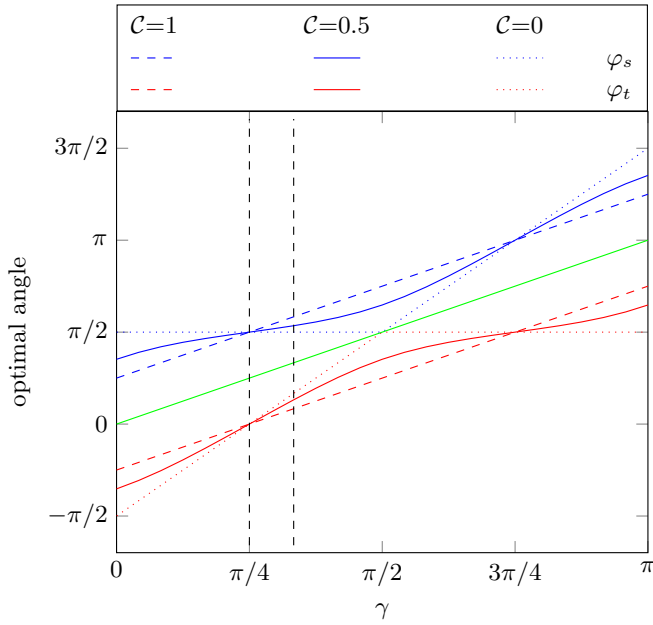
FIG. 4. Dependence of the optimal polar angles φ_s and φ_t on the angle γ maximizing the expectation value $\mathcal{S}$ for the three different concurrences $\mathcal{C} = 1$ (dashed), $\mathcal{C} = 0.5$ (solid), and $\mathcal{C} = 0$ (dotted). For a maximally entangled state (dashed parallel lines), the difference between the two angles is always $\varphi_s - \varphi_t = \pi/2$; that is, the measurements are orthogonal on each other. For a nonmaximally entangled state (solid lines), the difference between the optimal angles depends on γ. For a separable state (dotted lines), the difference between optimal angles is either $\varphi_s = \pi/2$ or $\varphi_t = \pi/2$. The average $(\varphi_s + \varphi_t)/2$ of the two angles is γ as given by Eqs. (28) and (29) and shown by the green solid line. The vertical dashed lines mark the special values of γ used in Fig. 7.

B. A parallelogram enclosed in an ellipse

We now use the optimal angles φ_s and φ_t to determine the value of the maximum of the correlation measure $\mathcal{S}$ and interpret the results in terms of our geometrical picture.

When we substitute the expression for φ_s and φ_t, Eqs. (28) and (29), into Eq. (26) we find that the expectation value $\mathcal{S}_m$ assumes its maximal value

$$\max |\mathcal{S}| = \max_{s_K, t_K} \mathcal{S}_m(s_K, t_K) = 2\sqrt{1 + \mathcal{C}^2}. \qquad (33)$$

In the case of $\mathcal{C} = 1$, that is for a maximally entangled state, Eq. (33) reduces to the well-known result [22]

$$\max |\mathcal{S}| = 2\sqrt{2} \qquad (34)$$

and for $\mathcal{C} = 0$, we get the classical bound [7]

$$\max |\mathcal{S}| = 2. \qquad (35)$$

Note that however small the entanglement of the pure state is, with the suitable measurement one can in principle violate the CHSH inequality.

We now discuss our results in the geometrical picture of a parallelogram inscribed by an ellipse. We show this behavior for three different concurrences $\mathcal{C}$ in Fig. 5.

For $\mathcal{C} = 1$, the ellipse reduces to a unit circle, as depicted in Fig. 5(a). In this case, the parallelogram assumes its maximal perimeter when it describes a square whose sides have length $\sqrt{2}$. Since the circle is invariant under rotations about its center, the square can also be rotated about this point without changing its perimeter. Therefore, all squares inscribed by the circle are possible solutions.

For $0 < \mathcal{C} < 1$ the geometry is more complicated, since there is no rotational symmetry anymore. The parallelograms associated with the optimal measurement strategy are no longer squares and not even necessarily rectangles. In Fig. 5(b), we present two parallelograms for maximizing the perimeter.

In the first example, shown in green, the vectors s_K and t_K point along the coordinate axes ξ and ζ; that is, the parallelogram connects the semiminor and the semimajor axes of the ellipse and becomes a rhombus whose perimeter is $4\sqrt{1 + \mathcal{C}^2}$.

The second realization depicted in Fig. 5(b) in orange is the optimal rectangle. This shape is optimal, when the tangents
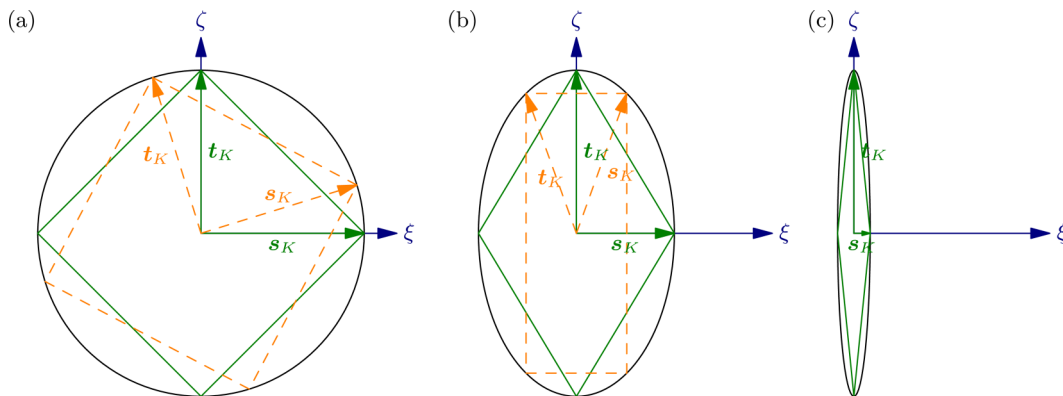

FIG. 5. Expectation value $\mathcal{S}$ as the perimeter of a parallelogram for different concurrences $\mathcal{C}$ based on Eq. (15). (a) For a maximally entangled state, that is, for $\mathcal{C} = 1$, the ellipse reduces to a unit circle and the parallelogram with the maximal perimeter is a square with edge length $\sqrt{2}$. Due to the rotational symmetry of the circle, the green square can be rotated inside the circle without changing the perimeter as exemplified by the dashed orange square. (b) For $0 < \mathcal{C} < 1$, the shape of the optimal parallelogram depends on the direction of one of the vectors s_K or t_K. Despite the different shapes, the perimeters are identical. (c) As $\mathcal{C}$ approaches zero, the ellipse becomes more elongated along the ζ axis. The optimal measurement is achieved for either s_K or t_K being almost parallel to this axis and the other one orthogonal. In the limit of $\mathcal{C} = 0$, the parallelogram reduces to a line along the ζ axis.

on the ellipse at the endpoints of s_K and t_K are perpendicular to each other [25]. This orthogonality condition is true for all possible solutions, as we show in Appendix B, and thus allows a geometric construction of all possible optimal parallelograms. We note that all other solutions in general have less symmetry and can be found from Eqs. (28) and (29).

As $\mathcal{C}$ approaches 0, leading to a separable state $|\Psi\rangle$, the ellipse collapses to the ζ axis, as suggested in Fig. 5(c). The parallelogram therefore is maximal if one of the two vectors s_K and t_K coincides with the semimajor axis, while the other vector is orthogonal to it. The parallelogram degenerates into a line between $\zeta = -1$ and $\zeta = 1$ on the ζ axis. It therefore has a perimeter of $2[1 - (-1)] = 4$, leading to max $\mathcal{S}_m = 2$.

C. Optimal measurement vectors

In the previous section, we have derived the optimal two-dimensional vectors s_K and t_K. We now construct the four optimal three-dimensional measurement vectors q, r, s, and t.

For this purpose, we obtain the vectors s and t by inserting the optimal angles φ_s and φ_t given by Eqs. (28) and (29) into the spherical representations, Eqs. (20) and (21), of s and t. They represent a family of measurement directions on the Bloch sphere of subsystem B parameterized by the average angle γ and the free parameter θ.

Next, we recall that the optimal measurement vectors q and r on the subsystem A are given by Eqs. (17) and (18). When we insert the explicit expressions, Eqs. (22) and (23), of s_K and t_K, together with the optimal angles, Eqs. (28) and (29), into the definitions, Eqs. (17) and (18), of q and r we find the representations

$$q(\gamma) = \mp \frac{1}{\sqrt{\mathcal{C}^2 \sin^2\gamma + \cos^2\gamma}} \begin{pmatrix} -\mathcal{C}\cos\theta\sin\gamma \\ -\mathcal{C}\sin\theta\sin\gamma \\ \cos\gamma \end{pmatrix} \quad (36)$$

and

$$r(\gamma) = \mp \frac{1}{\sqrt{\mathcal{C}^2 \cos^2\gamma + \sin^2\gamma}} \begin{pmatrix} \mathcal{C}\cos\theta\cos\gamma \\ \mathcal{C}\sin\theta\cos\gamma \\ \sin\gamma \end{pmatrix}. \quad (37)$$

Note that when, for instance, $q \parallel s$, the actual measurement directions are not necessarily parallel, as in the Schmidt representation of a given state $|\Psi\rangle$ the local three-dimensional coordinate systems in general are rotated relative to each other.

Since q and r are by definition unit vectors, we can rewrite them in spherical coordinates as

$$q(\gamma) = \mp \begin{pmatrix} \cos\theta\cos\varphi_q(\gamma) \\ \sin\theta\cos\varphi_q(\gamma) \\ \sin\varphi_q(\gamma) \end{pmatrix}, \quad r(\gamma) = \mp \begin{pmatrix} \cos\theta\cos\varphi_r(\gamma) \\ \sin\theta\cos\varphi_r(\gamma) \\ \sin\varphi_r(\gamma) \end{pmatrix}$$

$$(38)$$

with the γ-dependent polar angles

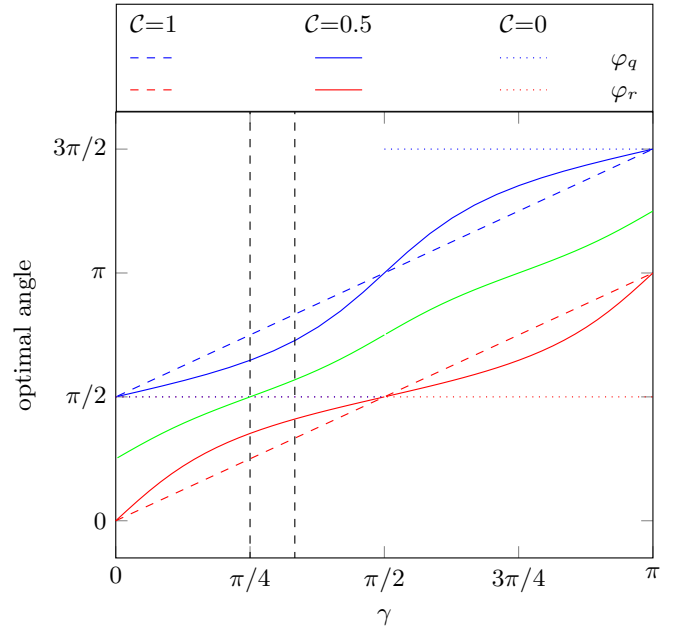$$\varphi_q(\gamma) \equiv \arctan\left(-\frac{1}{\mathcal{C}}\cot\gamma\right) + \pi \quad (39)$$

FIG. 6. Dependence of the optimal polar angles φ_q and φ_r on the angle γ maximizing the expectation value $\mathcal{S}$ for three different values of the concurrence $\mathcal{C}$. For a maximally entangled state, that is, $\mathcal{C} = 1$ (dashed lines), the difference between the two angles is always $\pi/2$. For $\mathcal{C} < 1$, the difference between the angles φ_q and φ_r decreases in the regime of $\gamma < \pi/2$ and increases for $\gamma > \pi/2$. In contrast to Fig. 4, the average $(\varphi_q + \varphi_r)/2$ of the two angles does not linearly depend on γ for $\mathcal{C} = 0.5$ (green solid line). The vertical dashed lines mark the special values of γ used in Fig. 7.

and

$$\varphi_r(\gamma) \equiv \arctan\left(\frac{1}{\mathcal{C}}\tan\gamma\right) + \Theta\left(\gamma - \frac{\pi}{2}\right)\pi. \quad (40)$$

Here, Θ denotes the Heaviside step function and we assume $-\pi/2 \leqslant \arctan x \leqslant \pi/2$.

The constant π in Eqs. (39) and (40) ensures that Eqs. (36) and (38), as well as Eq. (37) describe the same vectors, respectively. In Fig. 6, we present the dependence of φ_q and φ_r on the parameter γ. For $0 \leqslant \gamma < \pi/2$, the difference between the two angles φ_q and φ_r *decreases* with decreasing concurrence $\mathcal{C}$. In the regime of $\pi/2 < \gamma < \pi$, the angle difference *increases* with a decreasing entanglement.

When we compare this dependence of the angles φ_q and φ_r on γ, with the dependence of φ_s and φ_t on the same parameter, as shown in Fig. 4, we see that their dependence on the concurrence is shifted by an angle $\pi/4$. Furthermore, while γ is the angle bisection of the angle between s and t, this is not true for the angle between q and r. Indeed, the green line of Fig. 6 depends on γ.

D. Special examples

We conclude our discussions by considering three different examples that maximize the expectation value $\mathcal{S}$. For simplicity, we only consider the positive solutions of the vectors q and r in Eq. (38).

In Fig. 7, we depict the optimal vectors q, r, s, and t for these three different choices of the parameter γ. Here, we
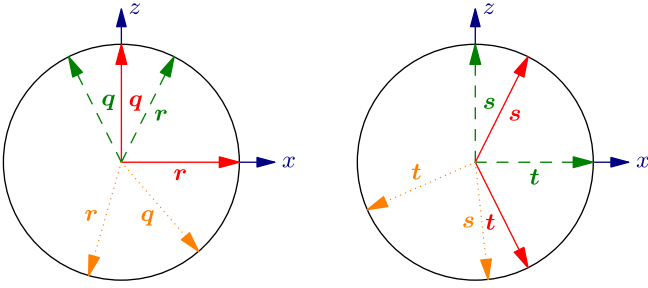
FIG. 7. Optimal measurement vectors q, r, s, and t for three different values of γ, but for a fixed concurrence $\mathcal{C} = 0.5$. For $\gamma = 0$ (red solid vectors), q and r are orthogonal, as indicated by Fig. 6, while the angle between the vectors s and t is maximal, as depicted in Fig. 4. In the case of $\gamma = \pi/4$ (green dashed vectors), s and t are orthogonal, while the angle between q and r is minimal. For $\gamma = 4\pi/3$ (orange dotted), the vectors in neither subsystem are orthogonal.

always choose a fixed azimuth $\theta = 0$, and hence the vectors lie in the x-z plane. In the following, we therefore provide only the x and z components of the vectors and suppress the y component, which always vanishes.

The most elementary example is $\gamma = 0$, where the vectors read

$$q = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad r = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{41}$$

and

$$s = \frac{1}{\sqrt{1+\mathcal{C}^2}} \begin{pmatrix} \mathcal{C} \\ 1 \end{pmatrix}, \quad t = \frac{1}{\sqrt{1+\mathcal{C}^2}} \begin{pmatrix} \mathcal{C} \\ -1 \end{pmatrix}. \tag{42}$$

This result shows again that the relative angle between the measurements depends on the entanglement on the state. For $\mathcal{C} = 1$, it reduces to the well-known settings [7], but for $\mathcal{C} < 1$, the angle 2δ between the vectors s and t is implicitly given by

$$\cos 2\delta = -\frac{1-\mathcal{C}^2}{1+\mathcal{C}^2}, \tag{43}$$

which is negative since $0 < \mathcal{C} < 1$. Therefore, this angle is always greater than $\pi/2$ and increases with decreasing concurrence $\mathcal{C}$.

A similar situation is found for $\gamma = \pi/4$. Here, the vectors read

$$q = \frac{1}{\sqrt{1+\mathcal{C}^2}} \begin{pmatrix} -\mathcal{C} \\ 1 \end{pmatrix}, \quad r = \frac{1}{\sqrt{1+\mathcal{C}^2}} \begin{pmatrix} \mathcal{C} \\ 1 \end{pmatrix}, \tag{44}$$

and

$$s = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \tag{45}$$

that is, now s and t of the subsystem B are orthogonal, while the angle η between the vectors q and r is given by

$$\cos \eta = \frac{1-\mathcal{C}^2}{1+\mathcal{C}^2}. \tag{46}$$

In contrast to the previous example $\gamma = 0$, this angle is smaller than $\pi/2$ for $\mathcal{C} < 1$ and decreases with decreasing concurrence $\mathcal{C}$.

The last example in Fig. 7 is $\gamma = 4\pi/3$, for which we obtain

$$q = \frac{1}{\sqrt{1+3\mathcal{C}^2}} \begin{pmatrix} \sqrt{3}\mathcal{C} \\ -1 \end{pmatrix}, \quad r = \frac{1}{\sqrt{3+\mathcal{C}^2}} \begin{pmatrix} -\mathcal{C} \\ -\sqrt{3} \end{pmatrix}, \tag{47}$$

for the subsystem A and

$$s = \begin{pmatrix} \cos\left(\frac{4\pi}{3} + \phi_0\right) \\ \sin\left(\frac{4\pi}{3} + \phi_0\right) \end{pmatrix}, \quad t = \begin{pmatrix} \cos\left(\frac{4\pi}{3} - \phi_0\right) \\ \sin\left(\frac{4\pi}{3} - \phi_0\right) \end{pmatrix} \tag{48}$$

for the subsystem B with the concurrence-dependent phase

$$\phi_0 = \arctan\sqrt{\frac{1+3\mathcal{C}^2}{3+\mathcal{C}^2}}, \tag{49}$$

which in general has to be calculated numerically.

In this case, neither the vectors q and r of the subsystem A nor the vectors s and t of B are orthogonal for $\mathcal{C} < 1$. This example illustrates the existence of solutions that are not easy to guess.

V. MIXED STATES

So far, we only considered the optimization of the correlation measure (4) for pure two-qubit states. We turn to the problem of mixed two-qubit states.

In order to do so, we generalize the correlation matrix K, Eq. (6), defined for pure states by the expression

$$K \equiv \text{tr}(\hat{\rho}_{AB} \cdot \hat{\sigma}_A \otimes \hat{\sigma}_B) \tag{50}$$

to two-qubit mixed states $\hat{\rho}_{AB}$.

For pure states, we demonstrated in Sec. III that the maximization of the expectation value $\mathcal{S}$ has a geometrical interpretation as the maximization of the circumference of a parallelogram inscribed by an ellipse. The deeper origin of the ellipse is the transformation of the Bloch sphere of one subsystem by the correlation matrix K into a prolate spheroid. This transformation directly follows from the diagonal form of the correlation matrix, Eq. (7), of any pure state.

The diagonal form of the correlation matrix for pure states was a consequence of writing the state in the Schmidt basis. Unfortunately, we cannot generalize this property directly to the mixed-state case. The question now is, if such a geometrical interpretation is still possible for the mixed two-qubit states, since it is not clear if the mixed state again transforms the Bloch sphere into a spheroid and our geometrical argument still holds for mixed states.

In this section, we first show that any mixed state in fact transforms the Bloch sphere of subsystem B into an ellipsoid, which in general is less symmetric than the prolate spheroid of the pure state case. We then exploit this result to show that a similar geometrical interpretation exists as for the pure state case, and derive the corresponding optimal angles and maximal expectation value $\mathcal{S}$. We conclude this section by considering two examples.

A. Transformed Bloch sphere: Still an ellipsoid

We show in the following that the correlation matrix K transforms the Bloch sphere of subsystem B into an ellipsoid with three in general different semiaxes a, b, and c, with

$a \geqslant b \geqslant c$. In addition, the orientation of the ellipsoid with respect to our coordinate system and the lengths a, b, and c depend on the state $\hat{\rho}_{AB}$.

We start by considering an arbitrary vector $\boldsymbol{v}$ which points from the origin to the surface of the Bloch sphere of subsystem B. As a consequence, the vector $\boldsymbol{v}$ is a unit vector. The correlation matrix K then defines a new vector

$$\boldsymbol{v}_K \equiv K\boldsymbol{v}. \qquad (51)$$

For any correlation matrix K, there exists a singular value decomposition [20]

$$K = U \Sigma V^T \qquad (52)$$

with two orthogonal matrices U and V and the diagonal matrix

$$\Sigma = \text{diag}(a, b, c), \qquad (53)$$

such that $a \geqslant b \geqslant c \geqslant 0$.

We note that since $|\det K| = abc$ all elements of the diagonal matrix Σ are strictly positive if the inverse of K exists, while at least one element vanishes if K is not invertible.

For simplicity, we choose $a, b, c > 0$, that is, all singular values to be positive. The case of at least one vanishing singular value can be discussed analogously.

When we insert the singular value decomposition into the definition, Eq. (51), of $\boldsymbol{v}_K$ we obtain

$$\boldsymbol{v}_K = U \Sigma V^T \boldsymbol{v} \qquad (54)$$

for the new vector.

We can interpret the action of the three matrices $U \Sigma V^T$ as follows: First, we rotate the coordinate system by the orthogonal transformation V. Then the matrix Σ deforms the Bloch sphere along the axes of the new coordinate system by scaling it with the respective singular value. Finally, the orthogonal matrix U rotates the new shape.

When we introduce the abbreviations $\boldsymbol{v}'_K \equiv U^T \boldsymbol{v}_K$ and $\boldsymbol{v}' \equiv V^T \boldsymbol{v}$, that is, directly perform the rotations of the two orthogonal matrices on the two vectors $\boldsymbol{v}_K$ and $\boldsymbol{v}$, respectively, we get

$$\boldsymbol{v}'_K = \Sigma \boldsymbol{v}'. \qquad (55)$$

As a next step, we introduce the inverse

$$\Sigma^{-1} = \text{diag}(a^{-1}, b^{-1}, c^{-1}) \qquad (56)$$

of the diagonal matrix Σ, which leads us to

$$\Sigma^{-1} \boldsymbol{v}'_K = \boldsymbol{v}'. \qquad (57)$$

When we take the absolute square of the vectors on both sides, we find

$$\boldsymbol{v}'^T_K (\Sigma^{-1})^T \Sigma^{-1} \boldsymbol{v}'_K = 1. \qquad (58)$$

Here, we made use of the norm $|\boldsymbol{v}'|^2 = |\boldsymbol{v}|^2 = 1$, which is a direct consequence of $\boldsymbol{v}$ being a unit vector by definition, as it points from the center of the Bloch sphere to its surface, and the fact that orthogonal matrices do not change the norm, such that $\boldsymbol{v}'$ is also a unit vector.

When we further use the relation

$$(\Sigma^{-1})^T \Sigma^{-1} = \text{diag}(a^{-2}, b^{-2}, c^{-2}), \qquad (59)$$

we finally arrive at

$$\left(\frac{v'_{K,1}}{a}\right)^2 + \left(\frac{v'_{K,1}}{b}\right)^2 + \left(\frac{v'_{K,3}}{c}\right)^2 = 1, \qquad (60)$$

which is the equation of an ellipsoid with semi axes a, b, and c.

We finally relate this result back to the vectors $\boldsymbol{v} = V\boldsymbol{v}'$ and $\boldsymbol{v}_K = U\boldsymbol{v}'_K$. Since the two matrices U and V are both orthogonal, the vectors $\boldsymbol{v}$ and $\boldsymbol{v}_K$ are only rotations of the vectors $\boldsymbol{v}'$ and $\boldsymbol{v}'_K$, respectively. Therefore, any vector $\boldsymbol{v}$ pointing from the center of the Bloch sphere to its surface is mapped onto an ellipsoid with semiaxes of length a, b, and c. The orientation of this ellipsoid with respect to the original coordinate system is defined by the matrices U and V. Hence, the family of states with the same singular values a, b, and c transforms the Bloch sphere into ellipsoids of identical shape but different orientation.

In our derivation, we only considered all singular values to be nonvanishing. In full analogy to the above derivation, we can also consider the case of vanishing singular values. The only difference is that for vanishing singular values the inverse Σ^{-1}, Eq. (56), has to be replaced by a pseudoinverse. Consequently, the dimensions of the problem are reduced. As a result, for one vanishing singular value, instead of an ellipsoid, the Bloch sphere is transformed into an area enclosed by an ellipse, while for two vanishing singular values the Bloch sphere is transformed into a line. In the case of all three singular values being vanishing, it is trivial to see that the Bloch sphere is mapped onto the origin of the coordinate system, that is, onto a point.

B. Geometry and maximization

In the preceding section, we have shown that the correlation matrix transforms the Bloch sphere into an ellipsoid. The vectors $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ point from the origin to the surface of this ellipsoid.

The intersection of the plane P, spanned by the vectors $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$, and the new ellipsoid is still an ellipse. The optimal plane P, that is, the plane which leads to the ellipse with the largest semiaxes, creates an ellipse with semimajor axis a and semiminor axis b. We note that for a general ellipsoid there exists only one distinct plane P, which cuts the ellipsoid into this optimal ellipse. In contrast, for a prolate spheroid there exist infinitely many such planes, due to the rotational symmetry of the spheroid. Therefore, going from a pure state to a mixed state breaks this symmetry.

The parallelogram inscribed by the ellipse is solely a result of the measurement vectors and does not change when changing the state of the system. Therefore, we now have a similar geometrical interpretation of maximizing the circumference of a parallelogram, inscribed by an ellipse. The only difference relevant for the present discussion between the pure and the mixed state cases therefore lies in the different length of the semiaxes of the ellipse.

The optimization strategy is thus the same as for the pure state case, discussed in Sec. IV. The only difference is that instead of semiaxes of length one and $\mathcal{C}$, we find a and b. As a result, the optimal angles are still given by Eqs. (28) and (29); however, the difference angle δ, Eq. (30), has to change. We recall from the definition of δ, Eq. (30), that for an ellipse with

semimajor axis of length l and semiminor axis of length $\mathcal{C}$, the angle is given by

$$\delta = \text{arccot}\sqrt{\frac{\mathcal{C}^2 \cos^2 \gamma + l^2 \sin^2 \gamma}{\mathcal{C}^2 \sin^2 \gamma + l^2 \cos^2 \gamma}}, \qquad (61)$$

where we have originally set $l = 1$. The extension to an arbitrary value of l can be easily made from the argument that the optimal angles φ_s and φ_t, and thus the difference angle δ, depend only on the ratio $\mathcal{C}/l$ of the two semiaxes but not on their absolute values. Indeed, stretching both axes by a constant factor will not change the optimal angles but only the perimeter of the associated parallelogram.

Thus, when we make the substitution $\mathcal{C} \to b$ and $l \to a$ in the difference angle δ, Eq. (61), we find the new difference angle

$$\delta = \text{arccot}\sqrt{\frac{b^2 \cos^2 \gamma + a^2 \sin^2 \gamma}{b^2 \sin^2 \gamma + a^2 \cos^2 \gamma}}, \qquad (62)$$

for the family of optimal angles $\varphi_s = \gamma + \delta$ and $\varphi_t = \gamma - \delta$, parameterized by the average angle γ.

All that is left to do is to find the maximal expectation value $\mathcal{S}$. In order to do so, we recall from our discussion in Sec. IV that one of the optimal parallelograms is the rhombus whose corners are determined by the intersection of the ellipse and its semiaxes [confer the green parallelogram in Fig. 5(b)]. By using the Pythagorean theorem, it is straightforward to see that each edge of this rhombus has length $\sqrt{a^2 + b^2}$. Therefore, the maximal correlation measure is

$$\max |\mathcal{S}| = 2\sqrt{a^2 + b^2}, \qquad (63)$$

since it is half of the perimeter of the parallelogram.

We note that from Eq. (63) it immediately follows that any mixed state can violate the CHSH inequality, Eq. (8), if and only if $a^2 + b^2 > 1$. Our result is in full agreement with the criterion found in Ref. [26] on the violation of the CHSH inequality for any mixed two-qubit state $\hat{\rho}$.

C. Mixture of entangled states

In the following, we consider an example of mixed two-qubit states. We apply the general concepts developed in the previous paragraphs in order to deduce the optimal correlation measures $\mathcal{S}$ and the optimal measurements for this states.

We consider the mixture

$$\hat{\rho} \equiv p|\Psi\rangle\langle\Psi| + (1 - p)|\Phi\rangle\langle\Phi| \qquad (64)$$

of the pure stateyyyyy

$$|\Psi\rangle \equiv \sqrt{\frac{1 + \sqrt{1 - \mathcal{C}^2}}{2}}|0\rangle_A|1\rangle_B - \sqrt{\frac{1 - \sqrt{1 - \mathcal{C}^2}}{2}}|1\rangle_A|0\rangle_B$$
$$\qquad (65)$$

with concurrence $\mathcal{C}$, that occurs with probability p, and

$$|\Phi\rangle \equiv \sqrt{\frac{1 + \sqrt{1 - \mathcal{D}^2}}{2}}|0\rangle_A|0\rangle_B + \sqrt{\frac{1 - \sqrt{1 - \mathcal{D}^2}}{2}}|1\rangle_A|1\rangle_B$$
$$\qquad (66)$$

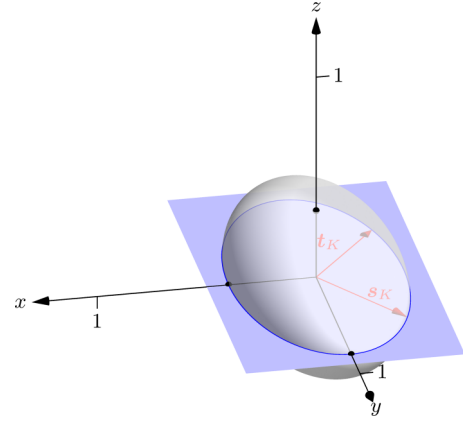with concurrence $\mathcal{D}$, which has probability $1 - p$.

FIG. 8. Geometrical interpretation of the correlation matrix K for the mixed state given by Eq. (64), for $p = 1/3$, $\mathcal{C} = 0.6$, and $\mathcal{D} = 0.9$. The correlation matrix is an ellipsoid with three different semiaxes along the x, y, and z coordinate, instead of a prolate spheroid of a pure state (cf. Fig. 3). The intersection of the spheroid with the plane P spanned by the vectors $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ is still an ellipse. The maximal correlation measure can be obtained when P cuts the ellipsoid in the ellipse with the largest semiaxes. In this case, the optimal plane P is the x-y plane.

By inserting the state ρ, Eq. (64), into the definition of the correlation matrix, Eq. (50), together with the definitions of the individual basis states, Eqs. (1) and (66), we find the correlation matrix

$$K = \begin{pmatrix} -p\mathcal{C} + (1 - p)\mathcal{D} & 0 & 0 \\ 0 & -p\mathcal{C} - (1 - p)\mathcal{D} & 0 \\ 0 & 0 & 1 - 2p \end{pmatrix}$$
$$\qquad (67)$$

of the mixed state $\hat{\rho}$, which is already diagonal. Thus, the transformation due to this matrix describes an ellipsoid with in general three different semiaxes that are oriented along the x, y, and z coordinates.

The lengths of these three semiaxes are the absolute values of the three diagonal entries of the correlation matrix, Eq. (67). We note that the three axes in general have different lengths. Furthermore, the semiaxis along the z coordinate is no longer of unit length, as for the pure state case. In fact, for $p = 1/2$ the axis even has vanishing length. This result originates from the different sign that the individual correlation matrices, Eq. (6), of the pure basis states $|\Psi\rangle$ and $|\Phi\rangle$ have in the zz component. In Fig. 8, we show the ellipsoid created by Eq. (67) for the choice $p = 1/3$, $\mathcal{C} = 0.6$, and $\mathcal{D} = 0.9$.

In order to find the maximal value of the correlation measure $\mathcal{S}$, we need to find the plane P spanned by the vectors $\boldsymbol{s}_K$ and $\boldsymbol{t}_K$ that maximizes the ellipse when cutting the ellipsoid.

Since $|p\mathcal{C} + (1 - p)\mathcal{D}| \geqslant |p\mathcal{C} - (1 - p)\mathcal{D}|$, the y axis of the ellipsoid is always longer than its x axis. Furthermore, if $|p\mathcal{C} + (1 - p)\mathcal{D}| \geqslant 1 - 2p$, the x axis is longer than the z axis, and the optimal plane is the x-y plane. This behavior is in contrast to the pure state case, where the optimal plane P always contained the z axis. Otherwise, if the z axis is longer than the x axis, the optimal plane is the y-z plane.

In the example depicted in Fig. 8, we find the optimal ellipse when the vectors s_K and t_K lie in the x-y plane. This ellipse has a semimajor axis $a = 0.8$ and $b = 0.4$. From Eq. (63), we immediately find the maximal expectation value $\mathcal{S} = 2\sqrt{4/5} < 2$. Therefore, this mixed state can never violate the CHSH inequality, even though it is composed of two entangled states.

We note that the optimal measurement vectors q, r, s, and t are found by the same techniques as for the pure state case, discussed in Sec. IV. In contrast to the pure state case, where we had the angle θ as a remaining degree of freedom, due to the rotational symmetry of the spheroid, for an ellipsoid the optimal plane is in general completely determined.

D. Entanglement and violation of the CHSH inequality for mixed states

As a second example of entangled two-qubit mixed states, we consider the famous Werner state [27]

$$\hat{\rho}_W \equiv p|\Psi^{(-)}\rangle\langle\Psi^{(-)}| + \frac{1-p}{4}\hat{\mathbb{1}}, \qquad (68)$$

which is a mixture of the maximally entangled singlet state $|\Psi^{(-)}\rangle$, as defined in Eq. (10), with probability p, and the completely mixed state $\hat{\mathbb{1}}/4$ in four dimensions with probability $1 - p$. When we insert Eq. (68) into the definition of the correlation matrix, Eq. (50), we obtain

$$K = \mathrm{diag}(-p, -p, -p) = -p\mathbb{1}, \qquad (69)$$

which transforms the Bloch sphere of subsystem B into a smaller sphere with radius p. This transformation is not surprising, since the correlation matrix for the singlet state $|\Psi^-\rangle$ does not change the shape of the Bloch sphere, while the completely mixed state transforms the Bloch sphere into a single point at the origin of the sphere. Thus, the convex combination of these two actions still preserves the rotational symmetry of the sphere but with a decreased radius.

When we apply our geometrical considerations, we see that the plane P cuts the sphere into a circle with radius p. The optimal measurement strategies are the same as for the maximally entangled pure state, and from Eq. (63), we directly obtain

$$\mathcal{S} \leqslant 2\sqrt{2}p \qquad (70)$$

for the maximal value of the correlation function. Hence, the Werner state allows to violate the CHSH inequality for $p > 1/\sqrt{2}$.

From the Peres-Horodecki positive partial transpose (PPT) criterion [28,29], it is known that the Werner state is nonseparable for $p > 1/3$. Hence, for $1/3 < p < 1/\sqrt{2}$ the state is entangled but cannot violate the CHSH inequality. This result is in full agreement with the literature [29], and in contrast to any pure state, which violates CHSH for an appropriate measurement if and only if it is entangled.

E. Nondiagonal correlation matrix

We conclude our discussion of the mixed state case with the remark that in both of our examples the correlation matrix K is already diagonal. For a general mixed state, an important step is to first find the diagonal representation of the matrix KK^T in order to obtain the principal axes of the ellipsoid. A nondiagonal matrix K leads to an ellipse whose semiaxes are not given by the coordinate axes anymore. Once these axes are known, the maximal expectation value $\mathcal{S}$ can easily be determined by Eq. (63), and the optimal measurement vectors are obtained by using the results of Sec. IV. We note that for a pure state this diagonalization is automatically done by using the Schmidt decomposition of the state.

VI. CONCLUSIONS AND OUTLOOK

In this article, we have demonstrated that the quantum mechanical expectation value $\mathcal{S}$ measuring the correlation in the CHSH inequality for a two-qubit entangled pure state can be interpreted geometrically as half of the perimeter of a parallelogram with its corners on an ellipse whose semiminor axis is determined by the entanglement of the pure bipartite state and whose circumference is maximal. The larger the entanglement characterized by the concurrence $\mathcal{C}$, the larger is the semiminor axis.

Moreover, this interpretation also allowed us to analytically find the optimal measurement directions for any pure entangled two-qubit state. We demonstrated that there exist infinitely many such directions.

Finally, we extended our considerations to mixed two-qubit states. The main difference compared to the pure state case are the resulting correlation matrices K, which in turn determine the curves enclosing the parallelogram. These curves remain ellipses, and the same geometrical interpretation as for pure states is possible. In contrast to the pure state case, the semimajor axes of the ellipse is no longer of unit length, and the semiminor axes is not determined by the concurrence anymore.

An remaining open question is whether our geometrical interpretation can also be applied to other types of Bell inequalities or similar equalities. In 1993, Hardy [30] introduced a test of nonclassicality of two qubits that exploits an equality instead of an inequality. Interestingly, this equality is applicable only for all nonmaximally entangled states, which indicates that a lack of symmetry in the underlying state is necessary. Therefore, it would be an interesting task to check whether a similar geometrical interpretation exists for this equality and how it brings out this symmetry argument.

APPENDIX A: ANALYTICAL DERIVATION OF THE MAXIMAL VIOLATION

In this Appendix, we analytically find the angles φ_s and φ_t which maximize the expectation value

$$
\mathcal{S}_m = \sqrt{\mathcal{C}^2(\cos\varphi_s - \cos\varphi_t)^2 + (\sin\varphi_s - \sin\varphi_t)^2} \\
+ \sqrt{\mathcal{C}^2(\cos\varphi_s + \cos\varphi_t)^2 + (\sin\varphi_s + \sin\varphi_t)^2}. \quad \text{(A1)}
$$

For this purpose, we introduce the average angle

$$
\gamma \equiv (\varphi_s + \varphi_t)/2 \quad \text{(A2)}
$$

and the difference angle

$$
\delta \equiv (\varphi_s - \varphi_t)/2, \quad \text{(A3)}
$$

yielding the expression

$$
\mathcal{S}_m = 2|\sin\delta|\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma} \\
+ 2|\cos\delta|\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}. \quad \text{(A4)}
$$

For the sake of simplicity, we restrict ourselves to the interval $0 \leqslant \delta \leqslant \pi/2$. This constraint can always be fulfilled, since

for $\delta > \pi/2$, we can simply exchange the two vectors s_K and t_K.

We are now able to maximize this expression by first calculating the gradient, that is, the derivatives

$$
\frac{\partial\mathcal{S}_m}{\partial\gamma} = 2\sin\delta\frac{\mathcal{C}^2\cos\gamma\sin\gamma - \sin\gamma\cos\gamma}{\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}} \\
- 2\cos\delta\frac{\mathcal{C}^2\sin\gamma\cos\gamma - \sin\gamma\cos\gamma}{\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}} = 0 \quad \text{(A5)}
$$

and

$$
\frac{\partial\mathcal{S}_m}{\partial\delta} = 2\cos\delta\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma} \\
- 2\sin\delta\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma} = 0 \quad \text{(A6)}
$$

of the expectation value and then equating both to zero.

In this way, we find from Eq. (A6) the relation

$$
\cos\delta = \sin\delta\frac{\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}}{\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}}, \quad \text{(A7)}
$$

which by inserting into Eq. (A5) leads us to

$$
\sin\delta\frac{\mathcal{C}^2\cos\gamma\sin\gamma - \sin\gamma\cos\gamma}{\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}} + \sin\delta\frac{\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}}{\sqrt{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}}\frac{-\mathcal{C}^2\sin\gamma\cos\gamma + \sin\gamma\cos\gamma}{\sqrt{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}} = 0. \quad \text{(A8)}
$$

Hence, a vanishing derivative of $\mathcal{S}_m$ with respect to δ also implies a vanishing derivative of $\mathcal{S}_m$ with respect to γ. In particular, Eq. (A8) is fulfilled for all γ and δ.

Therefore, for every γ we can choose according to Eq. (A7)

$$
\delta = \text{arccot}\sqrt{\frac{\mathcal{C}^2\cos^2\gamma + \sin^2\gamma}{\mathcal{C}^2\sin^2\gamma + \cos^2\gamma}}, \quad \text{(A9)}
$$

and the gradient of $\mathcal{S}_m$ vanishes.

When we insert this expression for δ into Eq. (A4), the expectation value becomes

$$
\max_{s,t}\mathcal{S}_m = 2\sqrt{1 + \mathcal{C}^2}, \quad \text{(A10)}
$$

independent of the parameter γ.

Hence, for every γ, we can find a parameter δ, such that the expectation value assumes its global maximum. We can finally invert the coordinate transformation, Eqs. (A2) and (A3), by $\varphi_s = \gamma + \delta$ and $\varphi_t = \gamma - \delta$, which together with Eq. (A9) defines the optimal measurements.

APPENDIX B: ORTHOGONALITY OF THE TANGENTS ON THE ELLIPSE

In this Appendix, we show that the optimal parallelograms in terms of maximizing the expectation value $\mathcal{S}$ are geometrically determined by the constraint that the tangents on the ellipse in the points S and T are orthogonal. In order to do so, we first note that the tangent vectors are determined by the

derivatives

$$
\frac{ds_K}{d\varphi_s} = \begin{pmatrix} \mathcal{C}\sin\varphi_s \\ -\cos\varphi_s \end{pmatrix} \quad \text{(B1)}
$$

and

$$
\frac{dt_K}{d\varphi_t} = \begin{pmatrix} \mathcal{C}\sin\varphi_t \\ -\cos\varphi_t \end{pmatrix} \quad \text{(B2)}
$$

of the vectors s_K and t_K, defined by Eqs. (24) and (25), respectively, leading to the scalar product

$$
\frac{ds_K}{d\varphi_s} \cdot \frac{dt_K}{d\varphi_t} = \mathcal{C}^2\sin\varphi_s\sin\varphi_t + \cos\varphi_s\cos\varphi_t. \quad \text{(B3)}
$$

When we insert the optimal angles $\varphi_s = \gamma + \delta$ and $\varphi_t = \gamma - \delta$, defined by Eqs. (28) and (29) into Eq. (B3), we find

$$
\frac{ds_K}{d\varphi_s} \cdot \frac{dt_K}{d\varphi_t} = \mathcal{C}^2\sin(\gamma + \delta)\sin(\gamma - \delta) \\
+ \cos(\gamma + \delta)\cos(\gamma - \delta), \quad \text{(B4)}
$$

which with the help of the trigonometric relations

$$
\cos(x + y)\cos(x - y) = \cos(2x) + \cos(2y) \quad \text{(B5)}
$$

and

$$
\sin(x + y)\sin(x - y) = \cos(2y) - \cos(2x), \quad \text{(B6)}
$$

simplifies to

$$
\frac{ds_K}{d\varphi_s} \cdot \frac{dt_K}{d\varphi_t} = \frac{1 + \mathcal{C}^2}{2}\cos(2\delta) + \frac{1 - \mathcal{C}^2}{2}\cos(2\gamma). \quad \text{(B7)}
$$

From the definition of δ, Eq. (30) and the relation

$$\cos(2\mathrm{arccot}x) = \frac{x^2 - 1}{x^2 + 1}, \tag{B8}$$

we find

$$\cos(2\delta) = \frac{\mathcal{C}^2(\cos^2\gamma - \sin^2\gamma) + (\sin^2\gamma - \cos^2\gamma)}{\mathcal{C}^2(\cos^2\gamma + \sin^2\gamma) + (\sin^2\gamma + \cos^2\gamma)}, \tag{B9}$$

which when we make use of the relations

$$\cos^2\gamma + \sin^2\gamma = 1 \tag{B10}$$

and

$$\cos^2\gamma - \sin^2\gamma = \cos(2\gamma) \tag{B11}$$

leads to

$$\cos(2\delta) = -\frac{1 - \mathcal{C}^2}{1 + \mathcal{C}^2}\cos(2\gamma). \tag{B12}$$

When we finally insert Eq. (B12) into the scalar product Eq. (B7), we find

$$\frac{d\boldsymbol{s}_K}{d\varphi_s} \cdot \frac{d\boldsymbol{t}_K}{d\varphi_t} = 0, \tag{B13}$$

proving the orthogonality of the two tangent vectors.

[1] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[2] N. Bohr, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **48**, 696 (1935).

[3] E. Schrödinger, Die gegenwärtige Situation in der Quantenmechanik, Naturwissenschaften **23**, 807 (1935).

[4] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics **1**, 195 (1964).

[5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[6] M. O. Scully and M. S. Zubairy, *Quantum Optics*, 6th ed. (Cambridge University Press, Cambridge, UK, 2008).

[7] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[8] S. J. Freedman and J. F. Clauser, Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. **28**, 938 (1972).

[9] E. S. Fry and R. C. Thompson, Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. **37**, 465 (1976).

[10] A. Aspect, J. Dalibard, and G. Roger, Experimental Test of Bell's Inequalities using Time-Varying Analyzers, Phys. Rev. Lett. **49**, 1804 (1982).

[11] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, Violation of Bell's Inequality Under Strict Einstein Locality Conditions, Phys. Rev. Lett. **81**, 5039 (1998).

[12] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán *et al.*, Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature (London) **526**, 682 (2015).

[13] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J. A. Larsson, C. Abellan *et al.*, Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. **115**, 250401 (2015).

[14] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman *et al.*, Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. **115**, 250402 (2015).

[15] S. Popescu and D. Rohrlich, Generic quantum nonlocality, Phys. Lett. A **166**, 293 (1992).

[16] N. Gisin, Bell's inequality holds for all non-product states, Phys. Lett. A **154**, 201 (1991).

[17] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[18] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature (London) **464**, 1021 (2010).

[19] A. Acín, S. Massar, and S. Pironio, Randomness Versus Nonlocality and Entanglement, Phys. Rev. Lett. **108**, 100402 (2012).

[20] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2001).

[21] J. Seiler, T. Strohm, and W. P. Schleich, Estimating the privacy of quantum-random numbers, New J. Phys. **22**, 093063 (2020).

[22] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[23] B. S. Tsirel'son, Quantum analogues of the Bell inequalities: The case of two spatially separated domains, J. Math. Sci. **36**, 557 (1987).

[24] J.-M. Richard, Safe domain and elementary geometry, Eur. J. Phys. **25**, 835 (2004).

[25] A. Connes and D. Zagier, A property of parallelograms inscribed in ellipses, Am. Math. Mon. **114**, 909 (2007).

[26] R. Horodecki, P. Horodecki, and M. Horodecki, Violating Bell inequality by mixed states: Necessary and sufficient condition, Phys. Lett. A **200**, 340 (1995).

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

[28] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[29] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: Necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[30] L. Hardy, Nonlocality for Two Particles without Inequalities for Almost All Entangled States, Phys. Rev. Lett. **71**, 1665 (1993).