Steering, incompatibility, and Bell-inequality violations in a class of probabilistic theories

Neil Stevens* and Paul Busch†
Department of Mathematics, University of York, York YO10 5DD, United Kingdom
(Received 5 December 2013; published 24 February 2014)

We show that connections between a degree of incompatibility of pairs of observables and the strength of violations of Bell's inequality found in recent investigations can be extended to a general class of probabilistic physical models. It turns out that the property of universal uniform steering is sufficient for the saturation of a generalized Tsirelson bound, corresponding to maximal violations of Bell's inequality. It is also found that a limited form of steering is still available and sufficient for such saturation in some state spaces where universal uniform steering is not given. The techniques developed here are applied to the class of regular polygon state spaces, giving a strengthening of known results. However, we also find indications that the link between incompatibility and Bell violation may be more complex than originally envisaged.

DOI: 10.1103/PhysRevA.89.022123 PACS number(s): 03.65.Ta, 03.65.Ud

I. INTRODUCTION

The Bell inequalities [1] provide constraints that certain families of joint probability distributions must satisfy to admit a common joint distribution. It is known that the satisfaction of a full set of Bell inequalities in a probabilistic system is equivalent to the existence of such a joint probability [2,3].[1] It was observed subsequently that joint measurability (in the sense that there exist joint probabilities of the usual quantum-mechanical form for every state) entails an operator form of Bell inequalities; therefore the Bell inequalities are satisfied whenever the observables involved in an Einstein-Podolsky-Rosen-Bell type experiment are mutually commutative [7]. In the case of "unsharp" observables, commutativity is not required for joint measurability and the degree of "unsharpness" of the observables required for joint measurability can be determined; this value is more restrictive than is needed for violations of the Bell inequalities to be eliminated in the case of the singlet state [8–11].

The connection between joint measurability and Bell inequalities—in the specific form of the Clauser-Holt-Shimony-Horne (CHSH) inequalities [12], which apply to experiments involving runs of measurements of two pairs of dichotomic observables on a bipartite system—has been further elucidated in two interesting recent publications by Wolf *et al.* [13] and Banik *et al.* [14]. The former have shown that for any pair of incompatible dichotomic observables in a finite-dimensional quantum system, a violation of a CHSH inequality will be obtained. Hence incompatibility is not only necessary but also sufficient for obtaining Bell-inequality violations. Wolf *et al.* [13] conclude that "if a hypothetical no-signaling theory is a refinement of quantum mechanics (but otherwise consistent with it), it cannot render possible the joint measurability of observables which are incompatible within quantum mechanics." With this result a tight link has been established between the availability of incompatible observables and the possibility of violating a CHSH inequality. It is natural to ask whether a quantitative connection can be found between a degree of incompatibility and the strength of these violations, and whether such a connection is specific to quantum mechanics or holds in a wider class of probabilistic physical theories.

It is a well-known fact that two incompatible quantum observables can be *approximately* measured together if some unsharpness in the measurement is allowed. A measure of the incompatibility of two observables can then be obtained by quantifying the degree of unsharpness required to obtain an approximate joint measurement. In the case of dichotomic observables this can be achieved by mixing each observable with a trivial observable [a positive operator-valued measure (POVM) whose positive operators are multiples of the identity],[2] with relative weights λ, $1-\lambda$. The mixing weight determines the degree of unsharpness of the resulting smeared observable.

Banik *et al.* have shown that the degree of incompatibility (they use the term complementarity) of two dichotomic observables, quantified by the largest smearing parameter λ for which the smeared versions are compatible, puts limitations on the maximum strength of CHSH inequality violations available in such a theory [14]. The Bell functional $\mathbb{B}$, a generalization of what is known as the Bell operator in the quantum case, then is bounded by the parameter λ_{opt} associated with the "most incompatible" pair of observables, so that $\mathbb{B} \leqslant 2/\lambda_{\text{opt}}$.

Here we study the connection between degrees of incompatibility and CHSH inequality violation in the context of general probabilistic physical theories by way of unifying the approaches of [13] and [14]. We will see that the degree of incompatibility used by Banik *et al.* is closely linked with an unnamed parameter used in [13] to characterize the joint measurability of two dichotomic observables. Under an additional assumption on the physical theory, namely, that it supports a sufficient degree of steering, the construction used to violate the CHSH inequality generalizes. This gives a sufficient condition under which the maximal violation can be saturated. This result can be rephrased by saying that probabilistic theories can be classified according to the value

*ns695@york.ac.uk
†paul.busch@york.ac.uk

[1] As observed by Pitowsky [4], Bell-type inequalities had already been formulated as early as 1854 by George Boole, who deduced them as conditions for the possibility of objective experience [5,6].

[2] Such mixing procedures and their connection with goal of achieving joint measurability are investigated systematically in [15]).

1050-2947/2014/89(2)/022123(7)

of the *generalized Tsirelson bound*, defined as the maximum value of the Bell functional, and this bound can (under said assumptions) be realized by suitable maximally incompatible observables (see Theorem 1).

Finally we illustrate the link between incompatibility and Bell violation in the class of regular polygon state spaces. It turns out that this connection appears to hold generally in the case of even-sided polygons but not, at least in the same form, for odd-sided cases.

II. GENERAL PROBABILISTIC MODELS

We begin by presenting the basic elements of the standard framework of probabilistic models. The framework was introduced in the 1960s by researchers in quantum foundations who used it to investigate axiomatic derivations of the Hilbert space formalism of quantum mechanics from operational postulates. Due to the emphasis on the convex structure of the set of states and the use of operations to model state transformations, the approach was called the *convex state approach* or *operational approach*. Some pioneering references are [16–20]. An overview of the literature and of relevant monographs can be obtained from [21] and [22]. Recently the approach has gained renewed interest from researchers in quantum information exploring the information theoretic foundations of quantum mechanics. Accessible recent introductions can be found in, e.g., [23–25].

The set of states Ω of a general probabilistic model is taken to be a compact convex subset of a finite-dimensional vector space V, where the convexity corresponds to the ability to define a preparation procedure as a probabilistic mixture of preparation procedures corresponding to other states. We write $A(\Omega)$ for the ordered linear space of affine functionals on Ω, with the ordering given pointwise: $f \geqslant 0$ if $f(\omega) \geqslant 0$ for all $\omega \in \Omega$. $A(\Omega)$ is also canonically an order unit space, with order unit u defined by $u(\omega) = 1$ for all states $\omega \in \Omega$. The (convex) set of effects on Ω is then taken to be the unit interval $[0,u]$ inside $A(\Omega)$, i.e.,

$$\mathscr{E}(\Omega) = \{e \in A(\Omega) | 0 \leqslant e(\omega) \leqslant 1, \forall\, \omega \in \Omega\}. \tag{1}$$

A discrete observable $\mathcal{O}$ is then a function from an outcome set X into $\mathscr{E}(\Omega)$ that satisfies the normalization condition $\sum_{x\in X} \mathcal{O}[x] = u$. The value (lying between 0 and 1) of $\mathcal{O}[x](\omega)$ denotes the probability of getting outcome x for a measurement of the observable $\mathcal{O}$ in state ω.

Under the assumption of tomographic locality [26], the state space of a composite system with local state spaces Ω_1 and Ω_2 naturally lives in the vector space $V_1 \otimes V_2$. We then write $\Omega = \Omega_1 \otimes \Omega_2 = (V_1 \otimes V_2)_+^1$, where the normalization is given by the order unit $u_1 \otimes u_2 \in V_1^* \otimes V_2^*$, but in general the positive cone is not unique [27].

Although there is much choice in general for the ordering on $V_1 \otimes V_2$, there are two canonical choices, the *maximal* and *minimal*. As a minimal demand it is reasonable to expect $v_1 \otimes v_2 \geqslant 0$ whenever $v_1, v_2 \geqslant 0$; therefore we make the definition

$$(V_1 \otimes_{\min} V_2)_+ = \left\{ \sum_{i,j} \lambda_{ij} v_1^{(i)} \otimes v_2^{(j)} \,\middle|\, \lambda_{ij} \in \mathbb{R}_+, v_k^{(i)} \in (V_k)_+ \right\}. \tag{2}$$

We can similarly make such demands on the order structure on $V_1^* \otimes V_2^*$, leading to the converse definition

$$(V_1 \otimes_{\max} V_2)_+ = (V_1^* \otimes_{\min} V_2^*)_+^*. \tag{3}$$

Any cone on $V_1 \otimes V_2$ which lies between the maximal and minimal cones is then admissible as a viable order structure. In general, the tensor product chosen is an important part in defining a theory; the only time when there is no choice (since maximal and minimal are the same) is when the local state spaces are simplexes [27]. The case where both Ω_1 and Ω_2 are quantum state spaces provides a prime example of a nonminimal, nonmaximal order structure, namely, the standard quantum-mechanical tensor product. By definition $\Omega_1 \otimes_{\min} \Omega_2$ contains only separable states, which form a proper subset of all bipartite states; by contrast, $\Omega_1 \otimes_{\max} \Omega_2$ contains not only the usual quantum states, but also all normalized entanglement witnesses.

A bipartite state $\omega \in \Omega_1 \otimes \Omega_2$ can also be viewed as a way to prepare states in Ω_1, via the measurement of an observable on Ω_2. In this way, for each state ω, we can define the corresponding linear map $\hat{\omega} : V_2^* \to V_1$ by

$$a(\hat{\omega}(b)) = \omega(a,b), \quad a \in V_1^*, \quad b \in V_2^*.$$

III. FUZZINESS AND JOINT MEASURABILITY

Consider a system represented by a probabilistic model, whose state space is given by the convex set Ω. Any dichotomic (or two-outcome) observable $\mathcal{O}$ on Ω is determined by an effect $e =: \mathcal{O}[+1] \in \mathscr{E}(\Omega)$, where for any $\omega \in \Omega$, the probability of getting the outcome labeled by +1 in the state ω is given by $e(\omega)$, and similarly for the outcome -1 associated with the complement effect $e' := u - e = \mathcal{O}[-1]$.

Two effects e and f are said to be jointly measurable if there exists $g \in A(\Omega)$, satisfying

$$0 \leqslant g, \quad g \leqslant e, \quad g \leqslant f, \quad e + f \leqslant g + u, \tag{4}$$

where u is the order unit on Ω. This condition is equivalent to the existence of a joint observable for the dichotomic observables corresponding to e and f. In fact, if the system of inequalities (4) is satisfied for some effect g, then the set of effects $g_{++} := g$, $g_{+-} := e - g$, $g_{-+} := f - g$, $g_{--} := u - e - f + g$ defines an observable that comprises e,e' and f,f' as marginals, in the sense that $e = g_{++} + g_{+-}$, $f = g_{++} + g_{-+}$, etc.[3]

Given a two-outcome observable A determined by effect e, one can introduce a corresponding fuzzy observable $\mathsf{A}^{(\lambda)}$ as a smearing (or fuzzy version) of A, whose defining effect is given by

$$e^{(\lambda)} = \frac{1+\lambda}{2} e + \frac{1-\lambda}{2} e' = \lambda e + \frac{1-\lambda}{2} u, \tag{5}$$

with smearing parameter $\lambda \in [0,1]$ and complement effect $e^{(\lambda)\prime} = e^{\prime(\lambda)}$.

Given any pair of two-outcome observables $\mathsf{A}_1, \mathsf{A}_2$, with corresponding effects e, f, we can use the parameter λ to

[3] For more detail on the notion of joint observables in probabilistic theories, we refer the reader to [28], where further relevant references can be found.

give a measure of how incompatible they are. First we note that for $\lambda = \frac{1}{2}$, the choice of effect $g = \frac{1}{4}(e + f)$ generates a joint observable for e and f since it satisfies (4), as is readily verified. Thus the set of values of λ which make $e^{(\lambda)}$ and $f^{(\lambda)}$ jointly measurable contains $\frac{1}{2}$. Further, if $e^{(\lambda)}$ and $f^{(\lambda)}$ are jointly measurable, then for any $\lambda' \leqslant \lambda$ so are $e^{(\lambda')}$ and $f^{(\lambda')}$. Hence the set lies inside the interval $[0,\lambda_{e,f}]$, where we define $\lambda_{e,f}$ to be the solution to the cone-linear program:

$$\begin{aligned} \text{maximize: } & \lambda \\ \text{subject to: } & g \leqslant e^{(\lambda)} \\ & g \leqslant f^{(\lambda)} \\ & 0 \leqslant g \\ e^{(\lambda)} + f^{(\lambda)} - u \leqslant{}& g. \end{aligned} \tag{6}$$

This measure of incompatibility of a pair of effects in turn leads to a measure of the degree of incompatibility of a given model by looking for the most incompatible pair:

$$\lambda_{\text{opt}} = \inf_{e,f \in \mathscr{E}(\Omega)} \lambda_{e,f}. \tag{7}$$

Following a path similar to [13], we can define a different parameter $t_{e,f}$, which we will see is closely linked with $\lambda_{e,f}$. For a given pair of effects e and f, we define $t_{e,f}$ to be the solution to the cone-linear program:

$$\begin{aligned} \text{minimize: } & t \\ \text{subject to: } & g \leqslant e + tu \\ & g \leqslant f + tu \\ & 0 \leqslant g \\ e + f - u \leqslant{}& g. \end{aligned} \tag{8}$$

As shown in [29], the optimal set for (8) is nonempty, so the minimum can be achieved; hence e and f are incompatible if and only if $t_{e,f} > 0$. Here we notice that the pair (λ, g) being feasible for the problem (6) is equivalent to the pair $(\frac{1-\lambda}{2\lambda}, \frac{g}{\lambda})$ being feasible for the problem (8). Combining this with the fact that the function $\frac{1-\lambda}{2\lambda}$ is monotonically decreasing for $\lambda \in [0,1]$ brings us to the promised link

$$t_{e,f} = \frac{1 - \lambda_{e,f}}{2\lambda_{e,f}}. \tag{9}$$

A. Examples

In a model of discrete classical probability theory, we take the state space to be the set of all probability measures on some countable set X, i.e.,

$$\Omega = \left\{ (\omega_x)_{x \in X} \,\middle|\, \omega_x \geqslant 0 \; \forall x \in X, \; \sum_x \omega_x = 1 \right\}. \tag{10}$$

A functional e on Ω with action $e(\omega) = \sum_x e_x \omega_x$ is easily seen to be positive if and only if $e_x \geqslant 0$ for all $x \in X$, and the order unit satisfies $u_x = 1$ for all $x \in X$.

Suppose we now have two effects $e, f \in \mathscr{E}(\Omega)$. Taking g to have components $g_x = \min\{e_x, f_x\}$, then since positivity is determined componentwise, the inequalities (4) are immediately satisfied and hence e and f are jointly measurable. Since this holds for arbitrary e and f, in this case we have $\lambda_{\text{opt}} = 1$.

As shown in [14], in any finite-dimensional Hilbert space the value of the joint measurability parameter for a pair of dichotomic observables is $\lambda_{\text{opt}} = 1/\sqrt{2}$.

A simple nonclassical, nonquantum example is that of the *squit*. The two-dimensional state space is given by a square, denoted $\square$; it contains all points $(x,y,1)$ with $-1 \leqslant x + y \leqslant 1$, $-1 \leqslant x - y \leqslant 1$, and takes the shape of a square. As we will see, the squit leads to maximally incompatible effects in the sense that it leads to the smallest possible value of λ_{opt}.

First, we note that for any probabilistic model, $\lambda = \frac{1}{2}$ provides a lower bound for λ_{opt}, since $e^{(\frac{1}{2})} = \frac{1}{2}e + \frac{1}{4}u$ and $f^{(\frac{1}{2})} = \frac{1}{2}f + \frac{1}{4}u$ are always jointly measurable. This can be seen explicitly by setting $g = \frac{1}{4}e + \frac{1}{4}f$, and then the corresponding equations (4) are satisfied.

As a convenient parametrization we can write a generic affine functional $g \in A(\square)$ as a vector $g = (a,b,c)$, with action given by the canonical inner product scaled by a factor of $\frac{1}{2}$. In this case the order unit is given by $u = (0,0,2)$. Since the positivity of a functional g on a compact convex set is equivalent to positivity on its extreme points, we can determine the structure of the set of effects by demanding that its elements g take values between 0 and 1 on the extreme points of the set of states. In the case of the squit, $\mathscr{E}(\square)$ is a convex polytope with defining inequalities given by

$$u \geqslant g \geqslant 0 \iff \begin{cases} 2 \geqslant c + a \geqslant 0, & 2 \geqslant c + b \geqslant 0, \\ 2 \geqslant c - a \geqslant 0, & 2 \geqslant c - b \geqslant 0. \end{cases} \tag{11}$$

We note the extreme points: $(0,0,2) = u$, $(0,0,0)$, $(1,1,1)$, $(1,-1,1)$, $(-1,1,1)$, $(-1,-1,1)$.

In an attempt to find the lowest possible value of $\lambda_{e,f}$, we consider the case of the two orthogonal extremal effects $e = (1,1,1)$ and $f = (1,-1,1)$. In order for $e^{(\lambda)}$ and $f^{(\lambda)}$ to be jointly measurable, we need to be able to find a g that satisfies all the inequalities in (4). This entails, in particular,

$$\begin{aligned} g - e^{(\lambda)} - f^{(\lambda)} + u = (a - 2\lambda, b, c) \geqslant 0, \\ \text{giving} \quad 2\lambda \leqslant a + c; \\ e^{(\lambda)} - g = (\lambda - a, \lambda - b, 1 - c) \geqslant 0, \\ \text{giving} \quad \lambda \leqslant 1 + a - c; \\ f^{(\lambda)} - g = (\lambda - a, -\lambda - b, 1 - c) \geqslant 0, \\ \text{giving} \quad \lambda \leqslant 1 - a - c; \\ g = (a,b,c) \geqslant 0, \\ \text{giving} \quad a \leqslant c. \end{aligned}$$

Combining these inequalities leads to $4\lambda \leqslant 2 + a - c \leqslant 2$, so for this choice of e and f we must have $\lambda_{e,f} \leqslant \frac{1}{2}$. Given that $\frac{1}{2}$ is the lowest possible value, we conclude that in the case of the squit, $\lambda_{\text{opt}} = \frac{1}{2}$.

IV. STEERING AND SATURATION OF THE GENERALIZED TSIRELSON BOUND

In order to give conditions on a generalized probabilistic model under which the bound on CHSH violations given in [14] can be achieved we need to introduce the notion of steering, as given in [30].

Given two systems A and B, with state spaces Ω_A and Ω_B, respectively, for any bipartite state $\omega \in \Omega_A \otimes \Omega_B$ we can define its A *marginal*, living in Ω_A in an analog to the quantum-mechanical partial trace:

$$\omega^A = \hat{\omega}(u_B), \tag{12}$$

where u_B is the order unit on B, with a similar definition for ω^B.

Following this we say that a state $\omega \in \Omega_A \otimes \Omega_B$ is *steering* for its A marginal if for any collection of subnormalized states that form a decomposition of that marginal, i.e., $\{\alpha_1,\dots,\alpha_n | \sum_i \alpha_i = \omega^A, 0 \leqslant u_A(\alpha_i) \leqslant 1\}$, there exists an observable $\{e_1,\dots,e_n\} \subset \mathscr{E}(\Omega_B)$ with $\alpha_i = \hat{\omega}(e_i)$.

It was observed by Schrödinger that this property holds in quantum mechanics for all pure bipartite states [31], originally coining the term "steering," which we generalize now following [30]: A general probabilistic model of a system A with state space Ω_A supports *uniform universal steering* if there is another system B with state space Ω_B, such that for any $\alpha \in \Omega_A$, there is a state $\omega_\alpha \in \Omega_A \otimes \Omega_B$, with $\omega_\alpha^A = \alpha$ that is steering for its A marginal, and supports *universal self-steering* if the above is satisfied with $B = A$. The existence of steering in this manner is similar to the idea of purification to be found, for example, in [32]. Indeed, any purification of a state will be steering for its marginals; however, steering states being pure is not required here.

The magnitude of maximal CHSH violations is quantified in quantum mechanics by the norm of the *Bell operator*. We take $\mathsf{A}_1,\mathsf{A}_2,\mathsf{B}_1$, and B_2 to be ±1-valued observables, and define following [14],

$$\mathbb{B} := \langle A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2\rangle_\omega,$$

where $A_1 := \mathsf{A}_1[+1] - \mathsf{A}_1[-1]$, etc., and $\langle X\rangle_\omega := X(\omega)$ for any affine functional X. We will call the map $\omega \mapsto \mathbb{B}$ the *Bell functional* and refer to $\sup_\omega \mathbb{B}$ as the (generalized) *Tsirelson bound*.

In order to see where steering enters the picture, we follow [14] to get a simple bound on the norm of $\mathbb{B}$. In order to do this we consider what effect smearing the observables of one party has by defining

$$\mathbb{B}^{(\lambda)} = \left\langle A_1^{(\lambda)}B_1 + A_1^{(\lambda)}B_2 + A_2^{(\lambda)}B_1 - A_2^{(\lambda)}B_2\right\rangle, \tag{13}$$

where $A_1^{(\lambda)} = \mathsf{A}_1^{(\lambda)}[+1] - \mathsf{A}_1^{(\lambda)}[-1]$ etc., with the smearing of the effects as defined as in (5). Because the choice of observable that is mixed to form the smearing is an unbiased, trivial observable, the resulting expectation scales with the smearing parameter:

$$A_1^{(\lambda)} = \lambda\mathsf{A}_1[+1] + \frac{1-\lambda}{2}u - \lambda\mathsf{A}_1[-1] - \frac{1-\lambda}{2}u = \lambda A_1. \tag{14}$$

Now since the Bell functional is bilinear and the same smearing parameter is being used on all functionals on the first system, the linear scaling carries over and we get $\mathbb{B}^{(\lambda)} = \lambda\mathbb{B}$.

As shown in the previous chapter, there always exists jointly measurable fuzzy versions of any pair of observables, as long as the value of the smearing parameter is small enough. Now if we take any λ such that $A_1^{(\lambda)}$ and $A_2^{(\lambda)}$ are jointly measurable, then we know that the corresponding Bell functional satisfies the usual Bell inequality, and thus its value is bounded by $\mathbb{B}^{(\lambda)} \leqslant 2$. Consequently, each such value of λ gives a bound on the Bell functional of $\mathbb{B} \leqslant \frac{2}{\lambda}$, and in order to obtain the lowest such upper bound, we take the largest smearing parameter which still results in joint measurability to get

$$\mathbb{B} \leqslant \frac{2}{\lambda_{\mathsf{A}_1[+1],\mathsf{A}_2[+1]}}. \tag{15}$$

Since every probabilistic model contains observables which are jointly measurable with no smearing, and thus satisfying the usual Bell inequality, knowing the above bound for a single pair of observables will not necessarily yield information about the structure of the system itself. A more general bound, however, can be written down by simply taking the most incompatible pair of observables:

$$\mathbb{B} \leqslant \frac{2}{\lambda_{\rm opt}}. \tag{16}$$

Theorem 1. In any probabilistic model of a system A that supports uniform universal steering, the Tsirelson bound is given by the tight inequality that can be saturated:

$$\mathbb{B} \leqslant \frac{2}{\lambda_{\rm opt}}, \tag{17}$$

with $\lambda_{\rm opt}$ defined in Eq. (7).

Proof. Suppose we have a model of a system A that supports uniform universal steering, and that we have two effects $e,f \in \mathscr{E}(\Omega_A)$. The parameter introduced earlier, $t_{e,f}$, can now also be calculated from the dual program to (8), which can be given as [33]

$$\begin{aligned} \text{maximize:}\quad & \mu_3(e+f-u_A) - \mu_1(e) - \mu_2(f) \\ \text{subject to:}\quad & (\mu_1+\mu_2)(u_A) = 1 \\ & \mu_1+\mu_2 = \mu_3+\mu_4 \\ & 0 \leqslant \mu_1,\mu_2,\mu_3,\mu_4, \end{aligned} \tag{18}$$

with $\mu_i \in A(\Omega_A)^*$.

Writing $\mu_1 + \mu_2 = \rho$, for the μ_i that achieve the optimal value for (18), we find that $\rho \geqslant 0$ and $u_A(\rho) = 1$, so $\rho \in \Omega_A$. By the assumption of uniform universal steering, therefore we can find a state $\omega \in \Omega_A \otimes \Omega_B$ with $\omega^A = \hat{\omega}(u_B) = \rho$; moreover, in $\{\mu_1,\mu_2\}$ and $\{\mu_3,\mu_4\}$ we have two different decompositions of ρ, and we can thus find effects $\tilde{e},\tilde{f} \in \mathscr{E}(\Omega_B)$ satisfying

$$\hat{\omega}(\tilde{e}) = \mu_1, \quad \hat{\omega}(\tilde{f}) = \mu_3. \tag{19}$$

To achieve the maximum CHSH violations, we take $\mathsf{A}_1,\mathsf{A}_2,\mathsf{B}_1$, and B_2 to be ±1-valued observables defined by effects $f',e,\tilde{e}'$, and $\tilde{f}'$, respectively; we then have

$$\begin{aligned} A_1 &= u_A - 2f, \quad B_1 = u_B - 2\tilde{e}, \\ A_2 &= 2e - u_A, \quad B_2 = u_B - 2\tilde{f}. \end{aligned} \tag{20}$$

The value of the Bell functional can now be evaluated:

$$\begin{aligned}\mathbb{B} &= \omega(u_A - 2f, 2u_B - 2\tilde{e} - 2\tilde{f}) + \omega(2e - u_A, 2\tilde{f} - 2\tilde{e})\\ &= 2\hat{\omega}(u_B)(u_A - 2f) + 4\hat{\omega}(\tilde{e})(f - e)\\ &\quad + 4\hat{\omega}(\tilde{f})(f + e - u_A)\\ &= 2 + 4[(\mu_1 + \mu_2)(-f) + \mu_1(f) - \mu_1(e)\\ &\quad + \mu_3(f + e - u_A)]\\ &= 2 + 4[\mu_3(e + f - u_A) - \mu_1(e) - \mu_2(f)]\\ &= 2(2t_{e,f} + 1) = \frac{2}{\lambda_{e,f}},\end{aligned}$$

thus saturating the generalized Tsirelson bound as claimed. ■

Not every probabilistic model may possess the property of supporting uniform universal steering, and although it is a sufficient condition to obtain the conclusion of the above theorem, as the following example will show, it is not a necessary one. Indeed, a model of "boxworld," which contains Popescu-Rohrlich (PR) box states exhibiting the maximum possible CHSH violations, uses local state spaces that are the squits introduced earlier, and composition is given by the maximal tensor product. Despite the saturation of the generalized Tsirelson bound, such a state space does not admit uniform universal steering.

To see this, we consider a bipartite state $\omega \in \square \otimes_{\max} \square$ with the corresponding map $\hat{\omega}$. Note that from the definition of ω being a state, $\hat{\omega}$ will automatically be a positive map sending V_+^* into V_+. Now suppose ω is steering for its marginal ρ, i.e., $\hat{\omega}(u) = \rho$, and choose a decomposition of ρ into pure states, $\rho = \sum_i \alpha_i$. Since the subnormalized states in the decomposition are pure and $\hat{\omega}$ is positive, the inverse images $\hat{\omega}^{-1}(\alpha_i)$ must lie on extremal rays of the cone V_+^*. Consider the extremal ray effect $e = (1,1,1)$ with its complement $e' = (-1,-1,1)$ (which is again extremal). With appropriate labeling of the α_i we can then write $\alpha_1 = \hat{\omega}(e)$ and $\alpha_2 = \hat{\omega}(e')$; however, since we have $e + e' = u$,

$$\alpha_1 + \alpha_2 = \hat{\omega}(e + e') = \hat{\omega}(u) = \rho,$$

and hence ρ can be written as a mixture of just two pure states. Since there are many points in a square that can only be written as a convex combination of a minimum of three extreme points, we conclude that such a boxworld model does not support universal uniform steering.

Remark 1. It is interesting to note that there is another set of conditions sufficient to obtain the conclusion of the above theorem. We say that a positive cone V_+ is *homogeneous* if the space of order automorphisms of V acts transitively on the interior of V_+, and (weakly) *self-dual* if there exists a linear map $\eta : V \to V^*$ that is an isomorphism of ordered linear spaces, i.e., $\eta(V_+) = V_+^*$. It is known that homogeneity follows from uniform universal steering. Conversely, if the positive cone V_+ generated by the state space Ω of the probabilistic model of a system A is homogeneous and weakly self-dual, then uniform universal self-steering follows if the maximal tensor product is adopted. Hence the conditions of Theorem 1 are fulfilled [30] and the Tsirelson bound in the inequality $\mathbb{B} \leqslant 2/\lambda_{\text{opt}}$ can be saturated.

In the quantum probabilistic model, the tensor product is not maximal but still uniform universal steering holds. The classical model (trivially) satisfies the conditions of weak self-duality and homogeneity, and the tensor product is maximal. The squit is weakly self-dual but does not satisfy uniform universal steering, so that homogeneity fails; but it allows enough self-steering so that the maximal Bell-Tsirelson bound of 4 can be realized.

V. GENERALIZED TSIRELSON BOUNDS FOR POLYGON STATE SPACES

Work in [34] suggests that there is a spectrum of values for the generalized Tsirelson bound in the case of two-dimensional polygon state spaces (given as the convex hulls of regular polygons). It is shown there that for a system composed of two identical polygon state spaces with an odd number of vertices, the maximally entangled state does not lead to a violation of the standard Tsirelson bound of $2\sqrt{2}$, whereas in the case of an even number of vertices this bound can be exceeded. This suggests that among the class of polygon state spaces, the generalized Tsirelson bound can be either smaller or greater than the standard Tsirelson bound.

Remark 2. We note that of the polygon state spaces, the only cases in which homogeneity holds are the $n = 3$ triangle and the $n \to \infty$ circle. Hence uniform universal steering is generally not available; however, it may still be possible to saturate the generalized Tsirelson bound in some cases, but in others this may not be possible.

As shown in [34], in the case of boxworld, where each local state space is a square, the maximally entangled state is a PR box and takes the maximum possible value for the Bell functional of 4. This agrees with the result that the squit does indeed lead to the maximum amount of incompatibility, and shows that in this case the generalized Tsirelson bound can be saturated. We have been able to show that this conclusion holds also in regular polygon state spaces where the number of vertices is a multiple of 8. We expect this result to extend to all even-sided cases. This strengthens the expectation, expressed in [34], that in these cases the Tsirelson bound is saturated with the maximally entangled state.

Moving to the $n = 5$ case makes things a lot more interesting, however. To see this we follow the notation in [34] and define the family of state spaces Ω_n to be the convex hull of the points

$$\omega_i = \begin{pmatrix} r_n \cos\left(\frac{2\pi i}{n}\right) \\ r_n \sin\left(\frac{2\pi i}{n}\right) \\ 1 \end{pmatrix}, \quad i = 1,\ldots,n,$$

with $r_n = \sqrt{\sec(\frac{\pi}{n})}$.

The qualitative difference between the state spaces of odd- and even-sided polygons first appears in the structure of the set of effects. For the case of even n, along with 0 and u, there are n extremal effects:

$$e_i = \frac{1}{2}\begin{pmatrix} r_n \cos\left(\frac{(2i-1)\pi}{n}\right) \\ r_n \sin\left(\frac{(2i-1)\pi}{n}\right) \\ 1 \end{pmatrix}, \quad i = 1,\ldots,n,$$

and in this case all the e_i lie on extremal rays of the cone V_+^*. This important fact occurs since for each of the e_i we can find another effect e_j, also extremal, which is its complement, i.e., $e_j = e_i' = u - e_i$, namely, for $j = i + \frac{n}{2} \bmod n$. For the case of odd n, a seemingly similar expression arises for the ray extremal effects:

$$e_i = \frac{1}{1+r_n^2}\begin{pmatrix} r_n \cos\left(\frac{2\pi i}{n}\right) \\ r_n \sin\left(\frac{2\pi i}{n}\right) \\ 1 \end{pmatrix}, \quad i = 1,\ldots,n.$$

On this occasion, however, the complements of the e_i are given by

$$e_i' = u - e_i = \frac{1}{1+r_n^2}\begin{pmatrix} -r_n \cos\left(\frac{2\pi i}{n}\right) \\ -r_n \sin\left(\frac{2\pi i}{n}\right) \\ r_n^2 \end{pmatrix}, \quad i = 1,\ldots,n,$$

which do not coincide with the e_i, and thus there are $2n$ nontrivial extreme points of $\mathscr{E}(\Omega_n)$.

Now we can pose the question of what the value is for $\lambda_{\rm opt}$ when the state space is Ω_5, and whether is it possible to achieve the corresponding Bell value $\mathbb{B} = 2/\lambda_{\rm opt}$. Since each extreme two-valued observable is determined by a ray effect, the largest value of incompatibility will come from one of the possible pairs of the e_i. However, due to the symmetry of the state space, the affine transformation of rotating by $\pi/5$ serves only to cyclically permute the indices of the e_i modulo 5. This means that there are only two possible values of λ_{e_i,e_j}, those for nearest neighbors and those for next-nearest neighbors. Calculation shows that these values are, for example,

$$\lambda_{e_1,e_2} = \frac{3+2\sqrt{5}}{11} \approx 0.679\,28, \quad \lambda_{e_1,e_3} = \frac{8+3\sqrt{5}}{19} \approx 0.774\,16,$$

and hence the value of $\lambda_{\rm opt}$ for the pentagon is $\frac{3+2\sqrt{5}}{11}$. From (16) this gives the bound on the Bell functional as $\mathbb{B} \leqslant 4\sqrt{5} - 6$. However, unlike in the case of the tensor product of two squits, the maximally entangled state between two pentagonal state spaces does not saturate the corresponding bound; instead we get a value of $\mathbb{B} = \frac{6}{\sqrt{5}}$, strictly below that coming from the level of incompatibility on one state space. This fact suggests that either the chosen way of evaluating the level of incompatibility in a system used does not capture everything, or that there is some structural obstruction that prevents such a link from holding that does not exist in other cases. Here we present some evidence towards the former.

In order to improve the measure of incompatibility used, we wish to modify the program used in Eq. (6). To do this we relax the method of smearing used, still mixing in multiples of the order unit corresponding to trivial observables, but we now allow them to be possibly biased as follows:

$$e^{(\lambda,p)} = \lambda e + p(1-\lambda)u. \tag{21}$$

This definition encompasses the old, with $e^{(\lambda)} = e^{(\lambda,\frac{1}{2})}$.

The updated measure of incompatibility of a pair of effects e and f, which we denote $\bar{\lambda}_{e,f}$, is now given by the optimal value of the optimization program:

$$\begin{aligned} \text{maximize:}\quad & \lambda \\ \text{subject to:}\quad & g \leqslant e^{(\lambda,p)} \\ & g \leqslant f^{(\lambda,q)} \\ & 0 \leqslant g \\ e^{(\lambda,p)} + f^{(\lambda,q)} - u \leqslant{} & g \\ & 0 \leqslant p,q \leqslant 1. \end{aligned} \tag{22}$$

Solving this updated problem in the case of the pentagon again gives the optimal value on, e.g., e_1 and e_2, with

$$\bar{\lambda}_{\rm opt} = \frac{5+\sqrt{5}}{10} \approx 0.723\,61,$$

which occurs for the values $p = q = 1$.

This is indeed a different value from earlier, but still we have that $\frac{2}{\bar{\lambda}_{\rm opt}} \neq \frac{6}{\sqrt{5}}$; however, in this case the unbiased nature of the observables mixed in means such a simple link is no longer expected, and indeed we can see that there is a link to the Bell value on the maximally entangled state as follows. As in the previous, we can define a smeared version of the Bell functional, where the smearing is all done on the functionals of one party:

$$\mathbb{B}^{(\lambda,1)} = \left\langle A_1^{(\lambda,1)}B_1 + A_1^{(\lambda,1)}B_2 + A_2^{(\lambda,1)}B_1 - A_2^{(\lambda,1)}B_2 \right\rangle. \tag{23}$$

But now instead of having the nice linear scaling in λ, we gain an extra expectation term $\mathbb{B}^{(\lambda,1)} = \lambda\mathbb{B} + 2(1-\lambda)\langle B_1\rangle$, and again under the assumption that λ is small enough to ensure joint measurability, and then taking the largest such value, we can write the inequality

$$\mathbb{B} \leqslant \frac{2[1-(1-\bar{\lambda}_{\rm opt})\langle B_1\rangle]}{\bar{\lambda}_{\rm opt}}. \tag{24}$$

The link to the maximally entangled state on two pentagons now comes from noting that the expectation of any observable B_1, defined by an extreme effect on the maximally entangled state, is $\langle B_1\rangle = \frac{5-2\sqrt{5}}{5}$. This means that if evaluated in the maximally entangled state, the inequality in (24), for the value of $\bar{\lambda}_{\rm opt}$ given above, is indeed saturated.

VI. CONCLUSION

By combining and developing ideas from the works of Wolf *et al.* [13] and Banik *et al.* [14], we have shown that probabilistic models can be classified according to their associated value of the generalized Tsirelson bound, which specifies the maximum possible violation of CHSH inequalities. We have given conditions (defined and studied in [30]), that probabilistic models may or may not satisfy, under which the maximal CHSH violations are attained for appropriate choices of maximally incompatible dichotomic observables. Here the degree of incompatibility of two observables is defined by the minimum amount of smearing of these observables necessary to turn them into jointly measurable observables.

The authors of [13] concluded that observables that are incompatible in quantum mechanics remain incompatible in any probabilistic model that serves as an extension of quantum mechanics. Here we have shown that this conclusion applies to extensions of any probabilistic model that allows for sufficient steering.

As an illustration of the general results, we have considered the squit system which underlies the PR box model, and have identified the pair of maximally incompatible extremal effects of the squit that give rise to the saturation of the largest possible value (i.e., 4) of the Tsirelson bound. In addition, we have obtained partial confirmation of the conjectured maximality of the Bell functional if evaluated on the maximally entangled state in the class of regular polygon state spaces considered in [34].

In the case of the pentagon state space, we discovered that the connection between incompatibility and Bell violation is not always of the simple form envisaged originally and used through most of this paper; this suggests that the definitive universal expression of this connection remains yet to be found.

The methods used here are taken from among some of the standard tools of quantum measurement and information theory used in [13] and [14], and we have shown that they apply equally well in a wide class of probabilistic models. This insight may prove valuable in future investigations into the characterization of quantum mechanics among all probabilistic models.

ACKNOWLEDGMENTS

We wish to thank Takayuki Miyadera for the suggestion to consider the generalization of the optimization problem for the incompatibility parameter that allows mixing with trivial observables that are not necessarily unbiased. We are also grateful to Manik Banik and Alexander Wilce for helpful comments on a draft version of the paper. N.S. gratefully acknowledges support through the award of an Annie Currie Williamson Ph.D. Bursary at the University of York.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[3] A. Fine, J. Math. Phys. **23**, 1306 (1982).
[4] I. Pitowsky, Br. J. Philos. Sci. **45**, 95 (1994).
[5] G. Boole, *The Laws of Thought* (Macmillan, Cambridge, 1854).
[6] G. Boole, Philos. Trans. R. Soc. London **152**, 225 (1862).
[7] W. M. de Muynck and O. Abu Zeid, Phys. Lett. A **100**, 486 (1984).
[8] P. Busch, in *Symposium on the Foundations of Modern Physics*, edited by P. J. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985), pp. 343–357.
[9] P. Busch, Phys. Rev. D **33**, 2253 (1986).
[10] G. Kar and S. Roy, Phys. Lett. A **199**, 12 (1995).
[11] E. Andersson, S. M. Barnett, and A. Aspect, Phys. Rev. A **72**, 042104 (2005).
[12] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[13] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).
[14] M. Banik, Md. R. Gazi, S. Ghosh, and G. Kar, Phys. Rev. A **87**, 052125 (2013).
[15] P. Busch and T. Heinosaari, Quantum Inf. Comput. **8**, 0797 (2008).
[16] G. Ludwig, Commun. Math. Phys. **4**, 331 (1967).
[17] G. Ludwig, Commun. Math. Phys. **9**, 1 (1968).
[18] B. Mielnik, Commun. Math. Phys. **15**, 1 (1969).
[19] E. B. Davies, J. Funct. Anal. **6**, 318 (1970).
[20] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
[21] S. P. Gudder, *Stochastic Methods in Quantum Mechanics* (North Holland, New York, 1979).
[22] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer-Verlag, Berlin, 1997).
[23] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, arXiv:quant-ph/0611295.
[24] H. Barnum and A. Wilce, Electronic Notes in Theoretical Computer Science **270**, 3 (2011).
[25] C. Pfister, arXiv:1203.5622.
[26] L. Hardy, arXiv:1104.2066.
[27] I. Namioka and R. R. Phelps, Pac. J. Math. **31**, 469 (1969).
[28] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, Europhys. Lett. **103**, 10002 (2013).
[29] R. Beneduci, P. Busch, and N. Stevens (unpublished).
[30] H. Barnum, C. P. Gaebler, and A. Wilce, arXiv:0912.5532.
[31] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936).
[32] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
[33] S. P. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).
[34] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, New J. Phys. **13**, 063024 (2011).