THE ROYAL SWEDISH ACADEMY OF SCIENCES
INSTITUT MITTAG-LEFFLER
Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

(Non-)Contextuality of Physical Theories as an Axiom

A. . Cabello, S. Severini and A. Winter

REPORT No. 8, 2010/2011, fall
ISSN 1103-467X
ISRN IML-R- -8-10/11- -SE+fall

(Non-)Contextuality of Physical Theories as an Axiom

Adán Cabello,[1,*] Simone Severini,[2,†] and Andreas Winter[3,4,‡]

[1] *Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain*
[2] *Department of Physics & Astronomy, University College London, WC1E 6BT London, United Kingdom*
[3] *Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom*
[4] *Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, Singapore 117542*

(Dated: 11 October 2010)

We show that the noncontextual inequality proposed by Klyachko *et al.* [Phys. Rev. Lett. **101**, 020403 (2008)] belongs to a broader family of inequalities, one associated to each compatibility structure of a set of events (a graph), and its independence number. These have the surprising property that the maximum quantum violation is given by the Lovász ϑ-function of the graph, which was originally proposed as an upper bound on its Shannon capacity. Furthermore, probabilistic theories beyond quantum mechanics may have an even larger violation, which is given by the so-called fractional packing number. We discuss in detail, and compare, the sets of probability distributions attainable by noncontextual, quantum, and generalized models; the latter two are shown to have semidefinite and linear characterizations, respectively. The implications for Bell inequalities, which are examples of noncontextual inequalities, are discussed. In particular, we show that every Bell inequality can be recast as a noncontextual inequality à la Klyachko *et al.*

PACS numbers: 03.65.Ud, 03.67.-a, 02.10.Ox

Introduction.—Recently, Klyachko, Can, Binicioğlu, and Shumovsky (KCBS) [1] have introduced a noncontextual inequality (*i.e.*, one satisfied by any noncontextual hidden variable theory), which is violated by quantum mechanics, and therefore can be used to detect quantum effects. The simplest physical system which exhibits quantum features in this sense is a three-level quantum system or qutrit [2–4]. The KCBS inequality is the simplest noncontextual inequality violated by a qutrit, in a similar way that the Clauser-Horne-Shimony-Holt (CHSH) inequality [5] is the simplest Bell inequality violated by a two-qubit system.

The KCBS inequality has been recently tested in the laboratory [6] and has stimulated many recent developments [7–13]. It can adopt two equivalent forms. Consider 5 *yes-no* questions P_i ($i = 0, \dots, 4$) such that P_j and P_{j+1} (with the sum modulo 5) are *compatible*: both questions can be jointly asked without mutual disturbance, so, when the questions are repeated, the same answers are obtained; and *exclusive*: not both can be true. One can represent each of these questions as a vertex of a pentagon (*i.e.*, a 5-cycle) where the edges denote compatibility and exclusiveness. What is the maximum number of *yes* answers one can get when asking the 5 questions to a physical system? Clearly, two, because of the exclusiveness condition [14]. If we denote *yes* and *no* by 1 and 0, respectively, then, even if one asks only one question to each one of an identically prepared collection of systems, and then count the average number of *yes* answers corresponding to each question, the following inequality holds:

$$\beta := \sum_{i=0}^{4} \langle P_i \rangle \leq 2, \tag{1}$$

if we assume that these answers are predetermined by a hidden variable theory. This is the first form of the KCBS inequality. What has (1) to do with noncontextuality? Noncontextual hidden variable theories are those in which the answer of P_j is independent of whether one ask P_j together with P_{j-1} (which is compatible with P_j), or together with P_{j+1} (which is also compatible with P_j). A set of mutually compatible questions is called a *context*. Since, P_{j+1} and P_{j-1} are not necessarily compatible, $\{P_j, P_{j-1}\}$ is one context and $\{P_j, P_{j+1}\}$ is a different one, and they are not both contained in a joint context. The assumption is that the answer to P_j will be the same in both.

Now, let us consider contexts instead of questions, *i.e.*, let us ask individual systems not one but two compatible and exclusive questions. In the pentagon, a context is represented by an edge connecting two vertices, so we have 5 different contexts. In order to study the correlations between the answers to these questions, it is useful to transform each question into a dichotomic observable with possible values -1 (no) or $+1$ (yes), so when both questions give the same answer the product of the results of the observables is $+1$, but when the answers are different then the product of the results of the observables is -1. For instance, this can be done by defining the observables $A_i = 2P_i - 1$. Then, inequality (1) is equivalent to the noncontextual correlation inequality, the second form of KCBS,

$$\beta' := \sum_{i=0}^{4} \langle A_i A_{i+1} \rangle \geq -3, \tag{2}$$

which can be derived independently based solely on the assumption that the observables A_i have noncontextual results -1 or $+1$. *I.e.*, we do not need to assume exclusiveness to derive it, effectively because the occurrence of correlation functions $\langle A_i A_{i+1}\rangle$ implements a penalty for violating exclusiveness.

For a qutrit, the maximum quantum violation of inequality (1) was shown to be $\beta_{\rm QM}(5) = \sqrt{5} \approx 2.236$, which is equivalent to a violation of inequality (2) of $\beta'_{\rm QM}(5) = 5 - 4\sqrt{5} \approx -3.94$. The maximum violation of the KCBS inequality occurs for the state $\langle\psi| = (0, 0, 1)$ and the questions $P_i = |v_i\rangle\langle v_i|$ or the observables $A_i = 2|v_i\rangle\langle v_i| - \mathbb{1}$, where

$$\begin{aligned}
\langle v_0| &= N_0 \left(1, 0, \sqrt{\cos(\pi/5)}\right),\\
\langle v_{1,4}| &= N_1 \left(\cos(4\pi/5), \pm\sin(4\pi/5), \sqrt{\cos(\pi/5)}\right), \quad (3)\\
\langle v_{2,3}| &= N_2 \left(\cos(2\pi/5), \mp\sin(2\pi/5), \sqrt{\cos(\pi/5)}\right),
\end{aligned}$$

the N_i being suitable normalization factors. These vectors connect the origin with the vertices of a regular pentagon. Interestingly, with this choice, $\langle A_i A_{i+1}\rangle = [-1 + 3\cos(\pi/5)]\sec^2(\pi/10)/2$, for $i = 0, \ldots, n-1$. Observe that $\langle v_i|v_{i+1}\rangle = 0$ and $\beta_{\rm QM}(5) = \sum_{i \bmod 5} |\langle\psi|v_i\rangle|^2$. The vectors that give $\beta_{\rm QM}(5)$ form an orthonormal representation of the 5-cycle.

General compatibility structures.— The KCBS inequality suggests itself a generalization to arbitrary graphs instead of the pentagon. Most generally and abstractly, Kochen-Specker (KS) theorems [4] are about the possibility of interpreting a given structure of compatibility of "events," and additional constraints such as exclusiveness, in a classical or nonclassical probabilistic theory. In this paper, these events are interpreted as *atomic* events, each of which can occur in different contexts. Formally, the events are labelled by a set V (in practice finite, and often just integer indices, $V = \{1, 2, \ldots, n\}$). The set of all valid contexts is a *hypergraph* Γ, which is simply a collection of subsets $C \subset V$; note that for hypergraphs of contexts, with each $C \in \Gamma$, all of the subsets of C are also valid contexts, and hence part of C. The interpretation is that there should exist (deterministic) events in a probabilistic model, one P_i for each $i \in V$, and for each context C a measurement among whose outcomes are the P_i $(i \in C)$. The events are hence mutually exclusive, as in the measurement postulated to exist for some $C \in \Gamma$, at most one outcome $i \in C$ can occur. For instance, a *classical (noncontextual) model* would be a measurable space Ω, with each P_i being the indicator function of a measurable set (an event, in fact) such that for all $C \in \Gamma$, $\sum_{i\in C} P_i \leq 1$ (*i.e.*, the supporting sets of the P_i should be pairwise disjoint).

In contrast, a *quantum model* requires a Hilbert space $\mathcal{H}$ and associates projection operators P_i to all $i \in V$, such that for all $C \in \Gamma$, $\sum_{i\in C} P_i \leq \mathbb{1}$ (*i.e.*, the P_i can be thought of as outcomes in a von Neumann measurement).

Thanks to KS we know that quantum models are strictly more powerful that classical ones; but they are still not the most general ones. A *generalized model* requires choosing a generalized probabilistic theory in which the P_i can be interpreted as measurement outcomes: following [15–19], formally it consists of a real vector space $\mathcal{A}$ of observables, with a distinguished unit element $u \in \mathcal{A}$ and a vector space order: the latter is given by the closed convex cone $\mathcal{P} \subset \mathcal{A}$ of positive elements containing u in its interior, such that $\mathcal{P}$ spans $\mathcal{A}$ and is pointed, meaning that, with the exception of 0, $\mathcal{P}$ is entirely on one side of a hyperplane. For two elements $X, Y \in \mathcal{A}$ we then say $X \leq Y$ if and only if $Y - X \in \mathcal{P}$. (We shall only discuss finite dimensional $\mathcal{A}$, otherwise there will be additional topological requirements.) The elements with $0 \leq E \leq u$ are called *effects*. This structure is enough to talk about measurements: they are collections of effects $(E_1, \ldots, E_k)$ such that $\sum_{j=1}^k E_j = u$.

[Observe how we recover quantum mechanics when $\mathcal{P}$ consists of the semidefinite matrices within the Hermitian ones over a Hilbert space, and $u = \mathbb{1}$. Classical probability instead, when $\mathcal{P}$ are the non-negative functions within the measurable ones over a measure space, u being the constant 1 function.]

Now, a generalized model for the hypergraph Γ is the association of an effect $P_i \in \mathcal{A}$ to each $i \in V$, such that each P_i is a sum of normalized extremal effects, and for all $C \in \Gamma$, $\sum_{i\in C} P_i \leq u$. The latter condition ensures that the family $(P_i : i \in C)$ can be completed to a measurement, possibly in a larger space $\widetilde{\mathcal{A}} \supset \mathcal{A}$. We finally demand that this can be done such that also $u - \sum_{i\in C} P_i$ is a sum of normalized extremal effects.

Notice that in all of the above we never require that any particular context should be associated to a complete measurement: the conditions only make sure that each context is a subset of outcomes of a measurement and that they are mutually exclusive. Thus, unlike the original KS theorem, it is clear that every context hypergraph Γ has always a classical noncontextual model, besides possibly quantum and generalized models. This is where noncontextual inequalities come in: note that all of the above types of models allow for the choice of a state (be it a probability density, a quantum density operator, or generalized state), under which all expectation values $\langle P_i\rangle$ make sense, and hence also the expression

$$\beta = \sum_{i\in V} \langle P_i\rangle. \quad (4)$$

Moreover, all probabilities $\langle P_i\rangle$ are independent of the context in which P_i occurs, as they depend only on the effect P_i and the underlying state. Since this is the condition underlying Gleason's theorem, we call it the *Gleason property*.

We can then ask for the set of all attainable vectors $\left(\langle P_i\rangle\right)_{i\in V}$ for given hypergraph Γ, over all models of a given sort (classical noncontextual, quantum mechanical, or generalized probabilistic theory) and states within it.

These are evidently convex subsets in $[0,1]^V \subset \mathbb{R}^V$; we denote the sets of noncontextual, quantum and generalized expectations by $\mathcal{E}_{\rm C}(\Gamma)$, $\mathcal{E}_{\rm QM}(\Gamma)$ and $\mathcal{E}_{\rm GPT}(\Gamma)$, respectively. The central task of the present theory is to characterize these convex sets and to compare them for various Γ. This is because a point $\vec{p} \in \mathcal{E}_{\rm X}(\Gamma)$ in any of these sets describes the outcome probabilities of any compatible set of events (*i.e.*, any context). Note that all of them are *corners* in the language of [20]: if $0 \leq q_i \leq p_i$ for all $i \in V$, then $\vec{p} \in \mathcal{E}_{\rm X}(\Gamma)$ implies also $\vec{q} \in \mathcal{E}_{\rm X}(\Gamma)$.

In particular, the extreme values of β over these sets are denoted $\beta_{\rm C}(\Gamma)$, $\beta_{\rm QM}(\Gamma)$, and $\beta_{\rm GPT}(\Gamma)$, respectively. It is clear that

$$\beta_{\rm C}(\Gamma) \leq \beta_{\rm QM}(\Gamma) \leq \beta_{\rm GPT}(\Gamma) \tag{5}$$

by definition.

Maximum values.—Prepared by the above discussion, for given hypergraph Γ, we can define the adjacency graph G on the vertex set V: two $i,j \in V$ are joined by an edge if and only if there exists a $C \in \Gamma$ such that both $i,j \in C$. Then,

$$\beta_{\rm C}(\Gamma) = \alpha(G), \quad \beta_{\rm QM}(\Gamma) = \vartheta(G), \tag{6}$$

where $\alpha(G)$ is the independence number of the graph, *i.e.* the maximum number of pairwise disconnected vertices, and $\vartheta(G)$ is the Lovász ϑ-function of G [20–22], defined as follows: First, an *orthonormal representation* (OR) of a graph is a set of unit vectors associated to the vertices such that two vectors are orthogonal if the corresponding vertices are adjacent. Then,

$$\vartheta(G) := \max \sum_{i=1}^{n} |\langle\psi|v_i\rangle|^2, \tag{7}$$

where the maximum is taken over all unit vectors $|\psi\rangle$ (in Eucledian space) and ORs $\{|v_i\rangle : i = 1,\dots,n\}$ of G [23]. Note that on the right hand side, we can get rid of $|\psi\rangle$ by observing

$$\max_{|\psi\rangle} \sum_{i=1}^{n} |\langle\psi|v_i\rangle|^2 = \left\| \sum_{i=1}^{n} |v_i\rangle\langle v_i| \right\|_\infty . \tag{8}$$

Furthermore, $\vartheta(G)$ is given by a semidefinite program [21], which explains the key importance of this number for combinatorial optimization and zero-error information theory – indeed $\vartheta(G)$ is an upper bound to the Shannon capacity of a graph [21].

Observe that this says in particular that when discussing classical and quantum models, we never need to consider contexts of more than two events. Indeed, it is a (nontrivial) property of these models that if in a set of events any pair is compatible and exclusive, then so is the whole set; more generalized probabilistic theories do not have this property, cf. [24].

To prove Eq. (6), we notice that for a given probabilistic model, the expectation is always maximized on an extremal, *i.e.* pure, state. In the classical case, this amounts to choosing a point $\omega \in \Omega$, so that $w_i := P_i(\omega)$ is a 0-1-valuation of the set V. By definition, it has the property that, in each hyperedge $C \in \Gamma$, at most one element is marked 1, and β is simply the number of marked elements. It is clear that the marked elements form an independent set in Γ (and equivalently in the graph G). In the quantum case, let the maximizing state be given by a unit vector $|\psi\rangle$, and for each i, $\langle\psi|P_i|\psi\rangle = |\langle\psi|v_i\rangle|^2$, for $|v_i\rangle := P_i|\psi\rangle/\sqrt{\psi|P_i|\psi\rangle}$. This clearly is an orthogonal representation of G, in fact the projectors $|v_i\rangle\langle v_i|$ form another quantum model of Γ, with the same maximum value of β, which by the definition we gave earlier is just Lovász' $\vartheta(G)$.

Each graph G where $\alpha(G) < \vartheta(G)$ thus exhibits a limitation of classical noncontextuality, which can be witnessed in experiments with an appropriate set of projectors, and on an appropriate state. In this sense, each such graph provides a proof of the KS theorem.

Taking $n \geq 5$ odd and applying a result from [21] to $G = C_n$, the n-cycle, one obtains the noncontextual quantum bounds

$$\beta_{\rm QM}(n) = \vartheta(C_n) = \frac{n\cos(\pi/n)}{1+\cos(\pi/n)}, \tag{9}$$

where C_n denotes the n-cycle. After some algebra, the quantum bound for the analogue of (2) can be written as

$$\beta'_{\rm QM}(n) = \frac{n}{2}\left[-1 + 3\cos\left(\frac{\pi}{n}\right)\right]\sec^2\left(\frac{\pi}{2n}\right), \tag{10}$$

for all state space dimensions larger or equal to 3; the same result was obtained recently by Liang, Spekkens, and Wiseman [24].

We remark here that there are also "state-independent" KS proofs [4, 25, 26]: these are given by quantum noncontextual models of a graph G such that $\sum_i \langle P_i\rangle > \alpha(G)$ for every state. The proofs in the literature typically have this property, as they are based on rank-one $P_i = |v_1\rangle\langle v_i|$, and for each $j \in V$ there exists $C \in \Gamma$ such that $j \in C$ and $\sum_{i\in C} P_i = \mathbb{1}$ (*i.e.*, each P_j is part of a context that is already a complete measurement; the $|v_i\rangle$ forming a complete orthonormal basis). Due to the symmetric structure of most KS proofs, $\sum_i P_i$ turns out to be proportional to the identity, so β is independent of the state.

It is known that $\vartheta(G)$ can be much larger than $\alpha(G)$; in particular, it is known that (for appropriate, arbitrarily large n) there are graphs G with $\vartheta(G) \approx \sqrt{n}$ but $\alpha(G) \approx 2\log n$, and others with $\vartheta(G) \approx \sqrt[4]{n}$ but $\alpha(G) = 3$ [27]. Hence, the quantum violation of noncontextual inequalities can be arbitrarily large.

Description of the probability sets.—We now show that arbitrary linear functions can be optimized over $\mathcal{E}_{\rm QM}(\Gamma)$

as semidefinite programs: for an arbitrary vector $\vec{\lambda} \in R^V$, let

$$\vec{\lambda}(\mathcal{E}_{\rm QM}(\Gamma)) = \max \sum_i \lambda_i p_i \text{ s.t. } \vec{p} \in \mathcal{E}_{\rm QM}(\Gamma). \tag{11}$$

First of all, without loss of generality, all λ_i are non-negative; this follows because $\mathcal{E}_{\rm QM}(\Gamma)$ is a corner and hence $\vec{\lambda}(\mathcal{E}_{\rm QM}(\Gamma))$ is unchanged when we replace all negative λ_i by 0. Now recall that $p_i = |\langle\psi|v_i\rangle|^2$ for some unit vector $|\psi\rangle$ and an orthonormal representation $\{|v_i\rangle \propto P_i|\psi\rangle\}$ of G. Hence,

$$\begin{aligned}
\vec{\lambda}\vec{p} &= \sum_{i\in V} \lambda_i \langle\psi|P_i|\psi\rangle \\
&= \langle\psi| \left(\sum_{i\in V} \lambda_i |v_i\rangle\langle v_i| \right) |\psi\rangle \\
&= \langle t| \left(\sum_{ij\in V} \sqrt{\lambda_i\lambda_j} \langle v_i|v_j\rangle |i\rangle\langle j| \right) |t\rangle \\
&= \sum_{ij\in V} \sqrt{\lambda_i\lambda_j}\, \overline{t_i} t_j \langle v_i|v_j\rangle \\
&= \operatorname{tr} T\Lambda.
\end{aligned} \tag{12}$$

for an appropriate vector $|t\rangle \in \mathbb{C}^V$, because the Hermitian matrices in the second and third line (the latter a Gram matrix) have the same spectrum. The matrices T and Λ in the last line are defined as follows:

$$\begin{aligned}
\Lambda_{ij} &= \sqrt{\lambda_j\lambda_i}, \\
T_{ij} &= \overline{t_i} t_j \langle v_i|v_j\rangle.
\end{aligned}$$

When varying over quantum models of G and states ψ, the matrix T varies over all semidefinite $T \geq 0$ such that $\operatorname{tr} T = 1$ and $T_{ij} = 0$ whenever $i \sim j$ are connected by and edge in G. *I.e.*,

$$\begin{aligned}
\vec{\lambda}(\mathcal{E}_{\rm QM}(\Gamma)) &= \max \operatorname{tr} \Lambda T \\
&\qquad \text{s.t. } T \geq 0,\ \operatorname{tr} T = 1,\ i\sim j \Rightarrow T_{ij} = 0,
\end{aligned} \tag{13}$$

which is indeed a semidefinite program. □

Closing this semidefinite discussion, the above primal SDP above has a dual, as follows:

$$\begin{aligned}
\vec{\lambda}(\mathcal{E}_{\rm QM}(G)) &= \min s \text{ s.t. } s\mathbb{1} \geq S,\ S = S^\dagger, \\
&\qquad (i \not\sim j \text{ or } i = j) \Rightarrow S_{ij} = \Lambda_{ij}.
\end{aligned} \tag{14}$$

The value $\vec{\lambda}(\mathcal{E}_{\rm QM}(\Gamma))$ is known as a *weighted Lovász number* (or *ϑ-function*) [20].

The previous discussion implies that not only function optimization, but also membership in $\mathcal{E}_{\rm QM}(\Gamma)$ is an efficient convex problem: there is a polynomial-time algorithm that, given a vector $\vec{p}$, tests whether it is in $\mathcal{E}_{\rm QM}(\Gamma)$ or not. This follows from general considerations of convex optimisation [29–31].

Does there exist such a nice and efficient description also for the classical set $\mathcal{E}_{\rm C}(\Gamma)$? The fact that the maximum of β over it is the independence number $\alpha(G)$, which is well-known to be NP complete, means that the answer is "no." In fact, $\mathcal{E}_{\rm C}(\Gamma)$ encodes the independence numbers $\alpha(G|_S)$ of all induced subgraphs of G on subsets $S \subset V$, and the best description that we have is as the following 0-1-polytope:

$$\mathcal{E}_{\rm C}(\Gamma) = \operatorname{conv}\{\vec{\sigma} : \sigma_i \in \{0,1\},\ i\sim j \Rightarrow \sigma_i\sigma_j = 0\}. \tag{15}$$

Turning to generalized probabilistic models, $\beta_{\rm GPT}(\Gamma)$ seems at first much harder to characterize, and we need to look at the full hypergraph structure. Indeed, it is this value that we should with good reason consider as the "algebraic bound" for β. After all, it is the largest value we can assign to it under the most general interpretation of the events $i \in V$ in a probabilistic model that obeys the Gleason property.

The difficulty in evaluating $\beta_{\rm GPT}(\Gamma)$ lies in capturing the constraint that the P_i have to be sums of extremal, normalized effects in the generalized probabilistic theory. If we relax this condition simply to P_i having to be an effect, we arrive at what we would like to call a *fuzzy model*, which formalizes the notion that all $\{P_i : i \in C\}$ are compatible, but not necessarily exclusive events: so we are left with Gleason's constraints $0 \leq \langle P_i\rangle \leq 1$ and for all $C \in \Gamma$, $\sum_{i\in C} \langle P_i\rangle \leq 1$. Denote the (convex) set of all expectations $\big(\langle P_i\rangle\big)_{i\in V}$ when varying over models and their states by $\mathcal{E}_{\rm F}(\Gamma)$.

$$\beta_{\rm GPT}(\Gamma) = \beta_{\rm F}(\Gamma) = \alpha^*(\Gamma), \tag{16}$$

where $\alpha^*(\Gamma)$ is the so-called *fractional packing number* of the hypergraph Γ, defined by the following intuitive linear program:

$$\begin{aligned}
\alpha^*(\Gamma) &= \max \sum_{i\in V} w_i \\
&\qquad \text{s.t. } \forall i\ 0 \leq w_i \leq 1 \text{ and } \forall C \in \Gamma\ \sum_{i\in C} w_i \leq 1.
\end{aligned} \tag{17}$$

The vectors $\vec{w}$ are known as *fractional packings of* Γ. To prove Eq. (16), observe on the one hand that, for given fuzzy model $\{P_i\}$ and a state ρ, the weights $w_i = \langle P_i\rangle$ form a fractional packing. Furthermore, a fractional packing $\{w_i\}$ *is* a fuzzy noncontextual model for the unique generalized probabilistic theory in $\mathbb{R}$, with the usual ordering and unit 1; the state is the identity. (In other words, $\mathcal{E}_{\rm F}(\Gamma)$ is precisely the polytope of fractional packings of Γ.)

Conversely, given a fractional covering $\vec{w}$, we now show that there is an appropriate generalized probabilistic model with effects P_i and a state, such that $w_i = \langle P_i \rangle$. Indeed, as the set of normalized states we choose $\mathcal{S} = 1 \oplus \mathcal{E}_{\mathrm{F}}(\Gamma)$, spanning a cone $\mathbb{R}_{\geq 0}\mathcal{S} \subset \mathbb{R} \oplus \mathbb{R}^V$. The dual cone (with respect to the usual Euclidean inner product) is the set of positive observables: $\mathcal{S}' =: \mathcal{P} \subset \mathbb{R} \oplus \mathbb{R}^V$ with unit element $u = 1 \oplus 0^V \in \mathcal{P}$, which is 1 precisely on the affine hyperplane spanned by $\mathcal{S}$. Now, for each $i \in V$, let $P_i = 0 \oplus \delta_i \in \mathcal{P}$ be the i-th standard basis vector. Clearly, for given fractional covering (*i.e.*, state) $\vec{w}$ and all $i \in V$, $\langle P_i \rangle = w_i$. Hence, all that remains to show is that these P_i and all $Q_C = u - \sum_{i \in C} P_i$ are extremal and normalized (assuming that $C \in \Gamma$ is a maximal element). Concerning normalization, observe that the fractional packings δ_i and 0 (the all-zero assignment) yield proper states. Regarding extremality, observe that on S, P_i and all Q_C are non-negative; furthermore, the equations $\langle P_i \rangle = 0$ and $\langle Q_C \rangle = 0$ each define hyperplanes intersecting $\mathbb{R}_{\geq 0}\mathcal{S}$ in a convex set of dimension $|V|$, *i.e.* these equations define facets of the cone $\mathbb{R}_{\geq 0}\mathcal{S}$, meaning that all $\mathbb{R}_{\geq 0}P_i$ and $\mathbb{R}_{\geq 0}Q_C$ are indeed extremal rays.

Note that by the above argument we proved in fact that $\mathcal{E}_{\mathrm{GPT}}(\Gamma) = \mathcal{E}_{\mathrm{F}}(\Gamma)$, the set of fractional packings. This means that *any* linear function of expectation values can be optimized over $\mathcal{E}_{\mathrm{GPT}}(\Gamma)$ as a linear program; likewise, checking whether $\vec{p}$ is in $\mathcal{E}_{\mathrm{GPT}}(\Gamma)$ is a linear programming feasibility problem. □

For an example, for the n-cycles above, $\alpha^*(C_n) = n/2$, regardless of the parity of n, which is strictly larger than $\vartheta(C_n)$ for all odd $n \geq 5$. Again, we know of arbitrarily large separations: there are hypergraphs Γ such that the adjacency graph G is the complete graph K_n, hence $\alpha(G) = \vartheta(G) = 1$, yet $\alpha^*(\Gamma) \gg 1$ [28].

Remark: Our $\mathcal{E}_{\mathrm{QM}}(\Gamma)$ equals Knuth's set $\mathtt{TH}(\overline{G})$ [20] for the adjacency graph G of Γ; likewise our $\mathcal{E}_{\mathrm{C}}(\Gamma)$ equals his $\mathtt{STAB}(G)$ and if Γ is the hypergraph of all cliques in G, also $\mathcal{E}_{\mathrm{GPT}}(\Gamma) = \mathtt{QSTAB}(G)$. Knuth introduced these sets in his treatment of the (weighted) Lovász ϑ-function, independence numbers and fractional packing numbers, in an attempt to explain the so-called "sandwich theorem" structurally.

Bell inequalities.—Where does nonlocality come into this? After all, Bell inequalities exploit locality in the form that one party's measurement is compatible with another party's, and that the former's outcomes are independent of the latter's choices (*i.e.*, insensitive to different contexts). We can model this also in our setting, by going to the atomic events, which are labelled by a list of settings and outcomes for each party. For instance, for bipartite scenarios, let Alice and Bob's settings be $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, respectively, and their respective outcomes be $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Then, we construct a graph with vertex set $V = \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y}$ and edges $abxy \sim a'b'x'y'$ if and only if ($x = x'$ and $a \neq a'$) or ($y = y'$ and $b \neq b'$), encoding precisely that two events in V are connected in the graph if and only if they are compatible and mutually exclusive (as events in the Bell experiment as a whole). Let Γ be the hypergraph of all cliques in G.

We can now discuss classical noncontextual, quantum and generalized models for this graph, and hence also noncontextual inequalities, restricting as above to linear functions $\vec{\lambda}\vec{p}$ of the vector of the probabilities $p_{ab|xy} = \langle P_{abxy} \rangle$, with with non-negative coefficient vector $\vec{\lambda}$. Note that any Bell inequality can always be rewritten in such a form, by removing negative coefficients using the identity $-p_{ab|xy} = -1 + \sum_{a'b' \neq ab} p_{a'b'|xy}$ for all x, y, a, and b. These equations are not automatically realized in the sets $\mathcal{E}_{\mathrm{X}}(\Gamma)$, $\mathrm{X} = \mathrm{C}, \mathrm{QM}, \mathrm{GPT}$ – as indeed in the underlying (classical, quantum or generalized) model it needs not hold that $\sum_{ab} P_{abxy}$ is the unit element, for any x, y. Hence, define for any class of models $\mathrm{X} = \mathrm{C}, \mathrm{QM}, \mathrm{GPT}$,

$$\mathcal{E}^1_{\mathrm{X}}(\Gamma) := \mathcal{E}_{\mathrm{X}}(\Gamma) \cap \left\{ \vec{p} : \forall xy \ \sum_{ab} p_{ab|xy} = 1 \right\}, \tag{18}$$

the set of probability assignments consistent with the contextuality structure Γ, and in addition satisfying normalization.

In the appendix we prove (which is not too difficult) that $\mathcal{E}^1_{\mathrm{C}}(\Gamma)$ is precisely the set of correlations explained by local hidden variable theories, and that $\mathcal{E}^1_{\mathrm{GPT}}(\Gamma)$ are exactly the no-signalling correlations. Furthermore, to calculate the local hidden variable value Ω_c of a given Bell inequality with non-negative coefficient vector $\vec{\lambda}$, it holds that

$$\Omega_c = \vec{\lambda}(\mathcal{E}^1_{\mathrm{C}}(\Gamma)) = \vec{\lambda}(\mathcal{E}_{\mathrm{C}}(\Gamma)). \tag{19}$$

In this sense, any Bell inequality is at the same time a noncontextual inequality for the underlying graph G.

With classical and no-signalling correlations taken care of, we turn our attention to the quantum case. Once again, we refer the reader to the appendix for a proof that the following subset of $\mathcal{E}^1_{\mathrm{QM}}(\Gamma)$ is precisely the set of correlations obtainable by local quantum measurements on a bipartite state (where "local" means that all operators of one party commute with all operators of another party):

$$\mathcal{E}^{\mathbb{1}}_{\mathrm{QM}}(\Gamma) = \left\{ \left(\langle P_{abxy} \rangle\right)_{abxy} : \forall xy \ \sum_{ab} P_{abxy} = \mathbb{1} \right\}. \tag{20}$$

I.e., we add the completeness relation for the measurements in the model. This of course also means that for a given Bell inequality with coefficients $\vec{\lambda}$, the maximum quantum value is

$$\Omega_q = \vec{\lambda}(\mathcal{E}^{\mathbb{1}}_{\mathrm{QM}}(\Gamma)). \tag{21}$$

For the time being we do not know whether the set of quantum correlations, *i.e.* $\mathcal{E}^{\mathbb{1}}_{\mathrm{QM}}(\Gamma)$, is efficient to characterize. It follows, however, from the above considerations and the general theory of convex optimization [29–31] that the – potentially larger – set $\mathcal{E}^1_{\rm QM}(\Gamma)$ can be decided efficiently. In fact, we shall see directly that the maximum values $\vec{\lambda}(\mathcal{E}^1_{\rm QM}(\Gamma))$ are computed to arbitrary precision by semidefinite programming, thus providing efficient upper bounds to Ω_q.

Namely, for $M \gg 1$, consider the linear function

$$\vec{\lambda}\cdot\vec{p}+M\sum_{xy}\left(-1+\sum_{ab}p_{ab|xy}\right) = \left(\vec{\lambda}+M\vec{1}\right)\cdot\vec{p}-M|\mathcal{X}\times\mathcal{Y}|, \tag{22}$$

which encodes $\vec{\lambda}$ plus a large negative penalty for any xy such that $\sum_{ab} p_{ab|xy} < 1$, and maximize it over the full set of quantum models, $\mathcal{E}_{\rm QM}(\Gamma)$. [Note that "$\leq 1$" is guaranteed by the Gleason property, which is valid in this set.] Clearly, all the values $(\vec{\lambda}+M\vec{1})(\mathcal{E}_{\rm QM}(\Gamma))$ are instances of the semidefinite programs discussed earlier, and as $M \to \infty$,

$$(\vec{\lambda}+M\vec{1})(\mathcal{E}_{\rm QM}(\Gamma)) - M|\mathcal{X}\times\mathcal{Y}| \longrightarrow \vec{\lambda}(\mathcal{E}^1_{\rm QM}(\Gamma)). \tag{23}$$

Implementing this for example for the CHSH inequality [5], we recover the Tsirelson bound $2\sqrt{2}$ [32] – see the appendix for details. On the other hand, for the I_{3322} inequality [33] the method yields the upper bound 0.251 47 on the quantum value; the currently best upper bound is slightly smaller: 0.250 875 56 [34], from which we conclude that in general, $\mathcal{E}^{\mathbb{1}}_{\rm QM}(\Gamma)$ is strictly contained in $\mathcal{E}^1_{\rm QM}(\Gamma)$ – once more, see the appendix for details. [As an aside, we note that in the latter case, maximizing over $\mathcal{E}_{\rm QM}(\Gamma)$ gives the even much larger bound 0.4114 – so, unlike classical models, in the quantum the probability normalization is not for granted.]

Conclusions.—Notice that the previous exposition bears striking similarity to the discussion of the no-signalling property in the context of classical, quantum, or more general correlations. Indeed, as it was observed by Popescu and Rohrlich [35], and Tsirelson [36], the no-signalling principle is not enough to explain the scope of quantum correlations; for instance, for the CHSH inequality, the classical bound is 2, the quantum bound is $2\sqrt{2}$, while the algebraic bound 4 is attainable under the most general no-signalling correlations. Likewise here: operational models obeying the Gleason constraint include classical and quantum ones, but they definitely go beyond these two. One might ask: why is nature not even more contextual than quantum mechanics?

Unlike Bell inequalities, here we see that the maximum quantum violation is always efficiently computable, as it is the solution to a semidefinite program, and these are solvable in polynomial time. Thanks to the general machinery of convex optimisation problems [29–31], this also means that membership of a probability assignment $\vec{p}$ in $\mathcal{E}_{\rm QM}(\Gamma)$ can be tested efficiently, despite the fact that the set is not itself defined directly by semidefinite constraints. Generalized models are captured instead entirely by linear inequalities and linear programming – in particular, also here all maximum violations of noncontextual inequalities can be computed efficiently, as linear programs. At the other end of the spectrum, the noncontextual set $\mathcal{E}_{\rm C}(\Gamma)$ is the convex hull of many, but easy to describe points, but its characterisation in terms of inequalities is computationally hard, and so are maximum values such as $\beta_{\rm C}(\Gamma)$, which can be as hard as NP complete.

The sets of probability assignments compatible with noncontextual, quantum and generalized operational models are different from each other even in the simplest nontrivial case, that of the pentagon, as witnessed by the values 2, $\sqrt{5}$, and $5/2$ for $\beta(5)$, respectively. Especially the gap between $\sqrt{5}$ for quantum and $5/2$ for generalized models is noteworthy, because the latter value is attained by putting weight $1/2$ to each vertex in a Gleason assignment of probabilities to each of the five vertices of C_5. It had been noted by other authors before, that the Gleason constraint on finite sets of vectors allows assignments incompatible with quantum theory [14]. We believe that here we clarified this observation further, since we showed that each such assignment originates in fact from a sound operational model based on generalized probabilistic theories. Each vertex is assigned an event such that, with respect to the given state, any adjacent pair is "complete" in the sense that the probabilities add up to 1. It is easy to see that quantum mechanics cannot yield this, as it would require successive subspace projectors to be orthogonal complements of each other.

We close by highlighting some open questions: Looking back, it is the insistence on exclusiveness of events, and the dropping of completeness relations, that made the KCBS inequalities and our generalizations possible; not insisting on effects having to sum to unity (always prominent in the "usual" KS proofs) also seems responsible for the fact that we obtain a semidefinite program for the maximum quantum value. On the other hand, how to incorporate this as an additional constraint in the SDP?

As this seems to mark exactly the difference between nonlocal quantum values and quantum violations of generalized KCBS inequalities, the question arises: how good is the latter as a bound on the former? And how does it relate to upper bounds obtained from the Navascués-Pironio-Acín hierarchy [37]?

Acknowledgments.—We thank P. Badziąg, J. Barrett, I. Bengtsson, T. Cubitt, A. Harrow, A. Klyachko, D. Leung, J.-Å. Larsson, W. Matthews, J. Oppenheim, and K. Svozil for conversations.

The present research was supported by the European Commission, the U.K. EPSRC, the Royal Society, the British Academy, the Royal Academy of Engineering, the Spanish MCI Project No. FIS2008-05596, and by the National Research Foundation as well as the Ministry of Education of Singapore.

---

* adan@us.es
† simoseve@gmail.com
‡ a.j.winter@bris.ac.uk

[1] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008).
[2] A. M. Gleason, J. Math. Mech. **6**(6), 885 (1957).
[3] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[4] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[5] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[6] R. Łapkiewicz *et al.* (unpublished).
[7] A. Cabello, Phys. Rev. Lett. **101**, 210401 (2008).
[8] P. Badziąg, I. Bengtsson, A. Cabello, and I. Pitowsky, Phys. Rev. Lett. **103**, 050401 (2009).
[9] G. Kirchmair *et al.*, Nature (London) **460**, 494 (2009).
[10] E. Amselem, M. Rådmark, M. Bourennane, and A. Cabello, Phys. Rev. Lett. **103**, 160405 (2009).
[11] A. Cabello, Phys. Rev. Lett. **104**, 220401 (2010).
[12] O. Gühne *et al.*, Phys. Rev. A **81**, 022121 (2010).
[13] A. Cabello, Phys. Rev. A **82**, 032110 (2010).
[14] R. Wright, in *Mathematical Foundations of Quantum Mechanics*, edited by A. R. Marlow (Academic Press, San Diego, 1978), p. 255.
[15] G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, New York, 1963).
[16] G. Ludwig, Z. Phys. **181**(3), 233 (1964).
[17] G. Ludwig, Comm. Math. Phys. **4**(5), 331 (1967).
[18] A. S. Holevo, *Statistical Structure of Quantum Theory* (Springer, Berlin, 2001).
[19] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[20] D. Knuth, Elec. J. Comb. **1**, 1 (1994).
[21] L. Lovász, IEEE Trans. Inf. Theory **25**, 1 (1979).
[22] J. Körner and A. Orlitsky, IEEE Trans. Inf. Theory, **44**, 2207 (1998).
[23] L. Lovász, *Geometric Representations of Graphs*, http://www.cs.elte.hu/∼lovasz/geomrep.pdf
[24] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, arXiv:1010.1273 [quant-ph].
[25] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993).
[26] A. Cabello, J. M. Estebaranz, and G. García-Alcaine, Phys. Lett. A **212**, 183 (1996).
[27] R. Peeters, Combinatorica **16**, 417 (1996).
[28] T. S. Cubitt, D. W. Leung, W. Matthews, and A. Winter, arXiv:1003.3195 [quant-ph].
[29] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization* (Springer, Berlin, 1988).
[30] D. Bertsimas and S. Vempala, J. ACM **51**, 540 (2004).
[31] Y.-K. Liu, arXiv:0712.3041 [quant-ph].
[32] B. S. Cirel'son [Tsirelson], Lett. Math. Phys. **4**, 93 (1980).
[33] N. Brunner and N. Gisin, Phys. Lett. A **327**, 3162 (2008).
[34] K. F. Pál and T. Vértesi, arXiv:1006.3032 [quant-ph].
[35] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[36] B. S. Tsirelson, Hadronic J. Suppl. **8**, 329 (1993).
[37] M. Navascués, S. Pironio, and A. Acín, New J. Phys. **10**, 073013 (2008).

Non-locality: proofs

Here we prove the claims in the Bell inequality section.

(i) Proof that $\mathcal{E}^1_{\rm C}(\Gamma)$ = local realistic correlations. If the A's and B's form a (deterministic) classical local hidden variable model, then the products $P^{ab}_{xy} = A^a_x B^b_y$ are a classical noncontextual model for the graph G. Since for each x and y there is exactly one a, b, respectively, such that $A^a_x = B^b_y = 1$, the normalization condition is fulfilled, too.

Vice versa, given any deterministic noncontextual model P^{ab}_{xy} for G we show how to construct local hidden variables A^a_x and B^b_y (taking values 0 and 1) such that $P^{ab}_{xy} \leq A^a_x B^b_y$; using the probability normalization, this must be an equality. Namely, assume $P_{abxy} = 1$ for any quadruple $abxy$. Then, thanks to the graph G, for any $a' \neq a$ and any y and b, $P_{a'b'xy'} = 0$. In other words, for every x, there is at most one a such that $P_{ab'xy'} = 1$ for any $b'y'$. Choose this a (or else an arbitrary one) to let $A^a_x = 1$ and all other $A^{a'}_x = 0$. Likewise for B^b_y, and we clearly obtain the claim. □

(ii) Proof that $\mathcal{E}^1_{\rm GPT}(\Gamma)$ = no-signalling correlations. Let $\vec{p} \in \mathcal{E}_{\rm GPT}(\Gamma)$ such that for all xy, $\sum_{ab} p_{ab|xy} = 1$. We have to show the no-signalling relations,

$$\forall ax \forall yy' \quad \sum_b p_{ab|xy} = \sum_b p_{ab|xy'},$$
$$\forall by \forall xx' \quad \sum_a p_{ab|xy} = \sum_a p_{ab|x'y}.$$

To prove this, note for fixed x, y and y', that the vertices

$$\{abxy : b \in \mathcal{B}\} \cup \{a'bxy' : a' \in \mathcal{A} \setminus a,\, b \in \mathcal{B}\}$$

form a clique in G, hence

$$\sum_b p_{ab|xy} + \sum_{a' \neq a, b} p_{a'bxy'} \leq 1,$$

which implies $\sum_b p_{ab|xy} \leq \sum_b p_{ab|xy'}$ for arbitrary y and y'. By symmetry, equality must hold. □

(iii) Proof that $\vec{\lambda}(\mathcal{E}^1_{\rm C}(\Gamma)) = \vec{\lambda}(\mathcal{E}_{\rm C}(\Gamma))$. Recall from (i) that we can find, for any deterministic noncontextual model P^{ab}_{xy}, local hidden variables A^a_x and B^b_y (taking values 0 and 1) such that $P^{ab}_{xy} \leq A^a_x B^b_y$. The right hand side is in evidently in $\mathcal{E}^1_{\rm C}(\Gamma))$. Hence, for the purpose of maximizing a objective function with non-negative coefficients $\vec{\lambda}$, we may restrict to $\mathcal{E}^1_{\rm C}(\Gamma))$. □

(iv) Proof that $\mathcal{E}^{\mathbb{1}}_{\rm QM}(\Gamma)$ = quantum correlations. We face a problem like in (i): given operators P_{abxy} forming a quantum model of G, we have to define projector valued

measurements $(A_x^a)_{a\in\mathcal{A}}$ and $(B_y^b)_{b\in\mathcal{B}}$ such that $[A_x^a, B_y^b] = 0$ and $P_{abxy} = A_x^a B_y^b$.

There are obvious candidates for these "local" measurements given as marginals of P_{abxy}:

$$A_x^a = \sum_{b'} P_{ab'xy} \quad \text{(for any } y\text{)},$$
$$B_y^b = \sum_{a'} P_{a'bxy} \quad \text{(for any } x\text{)},$$

which raises the immediate issue that, *a priori.* the right hand sides may not be independent of y and x, respectively. Denote the right hand sides above by A_{xy}^a and B_{xy}^b. We show that the assumption of completeness, $\sum_{ab} P_{abxy} = \mathbb{1}$, implies that A_{xy}^a is independent of y, B_{xy}^b independent of x. Indeed, observe that for any $a' \neq a$ and any y, y', b, and b', $P_{abxy} \perp P_{a'b'xy'}$, which by summation implies that

$$A_{xy}^a \perp \sum_{a'\neq a} A_{xy'}^{a'} = \mathbb{1} - A_{xy'}^a,$$

i.e., $A_{xy}^a \leq A_{xy'}^a$ for all y and y'. By symmetry we hence must have $A_{xy}^a = A_{xy'}^a$ and likewise $B_{xy}^b = B_{xy'}^b$.

Now, observe finally

$$A_x^a B_y^b = \sum_{a'b'} P_{ab'xy} P_{a'bxy} = P_{abxy} = B_y^b A_x^a,$$

and we are done. □

(v) Example CHSH. Here, $\mathcal{A} = \mathcal{B} = \mathcal{X} = \mathcal{Y} = \{0,1\}$ and $\vec{\lambda}$ encodes the winning condition for the CHSH (or PR) game:

$$\lambda_{abxy} = \begin{cases} 1 & : \ a \oplus b = xy, \\ 0 & : \ \text{otherwise.} \end{cases} \tag{24}$$

The CHSH inequality expresses the fact that $\Omega_{\rm c} = 3$ while $\Omega_{\rm q} = 2 + \sqrt{2}$.

Constructing the graph and the matrices Λ and T by hand is easy: G has 16 vertices, so the matrices are also 16×16. Since Λ is rather sparse, this allows us immediately to reduce it to a graph G' on 8 vertices with new Λ-matrix equal to J, the all-1-matrix. The graph is the $(1,4)$-circulant graph on 8 vertices; one can obtain it by joining antipodal vertices in the 8-cycle C_8. So, we find that $\lambda(\mathcal{E}_{\rm QM}(G)) = \vartheta(G')$, and the latter is easily evaluated to $2 + \sqrt{2}$, using the dual characterisation of Lovász (*i.e.* our dual SDP).

(vi) Example $I_{\mathbf{3322}}$. This is a Bell inequality for 3 settings for each Alice and Bob, each measurement having binary output. In the form found in [33] it reads

$$\begin{aligned} &-2\langle A_0^0\rangle - \langle A_1^0\rangle - \langle B_0^0\rangle \\ &\quad + \langle A_0^0 B_0^0\rangle + \langle A_0^0 B_1^0\rangle + \langle A_0^0 B_2^0\rangle \\ &\quad + \langle A_1^0 B_0^0\rangle + \langle A_1^0 B_1^0\rangle - \langle A_1^0 B_2^0\rangle \\ &\quad + \langle A_2^0 B_0^0\rangle - \langle A_2^0 B_1^0\rangle \leq 0, \end{aligned} \tag{25}$$

and the value 0 is the maximum attainable under local hidden variables. One form of the objective function with non-negative coefficients, using the above substitution trick, is $\vec{\lambda} \cdot \vec{p}$, with the vector $\vec{\lambda} \in \mathbb{R}^{36}$ being given by the following table:

$xa \setminus yb$	00	01	10	11	20	21
00	1	0	1	0	1	0
01	0	0	1	1	1	1
10	1	1	1	0	0	1
11	0	1	1	1	1	1
20	1	0	0	1	0	0
21	0	0	1	1	0	0

The classical bound is $\Omega_{\rm c} = 6$, while the best known quantum violation attains a value $6.250\,875\,384 \leq \Omega_{\rm q}$; on the other hand, it is known that $\Omega_{\rm q} \leq 6.250\,875\,56$, by going as far up in the Navascués-Pironio-Acín hierarchy [37] as was computationally feasible (almost the fourth level); the conjecture is that this is essentially the optimal value, although there is still disagreement from the 7th digit on. It is also conjectured that to attain the quantum limit, an infinitely large entangled state is required – in [34] a candidate sequence of larger and larger states and measurements is presented which give better and better values suggested to converge to the optimum.

The game context graph G on 36 vertices in not constructed explicitly here, though it is easy. Looking at the primal SDP, and noticing that only 20 out of 36 components of $\vec{\lambda}$ are populated, and then only by 1's, one sees – cf. the CHSH case – that, by constructing the induced subgraph G' of the context graph on the 20 vertices $abxy$ with $\lambda_{abxy} = 1$, we obtain $\vec{\lambda}(\mathcal{E}_{\rm QM}(G)) = \vartheta(G') \approx 6.4114$.

This is an instance of the probabilities simply not adding up to 1, in other words: $\vec{\lambda}(\mathcal{E}^1_{\rm QM}(G))$ is strictly smaller. Indeed, a calculation with on SeDuMi resulted in $\vec{\lambda}(\mathcal{E}^1_{\rm QM}(G)) \approx 6.251\,47$.