COMPLETELY POSITIVE SEMIDEFINITE RANK

A. PRAKASH [*], J. SIKORA [†], A. VARVITSIOTIS [‡], AND Z. WEI [§]

Abstract. An $n \times n$ matrix X is called completely positive semidefinite (cpsd) if there exist $d \times d$ Hermitian positive semidefinite matrices $\{P_i\}_{i=1}^n$ (for some $d \geq 1$) such that $X_{ij} = \text{Tr}(P_i P_j)$, for all $i, j \in \{1, \ldots, n\}$. The cpsd-rank of a cpsd matrix is the smallest $d \geq 1$ for which such a representation is possible. In this work we initiate the study of the cpsd-rank which we motivate twofold. First, the cpsd-rank is a natural non-commutative analogue of the completely positive rank of a completely positive matrix. Second, we show that the cpsd-rank is physically motivated as it can be used to upper and lower bound the size of a quantum system needed to generate a quantum behavior.

In this work we present several properties of the cpsd-rank. Unlike the completely positive rank which is at most quadratic in the size of the matrix, no general upper bound is known on the cpsd-rank of a cpsd matrix. In fact, we show that the cpsd-rank can be exponential in terms of the size. Specifically, for any $n \geq 1$, we construct a cpsd matrix of size $2n$ whose cpsd-rank is $2^{\Omega(\sqrt{n})}$. Our construction is based on Gram matrices of Lorentz cone vectors, which we show are cpsd. The proof relies crucially on the connection between the cpsd-rank and quantum behaviors. In particular, we use a known lower bound on the size of matrix representations of extremal quantum correlations which we apply to high-rank extreme points of the n-dimensional elliptope.

Lastly, we study cpsd-graphs, i.e., graphs G with the property that every doubly nonnegative matrix whose support is given by G is cpsd. We show that a graph is cpsd if and only if it has no odd cycle of length at least 5 as a subgraph. This coincides with the characterization of cp-graphs.

Key words. completely positive semidefinite cone, cpsd-rank, Lorentz cone, elliptope, Bell scenario, quantum behaviors, quantum correlations, cpsd-graphs

1. Introduction.

1.1. Setting the scene. Consider a family of vectors $\{v_i\}_{i=1}^n$ such that the angle between any pair of them is at most $\pi/2$. A necessary and sufficient condition for showing that the configuration $\{v_i\}_{i=1}^n$ admits an isometry to some nonnegative orthant is that the $n \times n$ matrix $(\langle v_i, v_j \rangle_{1 \leq i,j \leq n})$, formed by collecting all pairwise inner products of the vectors $\{v_i\}_{i=1}^n$, is *completely positive*. Formally, a symmetric $n \times n$ matrix X is called *completely positive* (cp) if there exist vectors $\{p_i\}_{i=1}^n \subseteq \mathbb{R}_+^d$, for some $d \geq 1$, such that $X_{ij} = \langle p_i, p_j \rangle$, for all $1 \leq i, j \leq n$.

The set of $n \times n$ completely positive matrices, denoted by $\mathcal{CP}^n$, forms a full-dimensional, pointed, closed convex cone whose structure has been extensively studied (e.g. see [3]). Linear conic programming over the $\mathcal{CP}$ cone is particularly interesting due to its expressive power. Specifically, any nonconvex quadratic program having both binary and continuous variables can be cast as a linear conic program over the $\mathcal{CP}$ cone [6]. In particular, this implies that optimization over the $\mathcal{CP}$ cone is intractable. On the positive side, there exist inner [20] and outer [28] semidefinite programming hierarchies that can be used to approximate the $\mathcal{CP}$ cone.

In this work we focus on a generalization of the embeddability question considered above: When can a family of vectors $\{v_i\}_{i=1}^n$ whose pairwise inner products are nonnegative be isometrically embedded into a cone of Hermitian positive semidefinite matrices? Throughout, we denote by $\mathcal{H}_+^d$ the cone of $d \times d$ Hermitian positive semidef-

---

[*]Centre for Quantum Technologies, National University of Singapore and Nanyang Technological University, Singapore. aprakash@ntu.edu.sg.

[†]Centre for Quantum Technologies, National University of Singapore, cqtjwjs@nus.edu.sg.

[‡]Centre for Quantum Technologies, National University of Singapore and Nanyang Technological University, Singapore, avarvits@gmail.com.

[§]Centre for Quantum Technologies, National University of Singapore and Nanyang Technological University, Singapore, weizhaohui@gmail.com.

inite (psd) matrices and by $\mathcal{S}_+^d$ the set of $d \times d$ symmetric psd matrices. Formally, we are asking for the existence of matrices $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$, for some $d \geq 1$, satisfying

$$\langle v_i, v_j \rangle = \mathrm{Tr}(P_i P_j), \text{ for all } 1 \leq i, j \leq n.$$

Since the direct sum of two psd matrices is again psd, the set of $n \times n$ matrices of the form $(\mathrm{Tr}(P_i P_i)_{1 \leq i,j \leq n})$, where $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ (for some $d \geq 1$), forms a convex cone. This set of matrices is denoted by $\mathcal{CS}_+^n$ and is known as the cone of *completely positive semidefinite (cpsd) matrices.*

The $\mathcal{CS}_+^n$ cone was introduced recently to provide linear conic formulations for various quantum graph parameters [21, 29]. Subsequently, it was shown in [30] that underlying these formulations is the fact that the set of quantum behaviors can be expressed as the projection of an affine section of the $\mathcal{CS}_+^n$ cone (cf. Theorem 2).

Clearly, for every $n \geq 1$ we have that $\mathcal{CP}^n \subseteq \mathcal{CS}_+^n \subseteq \mathcal{DNN}^n$, where we denote by $\mathcal{DNN}^n$ the set of $n \times n$ *doubly nonnegative* matrices, i.e., matrices that are positive semidefinite and entrywise nonnegative. For the rightmost inclusion recall that the trace inner product of two psd matrices is a nonnegative scalar. The leftmost inclusion holds since nonnegative vectors correspond to diagonal psd matrices.

It is known that $\mathcal{CP}^n = \mathcal{DNN}^n$ for $n \leq 4$ [24], whereas for $n \geq 5$, all inclusions given above are known to be strict. In particular, it follows from [12] that $\mathcal{CP}^6 \neq \mathcal{CS}_+^6$ and by [10] that $\mathcal{CP}^5 \neq \mathcal{CS}_+^5$. Furthermore, it was shown in [12] that $\mathcal{CS}_+^5 \neq \mathcal{DNN}^5$ and in [21] that $\mathrm{cl}(\mathcal{CS}_+^5) \subsetneq \mathcal{DNN}^5$, where $\mathrm{cl}(\mathcal{CS}_+^n)$ denotes the closure of $\mathcal{CS}_+^n$. Lastly, it was shown in [21] that for any matrix X whose support is a cycle we have that $X \in \mathcal{CP}$ if and only if $X \in \mathcal{CS}_+$. Furthermore, it is known that for every odd cycle C_{2t+1} ($t \geq 2$) there exists a matrix in $\mathcal{DNN} \setminus \mathcal{CP}$ whose support is given by C_{2t+1} (see [3, Theorem 2.12]). Combined with the above, this fact gives a family of matrices in $\mathcal{DNN} \setminus \mathcal{CS}_+$ that are supported by C_{2t+1}, for all $t \geq 2$.

Not many things are known concerning the structure of $\mathcal{CS}_+^n$. In particular it is not known whether $\mathcal{CS}_+^n$ is closed. The closure of $\mathcal{CS}_+^n$ was characterized in [7] as the set of doubly nonnegative matrices that admit a Gram factorization using positive elements in a certain finite von Neumann algebra, an infinite dimensional analogue of $\mathcal{CS}_+$-factorizations (cf. Section 6.1). Furthermore, combining results from [21] and [18] it follows that linear optimization over $\mathcal{CS}_+^n$ is NP-hard.

Given a completely positive matrix $X \in \mathcal{CP}^n$, the smallest integer $d \geq 1$ for which there exist vectors $\{p_i\}_{i=1}^n \subseteq \mathbb{R}_+^d$ satisfying $X_{ij} = \langle p_i, p_j \rangle$, for all $1 \leq i, j \leq n$ is called the *completely positive rank (cp-rank)* of X, and is denoted by cp-rank(X).

A very useful property of the cp-rank is that it admits an atomic reformulation. Specifically, the cp-rank(X) of a matrix $X \in \mathcal{CP}^n$ can be equivalently defined as the smallest $d \geq 1$ for which there exist vectors $\{x_i\}_{i=1}^d \subseteq \mathbb{R}_+^n$ satisfying $X = \sum_{i=1}^d x_i x_i^\mathsf{T}$.

Studying the properties of the cp-rank is a problem that has received significant attention. By the conic analogue of Carathéodory's Theorem (e.g. see [3, Theorem 1.34]) and the atomic reformulation of the cp-rank described above it follows that for any $X \in \mathcal{CP}^n$ we have cp-rank(X) $\leq \binom{n+1}{2}$. At present, the best upper bound is $\frac{n^2}{2} + O(n^{3/2})$, for any $X \in \mathcal{CP}^n$ [4]. Moreover, this upper bound is asymptotically tight with respect to the Drew-Johnson-Loewy lower bound of $\lfloor \frac{n^2}{4} \rfloor$, for $n \geq 4$ [9].

The definition of the $\mathcal{CS}_+$ cone suggests the following generalization of the notion of cp-rank, where nonnegative vectors are replaced by Hermitian psd matrices.

DEFINITION 1. *The* completely positive semidefinite rank *(cpsd-rank) of a matrix* $X \in \mathcal{CS}_+^n$, *denoted by* cpsd-rank(X)*, is defined as the least* $d \geq 1$ *for which there exist matrices* $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ *such that* $X_{ij} = \mathrm{Tr}(P_i P_j)$, *for all* $i, j \in [n]$.

Given a matrix $X \in \mathcal{CS}_+^n$, we refer to any family of matrices $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ such that $X_{ij} = \mathrm{Tr}(P_i P_j)$, for all $i, j \in [n]$, as a $\mathcal{CS}_+$-factorization of X. Furthermore, we call a $\mathcal{CS}_+$-factorization size-optimal if the size of each P_i is equal to cpsd-rank(X).

The notion of the cpsd-rank was introduced recently in [10] (as a variant of the psd-rank) although its properties were not studied there. Our goal in this work is to initiate the study of the cpsd-rank of a cpsd matrix.

The study of the cpsd-rank is motivated as follows. First, the cpsd-rank is a natural non-commutative generalization of the well-studied notion of cp-rank. Second, and most important, we show that the cpsd-rank enjoys strong physical motivation. Specifically, we show that some fundamental questions concerning the cpsd-rank are intimately related to long standing open problems on the foundations of quantum mechanics. This is explained in detail in the following section.

1.2. Physical motivation. A *Bell scenario* is a physical experiment involving two spatially separated parties, Alice and Bob, who perform local measurements on a shared physical system. For our purposes, imagine that Alice and Bob are individually given a closed box, whose inner workings are unknown to both parties. The boxes work as follows: Alice's (resp. Bob's) box has m_A (resp. m_B) different buttons. After each party presses a button, the box displays one out of o_A (resp. o_B) possible outcomes.

It is instructive to think of the boxes as measurement devices and the content of the boxes as a physical system that each party has in his possession. Furthermore, each button corresponds to a choice of measurement that can be performed on the system and the displayed outcome corresponds to the outcome of the measurement.

The object of interest in a Bell scenario are the statistics that can be obtained via such a pair of boxes. Specifically, suppose that Alice and Bob synchronize their clocks and distance themselves from each other so that they cannot communicate. After they are sufficiently far apart, they simultaneously press a button on their box (chosen randomly and independently) and record the button that they pressed and the displayed outcome. After repeating the whole process a sufficient number of times[1] Alice and Bob meet to calculate the joint conditional probabilities $p(ab|xy)$, i.e., the probability that upon pressing buttons $x \in [m_A]$, $y \in [m_B]$, they obtained the outcomes $a \in [o_A]$ and $b \in [o_B]$, respectively. These probabilities are arranged in a vector $\mathbf{p} = (p(ab|xy))$ of length $m_A m_B o_A o_B$ which we call a *behavior*.

Suppose that after the parties compare their statistics they note that for some x, y, a, b it is the case that $p_A(a|x)p_B(b|y) \neq p(ab|xy)$, where $p_A(a|x)$ and $p_B(b|y)$ denote the local marginal distributions of Alice and Bob, respectively. This indicates that the outcomes of the boxes are statistically dependent.

A *local hidden variable* (LHV) model would account for this dependence by asserting that the two systems have interacted at some point in the past, and as a result they both depend on some "hidden" variable k. Once the value of k is taken into account, then the probabilities decouple, i.e., $p(a|x, k)p(b|y, k) = p(ab|xyk)$. Formally, we say that a behavior $\mathbf{p} = (p(ab|xy))$ admits a LHV model (also referred to as being local) if there exist $k_i \geq 0$, $m_a^{s,i} \geq 0$, $n_b^{t,i} \geq 0$ satisfying $\sum_i k_i = 1$, and $\sum_a m_a^{x,i} = \sum_b n_b^{y,i} = 1$ for all x, y, i, such that $p(ab|xy) = \sum_i k_i m_a^{x,i} n_b^{y,i}$, for all a, b, x, y.

In this work we focus on the case where Alice and Bob share a *quantum mechanical system* (e.g. each box contains one of a pair of entangled particles). In this setting the system is governed by the laws of quantum mechanics. In particular, the outcome statistics can be calculated using the mathematical formalism of quantum mechanics

---

[1] To be precise, each time this is repeated each party should receive a new copy of the box.

which we sketch below (see also Section 2).

According to the postulates of quantum mechanics, the state of the quantum system shared by Alice and Bob corresponds to a Hermitian psd matrix ρ acting on $\mathbb{C}^d \otimes \mathbb{C}^d$, with trace equal to 1. The measurement process is described by two families of Hermitian psd operators $\{M_{a|x}\}_a$ and $\{N_{b|y}\}_b$, each acting on $\mathbb{C}^d$, such that $\sum_a M_{a|x} = \sum_b N_{b|y} = I_d$, for all x, y. We say the behavior $\mathbf{p} = (p(ab|xy))$ is *quantum* if there exists a quantum state ρ and measurement operators $\{M_{a|x}\}_a$ and $\{N_{b|y}\}_b$ such that $p(ab|xy) = \mathrm{Tr}\left((M_{a|x} \otimes N_{b|y})\rho\right)$, for all a, b, x, y.

In this work, we refer to a Bell scenario with m_A, m_B measurement settings and o_A, o_B measurement outcomes as an (m_A, m_B, o_A, o_B)-scenario. Furthermore, we denote by $\mathcal{Q}$ (resp. $\mathcal{L}$) the set of quantum (resp. local) behavior. To stress the dependence on the underlying Bell scenario we use the notation $\mathcal{Q}(m_A, m_B, o_A, o_B)$.

Clearly, $\mathcal{L} \subseteq \mathcal{Q}$ and it is one of the pillars of quantum information theory that there exist behaviors that are quantum but do not admit a local hidden variable explanation, i.e., $\mathcal{Q}$ is a strict superset of $\mathcal{L}$ [1, 2]. For an overview on Bell scenarios and the properties of quantum behaviors the reader is referred to [5].

Any quantum behavior $(p(ab|xy))$ is *no-signaling*, i.e., each party's local marginal distribution is independent of the other party's choice of measurement. Algebraically, this is expressed as $\sum_b p(ab|xy) = \sum_b p(ab|xy')$ for all $y \neq y'$, and symmetrically that $\sum_a p(ab|xy) = \sum_a p(ab|x'y)$ for all $x \neq x'$. This implies that the local marginal distributions $(p_A(a|x))$ and $(p_B(b|y))$ are well-defined. In a Bell scenario where all the measurements have binary outcomes, we call a behavior $(p(ab|xy))$ *unbiased* if $p_A(a|x) = p_B(b|y) = 1/2$, for all a, b, x, y.

Given a quantum behavior $\mathbf{p}$, we refer to any ensemble $\{\rho, \{M_{a|x}\}_a, \{N_{b|y}\}_b\}$ such that $p(ab|xy) = \mathrm{Tr}\left((M_{a|x} \otimes N_{b|y})\rho\right)$ for all a, b, x, y as a *quantum representation* of $\mathbf{p}$. A quantum behavior $\mathbf{p} = (p(ab|xy))$ admits multiple quantum representations. We say that $\mathbf{p} \in \mathcal{Q}$ admits a *d-dimensional (quantum) representation* if there exists a quantum representation $\{\rho, \{M_{a|x}\}_a, \{N_{b|y}\}_b\}$, where ρ acts on $\mathbb{C}^d \otimes \mathbb{C}^d$ and $\{M_{a|x}\}_a$ and $\{N_{b|y}\}_b$ each act on $\mathbb{C}^d$. We denote by $\mathcal{D}(\mathbf{p})$ the smallest integer $d \geq 1$ for which the quantum behavior $\mathbf{p} = (p(ab|xy))$ admits a d-dimensional representation.

The starting point for our work is a recent result from [30] which states that the set of quantum behaviors (resp. local) can be expressed as a projection of an affine section of the completely positive semidefinite cone (resp. completely positive cone).

THEOREM 2 ([30]). *Consider a behavior* $\mathbf{p} = (p(ab|xy))$ *and set* $n := m_A o_A + m_B o_B$. *The behavior* $\mathbf{p}$ *is quantum (resp. local) if and only if there exists a matrix* $R \in \mathcal{CS}_+^n$ *(resp. $\mathcal{CP}^n$) indexed by* $([m_A] \times [o_A]) \cup ([m_B] \times [o_B])$ *such that*

$$(1) \qquad \sum_{a=1}^{o_A} \sum_{a'=1}^{o_A} R_{xa,x'a'} = 1, \text{ for all } x, x' \in [m_A];$$

$$(2) \qquad \sum_{a=1}^{o_A} \sum_{b=1}^{o_B} R_{xa,yb} = 1, \text{ for all } x \in [m_A], \ y \in [m_B];$$

$$(3) \qquad \sum_{b=1}^{o_B} \sum_{b'=1}^{o_B} R_{yb,y'b'} = 1, \text{ for all } y, y' \in [m_B];$$

$$(4) \qquad R_{xa,yb} = p(ab|xy), \text{ for all } a \in [o_A], \ b \in [o_B], \ x \in [m_A], \ y \in [m_B].$$

For a fixed $\mathbf{p}$ we denote by $\mathcal{A}(\mathbf{p})$ the affine subspace of $\mathcal{S}^n$ consisting of matrices that satisfy (1), (2),(3) and (4), where $\mathcal{S}^n$ is the set of $n \times n$ symmetric matrices.

By combining the linear conic formulations from Theorem 2 with a reduction from [31] we have that $\mathcal{D}(\mathbf{p})$ corresponds to the smallest size of a $\mathcal{CS}_+$-factorization over all matrices $R \in \mathcal{CS}_+^n \cap \mathcal{A}(\mathbf{p})$. Using the notion of the cpsd-rank this can be equivalently expressed as follows.

THEOREM 3 ([30, 31]). *For any* $\mathbf{p} = (p(ab|xy)) \in \mathcal{Q}$ *we have that*

$$\mathcal{D}(\mathbf{p}) = \min \big\{ \text{cpsd-rank}(R) : R \in \mathcal{CS}_+^n \cap \mathcal{A}(\mathbf{p}) \big\}.$$

For a fixed Bell scenario, it is a fundamental problem to understand whether one can place a *finite* upper bound on the size of the quantum system necessary to generate all quantum behaviors. In mathematical terms, the question is to decide whether $\max\{\mathcal{D}(\mathbf{p}) : \mathbf{p} \in Q\}$ is finite or infinite, where, again, Q denotes the set of all quantum behaviors corresponding to this fixed Bell scenario.

There is no clear consensus whether finite dimensions always suffice. It follows from the work of Tsirelson [33] that in a $(m_A, m_B, 2, 2)$-scenario finite dimensions are sufficient to generate all *unbiased* behaviors (cf. Section 5.1). Furthermore, in a $(1, 1, o_A, o_B)$-scenario (i.e., exactly one measurement setting per party), the sets of local and quantum behaviors coincide and additionally, it was shown by Jain, Shi, Wei and Zhang [17] that $\mathcal{D}(\mathbf{p})$ is equal to the positive semidefinite rank of the nonnegative matrix $P = (p(ab)_{1 \leq a \leq o_A, 1 \leq b \leq o_B}) \in \mathbb{R}_+^{o_A \times o_B}$ (cf. Section 3.3.2). The latter quantity is upper bounded (e.g. by $\min\{o_A, o_B\}$), so again in this case the maximum of $\mathcal{D}(\mathbf{p})$ over all behaviors is finite. On the other hand, Pál and Vértesi in [27] provide numerical evidence that finite dimensional quantum systems do not suffice in the (3,3,2,2)-scenario, although this still remains to be proven analytically.

Our motivation for introducing and studying the cpsd-rank is that it provides a novel approach to address the finite vs. infinite representability problem of the set of quantum behaviors. Specifically, using Theorem 3 we immediately get two sufficient conditions, in terms of the cpsd-rank, that allow us to either prove or disprove that finite-dimensional systems suffice to generate all quantum behaviors.

PROPOSITION 4. *Fix a* (m_A, m_B, o_A, o_B)-*scenario, set* $n := m_A o_A + m_B o_B$ *and let* $\mathcal{Q}$ *be the corresponding set of quantum behaviors. We have that:*

(i) *If* $\max\{\text{cpsd-rank}(X) : X \in \mathcal{CS}_+^n\} < +\infty$ *then* $\max\{\mathcal{D}(\mathbf{p}) : \mathbf{p} \in \mathcal{Q}\} < +\infty$;

(ii) *Say that for every* $d \geq 1$ *there exists* $p_d \in \mathcal{Q}$ *such that for any* $R \in \mathcal{CS}_+^n \cap \mathcal{A}(\mathbf{p}_d)$ *we have* $\text{cpsd-rank}(R_d) > d$. *Then* $\max\{\mathcal{D}(\mathbf{p}) : \mathbf{p} \in \mathcal{Q}\} = +\infty$.

The value of Proposition 4 is that it identifies a concrete mathematical problem, stripped off all quantum mechanical context, whose resolution would settle the question of finite vs. infinite dimensionality of the set of quantum behaviors:

Question: Is $\max\{\text{cpsd-rank}(X) : X \in \mathcal{CS}_+^n\}$ finite or infinite?

The question concerning the finiteness of the cpsd-rank was already stated in [10]. As already mentioned, if we pose the same question but replace $\mathcal{CS}_+^n$ by the cone of completely positive matrices, the answer is known: The cp-rank can be at most quadratic in the size of the matrix. The proof of this fact relies on the atomic reformulation for the cp-rank. On the other hand, we are not aware of such an atomic reformulation for the cpsd-rank and this limits the analogies with the $\mathcal{CP}^n$ case.

1.3. Contributions and paper organization. In this work we initiate the systematic study of the cpsd-rank and by establishing a connection to quantum behaviors, we make the case that it admits significant physical motivation.

In Section 2 we introduce all necessary notation, definitions and background material on Linear Algebra, Quantum Mechanics, Convexity and Graph Theory. We begin our study of the cpsd-rank in Section 3 where our goal is to give a general introduction and collect basic properties. Specifically, in Section 3.1 we consider matrix operations that preserve the property of being cpsd and examine how they affect the cpsd-rank. In Section 3.2 we identify an analytic and support based lower bound on the cpsd-rank and note that both the bounds never exceed the size of the matrix. Lastly, in Section 3.3 we relate the cpsd-rank to other notions of matrix ranks.

As was already mentioned, no general upper bound is currently known on the cpsd-rank of a cpsd matrix. In view of this, there are two natural research directions: First, identify families of cpsd matrices for which we can place an upper bound on the cpsd-rank and second, identify cpsd matrices with high cpsd-rank. As we describe below, in this work we make progress in both directions.

In Section 4 we consider the question of upper bounding the cpsd-rank for certain families of $\mathcal{CS}_+$ matrices. We focus on *Gram-Lorentz* matrices, denoted by $\mathcal{GL}$, defined as the set of Gram matrices of Lorentz cone vectors (also known as the second-order cone) which we introduce and study in Section 4.1. Furthermore, in Section 4.1 we revisit and give a simplified proof of a construction from [12], where it is shown that the m-dimensional Lorentz cone can be isometrically embedded into a psd cone of size $2^{\Omega(m)}$. This implies that Gram-Lorentz matrices are cpsd. Furthermore, in Section 4.2 we show that for any $X \in \mathcal{GL}$ we have that cpsd-rank$(X) \leq 2^{O(\text{rank}(X))}$.

As it turns out, Gram-Lorentz matrices are also useful to construct matrices that are cpsd but not completely positive. The first such separation $\mathcal{CP}^6 \subsetneq \mathcal{CS}_+^6$ was in fact shown using $\mathcal{GL}$ matrices [12]. In Section 4.3 we generalize the construction from [12] and identify a sufficient condition for constructing matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$.

Lastly, Gram-Lorentz matrices are also relevant in the context of quantum behaviors. In view of Theorem 2 given above, any $\mathcal{K} \subseteq \mathcal{CS}_+$ corresponds to a subset of the set of quantum behaviors. In Section 4.4 we introduce and study *Gram-Lorentz behaviors*, i.e., the quantum behaviors that correspond to $\mathcal{K} = \mathcal{GL}$. Since $\mathcal{GL}$ matrices have bounded cpsd-rank, all $\mathcal{GL}$ behaviors can be generated using a finite-dimensional quantum system. This is again very interesting since, as we mentioned in Section 1.2, it is not known whether there exists a finite upper bound on the size of a quantum system necessary to generate all quantum behaviors corresponding to a Bell scenario.

In Section 5 we turn to the problem of constructing cpsd matrices with high cpsd-rank. Interestingly, Gram-Lorentz matrices turn out to be the right tool to address this problem. Indeed, for our main result in Section 5 (cf. Theorem 33) we construct a family of $\mathcal{GL}$ matrices whose cpsd-rank is exponential in terms of their size.

RESULT 1. *For any integer $n \geq 1$ there exists a matrix $X_n \in \mathcal{GL}^{2n}$ such that*

$$(5) \qquad \text{cpsd-rank}(X_n) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}, \quad \text{where } r_{\max}(n) := \left\lfloor (\sqrt{1+8n}-1)/2 \right\rfloor.$$

In particular, if we take C_n to be an extreme point of the n-dimensional elliptope $\mathcal{E}_n := \{X \in \mathcal{S}_+^n : X_{ii} = 1, \ \forall i \in [n]\}$ satisfying $\text{rank}(C_n) = r_{\max}(n)$, then (5) holds for

$$X_n := \begin{pmatrix} J + C_n & J - C_n \\ J - C_n & J + C_n \end{pmatrix},$$

where J is the $n \times n$ matrix of all 1's.

The starting point for proving Result 1 is Theorem 3. Specifically, it follows by Theorem 3 that given a quantum behavior $\mathbf{p} \in \mathcal{Q}$, for any $R \in \mathcal{CS}_+^n \cap \mathcal{A}(\mathbf{p})$ we have that cpsd-rank$(R) \geq \mathcal{D}(\mathbf{p})$. Consequently, in order to derive Result 1 it suffices to identify a sequence of Gram-Lorentz behaviors $(\mathbf{p}_n)_{n \in \mathbb{N}}$ for which all quantum representations require a quantum system of size exponential in n. We show that for any $n \geq 1$, there exists a Gram-Lorentz behavior $\mathbf{p}_n$ corresponding to the $(n, n, 2, 2)$-scenario with the property that $\mathcal{D}(\mathbf{p}_n) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}$ (cf. Theorem 32). This is the main step in the proof of Result 1.

The first step towards constructing these Gram-Lorentz behaviors is to restrict to Bell scenarios where each party has two possible outcomes, which we label by $\{\pm 1\}$. In this case, instead of working with quantum behaviors we can equivalently work with the corresponding *correlation vectors*. These are just the vectors that correspond to the expected value of the product of the player's individual outcomes. This correspondence is explained in detail in Section 5.1. Quantum correlation vectors turn out to be extremely important for this work due to a lower bound on the size of operator representations of extremal quantum correlations. This result is implicit in [33] and is explained in detail in Section B in the Appendix.

In Section 5.2 we construct a family Gram-Lorentz behaviors $(\mathbf{p}_n)_{n \in \mathbb{N}}$ satisfying $\mathcal{D}(\mathbf{p}_n) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}$ (cf. Theorem 32). To do this, in Section 5.2.1, we translate the aforementioned lower bound in terms of Gram-Lorentz behaviors. Specifically, we show that to any extremal quantum correlation, represented as a matrix C, we can associate a Gram-Lorentz behavior $\mathbf{p}_C$ such that $\mathcal{D}(\mathbf{p}_C) \geq 2^{\Omega(\mathrm{rank}(C))}$. In view of this, it suffices to identify high-rank extremal quantum correlations. In Section 5.2.2 we focus on the case $m_A = m_B =: n$ and show that the extreme points of the n-dimensional elliptope $\mathcal{E}_n$ are also extreme points of the set of quantum correlations. This allows us to conclude the proof as it is well-known that for any $n \geq 1$ there exist extreme points of $\mathcal{E}_n$ whose rank is equal to $r_{\max}(n)$. In Section 5.2.3 we put everything together, and also provide an explicit family of Gram-Lorentz behaviors realizing this exponential lower bound. Lastly, Section 5.3 is dedicated to the proof of Result 1 where we construct cpsd matrices with exponential cpsd-rank.

In Section 6 we study cpsd-graphs, i.e., graphs G with the property that every $\mathcal{DNN}$ matrix whose support is given by G is also in $\mathcal{CS}_+$. The analogous notion of cp-graphs has been extensively studied (e.g. see [3, Section 2.5]). In particular, the class of cp-graphs admits an exact combinatorial characterization: A graph is cp if and only if it does not contain an odd cycle C_{2t+1} $(t \geq 2)$ as a subgraph [19].

We show that the same characterization extends to cpsd-graphs:

RESULT 2. *A graph is cpsd if and only if it has no C_{2t+1}-subgraph $(t \geq 2)$.*

To prove Result 2, in Section 6.1 we generalize a construction from [12] and [21] and identify a sufficient condition for constructing doubly-nonnegative matrices that do not admit a Gram factorization using positive elements in any tracial von Neumann algebra. On the other hand, the closure of $\mathcal{CS}_+^n$ was characterized in [7, Theorem 4.6] as the set of psd matrices that admit a Gram factorization using positive elements in a certain tracial von Neumann algebra. Thus, our sufficient condition can be used to construct matrices in $\mathcal{DNN} \setminus \mathrm{cl}(\mathcal{CS}_+)$. Using these matrices, in Section 6.2 (cf. Theorem 36) we give the proof of Result 2.

2. Preliminaries.

Linear Algebra. We denote by $[d]$ the set $\{1, \ldots, d\}$. The standard orthonormal basis of $\mathbb{C}^d$ is denoted by $\{e_i\}_{i=1}^d$, which we consider as column vectors. The linear span of the vectors $\{x_i\}_{i=1}^n$ is denoted by $\operatorname{span}(\{x_i\}_{i=1}^n)$. We write $x \circ y$ for the entrywise product of two vectors x, y.

We denote the set of $d \times d$ Hermitian (resp. symmetric) matrices by $\mathcal{H}^d$ (resp. $\mathcal{S}^d$). An operator X is called an (orthogonal) *projector* if it satisfies $X = X^* = X^2$, where X^* denotes the conjugate transpose of X. The entrywise product of two matrices X, Y is denoted by $X \circ Y$ and their Kronecker product by $X \otimes Y$. Throughout this work we equip $\mathcal{H}^d$ with the Hilbert-Schmidt inner product $\langle X, Y \rangle := \operatorname{Tr}(XY^*)$. The *direct sum* of two matrices X, Y is the matrix $\left(\begin{smallmatrix} X & 0 \\ 0 & Y \end{smallmatrix} \right)$ which we denote by $X \oplus Y$. The matrix with all entries equal to 1 is denoted by J and the identity matrix by I.

A matrix $X \in \mathcal{H}^d$ is called *positive semidefinite* (psd) if $\psi^* X \psi \geq 0$ for all $\psi \in \mathbb{C}^d$. The set of $d \times d$ Hermitian psd (resp. symmetric psd) matrices forms a closed convex cone denoted by $\mathcal{H}_+^d$ (resp. $\mathcal{S}_+^d$).

Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an inner product space. The *Gram matrix* of a family of vectors $\{x_i\}_{i=1}^n \subseteq \mathcal{V}$, denoted by $\operatorname{Gram}(\{x_i\}_{i=1}^n)$, is the $n \times n$ matrix whose (i, j) entry is given by $\langle x_i, x_j \rangle$, for all $i, j \in [n]$. Lastly, note that for all $\{x_i\}_{i=1}^n \subseteq \mathcal{V}$ we have that $\operatorname{Gram}(\{x_i\}_{i=1}^n)$ is psd and moreover, $\operatorname{rank}(\operatorname{Gram}(\{x_i\}_{i=1}^n)) = \dim(\operatorname{span}(\{x_i\}_{i=1}^n))$, where $\dim(\mathcal{V})$ denotes the dimension of vector space $\mathcal{V}$.

Quantum mechanics. In this section we briefly introduce some notions from quantum mechanics that are of relevance to this work. For a detailed introduction the interested reader is referred to [26].

According to the axioms of quantum mechanics, the *state* of a d-dimensional quantum system is specified by a Hermitian psd operator $\rho \in \mathcal{H}_+^d$ (for some $d \geq 1$) such that $\operatorname{Tr}(\rho) = 1$, where $\operatorname{Tr}(\rho)$ is the trace of ρ. In order to extract information from a quantum system we need to *measure* it. Measurements on a quantum system are described by the Positive Operator Valued Measure (POVM) formalism. A POVM is a family of psd matrices $\{M_i\}_{i=1}^m \subseteq \mathcal{H}_+^d$ that sum to the identity operator, i.e., $\sum_{i=1}^m M_i = I_d$. If the measurement $\{M_i\}_{i=1}^m$ is performed on a quantum system which is in state ρ then the outcome i is observed with probability $p_i := \operatorname{Tr}(\rho M_i)$. Note that from the definitions above $\{p_i\}_{i=1}^m$ is a valid probability distribution.

We also use a second (equivalent) mathematical formalism describing a quantum measurement. Given a d-dimensional quantum system, an *observable* is any Hermitian operator H acting on $\mathbb{C}^d$. By the spectral theorem we know that $H = \sum_{i=1}^k \lambda_i P_i$, where $\{\lambda_i\}_{i=1}^k$ ($k \leq d$) are the eigenvalues of H and $\{P_i\}_{i=1}^k$ are the projectors onto the corresponding eigenspaces. The observable H describes the POVM measurement $\{P_i\}_{i=1}^k$ with outcomes $\{\lambda_i\}_{i=1}^k$, i.e., upon measuring state ρ, the probability of outcome λ_i is given by $\operatorname{Tr}(\rho P_i)$. We say that H is a ± 1 *observable* if it has ± 1 eigenvalues.

Consider two quantum systems S_1 and S_2 and say that S_1 is in state $\rho_1 \in \mathcal{H}_+^{d_1}$ and S_2 is in state $\rho_2 \in \mathcal{H}_+^{d_2}$. In this case, the state of the joint system is given by the density matrix $\rho_1 \otimes \rho_2 \in \mathcal{H}_+^{d_1 d_2}$. If $\{M_i\}_{i=1}^{m_1} \subseteq \mathcal{H}_+^{d_1}$ and $\{N_j\}_{j=1}^{m_2} \subseteq \mathcal{H}_+^{d_2}$ are POVMs on the individual systems S_1 and S_2, the operators $\{M_i \otimes N_j : i \in [m_1], j \in [m_2]\} \subseteq \mathcal{H}_+^{d_1 d_2}$ define a *joint measurement* on the joint system. Note that not all states and measurements are of this form. In particular, states that are not convex combinations of states of the form $\rho_1 \otimes \rho_2$ are said to be *entangled*.

We frequently consider *rank* 1 quantum states which can be written as the outer product $\psi \psi^*$ for some vector ψ (which must have unit norm since its outer product

must have unit trace). Such quantum states are called *pure* and there is one such pure quantum state we use frequently in this paper. We denote by Ψ_d the canonical *maximally entangled state* given by

$$
(6) \qquad \frac{1}{\sqrt{d}} \sum_{i=1}^{d} e_i \otimes e_i \in \mathbb{C}^d \otimes \mathbb{C}^d.
$$

One can check that it is indeed entangled. We make repeated use of the fact that

$$
(7) \qquad \Psi_d^*(A \otimes B)\Psi_d = \frac{1}{d} \operatorname{Tr}\left(AB^\mathsf{T}\right), \text{ for all } A, B \in \mathbb{C}^{d \times d}.
$$

The *Pauli matrices* are given by

$$
I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{ and } Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
$$

Note that the (non-identity) Pauli matrices are Hermitian, their trace is equal to zero, they have ± 1 eigenvalues and they pairwise anticommute. Many of the explicit observables we consider in this paper are constructed using the Pauli matrices.

Convexity. A set $C \subseteq \mathbb{R}^n$ is *convex* if for all $a, b \in C$ and $\lambda \in [0, 1]$ we have that $\lambda a + (1 - \lambda)b \in C$. A subset $F \subseteq C$ is called a *face* of C if $\lambda c_1 + (1 - \lambda)c_2 \in F$ implies that $c_1, c_2 \in F$, for all $c_1, c_2 \in C$ and $\lambda \in [0, 1]$. We say that c is an *extreme point* of the convex set C if the set $\{c\}$ is a face of C. We denote by $\operatorname{ext}(C)$ the set of extreme points of the convex set C.

Graph theory. A graph G is an ordered pair of sets $([n], E(G))$, where $E(G)$ is a collection of 2-element subsets of $[n]$. The elements of $[n]$ are called the *vertices* of the graph and the elements of $E(G)$ its *edges*. For every edge $e = \{u, v\} \in E(G)$ we say that u and v are *adjacent* and write $u \sim v$. A *subgraph* of G is a graph whose vertex and edge sets are subsets of the vertex and edge sets of G, respectively. The *adjacency matrix* of G is the $n \times n$ matrix

$$
A := \sum_{u \sim v}(e_u e_v^\mathsf{T} + e_v e_u^\mathsf{T}).
$$

Note that the smallest eigenvalue of A is negative. The *support graph* of a matrix $X \in \mathcal{S}_n$, denoted by $S(X)$, is the graph with vertex set $[n]$, and $u \sim v$ if and only if $X_{uv} \neq 0$ (and $u \neq v$). The *n-cycle*, denoted C_n, is the graph with vertex set $[n]$ where $u \sim v$ if $(u - v) \equiv 1 \mod n$.

3. Studying the cpsd-rank.

3.1. Basic properties. Our goal in this section is to determine basic properties of the cpsd-rank that we use throughout this work.

Note that in the definition of cpsd-rank we only consider $\mathcal{CS}_+$-factorizations using Hermitian (i.e., complex valued) psd matrices. If we restrict to $\mathcal{CS}_+$-factorizations using symmetric (i.e., real valued) psd matrices, we arrive at the notion of *real cpsd-rank*. Nevertheless, the real cpsd-rank can differ at most by a factor of two from the cpsd-rank. To see this, for any $X \in \mathbb{C}^{d \times d}$ set

$$
(8) \qquad T(X) := \frac{1}{\sqrt{2}} \begin{pmatrix} \mathcal{R}(X) & -\mathcal{I}(X) \\ \mathcal{I}(X) & \mathcal{R}(X) \end{pmatrix},
$$

9

and note that T is a bijection between $\mathbb{C}^{d \times d}$ and $\mathbb{R}^{2d \times 2d}$. Furthermore, $X \in \mathcal{H}_+^n$ if and only if $T(X) \in \mathcal{S}_+^{2n}$ and moreover $\langle X, Y \rangle = \langle T(X), T(Y) \rangle$ for all $X, Y \in \mathcal{H}_+^n$.

In our first result in this section we collect several simple properties concerning the psd matrices in a $\mathcal{CS}_+$-factorization.

LEMMA 5. *Let $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ be a $\mathcal{CS}_+$-factorization for $X \in \mathcal{CS}_+^n$.*

(i) *For any $d \times d$ unitary matrix U, the matrices $\{U^* P_i U\}_{i=1}^n \subseteq \mathcal{H}_+^d$ are a $\mathcal{CS}_+$-factorization of X.*

(ii) *We have that $\mathrm{cpsd\text{-}rank}(X) \leq \mathrm{rank}(\sum_{i=1}^n P_i)$. In particular, if $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ is a size-optimal $\mathcal{CS}_+$-factorization then $\mathrm{rank}(\sum_{i=1}^n P_i) = d$.*

Proof. Part (i) is clear. For (ii) define the psd matrix $P := \sum_{i=1}^n P_i$ and set $r := \mathrm{rank}(P)$. Clearly, P is unitarily equivalent to a diagonal matrix with exactly r positive entries. By restricting to the support of P, we get a $\mathcal{CS}_+$-factorization of X using $r \times r$ psd matrices. $\qquad\square$

Recall that any family of pairwise-commuting Hermitian matrices is simultaneously diagonalizable by a unitary matrix (e.g. see [16, Theorem 2.5.5]). Consider $X \in \mathcal{CS}_+^n$ and let $\mathcal{I} \subseteq [n]$ so that the principal submatrix corresponding to $\mathcal{I}$ is diagonal (with positive diagonal entries). In view of Lemma 5 (i), we may assume that in any $\mathcal{CS}_+$-factorization $\{P_i\}_{i=1}^n$ of X, the matrices $\{P_i\}_{i \in \mathcal{I}}$ can be taken to be diagonal psd. This immediately implies $\mathrm{cpsd\text{-}rank}(I_n) \geq n$, for any n, which can be easily seen to hold with equality. We proceed with a second example.

EXAMPLE 3.1. *We prove that the cpsd-rank of the matrix* $\begin{pmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 & 3 \end{pmatrix}$ *is*

equal to 4 and a size-optimal $\mathcal{CS}_+$-factorization is given by

$$\begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Assume that $\mathrm{cpsd\text{-}rank}(X) \leq 3$ and let $\{P_i\}_{i=1}^5 \subseteq \mathcal{H}_+^3$ be a $\mathcal{CS}_+$-factorization. Since P_1, P_2 and P_3 commute pairwise, by applying an appropriate change of basis, we may assume by Lemma 5 (i) that they are diagonal. Furthermore, as $\langle P_i, P_j \rangle = 0$, for $i \neq j \in [3]$ it follows that $P_i = \sqrt{2} e_i e_i^\mathsf{T}$, for $1 \leq i \leq 3$. Since P_4, P_5 are 3×3 orthogonal psd matrices (and nonzero) one of them has rank 1. Suppose without loss of generality $P_4 = xx^*$ for some $x = (x_i) \in \mathbb{C}^3$. Note that $|x_i|^2 = 1/\sqrt{2}$, for all $1 \leq i \leq 3$. Thus, $\langle P_4, P_4 \rangle = (\sum_{i=1}^3 |x_i|^2)^2 = 9/2 \neq 3$, a contradiction.

In the remaining part of this section we focus on matrix operations that preserve the property of being cpsd and we investigate in what way they affect the cpsd-rank.

LEMMA 6. *Consider $X \in \mathcal{CS}_+^n$. We have that:*

(i) *For any $n \times n$ diagonal matrix D with strictly positive diagonal entries we have*

$$DXD \in \mathcal{CS}_+^n, \text{ and } \mathrm{cpsd\text{-}rank}(X) = \mathrm{cpsd\text{-}rank}(DXD).$$

(ii) *For any $n \times n$ permutation matrix P we have*

$$PXP^\mathsf{T} \in \mathcal{CS}_+^n, \text{ and } \mathrm{cpsd\text{-}rank}(X) = \mathrm{cpsd\text{-}rank}(PXP^\mathsf{T}).$$

We now determine how the cpsd-rank behaves under matrix sums.

LEMMA 7. *For any $X, Y \in \mathcal{CS}_+^n$ we have that $X + Y \in \mathcal{CS}_+^n$ and furthermore,*

$$\text{cpsd-rank}(X + Y) \leq \text{cpsd-rank}(X) + \text{cpsd-rank}(Y).$$

Proof. Let $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^{d_1}$ and $\{Q_j\}_{j=1}^m \subseteq \mathcal{H}_+^{d_2}$ be size-optimal $\mathcal{CS}_+$-factorizations for X and Y, respectively. For all $i \in [n]$ define $Z_i := P_i \oplus Q_i \in \mathcal{H}_+^{d_1+d_2}$ and note that the matrices $\{Z_i\}_{i=1}^n$ are a $\mathcal{CS}_+$-factorization for $X + Y$. $\square$

REMARK 3.1. *As it turns out, the cpsd-rank of the sum of a family of cpsd matrices can be exponentially smaller compared to any of the individual cpsd-ranks. To see this, let $X \in \mathcal{CS}_+^n$ and define $X_{sym} := \sum_{P \in P_n} PXP^\mathsf{T}$, where P_n is the set of $n \times n$ permutation matrices. By Lemma 6 we have $X_{sym} \in \mathcal{CS}_+^n$ and by its definition we have $X_{sym} = (a - b)I + bJ$, for appropriate constants a, b where $a \geq b \geq 0$. By Lemma 7 we have $\text{cpsd-rank}(X_{sym}) \leq n + 1$, since $\text{cpsd-rank}(I_n) = n$ and $\text{cpsd-rank}(J) = 1$. On the other hand, in Section 5 we show that for any $n \geq 1$, there exists a matrix in $\mathcal{CS}_+^{2n}$ with cpsd-rank $2^{\Omega(\sqrt{n})}$.*

In our next result we determine how the cpsd-rank behaves under direct sums.

LEMMA 8. *For any $X \in \mathcal{CS}_+^n$ and $Y \in \mathcal{CS}_+^m$ we have that $X \oplus Y \in \mathcal{CS}_+^{n+m}$ and furthermore, $\text{cpsd-rank}(X \oplus Y) = \text{cpsd-rank}(X) + \text{cpsd-rank}(Y)$.*

Proof. Let $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^{d_1}$ and $\{Q_j\}_{j=1}^m \subseteq \mathcal{H}_+^{d_2}$ be size-optimal $\mathcal{CS}_+$-factorizations for X and Y, respectively. For $i \in [n]$, set $\tilde{P}_i := P_i \oplus 0_{d_2} \in \mathcal{H}_+^{d_1+d_2}$ and for $j \in [m]$ set $\tilde{Q}_j := 0_{d_1} \oplus Q_j \in \mathcal{H}_+^{d_1+d_2}$. Clearly the matrices $\{\tilde{P}_i\}_{i=1}^n \cup \{\tilde{Q}_j\}_{j=1}^m$ form a $\mathcal{CS}_+$-factorization for $X \oplus Y$. Thus we get that $X \oplus Y \in \mathcal{CS}_+^{n+m}$ and furthermore, $\text{cpsd-rank}(X \oplus Y) \leq \text{cpsd-rank}(X) + \text{cpsd-rank}(Y)$.

For the reverse inequality set $d := \text{cpsd-rank}(X \oplus Y)$ and let $\{P_i\}_{i=1}^n \cup \{Q_j\}_{j=1}^m \subseteq \mathcal{H}_+^d$ be a size-optimal $\mathcal{CS}_+$-factorization for $X \oplus Y$. Moreover, set $P := \sum_{i=1}^n P_i$ and $Q := \sum_{j=1}^m Q_j$. By Lemma 5 we have $\text{rank}(P+Q) = d$. Furthermore, by the structure of $X \oplus Y$ we have $\langle P_i, Q_j \rangle = 0$, for all i, j and thus $\langle P, Q \rangle = 0$. As P, Q are psd this implies that $d = \text{rank}(P+Q) = \text{rank}(P) + \text{rank}(Q)$. Since the matrices $\{P_i\}_{i=1}^n$ form a $\mathcal{CS}_+$-factorization of X, by Lemma 5 (ii) we have $\text{rank}(P) \geq \text{cpsd-rank}(X)$ and similarly that $\text{rank}(Q) \geq \text{cpsd-rank}(Y)$. Putting everything together, the claim follows. $\square$

Our next goal is to show that there exist $\mathcal{CS}_+$ matrices that do not admit $\mathcal{CS}_+$-factorizations using only rank-one factors. In contrast to this, restricting to factorizations using rank-one psd matrices has been a useful approach to provide upper bounds on the positive semidefinite rank (cf. Section 3.3.2) [22, 10].

We denote by $\mathcal{CS}_+^{n,1}$ the set of matrices in $\mathcal{CS}_+^n$ that admit $\mathcal{CS}_+$-factorizations using rank-one factors. Furthermore, we call a *Hadamard square root* of $X \in \mathbb{R}_+^{n \times m}$ any matrix obtained by replacing each entry of X by one of its two square roots. We have the following result whose proof is straightforward and is omitted.

LEMMA 9. *For any matrix $X \in \mathcal{H}_+^n$ we have that $X \circ X^* \in \mathcal{CS}_+^{n,1}$ and moreover $\text{cpsd-rank}(X \circ X) \leq \text{rank}(X)$. In particular, if $X \in \mathcal{H}_+^n$ is a matrix with 0/1 entries then $X \in \mathcal{CS}_+^n$ and $\text{cpsd-rank}(X) \leq \text{rank}(X)$. Conversely, if $X \in \mathcal{CS}_+^{n,1}$ then X has a psd Hadamard square root.*

As a concrete example of a matrix in $\mathcal{CS}_+ \setminus \mathcal{CS}_+^1$, consider

$$(9) \qquad X = \begin{pmatrix} 1 & \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 & 1/10 \\ \sqrt{2}/2 & 1/10 & 1 \end{pmatrix}.$$

Clearly $X \in \mathcal{CS}_+^3 = \mathcal{DNN}^3$, but no Hadamard square root of X is psd.

3.2. Lower bounds. In this section we derive two general lower bounds on the cpsd-rank. The first one is analytic and the second one is based on the support of the matrix. We show that in both cases, our bounds never exceed the size of the matrix.

3.2.1. Analytic lower bound. We start with the following result.

THEOREM 10. *For any matrix $X \in \mathcal{CS}_+^n$ we have that*

$$\text{(10)} \qquad \text{cpsd-rank}(X) \geq \frac{\left(\sum_{i=1}^n \sqrt{X_{ii}}\right)^2}{\sum_{i,j=1}^n X_{ij}}.$$

Proof. Set $d := \text{cpsd-rank}(X)$ and let $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ be a size-optimal $\mathcal{CS}_+$-factorization. By Lemma 5 (ii) we have that $P := \sum_{i=1}^n P_i \in \mathcal{H}_+^d$ has full-rank. By the Cauchy-Schwartz inequality we have that $d \geq \text{Tr}(P)^2/\text{Tr}(P^2)$. Note that $\text{Tr}(P^2) = \sum_{i,j=1}^n X_{ij}$. Lastly,

$$\text{Tr}(P)^2 = \left(\sum_{i=1}^n \text{Tr}(P_i)\right)^2 \geq \left(\sum_{i=1}^n \sqrt{\text{Tr}(P_i^2)}\right)^2 = \left(\sum_{i=1}^n \sqrt{X_{ii}}\right)^2,$$

where we used $\text{Tr}(P_i) \geq \sqrt{\text{Tr}(P_i^2)}$, since $P_i \in \mathcal{H}_+^d$, for the last inequality. □

In view of Theorem 10, two remarks are in order. First, it follows by (10) that cpsd-rank$(I_n) \geq n$ and this is obviously tight. Second, the Cauchy-Schwartz inequality combined with the fact that any cpsd matrix is entrywise nonnegative implies that the lower bound (10) can never exceed the size of the matrix.

3.2.2. Support-based lower bound. To study support-based lower bounds on the cpsd-rank we introduce the following graph parameter:

$$\text{(11)} \qquad f(G) := \min\{d \geq 1 : \exists \text{ subspaces } \{\mathcal{L}_i\}_{i=1}^n \subseteq \mathbb{C}^d \text{ s.t. } \mathcal{L}_i \perp \mathcal{L}_j \iff i \nsim j\}.$$

To see $f(G)$ is well-defined let A be the adjacency matrix of G and let τ be its least eigenvalue with multiplicity m. Since $A - \tau I \in \mathcal{S}_+^n$ and $\text{rank}(A - \tau I) = n - m$, there exist vectors $\{x_i\}_{i=1}^n \subseteq \mathbb{R}^{n-m}$ such that $A - \tau I = \text{Gram}(\{x_i\}_{i=1}^n)$. For $i \in [n]$, set $\mathcal{L}_i := \text{span}(\{x_i\})$ and note this is a feasible solution for (11) yielding $f(G) \leq n-m$.

THEOREM 11. *For any graph $G = ([n], E)$ we have that $f(G)$ is equal to*

$$\text{(12)} \qquad \min\{\text{cpsd-rank}(X) : X \in \mathcal{CS}_+^n \text{ and } S(X) = G\}.$$

Proof. By Lemma 9, the 0/1 matrix $A - \tau I$ defined in the previous paragraph is cpsd. This shows that (12) is feasible. Let X be optimal for (12) and let $\{P_i\}_{i=1}^n \subseteq \mathcal{H}_+^d$ be a size-optimal $\mathcal{CS}_+$-factorization for X. For $i \in [n]$, define $\mathcal{L}_i := \text{Range}(P_i) \subseteq \mathbb{C}^d$ and note this is feasible for (11). Conversely, let $\{\mathcal{L}_i\}_{i=1}^n \subseteq \mathbb{C}^d$ be a family of subspaces feasible for (11) and for $i \in [n]$ define P_i to be the orthogonal projector onto $\mathcal{L}_i$. Lastly, note that the matrix $X := \text{Gram}(\{P_i\}_{i=1}^n) \in \mathcal{CS}_+^n$ is feasible for (12) and satisfies cpsd-rank$(X) \leq d$. □

By Theorem 11 and the fact that $f(G)$ is upper bounded by n it follows that support-based lower bounds on the cpsd-rank never exceed the size of the matrix.

3.3. Comparisons with other notions of rank. In this section we investigate further the relationships between the cpsd-rank and other notions of matrix ranks.

3.3.1. cpsd-rank vs. rank. As $\mathcal{H}^d$ is isometrically isomorphic to $\mathbb{R}^{d^2}$, we have

$$\sqrt{\mathrm{rank}(X)} \leq \mathrm{cpsd\text{-}rank}(X), \tag{13}$$

for any $X \in \mathcal{CS}_+$. We provide an example that illustrates that the above can be tight up to a constant factor. Let $r \geq 2$ be an integer and let $E_{i,j} := I_r + e_i e_j^\mathsf{T} + e_j e_i^\mathsf{T} \in \mathcal{H}_+^r$ for all $i, j \in [r]$. The matrix $X := \mathrm{Gram}(\{E_{i,j}\}_{i,j}) \in \mathcal{CS}_+^{r(r-1)/2}$ has cpsd-rank$(X) \leq r$, by construction, while X can be easily seen to have full rank. On the other hand, no upper bound for cpsd-rank(X) in terms of rank(X) is known.

3.3.2. cpsd-rank vs. psd-rank. Given any entrywise nonnegative matrix $X \in \mathbb{R}_+^{n \times m}$, its *positive semidefinite rank* (psd-rank), denoted by $\mathrm{rank}_{psd}(X)$, is defined as the least integer $d \geq 1$ for which there exist $\{A_i\}_{i=1}^n, \{B_j\}_{j=1}^m \subseteq \mathcal{H}_+^d$ such that $X_{ij} = \mathrm{Tr}(A_i B_j)$ for all $i \in [n], j \in [m]$. Generalizing a theorem by Yannakakis [35], it was shown in [11] and [14] that $\mathrm{rank}_{psd}(S)$ where S is a slack matrix for polytope P corresponds to the smallest size of a spectrahedron that projects onto P. For further properties of the rank_{psd} the reader is referred to [10] and Section 3.3.2.

Clearly, for any $X \in \mathcal{CS}_+^n$ we have that $\mathrm{rank}_{psd}(X) \leq \mathrm{cpsd\text{-}rank}(X)$. Furthermore, since $\mathrm{rank}_{psd}(X) \leq n$, for any $X \in \mathcal{CS}_+^n$, the example of the matrix X with cpsd-rank$(X) = 2^{\Omega(\sqrt{n})}$ given in Section 5 provides an exponential separation between $\mathrm{rank}_{psd}(X)$ and cpsd-rank(X).

We conclude this section by determining the exact relation between $\mathrm{rank}_{psd}(X)$ and cpsd-rank(X). This follows from the connection to Bell scenarios. As both quantities are invariant under scaling by a positive constant, without loss of generality we can assume that $\sum_{i,j} X_{i,j} = 1$ so that $\mathbf{p} := (X_{ij})_{ij}$ is a probability distribution. We can think of $\mathbf{p}$ as a behavior corresponding to a $(1, 1, m, n)$ Bell scenario where each party has a unique POVM. As mentioned in the introduction, in this case the behavior $\mathbf{p}$ is quantum and moreover, $\mathcal{D}(\mathbf{p}) = \mathrm{rank}_{psd}(X)$ [17]. This fact combined with Theorem 3 implies that $\mathrm{rank}_{psd}(X)$ is equal to

$$\min\left\{ \mathrm{cpsd\text{-}rank}(R) : R = \begin{pmatrix} A & X \\ X^\mathsf{T} & B \end{pmatrix} \in \mathcal{CS}_+^{n+m} \ \text{ and } \ \sum_{i,j=1}^n A_{ij} = \sum_{i,j=1}^n B_{ij} = 1 \right\}.$$

In turn, this is equal to the smallest integer $d \geq 1$ for which there exist Hermitian psd matrices $\{A_i\}_{i=1}^n, \{B_j\}_{j=1}^m \subseteq \mathcal{H}_+^d$ such that $X_{ij} = \mathrm{Tr}(A_i B_j)$, for all $i \in [n], j \in [m]$ and $\sum_{i=1}^n A_i = \sum_{j=1}^m B_j$. As a corollary we also get that in any psd-factorization of X we may assume that the psd factors satisfy $\sum_{i=1}^n A_i = \sum_{j=1}^m B_j$.

3.3.3. cpsd-rank vs. cp-rank. For all matrices $X \in \mathcal{CP}$ we clearly have that

$$\Omega(\mathrm{cp\text{-}rank}(X)^{1/4}) \leq \mathrm{cpsd\text{-}rank}(X) \leq \mathrm{cp\text{-}rank}(X).$$

The lower bound follows from the fact that cp-rank$(X) \leq \binom{\mathrm{rank}(X)+1}{2} - 1$, for all $X \in \mathcal{CP}$ (e.g. see [3, Theorem 3.5]) combined with (13).

We now give an example where cpsd-rank$(X) = $ cp-rank(X). For this, let $a \in (0, 3/4)$ and set $X_a := I_3 + a e_1 e_3^\mathsf{T} + a e_3 e_1^\mathsf{T}$. Recall that $\mathcal{CS}_+^3 = \mathcal{CP}^3 = \mathcal{DNN}^3$. By [3, Theorem 3.2] we have that cp-rank$(X_a) = $ rank$(X_a) = 3$. From Theorem 10 it follows that cpsd-rank$(X_a) \geq 3$, thus cpsd-rank$(X) = $ cp-rank(X) for this case.

Lastly, the example given in Section 3.3.1 also provides a quadratic separation between the cp-rank and the cpsd-rank. The matrix $X \in \mathcal{CP}$ as it is the Gram matrix of $E_{i,j}$ which has nonnegative entries. Further, cp-rank$(X) \geq$ rank$(X) = \binom{r}{2}$ while cpsd-rank$(X) \leq r$, by construction.

13

4. Gram-Lorentz matrices. As already mentioned, it is currently not known whether the cpsd-rank of all matrices in $\mathcal{CS}_+^n$ admits a finite upper bound. In this section we identify a family of $\mathcal{CS}_+$ matrices for which it is possible to prove a finite upper bound. These are the *Gram-Lorentz matrices*, i.e., Gram matrices of Lorentz cone vectors. In Section 4.1 we recall a construction from [12] where it is shown that the Lorentz cone can be isometrically embedded into a psd cone of an appropriate size. This implies that Gram-Lorentz matrices are cpsd. In Section 4.2 we show that the cpsd-rank of a Gram-Lorentz matrix is upper bounded in terms of its rank. Lastly, in Section 4.3 we use Gram-Lorentz matrices to construct matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$, generalizing a construction from [12].

4.1. Embedding the Lorentz cone isometrically into $\mathcal{H}_+^d$. Underlying the results in this section is a linear embedding of vectors in $\mathbb{R}^n$ into traceless Hermitian operators of size $2^{\lfloor \frac{n}{2} \rfloor}$, so that inner products are preserved up to a constant factor and unit vectors get mapped to ± 1 observables.

In fact, this embedding corresponds to a complex representation of the Clifford algebra over $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ defined in terms of the so-called Brauer-Weyl matrices. For more details see [13] and Section A in the Appendix. This embedding is also the main ingredient in Tsirelson's characterization of binary outcome correlations [33].

THEOREM 12. *There exists a linear map $\gamma : \mathbb{R}^n \to \mathcal{H}^d$, where $d = 2^{\lfloor \frac{n}{2} \rfloor}$ such that:*
  (i) For all $x \in \mathbb{R}^n$ we have that $\mathrm{Tr}\,(\gamma(x)) = 0$;
  (ii) For all $x \in \mathbb{R}^n$ with $\|x\| = 1$ we have $\gamma(x)^2 = I_d$;
  (iii) For all $x, y \in \mathbb{R}^n$ we have $d \cdot \langle x, y \rangle = \mathrm{Tr}\,(\gamma(x)\gamma(y))$.

Specifically, when $n = 2\ell$ we define:

$$(14) \qquad \gamma(e_i) = Z^{\otimes(i-1)} \otimes X \otimes I_2^{\otimes(\ell-i)} \in \mathcal{H}^d, \ (i \in [\ell]),$$

and

$$(15) \qquad \gamma(e_{i+\ell}) = Z^{\otimes(i-1)} \otimes Y \otimes I_2^{\otimes(\ell-i)} \in \mathcal{H}^d, \ (i \in [\ell]).$$

When $n = 2\ell + 1$ we define $\{\gamma(e_i)\}_{i=1}^{2\ell}$ as in (14) and (15) and set $\gamma(e_{2\ell+1}) = Z^{\otimes\ell}$. Lastly, we extend γ linearly, i.e., $\gamma(x) = \sum_{i=1}^n x_i \gamma(e_i)$, for any $x = \sum_{i=1}^n x_i e_i \in \mathbb{R}^n$.

By definition of γ it follows that $\mathrm{Tr}(\gamma(x)) = 0$ for all $x \in \mathbb{R}^n$. Furthermore, note that $\gamma(e_i)\gamma(e_j) + \gamma(e_j)\gamma(e_i) = 2\delta_{ij}I_d$, for all i, j, which implies that

$$(16) \qquad \gamma(x)\gamma(y) + \gamma(y)\gamma(x) = 2\langle x, y \rangle I_d, \text{ for all } x, y \in \mathbb{R}^n.$$

For any $x \in \mathbb{R}^n$ with $\|x\| = 1$, setting $x = y$ in (16) we get that $\gamma(x)^2 = I_d$. Lastly, taking traces in (16) we see that $d \cdot \langle x, y \rangle = \mathrm{Tr}\,(\gamma(x)\gamma(y))$ for all $x, y \in \mathbb{R}^n$.

Next we introduce a convex cone which plays a central role in this work.

DEFINITION 13. *The m-dimensional* Lorentz cone, *denoted $\mathcal{L}_m$, is defined as the set of vectors in $\mathbb{R}^m$ whose angle with the vector $e_1 \in \mathbb{R}^m$ does not exceed $\pi/4$, i.e.,*

$$\mathcal{L}_m = \left\{ (c, x) \in \mathbb{R} \times \mathbb{R}^{m-1} : c \geq \|x\| \right\}.$$

It was shown in [12] that the Lorentz cone can be isometrically embedded into the cone of psd matrices of an appropriate dimension. For the convenience of the reader, we include a short new proof of the existence of the isometric embedding.

14

THEOREM 14 ([12]). *Set* $d := 2^{\lfloor \frac{m-1}{2} \rfloor}$. *There exists an isometry* $\Gamma : \mathbb{R}^m \to \mathcal{H}^d$, *such that*

$$\mathcal{L}_m = \{(c, x) \in \mathbb{R} \times \mathbb{R}^{m-1} : \Gamma((c, x)) \in \mathcal{H}^d_+\}.$$

Proof. For $(c, x) \in \mathbb{R}^m = \mathbb{R} \times \mathbb{R}^{m-1}$ define

(17)
$$\Gamma((c, x)) = \frac{1}{\sqrt{d}} \left(cI_d + \gamma(x) \right).$$

To see that Γ defines an isometry note that

$$\mathrm{Tr}(\gamma((c, x))\gamma((c', x'))) = cc' + \langle x, x' \rangle = \langle (c, x), (c', x') \rangle.$$

Lastly, recall that $\gamma(x)^2 = \|x\|^2 I_d$ and thus the eigenvalues of $\gamma(x)$ are given by $\pm \|x\|$. Consequently, $\Gamma((c, x)) \in \mathcal{H}^d_+$ if and only if $c \geq \|x\|$. □

For concreteness, below we explicitly describe the isometry $\Gamma : \mathbb{R}^3 \to \mathcal{H}^2$.

EXAMPLE 4.1. *Let* $(c, v, w) \in \mathcal{L}_3$ *(so* $v, w \in \mathbb{R}$). *By Theorem* 12 *we have that*

$$\gamma(v, w) = vX + wY = \begin{pmatrix} 0 & v - iw \\ v + iw & 0 \end{pmatrix}.$$

Thus, substituting this into (17) *we see that*

(18)
$$\Gamma(c, v, w) = \frac{1}{\sqrt{2}} \begin{pmatrix} c & v - iw \\ v + iw & c \end{pmatrix}.$$

Note that $\Gamma((c, v, w)) \in \mathcal{H}^2_+$ *since* $c \geq 0$ *and the determinant* $c^2 - (|v|^2 + |w|^2) \geq 0$ *since* $(c, v, w) \in \mathcal{L}_3$. *Lastly, notice that if* (c, v, w) *lies on the boundary of* $\mathcal{L}_3$, *then the determinant is* 0 *and thus* $\Gamma((c, v, w))$ *has rank 1.*

4.2. Gram-Lorentz matrices. Theorem 14 suggests the following definition.

DEFINITION 15. *A matrix* $X \in \mathcal{S}^n$ *is called* Gram-Lorentz *if there exist vectors* $\{\ell_i\}_{i=1}^n \subseteq \mathcal{L}_m$ *(for some* $m \geq 1$*) such that* $X = \mathrm{Gram}(\{\ell_i\}_{i=1}^n)$. *We denote the set of* $n \times n$ *Gram-Lorentz matrices by* $\mathcal{GL}^n$.

The study of the set of Gram-Lorentz matrices is motivated as follows. Firstly, in view of Theorem 14, we have that $\mathcal{GL}^n \subseteq \mathcal{CS}^n_+$. Identifying matrices in $\mathcal{GL} \setminus \mathcal{CP}$ therefore provides a systematic approach for finding matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$. All known examples of matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$ are constructed exactly in this manner [12]. Secondly, as we show in this section, the cpsd-rank of a Gram-Lorentz matrix can be upper bounded in terms of its rank. Since it is currently unknown whether the cpsd-rank of all matrices in $\mathcal{CS}^n_+$ admits a finite upper bound, it is instructive to identify families of matrices in $\mathcal{CS}^n_+$ for which there is one.

It is not clear from its definition whether $\mathcal{GL}$ is convex. In fact, the following holds.

LEMMA 16. *The set* $\mathcal{GL}^n$ *is convex if and only if* $n \leq 2$.

Proof. First we show that $\mathcal{GL}^2 = \mathcal{DNN}^2$. For this, let $A := \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathcal{DNN}^2$. We now show how to write it as the Gram matrix of vectors in $\mathcal{L}_3$. Note that $ac \geq b^2$ and wlog assume that $a \geq c > 0$. Set $v_1 := \sqrt{\frac{a}{2}}(1, 1, 0)$ and $v_2 := \sqrt{\frac{c}{2}}(1, d, \sqrt{1 - d^2})$, where $d := (2b - \sqrt{ac})/\sqrt{ac}$. Lastly, note that $A = \mathrm{Gram}(\{v_1, v_2\})$ and that $\{v_i\}_{i=1}^2 \subseteq \mathcal{L}_3$. Since $\mathcal{GL}^2 \subseteq \mathcal{DNN}^2$, the two sets are equal, and thus $\mathcal{GL}^2$ is convex.

Next we show that $\mathcal{GL}^n$ is not convex for $n \geq 3$. Since $\{e_i e_i^\mathsf{T}\}_{i=1}^n \subseteq \mathcal{GL}^n$ we have that $2I_n$ is in the convex hull of $\mathcal{GL}^n$. It is sufficient to show that $2I_n \notin \mathcal{GL}^n$

for $n \geq 3$. To this end, suppose that $2I_n$ is the Gram matrix of the Lorentz vectors $\{(t_i, u_i)\}_{i=1}^n$. This implies that $t_i^2 + \|u_i\|_2^2 = 2$, for all $i \in [n]$ and $t_i t_j + \langle u_i, u_j \rangle = 0$ for all $i \neq j \in [n]$. Since $t_i \geq \|u_i\|$ for all $i \in [n]$, the Cauchy-Schwarz inequality implies that $t_i t_j \geq \|u_i\|_2 \|u_j\|_2 \geq |\langle u_i, u_j \rangle| = t_i t_j$, for all $i \neq j \in [n]$. Thus, equality holds throughout which shows that $t_i = \|u_i\|_2 = 1$, for all $i \in [n]$, and that $u_i = -u_j$ for all $i \neq j \in [n]$. This gives a contradiction since $n \geq 3$. □

We now show that for any $\mathcal{GL}$ matrix we can place an upper bound on the dimension of the Lorentz cone we need to generate it.

LEMMA 17. *Any $X \in \mathcal{GL}^n$ has a $\mathcal{GL}$-factorization using vectors in $\mathcal{L}_{\text{rank}(X)+2}$.*

Proof. Since $X \in \mathcal{GL}^n$ there exist vectors $\{\ell_i\}_{i=1}^n \subseteq \mathcal{L}_m$ (for some $m \geq 1$) such that $X = \text{Gram}(\{\ell_i\}_{i=1}^n)$. For $i \in [n]$ set $\ell_i := (t_i, u_i)$, where $u_i \in \mathbb{R}^{m-1}$ and $\|u_i\| \leq t_i$. Define $U := \text{Gram}(\{u_i\}_{i=1}^n)$ and $t := \sum_{i=1}^n t_i e_i$ and note that $X = U + tt^{\mathsf{T}}$. Since U is psd of rank at most $r := \text{rank}(X) + 1$, there exists a family of vectors $\{\tilde{u}_i\}_{i=1}^n \subseteq \mathbb{R}^r$ such that $U = \text{Gram}(\{\tilde{u}_i\}_{i=1}^n)$. Lastly, since $\|\tilde{u}_i\| = \|u_i\|$ for all $i \in [n]$ it follows that the vectors $\tilde{\ell}_i := (t_i, \tilde{u}_i)$ lie in $\mathcal{L}_{r+1}$ and satisfy $X = \text{Gram}(\{\tilde{\ell}_i\}_{i=1}^n)$. □

THEOREM 18. *For any matrix $X \in \mathcal{GL}^n$ we have that $X \in \mathcal{CS}_+^n$ and*

(19)
$$\text{cpsd-rank}(X) \leq 2^{\lfloor (\text{rank}(X)+1)/2 \rfloor}.$$

Proof. The proof follows by combining Theorem 14 with Lemma 17. □

In Section 5 we show that this bound is essentially tight (cf. Remark 5.4).

4.3. Matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$. In this section we use Gram-Lorentz matrices to present a new family of matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$. To this end, we make use of the following technical lemma.

LEMMA 19. *Consider vectors $\{p_i\}_{i=1}^n \subseteq \mathbb{R}^d$ and scalars $\{\lambda_i^j : i \in [n], j \in [m]\}$ such that $c := \sum_{i=1}^n \lambda_i^j p_i$ for all $j \in [m]$. Consider vectors $\{q_i\}_{i=1}^n \subseteq \mathbb{R}^{d'}$ satisfying $\langle p_i, p_j \rangle = \langle q_i, q_j \rangle$ for all $i, j \in [n]$. Then there exists $c' \in \mathbb{R}^{d'}$ such that $\|c\| = \|c'\|$ and moreover $c' = \sum_{i=1}^n \lambda_i^j q_i$ for all $j \in [m]$.*

Proof. For $j \in [m]$ set $c_j' := \sum_{i=1}^n \lambda_i^j q_i$. For $j \neq j' \in [m]$ we have $\|c_j' - c_{j'}'\|^2 = 0$ and thus $c_j' = c_{j'}'$. Lastly, set c' to be this common value and note that $\|c'\|^2 = \|c\|^2$. □

We now give a sufficient condition for constructing matrices in $\mathcal{DNN} \setminus \mathcal{CP}$, generalizing the construction in [12].

THEOREM 20. *Consider vectors $\mathcal{F} := \cup_{i \in I} \{p_i, p_i'\}$ with the following properties:*
  (i) There exists a nonzero vector c such that $(p_i + p_i')/2 = c$, for all $i \in I$;
  (ii) For all $i \in I$ we have $\langle p_i, p_i' \rangle = 0$;
  (iii) There exists $J \subseteq I$ that has odd cardinality and $\sum_{j \in J} p_j = c \cdot |J|$;
  (iv) The pairwise inner products of all vectors in $\mathcal{F}$ are nonnegative.
Then we have that $\text{Gram}(\mathcal{F}) \in \mathcal{DNN} \setminus \mathcal{CP}$.

Proof. By (iv) we have $\text{Gram}(\mathcal{F}) \in \mathcal{DNN}$. For a contradiction, assume that $\text{Gram}(\mathcal{F}) \in \mathcal{CP}$ and let $\{a_f\}_{f \in \mathcal{F}} \subseteq \mathbb{R}_+^d$ be a nonnegative Gram factorization.

By Lemma 19 there exists a vector $a \in \mathbb{R}^d$ with $\|a\| = \|c\|$ satisfying $(a_i + a_i')/2 = a$, for all $i \in I$ and $\sum_{j \in J} a_j = |J|a$. This implies that for all $i \in I$ we have $a_i - a = a - a_i'$, and we call this common value b_i. Notice that

(20)
$$\|b_i\|^2 = \langle a_i - a, a - a_i' \rangle = \|a\|^2,$$

16

where we use $\langle a_i, a_i' \rangle = 0$ (this follows from (ii)) and the definition of a. For all $i \in I$ the vectors $a \pm b_i$ are entrywise nonnegative which implies that $|b_i(k)| \leq a(k)$ for all $k \in [d]$ and $i \in I$. This fact combined with (20) implies that $b_i = s_i \circ a$, for some $s_i \in \{\pm 1\}^d$. Substituting $a_j = b_j + a$ in $\sum_{j \in J} a_j = |J| a$ it follows that $\sum_{j \in J} b_j = 0$, which in turn implies that $\sum_{j \in J} s_j \circ a = 0$. For $k \in [d]$ with $a(k) \neq 0$ we get $\sum_{j \in J} s_i(k) = 0$, a contradiction since $s_i \in \{\pm 1\}^d$ and $|J|$ is odd. As $\|a\| = \|c\| > 0$ (since $c \neq 0$ by assumption) there must exist a k such that $a(k) \neq 0$. □

Using Theorem 20 we now give a new family of matrices in $\mathcal{CS}_+ \setminus \mathcal{CP}$.

COROLLARY 21. *Let $n = 2\ell$, where $\ell \geq 3$ is odd. For $0 \leq k \leq n - 1$ define the Lorentz cone vectors $p_k := (1, \cos \frac{2\pi k}{n}, \sin \frac{2\pi k}{n})$. Clearly, we have that*

$$(p_k + p_{k+\ell})/2 = (1, 0, 0), \ \text{and} \ \langle p_k, p_{k+\ell} \rangle = 0, \ \text{for all} \ \ 0 \leq k \leq \ell - 1.$$

Furthermore, we have that $\langle p_k, p_{k'} \rangle \geq 0$ for all $0 \leq k, k' \leq n - 1$. Lastly, note that

$$\sum_{k=0}^{\ell-1} p_{2k} = \sum_{k=0}^{\ell-1} p_{2k+1} = \ell \cdot (1, 0, 0).$$

Since ℓ is odd, it follows from Theorem 20 that $X := \mathrm{Gram}(\{p_k\}_{k=0}^{n-1})$ is not completely positive. Moreover, as $\{p_k\}_{k=0}^{n-1} \subseteq \mathcal{L}_3$ it follows that $X \in \mathcal{GL}^n \setminus \mathcal{CP}^n$. In particular we have that $X \in \mathcal{CS}_+^n \setminus \mathcal{CP}^n$.

4.4. Gram-Lorentz behaviors. In view of Theorem 2, to any set $\mathcal{K} \subseteq \mathcal{CS}_+$ we can associate a family of quantum behaviors which we denote by $\mathcal{Q}_{\mathcal{K}}$. We refer to the quantum behaviors $\mathcal{Q}_{\mathcal{GL}}$ corresponding to $\mathcal{K} = \mathcal{GL}$ as *Gram-Lorentz behaviors*.

As it turns out Gram-Lorentz behaviors are quite interesting from a physical point of view. First of all, by Theorem 18 it follows that we can place an upper bound on the size of a quantum system necessary to generate all Gram-Lorentz behaviors, i.e.,

$$(21) \qquad \qquad \max\{\mathcal{D}(\mathbf{p}) : \mathbf{p} \in \mathcal{Q}_{\mathcal{GL}}\} < +\infty.$$

Note that (21) is in stark contrast to the case of arbitrary quantum behaviors, where no finite bound is currently known (recall Proposition 4 and the discussion preceding it). In fact, as was already mentioned in the introduction, the only quantum behaviors for which we can a piori bound the size of a quantum system necessary to generate them are the unbiased behaviors corresponding to a Bell scenario with binary outcomes [33]. In fact, we can recover this by combining (21) with the following result.

THEOREM 22. *In any $(m_A, m_B, 2, 2)$-scenario, all unbiased quantum behaviors are Gram-Lorentz behaviors.*

The proof of Theorem 22 is deferred to Section 5.2.1 (cf. Remark 5.3).

A second interesting fact is that there exist Gram-Lorentz behaviors for which any quantum representation has size exponential in m_A and m_B. Specifically, our main result in Section 5.2 (cf. Theorem 32) is that for any $n \geq 1$ there exists a Gram-Lorentz behavior $\mathbf{p}_n$ corresponding to the $(n, n, 2, 2)$-scenario satisfying $\mathcal{D}(\mathbf{p}_n) \geq 2^{\Omega(\sqrt{n})}$.

As an immediate consequence of this fact we get that no finite dimension suffices to generate all behaviors in $\cup_{n \geq 1} \mathcal{Q}(n, n, 2, 2)$. This was the main result in [34].

Lastly, the existence of Gram-Lorentz behaviors for which every quantum representation has exponential size is our crucial step for constructing Gram-Lorentz matrices whose cpsd-rank is exponential in terms of their size (cf. Section 5.3).

5. $\mathcal{CS}_+$ matrices whose cpsd-rank is exponential in terms of their size. This section is dedicated to the proof of Result 1, i.e., we show that for any $n \geq 1$ there exists a matrix $X_n \in \mathcal{GL}^{2n}$ such that cpsd-rank$(X_n) \geq 2^{\Omega(\sqrt{n})}$. The proof is given in Section 5.3 (cf. Theorem 33) and relies on Theorem 3. Specifically, given a quantum behavior $\mathbf{p} \in \mathcal{Q}$ it follows by Theorem 3 that cpsd-rank$(R) \geq \mathcal{D}(\mathbf{p})$, for any $R \in \mathcal{CS}_+^n \cap \mathcal{A}(\mathbf{p})$. Consequently, in order to derive Result 1 it suffices to identify a sequence of Gram-Lorentz behaviors $(\mathbf{p}_n)_{n \in \mathbb{N}}$ for which all quantum representations require a quantum system of size exponential in n. We show that for every $n \geq 1$ there exists a Gram-Lorentz behavior $\mathbf{p}_n$ corresponding to the $(n, n, 2, 2)$-scenario with the property that $\mathcal{D}(\mathbf{p}_n) \geq 2^{\Omega(\sqrt{n})}$. This is the main step for showing Result 1 and its proof is given in Section 5.2 (cf. Theorem 32). To prove this, instead of working with quantum behaviors we take the equivalent viewpoint of quantum correlations. This allows us to use a lower bound on the size of matrix representations of extremal quantum correlations, which is implicit in [33]. This is explained in Section 5.1 and Section B in the Appendix. In Section 5.2.1 we show that to any extremal quantum correlation C we can associate a Gram-Lorentz behavior $\mathbf{p}_C$ with the property that $\mathcal{D}(\mathbf{p}_C) \geq 2^{\Omega(\mathrm{rank}(C))}$. In Section 5.2.2 we focus on the $(n, n, 2, 2)$-scenario and show that any extreme point of the n-dimensional elliptope $\mathcal{E}_n$ is also an extreme point of the corresponding set of quantum correlations. It is well-known that for any $n \geq 1$ there exist extreme points of $\mathcal{E}_n$ with rank $\Theta(\sqrt{n})$. Thus, quantum behaviors corresponding to high-rank extreme points of the elliptope have the required properties. Furthermore, in Section 5.2.3 we give an explicit family of Gram-Lorentz behaviors achieving the exponential lower bound. We conclude the proof of Result 1 in Section 5.3 and give an explicit family of matrices with exponentially large cpsd-rank.

5.1. Quantum correlations. Throughout this section, for notational convenience we set $n := m_A$ and $m := m_B$. Furthermore, we focus on the $(n, m, 2, 2)$-scenario and we assume that the measurement outcomes are given by $\{\pm 1\}$. We denote by $\mathcal{Q}$ the corresponding set of quantum behaviors.

For the reader's convenience we have collected in this subsection some facts we use in later parts of this work. These results are well-known in the quantum information community but much less so in the mathematical optimization community.

We first describe a well-known equivalent parametrization of the set of quantum behaviors, which is the appropriate language for stating Tsirelson's theorem (cf. Theorem 24). To describe this we use the map $f : \mathbb{R}^{4nm} \to \mathbb{R}^{n+m+nm}$, which maps the behavior $\mathbf{p} = (p(ab|xy))$ to the vector $\mathbf{c} = (c_x, c_y, c_{xy})$ where

$$(22) \quad c_x := \sum_{a \in \{\pm 1\}} a \, p_A(a|x), \; c_y := \sum_{b \in \{\pm 1\}} b \, p_B(b|y), \text{ and } c_{xy} := \sum_{a,b \in \{\pm 1\}} ab \, p(ab|xy).$$

Note that c_{xy} corresponds to the expected value of the product of the players' outcomes, given that they performed measurements x and y, respectively. Similarly, c_x and c_y correspond to the expected values of the player's individual outcomes.

The map f is linear and injective. Consequently, the set $\mathcal{Q}$ of quantum behaviors is in one-to-one correspondence with $f(\mathcal{Q})$, i.e., the image of $\mathcal{Q}$ via the map f. We refer to $f(\mathcal{Q})$ as the set of *full quantum correlations*. A full quantum correlation (c_x, c_y, c_{xy}) is called *unbiased* if $c_x = c_y = 0$, for all x, y. Lastly, note that the inverse of f is the map $g : \mathbb{R}^{n+m+nm} \to \mathbb{R}^{4nm}$, which maps a full quantum correlation $\mathbf{c} = (c_x, c_y, c_{xy})$ to the behavior $\mathbf{p} := g(\mathbf{c})$ defined as

$$(23) \qquad\qquad p(ab|xy) = \frac{1 + a \, c_x + b \, c_y + ab \, c_{xy}}{4}.$$

18

The following lemma gives a characterization of the set of full quantum correlations. We have included a proof for completeness, which we also use in Remark 5.1.

LEMMA 23. *The vector* $\mathbf{c} = (c_x, c_y, c_{xy}) \in [-1,1]^{n+m+nm}$ *is a full quantum correlation if and only if there exist Hermitian operators* $\{M_x\}_x, \{N_y\}_y$ *with eigenvalues in* $[-1,1]$ *and a quantum state* ρ *such that, for all* x, y, *we have*

(24) $\qquad c_x = \mathrm{Tr}((M_x \otimes I)\rho), \; c_y = \mathrm{Tr}((I \otimes N_y)\rho), \; and \; c_{xy} = \mathrm{Tr}((M_x \otimes N_y)\rho).$

Proof. Consider $\mathbf{p} \in \mathcal{Q}$ such that $f(\mathbf{p}) = \mathbf{c}$ and let $\left\{ \{M_{a|x}\}_a, \{N_{b|y}\}_b, \rho \right\}$ be a quantum representation for $\mathbf{p}$. For $x \in [n]$ set $M_x := M_{1|x} - M_{-1|x}$ and for $y \in [m]$ set $N_y := N_{1|y} - N_{-1|y}$. Since $I = M_{1|x} + M_{-1|x} = N_{1|y} + N_{-1|y}$, for all x, y it follows that M_x and N_y have eigenvalues in $[-1,1]$. Lastly, using (22), an easy calculation shows that (24) is satisfied.

Conversely, let $\{M_x\}_x, \{N_y\}_y$ be Hermitian operators with eigenvalues in $[-1,1]$ and ρ a quantum state satisfying (24). For $x \in [n]$ and $a \in \{\pm 1\}$ set $M_{a|x} = \frac{I + aM_x}{2}$ and similarly, for $y \in [m]$ and $b \in \{\pm 1\}$ set $N_{b|y} = \frac{I + bN_y}{2}$. Note that $\{M_{a|x}\}_a$ and $\{N_{b|y}\}_b$ are valid POVMs. Lastly, defining the quantum behavior $\mathbf{p}$ where $p(ab|xy) = \mathrm{Tr}((M_{a|x} \otimes N_{b|y})\rho)$, it follows that $\mathbf{c} = f(\mathbf{p})$ and is thus a full quantum correlation.□

Given a full quantum correlation $\mathbf{c} = (c_x, c_y, c_{xy})$ we refer to any ensemble of Hermitian operators $\{\{M_x\}_x, \{N_y\}_y, \rho\}$ as defined in Lemma 23 as a *quantum representation* of $\mathbf{c}$. We say that a quantum representation of $\mathbf{c}$ is *d-dimensional* if $\{M_x\}_x, \{N_y\}_y \subseteq \mathcal{H}^d$ and $\rho \in \mathcal{H}_+^{d^2}$.

REMARK 5.1. *From the proof of Lemma 23, we see we have that* $\mathbf{p} \in \mathcal{Q}$ *has a d-dimensional quantum representation (as a bevavior) if and only if* $f(\mathbf{p})$ *has a d-dimensional quantum representation (as a full quantum correlation).*

We denote by $\mathrm{Cor}(n,m)$ the projection of the set of full quantum correlations onto $\mathbb{R}^{nm}$, that is, we only keep the entries $(c_{xy})_{xy}$, and refer to its elements as *quantum correlations*. It is sometimes useful to arrange the entries of a quantum correlation $\mathbf{c} \in \mathrm{Cor}(n,m)$ as a matrix C in $[-1,1]^{n \times m}$, in which case we write $C \in \mathrm{Cor}(n,m)$. Throughout this section we use these two forms interchangeably.

Tsirelson's theorem [33] given below has two important consequences: First, it characterizes the set of quantum correlations as the feasible region of a semidefinite program (cf. condition (iii)). Second, condition (ii) implies that all unbiased quantum behaviors can be generated using quantum systems of finite dimension.

THEOREM 24 ([33]). *For any* $C = (c_{xy}) \in [-1,1]^{n \times m}$ *the following are equivalent:*

(i) C *is a quantum correlation, i.e., there exist Hermitian operators* $\{M_x\}_x$, $\{N_y\}_y$ *with eigenvalues in* $[-1,1]$ *and a quantum state* ρ *satisfying*

$$c_{xy} = \mathrm{Tr}((M_x \otimes N_y)\rho), \;\; for \; all \;\; x \in [n], y \in [m].$$

(ii) *There exist unit vectors* $\{u_x\}_x$ *and* $\{v_y\}_y$ *in* $\mathbb{R}^{n+m}$ *such that*
   (a) $c_{xy} = \Psi_d^*(\gamma(u_x) \otimes \gamma(v_y)^\top)\Psi_d, \; for \; all \; x \in [n], y \in [m];$
   (b) $\Psi_d^*(\gamma(u_x) \otimes I)\Psi_d = 0, \; for \; all \; x \in [n];$
   (c) $\Psi_d^*(I \otimes \gamma(v_y)^\top)\Psi_d = 0, \; for \; all \; y \in [m],$
   where $d := 2^{\lfloor \frac{n+m}{2} \rfloor}$, Ψ_d *is the d-dimensional maximally entangled state defined in (6) and the map* γ *is defined in Theorem 12.*

(*iii*) *There exist unit vectors* $\{u_x\}_x$ *and* $\{v_y\}_y$ *in* $\mathbb{R}^{n+m}$ *such that*

$$c_{xy} = \langle u_x, v_y \rangle, \quad \text{for all} \ \ x \in [n], y \in [m].$$

The next result, which is implicit in Tsirelson's work [33], gives a lower bound on the size of a quantum representation for any extreme point of the set of quantum correlations. Since this is not stated explicitly in [33], for completeness we have included a short proof in Appendix B.

THEOREM 25 ([33]). *Let* $C = (c_{xy}) \in \text{ext}(\text{Cor}(n,m))$ *and consider a family of Hermitian operators* $\{M_x\}_x, \{N_y\}_y \subseteq \mathcal{H}^d$ *with eigenvalues in* $[-1,1]$ *and a quantum state* $\rho \in \mathcal{H}_+^{d^2}$ *satisfying* $c_{xy} = \text{Tr}((M_x \otimes N_y)\rho)$, *for all* x, y. *Then we have that*

$$d \geq \sqrt{2}^{\lfloor \text{rank}(C)/2 \rfloor}.$$

We note that Slofstra [32] generalized Tsirelson's lower bound given above by considering *near-extremal* quantum correlations and their *approximate representations*.

As we explain in the next section, Theorem 25 turns out to be the main ingredient for constructing cpsd matrices whose cpsd-rank is exponential in terms of their sizes.

5.2. Gram-Lorentz behaviors with large quantum representations. In this section we show that for every $n \geq 1$ there exists a Gram-Lorentz behavior $\mathbf{p}_n$ corresponding to the $(n, n, 2, 2)$-scenario such that $\mathcal{D}(\mathbf{p}_n) \geq 2^{\Omega(\sqrt{n})}$ (cf. Theorem 32).

5.2.1. Going from quantum correlations to Gram-Lorentz behaviors. By Theorem 24 we can associate a quantum behavior to any quantum correlation.

DEFINITION 26. *For any* $C \in \text{Cor}(n,m)$ *we denote by* $\mathbf{p}_C = (p_C(ab|xy))$ *the quantum behavior given by* $g((0,0,C))$. *Concretely, by* (23) *we have that*

$$(25) \qquad p_C(ab|xy) = \frac{1 + ab\, c_{xy}}{4}, \quad \text{for all} \ a, b, x, y.$$

It is also useful to arrange the entries of $\mathbf{p}_C$ into a $2n \times 2m$ matrix given by

$$(26) \qquad P_C := \sum_{a,b \in \{\pm 1\}, x, y \in [n]} p_C(ab|xy)\, e_a e_b^\mathsf{T} \otimes e_x e_y^\mathsf{T} = \frac{1}{4}\begin{pmatrix} J + C & J - C \\ J - C & J + C \end{pmatrix}.$$

REMARK 5.2. *Note that the behavior* $\mathbf{p}_C$ *is well-defined. This follows by Theorem* 24 (*ii*) *as* $(0, 0, C)$ *is a full quantum correlation vector for any* $C \in \text{Cor}(n,m)$.

As it turns out, behaviors constructed in this manner have interesting properties.

LEMMA 27. *For any* $C = (c_{x,y}) \in \text{Cor}(n,m)$ *the behavior* $\mathbf{p}_C$ *is Gram-Lorentz. In particular, consider unit vectors* $\{u_x\}_x$ *and* $\{v_y\}_y$ *in* $\mathbb{R}^{n+m}$ *such that* $c_{xy} = \langle u_x, v_y \rangle$, *for all* x, y (*these exist by Theorem* 24 (*iii*)). *Then we have that*

$$(27) \qquad p_C(ab|xy) = \langle \ell_a^x, \tilde{\ell}_b^y \rangle, \quad \text{for all} \ a, b, x, y, \ \text{where}$$

$$(28) \qquad \ell_a^x = \frac{1}{2}(1, au_x), \ \forall x \in X, a \in \{\pm 1\}, \quad \text{and} \quad \tilde{\ell}_b^y = \frac{1}{2}(1, bv_y), \ \forall y \in Y, b \in \{\pm 1\}.$$

20

Proof. By (25) we have that $p_C(ab|xy) = (1 + ab\, c_{xy})/4$, for all a, b, x, y. By Theorem 12 *(iii)* we get $c_{xy} = \langle u_x, v_y \rangle = \Psi_d^* \left(\gamma(u_x) \otimes \gamma(v_y)^\mathsf{T} \right) \Psi_d$, for all x, y, where $d := 2^{\lfloor \frac{n+m}{2} \rfloor}$. This gives

$$(29) \qquad p_C(ab|xy) = \Psi_d^* \left(\frac{I + a\gamma(u_x)}{2} \right) \otimes \left(\frac{I + b\gamma(v_y)^\mathsf{T}}{2} \right) \Psi_d, \text{ for all } a, b, x, y.$$

Set

$$\Gamma(\ell_a^x) = \frac{1}{\sqrt{d}} \left(\frac{I + a\, \gamma\,(u_x)}{2} \right) \in \mathcal{H}_+^d, \text{ for } a \in \{\pm 1\},$$

and

$$\Gamma(\tilde{\ell}_b^y) = \frac{1}{\sqrt{d}} \left(\frac{I + b\, \gamma\,(v_y)}{2} \right) \in \mathcal{H}_+^d, \text{ for } b \in \{\pm 1\},$$

where Γ was defined in (17). Using (7), it follows by (29) that

$$(30) \qquad p_C(ab|xy) = \langle \Gamma(\ell_a^x), \Gamma(\tilde{\ell}_b^y) \rangle = \langle \ell_a^x, \tilde{\ell}_b^y \rangle, \text{ for all } a, b, x, y,$$

where we used the fact that Γ is an isometry. Since the vectors $\{u_x\}_x$ and $\{v_y\}_y$ are unit it follows that the vectors $\{\ell_a^x\}_{a,x}$, $\{\tilde{\ell}_b^y\}_{b,y}$ belong to the Lorentz cone $\mathcal{L}_{m+n+1}$. Furthermore, by (28) we have that $\ell_1^x + \ell_{-1}^x = \tilde{\ell}_1^y + \tilde{\ell}_{-1}^y = e_1$, for all x, y implying $\mathrm{Gram}(\{\ell_a^x\}_{a,x}, \{\tilde{\ell}_b^y\}_{b,y}) \in \mathcal{A}(\mathbf{p}_C)$. Thus the behavior $\mathbf{p}_C$ is Gram-Lorentz. $\qquad \square$

REMARK 5.3. *As an immediate consequence of Lemma 27 it follows that every unbiased quantum behavior is Gram-Lorentz.*

We are now ready to translate Theorem 25 to Gram-Lorentz behaviors.

THEOREM 28. *For any $C \in \mathrm{ext}(\mathrm{Cor}(n, m))$ we have that $\mathbf{p}_C$ is Gram-Lorentz and*

$$\mathcal{D}(\mathbf{p}_C) \geq \sqrt{2}^{\lfloor \mathrm{rank}(C)/2 \rfloor}.$$

Proof. Fix $C \in \mathrm{ext}(\mathrm{Cor}(n, m))$ and let $\mathbf{p}_C = g((0, 0, C))$. We already determined in Lemma 27 that $\mathbf{p}_C$ is Gram-Lorentz. By definition, we have that $\mathcal{D}(\mathbf{p}_C)$ is equal to the least integer $d \geq 1$ for which $\mathbf{p}_C$ admits a d-dimensional representation. Since $(0, 0, C) = f(\mathbf{p}_C)$, by Remark 5.1 we know that $\mathcal{D}(\mathbf{p}_C)$ is also equal to the least integer $d \geq 1$ for which $(0, 0, C)$ admits a d-dimensional representation. By Theorem 25, the latter quantity is lower bounded by $\sqrt{2}^{\lfloor \mathrm{rank}(C)/2 \rfloor}$ as desired. $\qquad \square$

In view of Theorem 28, to construct Gram-Lorentz behaviors all of whose quantum representations require exponential size, it suffices to identify high-rank extreme points of $\mathrm{Cor}(n, m)$. In the next section we consider this problem for the case $n = m$.

5.2.2. High-rank extremal quantum correlations. Throughout this section we set $n = m$ and we view any $C \in \mathrm{Cor}(n, n)$ as a square $n \times n$ matrix.

Of special interest to us are the elements of $\mathrm{Cor}(n, n)$ whose diagonal entries are all equal to 1. Specifically, in our next lemma below we show they coincide with the n-dimensional *elliptope*, denoted by $\mathcal{E}_n$, which is defined as the set of $n \times n$ symmetric psd matrices with diagonal entries equal to 1. The elliptope is a spectrahedral set whose structure has been extensively studied (e.g. see [8] and references therein).

We begin this section by determining a useful relation between $\mathcal{E}_n$ and $\mathrm{Cor}(n, n)$.

PROPOSITION 29. *We have that $\mathrm{ext}(\mathcal{E}_n) \subseteq \mathrm{ext}(\mathrm{Cor}(n, n))$.*

Proof. Fix $X \in \text{ext}(\mathcal{E}_n)$ and let $X = \lambda A + (1 - \lambda)B$, where $A, B \in \text{Cor}(n,n)$ and $\lambda \in [0,1]$. For all $i \in [n]$ we have that $1 = \lambda A_{ii} + (1 - \lambda)B_{ii}$ and since $A_{ii}, B_{ii} \in [-1,1]$ it follows that $1 = A_{ii} = B_{ii}$, for all $i \in [n]$.

We now show that $A, B \in \mathcal{E}_n$, and the proof is concluded by the extremality assumption. By Theorem 24 (*iii*) there exist unit vectors $\{u_i\}_i$ and $\{v_j\}_j$ such that $A_{ij} = \langle u_i, v_j \rangle$, for all $i,j \in [n]$. By the Cauchy-Schwartz inequality we have that $1 = A_{ii} = \langle u_i, v_i \rangle \leq 1$, for all $i \in [n]$. Thus, equality holds throughout which implies that u_i is parallel to v_i, for all $i \in [n]$. Lastly, as all vectors have unit norm we have $u_i = v_i$ for all $i \in [n]$ and thus $A \in \mathcal{E}_n$. Similarly, we have $B \in \mathcal{E}_n$. $\square$

Using Proposition 29 we can construct extreme points of $\text{Cor}(n,n)$ using extreme points of $\mathcal{E}_n$. This is extremely useful as the extreme points of the elliptope are completely understood. We explain this in the remaining part of this section.

Throughout, we denote by $r_{\max}(n)$ the greatest integer satisfying $\binom{r+1}{2} \leq n$, i.e.,

$$r_{\max}(n) = \left\lfloor \frac{\sqrt{1 + 8n} - 1}{2} \right\rfloor.$$

We now state two well-known results concerning properties of extreme points of the elliptope that we use in the next section. The first one due to [23] (see also [8, Corollary 31.5.4]) allows one to easily check whether a matrix $X \in \mathcal{E}_n$ is an extreme point.

THEOREM 30 ([23]). *Let $X \in \mathcal{E}_n$ with $\text{rank}(X) = r$ and let $\{u_i\}_{i=1}^n \in \mathbb{R}^r$ be a Gram representation of X. Then $X \in \text{ext}(\mathcal{E}_n)$ if and only if*

$$\dim(\text{span}(\{u_i u_i^\mathsf{T} : i \in [n]\})) = \binom{r+1}{2}.$$

The second result due to [15] (see also [8, Proposition 31.5.7]) specifies the range of possible ranks for the extreme points of the elliptope and moreover shows that every value in that range is achievable.

THEOREM 31 ([15]). *For any $X \in \text{ext}(\mathcal{E}_n)$ we have that $\text{rank}(X) \leq r_{\max}(n)$. Furthermore, for any integer r in the range $1 \leq r \leq r_{\max}(n)$ there exists $X_r \in \text{ext}(\mathcal{E}_n)$ with $r = \text{rank}(X_r)$.*

EXAMPLE 5.1. *We now describe the constructive part of Theorem 31 which we use in the next section. Fix an integer r satisfying $1 \leq r \leq r_{\max}(n)$. In particular we have that $\binom{r+1}{2} < n + 1$. Let $\{e_i\}_{i=1}^r$ be the standard basis in $\mathbb{R}^r$. For $i,j \in [r]$ define $w_{i,j} := \frac{1}{\sqrt{2}}(e_i + e_j)$. Define X_r to be the Gram matrix of the following family of vectors: we use e_1 repeated $n + 1 - \binom{r+1}{2}$ times, followed by $e_2, \ldots, e_r$ one time each and lastly, we use w_{ij} for all $1 \leq i < j \leq [r]$. Clearly, we have that $\text{rank}(X_r) = r$. Furthermore, since the matrices $e_i e_i^\mathsf{T}, w_{ij} w_{ij}^\mathsf{T}$ are linearly independent it follows that*

$$\dim\left(\text{span}\left(\{e_i e_i^\mathsf{T}\}_{i=1}^r, \{w_{ij} w_{ij}^\mathsf{T}\}_{1 \leq i < j \leq [r]}\right)\right) = r + \binom{r}{2} = \binom{r+1}{2}.$$

By Theorem 30 it follows that $X_r \in \text{ext}(\mathcal{E}_n)$.

5.2.3. Putting everything together. Combining the results given in Sections 5.2.1 and 5.2.2 we now show that for every $n \geq 1$ there exists a Gram-Lorentz behavior $\mathbf{p}_n$ corresponding to the $(n,n,2,2)$-scenario satisfying $\mathcal{D}(\mathbf{p}_n) \geq 2^{\Omega(\sqrt{n})}$.

THEOREM 32. *Fix $n \geq 1$ and let $C_n \in \text{ext}(\mathcal{E}_n)$ with $\text{rank}(C_n) = r_{\max}(n)$. Then*

$$\mathcal{D}(\mathbf{p}_{C_n}) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}.$$

Proof. By Theorem 31 there exists $C_n \in \text{ext}(\mathcal{E}_n)$ with $\text{rank}(C_n) = r_{\max}(n)$. By (29) it follows that $C_n \in \text{ext}(\text{Cor}(n,n))$. The proof is concluded by Theorem 28. □

We conclude this section with an explicit example. To ease the exposition we only consider matrices of size $N := 2n^2 + n$, for any $n \geq 1$. In this case $r_{\max}(N) = 2n$.

By Theorem 31 there exists $C_n \in \text{ext}(\mathcal{E}_N)$ with $\text{rank}(C_n) = 2n$. As described in Example 5.1, the matrix C_n is defined as the Gram matrix of the vectors

(31) $\qquad w_{ii} := e_i, \text{ for } i \in [2n] \text{ and } w_{ij} := \frac{1}{\sqrt{2}}(e_i + e_j), \text{ for } 1 \leq i < j \leq [2n].$

It is instructive to think of the underlying Bell scenario as each player having $\binom{2n+1}{2}$ questions that are indexed by the 2-element *multisets* of $[2n]$. In particular, the first $2n$ questions correspond to the multisets $\{\{i,i\} : i \in [2n]\}$ and the remaining $\binom{2n}{2}$ questions correspond to $\{\{i,j\} : 1 \leq i < j \leq 2n\}$.

By construction, the entries of C_n are given by

(32) $$C_n = \begin{pmatrix} I_n & A_n \\ A_n^\mathsf{T} & B_n \end{pmatrix}, \text{where}$$

(33) $\qquad A_n[ii, kl] = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } i \in \{k,l\}, \\ 0, & \text{otherwise,} \end{cases}$ and $B_n[ij, kl] = \frac{1}{2}|\{i,j\} \cap \{k,l\}|.$

Lastly, using (26) we have that

(34) $$P_{C_n} = \frac{1}{4}\begin{pmatrix} J + C_n & J - C_n \\ J - C_n & J + C_n \end{pmatrix}.$$

5.3. cpsd-matrices with high cpsd-rank. In this section we give the proof of Result 1, i.e., we show that for any $n \geq 1$ there exists $X_n \in \mathcal{GL}^{2n}$ such that $\text{cpsd-rank}(X_n) \geq 2^{\Omega(\sqrt{n})}$. This follows by combining Theorem 32 with Theorem 3.

THEOREM 33. *Fix $n \geq 1$ and let $C_n \in \text{ext}(\mathcal{E}_n)$ with $\text{rank}(C_n) = r_{\max}(n)$. Then*

(35) $$P_{C_n} := \frac{1}{4}\begin{pmatrix} J + C_n & J - C_n \\ J - C_n & J + C_n \end{pmatrix}$$

is a $2n \times 2n$ Gram-Lorentz matrix satisfying

(36) $$\text{cpsd-rank}(P_{C_n}) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}.$$

Proof. By Lemma 27 we get that $P_{C_n} \in \mathcal{GL}^{2n}$. Furthermore, as $P_{C_n} \in \mathcal{GL}^{2n} \subseteq \mathcal{CS}_+^{2n}$ we have that $X_n := \begin{pmatrix} P_{C_n} & P_{C_n} \\ P_{C_n} & P_{C_n} \end{pmatrix} \in \mathcal{CS}_+^{4n}$, since the psd matrices in the $\mathcal{CS}_+$-factorization can be repeated. Also, we clearly have that $X_n \in \mathcal{A}(\mathbf{p}_{C_n})$. Thus, by Theorem 3, we get $\text{cpsd-rank}(X_n) \geq \mathcal{D}(\mathbf{p}_{C_n})$. It is easy to verify that $\text{cpsd-rank}(X_n) = \text{cpsd-rank}(P_{C_n})$. Lastly, since $C_n \in \text{ext}(\text{Cor}(n,n))$, by Theorem 32 we have $\mathcal{D}(\mathbf{p}_{C_n}) \geq \sqrt{2}^{\lfloor r_{\max}(n)/2 \rfloor}$ and the proof is concluded. □

REMARK 5.4. *In Theorem 18 we determined that for any $X \in \mathcal{GL}^n$ we have that* $\text{cpsd-rank}(X) \leq 2^{\lfloor(\text{rank}(X)+1)/2\rfloor}$. *Since* $\text{rank}(P_{C_n}) \leq \text{rank}(C_n) + 1$, *this upper bound applied to the matrices* P_{C_n} *defined in (35) shows that for all* $n \geq 1$ *we have* $\text{cpsd-rank}(P_{C_n}) \leq 2^{\lfloor(r_{\max}(n)+2)/2\rfloor} = 2^{\lfloor r_{\max}(n)/2\rfloor+1}$. *Thus, in view of (36), the upper bound on the* cpsd-rank *of* $\mathcal{GL}$ *matrices given in Theorem 18 is essentially tight.*

Returning to the example (34) from Section 5.2.3 it follows that $P_{C_n} \in \mathcal{GL}^{2(2n^2+n)}$ and $\text{cpsd-rank}(P_{C_n}) \geq \sqrt{2}^{\lfloor(2n-1)/2\rfloor}$. In particular, Lemma 27 implies that the vectors

$$\ell_a^{ij} := \left(\frac{1}{2}, \frac{a\, w_{ij}}{2}\right), \ a \in \{\pm 1\}, \ 1 \leq i \leq j \leq 2n,$$

lie in $\mathcal{L}_{2n+1}$ and give a $\mathcal{GL}$-factorization of P_{C_n} (for the definition of the w_{ij}'s see (31)). The corresponding $\mathcal{CS}_+$-factorization is given by the psd matrices $\{\Gamma(\ell_a^{ij})\}_{ij,a}$, where

$$\Gamma(\ell_a^{ij}) = \frac{1}{2^{n/2}}\left(\frac{I + a\,\gamma(w_{ij})}{2}\right) \in \mathcal{H}_+^{2^n}.$$

6. cpsd-graphs. We say that $G = ([n], E)$ is a *cpsd-graph* if for any matrix $X \in \mathcal{DNN}^n$ whose support is given by G, i.e., $S(X) = G$, we have that $X \in \mathcal{CS}_+^n$. The analogous notion of *cp-graphs* has been studied extensively (e.g. see [3, Section 2.5]). In fact, the class of cp-graphs admits an exact characterization: A graph is cp if and only if it does not contain an odd cycle C_{2t+1} ($t \geq 2$) as a subgraph [19]. In this section we show that the same characterization extends to cpsd-graphs (cf. Theorem 36).

To arrive at the characterization of cpsd-graphs, we generalize a sufficient condition from [12] for constructing doubly-nonnegative matrices that are not cpsd. As noted in [21], the example of the matrix in $\mathcal{DNN}^5 \setminus \mathcal{CS}_+^5$ given in [12] does not admit a Gram factorization by positive elements in any finite von Neumann algebra. Our sufficient condition given in Theorem 34 below generalizes this construction.

6.1. $\mathcal{DNN}$ matrices with no $\mathcal{N}^+$-factorizations. First, we introduce some necessary background on von Neumann algebras. We keep the discussion to a minimum and refer the interested reader to [25] for a comprehensive introduction.

A *von Neumann algebra* is a unital $*$-subalgebra of the C^*-algebra of bounded operators on a Hilbert space H, that is closed in the weak operator topology. A von Neumann algebra $\mathcal{N}$ is called *tracial* if it is equipped with a linear functional $\tau : \mathcal{N} \to \mathbb{C}$ satisfying: (i) $\tau(x^*x) \geq 0$ for all $x \in \mathcal{N}$ and $\tau(1) = 1$ (ii) $\tau(x^*x) = 0 \Longrightarrow x = 0$ (iii) $\tau(xy) = \tau(yx)$, for all $x, y \in \mathcal{N}$ and (iv) the restriction of τ to the unit ball is continuous with respect to the weak operator topology.

An element $p \in \mathcal{N}$ is called *positive* if $p = x^*x$, for some $x \in \mathcal{N}$. We denote by $\mathcal{N}^+$ the set of positive elements in $\mathcal{N}$. We make use of the fact that any $p \in \mathcal{N}^+$ has a unique positive square root (e.g. see [25, Theorem 2.2.1]).

REMARK 6.1. *Let $(\mathcal{N}, \tau)$ be a tracial von Neumann algebra. Let $\{x_i\}_{i=1}^n \subseteq \mathcal{N}$ such that $x_i^* = x_i$ for all $i \in [n]$ and set $X := (\tau(x_ix_j)_{1\leq i,j\leq n})$. For any $u \in \text{Ker}X$ we have that $\sum_{i=1}^n x_iu_i = 0$. Indeed, note that*

$$0 = u^*Xu = \sum_{i,j=1}^n \bar{u}_i u_j \tau(x_ix_j) = \tau\left(\left(\sum_{i=1}^n u_ix_i\right)^*\left(\sum_{i=1}^n u_ix_i\right)\right),$$

which by (ii) implies that $\sum_{i=1}^n u_ix_i = 0$.

Moreover, if $\tau(pq) = 0$ where p, q are positive elements of $\mathcal{N}$ then we have that $pq = 0$. To see this let $p = a^*a$ and $q = b^*b$ and note that $\tau(pq) = \tau(a^*ab^*b) = \tau((ab^*)^*ab^*) = 0$ which by (ii) implies that $ab^* = 0$. This shows that $pq = 0$.

Let $(\mathcal{N}, \tau)$ be a tracial von Neumann algebra. We say that a matrix $X \in \mathcal{DNN}^n$ admits an $\mathcal{N}^+$-*factorization* if there exist positive elements $\{p_i\}_{i=1}^n \subseteq \mathcal{N}^+$ such that $X = (\tau(p_i p_j)_{1 \le i \le j \le n})$. Next we give a sufficient condition for constructing $\mathcal{DNN}$ matrices for which no $\mathcal{N}^+$-factorization exists, generalizing a construction from [12].

THEOREM 34. *Consider nonzero vectors $\{u_i\}_{i=1}^n \subseteq \mathbb{R}^d$ such that $\langle u_i, u_j \rangle \ge 0$ for all $i, j \in [n]$. Assume that there exist subsets $I, J \subseteq [n]$ with the following properties:*
 (i) $\text{span}(\{u_i : i \in I\}) = \text{span}(\{u_j : j \in J\}) = \text{span}(\{u_i : i \in [n]\})$;
 (ii) There exists $i^* \in I$ such that $\langle u_{i^*}, u_i \rangle = 0$, for all $i \in I \setminus \{i^*\}$;
 (iii) There exists $j^* \in J$ such that $\langle u_{j^*}, u_j \rangle = 0$, for all $j \in J \setminus \{j^*\}$;
 (iv) The vector u_{i^*} is not parallel to u_{j^*};
 (v) We have $\langle u_{i^*}, u_{j^*} \rangle \ne 0$.
Then the matrix $\text{Gram}(\{u_i\}_{i=1}^n)$ does not admit an $\mathcal{N}^+$-factorization for any tracial von Neumann algebra $(\mathcal{N}, \tau)$.

Proof. Let $(\mathcal{N}, \tau)$ be a tracial von Neumann algebra and let $\text{Gram}(\{u_i\}_{i=1}^n) = (\tau(p_i p_j)_{1 \le i \le j \le n})$, for some positive elements $\{p_i\}_{i=1}^n \subseteq \mathcal{N}$. By (i) we have that $u_{i^*} \in \text{span}(\{u_j : j \in J\})$ so Remark 6.1 implies that $p_{i^*} \in \text{span}(\{p_j : j \in J\})$. Pre-multiplying this by p_{j^*}, it follows from (iii) that $p_{j^*} p_{i^*} \in \text{span}(\{p_{j^*}^2\})$, where we have utilized the fact that $\tau(p_{j^*} p_{i^*}) = 0$ implies $p_{j^*} p_{i^*} = 0$. Analogously, (i) implies that $p_{j^*} \in \text{span}(\{p_i : i \in I\})$ and post-multiplying by p_{i^*} we get from (ii) that $p_{j^*} p_{i^*} \in \text{span}(\{p_{i^*}^2\})$. By (v) we get $p_{j^*} p_{i^*} \ne 0$ and combining the two equations, there exists a scalar $c \ne 0$ such that $p_{i^*}^2 = c p_{j^*}^2$. Also note that $c > 0$ since $0 < \tau(p_{i^*}^2) = c\tau(p_{j^*}^2)$ and $\tau(p_{j^*}^2) > 0$. Since each positive element of a C^*-algebra has a unique positive square root we have $p_{i^*} = \sqrt{c} p_{j^*}$. This contradicts (iv). $\square$

Based on Theorem 34, we now give a family of $\mathcal{DNN}$ matrices supported by C_{2t+1} (for all $t \ge 1$) that do not admit a Gram factorization with positive elements in any tracial von Neumann algebra.

LEMMA 35. *Let A_t denote the adjacency matrix of C_{2t+1}, $(t \ge 2)$, and let λ_t be its least eigenvalue. The matrix $A_t - \lambda_t I$ is doubly-nonnegative, its support is C_{2t+1}, and it does not admit an $\mathcal{N}^+$-factorization for any tracial von Neumann algebra $(\mathcal{N}, \tau)$.*

Proof. Set $n := 2t+1$ and $X := A_t - \lambda_t I$. Clearly, $X \in \mathcal{DNN}^n$ and $S(X) = C_{2t+1}$. Note that $\lambda_t = 2\cos(\frac{2\pi t}{2t+1})$ with multiplicity 2. In particular $\text{rank}(X) = n - 2$. Let $X = \text{Gram}(\{u_i\}_{i=1}^n)$ where $\{u_i\}_{i=1}^n \subseteq \mathbb{R}^{n-2}$ and $\text{span}(\{u_i\}_{i=1}^n) = \mathbb{R}^{n-2}$. We show that the assumptions of Theorem 34 are satisfied for $I := [n] \setminus \{2, n\}$ and $J := [n] \setminus \{1, 3\}$. For (i) note that $\dim(\text{span}(\{u_i : i \in I\})) = \dim(\text{span}(\{u_j : j \in J\})) = n - 2$ and since $\{u_i\}_{i=1}^n \subseteq \mathbb{R}^{n-2}$, we have $\text{span}(\{u_i : i \in I\})) = \dim(\text{span}(\{u_j : j \in J\}))$. Moreover, setting $i^* := 1$ and $j^* := 2$, we see that (ii) and (iii) are satisfied. For (iv) note that $\det(X[1,2]) = \lambda_t^2 - 1 \ne 0$, where $X[1,2]$ denotes the principal submatrix of X corresponding to the first two rows and columns. Lastly, (v) holds as $\langle u_1, u_2 \rangle = +1$. $\square$

REMARK 6.2. *It was shown in [7] that there exists a tracial von Neumann algebra $(\mathcal{N}, \tau)$ such that any element in the closure of $\mathcal{CS}_+$ admits an $\mathcal{N}^+$-factorization. Consequently, the matrices $A_t - \lambda_t I$ constructed in Lemma 35 are doubly-nonnegative and do not belong to the closure of $\mathcal{CS}_+$.*

6.2. Characterizing cpsd-graphs. Using the family of matrices constructed in Lemma 35 we are now ready to complete our characterization of cpsd-graphs.

THEOREM 36. *A graph is cpsd if and only if it has no C_{2t+1}-subgraph ($t \geq 2$).*

Proof. Consider a graph G and suppose it has no C_{2t+1}-subgraph for all $t \geq 2$. Then G is a cp-graph and thus, also a cpsd-graph. Conversely, consider a graph G that contains a C_{2t+1}-subgraph, for some $t \geq 2$. We show that G is not a cpsd-graph. First, suppose that $G = C_{2t+1}$ for some $t \geq 2$. It follows from Lemma 35 and Remark 6.2 that odd cycles of length at least 5 are not cpsd-graphs so we are done. Next suppose that $G = ([n], E)$ contains C_{2t+1} (for some $t \geq 2$) as a proper subgraph. Let A_t and λ_t be as in Lemma 35. Recall that $X = A_t - \lambda_t I \in \mathcal{DNN} \setminus \mathrm{cl}(\mathcal{CS}_+)$. Let $\tilde{X}$ be the $n \times n$ matrix whose principal submatrix corresponding to the vertices of C_{2t+1} is given by X, and all other entries are equal to 0. For any $a > 0$, since $\tilde{X} + aI$ is positive definite, we can find $0 < b < a$ such that $X_a := \tilde{X} + aI + bA_G \in \mathcal{DNN}$, where A_G is the adjacency matrix of G. By a continuity argument we see that $\mathrm{cl}(\mathcal{CS}_+)$ is closed under taking principal submatrices. Thus, as $\lim_{a \to 0} X_a = \tilde{X}$ and $\tilde{X} \notin \mathrm{cl}(\mathcal{CS}_+)$, there exists $a^* > 0$ such that $X_{a^*} \in \mathcal{DNN} \setminus \mathrm{cl}(\mathcal{CS}_+)$. In particular, we have that $X_{a^*} \in \mathcal{DNN} \setminus \mathcal{CS}_+$. As $S(X_{a^*}) = G$, it follows that G is not a cpsd-graph. □

REFERENCES

[1] J. S. BELL, *On the Einstein Podolsky Rosen paradox*, Physics, 1 (1964), pp. 195–200.

[2] J. S. BELL, *On the problem of hidden variables in quantum mechanics*, Rev. Mod. Phys., 38 (1966).

[3] A. BERMAN AND N. SHAKED-MONDERER, *Completely Positive Matrices*, World Scientific, 2003.

[4] I. N. BOMZE, W. SCHACHINGER, AND R. ULLRICH, *New lower bounds and asymptotics for the cp-rank*, SIAM. J. Matrix Anal. A., 36 (2015), pp. 20–37.

[5] N. BRUNNER, D. CAVALCANTI, S. PIRONIO, V. SCARANI, AND S. WEHNER, *Bell nonlocality*, Rev. Mod. Phys., 86 (2014), p. 419.

[6] S. BURER, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Math. Program. Ser. A, 120 (2009), pp. 479–495.

[7] S. BURGDORF, M. LAURENT, AND T. PIOVESAN, *On the closure of the completely positive semidefinite cone and linear approximations to quantum colorings.* arXiv:1502.02842, 2015.

[8] M. DEZA AND M. LAURENT, *Geometry of cuts and metrices*, Springer, 1997.

[9] J. DREW, C. JOHNSON, AND R. LOEWY, *Completely positive matrices associated with M-matrices*, Linear Multilinear A., 37 (1994), pp. 303–310.

[10] H. FAWZI, J. GOUVEIA, P. PARRILO, R. Z. ROBINSON, AND R. THOMAS, *Positive semidefinite rank*, Math. Program., 153 (2015), pp. 133–177.

[11] S. FIORINI, S. MASSAR, S. POKUTTA, H. TIWARY, AND R. DE WOLF, *Exponential lower bounds for polytopes in combinatorial optimization*, J. ACM, 62 (2015), pp. 1–23.

[12] P. E. FRENKEL AND M. WEINER, *On vector configurations that can be realized in the cone of positive matrices*, Linear Alg. Appl., 459 (2014), pp. 465–474.

[13] R. GOODMAN AND N. R. WALLACH, *Symmetry, Representations, and Invariants*, Springer, 2009.

[14] J. GOUVEIA, P. PARRILO, AND R. THOMAS, *Lifts of convex sets and cone factorizations*, Math. Oper. Res., 38 (2013), pp. 248–264.

[15] R. GRONE, S. PIERCE, AND W. WATKINS, *Extremal correlation matrices*, Linear Alg. Appl., 132 (1990), pp. 63–70.

[16] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[17] R. Jain, Y. Shi, Z. Wei, and S. Zhang, *Efficient protocols for generating bipartite classical distributions and quantum states.*, IEEE Trans. Inform. Theory, 59 (2013), pp. 5171–5178.

[18] Z. Ji, *Binary constraint system games and locally commutative reductions.* arXiv:1310.3794, 2013.

[19] N. Kogan and A. Berman, *Characterization of completely positive graphs*, Discrete Math., 114 (1993), pp. 298–304.

[20] J. B. Lasserre, *New approximations for the cone of copositive matrices and its dual*, Mathematical Programming, 144 (2013), pp. 265–276.

[21] M. Laurent and T. Piovesan, *Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone*, SIAM J. Optimiz., 25 (2015), pp. 2461–2493.

[22] T. Lee and Z. Wei, *The square root rank of the correlation polytope is exponential.* arXiv:1411.6712, 2014.

[23] C.-K. Li and B.-S. Tam, *A note on extremal correlation matrices*, SIAM. J. Matrix Anal. A., 15 (1994), pp. 903–908.

[24] J. E. Maxfield and H. Minc, *On the matrix equation $X'X = A$*, P. Edinburgh Math. So. (Series 2), 13 (1962), pp. 125–129.

[25] G. J. Murphy, *C^*-Algebras and Operator Theory*, Academic Press, 1990.

[26] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, 2000.

[27] K. F. Pál and T. Vértesi, *Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems*, Phys. Rev. A, 82 (2010), p. 022116.

[28] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, PhD thesis, California Institute of Technology, 2000.

[29] D. E. Roberson, *Conic formulations of graph homomorphisms*, J. Algebr. Comb., (2016), pp. 1–37.

[30] J. Sikora and A. Varvitsiotis, *Linear conic formulations for two-party correlations and values of nonlocal games.* arXiv:1506.07297, 2015.

[31] J. Sikora, A. Varvitsiotis, and Z. Wei, *On the minimum dimension of a Hilbert space needed to generate a quantum correlation.* arXiv:1507.00213, 2015.

[32] W. Slofstra, *Lower bounds on the entanglement needed to play xor non-local games*, J. Math. Phys., 52 (2011), p. 102202.

[33] B. S. Tsirelson, *Quantum analogues of the Bell inequalities: The case of two spatially separated domains*, Journal of Soviet Mathematics, 36 (1987), pp. 557–570.

[34] T. Vértesi and K. F. Pál, *Bounding the dimension of bipartite quantum systems*, Phys. Rev. A, 79 (2009), p. 042106.

[35] M. Yannakakis, *Expressing combinatorial optimization problems by linear programs*, J. Comput. System Sci., 43 (1991), pp. 441–466.

Appendix A. Clifford algebras. Our goal in this section is to briefly introduce Clifford algebras. For additional details the reader is referred to [13, Chapter 6].

Consider a real vector space V equipped with a bilinear form $\beta : V \times V \to \mathbb{R}$ such that (i) $\beta(x,y) = \beta(y,x), \forall x, y \in V$ and (ii) β is *non-degenerate*, i.e., $\forall x \in V, \beta(x,y) = 0 \implies y = 0$. A *Clifford algebra* for (V, β) consists of a real unital associative algebra denoted $\mathrm{Cl}(V, \beta)$ together with a linear map $e : V \to \mathrm{Cl}(V, \beta)$ satisfying:

(i) $e(u)e(v) + e(v)e(u) = \beta(u,v)1$, for all $u, v \in V$;

(ii) $\mathrm{Cl}(V, \beta)$ is generated by $e(V)$ as an algebra;

(iii) Given a real unital associative algebra $\mathcal{A}$ and linear map $f : V \to \mathcal{A}$ satisfying

$$f(u)f(v) + f(v)f(u) = \beta(u,v)1, \text{ for all } u, v \in V,$$

there exists a unique algebra homomorphism $h : \mathrm{Cl}(V, \beta) \to \mathcal{A}$ where $f = h \circ e$.

A Clifford algebra for (V, β) can be explicitly defined as the quotient algebra $\mathcal{T}(V)/\mathcal{I}(V)$, where $\mathcal{T}(V) := \oplus_{k \geq 0} V^{\otimes k}$ is the tensor algebra over V and $\mathcal{I}(V)$ is the two-sided ideal in $\mathcal{T}(V)$ generated by the elements of the form $u \otimes v + v \otimes u - \beta(u,v)1$, for all $u, v \in V$. Any two algebras satisfying conditions $(i), (ii), (iii)$ above are isomorphic. Thus we refer to $\mathrm{Cl}(V, \beta)$ as *the* Clifford algebra over V.

A *representation* of an associative algebra $\mathcal{A}$ consists of a vector space W together with an algebra homomorphism $\Gamma : \mathcal{A} \to \mathrm{End}(W)$, i.e., a linear map preserving multiplication and the unit element, where $\mathrm{End}(W)$ is the set of all endomorphisms of W. The *dimension* of a representation (Γ, W) is the dimension of W as a vector space. A *subrepresentation* of a representation (Γ, W) is a subspace $U \subseteq W$ such that $\Gamma(a)(U) \subseteq U$, for all $a \in \mathcal{A}$. A representation is called *irreducible* if its only subrepresentations are itself and the trivial vector space.

It is well-known that the irreducible representations of $\mathrm{Cl}(V, \beta)$ have exponential size in terms of the dimension of V. This is the source of our exponential lower bound in this paper. Specifically, it is known that:

THEOREM 37. *Let β be a nondegenerate bilinear form on V.*
(i) *If $\dim V = 2\ell$ then (up to isomorphism) there exists a unique irreducible representation of $\mathrm{Cl}(V, \beta)$ which has dimension 2^ℓ;*
(ii) *If $\dim V = 2\ell + 1$ then there exist two nonisomorphic irreducible representations of $\mathrm{Cl}(V, \beta)$. Both representations have dimension 2^ℓ.*

For a proof of this fact the reader is referred to [13, Theorem 6.1.3].

REMARK A.1. *Consider a real vector space V equipped with a symmetric and nondegenerate bilinear form $\beta : V \times V \to \mathbb{R}$. Let f be a linear map $f : V \to \mathrm{End}(W)$ satisfying $f(u)f(v) + f(v)f(u) = \beta(u, v)1_W$, for all $u, v \in V$. Using the three defining axioms for $\mathrm{Cl}(V, \beta)$ it follows that f can be extended to a representation for $\mathrm{Cl}(V, \beta)$.*

Appendix B. Proof of Theorem 25.
In this section we give for completeness a proof of Theorem 25, as this is not stated explicitly in [33]. We start with a definition.

DEFINITION 38. *Given $C = (c_{xy}) \in \mathrm{Cor}(n, m)$, we say that a family of real vectors $\{u_x, v_y\}_{x,y}$ forms a C-system if they satisfy*

$$\|u_x\| \leq 1, \ \forall x, \ \|v_y\| \leq 1, \ \forall y, \ \text{and } c_{xy} = \langle u_x, v_y \rangle, \forall x, y.$$

As it turns out, C-systems of vectors corresponding to extremal quantum correlations have interesting properties. For our purposes we only need the following result:

LEMMA 39. *[33, Lemma 3.1] Let $C \in \mathrm{ext}(\mathrm{Cor}(n, m))$. Then, for any C-system of vectors $\{u_x, v_y\}_{x,y}$ we have that*

$$(37) \qquad \mathrm{span}(\{u_x\}_x) = \mathrm{span}(\{v_y\}_y).$$

Furthermore, there exists an integer $\tau_C \geq 1$, depending only on C, such that for any C-system of vectors $\{u_x, v_y\}_{x,y}$ we have that

$$(38) \qquad \dim\left(\mathrm{span}(\{u_x\}_x)\right) = \dim\left(\mathrm{span}(\{v_y\}_y)\right) = \tau_C.$$

Also, we can find C-systems $\{u_x, v_y\}_{x,y}$ that lie in $\mathbb{R}^{\tau_C}$ (and thus span $\mathbb{R}^{\tau_C}$). For this, let $\{a_x, b_y\}_{x,y}$ be an arbitrary C-system and consider the matrix $\mathrm{Gram}(\{a_x\}_x, \{b_y\}_y)$. By (38) and (37), this is a real psd matrix of rank τ_C and thus any Gram factorization with vectors in $\mathbb{R}^{\tau_C}$ gives a C-system with the required properties. Lastly, note that $\mathrm{rank}(C) \leq \tau_C$.

We continue by stating another result due to Tsirelson which shows that the operators in a quantum representation of an extremal quantum correlation correspond to a representation of an appropriate Clifford algebra. This is the essential ingredient in the proof of Theorem 25 given below.

THEOREM 40. *[33, Theorem 3.1] Let $C = (c_{xy}) \in \text{ext}(\text{Cor}(n,m))$ and consider a family of Hermitian operators $\{A_x\}_x, \{B_y\}_y, \rho$ in $\mathcal{H}^d$ such that:*

    (i) $c_{xy} = \text{Tr}(A_x B_y \rho)$ for all x, y;

    (ii) $A_x B_y = B_y A_x$;

    (iii) ρ is a density matrix;

    (iv) The eigenvalues of A_x, B_y are in $[-1, 1]$;

    (v) There does not exist an orthogonal projector $P \neq I$ such that

$$(39) \qquad PA_x = A_x P, \ PB_y = B_y P \text{ and } P\rho P = \rho.$$

Then, for any C-system of vectors $\{u_x\}_x, \{v_y\}_y$ we have that

$$(40) \qquad \{A_x, A_{x'}\} = 2\langle u_x, u_{x'}\rangle I_d, \ \forall x, x' \text{ and } \{B_y, B_{y'}\} = 2\langle v_y, v_{y'}\rangle I_d, \ \forall y, y',$$

where $\{A, B\} := AB + BA$ is the anticommutator *of A and B.*

Using Theorem 40 we are now ready to give a proof for Theorem 25.

THEOREM. *Let $C = (c_{xy}) \in \text{ext}(\text{Cor}(n,m))$ and consider a family of Hermitian operators $\{M_x\}_x, \{N_y\}_y \subseteq \mathcal{H}^d$ with eigenvalues in $[-1, 1]$ and a quantum state $\rho \in \mathcal{H}_+^{d^2}$ satisfying $c_{xy} = \text{Tr}((M_x \otimes N_y)\rho)$, for all x, y. Then we have that*

$$d \geq \sqrt{2}^{\lfloor \text{rank}(C)/2 \rfloor}.$$

Proof. For all x set $A_x := M_x \otimes I_d \in \mathcal{H}^{d^2}$ and for all y set $B_y := I_d \otimes N_y \in \mathcal{H}^{d^2}$. Note that conditions $(i) - (iv)$ of Theorem 40 are satisfied. Furthermore, if there exists an orthogonal projector $P \neq I$ satisfying (39), by restricting on the support of the matrices $\{PA_x P\}_x, \{PB_y P\}_y$ and $P\rho P$ we get a new family of operators that satisfy conditions $(i) - (iv)$ from Theorem 40 that have smaller size. This process can be repeated to obtain matrices satisfying conditions $(i) - (v)$ whose size is at most d^2.

By Lemma 39 there exists a C-system of vectors $\{u_x, v_y\}_{x,y}$ satisfying

$$\text{span}(\{u_x\}_x) = \text{span}(\{v_y\}_y) = \mathbb{R}^{\tau_C}.$$

Furthermore, by Theorem 40 we have

$$(41) \qquad \{A_x, A_{x'}\} = 2\langle u_x, u_{x'}\rangle I_{d^2}, \ \forall x, x' \text{ and } \{B_y, B_{y'}\} = 2\langle v_y, v_{y'}\rangle I_{d^2}, \ \forall y, y'.$$

To ease notation set $\tau := \tau_C$ and without loss of generality assume that $\{u_x\}_{x=1}^{\tau}$ is a basis for $\mathbb{R}^{\tau}$. Set $f : \mathbb{R}^{\tau} \to \mathcal{H}^{d^2}$ where $f(u_x) := A_x$, for $1 \leq x \leq \tau$ and extend linearly, i.e., $f(\lambda) = \sum_{x=1}^{\tau} \lambda_x A_x$, for all $\lambda \in \mathbb{R}^{\tau}$, where $\lambda = (\lambda_x)$ are the coordinates with respect to the $\{u_x\}_{x=1}^{\tau}$ basis. Using (41) it follows that for $\lambda, \mu \in \mathbb{R}^{\tau}$ we have

$$(42) \qquad \{f(\lambda), f(\mu)\} = 2\lambda^{\mathsf{T}}\text{Gram}(\{u_x\}_{x=1}^{\tau})\mu \cdot I_{d^2}.$$

Define the bilinear form $\beta : \mathbb{R}^{\tau} \times \mathbb{R}^{\tau} \to \mathbb{R}$ by

$$\beta(\lambda, \mu) = 2\lambda^{\mathsf{T}}\text{Gram}(\{u_x\}_{x=1}^{\tau})\mu.$$

Note that β is symmetric and furthermore, since $\text{Gram}(\{u_x\}_{x=1}^{\tau})$ is full-rank, β is also nondegenerate. By (42), the map f can be extended to a representation of the Clifford algebra $\text{Cl}(\mathbb{R}^{\tau}, \beta)$ (cf. Remark A.1). Any representation of $\text{Cl}(\mathbb{R}^{\tau}, \beta)$ can be decomposed as a direct sum of irreducible ones, which by Theorem 37 have size at least $2^{\lfloor \tau/2 \rfloor}$. This implies that $d^2 \geq 2^{\lfloor \tau/2 \rfloor}$ and thus $d \geq \sqrt{2}^{\lfloor \tau/2 \rfloor}$. Lastly, by Lemma 39 we have that $\tau \geq \text{rank}(C)$ and the proof is concluded. $\square$