Polytope compatibility-From quantum measurements to magic squares

Andreas Bluhm, Ion Nechita, Simon Schmidt

POLYTOPE COMPATIBILITY

— FROM QUANTUM MEASUREMENTS TO MAGIC SQUARES —

ANDREAS BLUHM, ION NECHITA, AND SIMON SCHMIDT

ABSTRACT. Several central problems in quantum information theory (such as measurement compatibility and quantum steering) can be rephrased as membership in the minimal matrix convex set corresponding to special polytopes (such as the hypercube or its dual). In this article, we generalize this idea and introduce the notion of polytope compatibility, by considering arbitrary polytopes. We find that semiclassical magic squares correspond to Birkhoff polytope compatibility. In general, we prove that polytope compatibility is in one-to-one correspondence with measurement compatibility, when the measurements have some elements in common and the post-processing of the joint measurement is restricted. Finally, we consider how much tuples operators with appropriate joint numerical range have to be scaled in the worst case in order to become polytope compatible and give both analytical sufficient conditions and numerical ones based on linear programming.
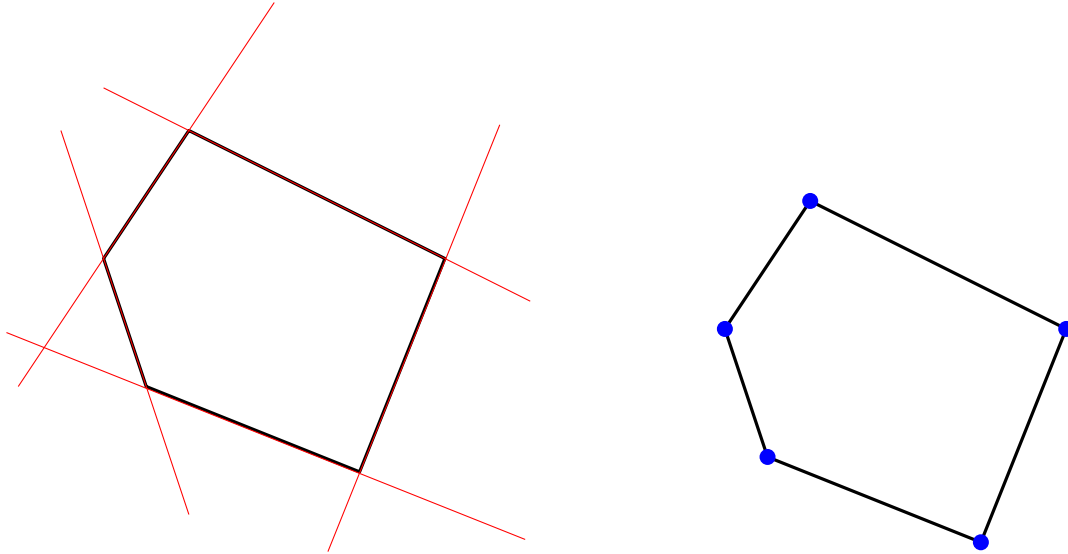
CONTENTS

FIGURE 1. A polytope $\mathcal{P}$ described by its facets as an intersection of half-spaces (left) and as the convex hull of its extreme points (right).

1. INTRODUCTION

A polytope $\mathcal{P}$ containing the origin can be characterized in two different but equivalent ways:

- by its facets, as an intersection of half-spaces (the "H" representation):

$$(1) \qquad \mathcal{P} = \bigcap_{i=1}^{f} \{x \in \mathbb{R}^g \,:\, \langle x, h_i \rangle \leq 1\},$$

- by its extreme points, as a convex hull (the "V" representation):

$$(2) \qquad \mathcal{P} = \mathrm{conv}\{v_i\}_{i=1}^{k}.$$

These two different points of view are graphically represented in Figure 1.

If we want to allow the elements of the polytope to be tuples of matrices instead of tuples of scalars, these two conditions give rise to two different and *inequivalent* (in general) such matricization, which are both special cases of so-called *matrix convex sets*

- the facet description from Eq. (1) generalizes to the set

$$(3) \qquad \mathcal{P}_{\max}(d) := \{(A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g \,:\, \langle A, h_i \otimes \rho \rangle \leq 1 \quad \forall i \in [f], \forall \rho \in \mathcal{M}_d^{1,+}(\mathbb{C})\},$$

- and the extreme points description from Eq. (2) generalizes to the set

$$(4) \qquad \mathcal{P}_{\min}(d) := \left\{ (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g \,:\, \exists\, \mathrm{POVM}\, C \,\text{s.t.}\, A_x = \sum_{i=1}^{k} v_i(x) C_i, \forall x \in [g] \right\}.$$

We refer the reader to Section 2 for the definition of the set of density matrices $\mathcal{M}_d^{1,+}(\mathbb{C})$, which describe quantum states, and that of a positive operator valued measure (POVM), which describe quantum measurements. As our notation suggests, $\mathcal{P}_{\min}$ is the smallest matrix convex set arising from $\mathcal{P}$ and $\mathcal{P}_{\max}$ is the largest.

The appearance of density matrices and POVMs in the definition of the sets $\mathcal{P}_{\min}, \mathcal{P}_{\max}$ suggest that there might be a link between such matrix convex sets and quantum information theory. Indeed, in the articles [BN18, BN20, BN22b], some of the present authors realized that if one takes

$\mathcal{P}$ to be the hypercube $[-1, 1]^g$, then the following correspondence holds:

$$(2E_1 - I, \ldots, 2E_g - I) \in ([-1, 1]^g)_{\max} \iff \{E_i, I - E_i\} \text{ POVMs } \forall i.$$

What about $([-1, 1]^g)_{\min}$? One of the defining properties that distinguish quantum mechanics from our everyday experience based on classical mechanics is the existence of *incompatible measurements*, i.e., measurements that cannot be performed at the same time [Hei27, Boh28]. Such measurements are indispensable for detecting quantum non-locality [Fin82] and can therefore be seen as a resource for many quantum information theoretic tasks similar to entanglement [BCP+14, HKR15]. For the measurements that are compatible, a joint measurement exists such that their outcomes can be recovered by classically post-processing the outcomes of the joint measurement. It turns out that membership in $([-1, 1]^g)_{\min}$ is related to measurement compatibility

$$(2E_1 - I, \ldots, 2E_g - I) \in ([-1, 1]^g)_{\min} \iff \{E_i, I - E_i\} \text{ compatible POVMs } \forall i.$$

The reformulation of measurement compatibility as minimal and matrix convex sets has been instrumental in finding new bounds on the maximal amount of incompatibility available in different situations.

The success of the study of minimal and maximal matrix convex sets for the hypercube suggests the natural question: *What tasks in quantum information theory can be formulated as membership in $\mathcal{P}_{\min}$, $\mathcal{P}_{\max}$ for polytopes $\mathcal{P}$?* This is the task this paper sets out to solve.

Motivated by the example of measurement compatibility, we define a notion of *polytope operators* and *polytope compatibility*. A tuple of matrices is a $\mathcal{P}$-operator if it is in $\mathcal{P}_{\max}$ and it is $\mathcal{P}$-compatible if it is in $\mathcal{P}_{\min}$. We study equivalent formulations and implications of polytope compatibility in Section 3. In particular, we characterize the elements which are $\mathcal{P}$-compatible if and only if they are $\mathcal{P}$-operators in Theorem 3.10. An informal version is the following:

Theorem. *Let A be a g-tuple of self-adjoint operators. Then, A is $\mathcal{P}$-compatible for* all *polytopes $\mathcal{P}$ such that A are $\mathcal{P}$-operators if and only if the operators A admit a pairwise commuting dilation N with essentially the same numerical range.*

Then we discuss some examples before we treat the general case. In Section 4, we review in detail which implications arise for measurement compatibility from our results on polytope compatibility.

In Section 5, we show that another well-known problem from quantum information theory can be formulated as polytope compatibility, namely the study of *quantum magic squares*. An $N \times N$ block matrix $(A_{ij})_{i,j \in [N]}$ with d-dimensional matrices A_{ij} is a quantum magic square if both its rows $\{A_{ij}\}_{j \in [N]}$ and columns $\{A_{ij}\}_{i \in [N]}$ form POVMs. This can be expressed with the help of the Birkhoff polytope (the set of bistochastic matrices), projected onto its supporting affine subspace. Calling this polytope $\mathcal{B}_N$, we arrive at the following equivalence:

$$A \in (\mathcal{B}_N)_{\max} \iff A \text{ is a quantum magic square.}$$

A quantum magic square is especially simple if it has a hidden structure in terms of a tensor product of permutation matrices and a POVM. In [DlCDN20], such a quantum magic square is called *semiclassical*, whereas [GB19] uses the term doubly normalised tensor of positive semidefinite operators. The interest in such objects comes from the Birkhoff-von Neumann theorem, which states that the bistochastic matrices are the convex hull of the permutation matrices. The semiclassical magic squares can be seen as a matricization of this idea. It is known that not all quantum magic squares are semiclassical, but we can characterize the ones that are:

$$A \in (\mathcal{B}_N)_{\min} \iff A \text{ is a semiclassical quantum magic square.}$$

One might be tempted to conjecture that semiclassical quantum magic squares are the ones in which the row and column POVMs are compatible. However, we give an explicit example in Section 5 of a quantum magic square with compatible POVMs which is *not* semiclassical. Using one of our reformulations of $\mathcal{P}$-compatibility as factorization of an associated map through a simplex,

Proposition 3.8, we recover being a semiclassical quantum magic square does not only require the POVMs to be compatible, but also that the post-processing via which they arise from a joint measurement is symmetric (previously observed in [GB19]).

In Section 6 we find that polytope compatibility corresponds in general to the compatibility of POVMs with common elements under restricted post-processing. Any collection of POVMs which share elements can be represented as a hypergraph, where each POVM element is a vertex and which POVM elements belong to the same POVM is represented by hyperedges. Such hypergraphs are in one-to-one correspondence with polytopes $\mathcal{P}$ having vertices with rational coordinates. Being a $\mathcal{P}$-operator then corresponds to being a POVM with the desired common elements. These POVMs are $\mathcal{P}$-compatible if and only if the POVMs are compatible and have a joint measurement from which they arise under restricted post-processing. We illustrate this in Proposition 6.8 where we consider two POVMs with a common element A, such that the POVMs become

$$(5) \qquad (A, B, I - A - B) \qquad (A, C, I - A - C).$$

The polytope to which this compatibility structure corresponds is a pyramid with square basis. (A, B, C) is in the minimal matrix convex set corresponding to the pyramid if and only if the two POVMs above are compatible and have a joint POVM Q with five elements from which they arise as

$$A = Q_1$$
$$B = Q_4 + Q_5$$
$$C = Q_3 + Q_5.$$

We conclude the section with an explicit counterexample that not all compatible POVMs as in Eq. (5) have a joint POVM of the form above, which shows that the restricted post-processing is indeed necessary.

Finally, we consider in Section 7 the question of how much we need to shrink level-d of the maximal matrix convex set for a polytope to fit it into the minimal matrix convex set for the same polytope. This leads us to the set of *inclusion constants*:

$$\Delta_{\mathcal{P}}(d) := \left\{ s \in \mathbb{R}^g : s \cdot \mathcal{P}_{\max}(d) \subseteq \mathcal{P}_{\min}(d) \right\}.$$

It has been shown in [BN18] that

$$s \in \Delta_{[-1,1]^g}(d) \cap [0,1]^g \iff \{\widetilde{E}_i, I - \widetilde{E}_i\} \text{ compatible } \forall d - \text{dimensional POVMs } \{E_i, I - E_i\}.$$

Here, the $\widetilde{E}_i$ are the noisy versions of the E_i under added white noise according to s, i.e.,

$$\widetilde{E}_i := s_i E_i + (1 - s_i)\frac{I_d}{2}.$$

The smaller s_i is, the noisier the measurement $\{\widetilde{E}_i, I - \widetilde{E}_i\}$. Thus, the set $\Delta_{[-1,1]^g}(d) \cap [0,1]^g$ is quantifying the maximal incompatibility in d-dimensional dichotomic measurements.

We give several sufficient conditions in this article for s to be in $\Delta_{\mathcal{P}}(d)$. In Section 7.1, we show how to obtain inclusion constant from comparing the polytope of interest to another polytope for which the set of inclusion constants is known. In Section 7.2, we prove sufficient conditions based on the symmetrization of the polytope inspired by previous work [BN20]. In particular, Proposition 7.9 generalizes findings from [DDOSS17, HKMS19] and could be of independent interest. In particular, the proposition implies for polytopes such that $\mathcal{P} = -\mathcal{P}$

$$\frac{1}{2d - 1} \in \Delta_{\mathcal{P}}(d) \text{ for } d \text{ even}, \qquad \frac{1}{2d + 1} \in \Delta_{\mathcal{P}}(d) \text{ for } d \text{ odd}.$$

We put forward a linear program in Section 7.3 and show that we can use it to compute vectors $s \in \Delta_{\mathcal{P}}(d)$ for *all* d efficiently. Informally, the linear program is the following feasibility problem (see Theorem 7.14):

Theorem. *Given $s \in \mathbb{R}^g$, if there exists an entrywise non-negative matrix T such that*

$$\mathrm{diag}(s_1, s_2, \ldots, s_g, 1) = \hat{V}T\hat{H},$$

then $s \in \Delta_{\mathcal{P}}(d)$, for all $d \geq 1$. Here, $\hat{V}$ is a matrix containing the vertices and $\hat{H}$ a matrix containing the facets of the polytope $\mathcal{P}$.

We showcase the usefulness of our methods by applying them to the Birkhoff polytope and the pyramid corresponding to POVMs with shared element as in Eq. (5).

Thus, in summary, this article generalized the correspondence between measurement incompatibility and the minimal and maximal matrix convex sets of the hypercube. We find that polytope compatibility is in one-to-one correspondence with measurement compatibility with common elements and restricted post-processing. As an example, we find that being a semiclassical magic square corresponds to being Birkhoff-polytope compatible.

2. Preliminaries

2.1. **Notation.** For simplicity, we use the notation $[n] := \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$. By $\mathcal{S}_n$, we denote the group of permutations on n elements. For a convex set $\mathcal{P} \subset \mathbb{R}^n$, we will write $\mathcal{P}^\circ := \{h \in \mathbb{R}^n : \langle h, x \rangle \leq 1 \ \forall x \in \mathcal{P}\}$ for its polar dual.

For complex $n \times n$ matrices, we write $\mathcal{M}_n(\mathbb{C})$. If we restrict to Hermitian matrices, we will write $\mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})$. To indicate that such a matrix A is positive semidefinite, we will write $A \geq 0$, and we will denote its trace by $\mathrm{Tr}[A]$. If we restrict to density matrices, we will write $\mathcal{M}_n^{1,+}(\mathbb{C}) := \{\rho \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C}) : \rho \geq 0, \ \mathrm{Tr}[\rho] = 1\}$. We write I_n for the identity matrix, but we will sometimes drop subscript. We will write $\mathcal{B}(\mathcal{H})$ for the bounded operators on a Hilbert space $\mathcal{H}$.

Given a g-tuple $A = (A_1, \ldots, A_g)$ of self-adjoint operators in $\mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})$, their *joint numerical range* is defined by

$$(6) \qquad \mathcal{W}(A) := \{(\langle x, A_1 x\rangle, \ldots, \langle x, A_g x\rangle) \ : \ x \in \mathbb{C}^d, \|x\|_2 = 1\} \subseteq \mathbb{R}^g,$$

where $\|\cdot\|_2$ is the Euclidean norm. For $g = 1$, one recovers the usual numerical range of a matrix, which is a convex set by the celebrated Toeplitz–Hausdorff theorem; however, for $g \geq 4$, the joint numerical range is, in general, not convex [LP00, GJK04]. The convex hull of the joint numerical range is obtained by going from unit vectors (i.e. pure quantum states) to density matrices (i.e. mixed quantum states):

$$\mathrm{conv}\,\mathcal{W}(A) = \{(\mathrm{Tr}[A_1 \rho], \ldots \mathrm{Tr}[A_g \rho]) \ : \ \rho \in \mathcal{M}_d^{1,+}(\mathbb{C})\}.$$

2.2. **Convexity.** A well known way to combine two polytopes $\mathcal{P}_1 \subseteq \mathbb{R}^{k_1}$, $\mathcal{P}_2 \subseteq \mathbb{R}^{k_2}$ into another one is taking their Cartesian product:

$$\mathcal{P}_1 \times \mathcal{P}_2 := \left\{ (x, y) \in \mathbb{R}^{k_1 + k_2} : x \in \mathcal{P}_1, y \in \mathcal{P}_2 \right\}.$$

Another one is the direct sum:

$$\mathcal{P}_1 \oplus \mathcal{P}_2 := \mathrm{conv}\left(\left\{ (x, 0) \in \mathbb{R}^{k_1 + k_2} : x \in \mathcal{P}_1 \right\} \cup \left\{ (0, y) \in \mathbb{R}^{k_1 + k_2} : y \in \mathcal{P}_2 \right\}\right),$$

which is again a polytope. If both polytopes contain 0, it holds that [Bre97, Lemma 2.4]

$$\mathcal{P}_1 \oplus \mathcal{P}_2 = (\mathcal{P}_1^\circ \times \mathcal{P}_2^\circ)^\circ.$$

See also [BN20, Lemma 3.6] for a short proof. Moreover, if both polytopes contain 0, clearly

$$(7) \qquad \mathcal{P}_1 \oplus \mathcal{P}_2 \subseteq \mathcal{P}_1 \times \mathcal{P}_2,$$

as $0 \in \mathcal{P}_i$ for $i \in [2]$ implies $(p_1, 0), (0, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2$ for all $p_i \in \mathcal{P}_i$.

2.3. **Maximal and minimal matrix convex sets.** We will now review some basic results about matrix convex sets, with a focus on minimal and maximal matrix convex sets. For more details, we refer the reader to [DDOSS17].

Definition 2.1. *Let* $g \in \mathbb{N}$. *Moreover, let* $\mathcal{F}(n) \subseteq \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})^g$ *for all* $n \in \mathbb{N}$. *Then, we call* $\mathcal{F} = \bigsqcup_{n \in \mathbb{N}} \mathcal{F}(n)$ *a* free set. *Moreover,* $\mathcal{F}$ *is a* matrix convex set *if it satisfies the following two properties for any* $m, n \in \mathbb{N}$:

(1) *If* $X = (X_1, \ldots, X_g) \in \mathcal{F}(m)$, $Y = (Y_1, \ldots, Y_g) \in \mathcal{F}(n)$, *then* $X \oplus Y := (X_1 \oplus Y_1, \ldots, X_g \oplus Y_g) \in \mathcal{F}(m + n)$.
(2) *If* $X = (X_1, \ldots, X_g) \in \mathcal{F}(m)$ *and* $\Psi : \mathcal{M}_m(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ *is a unital completely positive (UCP) map, then* $(\Psi(X_1), \ldots, \Psi(X_g)) \in \mathcal{F}(n)$.

That is, a matrix convex set is a free set that is closed under direct sums and UCP maps.

In particular, it follows from the definition that all sets $\mathcal{F}(n)$ are convex. A matrix convex set $\mathcal{F}$ is open/closed/bounded/compact if all levels $\mathcal{F}(n)$ have this property.

Let $\mathcal{C} \subseteq \mathbb{R}^g$ be a closed convex set. Fixing $\mathcal{F}(1) = \mathcal{C}$, in most cases there are infinitely many matrix convex sets $\mathcal{F}$ with the same $\mathcal{F}(1)$. However, we can find a maximal and a minimal matrix convex set which have $\mathcal{C}$ as their first level. First, we consider the maximal matrix convex set for $\mathcal{C}$ [DDOSS17, Definition 4.1]:

$$\mathcal{C}_{\max}(n) :=$$
$$\left\{ X \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})^g : \sum_{i=1}^{g} c_i X_i \leq \alpha I_n, \quad \forall c \in \mathbb{R}^g, \forall \alpha \in \mathbb{R} \text{ s.t. } \mathcal{C} \subseteq \{ x \in \mathbb{R}^g : \langle c, x \rangle \leq \alpha \} \right\}.$$

Note that if $\mathcal{C}$ is a polyhedron, only finitely many hyperplanes need to be considered.

The minimal matrix convex set associated with $\mathcal{C}$ is defined as [PSS18, Eq. (1.4)]:

$$(8) \qquad \mathcal{C}_{\min}(n) := \left\{ X \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})^g : X = \sum_j z_j \otimes Q_j, z_j \in \mathcal{C} \ \forall j, Q_j \geq 0 \ \forall j, \sum_j Q_j = I_n \right\}.$$

Note that if $\mathcal{C}$ is a polytope, i.e. it has finitely many extreme points, the number of terms in the decomposition above can be taken to be the number of extreme points of $\mathcal{C}$.

An equivalent definition of the minimal matrix convex set is the one used in [DDOSS17] as

$$(9) \quad \mathcal{C}_{\min}(n) := \{X \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})^g : \exists \text{ pairwise-commuting normal dilation } N \text{ of } X \text{ s.t. } \sigma(N) \subseteq \mathcal{C}\},$$

where $\sigma(N)$ is the joint spectrum of the pairwise-commuting normal dilation N. We recall that $N \in \mathcal{B}(\mathcal{H})^g$ is a dilation of $X \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})^g$ if there exists an isometry $V : \mathbb{C}^n \hookrightarrow \mathcal{H}$ such that $X_i = V^* N_i V$ for all $i \in [g]$.

Remark 2.2. *To go from Eq.* (9) *to Eq.* (8), *we can use Naimark's dilation theorem. In order to go from Eq.* (8) *to Eq.* (9), *we can use the construction used in the proof of Theorem 7.1 in* [DDOSS17]. *Theorem 7.1 in* [DDOSS17] *also implies that we can restrict the dilation* N *to be a tuple of self-adjoint* nm-*dimensional matrices and the POVM in Eq.* (9) *to have* m *outcomes, where* $m = 2n^2(g + 1) + 1$.

We will also use the notion of *inclusion constants*, i.e., constants for which the inclusion

$$s \cdot \mathcal{C}_{\max} \subseteq \mathcal{C}_{\min}$$

holds. Allowing for different scalings in each direction, the *(asymmetrically) scaled* matrix convex set is

$$s \cdot \mathcal{C}_{\max} := \{ (s_1 X_1, \ldots, s_g X_g) : X \in \mathcal{C}_{\max} \}.$$

Definition 2.3. *Let d, $g \in \mathbb{N}$ and $\mathcal{C} \subset \mathbb{R}^g$. The* inclusion set *is defined as*

$$\Delta_{\mathcal{C}}(d) := \{\, s \in \mathbb{R}^g : s \cdot \mathcal{C}_{\max}(d) \subseteq \mathcal{C}_{\min}(d) \,\}.$$

Note that $\Delta_{\mathcal{C}}(d)$ is a convex set, because both $\mathcal{C}_{\min}(d)$ and $\mathcal{C}_{\max}(d)$ are.

Finally, we will show that if $\mathcal{C}$ is compact, then its corresponding minimal matrix convex set is also compact. This has already been pointed out in [PP21] and we include the proof here for convenience.

Lemma 2.4. *Let $g \in \mathbb{N}$ and let $\mathcal{C} \subset \mathbb{R}^g$ be a compact convex set. Then, $\mathcal{C}_{\min}$ is compact (i.e., $\mathcal{C}_{\min}(n)$ is compact for all $n \in \mathbb{N}$).*

Proof. For $X \in \mathcal{C}_{\min}(n)$, Remark 2.2 implies that we can write

$$X = \sum_{i=1}^{2n^2(g+1)+1} z_i \otimes Q_i$$

for a POVM Q and $z_i \in \mathcal{C}$ for all $i \in [2n^2(g+1)+1]$. As $\mathcal{C}$ and the set of POVMs with a fixed number of outcomes are both compact, $\mathcal{C}_{\min}(n)$ is compact as the image of a compact set under a continuous function. $\square$

2.4. **Incompatible measurements in quantum mechanics.** As this will be our guiding example in this work, we will give a short introduction to measurement incompatibility in quantum mechanics. For background concerning the mathematics of quantum mechanics, see [HZ11, Wat18]. A quantum mechanical measurement with k outcomes on a quantum system of dimension d is described by a collection of positive operators $\{\, E_j \,\}_{j \in [k]} \subset \mathcal{M}_d^{\mathrm{sa}}$, $E_j \geq 0$ for all $j \in [k]$, such that

$$\tag{10} \sum_{j=1}^{k} E_j = I_d.$$

The set $\{\, E_j \,\}_{j \in [k]}$ is called a *positive operator-valued measure (POVM)* and its elements E_j are referred to as *effects*. If the condition Eq. (10) is replaced by the sub-normalization

$$\sum_{j=1}^{k} E_j \leq I_d,$$

then $\{\, E_j \,\}_{j \in [k]}$ is called a *sub-POVM*.

An important characteristic distinguishing quantum mechanics from classical mechanics is that measurements can be *incompatible* (see [HMZ16] for an introduction). A collection of POVMs is compatible if they arise as marginals from a joint POVM:

Definition 2.5 (Compatible POVMs). *Let $\left\{\, E_j^{(i)} \,\right\}_{j \in [k_i]}$ be a collection of d-dimensional POVMs, where $k_i \in \mathbb{N}$ for all $i \in [g]$, d, $g \in \mathbb{N}$. The POVMs are* compatible *if there is a d-dimensional joint POVM $\left\{\, R_{j_1,\dots,j_g} \,\right\}$ with $j_i \in [k_i]$ such that for all $u \in [g]$ and $v \in [k_u]$,*

$$E_v^{(u)} = \sum_{\substack{j_i \in [k_i] \\ i \in [g] \setminus \{\, u \,\}}} R_{j_1,\dots,j_{u-1},v,j_{u+1},\dots j_g}.$$

Not all measurements in quantum mechanics are compatible. For projective measurements (POVMs in which all effects are orthogonal projections), compatibility is equivalent to their effects pairwise commuting. There is an equivalent definition of joint measurability [HMZ16, Eq. 16], formulated in terms of classical post-processing, which we will also use in this work. Measurements are compatible if and only if there is a joint measurement from which their outcomes can be obtained with the help of classical randomness, i.e., *classical post-processing*.

Lemma 2.6. *Let $E^{(i)} \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^{k_i}$, $i \in [g]$, be a collection of POVMs, where $k_i \in \mathbb{N}$ for all $i \in [g]$, d, $g \in \mathbb{N}$. These POVMs are compatible if and only if there is some $m \in \mathbb{N}$ and a POVM $M \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^m$ such that*

$$E_j^{(i)} = \sum_{\lambda=1}^{m} p_\lambda(j|i) M_\lambda$$

for all $j \in [k_i]$, $i \in [g]$ and some conditional probabilities $p_\lambda(j|i)$.

For a collection of dichotomic measurements $\{E_i, I - E_i\}$, $i \in [g]$, we will for simplicity say that the effects $\{E_i\}_{i \in [g]}$ are compatible, since they completely determine the corresponding measurements due to normalization.

2.5. **General probabilistic theories.** In this section, we briefly introduce the formalism to describe a class of generalization of quantum mechanics, the *general probabilistic theories* (GPTs). For more background, see [Lam18].

Any GPT corresponds to a triple $(V, V^+, \mathbb{1})$, where V is a vector space with a proper cone V^+ and $\mathbb{1}$ is an order unit in the dual cone $A^+ = (V^+)^* \subset V^* = A$. We assume here that V is finite dimensional. The set of states of the system is identified as the subset

$$K := \{v \in V^+, \ \langle \mathbb{1}, v \rangle = 1\}.$$

It holds that K is compact and convex and is a base of the cone V^+. It is also possible to start with any compact convex set K as a state space and construct the corresponding GPT $(V(K), V(K)^+, \mathbb{1}_K)$ from there. Indeed, let K be such a set and let $A(K)$ be the set of affine functions $K \to \mathbb{R}$. Then $A(K)$ is a finite dimensional vector space and the subset of affine functions which are positive on K, $A(K)^+$, is a proper cone in $A(K)$. Let $\mathbb{1}_K$ be the constant function, then $\mathbb{1}_K$ is an order unit. We put $V(K) = A(K)^*$, $V(K)^+ = (A(K)^+)^*$. Then $V(K)^+$ is a proper cone in $V(K)$ and K is affinely isomorphic to the base of V^+, determined by $\mathbb{1}_K$.

Example 2.7. *Any* classical system *is described by the triple* $\mathrm{CM}_d := (\mathbb{R}^d, \mathbb{R}_+^d, 1_d)$, $d \in \mathbb{N}$, *where* $\mathbb{R}_+^d$ *denotes the set of elements with non-negative coordinates and* $1_d = (1, 1, \ldots, 1) \in \mathbb{R}^d$. *Then* $(\mathbb{R}^d)^* = \mathbb{R}^d$ *with duality given by the standard inner product and the simplicial cone* $\mathbb{R}_+^d$ *is self-dual. The classical state space is the probability simplex*

$$\Delta_d = \left\{ x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \ x_i \geq 0, \ \sum_{i=1}^{d} x_i = 1 \right\} = \{x \in \mathbb{R}_+^d, \ \langle x, 1_d \rangle = 1\}.$$

Example 2.8. Quantum mechanics *corresponds to the triple* $\mathrm{QM}_d := (\mathcal{M}_d^{\mathrm{sa}}(\mathbb{C}), \mathrm{PSD}_d, \mathrm{Tr})$, $d \in \mathbb{N}$, *where* PSD_d *is the cone of $d \times d$ positive semidefinite complex, self-adjoint matrices, and* Tr *is the usual, un-normalized, trace. As in the case of classical systems described above, the* PSD_d *cone is self-dual. The quantum state space is the set of density matrices* $\mathcal{M}_n^{1,+}(\mathbb{C})$.

A *channel* between GPTs $(V_i, V_i^+, \mathbb{1}_i)$ with state spaces K_i, $i \in [2]$, is a linear map $\Phi : V_1 \to V_2$ such that $\Phi(V_1^+) \subseteq V_2^+$ (i.e., the map is a positive map between the corresponding ordered vector spaces) and such that it maps states of one GPT to states of the other: $\Phi(K_1) \subseteq K_2$. Note that $\Phi(K_1) \subseteq K_2$ implies $\Phi(V_1^+) \subseteq V_2^+$, since K_i is a base of V_i^+ and the map Φ is linear.

3. Polytope compatibility

In this section, we introduce the notion of polytope compatibility, or $\mathcal{P}$-compatibility for a fixed polytope $\mathcal{P}$.

3.1. **Equivalent characterizations.** We start with the definition of tuples of matrices being $\mathcal{P}$-operators and $\mathcal{P}$-compatible.

Definition 3.1. *Let d, g, $k \in \mathbb{N}$ and let $\mathcal{P}$ be a polytope with k extreme points $v_1, \ldots, v_k \in \mathbb{R}^g$ such that $0 \in \operatorname{int} \mathcal{P}$. Let $A = (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g \cong \mathbb{R}^g \otimes \mathcal{M}_d^{sa}(\mathbb{C})$ be a g-tuple of Hermitian matrices. We say that*

- *A are $\mathcal{P}$-operators if, for any dual hyperplane $h \in \mathcal{P}^\circ$ and any density matrix $\rho \in \mathcal{M}_d^{1,+}(\mathbb{C})$,*

$$\langle A, h \otimes \rho \rangle \leq 1.$$

- *A are $\mathcal{P}$-compatible if there exists a POVM $C = (C_1, \ldots C_k)$ in $\mathcal{M}_d(\mathbb{C})^k$ such that*

$$A = \sum_{i=1}^k v_i \otimes C_i.$$

We can express these definitions naturally in the language of minimal and maximal matrix convex sets arising from $\mathcal{P}$.

Proposition 3.2. *Let d, g, $k \in \mathbb{N}$ and let $\mathcal{P}$ be a polytope with k extreme points $v_1, \ldots, v_k \in \mathbb{R}^g$ such that $0 \in \operatorname{int} \mathcal{P}$. Let $A = (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g \cong \mathbb{R}^g \otimes \mathcal{M}_d^{sa}(\mathbb{C})$ be a g-tuple of Hermitian matrices. Then, A are $\mathcal{P}$-operators if and only if $A \in \mathcal{P}_{\max}(d)$. Moreover, A are $\mathcal{P}$-compatible if and only if $A \in \mathcal{P}_{\min}(d)$.*

Proof. The proof of the first assertion is straightforward. For the second assertion, we note the following: Let $(C_1', \ldots, C_n')$ be a POVM for $n \in \mathbb{N}$ and let $z_j \in \mathcal{P}$ for all $j \in [n]$. Then, every z_j is a convex combination of extreme points, i.e., $z_j = \sum_{i=1}^k w_i^{(j)} v_i$ for all $j \in [n]$, where $w_i^{(j)} \geq 0$ for all $i \in [k]$ and $\sum_{i=1}^k w_i^{(j)} = 1$ for all $j \in [n]$. Thus, we can always write

$$\sum_{j=1}^n z_j \otimes C_j' = \sum_{j=1}^n \left(\sum_{i=1}^k w_i^{(j)} v_i \right) \otimes C_j' = A = \sum_{i=1}^k v_i \otimes C_i,$$

where $C_i = \sum_{j=1}^n w_i^{(j)} C_j'$. It can readily be verified that $(C_1, \ldots, C_k)$ is also a POVM. $\qquad \square$

Remark 3.3. *In the scalar case $d = 1$, the tuple A being $\mathcal{P}$-operators and A being $\mathcal{P}$-compatible coincide and the two notions are equivalent to $A \in \mathcal{P}$. They correspond, respectively, to the hyperplane (Eq. (1)) and to the extreme point (Eq. (2)) definition of the polytope $\mathcal{P}$. They can therefore be thought of as two different ways of "quantizing" what it means to be a polytope.*

A polytope $\mathcal{P}$ containing 0 in its interior can be characterized (in the hyperplane representation) as the intersection of finitely many halfspaces

$$\mathcal{P} = \bigcap_{j=1}^r \{ x \in \mathbb{R}^g \ : \ \langle h_j, x \rangle \leq 1 \},$$

where the vectors $h_1, \ldots, h_r \in \mathbb{R}^g$ define the facets of $\mathcal{P}$. In this picture, one can state the following nice description of the set $\mathcal{P}_{\max}$.

Proposition 3.4. *Let d, g, $r \in \mathbb{N}$ and let $\mathcal{P}$ be a polytope in $\mathbb{R}^g$, containing 0 in its interior, defined by the facet vectors $h_1, \ldots, h_r \in \mathbb{R}^g$ as above. Then, for all $d \geq 1$,*

$$\mathcal{P}_{\max}(d) = \left\{ A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g \ : \ \forall j \in [r], \ \sum_{x=1}^g h_j(x) A_x \leq I_d \right\}.$$

Proof. The proof follows immediately from the fact that the extreme points of the dual polytope $\mathcal{P}^\circ$ are a subset of the h_j's. $\qquad \square$

There are ways to rewrite the condition that $A \in \mathcal{P}_{\min}(d)$ which will be useful later on.

Remark 3.5. *If $A \in \mathcal{P}_{\min}(d)$ for a polytope $\mathcal{P}$ with $0 \in \operatorname{int} \mathcal{P}$, then for all $\rho \in \mathcal{M}_d^{1,+}(\mathbb{C})$*

$$(\langle A_x, \rho \rangle)_{x \in [g]} = V \cdot (\langle C_i, \rho \rangle)_{i \in [k]} \in \mathcal{P},$$

where $(C_1, \ldots, C_k)$ is a POVM and $V \in \mathcal{M}_{g \times k}(\mathbb{R})$ is the matrix having the extreme points of $\mathcal{P}$ as columns:

$$\forall x \in [g], \forall i \in [k], \qquad V_{x,i} = v_i(x).$$

One can relax the POVM condition in the definition of P-compatibility to C being a sub-POVM.

Proposition 3.6. *Let d, g, $k \in \mathbb{N}$ and let $\mathcal{P}$ be a polytope with k extreme points $v_1, \ldots, v_k \in \mathbb{R}^g$ such that $0 \in \operatorname{int} \mathcal{P}$. Let $A = (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ be a g-tuple of Hermitian matrices. Then, A are $\mathcal{P}$-compatible if and only if there exists a sub-POVM $C = (C_1, \ldots C_k)$ in $\mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^k$ such that*

$$A = \sum_{i=1}^k v_i \otimes C_i.$$

Proof. Assume $A = \sum_i v_i \otimes C_i$ for a sub-POVM C and denote $C_0 := I_d - \sum_i C_i \geq 0$. Since $0 \in P$, there is a probability vector $\pi \in \mathbb{R}^k$ such that

$$\sum_{i=1}^k \pi_i v_i = 0.$$

Write

$$A = \sum_{i=1}^k v_i \otimes C_i + \left(\sum_{i=1}^k \pi_i v_i \right) \otimes C_0 = \sum_i v_i \otimes C_i',$$

where $C_i' := C_i + \pi_i C_0$ for all $i \in [k]$ forms a POVM. $\qquad\qquad\square$

Proposition 3.7. *Let d, $g \in \mathbb{N}$. A tuple $A = (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ consists of $\mathcal{P}$-operators for a polytope $\mathcal{P}$ with $0 \in \operatorname{int} \mathcal{P}$ if and only if for all density matrices $\rho \in \mathcal{M}_d^{1,+}(\mathbb{C})$,*

$$(\langle A_x, \rho \rangle)_{x \in [g]} \in \mathcal{P}.$$

In other words, A are $\mathcal{P}$-operators if and only if the joint numerical range $\mathcal{W}(A)$ of A is contained in $\mathcal{P}$:

$$\mathcal{W}(A) \subseteq \mathcal{P} \iff \operatorname{conv} \mathcal{W}(A) = \left\{ (\langle A_x, \rho \rangle)_{x \in [g]} \right\}_{\rho \in \mathcal{M}_d^{1,+}(\mathbb{C})} \subseteq \mathcal{P}.$$

Proof. This follows from the definition as we can write

$$\langle A, h \otimes \rho \rangle = \sum_{i=1}^g h_i \operatorname{Tr}[A_i \rho]$$

and use the bipolar theorem [Bar02, Theorem IV.1.2] to conclude that $\mathcal{P}^{\circ\circ} = \mathcal{P}$. $\qquad\square$

Another way to understand polytope compatibility is as factorization of positive maps through ordered vector spaces.

Proposition 3.8. *Let d, g, $k \in \mathbb{N}$ and let $\mathcal{P}$ be a polytope with k extreme points $v_1, \ldots, v_k \in \mathbb{R}^g$ such that $0 \in \operatorname{int} \mathcal{P}$. Let $A = (A_1, \ldots, A_g) \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ be a g-tuple of Hermitian matrices. Let us consider the map $\mathcal{A} : \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C}) \to \mathbb{R}^g$,*

$$\mathcal{A}(X) = (\operatorname{Tr}[A_1 X], \ldots, \operatorname{Tr}[A_g X]).$$

Then,

(1) *A are $\mathcal{P}$-operators if and only if $\mathcal{A}$ is a channel between $(\mathcal{M}_d^{\mathrm{sa}}, \mathrm{PSD}_d, \operatorname{Tr})$ and $(V(\mathcal{P}), V(\mathcal{P})^+, \mathbb{1}_\mathcal{P})$.*
(2) *A are $\mathcal{P}$-compatible if and only if in addition $\mathcal{A}$ factors through Δ_k.*

Proof. For the first point, it is clear that if A are $\mathcal{P}$-operators, then $\mathcal{A}$ is a channel of the required form. Conversely, if $\mathcal{A}$ is a channel of the form in the assertion, then in particular $\mathcal{W}(A) \subseteq \mathcal{P}$. The first assertion then follows from Proposition 3.7.

For the second point, if the A are $\mathcal{P}$-compatible, the map $\mathcal{Q} : \mathcal{M}_d^{\mathrm{sa}} \to \mathbb{R}^k$,

$$(11) \qquad \mathcal{Q}(X) = (\mathrm{Tr}[Q_1 X], \ldots, \mathrm{Tr}[Q_k X]),$$

sends $\mathcal{M}_d^{1,+}(\mathbb{C})$ to Δ_k. The matrix V as defined in Remark 3.5 then maps Δ_k into $\mathcal{P}$ as required. Conversely, it can be checked that any channel mapping $\mathcal{M}_d^{1,+}(\mathbb{C})$ to Δ_k gives rise to a POVM $(Q_1, \ldots, Q_k)$ as in Eq. (11). Any map $\nu : \Delta_k \to \mathcal{P}$ satisfies $\nu(\delta_i) = z_i \in \mathcal{P}$ for the vertices δ_i of Δ_k. Thus, the factorization implies

$$(\mathrm{Tr}[A_1 \rho], \ldots, \mathrm{Tr}[A_g \rho]) = \sum_{i=1}^{k} z_i \mathrm{Tr}[Q_i \rho] \qquad \forall \rho \in \mathcal{M}_d^{1,+}(\mathbb{C}),$$

hence $A_j = \sum_{i=1}^{k} z_i(j) Q_i$. Therefore, $A \in \mathcal{P}_{\min}$ and the assertion follows from Proposition 3.2. $\square$

3.2. Tuples that are $\mathcal{P}$-compatible if and only if they are $\mathcal{P}$-operators.
In this subsection, we give a characterization of tuples for which $\mathcal{P}$-compatibility is equivalent to being $\mathcal{P}$-operators. We first need the following lemma.

Lemma 3.9. *Let d, $g \in \mathbb{N}$ and let $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ such that $0 \in \mathrm{int}(\mathrm{conv}\,\mathcal{W}(A))$. If A are $\mathcal{P}$-compatible for all polytopes $\mathcal{P}$ such that A are $\mathcal{P}$-operators, then $A \in (\mathrm{conv}(\mathcal{W}(A)))_{\min}$.*

Proof. The assertion is not obvious, because $\mathrm{conv}(\mathcal{W}(A))$ might not be a polytope. It is easy to see that $\mathcal{W}(A)$ is compact as the image of the unit sphere under a continuous map, thus $\mathrm{conv}(\mathcal{W}(A))$ is compact as well [AB06, Theorem 5.35]. Since every convex body can be approximated to arbitrary precision by a polytope, for any $\varepsilon > 0$, there is a polytope $\mathcal{P}^\varepsilon$ such that $\mathrm{conv}(\mathcal{W}(A)) \subseteq \mathcal{P}^\varepsilon$ and such that $\max_{x \in \mathcal{P}^\varepsilon} \min_{y \in \mathrm{conv}(\mathcal{W}(A))} \|x - y\|_2 \leq \varepsilon$ [Bro08, Section 4].

Let us assume that $A \notin (\mathrm{conv}(\mathcal{W}(A)))_{\min}$. Since $\mathrm{conv}(\mathcal{W}(A))$ is compact, Lemma 2.4 implies that $(\mathrm{conv}(\mathcal{W}(A)))_{\min}$ is compact as well. Hence, there is an $\eta > 0$ such that

$$(12) \qquad \min_{B \in (\mathrm{conv}(\mathcal{W}(A)))_{\min}(d)} \|A - B\| = \eta,$$

where we take the norm to be $\|(X_1, \ldots, X_g)\| = \sum_{i=1}^{g} \|X_i\|_\infty$ for later convenience.

Now we will argue that $A \notin \mathcal{P}_{\min}^{\eta/(2g)}$. If the tuple A was in $\mathcal{P}_{\min}^{\eta/(2g)}$, then we could write

$$A = \sum_i p_i \otimes Q_i$$

with $p_i \in \mathcal{P}^{\eta/(2g)}$ and a POVM Q. From the definition of the approximating polytope, for any i, there exists a $z_i \in \mathrm{conv}(\mathcal{W}(A))$ such that $\|z_i - p_i\|_2 \leq \frac{\eta}{2g}$. Defining

$$B = \sum_i z_i \otimes Q_i,$$

it holds that $B \in \mathrm{conv}(\mathcal{W}(A)))_{\min}(d)$ by construction, but also $\|A - B\| \leq \eta/2$, as for all $j \in [g]$,

$$-\frac{\eta}{2g} I \leq -\sum_i |p_i(j) - z_i(j)| Q_i \leq A_j - B_j \leq \sum_i |p_i(j) - z_i(j)| Q_i \leq \frac{\eta}{2g} I.$$

This contradicts Eq. (12), hence $A \notin \mathcal{P}_{\min}^{\eta/(2g)}$.

However, $\mathrm{conv}(\mathcal{W}(A)) \subseteq \mathcal{P}^{\eta/(2g)}$ by construction, which implies that A is a $\mathcal{P}^{\eta/(2g)}$-operator by Proposition 3.7. Using the assumption that A is $\mathcal{P}$-compatible for all polytopes $\mathcal{P}$ such that A are $\mathcal{P}$-operators, it follows that $A \in \mathcal{P}_{\min}^{\eta/(2g)}$, which is a contradiction. Hence, we have to conclude that $A \in (\mathrm{conv}(\mathcal{W}(A)))_{\min}$. $\square$

Now we can give a characterization of the operators for which being $\mathcal{P}$-operators and $\mathcal{P}$-compatible is the same.

Theorem 3.10. *Let d, $g \in \mathbb{N}$ and let $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ such that $0 \in \mathrm{int}(\mathrm{conv}\,\mathcal{W}(A))$. Then, A is $\mathcal{P}$-compatible for all polytopes $\mathcal{P}$ such that they are $\mathcal{P}$-operators if and only if the operators A admit a pairwise commuting dilation N such that $\mathrm{conv}(\mathcal{W}(A)) = \mathcal{W}(N)$.*

Proof. For the converse, we note that $\mathcal{W}(N)$ is a polytope for a commuting tuple N. Thus, by assumption $\mathrm{conv}\,\mathcal{W}(A)$ is a polytope. Proposition 3.7 implies moreover that A are $\mathcal{P}$-operators for a polytope $\mathcal{P}$ if and only if $\mathrm{conv}(\mathcal{W}(A)) \subseteq \mathcal{P}$. Using the equivalent definition of $\mathcal{P}_{\min}$ from [DDOSS17], see Eq. (9), it follows thus that $A \in \mathcal{P}_{\min}$ for any polytope $\mathcal{P}$ such that A are $\mathcal{P}$-operators, because the joint spectrum $\sigma(N) = \mathrm{conv}(\mathcal{W}(A)) \subseteq \mathcal{P}$ for any such polytope. Hence, the A are $\mathcal{P}$-compatible if and only if they are $\mathcal{P}$-operators.

For the remaining direction, it follows from Lemma 3.9 that A being $\mathcal{P}$-compatible for all polytopes $\mathcal{P}$ such that the A are $\mathcal{P}$-operators implies that $A \in (\mathrm{conv}(\mathcal{W}(A)))_{\min}$. Thus,

$$A = \sum_i z_i \otimes Q_i$$

for a POVM Q and $z_i \in \mathrm{conv}(\mathcal{W}(A))$. Naimark's dilation theorem implies that we can find pairwise-commuting Hermitian operators $N_j = \sum_i z_i(j) P_i$, where P is the Naimark dilation of Q. By Remark 2.2, we can assume that the sum is finite and that the N_j are finite dimensional. We can verify that N is a dilation of A and that $\{z_i\}_i = \sigma(N) \subseteq \mathcal{W}(N) \subseteq \mathrm{conv}(\mathcal{W}(A))$, because $\mathrm{conv}(\sigma(N)) = \mathcal{W}(N)$ for pairwise-commuting operators. On the other hand, as the N are dilations of A, we have the reverse inclusion $\mathcal{W}(N) \supseteq \mathcal{W}(A)$. Noting that $\mathcal{W}(N)$ is convex as the tuple is pairwise commuting, it follows that $\mathrm{conv}\,\mathcal{W}(A) = \mathcal{W}(N)$. $\square$

Remark 3.11. *We give an example of A being $\mathcal{P}$-operators, but not $\mathcal{P}$-compatible, see [LPW20, Example 3.2]. Take*

$$A_1 := \mathrm{diag}(1, 1, -1, -1) \oplus \sigma_Z \quad and \quad A_2 := \mathrm{diag}(1, -1, 1, -1) \oplus \sigma_Y.$$

Their joint numerical range is the square $[-1, 1]^2$, while the matrices do not commute. The diagonal part is responsible for the square, while the off-diagonal part (which is non-commutative) has a unit disk as a joint numerical range, which is hidden by the larger square. This shows that having a numerical range which is a polytope is not equivalent to commuting, which indicates that the previous theorem cannot be significantly strengthened. From joint measurability (see Section 4), we know that (A_1, A_2) are $[-1, 1]^2$-operators, but not $[-1, 1]^2$-compatible.

Question 3.12. *To show that Theorem 3.10 cannot be simplified at all, we would like to find a tuple $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ which is not pairwise commuting, but which has a commuting dilation N such that $\mathrm{conv}(\mathcal{W}(A)) = \mathcal{W}(N)$. We leave the construction of such an example as an open question.*

In view of Proposition 3.2, one might ask when $\mathcal{P}_{max}(d) = \mathcal{P}_{min}(d)$ holds. This is known to be the case if and only if $\mathcal{P}$ is a simplex.

Proposition 3.13 ([ALPP21]). *Let g, $d \in \mathbb{N}$ and let $\mathcal{P} \subseteq \mathbb{R}^g$ be a polytope and $d \geq 2$. Then*

$$\mathcal{P}_{\min}(d) = \mathcal{P}_{\max}(d) \iff \mathcal{P} \text{ is a simplex.}$$

Proof. This follows from Corollary 2 of [ALPP21] and Section 7 of [PSS18]. $\square$

3.3. **Additional results.** The notions of being $\mathcal{P}$-operators and $\mathcal{P}$-compatible behave well with respect to inclusion of polytopes.

Proposition 3.14. *Let $\mathcal{P} \subseteq \mathcal{Q}$ be two polytopes with $0 \in \mathrm{int}\,\mathcal{P}$. If A are $\mathcal{P}$-operators (resp. $\mathcal{P}$-compatible), then A are also $\mathcal{Q}$-operators (resp. $\mathcal{Q}$-compatible).*

Proof. The first claim is trivial: if A are $\mathcal{P}$-operators, then by Proposition 3.7

$$\mathcal{W}(A) \subseteq \mathcal{P} \subseteq \mathcal{Q} \implies A \text{ are } \mathcal{Q}\text{-operators.}$$

For the second claim, let $v_1, \ldots, v_k$ (resp. $w_1, \ldots, w_l$) be the extreme points of $\mathcal{P}$ (resp. $\mathcal{Q}$). Since A are $\mathcal{P}$-compatible, there exists a (sub-)POVM C such that

$$A = \sum_{i=1}^{k} v_i \otimes C_i.$$

Since $\mathcal{P} \subseteq \mathcal{Q}$, by convexity there exists a conditional probability kernel $p(\cdot|\cdot)$ such that

$$\forall i \in [k], \qquad v_i = \sum_{j=1}^{l} p(j|i) w_j.$$

It follows that

$$A = \sum_{i=1}^{k} v_i \otimes C_i = \sum_{j=1}^{l} w_j \otimes \left[\sum_{i=1}^{k} p(j|i) C_i \right] = \sum_{j=1}^{l} w_j \otimes D_j,$$

where $D_j := \sum_{i=1}^{k} p(j|i) C_i$ is a (sub-)POVM. $\qquad\square$

Example 3.15. *We give an example of A being $\mathcal{Q}$-compatible, but not $\mathcal{P}$-compatible for $\mathcal{P} \subseteq \mathcal{Q}$. Let $A = (\sigma_X, \sigma_Y, \sigma_Z) \in \mathcal{M}_2^{\mathrm{sa}}(\mathbb{C})^3$ be the triplet of Pauli matrices. Then:*

- *The joint numerical range of A is the unit sphere of $\mathbb{R}^3$.*
- *A are $[-1,1]^3$-operators.*
- *A are not $[-1,1]^3$-compatible, but are $\left(s_1^{-1}[-1,1]\right) \times \left(s_2^{-1}[-1,1]\right) \times \left(s_3^{-1}[-1,1]\right)$-compatible, for all $s_{1,2,3} > 0$ with $s_1^2 + s_2^2 + s_3^2 \leq 1$.*

The last point follows from Section 4 and known measurement compatibility results, see [BN18], for example.

Recall from Section 2.2 the definitions of the direct sum and the direct product of polytopes containing 0. We gather next some results about the behavior of the Cartesian product and the direct sum operations when considering matrix levels.

Proposition 3.16. *Let $g_1, g_2 \in \mathbb{N}$. Let A_i be g_i-tuples of operators, and $\mathcal{P}_i \in \mathbb{R}^{g_i}$ be polytopes with $0 \in \mathrm{int}\,\mathcal{P}_i$, $i = 1, 2$. Then:*

1. *$(A_1, A_2) \in (\mathcal{P}_1 \times \mathcal{P}_2)_{\max} \iff A_1 \in (\mathcal{P}_1)_{\max}$ and $A_2 \in (\mathcal{P}_2)_{\max}$.*
2. *$(A_1, A_2) \in (\mathcal{P}_1 \times \mathcal{P}_2)_{\min} \implies A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (\mathcal{P}_2)_{\min}$, but the converse does not hold in general.*
3. *$A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (\mathcal{P}_2)_{\min} \implies (A_1 \oplus 0, 0 \oplus A_2) \in (\mathcal{P}_1 \times \mathcal{P}_2)_{\min}$ and $(A_1 \otimes I, I \otimes A_2) \in (\mathcal{P}_1 \times \mathcal{P}_2)_{\min}$; this holds even for tuples of operators having different dimensions.*
4. *If $q_1, q_2 \geq 0$ with $q_1 + q_2 \leq 1$, then $A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (P_2)_{\min} \implies (q_1 A_1, q_2 A_2) \in (\mathcal{P}_1 \times \mathcal{P}_2)_{\min}$.*
5. *$(A_1, A_2) \in (\mathcal{P}_1 \oplus \mathcal{P}_2)_{\min} \implies A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (\mathcal{P}_2)_{\min}$, but the converse does not hold in general.*
6. *$A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (\mathcal{P}_2)_{\min} \implies (A_1 \oplus 0, 0 \oplus A_2) \in (\mathcal{P}_1 \oplus \mathcal{P}_2)_{\min}$; this holds even for tuples of operators having different dimensions.*
7. *If $q_1, q_2 \geq 0$ with $q_1 + q_2 \leq 1$, then $A_1 \in (\mathcal{P}_1)_{\min}$ and $A_2 \in (P_2)_{\min} \implies (q_1 A_1, q_2 A_2) \in (\mathcal{P}_1 \oplus \mathcal{P}_2)_{\min}$.*

Proof. For the first point, note that the condition $\langle (A_1, A_2), h_{12} \otimes \rho \rangle \leq 1$ has to be checked only for extreme points $h_{12} \in (\mathcal{P}_1 \times \mathcal{P}_2)^{\circ} = \mathcal{P}_1^{\circ} \oplus \mathcal{P}_2^{\circ}$. Such extreme points are either of the form $(h_1, 0)$ or of the form $(0, h_2)$, with $h_i \in \mathrm{ext}\,\mathcal{P}_i^{\circ}$, see [BN20, Section 3.1].

For the second point, let $v_j^{(i)}{}_{j\in[l_i]}$ be the extreme points of $\mathcal{P}_i$. Then, we can write

$$(A_1, A_2) = \sum_{i\in[l_1]}\sum_{j\in[l_2]} (v_i^{(1)}, v_j^{(2)}) \otimes C_{ij}$$
$$= \sum_{i\in[l_1]} (v_i^{(1)}, 0) \otimes C_i^{(1)} + \sum_{j\in[l_2]} (0, v_j^{(2)}) \otimes C_j^{(2)}$$

with POVMs $C_i^{(1)} = \sum_{j\in[l_2]} C_{ij}$ and $C_j^{(2)} = \sum_{i\in[l_1]} C_{ij}$. Thus,

$$A_i = \sum_{j\in[l_i]} v_j^{(i)} \otimes C_j^{(i)}.$$

For the third point, start with decompositions

$$(13) \qquad\qquad A_1 = \sum_{i=1}^{k} v_i \otimes C_i$$

$$(14) \qquad\qquad A_2 = \sum_{j=1}^{l} w_j \otimes D_j,$$

with $\{v_i\}_{i\in[k]} \subset \mathcal{P}_1$ and $\{v_j\}_{j\in[l]} \subset \mathcal{P}_2$ and $\{C_i\}_{i\in[k]}$, $\{D_j\}_{j\in[l]}$ POVMs. From Eq. (7) we know that $\mathcal{P}_1 \oplus \mathcal{P}_2 \subseteq \mathcal{P}_1 \times \mathcal{P}_2$ and the first claim follows from the sixth point and Proposition 3.14. Finally, we can put

$$(A_1 \otimes I, I \otimes A_2) = \sum_{i=1}^{k}\sum_{j=1}^{l} (v_i, w_j) \otimes (C_i \otimes D_j)$$

and note that $\{C_i \otimes D_j\}_{i\in[k], i\in[k]}$ forms a POVM.

The fourth point follows again from the seventh point, Eq. (7), and Proposition 3.14.

For the fifth point, using again the general form of the extreme points of $P_1 \oplus P_2$, we have

$$(A_1, A_2) = \sum_{i=1}^{k} (v_i, 0) \otimes C_i + \sum_{j=1}^{l} (0, w_j) \otimes D_j,$$

for positive semidefinite operators C_i, D_j such that $\sum_i C_i + \sum_j D_j \leq I_d$. Hence, both $C = (C_i)_{i\in[k]}$ and $D = (D_j)_{j\in[l]}$ are sub-POVMs and the conclusion follows from Proposition 3.6.

For the sixth point, start with the decompositions in from Eqs. (13)-(14). Then write

$$(A_1 \oplus 0, 0 \oplus A_2) = \sum_{i=1}^{k} (v_i, 0) \otimes (C_i \oplus 0) + \sum_{j=1}^{l} (0, w_j) \otimes (0 \oplus D_j),$$

where $\{C_i \oplus 0\}_{i\in[k]} \sqcup \{0 \oplus D_j\}_{j\in[l]}$ forms itself a POVM.

For the seventh point, start again from Eqs. (13)-(14), and write

$$(q_1 A_1, q_2 A_2) = \sum_{i=1}^{k} (v_i, 0) \otimes q_1 C_i + \sum_{j=1}^{l} (0, w_j) \otimes q_2 D_j,$$

where $\{q_1 C_i\}_{i\in[k]} \sqcup \{q_2 D_j\}_{j\in[l]}$ forms itself a sub-POVM. Proposition 3.6 yields the conclusion. $\quad\square$

Remark 3.17. *The results above generalize in the obvious manner to more than two polytopes and tuples of operators.*

4. Measurement compatibility

The first example we will consider concerns the compatibility of dichotomic quantum measurements. Remember that in this case, we identify the POVM $\{E_i, I - E_i\}$ simply with the effect E_i. Our aim is to rephrase the compatibility of g dichotomic measurements in dimension d as $\mathcal{P}$-compatibility with $\mathcal{P} = [-1, 1]^g$. This recovers results from [BN22b] with alternative proofs. Taking a different $\mathcal{P}$, the results will extend to measurements with more outcomes. We will discuss this at the end of this section.

Proposition 4.1. *Let g, $d \in \mathbb{N}$ and let $A \in (M_d^{\mathrm{sa}}(\mathbb{C}))^g$. Then, the A are $[-1, 1]^g$-operators if and only if the $E_i = \frac{1}{2}(A_i + I_d)$ are effects for all $i \in [g]$.*

Proof. By Proposition 3.7, the A are $[-1, 1]^g$-operators if and only if

$$\mathrm{Tr}[A_i \rho] \in [-1, 1] \qquad \forall \rho \in \mathcal{M}_d^{1,+}(\mathbb{C}), \forall i \in [g].$$

This is equivalent to $-I_d \leq A_i \leq I_d \; \forall i \in [g]$, from which the assertion follows. $\square$

Proposition 4.2. *Let g, $d \in \mathbb{N}$ and let $A \in (M_d^{\mathrm{sa}}(\mathbb{C}))^g$. Then, the A are $[-1, 1]^g$-compatible if and only if the $E_i = \frac{1}{2}(A_i + I_d)$ are compatible effects for all $i \in [g]$.*

Proof. We note that the extreme points of $[-1, 1]^g$ are the sign vectors $\varepsilon_j \in \{\pm 1\}^g$, $j \in [2^g]$. From Definition 3.1, the A are $[-1, 1]^g$-compatible if and only if

$$A_i = \sum_{\varepsilon \in \{\pm 1\}^g} \varepsilon(i) C_\varepsilon \qquad \forall i \in [g]$$

and for some POVM $\{C_\varepsilon\}_{\varepsilon \in \{\pm 1\}^g}$. As $\sum_{\varepsilon \in \{\pm 1\}^g} C_\varepsilon = I_d$ and $\frac{1}{2}(\varepsilon(i) + 1) \in \{0, 1\}$, we obtain

$$E_i = \sum_{\substack{\varepsilon \in \{\pm 1\}^g \\ \varepsilon(i) = 1}} C_\varepsilon \qquad \forall i \in [g].$$

Thus, $\{C_\varepsilon\}_{\varepsilon \in \{\pm 1\}^g}$ is a joint POVM for the measurements defined by $\{E_i\}_{i \in [g]}$. $\square$

In fact, this example has motivated our terminology of $\mathcal{P}$-operators and $\mathcal{P}$-compatibility.

Remark 4.3. *We can extract a few easy consequences about compatible measurements from our theory of $\mathcal{P}$-compatibility. Most of them are easy to check directly and should be seen primarily as sanity checks and as providing a way to think about the propositions from which they follow in a more abstract setting.*

Point (1) of Proposition 3.16 implies that a collection of matrices is a collection of effects if and only if each matrix is an effect individually.

Point (2) of Proposition 3.16 implies that if you have a collection of compatible effects, taking any subset of these effects is still a collection of compatible effects. The proof of this point shows that the corresponding joint POVM arises from taking marginals of the joint POVM for the larger collection of compatible effects. It is known that you cannot simply combine sets of compatible effects to result in a larger set of compatible effects. This is easy to see, since any effect is compatible with itself, such that there would not be incompatible effects otherwise.

Point (3) of Proposition 3.16 implies that if you insert two collections of compatible measurements into different blocks of a larger matrix, their union is compatible. The same is true if you combine measurements on different subsystems.

Point (4) of Proposition 3.16 implies that for two collections of compatible measurements, adding a certain amount of noise to each of them implies that their union remains compatible. More concretely, if you have, for example, two collections of dichotomic measurements $\{E_1, \ldots, E_g\}$

and $\{F_1, \ldots, F_{g'}\}$, where the measurements $\{E_1, \ldots, E_g\}$ are compatible and the measurements $\{F_1, \ldots, F_{g'}\}$ are compatible, then the collection of noisy measurements

$$\{q_1 E_1 + (1 - q_1)I/2, \ldots, q_1 E_g + (1 - q_1)I/2, q_2 F_1 + (1 - q_2)I/2, \ldots, q_2 F_{g'} + (1 - q_2)I/2\}$$

are compatible as well if $q_1 + q_2 \leq 1$, since $2(q_1 E_1 + (1 - q_1)I/2) - I = q_1(2E_i - I)$. The latter shows that adding noise with parameter q_1 to the measurements is equivalent to multiplying the corresponding tensor A as in Proposition 4.1 by q_1. This can be interpreted in terms of coin tossing and mixing as in [HMZ16, Section 2.3].

From Proposition 3.13, it follows that there is no incompatibility if one considers only a single POVM. Conversely, the proposition implies that if the dimension is at least 2, we can always find a collection of incompatible measurements if we consider at least two measurements with at least two outcomes.

Remark 4.4. We could have also used Proposition 3.8 for the proofs in this section, since any collection of g measurements $E^{(i)}$ with k_i outcomes each, $i \in [g]$, can be seen as a measurement map $\mathcal{E}$ from $\mathcal{M}_d^{1,+}(\mathbb{C}) \to \Delta_\mathbf{k}$, given as

$$\mathcal{E} : \rho \mapsto (\mathrm{Tr}[E_1^{(i)}\rho]\delta_1^{(i)} + \ldots + \mathrm{Tr}[E_1^{(i)}\rho]\delta_{k_i}^{(i)})_{i \in [g]}.$$

Here, $\Delta_\mathbf{k}$ is the polysimplex, i.e., the GPT generated by the Cartesian product of simplices $\Delta_{k_1} \times \ldots \times \Delta_{k_g}$, and $\delta_1^{(i)}, \ldots, \delta_{k_i}^{(i)}$ are the vertices of the simplex Δ_{k_i}. We refer the reader to [Jen18] for details. In Theorem 1 of [Jen18], it was shown that the measurements $E^{(i)}$ are compatible if and only if $\mathcal{E}$ factors through a simplex (see also [BJN22]). This fact could also have been proven from Proposition 3.8 combined with Propositions 4.1 and 4.2.

Finally, we could have proven Propositions 4.1 and 4.2 using the techniques in [BN22b] based on tensor norms on Banach spaces and their link to matrix convex sets.

In this example, we have focused on dichotomic measurements for simplicity, but the correspondence works for general POVMs. Let $k \in \mathbb{N}$. Defining

(15) $$\mathcal{P}_k = \{x \in \mathbb{R}^{k-1} : \langle -ke_j, x \rangle \leq 1 \ \forall j \in [k-1], \langle k(1, \ldots, 1), x \rangle \leq 1\},$$

where the e_j are the standard basis vectors, we can check that $\mathcal{P}_k$ is a polytope with extreme points

$$\left\{ -\frac{1}{k}(1, \ldots, 1) + e_j \ \forall j \in [k] \right\} \cup \left\{ -\frac{1}{k}(1, \ldots, 1) \right\}$$

Then, we obtain the following statement that generalizes Propositions 4.1 and 4.2:

Proposition 4.5. Let g, d, $k_i \in \mathbb{N}$ for all $i \in [g]$ and let $A \in (\mathcal{M}_d^{\mathrm{sa}}(\mathbb{C}))^{k_1 + \ldots + k_g - g}$. Let $E_i^{(j)} = A_{k_1 + \ldots + k_{j-1} + i - j + 1} + \frac{1}{k_j}I_d$ for all $i \in [k_j - 1]$ and $E_{k_j}^{(j)} = I_d - E_1^{(j)} - \ldots - E_{k_j-1}^{(j)}$ for all $j \in [g]$. Let $\mathcal{P}_\mathbf{k} := \mathcal{P}_{k_1} \times \ldots \times \mathcal{P}_{k_g}$. Then,

  (1) The A are $\mathcal{P}_\mathbf{k}$-operators if and only if the tuples $(E_1^{(j)}, \ldots, E_{k_j}^{(j)})$ are POVMs for all $j \in [g]$.

  (2) The A are $\mathcal{P}_\mathbf{k}$-compatible if and only if the tuples $(E_1^{(j)}, \ldots, E_{k_j}^{(j)})$ are compatible POVMs for all $j \in [g]$.

The proof proceeds analogously to the proofs of Propositions 4.1 and 4.2. The statements of Remark 4.3 carry over with the necessary adjustments. We note that $\mathcal{P}_k$ is the polar of the matrix jewel base at the level 1, called $\mathcal{D}_{\oplus,k}(1)$ in [BN20]; note that there is a difference in normalization between the definitions above and the ones in [BN20] by a factor of 2. This set was denoted as $\mathcal{D}_{\oplus,k}(1)$ in that paper. Furthermore, $\mathcal{P}_\mathbf{k}$ is the polar of the matrix jewel at level 1, called $\mathcal{D}_{\oplus,\mathbf{k}}(1)$ in [BN20]. This set was denoted as $\mathcal{D}_{\oplus,\mathbf{k}}(1)$ in that paper. Contrary to the dichotomic case this proposition does not follow from [BN22b], since we cannot define tensor norms corresponding to these matrix convex sets as they are asymmetric.

5. MAGIC SQUARES

In this section we discuss the case of *magic squares* by relating them to the Birkhoff polytope. This is one of the main examples we discuss in detail, and was the starting point of our investigation.

We recall the following definitions from [DlCDN20] and from the quantum groups literature [BBC07], restricting ourselves to the matrix algebra setting.

Definition 5.1. *Let d, $N \in \mathbb{N}$. A block matrix $A \in \mathcal{M}_N(\mathcal{M}_d(\mathbb{C}))$ having positive semidefinite blocks $A_{ij} \geq 0 \ \forall i, j \in [N]$ is called*

- *a* quantum magic square *if it is* bistochastic*:*

$$\forall i \in [N], \sum_{j=1}^{N} A_{ij} = I_d \quad and \quad \forall j \in [N], \sum_{i=1}^{N} A_{ij} = I_d;$$

- *a* semiclassical magic square *if there exist a POVM $(Q_\pi)_{\pi \in \mathcal{S}_N}$ such that*

$$A = \sum_{\pi \in \mathcal{S}_N} P_\pi \otimes Q_\pi,$$

where P_π is the permutation matrix associated to π.

Remark 5.2. *It is easy to verify that a semiclassical magic square is a magic square. Indeed, P_π being a permutation matrix means that $(P_\pi)_{ij} = \delta_{\pi(i),j}$. Thus,*

$$A_{ij} = \sum_{\pi \in \mathcal{S}_n} \delta_{\pi(i),j} Q_\pi,$$

such that $A_{ij} \geq 0$. Moreover, for any $i \in [N]$, using that the Q_π form a POVM,

$$\sum_{j=1}^{N} A_{ij} = \sum_{\pi \in \mathcal{S}_n} \sum_{j=1}^{N} \delta_{\pi(i),j} Q_\pi = \sum_{\pi \in \mathcal{S}_n} Q_\pi = I_d.$$

Likewise, for any $j \in [N]$,

$$\sum_{i=1}^{N} A_{ij} = \sum_{\pi \in \mathcal{S}_n} \sum_{i=1}^{N} \delta_{\pi(i),j} Q_\pi = \sum_{\pi \in \mathcal{S}_n} \sum_{i=1}^{N} \delta_{i,\pi^{-1}(j)} Q_\pi = \sum_{\pi \in \mathcal{S}_n} Q_\pi = I_d.$$

5.1. The Birkhoff polytope.

In this subsection we shall introduce and study the basic properties of matrix convex sets built upon the celebrated Birkhoff polytope. These objects will be related to (semiclassical) magic squares in Subsection 5.2.

The *Birkhoff polytope* Birk_N is the convex set of bistochastic matrices from $\mathcal{M}_N(\mathbb{R})$ [Zie12, Example 0.12]. Famously, its vertices are the permutation matrices P_π, for $\pi \in \mathcal{S}_N$. Since this polytope lives in an $(N-1)^2$ dimensional affine hyperplane of $\mathbb{R}^{N^2}$, we shall consider the convex body version of the Birkhoff polytope, defined as follows. For an arbitrary matrix $X \in \mathcal{M}_N(\mathbb{C})$, we shall denote by $X^{(N-1)}$ the principal submatrix obtained by deleting the last row and the last column from X. We consider the convex body obtained by truncating bistochastic matrices, after centering the Birkhoff polytope at J/N, where J is the matrix in which all entries are 1.

Definition 5.3. *For a given $N \geq 2$, the* Birkhoff body *$\mathcal{B}_N$ is defined as the set of $(N-1) \times (N-1)$ truncations of $N \times N$ bistochastic matrices, shifted by J_{N-1}/N:*

$$\mathcal{B}_N = \{A^{(N-1)} - J_{N-1}/N \ : \ A \in \mathcal{M}_N(\mathbb{R}) \ bistochastic\} \subset \mathcal{M}_{N-1}(\mathbb{R}) \cong \mathbb{R}^{(N-1)^2}.$$

For example, it is easy to see that $\mathcal{B}_2 = [-1/2, 1/2]$. Let us now introduce an important notation: to a matrix $X \in \mathcal{M}_{N-1}(\mathcal{M}_d(\mathbb{C}))$, we associate the matrix $\tilde{X} \in \mathcal{M}_N(\mathcal{M}_d(\mathbb{C}))$ given by

$$\tilde{X}_{ij} = \frac{I_d}{N} + \begin{cases} X_{ij}, & \text{if } i, j \in [N-1] \\ -\sum_{k=1}^{N-1} X_{ik}, & \text{if } i \in [N-1], \, j = N \\ -\sum_{k=1}^{N-1} X_{kj}, & \text{if } j \in [N-1], \, i = N \\ \sum_{k,l=1}^{N-1} X_{kl}, & \text{if } i, j = N. \end{cases}$$

The matrix $\tilde{X}$ agrees with $X + J_{N-1}/N \otimes I_d$ on the top $(N-1) \times (N-1)$ corner, and has row and column sums equal to I_d.

The convex and combinatorial properties of the Birkhoff polytope have been studied in a series of papers by Brualdi and Gibson, see [BG77].

Proposition 5.4. *Let* $N \in \mathbb{N}$. *The Birkhoff body* $\mathcal{B}_N$ *has* $N!$ *extreme points* $P_\pi^{(N-1)} - J_{N-1}/N$ *and is described by the following inequalities:*

$$(16) \qquad\qquad \forall i, j \in [N-1], \quad A_{ij} \geq -1/N$$

$$(17) \qquad\qquad \forall i \in [N-1], \quad \sum_{j=1}^{N-1} A_{ij} \leq 1/N$$

$$(18) \qquad\qquad \forall j \in [N-1], \quad \sum_{i=1}^{N-1} A_{ij} \leq 1/N$$

$$(19) \qquad\qquad\qquad\qquad \sum_{i,j=1}^{N-1} A_{ij} \geq -1/N.$$

For $N \geq 3$, *its* N^2 *facets are given by replacing one of the* N^2 *inequalities by an equality. For* $N = 2$, *there are only* 2 *facets given by* $A_{11} = -1/2$ *and* $A_{11} = 1/2$, *respectively. Finally, it has* 0 *in its interior.*

Proof. We build on the findings on the extreme points and facets of the Birkhoff polytope in [BG77]. First, let us point out that, given the $(N-1) \times (N-1)$ shifted truncation of a $N \times N$ bistochastic matrix, there is an unique way to recover the bistochastic matrix, given by the $X \mapsto \tilde{X}$ mapping defined above. This fact settles the claim about extreme points. Regarding the facets, note that the second and the third type of inequalities above ensure that one can fill the last element in the first $(N-1)$ rows and columns with a non-negative number. The last inequality ensures the non-negativity of the bottom-right entry of $\tilde{A}$. The fact that $0 \in \text{int} \, \mathcal{B}_N$ holds, since 0 fulfills all inequalities strictly. $\qquad\square$

5.2. **Connecting magic squares to matrix convex sets.** This subsection contains the main insight of this section: there is an intimate relation between the Birkhoff body $\mathcal{B}_N$ introduced in Section 5.1 and the (semiclassical property for) magic squares introduced in Section 5.

Theorem 5.5. *Let* d. $N \in \mathbb{N}$ *Consider* $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^{(N-1)^2}$ *and the corresponding matrix* $\tilde{A} \in \mathcal{M}_N(\mathcal{M}_d(\mathbb{C}))$. *Then:*

(1) *the matrix* $\tilde{A}$ *is a magic square if and only if* $A \in (\mathcal{B}_N)_{\max}$;
(2) *the matrix* $\tilde{A}$ *is a semiclassical magic square if and only if* $A \in (\mathcal{B}_N)_{\min}$.

Proof. Requiring that the tuple A satisfies the inequalities from Proposition 5.4 is easily seen to be equivalent to the fact that $\tilde{A}$ is a magic square, establishing the first point. Similarly, the

second point follows from the form of the extreme points of the Birkhoff body, which are related to permutation matrices. Indeed, for $i, j \in [N-1]$, we have

$$A_{ij} = \sum_{\pi \in \mathcal{S}_N} (P_\pi(i,j) - 1/N)Q_\pi$$

for a POVM $Q = (Q_\pi)_{\pi \in \mathcal{S}_N}$. In turn, we have

$$\tilde{A}_{ij} = I/N + \sum_{\pi \in \mathcal{S}_N} (P_\pi(i,j) - 1/N)Q_\pi = \sum_{\pi \in \mathcal{S}_N} P_\pi(i,j)Q_\pi.$$

For $j = N$ and $i \in [N-1]$, we have

$$\tilde{A}_{iN} = I - \sum_{j=1}^{N-1} \tilde{A}_{ij} = I - \sum_{j=1}^{N-1} \sum_{\pi \in \mathcal{S}_N} P_\pi(i,j)Q_\pi$$

$$= \sum_{\pi \in \mathcal{S}_N} \left(1 - \sum_{j=1}^{N-1} P_\pi(i,j)\right) Q_\pi = \sum_{\pi \in \mathcal{S}_N} P_\pi(i,N)Q_\pi.$$

Similar computations yield the cases $i = N, j \in [N-1]$ and $i = j = N$, finishing the proof. $\qquad \square$

5.3. **Semiclassicality vs. compatibility.** The magic square condition from Definition 5.1 can be equivalently stated that the columns, respectively the rows of the square form quantum measurements:

$$R^{(i)} := \{A_{ij}\}_{j \in [N]} \qquad \text{and} \qquad C^{(j)} := \{A_{ij}\}_{i \in [N]}.$$

We have the following observation, see also [GB19].

Proposition 5.6. *Let* $N \in \mathbb{N}$. *For* $i, j \in [N]$, *the* $2N$ *measurements* $R^{(i)}$ *and* $C^{(j)}$ *(defined as above) of a* semiclassical *magic square are compatible.*

Proof. The column and row POVMs are post-processings in the sense of Lemma 2.6 of the measurement Q_π from Definition 5.1. For example,

$$C_i^{(j)} = A_{ij} = \sum_{\pi \in \mathcal{S}_N} \delta_{\pi(i),j} Q_\pi.$$

Above, we have used the post-processing

$$p_\pi(i|C^{(j)}) = \delta_{\pi(i),j} = \delta_{i,\pi^{-1}(j)}.$$

Similar expressions can be written for the row measurements, finishing the proof. $\qquad \square$

We now show, via an example, that the converse to the proposition above does not hold, i.e., there exist magic squares with compatible measurements $C^{(i)}$ and $R^{(j)}$ without the magic square being semiclassical. Consider the magic square in Table 1. Here, e_1, e_2 are the standard basis vectors in $\mathbb{C}^2$ and $f_1 = 1/\sqrt{2}(e_1 + e_2)$, $f_2 = 1/\sqrt{2}(e_1 - e_2)$.

$\frac{1}{2}e_1e_1^*$	$\frac{1}{2}e_2e_2^*$	0	$\frac{1}{2}I_2$
$\frac{1}{2}e_2e_2^*$	$\frac{1}{2}e_1e_1^*$	$\frac{1}{2}I_2$	0
0	$\frac{1}{2}I_2$	$\frac{1}{2}f_1f_1^*$	$\frac{1}{2}f_2f_2^*$
$\frac{1}{2}I_2$	0	$\frac{1}{2}f_2f_2^*$	$\frac{1}{2}f_1f_1^*$

TABLE 1. Example for 8 qubit POVMs arranged in a magic square which are compatible, but which do not form a semiclassical magic square.

It is easy to see that the measurements in the rows and columns of Table 1 in fact reduce to only two different POVMs (since all the other ones arise only as relabeling of outcomes). These two POVMs are

$$(20) \qquad \left(\frac{1}{2}e_1e_1^*, \frac{1}{2}e_2e_2^*, \frac{1}{2}I_2, 0\right) \quad \text{and} \quad \left(\frac{1}{2}f_1f_1^*, \frac{1}{2}f_2f_2^*, \frac{1}{2}I_2, 0\right)$$

Moreover, it is straightforward to verify that

$$\left(\frac{1}{2}e_1e_1^*, \frac{1}{2}e_2e_2^*, \frac{1}{2}f_1f_1^*, \frac{1}{2}f_2f_2^*\right)$$

is a joint POVM from which the POVMs in Eq. (20) arise via classical post-processing. Hence, all the POVMs shown in the rows and columns of Table 1 are compatible.

It is a bit tedious to show that the magic square in Table 1 is not semiclassical: if it was semiclassical, it would have an expression as

$$\sum_{\pi \in \mathcal{S}_4} P_\pi \otimes Q_\pi, \qquad Q_\pi \geq 0 \; \forall \pi \in \mathcal{S}_4, \qquad \sum_\pi Q_\pi = I_2$$

where $P_\pi = (\delta_{\pi(i),j})_{i,j\in[4]}$ are the permutation matrices representing π. Let us denote by $A_{i,j}$ the entries of the magic square. Since $A_{1,3} = A_{2,4} = A_{3,1} = A_{4,2} = 0$, only the Q_π with $\pi(1) \neq 3, \pi(2) \neq 4, \pi(3) \neq 1, \pi(4) \neq 2$ can be nonzero. Thus out of the 24 Q_π, 15 have to be zero. Since $A_{1,1}, A_{1,2}, A_{2,1}, A_{2,2}, A_{3,3}, A_{3,4}, A_{4,3}, , A_{4,4}$ are rank one, it follows that Q_π summing to these entries have to be proportional to the corresponding rank one projectors. For example $Q_{(1,2,3,4)}$ has to be proportional to both $e_1e_1^*$ and $f_1f_1^*$ and therefore must be 0. This reasoning sets another 8 of the Q_π to 0. The only non-zero element is thus $Q_{(4,3,2,1)}$, which needs to be $I/2$. However, this choice of Q_π does not generate the required magic square and contradicts the requirement of Q being a POVM.

Clearly, the measurements in Table 1 are not linearly independent, so we are not in the setting of [GB19, Corollary 5]. Similarly, the post-processing map used here is the "marginal" one, which is not symmetric in the sense of [GB19, Proposition 3]; these considerations show that the counterexample above do not contradict the results in [GB19].

Using our Proposition 3.8, we can recover the characterization in [GB19, Theorem 4] of semiclassical magic squares as compatibility plus symmetric post-processing:

Proposition 5.7. $A \in (\mathcal{B}_N)_{\min}$ if and only if the tuples $C^{(j)} := \{\tilde{A}_{ij}\}_{i\in[N]}$, $R^{(i)} := \{\tilde{A}_{ij}\}_{j\in[N]}$ where $i, j \in [N]$ are compatible POVMs with the post-processing in Lemma 2.6 satisfying $p_\lambda(i|R^{(j)}) = p_\lambda(j|C^{(i)})$ for all $i, j \in [N]$ and all λ.

Proof. Following the strategy for the proof of Theorem 5.5, using Proposition 3.8, one can verify that being a semiclassical magic square is equivalent to the map $\tilde{\mathcal{A}} : \mathcal{M}_d^{1,+}(\mathbb{C}) \to \mathrm{Birk}_N$,

$$\tilde{\mathcal{A}}(X) = \mathrm{Tr}[\tilde{A}_{ij}X]_{i,j\in[N]},$$

factoring through a k-simplex. The factorization is equivalent to the existence of a POVM $Q = \{Q_i\}_{i\in[k]}$ and a map $\nu : \Delta_k \to \mathrm{Birk}_N$ such that $\nu(\delta_l) = z_l$, where the δ_l are the vertices of the simplex for $l \in [k]$. Then, we can identify $p_\lambda(i|C^{(j)}) = \nu(\delta_\lambda)(i,j) = p_\lambda(j|R^{(i)})$ if we see z_λ as a bistochastic matrix with entries $z_\lambda(i,j)$. It is easy to check that the $p_\lambda(j|R^{(i)})$, $p_\lambda(i|C^{(j)})$ are conditional probabilities because the z_λ are bistochastic matrices and that

$$A_{ij} = \sum_{\lambda\in[k]} z_\lambda(i,j)Q_\lambda = \sum_{\lambda\in[k]} p_\lambda(i|C^{(j)})Q_\lambda = \sum_{\lambda\in[k]} p_\lambda(j|R^{(i)})Q_\lambda.$$

Since $C_i^{(j)} = A_{ij} = R_j^{(i)}$ and our identification between conditional probabilities and bistochastic matrices can be reversed, the assertion follows. $\square$

6. POVMs with common elements

We discuss in this section a very general way of defining polytopes for which the notions of $\mathcal{P}$-operators and $\mathcal{P}$-compatibility have a very clear physical interpretation. Interestingly, the mathematical framework is that of (measurement) hypergraphs, which is precisely the one used in the combinatorial approach to contextuality [AFLS15]. We would like to point out that, beyond the mathematical correspondence, there is no clear physical relation between contextuality and the notion of compatibility discussed here.

We start by recalling that a hypergraph is a pair $G = (V, E)$, where V is a non-empty set and E is a set of non-empty subsets $\emptyset \neq e \subseteq V$, called *hyperedges*. In this section, we shall assume that our hypergraphs have *no isolated vertices*, i.e.

$$\forall v \in V \quad \exists e \in E \text{ s.t. } v \in e \iff \bigcup_{e \in E} e = V.$$

Definition 6.1. *A hypergraph G (with no isolated vertices) is called a* probability hypergraph *if there exists a function $\pi : V \to (0, 1]$ such that*

$$(21) \qquad \forall e \in E, \quad \sum_{v \in e} \pi(v) = 1.$$

In other words, given a probability hypergraph, we can associate to each vertex a positive number in such a way that all hyperedges correspond to probability vectors. We denote by $\Pi(G)$ the set of all functions $\pi : V \to [0, 1]$ such that Eq. (21) holds.

We display in Figure 2 three hypergraphs. The first two are probability hypergraphs, since assigning weight $\pi(v) = 1/2$ to each vertex of the first graph and $\pi(v) = 1/3$ to each vertex of the second satisfies the condition from Definition 6.1. However, the third one is *not* a probability hypergraph: hyperedge $\{1\}$ implies $\pi(1) = 1$, while the other two hyperedges $\{1, 2\}$ and $\{1, 3\}$ force $\pi(2) = \pi(3) = 0$, contradicting $\pi(v) > 0$.
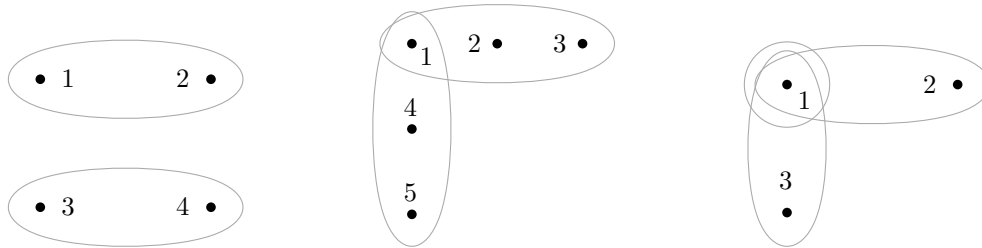

FIGURE 2. Three hypergraphs. The first two are probability hypergraphs, while the third one is not.

The incidence relation between vertices and hyperedges generate an equivalence relation on E, as the closure of the binary relation $e \cap e' \neq \emptyset \implies e \sim e'$.

The following result establishes a large family of hypergraphs as probability hypergraphs.

Proposition 6.2. *All hypergraphs $G = (V, E)$ having the property that edges in the same incidence equivalence class have the same cardinality*

$$e \sim e' \implies |e| = |e'|$$

are probability hypergraphs.

Proof. The result follows by setting

$$\forall v \in V, \qquad \pi(v) := \frac{1}{|e|} \quad \text{for } v \in e.$$

The assignment above is independent of the choice of the hyperedge $e \ni v$ by the hypothesis. Clearly, $\pi \in \Pi(G)$ and $\pi(v) > 0$, proving the claim. $\qquad\square$

In particular, hypergraphs with disjoint hyperedges ($e \neq e' \implies e \cap e' = \emptyset$) are probability hypergraphs. The first two examples in Figure 2 satisfy the assumptions of the proposition above.

In what follows, we shall associate to a probability hypergraph G a polytope in an essentially unique manner. To start, consider the set

$$\Pi^0(G) := \{\pi : V \to \mathbb{R} \, : \, \forall e \in E, \sum_{v \in e} \pi(v) = 0\}.$$

Clearly, $\Pi^0(G)$ is a vector space, and we set $g := \dim \Pi^0(G)$.

Let G be a probability hypergraph, $\pi_* \in \Pi(G)$, and consider a basis $\pi_1, \ldots, \pi_g$ of the vector space $\Pi^0(G)$. Define the set

$$\mathcal{P} := \{a \in \mathbb{R}^g \, : \, \pi_* + \sum_{x=1}^{g} a_x \pi_x \in \Pi(G)\},$$

which depends on the choice of the functions $\pi_*, \pi_1, \ldots, \pi_g$.

Proposition 6.3. *Let $g \in \mathbb{N}$. The set $\mathcal{P}$ as above is a polytope in $\mathbb{R}^g$, containing 0 in its interior. One can recover the set $\Pi(G)$ (see Definition 6.1) from the polytope $\mathcal{P}$:*

$$\Pi(G) = \left\{ \pi_* + \sum_{x=1}^{g} a_x \pi_x \, : \, a \in \mathcal{P} \right\}.$$

A different choice $\pi_*', \pi_1', \ldots, \pi_g'$ yields a polytope $\mathcal{P}'$ which is an affine transformation of $\mathcal{P}$.

Proof. A point $a \in \mathbb{R}^g$ is an element of $\mathcal{P}$ if and only if the following conditions are satisfied:

$$\forall v \in V, \qquad \pi_*(v) + \sum_{x=1}^{g} a_x \pi_x(v) \geq 0.$$

Note that these conditions are affine in a and define the facets of the polytope $\mathcal{P}$. The normalization condition Eq. (21) is automatically satisfied:

$$\forall e \in E \qquad \sum_{v \in e} \pi_*(v) + \sum_{x=1}^{g} \sum_{v \in e} a_x \pi_x(v) = \sum_{v \in e} \pi_*(v) = 1.$$

For the final claim, changing the base point π_* to a different one amounts to a translation of the polytope, while a basis change for the vector space $\Pi^0(G)$ amounts to an invertible linear transformation of $\mathcal{P}$. $\qquad\square$

It is remarkable that all reasonable polytopes can be obtained in the way described above, a result due to Shultz [Shu74].

Theorem 6.4 ([Shu74]). *Any polytope $\mathcal{P}$ having vertices with rational coefficients can be obtained from a probability hypergraph G.*

We consider now a matrix version of the set $\Pi(G)$ from Definition 6.1, by requiring that the function π is matrix-valued.

Definition 6.5. *Let $d \in \mathbb{N}$. Given G a probability hypergraph, define, for all $d \geq 1$*

$$\Pi_d(G) := \left\{ A : V \to \mathrm{PSD}_d \, : \, \forall e \in E, \sum_{v \in E} A(v) = I_d \right\}.$$

We refer to the elements of $\Pi_d(G)$ as G-operators.

Clearly, $\Pi_d(G)$ is the set of POVMs corresponding to the hyperedges of G, where common vertices correspond to common effects.

The set of semiclassical POVMs can be seen as coming from operators obtained from the extreme points of the polytope $\mathcal{P}$. Indeed, let $\sigma_1, \ldots, \sigma_k$ be the extreme points of $\Pi(G)$, and $C_1, \ldots, C_k$ be a POVM. The operators

$$\forall v \in V \qquad A_v := \sum_{i=1}^{k} \sigma_i(v) C_i$$

are called semiclassical. In other words, operators $\{A_v\}_{v \in V}$ are called G-compatible if there exits a POVM $\{C_\sigma\}$ indexed by the extreme points $\sigma \in \operatorname{ext} \Pi(G)$ such that

$$\forall v \in V \qquad A_v = \sum_{\sigma \in \operatorname{ext} \Pi(G)} \sigma(v) C_\sigma.$$

One can easily show the following result, connecting the formalism introduced in this section to the one from Section 3.

Proposition 6.6. *Let G be a probability hypergraph. To a g-tuple of operators $(B_x)_{x \in [g]}$ associate a tuple of self-adjoint operators $(A_v)_{v \in V}$ defined by*

$$\forall v \in V, \qquad A_v := \pi_*(v) I_d + \sum_{x=1}^{g} \pi_x(v) B_x \in \mathcal{M}_d^{sa}(\mathbb{C}),$$

where $\pi_*, \pi_1, \ldots, \pi_g$ define the polytope $\mathcal{P}$ associated to G. Then, we have the following equivalences:
- *$(B_x)_{x \in [g]}$ are $\mathcal{P}$-operators $\iff$ $(A_v)_{v \in V}$ are G-operators;*
- *$(B_x)_{x \in [g]}$ are $\mathcal{P}$-compatible $\iff$ $(A_v)_{v \in V}$ are G-compatible.*

We discuss next the relation between the notion of hypergraph compatibility introduced in this section, and the standard notion of compatibility for the quantum measurements associated to the hyperedges. In the spirit of [GB19], hypergraph compatibility is equivalent to standard compatibility plus the requirement that there exists a post-processing respecting the structure of the hypergraph.

Theorem 6.7. *Let $G = (V, E)$ be a probability hypergraph, and consider a tuple of G-operators $(A_v)_{v \in V} \subseteq \mathcal{M}_d^{sa}(\mathbb{C})$, $d \in \mathbb{N}$. Consider also the POVMs $\hat{A}_{\cdot|e} = (A_v)_{v \in e}$ indexed by the hyperedges of G. The following assertions are equivalent:*
- *The tuple $(A_v)_{v \in V}$ is G-compatible.*
- *The measurements $\hat{A}_{\cdot|e}$ are compatible (in the standard sense of quantum mechanics), with the additional constraint that they can be post-processed from a single POVM $B = (B_\lambda)_{\lambda \in \Lambda}$*

$$\forall e \in E, \forall v \in e, \qquad \hat{A}_{v|e} = A_v = \sum_{\lambda \in \Lambda} p(v|e, \lambda) B_\lambda$$

using a post-processing p respecting the symmetry of G:

(22) $$\forall e, f \in E, \forall v \in e \cap f, \forall \lambda \in \Lambda, \qquad p(v|e, \lambda) = p(v|f, \lambda).$$

Proof. The fact that the first point implies the second one is immediate: the tuple A being G-compatible implies that there exists a POVM C indexed by the extreme points of $\Pi(G)$ such that

$$\forall v \in V \qquad A_v = \sum_{\sigma \in \operatorname{ext} \Pi(G)} \sigma(v) C_\sigma.$$

Since the functions σ have the property that

$$\forall e \in E, \qquad \sum_{v \in e} \sigma(v) = 1,$$

this yields the conclusion by setting $\Lambda = \mathrm{ext}\,\Pi(G)$ and $p(v|e,\sigma) = \sigma(v)$.

For the reverse implication, the symmetry condition Eq. (22) implies that we can unambiguously define

$$\forall v \in V, \forall \lambda \in \Lambda, \qquad \pi(v|\lambda) := p(v|e,\lambda) \quad \text{for any } e \ni v.$$

With this notation, we rewrite the (standard) compatibility of the POVMs $\hat{A}_{\cdot|e}$ as

$$\forall v \in V, \qquad A_v = \sum_{\lambda \in \Lambda} \pi(v|\lambda) B_\lambda.$$

Note that since p is a post-processing, we have, for all $\lambda \in \Lambda$,

$$\forall e \in E \qquad \sum_{v \in e} \pi(v|\lambda) = 1,$$

hence $\pi(\cdot|\lambda) \in \Pi(G)$. We now decompose these elements in $\Pi(G)$ in terms of the extreme points:

$$\forall \lambda \in \Lambda, \qquad \pi(\cdot|\lambda) = \sum_{\sigma \in \mathrm{ext}\,\Pi(G)} q(\sigma|\lambda)\sigma(\cdot),$$

for some conditional probabilities $q(\cdot|\cdot)$. Putting everything together, we have

$$\forall v \in V, \qquad A_v = \sum_{\sigma \in \mathrm{ext}\,\Pi(G)} \sigma(v) \underbrace{\sum_{\lambda \in \Lambda} q(\sigma|\lambda) B_\lambda}_{C_\sigma}.$$

It is easy to check that the C_σ defined above form a POVM, concluding the proof. $\qquad\square$

Let us now emphasize the considerations above with an example, which corresponds to the middle hypergraph in Figure 2, also reproduced below for convenience. This hypergraph G, having 5 vertices and 2 hyperedges, corresponds to two 3-outcome POVMs sharing one effect. The set of probability vectors on G can be easily computed:

$$\Pi(G) = \{(x, y, 1-x-y, z, 1-x-z) : 0 \le x \le 1 \text{ and } 0 \le y \le 1-x \text{ and } 0 \le z \le 1-x\}.$$

The allowed triples (x, y, z) form a pyramid with a square base

$$\mathrm{conv}\{((1,0,0), (0,0,0), (0,0,1), (0,1,0), (0,1,1)\},$$

depicted below (right panel).
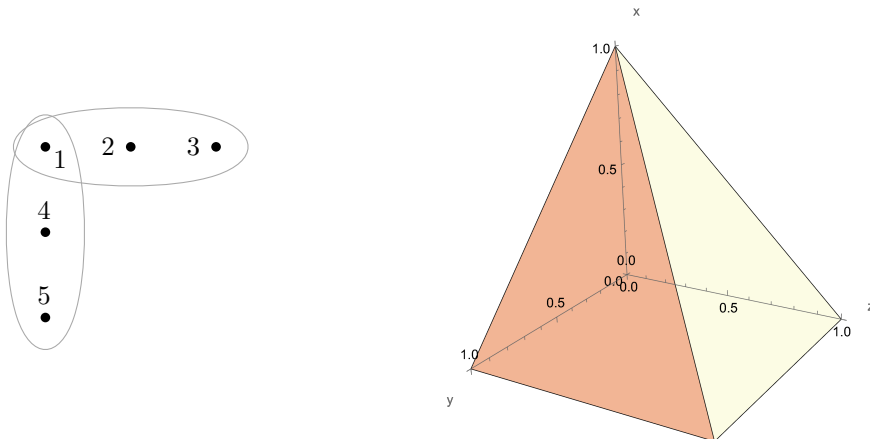
Considering the distinguished element $\pi_* := (1,1,1,1,1)/3 \in \Pi(G)$ and the basis elements

$$\pi_1 = (1, 0, -1, 0, -1)$$
$$\pi_2 = (0, 1, -1, 0, 0)$$
$$\pi_3 = (0, 0, 0, 1, -1),$$

we obtain the associated polytope

(23) $$\mathcal{P} = (-1/3, -1/3, -1/3) + \operatorname{conv}\{((1,0,0),(0,0,0),(0,0,1),(0,1,0),(0,1,1)\}.$$

The following result is an application of Proposition 6.6 and Theorem 6.7

Proposition 6.8. *Consider a triple of self-adjoint matrices $(A, B, C) \in (\mathcal{M}_d^{sa}(\mathbb{C}))^3$, $d \in \mathbb{N}$. Then*

- *$(A, B, C) \in 1/3(I, I, I) + \mathcal{P}_{\max}(d)$ if and only if $(A, B, I_d - A - B, C, I_d - A - C)$ are G-operators if and only if both triples $(A, B, I_d - A - B)$ and $(A, C, I_d - A - C)$ are POVMs.*
- *$(A, B, C) \in 1/3(I, I, I) + \mathcal{P}_{\min}(d)$ if and only if $(A, B, I_d - A - B, C, I_d - A - C)$ are G-compatible if and only if the two POVMs $(A, B, I_d - A - B)$ and $(A, C, I_d - A - C)$ are compatible and the post-processing in Lemma 2.6 satisfies $p_\lambda(1|1) = p_\lambda(1|2)$, where $p_\lambda(\cdot|1)$, $p_\lambda(\cdot|2)$ are conditional probabilities for all λ.*

Note that the last point above corresponds to the existence of a joint POVM of the form

Q_1	0	0	$= A$
0	Q_5	Q_4	$= B$
0	Q_3	Q_2	$= I_d - A - B$
$= A$	$= C$	$= I_d - A - C$	$\sum = I_d$

One can ask a question similar to the one in Section 5.3. We present now an example of compatible POVMs which do not admit a joint POVM of the form above. Again, e_1, e_2 are the standard basis vectors in $\mathbb{C}^2$ and $f_1 = 1/\sqrt{2}(e_1 + e_2)$, $f_2 = 1/\sqrt{2}(e_1 - e_2)$. Consider the following two 3-outcome qubit measurements:

(24) $$\left(\frac{1}{2}I_2, \frac{1}{2}e_1 e_1^*, \frac{1}{2}e_2 e_2^*\right) \quad \text{and} \quad \left(\frac{1}{2}I_2, \frac{1}{2}f_1 f_1^*, \frac{1}{2}f_2 f_2^*\right).$$

These POVM are compatible, as they are the marginals of the following joint POVM:

	0	$f_1 f_1^*$	$f_2 f_2^*$
$e_1 e_1^*$	0	0	
$e_2 e_2^*$	0	0	

with $\frac{1}{2} \cdot$ prefactor.

However, when trying to write

$$\frac{1}{2}I_2 = Q_1$$
$$\frac{1}{2}e_1 e_1^* = Q_4 + Q_5$$
$$\frac{1}{2}e_2 e_2^* = Q_2 + Q_3$$
$$\frac{1}{2}f_1 f_1^* = Q_3 + Q_5$$
$$\frac{1}{2}f_2 f_2^* = Q_2 + Q_4,$$

we infer, for example, $Q_2 \sim e_2 e_2^*$ and $Q_2 \sim f_2 f_2^*$, hence $Q_2 = 0$. The same reasoning yields $Q_3 = Q_4 = Q_5 = 0$, implying $Q_1 = I_2$, which contradicts the first equality above. In conclusion, the POVMs from Eq. (24) are compatible but the corresponding effects do not belong to $\mathcal{P}_{\min}(2)$.

7. INCLUSION CONSTANTS

The main theme in this paper is understanding the relation between the sets of $\mathcal{P}$-operators and $\mathcal{P}$-compatible operators for a given polytope $\mathcal{P}$. Since the inclusion $\mathcal{P}_{\min}(d) \subseteq \mathcal{P}_{\max}(d)$ always holds, we would like to have a measure of how restrictive the compatibility condition is. This motivates the following definition, which is a specialization of Definition 2.3 to matrix convex sets defined by polytopes.

Definition 7.1. *Let g, $d \in \mathbb{N}$. For a polytope $\mathcal{P} \subseteq \mathbb{R}^g$ having 0 in its interior, we define the set of inclusion constants of $\mathcal{P}$ at level d by*

$$\Delta_{\mathcal{P}}(d) := \{s \in \mathbb{R}^g \, : \, s \cdot \mathcal{P}_{\max}(d) \subseteq \mathcal{P}_{\min}(d)\}.$$

In words, $s \in \Delta_{\mathcal{P}}(d)$ if and only if for any *g-tuple of $\mathcal{P}$-operators $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$, the scaled tuple $(s_1 A_1, s_2 A_2, \ldots, s_g A_g)$ is $\mathcal{P}$-compatible.*

Computing the set of inclusion constants for a given polytope at a given level is in general a difficult problem, having many applications. For example, in the case of the hypercube, inclusion constants are related to the robustness of incompatibility of quantum effects [BN18]. More generally, the case of direct products of simplices is related to compatibility of quantum measurements [BN20]. The dual case, that of the cross polytope, is related to quantum steering [BN22a].

Clearly, the set of inclusion constants of a polytope is convex. We can also show that it contains 0 in its interior. The following lemma is most likely known:

Lemma 7.2. *Let $\mathcal{F}$ be a matrix convex set consisting of g-tuples, $g \in \mathbb{N}$. If $0 \in \mathrm{int}\,\mathcal{F}_1$, then $0 \in \mathrm{int}\,\mathcal{F}_n$ for all $n \in \mathbb{N}$.*

Proof. Let us consider vectors of the form xe_i, where e_i is the i-th standard basis vector and $x \in \mathbb{R}$. As $0 \in \mathrm{int}\,\mathcal{F}_1$, there exists a $C > 0$ such that $xe_i \in \mathcal{F}_1$ for all x with $|x| \leq C$ and all $i \in [g]$. As matrix convex sets are closed under direct sums, this implies that $(0, \ldots, 0, \mathrm{diag}[x_1, \ldots, x_n], 0, \ldots, 0) \in \mathcal{F}_n$ for all $x_1, \ldots, x_n$ such that $|x_j| \leq C$ for all $j \in [n]$, irrespective of the position of the diagonal matrix in the g-tuple. As the matrix convex set is closed under UCP maps, it is in particular closed under unitary conjugation, such that $(0, \ldots, 0, X, 0, \ldots, 0) \in \mathcal{F}_n$ for all $X \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})$ with $\|X\|_\infty \leq C$, where the norm is the operator norm. As every level of the matrix convex set is convex, we infer that $(X_1, \ldots, X_g) \in \mathcal{F}_n$ if $X_j \in \mathcal{M}_n^{\mathrm{sa}}(\mathbb{C})$ and $\|X_j\|_\infty \leq C/g$ for all $j \in [j]$. This proves the claim. $\square$

Corollary 7.3. *Let $\mathcal{P}$ be a polytope with 0 in its interior. Then, $0 \in \mathrm{int}\,\Delta_{\mathcal{P}}(d)$ for all $d \in \mathbb{N}$.*

Proof. Using Lemma 7.2, it holds that $0 \in \mathcal{P}_{\min}(d)$. Since $\mathcal{P}$ is a polytope, it is easy to see that $\mathcal{P}_{\max}(d)$ is bounded. Hence, there is a $C > 0$ such that $s \cdot \mathcal{P}_{\max}(d) \subseteq \mathcal{P}_{\min}(d)$ for all vectors s with $\|s\|_2 \leq C$. $\square$

The sets of inclusion constants are decreasing with the dimension:

Proposition 7.4. *Let $\mathcal{P} \subseteq \mathbb{R}^g$ be a polytope containing 0 in their interior, and $d \leq d'$, where g, d, $d' \in \mathbb{N}$. We have then*

$$\Delta_{\mathcal{P}}(d) \supseteq \Delta_{\mathcal{P}}(d').$$

Proof. Let $s \in \Delta_{\mathcal{P}}(d')$ and $A \in \mathcal{P}_{\max}(d)$. We claim that $A \oplus 0_{d'-d} \in \mathcal{P}_{\max}(d')$. To see this, consider some fixed defining hyperplane h_j and compute

$$\sum_{x=1}^{g} h_j(x)(A_x \oplus 0_{d'-d}) = \left(\sum_{x=1}^{g} h_j(x) A_x\right) \oplus 0_{d'-d} \leq I_d \oplus 0_{d'-d} \leq I_{d'}.$$

Alternatively, we could have used that matrix convex sets are closed under direct sums. From the fact that s is an inclusion constant at level d', it follows that $s \cdot (A \oplus 0_{d'-d}) \in \mathcal{P}_{\min}(d')$. Hence, there exist elements $v_1, \ldots, v_k \in \mathcal{P}$ and a POVM $C_1, \ldots, C_k$ of size d' such that, for all $x \in [g]$,

$$s_x A_x \oplus 0_{d'-d} = \sum_{i=1}^{k} v_i(x) C_i.$$

Defining $\tilde{C}_i$ to be the $d \times d$ corner of C_i, which is still a POVM, shows that

$$s_x A_x = \sum_{i=1}^{k} v_i(x) \tilde{C}_i,$$

finishing the proof. $\qquad\square$

The following result shows that the set of inclusion constants behaves nicely with respect to the Cartesian product and direct sum operations. Note that one can easily generalize the inclusions below to more than two polytopes.

Proposition 7.5. *Let $\mathcal{P} \subseteq \mathbb{R}^g$ and $\mathcal{Q} \subseteq \mathbb{R}^n$ be two polytopes containing 0 in their interior. Then, for all $d \geq 1$, $\Delta_{\mathcal{P} \times \mathcal{Q}}(d) \supseteq \Delta_{\mathcal{P}}(d) \oplus \Delta_{\mathcal{Q}}(d)$*

Proof. Consider $(A, B) \in (\mathcal{P} \times \mathcal{Q})_{\max}(d)$, and $s \in \Delta_{\mathcal{P}}(d)$. From Proposition 3.16, we have $A \in \mathcal{P}_{\max}(d)$, $B \in \mathcal{Q}_{\max}(d)$, and also $s \cdot A \in \mathcal{P}_{\min}(d)$. Since $0 \in \mathcal{Q}_{\min}(d)$, by using Proposition 3.16 again, we have $(s \cdot A, 0) \in (\mathcal{P} \times \mathcal{Q})_{\min}(d)$, proving that $(s, 0) \in \Delta_{\mathcal{P} \times \mathcal{Q}}(d)$. A similar reasoning shows that

$$t \in \Delta_{\mathcal{Q}}(d) \implies (0, t) \in \Delta_{\mathcal{P} \times \mathcal{Q}}(d),$$

concluding the proof of the claim. $\qquad\square$

Corollary 7.6. *Let $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be polytopes containing 0 in their interior, $n \in \mathbb{N}$, $s_i \in \Delta_{\mathcal{P}_i}(d)$ for all $i \in [n]$, and $\lambda = (\lambda_1, \ldots, \lambda_n)$ a probability vector. Then*

$$(\lambda_1 s_1, \lambda_2 s_2, \ldots, \lambda_n s_n) \in \Delta_{\times_{i=1}^n \mathcal{P}_i}(d).$$

In particular, if the $\mathcal{P}_i$ are simplices, then, for all $d \geq 1$,

$$n^{-1}(1, 1, \ldots, 1) \in \Delta_{\times_{i=1}^n \mathcal{P}_i}(d).$$

Proof. The claim about the polysimplex follows from the fact that for any simplex $\mathcal{P}$, $\mathcal{P}_{\min} = \mathcal{P}_{\max}$ (see Proposition 3.13). This fact corresponds to the fact that mixing quantum measurements with white noise using parameter $1/n$ yields compatible measurements [HMZ16]. $\qquad\square$

We derive in the remaining part of this section some general bounds for the set of inclusion constants of polytopes, applying them to the case of the Birkhoff body $\mathcal{B}_N$ (and the pyramid polytope discussed at the end of Section 6 in the end of the paper). We shall discuss three different methods:

- comparing with other polytopes
- symmetrization
- linear programming bounds

7.1. **Relating different polytopes.** We start with the first method, where one relates the inclusion constants for two polytopes.

Proposition 7.7. *Let $g \in \mathbb{N}$. Give two polytopes $\mathcal{P}, \mathcal{Q} \subseteq \mathbb{R}^g$ containing 0 in their interior, define the sets*

$$\mathcal{I}_{\mathcal{P} \to \mathcal{Q}} := \{s \in \mathbb{R}^g \; : \; s \cdot \mathcal{P} \subseteq \mathcal{Q}\}$$
$$\mathcal{I}_{\mathcal{Q} \to \mathcal{P}} := \{u \in \mathbb{R}^g \; : \; u \cdot \mathcal{Q} \subseteq \mathcal{P}\}.$$

Then, for all $d \geq 1$, $d \in \mathbb{N}$,

$$\mathcal{I}_{\mathcal{P} \to \mathcal{Q}} \cdot \Delta_{\mathcal{Q}}(d) \cdot \mathcal{I}_{\mathcal{Q} \to \mathcal{P}} \subseteq \Delta_{\mathcal{P}}(d).$$

Proof. Consider a tuple $A \in \mathcal{M}_d^{\mathrm{sa}}(\mathbb{C})^g$ of $\mathcal{P}$ operators, and scaling vectors $s \in \mathcal{I}_{\mathcal{P} \to \mathcal{Q}}$, $t \in \Delta_{\mathcal{Q}}(d)$, $u \in \mathcal{I}_{\mathcal{Q} \to \mathcal{P}}$. Since A are $\mathcal{P}$-operators, we have

$$\forall \rho \in \mathcal{M}_d^{1,+}(\mathbb{C}) \quad (\langle A_x, \rho \rangle)_{x \in [g]} \in \mathcal{P} \implies (\langle s_x A_x, \rho \rangle)_{x \in [g]} \in \mathcal{Q},$$

that is $s \cdot A \in \mathcal{Q}_{\max}(d)$. In turn, using the fact that t is an inclusion constant for $\mathcal{Q}$, we have that $t \cdot s \cdot A \in \mathcal{Q}_{\min}(d)$. In particular, this means that there exists a POVM $C_1, \ldots, C_k$ such that

$$t \cdot s \cdot A = \sum_{i=1}^{k} w_i \otimes C_i,$$

where $w_1, \ldots, w_k \in \mathbb{R}^g$ are elements (e.g. the extreme points) of $\mathcal{Q}$. Scaling by u gives

$$u \cdot t \cdot s \cdot A = \sum_{i=1}^{k} (u \cdot w_i) \otimes C_i,$$

with $u \cdot w_i \in \mathcal{P}$ now, proving the claim. $\qquad\square$

We shall now apply the result above to the case of the Birkhoff body $\mathcal{B}_N$ and a *polysimplex*, i.e. a Cartesian product of simplices [Jen18, Section III], [BJN22, Section 5]. Note that this method has been used previously to derive inclusion constants for free spectrahedra by comparing them to more symmetric matrix convex sets for which the set of inclusion constants was better known [BN20, Section 7]. Starting from this point, we are focusing on "flat" inclusion constants of the form $s(1, 1, \ldots, 1)$, so we are going to simply identify the scalar s with the corresponding flat vector.

Consider the polysimplex

$$\mathcal{Q} := \bigtimes_{i=1}^{N-1} \mathcal{P}_N \subseteq \mathbb{R}^{(N-1)^2} \cong \mathcal{M}_{N-1}(\mathbb{R}),$$

where $\mathcal{P}_N \subseteq R^{N-1}$ is the simplex defined in Eq. (15). Notice that $\mathcal{B}_N \subseteq \mathcal{Q}$, since the extreme points of $\mathcal{B}_N$ are a subset of those of $\mathcal{Q}$; this shows that $1 \in \mathcal{I}_{\mathcal{B}_N \to \mathcal{Q}}$. In order to find the largest scalar $u \geq 0$ such that $u \in \mathcal{I}_{\mathcal{Q} \to \mathcal{B}_N}$, note that the extreme points of $\mathcal{Q}$ are of the form

$$q_f := -\frac{1}{N} J_{N-1} + \begin{bmatrix} e_{f(1)} \\ e_{f(2)} \\ \vdots \\ e_{f(N-1)} \end{bmatrix},$$

where $f : [N-1] \to [N]$ is an arbitrary function, and where we set $e_N = 0$. We have to determine by how much we have to scale these extreme points in order for them to satisfy the hyperplane equations Eqs. (16)-(19). Eqs. (16) and (17) are satisfied without scaling ($u \leq 1$). For Eq. (18), fix $j \in [N-1]$ and compute

$$\sum_{i=1}^{N-1} q_f(i, j) = -\frac{N-1}{N} + |\{i \in [N-1] : f(i) = j\}| \leq -\frac{N-1}{N} + N - 1 = \frac{(N-1)^2}{N}.$$

Hence, for $u \cdot q_f \in \mathcal{B}_N$ to hold, we need, in the worst case, $u \leq 1/(N-1)^2$. We leave to the reader the case of the condition Eq. (19), which yields the same inequality, $u \leq 1/(N-1)^2$. We conclude that

$$\frac{1}{(N-1)^2} \in \mathcal{I}_{\mathcal{Q} \to \mathcal{B}_N}.$$

From Corollary 7.6 we have, for all $d \geq 1$, $1/(N-1) \in \Delta_{\mathcal{Q}}(d)$. Putting all these facts into Proposition 7.7, we arrive at the following result.

Corollary 7.8. *For any $N \geq 2$ and $d \geq 1$,*

$$\frac{1}{(N-1)^3} \in \Delta_{\mathcal{B}_N}(d).$$

7.2. Symmetrization. In this section we show that matrix convex sets with *symmetric* first level admit a non-trivial, dimension-dependent, inclusion constant vector. This vector does not depend on the length of the tuple. We start with a general result regarding matrix convex sets, specialize to polytope inclusion constants, and then derive constants for the Birkhoff body $\mathcal{B}_N$.

The following result is inspired by [HKMS19, Proposition 8.1] for free spectrahedra, which are a special class of matrix convex sets (see also [BN18, Proposition 7.2]).

Proposition 7.9. *Let $d \in \mathbb{N}$ and let $\mathcal{F}, \mathcal{G}$ be matrix convex sets such that $\pm \mathcal{F}(1) \subseteq \mathcal{G}(1)$. Then,*

$$\frac{1}{2d-1}\mathcal{F}(d) \subseteq \mathcal{G}(d) \quad \text{if } d \text{ is even,}$$

$$\frac{1}{2d+1}\mathcal{F}(d) \subseteq \mathcal{G}(d) \quad \text{if } d \text{ is odd.}$$

Proof. Let e_i, $i \in [d]$, be the standard basis. We define for $s \neq t$

$$e_{s,t}^{\pm} = \frac{1}{\sqrt{2}}(e_s \pm e_t),$$

$$\varphi_{s,t}^{\pm} = \frac{1}{\sqrt{2}}(e_s \pm ie_t).$$

It is easy to see that $X \mapsto (e_{s,t}^{\pm})^* X e_{s,t}^{\pm}$, $X \mapsto e_t^* X e_t$ and $X \mapsto (\varphi_{s,t}^{\pm})^* X \varphi_{s,t}^{\pm}$ are UCP maps from $\mathcal{M}_d(\mathbb{C})$ to $\mathbb{C}$.

Given $(Y_1, \ldots, Y_g) \in \mathcal{F}(d)$, it follows that, for $s \neq t \in [d]$,

$$\pm [e_s^* Y_i e_s]_{i \in [g]} \in \pm \mathcal{F}(1) \subseteq \mathcal{G}(1)$$

$$\pm \left[(e_{s,t}^+)^* Y_i e_{s,t}^+\right]_{i \in [g]} \in \pm \mathcal{F}(1) \subseteq \mathcal{G}(1)$$

$$\pm \left[(e_{s,t}^-)^* Y_i e_{s,t}^-\right]_{i \in [g]} \in \pm \mathcal{F}(1) \subseteq \mathcal{G}(1)$$

$$\pm \left[(\varphi_{s,t}^+)^* Y_i \varphi_{s,t}^+\right]_{i \in [g]} \in \pm \mathcal{F}(1) \subseteq \mathcal{G}(1)$$

$$\pm \left[(\varphi_{s,t}^-)^* Y_i \varphi_{s,t}^-\right]_{i \in [g]} \in \pm \mathcal{F}(1) \subseteq \mathcal{G}(1).$$

Note that we can recover the real and imaginary parts of the entries of an arbitrary matrix X from the quantities above:

$$X_{ss} = e_s^* X e_s, \quad \operatorname{Re} X_{st} = \frac{1}{2}\left[(e_{s,t}^+)^* X e_{s,t}^+ - (e_{s,t}^-)^* X e_{s,t}^-\right], \quad \operatorname{Im} X_{st} = \frac{1}{2}\left[(\varphi_{s,t}^-)^* X \varphi_{s,t}^- - (\varphi_{s,t}^+)^* X \varphi_{s,t}^+\right].$$

We conclude that, for all s, $t \in [d]$,

$$\pm (\operatorname{Re}(Y_1)_{s,t}, \ldots, \operatorname{Re}(Y_g)_{s,t}) \in \mathcal{G}(1),$$

$$\pm (\operatorname{Im}(Y_1)_{s,t}, \ldots, \operatorname{Im}(Y_g)_{s,t}) \in \mathcal{G}(1).$$

Let us consider the set

$$\mathcal{I} := \{(s,t) : s, t \in [d], s < t\}.$$

Without loss of generality, we can assume that the dimension d is even by going to $d+1$ if necessary and using Proposition 7.4. Then, we can partition $\mathcal{I}$ into $d-1$ subsets $\mathcal{J}_k$ of $d/2$ tuples each such that $\cup_{(s,t) \in \mathcal{J}_k}\{s,t\} = [d]$ for all $k \in [d-1]$, i.e., no index appears twice. This partitioning is possible, since it is equivalent to an edge-coloring of the complete graph K_d such that the edges of each color

form a perfect matching. The edges of each color k then correspond to $\mathcal{J}_k$. Such a coloring exists for even d, e.g., by Baranyai's theorem (see, e.g., [VLW01]).

Let us fix $k \in [d-1]$ for now and consider $\mathcal{J} = \{(s_1, t_1), \ldots, (s_{d/2}, t_{d/2})\}$. As matrix convex sets are closed under direct sums, it follows that

$$(25) \qquad (\mathrm{Re}(Y_i)_{s_1,t_1} \oplus \ldots \oplus \mathrm{Re}(Y_i)_{s_{d/2},t_{d/2}} \oplus -\mathrm{Re}(Y_i)_{s_1,t_1} \oplus \ldots \oplus -\mathrm{Re}(Y_i)_{s_{d/2},t_{d/2}})_{i \in [g]} \in \mathcal{G}(d),$$

$$(26) \qquad (-\mathrm{Im}(Y_i)_{s_1,t_1} \oplus \ldots \oplus -\mathrm{Im}(Y_i)_{s_{d/2},t_{d/2}} \oplus \mathrm{Im}(Y_i)_{s_1,t_1} \oplus \ldots \oplus \mathrm{Im}(Y_i)_{s_{d/2},t_{d/2}})_{i \in [g]} \in \mathcal{G}(d)$$

$$((Y_i)_{1,1} \oplus \ldots \oplus (Y_i)_{d,d})_{i \in [g]} \in \mathcal{G}(d).$$

Furthermore, matrix convex sets are closed under unitary conjugation. Let us define a unitary U that maps

$$e_j \mapsto e^+_{s_j,t_j} \quad \forall j \in [d/2]$$

$$e_{d/2+j} \mapsto e^-_{s_j,t_j} \quad \forall j \in [d/2]$$

and a unitary V that maps

$$e_j \mapsto \varphi^+_{s_j,t_j} \quad \forall j \in [d/2]$$

$$e_{d/2+j} \mapsto \varphi^-_{s_j,t_j} \quad \forall j \in [d/2]$$

Thus, conjugating Eq. (25) by U and Eq. (26) by V, we obtain that

$$\left(\sum_{j \in [d/2]} \mathrm{Re}(Y_\ell)_{s_j,t_j} (e_{s_j} e^*_{t_j} + e_{t_j} e^*_{s_j}) \right)_{\ell \in [g]} \in \mathcal{G}(d),$$

$$\left(\sum_{j \in [d/2]} i \, \mathrm{Im}(Y_\ell)_{s_j,t_j} (e_{s_j} e^*_{t_j} - e_{t_j} e^*_{s_j}) \right)_{\ell \in [g]} \in \mathcal{G}(d),$$

because $e^+_{s,t}(e^+_{s,t})^* - e^-_{s,t}(e^-_{s,t})^* = e_s e^*_t + e_t e^*_s$ and $\varphi^+_{s,t}(\varphi^+_{s,t})^* - \varphi^-_{s,t}(\varphi^-_{s,t})^* = i(-e_s e^*_t + e_t e^*_s)$. This construction can be repeated for all $\mathcal{J}_k$, $k \in [d-1]$. Taking uniform convex combinations of these elements, we infer that for even d

$$\frac{1}{2d-1}(Y_1, \ldots, Y_g) \in \mathcal{G}(d).$$

As $(Y_1, \ldots, Y_g) \in \mathcal{F}(d)$ was arbitrary, the assertion follows. $\square$

Note that the inclusion constant obtained for free spectrahedra in [BN18] is $2d$, which is slightly worse than our result for even d. Remark 7.5 in [BN18] indicates that our result for even d is optimal.

Corollary 7.10. *Let $\mathcal{P} \in \mathbb{R}^g$ be a* symmetric *polytope, i.e. $\mathcal{P} = -\mathcal{P}$. Then, for all $d \geq 2$,*

$$\frac{1}{\delta}\Big(\underbrace{1, 1, \ldots, 1}_{g \ times} \Big) \in \Delta_{\mathcal{P}}(d),$$

where

$$(27) \qquad \delta = \begin{cases} 2d-1 & \textit{if } d \textit{ is even} \\ 2d+1 & \textit{if } d \textit{ is odd.} \end{cases}$$

In order to apply the corollary above to the Birkhoff body $\mathcal{B}_N$, we first have to symmetrize it: we seek the best (i.e. largest) constant $s \in [0, 1]$ such that

$$(28) \qquad s(-\mathcal{B}_N) \subseteq \mathcal{B}_N \iff s \underbrace{\mathrm{conv}(-\mathcal{B}_N \cup \mathcal{B}_N)}_{=:\mathcal{P}} \subseteq \mathcal{B}_N.$$

In this way, we have, using Propositions 7.9 and 3.14:

$$X \in (\mathcal{B}_N)_{\max}(d) \implies X \in \mathcal{P}_{\max}(d) \implies \frac{1}{\delta} \cdot X \in \mathcal{P}_{\min}(d) \implies \frac{s}{\delta} \cdot X \in (\mathcal{B}_N)_{\min}(d).$$

Let us now compute the best constant s in Eq. (28). To do this, we have to find the largest value the facet inequalities from Proposition 5.4 attain on the negative of the extremal points of $\mathcal{B}_N$. The facets of $\mathcal{B}_N$ from Eqs. (16)-(19) correspond to the matrices

$$h_{ij} = \begin{cases} -N e_i e_j^* & \text{if } i, j \in [N-1] \\ N \sum_{k=1}^{N-1} e_i e_k^* & \text{if } i \in [N-1], j = N \\ N \sum_{k=1}^{N-1} e_k e_j^* & \text{if } j \in [N-1], i = N \\ -N \sum_{i,j=1}^{N-1} e_i e_j^* & \text{if } i, j = N. \end{cases}$$

It is easy to see that the maximum value of the quantities

$$\left\langle h_{ij}, -\left(P_\sigma^{(N-1)} - \frac{J}{N} \right) \right\rangle$$

is $N - 1$, for all $i, j \in [N]$ and σ a permutation of N elements. Hence, the largest s for which Eq. (28) holds is $s = 1/(N-1)$. We have thus the following corollary regarding the Birkhoff body.

Corollary 7.11. *For any $N \geq 2$ and $d \geq 1$,*

$$\frac{1}{(N-1)\delta} \in \Delta_{\mathcal{B}_N}(d),$$

where δ is the dimension dependent constant from Eq. (27).

7.3. **Inclusion constants from linear programming.** Recall that a polytope $\mathcal{P}$ has two equivalent representations: one in terms of vertices (the "V" representation) and one in terms of supporting hyperplanes (the "H" representation). To the k extreme points $v_1, \ldots, v_k \in \mathbb{R}^g$ of a polytope $\mathcal{P}$ we associate the matrix

$$V := \sum_{i=1}^{k} v_i e_i^* \in \mathcal{M}_{g \times k}(\mathbb{R})$$

having the v_i as columns. Similarly, if $\{x : \langle h_j, x \rangle \leq 1\}_{j \in [r]}$ are the halfspaces defining $\mathcal{P}$ (recall that $\mathcal{P}$ contains 0 in its interior), we introduce the matrix

$$H := \sum_{j=1}^{r} e_j^* h_j \in \mathcal{M}_{r \times g}(\mathbb{R})$$

having the h_j as columns. We extend these matrices by appending ones (we denote by $1_n \in \mathbb{R}^n$ the all-1 vector):

$$\hat{V} := \begin{bmatrix} V \\ 1_k^\top \end{bmatrix} \in \mathcal{M}_{(g+1) \times k}(\mathbb{R})$$

and

$$\hat{H} := \begin{bmatrix} -H & 1_r \end{bmatrix} \in \mathcal{M}_{r \times (g+1)}(\mathbb{R}).$$

The matrices $\hat{V}$ and $\hat{H}$ of the polytope $\mathcal{P}$ are associated to the *slack matrix* of $\mathcal{P}$ [Yan91]: $S_\mathcal{P} = \hat{H}\hat{V}$.

In the same vein, to a g-tuple A of self-adjoint operators, we associate the $(g+1)$-tuple

$$\hat{A} := \begin{bmatrix} A \\ I_d \end{bmatrix} \in \mathbb{R}^{g+1} \otimes \mathcal{M}_d^{sa}(\mathbb{C}).$$

Lemma 7.12. *Let g, $r \in \mathbb{N}$. A g-tuple of self-adjoint operators A are $\mathcal{P}$-operators iff $\hat{H}\hat{A}$ is entrywise positive semidefinite. In particular, for a vector $x \in \mathbb{R}^g$, $x \in \mathcal{P} \iff \hat{H}[\begin{smallmatrix} x \\ 1 \end{smallmatrix}] \in \mathbb{R}_+^r$.*

Proof. Requiring that the r blocks of $\hat{H}\hat{A}$ are positive semidefinite is equivalent to:

$$\forall j \in [r] \qquad \sum_{x=1}^{g} \hat{H}_{j,x} A_x + 1 \cdot I_d = -\sum_{x=1}^{g} h_j(x) A_x + I_d \geq 0,$$

which is precisely the condition that $A \in \mathcal{P}_{\max}(d)$ from Proposition 3.4. $\qquad\square$

Lemma 7.13. *Let g, $k \in \mathbb{N}$. A g-tuple of self-adjoint operators A are P-compatible iff there exists an entrywise positive k-tuple C such that $\hat{V}C = \hat{A}$.*

Proof. The equation for the extended vectors is equivalent to

$$VC = A \quad \text{and} \quad 1_k^\top C = I_d.$$

While the latter equation is equivalent to the normalization condition $\sum_{i=1}^{k} C_i = I_d$, the former is equivalent to

$$\sum_{i=1}^{k} v_i \otimes C_i = \sum_{x=1}^{g} e_x \otimes A_x,$$

where we recall that $v_1, \ldots, v_k \in \mathbb{R}^g$ are the extreme points of $\mathcal{P}$. This, in turn, gives

$$\forall x \in [g] \qquad A_x = \sum_{i=1}^{k} v_i(x) C_i$$

which is precisely the condition that $A \in \mathcal{P}_{\min}(d)$ from Definition 3.1. $\qquad\square$

Note that polytopes $P \subseteq \mathbb{R}^g$ containing zero in their interior have at least $g+1$ extreme points.

Theorem 7.14. *Let d, g, k, $r \in \mathbb{N}$. Given $s \in \mathbb{R}^g$, if there exists an entrywise non-negative matrix $T \in \mathcal{M}_{k \times r}(\mathbb{R}_+)$ such that*

$$\operatorname{diag}(s_1, s_2, \ldots, s_g, 1) =: \hat{D}_s = \hat{V} T \hat{H},$$

then $s \in \Delta_{\mathcal{P}}(d)$, for all $d \geq 1$.

Proof. Consider a scaling vector s satisfying the hypotheses of the statement, and let $A \in \mathbb{R}^g \otimes \mathcal{M}_d^{sa}(\mathbb{C})$ a g-tuple of P-operators. From the hypothesis, we have

$$\hat{V}T\hat{H} = \hat{D}_s \implies \hat{V}\underbrace{T\hat{H}\hat{A}}_{C} = \hat{D}_s\hat{A} = \widehat{s \cdot A}.$$

Since A are $\mathcal{P}$-operators, it follows from Lemma 7.12 that $\hat{H}\hat{A}$ is entrywise positive semidefinite, hence so is C. We conclude that $s \cdot A$ are $\mathcal{P}$-compatible by applying Lemma 7.13. $\qquad\square$

Note that the existence of the non-negative matrix T can be formulated as a *linear program*, so the theorem above provides a computationally tractable way to produce elements of the inclusion constants set

$$\bigcap_{d \geq 1} \Delta_{\mathcal{P}}(d).$$

Let us now apply this to the Birkhoff body $\mathcal{B}_N$. Recall that $\mathcal{B}_N$ has $k = N!$ extreme points and $r = N^2$ facets.

Proposition 7.15. *Let s be the constant vector with entries $1/(N-1)$ and $T = [\mathbb{1}_{\langle h_j, v_i \rangle \neq 1}/(N \cdot N!)]_{ij}$. Then $\hat{V}T\hat{H} = \hat{D}_s$, proving that $1/(N-1) \in \Delta_{\mathcal{B}_N}(d)$ for all $d \geq 1$.*

Proof. The proof follows from direct calculation. First, note that the condition $\hat{V}T\hat{H} = \hat{D}_s$ is equivalent to

$$-VTH = \mathrm{diag}(s) =: D_s, \quad 1_k^\top TH = 0, \quad VT1_r = 0, \quad \langle 1_k, T1_r \rangle = 1.$$

In other words, T is a bistochastic matrix, with marginals in the kernels of V, resp. $H^\top$, which satisfies $-VTH = D_s$. The columns of V are the extreme points of $\mathcal{B}_N$, and correspond to shifted truncated permutation matrices. They have coordinates

$$\forall \pi \in \mathcal{S}_N, \forall x, y \in [N-1], \qquad v_\pi(x, y) = -\frac{1}{N} + 1_{\pi(x)=y}.$$

The rows of H correspond to the inequalities (16)-(19) defining the Birkhoff body, and they can be indexed by $[N]^2$; they have coordinates

$$\begin{aligned}
\forall i, j \in [N-1] \qquad & h_{ij}(x, y) = -N 1_{i=x, j=y} \\
j = N, \forall i \in [N-1] \qquad & h_{ij}(x, y) = N 1_{i=x} \\
i = N, \forall j \in [N-1] \qquad & h_{ij}(x, y) = N 1_{j=y} \\
i, j = N \qquad & h_{ij}(x, y) = -N.
\end{aligned}$$

In particular, we have the crucial relation

$$\langle h_{ij}, v_\pi \rangle = \begin{cases} 1 & \text{if } \pi(i) \neq j \\ 1 - N & \text{if } \pi(i) = j. \end{cases}$$

It follows that a given extreme point v_π belongs to all the facets $\{x \in \mathbb{R}^g : \langle h_{ij}, x \rangle = 1\}$ such that $\pi(i) \neq j$; there are $N^2 - N$ such facets. In other words, the π-th row of the matrix T has precisely N non-zero entries. It follows that the row sums of T are all equal to $1/N!$. The condition

$$VT1_r = \frac{1}{N!}V1_{N^2} = 0$$

follows from the fact that the average of the extreme points v_π is 0; equivalently, this can be seen to follow from the fact that the average of the permutation matrices is the matrix J/N. A similar reasoning yields the condition $1_{N!}^\top TH = 0$. Finally, the main condition $-VTH = I/(N-1)$ follows from a simple (but tedious) analysis of the expression:

$$(VTH)_{(x_1, y_1),(x_2, y_2)} = \sum_{\pi \in \mathcal{S}_N} \sum_{i,j=1}^{N} \frac{1_{\pi(i)=j}}{N \cdot N!} v_\pi(x_1, y_1) h_{ij}(x_2, y_2).$$

$\square$

Remark 7.16. *Note how the Birkhoff body inclusion constant $1/(N-1)$ derived in this section is better than the ones obtained using the comparison technique (Corollary 7.8) and the symmetrization technique (Corollary 7.11).*

Corollary 7.17. *Let N, $d \in \mathbb{N}$. For any magic square $A \in \mathcal{M}_N(\mathcal{M}_d^{sa}(\mathbb{C}))$, the convex combination*

$$(29) \qquad\qquad B := \frac{1}{N-1}A + \frac{N-2}{N-1} \cdot \frac{J_N}{N}$$

is a semiclassical magic square.

Proof. This follows from Proposition 7.15 above and from the identity

$$B^{(N-1)} - \frac{J_{N-1}}{N} = \frac{1}{N-1}\left(A^{(N-1)} - \frac{J_{N-1}}{N}\right).$$

$\square$

In [DlCDN20, Theorem 12-(ii)], the authors have shown that a magic square A having the property that

$$(30) \qquad \forall \pi \in \mathcal{S}_N \qquad \sum_{k=1}^{N} A_{k,\pi(k)} \geq \frac{N-2}{N-1} I_d$$

is semiclassical. This fact can obtained from the following rewriting of A:

$$A = \sum_{i,j=1}^{N} e_i e_j^* \otimes A_{ij}$$

$$= \sum_{\pi \in S_N} \pi \otimes \frac{1}{(N-2)!N} \left[\sum_{k=1}^{N} A_{k,\pi(k)} - \frac{N-2}{N-1} I_d \right].$$

Indeed, if the condition Eq. (30) is satisfied, then one can take

$$C_\pi := \frac{1}{(N-2)!N} \left[\sum_{k=1}^{N} A_{k,\pi(k)} - \frac{N-2}{N-1} I_d \right] \geq 0$$

as the POVM certifying semiclassicality in Definition 5.1. We would like to end this section by emphasizing the close relation between this condition and the one from Corollary 7.17. A magic square B can we written as in Eq. (29) if and only if $B_{ij} \geq (N-2)/(N(N-1))I_d$ for all $i,j \in [N]$. This condition implies the one in Eq. (30), but the converse is not true. Indeed, consider the bistochastic matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

which clearly satisfies Eq. (30) but which cannot be written as in Eq. (29) because of its 0 entries.

Finally, let us apply Theorem 7.14 to the case of the pyramid polytope from Eq. (23) corresponding to two POVMs with three outcomes sharing one effect. Recall from Section 6 that this polytope $\mathcal{P}$ is a pyramid with a square basis, having defining matrices $\hat{V}$ and $\hat{H}$ given respectively by

$$\hat{V} = \begin{bmatrix} 2/3 & -1/3 & -1/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & -1/3 & 2/3 & 2/3 \\ -1/3 & -1/3 & 2/3 & -1/3 & 2/3 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \hat{H} = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ -3 & -3 & 0 & 1 \\ -3 & 0 & -3 & 1 \end{bmatrix}.$$

A simple calculation shows that taking

$$T := \frac{1}{30} \begin{bmatrix} 6 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 2 & 2 \\ 1 & 0 & 2 & 2 & 0 \\ 1 & 2 & 0 & 0 & 2 \\ 1 & 2 & 2 & 0 & 0 \end{bmatrix}$$

yields $\hat{V}T\hat{H} = \mathrm{diag}(2/5, 2/5, 2/5, 1)$, showing that $(2/5, 2/5, 2/5) \in \Delta_{\mathcal{P}}(d)$ for all dimensions d. This implies that the two POVMs $(\frac{2}{5}A + \frac{1}{5}I_d, \frac{2}{5}B + \frac{1}{5}I_d, \frac{3}{5}I_d - \frac{2}{5}A - \frac{2}{5}B)$ and $(\frac{2}{5}A + \frac{1}{5}I_d, \frac{2}{5}C + \frac{1}{5}I_d, \frac{3}{5}I_d - \frac{2}{5}A - \frac{2}{5}C)$ are compatible. In addition, the proof of Theorem 7.14 shows that the joint measurement has the form

Q_1	0	0	$= \frac{2}{5}A + \frac{1}{5}I_d$
0	Q_5	Q_4	$= \frac{2}{5}B + \frac{1}{5}I_d$
0	Q_3	Q_2	$= \frac{3}{5}I_d - \frac{2}{5}A - \frac{2}{5}B$
$= \frac{2}{5}A + \frac{1}{5}I_d$	$= \frac{2}{5}C + \frac{1}{5}I_d$	$= \frac{3}{5}I_d - \frac{2}{5}A - \frac{2}{5}C$	

with elements

$$Q_1 = \frac{2}{5}A + \frac{1}{3}I_d,$$

$$Q_2 = -\frac{3}{10}A - \frac{1}{5}B - \frac{1}{5}C + \frac{2}{5}I_d,$$

$$Q_3 = -\frac{1}{10}A - \frac{1}{5}B + \frac{1}{5}C + \frac{1}{5}I_d,$$

$$Q_4 = -\frac{1}{10}A + \frac{1}{5}B - \frac{1}{5}C + \frac{1}{5}I_d,$$

$$Q_5 = \frac{1}{10}A + \frac{1}{5}B + \frac{1}{5}C.$$

For the example in Eq. (24), this means that after adding sufficient noise, the joint POVM has elements

Q_1	0	0	$= \frac{2}{5}I_2$
0	Q_5	Q_4	$= \frac{1}{5}e_1 e_1^* + \frac{1}{5}I_2$
0	Q_3	Q_2	$= \frac{1}{5}e_2 e_2^* + \frac{1}{5}I_2$
$= \frac{2}{5}I_2$	$= \frac{1}{5}f_1 f_1^* + \frac{1}{5}I_2$	$= \frac{1}{5}f_2 f_2^* + \frac{1}{5}I_2$	

where

$$Q_1 = \frac{2}{5}I_2,$$

$$Q_2 = \frac{1}{10}\left(\frac{1}{2}I_2 + f_2 f_2^* + e_2 e_2^*\right),$$

$$Q_3 = \frac{1}{10}\left(\frac{1}{2}I_2 + f_1 f_1^* + e_2 e_2^*\right),$$

$$Q_4 = \frac{1}{10}\left(\frac{1}{2}I_2 + f_2 f_2^* + e_1 e_1^*\right),$$

$$Q_5 = \frac{1}{10}\left(\frac{1}{2}I_2 + f_1 f_1^* + e_1 e_1^*\right).$$

Again, we have written e_1, e_2 for the standard basis vectors in $\mathbb{C}^2$ and $f_1 = 1/\sqrt{2}(e_1 + e_2)$, $f_2 = 1/\sqrt{2}(e_1 - e_2)$. In this case, it is easy to see that the Q_i indeed form a POVM. However, it can be checked with a semidefinite program inspired by [WPGF09] that we could have taken $s = 3/4 > 2/5$ in this case. Hence the question remains open if $s = 2/5$ is optimal in this setup. Note that if s is an inclusion constant for the polytope with square basis, $s \leq \frac{1}{\sqrt{2}}$ is a necessary requirement by the results in [BN18], since for $A = 0$, the problem reduces to the compatibility of two dichotomic POVMs.

Acknowledgements.

I.N. was supported by the ANR projects ESQuisses, grant number ANR-20-CE47-0014-01 and STARS, grant number ANR-20-CE40-0008, and by the PHC program *Star* (Applications of random matrix theory and abstract harmonic analysis to quantum information theory). S.S. has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101030346.

REFERENCES

[AB06]   Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer, 3rd edition, 2006. 11

[AFLS15]   Antonio Acín, Tobias Fritz, Anthony Leverrier, and Ana Belén Sainz. A combinatorial approach to nonlocality and contextuality. *Communications in Mathematical Physics*, 334:533–628, 2015. 21

[ALPP21]   Guillaume Aubrun, Ludovico Lami, Carlos Palazuelos, and Martin Plávala. Entangleability of cones. *Geometric and Functional Analysis*, 31:181–205, 2021. 12

[Bar02]   Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. 10

[BBC07]   Teodor Banica, Julien Bichon, and Benoît Collins. Quantum permutation groups: a survey. *Banach Center Publications*, 78(1):13–34, 2007. 17

[BCP+14]   Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell non-locality. *Reviews of Modern Physics*, 86:419–478, 2014. 3

[BG77]   Richard A. Brualdi and Peter M. Gibson. Convex polyhedra of doubly stochastic matrices. I. Applications of the permanent function. *Journal of Combinatorial Theory, Series A*, 22(2):194–230, 1977. 18

[BJN22]   Andreas Bluhm, Anna Jenčová, and Ion Nechita. Incompatibility in general probabilistic theories, generalized spectrahedra, and tensor norms. *Communications in Mathematical Physics*, 393(3):1125–1198, 2022. 16, 28

[BN18]   Andreas Bluhm and Ion Nechita. Joint measurability of quantum effects and the matrix diamond. *Journal of Mathematical Physics*, 59(11):112202, 2018. 2, 4, 13, 26, 29, 30, 35

[BN20]   Andreas Bluhm and Ion Nechita. Compatibility of quantum measurements and inclusion constants for the matrix jewel. *SIAM Journal on Applied Algebra and Geometry*, 4(2):255–296, 2020. 2, 4, 5, 13, 16, 26, 28

[BN22a]   Andreas Bluhm and Ion Nechita. Maximal violation of steering inequalities and the matrix cube. *Quantum*, 6:656, 2022. 26

[BN22b]   Andreas Bluhm and Ion Nechita. A tensor norm approach to quantum compatibility. *Journal of Mathematical Physics*, 63:062201, 2022. 2, 15, 16

[Boh28]   Niels Bohr. The quantum postulate and the recent development of atomic theory. *Nature*, 121(3050):580–590, 1928. 3

[Bre97]   David D. Bremner. On the complexity of vertex and facet enumeration for complex polytopes. *Ph.D. thesis, School of Computer Science, McGill University, Monréal, Canada*, 1997. 5

[Bro08]   Efim M. Bronstein. Approximation of convex sets by polytopes. *Journal of Mathematical Sciences*, 153(6):727–762, 2008. 11

[DDOSS17]   Kenneth R. Davidson, Adam Dor-On, Orr Moshe Shalit, and Baruch Solel. Dilations, inclusions of matrix convex sets, and completely positive maps. *International Mathematics Research Notices*, 2017(13):4069–4130, 2017. 4, 6, 12

[DlCDN20]   Gemma De las Cuevas, Tom Drescher, and Tim Netzer. Quantum magic squares: Dilations and their limitations. *Journal of Mathematical Physics*, 61(11):111704, 2020. 3, 17, 34

[Fin82]   Arthur Fine. Hidden variables, joint probability, and the Bell inequalities. *Physical Review Letters*, 48(5):291–295, 1982. 3

[GB19]   Leonardo Guerini and Alexandre Baraviera. Joint measurability meets Birkhoff-von Neumann's theorem. *arXiv preprint arXiv:1809.07366v3*, 2019. 3, 4, 19, 20, 23

[GJK04]   Eugene Gutkin, Edmond A Jonckheere, and Michael Karow. Convexity of the joint numerical range: topological and differential geometric viewpoints. *Linear Algebra and its Applications*, 376:143–171, 2004. 5

[Hei27]   Werner Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3):172–198, 1927. 3

[HKMS19]   J. William Helton, Igor Klep, Scott McCullough, and Markus Schweighofer. Dilations, linear matrix inequalities, the matrix cube problem and beta distributions. *Memoirs of the American Mathematical Society*, 257(1232), 2019. 4, 29

[HKR15]   Teiko Heinosaari, Jukka Kiukas, and Daniel Reitzner. Noise robustness of the incompatibility of quantum measurements. *Physical Review A*, 92:022115, 2015. 3

[HMZ16]   Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. An invitation to quantum incompatibility. *Journal of Physics A: Mathematical and Theoretical*, 49(12):123001, 2016. 7, 16, 27

[HZ11]   Teiko Heinosaari and Mário Ziman. *The Mathematical Language of Quantum Theory*. Cambridge University Press, 2011. 7

[Jen18]   Anna Jenčová. Incompatible measurements in a class of general probabilistic theories. *Physical Review A*, 98(1):012133, 2018. 16, 28

[Lam18]   Ludovico Lami. Non-classical correlations in quantum mechanics and beyond. *PhD thesis. arXiv preprint arXiv:1803.02902*, 2018. 8

[LP00]   Chi-Kwong Li and Yiu-Tung Poon. Convexity of the joint numerical range. *SIAM Journal on Matrix Analysis and Applications*, 21(2):668–678, 2000. 5

[LPW20]   Chi-Kwong Li, Yiu-Tung Poon, and Ya-Shu Wang. Joint numerical ranges and commutativity of matrices. *Journal of Mathematical Analysis and Applications*, 491(1):124310, 2020. 12

[PP21]   Benjamin Passer and Vern I. Paulsen. Matrix range characterizations of operator system properties. *Journal of Operator Theory*, 85:547–568, 2021. 7

[PSS18]   Benjamin Passer, Orr Moshe Shalit, and Baruch Solel. Minimal and maximal matrix convex sets. *Journal of Functional Analysis*, 274:3197–3253, 2018. 6, 12

[Shu74]   Frederic W. Shultz. A characterization of state spaces of orthomodular lattices. *Journal of Combinatorial Theory, Series A*, 17(3):317–328, 1974. 22

[VLW01]   Jacobus Hendricus Van Lint and Richard Michael Wilson. *A course in combinatorics*. Cambridge University Press, second edition, 2001. 30

[Wat18]   John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. 7

[WPGF09] Michael M. Wolf, David Pérez-García, and Carlos Fernández. Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Physical Review Letters*, 103:230402, 2009. 35

[Yan91]   Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991. 31

[Zie12]   Günter M Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012. 17

Email address: andreas.bluhm@univ-grenoble-alpes.fr

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

Email address: ion.nechita@univ-tlse3.fr

Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, France

Email address: sisc@math.ku.dk

QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark