Birkhoff–von Neumann’s theorem, doubly normalized tensors, and joint measurability

Leonardo Guerini\textsuperscript{a,b} and Alexandre Baraviera\textsuperscript{c}

\textsuperscript{a}International Centre for Theoretical Physics – South American, Institute for Fundamental Research & Instituto de Física Teórica – UNESP, São Paulo, Brazil; \textsuperscript{b}Centro de Ciências Naturais e Exatas – Universidade Federal de Santa Maria, Santa Maria, Brazil; \textsuperscript{c}Instituto de Matemática e Estatística – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

ABSTRACT
Quantum measurements can be interpreted as a generalization of probability vectors, in which non-negative real numbers are replaced by positive semi-definite operators. We extrapolate this analogy to define a generalization of doubly stochastic matrices that we call doubly normalized tensors (DNTs), and investigate a corresponding version of Birkhoff–von Neumann’s theorem, which states that permutations are the extremal points of the set of doubly stochastic matrices. We prove that joint measurability appears naturally as a mathematical feature of DNTs in this context and that this feature is necessary and sufficient for a characterization similar to Birkhoff–von Neumann’s. Conversely, we also show that DNTs arise from a particular instance of a joint measurability problem, remarking the relevance of this quantum-theoretical property in general operator theory.

1. Introduction
The purpose of this text is to establish some properties of some families of matrices, starting from matrices whose elements form probability vectors. For pedagogical reasons, we start from basic definitions so that some analogies are explicit. A probability vector of $n$ components $\mathbf{p}$ is given by

$$\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n; \quad p_i \geq 0, \quad \sum_i p_i = 1. \quad (1)$$

Let us denote the set of $n$-component probability vectors by $\mathbb{S}_n$. An $n \times n$ doubly stochastic matrix is a matrix $D \in \mathbb{R}^{n \times n}$ such that each column and each row is a probability vector,

$$D = \begin{bmatrix}
p_{11} & \cdots & p_{1n} \\
\vdots & \ddots & \vdots \\
p_{n1} & \cdots & p_{nn}
\end{bmatrix}; \quad (2a)$$

CONTACT Alexandre Baraviera at baraviera@gmail.com Instituto de Matemática e Estatística – Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500 Porto Alegre, Brazil

© 2022 Informa UK Limited, trading as Taylor & Francis Group
\( \mathbf{c}^1 = (p_{i1})_i, \ldots, \mathbf{c}^n = (p_{in})_i \in \mathcal{S}_n, \) (2b)

\( \mathbf{r}^1 = (p_{1j})_j, \ldots, \mathbf{r}^n = (p_{nj})_j \in \mathcal{S}_n. \) (2c)

The set of doubly stochastic matrices is convex. Among the most important results on this topic lies Birkhoff–von Neumann’s (BvN’s) Theorem [1], which states that the extremal points of this set are the permutation matrices. This implies that every doubly stochastic matrix \( D \) can be written as a convex combination of permutation matrices \( \Pi_i, \)

\[
D = \sum_{l=1}^{n!} q_l \Pi_l,
\] (3)

where \( \mathbf{q} = (q_l)_l \in \mathcal{S}_{n!} \). BvN’s Theorem can be generalized in some distinct directions: for example, when one considers multistochastic tensors it is still possible to describe some properties of the extreme points of the set [2]. Here we present one such generalization motivated by its interpretation in quantum mechanics.

Quantum theory is inherently probabilistic, in the sense that the result of a measurement on a quantum system cannot be predicted deterministically; we rather have to cope with a probability distribution over the set of possible outcomes. A quantum system is described by a Hilbert space and its states are given by density matrices, which are positive semi-definite operators of unit trace [3]. Due to this positivity and normalization features, the density matrix is sometimes interpreted as a quantum analogue of a probability vector, from which we extract information about the system.

On the other hand, a quantum measurement of \( n \) outcomes acting on a Hilbert space \( \mathcal{H} \) is modelled by a positive-operator valued measure (POVM) \( \mathbf{A} \), described by

\[
\mathbf{A} = (A_1, \ldots, A_n) \in \mathcal{L}(\mathcal{H})^n; \quad A_i \succeq 0, \quad \sum_i A_i = I,
\] (4)

where \( \mathcal{L}(\mathcal{H}) \) is the space of linear operators acting in \( \mathcal{H} \), \( \succeq \) is the partial order that defines positive semi-definite operators and \( I \) is the identity operator. Notice that, in analogy to (1), a POVM is an operator-version of a probability vector, obtained by enlarging the dimension of the entries. Hence, there is an even more direct analogy between probability vectors and quantum measurements, given by the natural connection between non-negative real numbers and positive semi-definite Hermitian operators together with a normalization constraint on both of them.

In this work we explore this parallel, investigating the correspondence of standard features of probability vectors and their implications in terms of quantum theory. We introduce an operator-version of doubly stochastic matrices where probabilities are substituted by quantum measurements, which we call doubly normalized tensors (DNTs). Following BvN’s Theorem, our goal is to study the extremal points of such a set. While a complete characterization eludes us, we find, perhaps surprisingly, that in this richer context not all DNTs admit a decomposition in terms of permutations, as their counterparts do. We prove that such decomposition exists if and only if the involved POVMs are jointly measurable [4] – a property that plays a central role in several quantum information topics, such as Bell nonlocality [5] and uncertainty relations [6].
Finally, we present yet another connection to joint measurability: Starting from quantum theory and studying the plurality of joint measurements, we see the emergence of doubly normalized tensors arising naturally in this context.

1.1. Related results

The problem of considering matrices whose entries are not necessarily numbers is not new and appeared recently in some contexts involving quantum objects. For example, in the so-called quantum Latin squares, the entries are vectors and each column/row is an orthonormal basis of a certain complex vector space, see [7]–[9]. Those are applied, among other topics, in quantum codes and in the construction of unitary error bases. Note that each orthonormal basis correspond to a projective measurement, hence quantum Latin squares can be seen as corresponding to projective DNTs.

Moving closer to the case considered here, the appendix of Ref. [10] also considers DNTs, called there block-bistochastic matrices. They present an algorithm that starts with a matrix whose entries are positive matrices and outputs another one whose sum of rows/columns is $\epsilon$—close to the identity matrix.

In a similar direction, Ref. [11] introduced so-called quantum magic squares, which in our language translates as generalized DNTs whose entries are positive elements of some algebra. Whenever this is taken as the matrix algebra over $\mathbb{C}$, the quantum magic squares reduce exactly to a DNT. Also inspired by BvN’s Theorem, in the context of operator algebras, Ref. [11] shows how quantum magic squares relate to the matrix convex hull so-called quantum permutation matrices (or, in our terms, projective DNTs). As these are also compared to decompositions in permutation tensors, there are various connections we can make. We comment on how this relates to our results in the last section of the paper.

At last, another quantum-motivated generalization of BvN’s Theorem is given in terms of quantum channels [12]. If a quantum state can be seen as an analogue to a probability vector, one can argue that doubly stochastic matrices correspond unital quantum channel. Thus, one can analyse and prove close connections between the structure of these channels and of mixtures of unitary channels.

2. Doubly normalized tensors

Let us denote the set of $n$-outcome quantum measurements on $\mathcal{H}$ by $\mathbb{P}_n$. We now present the main object of this work.

**Definition 2.1:** A doubly normalized tensor of positive semi-definite operators (DNT, for short) is a tensor $A \in \mathcal{L}(\mathcal{H})^{n \times n}$ which, in analogy to (2a), each element is positive semi-definite, and each column and each row sums up to the identity,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}. \quad (5)$$
This is to say that each row \( R^{(i)} \) and each column \( C^{(j)} \) of \( A \) is a \( n \)-outcome quantum measurement,

\[
R^{1} = (A_{1})_{j}, \ldots, R^{n} = (A_{n})_{j} \in \mathbb{P}_{n}, \tag{6a}
\]

\[
C^{1} = (A_{i})_{i}, \ldots, C^{n} = (A_{in})_{i} \in \mathbb{P}_{n}, \tag{6b}
\]

The set of DNTs (related to \( n \)-outcome quantum measurements) over a Hilbert space \( \mathcal{H} \) is denoted by \( D_{n}(\mathcal{H}) \), or simply \( D_{n} \) whenever the space is obvious. Notice that we can write

\[
\mathcal{A} = \sum_{i,j=1}^{n} E_{i,j} \otimes A_{ij}, \tag{7}
\]

where \( E_{i,j} = |i\rangle\langle j| \) is the \( n \times n \) matrix with entries \( e_{ab} = \delta_{a,i}\delta_{b,j} \), that belongs to the canonical basis of \( \mathbb{P}^{n\times n} \). A DNT will be denoted by \( \mathcal{A} = [A_{ij}] \).

Let us investigate the extremal elements of \( D_{n} \). These form the set \( \text{Ext}(D_{n}) \), where the convex combinations are considered element-wise, i.e. \( \mathcal{A} = \alpha \mathcal{B} + (1 - \alpha) \mathcal{C} \) is the DNT \( \mathcal{A} = [\alpha B_{ij} + (1 - \alpha) C_{ij}] \). It is straightforward to see that a non-trivial convex decomposition like this implies a non-trivial decomposition for at least one row-POVM of \( \mathcal{A} \).

Therefore, if \( n - 1 \) row-POVMs of \( \mathcal{A} \) are extremal POVMs, say the \( n - 1 \) first ones, then \( \mathcal{A} \in \text{Ext}(D_{n}) \), since

\[
\begin{bmatrix}
A_{11} & \ldots & A_{1n} \\
\vdots & \ddots & \vdots \\
A_{n1} & \ldots & A_{nn}
\end{bmatrix} = \alpha
\begin{bmatrix}
B_{11} & \ldots & B_{1n} \\
\vdots & \ddots & \vdots \\
B_{n1} & \ldots & B_{nn}
\end{bmatrix} + (1 - \alpha)
\begin{bmatrix}
C_{11} & \ldots & C_{1n} \\
\vdots & \ddots & \vdots \\
C_{n1} & \ldots & C_{nn}
\end{bmatrix} \tag{8}
\]

implies that \( [A_{i1} \ldots A_{in}] = [B_{i1} \ldots B_{in}] = [C_{i1} \ldots C_{in}] \) for \( i = 1, \ldots, n - 1 \) and the normalization of the columns implies \( A_{nj} = B_{nj} = C_{nj} \) for \( j = 1, \ldots, n \). An analogous argument shows the same is true for column-POVMs.

The next example shows that the converse does not hold: there are extremal DNTs with more than one non-extremal rows.

**Example 2.1:** Take \( (Q_{1}, Q_{2}, Q_{3}) \) an extremal POVM such that \( (Q_{2}, Q_{1} + Q_{3}, 0) \) and \( (Q_{3}, 0, Q_{1} + Q_{2}) \) are not extremal (one such example is defined in Equation (27)). Then it is simple to check that

\[
\mathcal{A} = \begin{bmatrix}
Q_{1} & Q_{2} & Q_{3} \\
Q_{2} & Q_{1} + Q_{3} & 0 \\
Q_{3} & 0 & Q_{1} + Q_{2}
\end{bmatrix} = \alpha \mathcal{B} + (1 - \alpha) \mathcal{C} \tag{9}
\]

implies \( \mathcal{A} = \mathcal{B} = \mathcal{C} \). Hence, \( \mathcal{A} \) is an extremal DNT, with two rows and two columns that are not extremal.

At this point, one might feel tempted to claim that a single extremal row and a single extremal column are sufficient for guaranteeing DNT-extremality. The next example shows otherwise.
Table 1. Summary of relations between the extremality of rows, columns, and the whole DNT.

| Rows          | Columns          | DNT                        |
|---------------|------------------|----------------------------|
| \( n - 1 \) ext. in \( P_n \) | \( n - 1 \) ext. in \( P_n \) | \( \Rightarrow \) ext. in \( \mathbb{D}_n \) Equation (8) |
| all ext. in \( P_n \) | all ext. in \( P_n \) | \( \Leftrightarrow \) ext. in \( \mathbb{D}_n \) Example (2.1) |
| one ext. in \( P_n \) ^\wedge | one ext. in \( P_n \) | \( \Leftrightarrow \) ext. in \( \mathbb{D}_n \) Example (2.1) |
| all ext. in \( P_n \) | all ext. in \( P_n \) | \( \iff \) ext. in \( \mathbb{D}_n \) Example (2.2) |

Example 2.2: Take \((A, B, 0)\) an extremal POVM. Then

\[
\begin{bmatrix}
A & B & 0 \\
B & \frac{1}{2}A & \frac{1}{2}A \\
0 & \frac{1}{2}A & \frac{1}{2}A + B
\end{bmatrix}
= \frac{1}{2}
\begin{bmatrix}
A & B & 0 \\
B & A & 0 \\
0 & 0 & I
\end{bmatrix}
+ \frac{1}{2}
\begin{bmatrix}
A & B & 0 \\
B & 0 & A \\
0 & A & B
\end{bmatrix}
\] (10)

In this case, the DNT has one row and one column that are extremal, but it is a convex combination of two distinct DNTs and hence not extremal.

These simple examples, organized in Table 1, reveal that the inner relations of extremality between a DNT and its rows and columns are not obvious. In fact, it is not even clear whether the extremality of every row is equivalent to the extremality of every column (assuming an example of such a DNT actually exists). This already shows a contrast with Birkhoff–von Neumann’s result: there, the rows/columns of extremal doubly stochastic matrices are extremal probability vectors. Hence, we are left to search for a generalization in another direction.

3. Permutation tensors

Inspired by BvN’s decomposition (3), we might conjecture that the extremal points of the set of DNTs are the operator-versions of permutations, such as

\[
\begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}
\leftrightarrow
\begin{bmatrix}
0 & I \\
I & 0
\end{bmatrix}
\] (11)

Nevertheless, convex combinations of DNTs like the one in the right-hand side above yield DNTs where all entries are proportional to \(I\). Clearly, this does not recover the whole set, as for any operator \(0 \leq A \leq I\) (a rank-one projector, for example) we can construct the DNT

\[
\begin{bmatrix}
A & I - A \\
I - A & A
\end{bmatrix}
\] (12)

Notice this DNT can be written as

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \otimes A + \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix} \otimes (I - A),
\] (13)

which suggests us to consider combinations of permutation matrices in which the convex weights are associated to operators that form a quantum measurement (instead of scalars that form a probability vector) attached to each term via tensor product. We formalize this idea in the following way.
**Definition 3.1:** Let \( \{ \Pi_l \} \) be the set of permutations \( n \times n \) and consider a quantum measurement \( Q = (Q_l) \in \mathbb{P}_{n!} \). We call

\[
\sum_{l=1}^{n!} \Pi_l \otimes Q_l,
\]

(14)
a decomposition into permutation tensors.

In fact, due to normalization, DNT (12) is the most general one for \( n = 2 \). In this case, expression (13) attests that every DNT in \( \mathbb{D}_2 \) can be decomposed into permutation tensors, and clearly its extremality is equivalent to the extremality of the coefficient POVM \( (A, I - A) \). Since extremal dichotomic POVMs are projective [13], we have the following.

**Theorem 3.1:** Any \( 2 \times 2 \) doubly normalized tensor \( A \in \mathbb{D}_2(\mathcal{H}) \) can be decomposed into permutation tensors, \( A = \sum_{l=1}^{n!} \Pi_l \otimes Q_l \), and the extremal points of this set are given by projective dichotomic coefficient measurements \( Q = (Q_l) \in \mathbb{P}_{n!} \).

Our goal is to study the relation between DNTs and decompositions into permutation tensors for arbitrary \( n \). We start by proving that every combination (14) is a DNT.

**Proposition 3.2:** If \( B = \sum_{l=1}^{n!} \Pi_l \otimes Q_l \), where \( Q = (Q_l) \) is a POVM and \( \{ \Pi_l \} \) is the set of \( n \)-dimensional permutation matrices, then \( B \) is a DNT.

**Proof:** Note that the permutation \( \Pi_l \) acting on the canonical basis can be written as

\[
\Pi_l = \sum_i E_{\Pi_l(i),i}.
\]

(15)

We have

\[
B = \sum_l \left( \sum_i E_{\Pi_l(i),i} \right) \otimes Q_l
\]

(16a)

\[
= \sum_{ab} E_{ab} \otimes \sum_{l:a=\Pi_l(b)} Q_l.
\]

(16b)

Thus, defining

\[
B_{ab} := \sum_{l:a=\Pi_l(b)} Q_l,
\]

(17)

it satisfies \( B = [B_{ab}] \) and \( B_{ab} \geq 0 \). Also, for all \( a, b \) we have

\[
\sum_b B_{ab} = \sum_a B_{ab} = \sum_l Q_l = I,
\]

(18)

hence each row and column of \( B \) is in \( \mathbb{P}_{n!} \).

As aforementioned, DNTs that admit a decomposition into permutation tensors have their extremality decided by the extremality of its coefficient-POVM, and the characterizations of extremal points of this set are well-known [14]. The question left is whether
tensors $\mathcal{B}$ that admit a decomposition into permutation tensors are the only possible DNTs. To address this question we need to introduce the notion of joint measurability [4].

**Definition 3.2:** A set of $m$ quantum measurements $\{A^{(1)}, \ldots, A^{(m)}\} \subset \mathbb{P}_n^m$ is said to be jointly measurable if there exists a so-called mother measurement $M \in \mathbb{P}_m$, for some positive integer $m$, from which we can recover each measurement of the set by post-processing it, i.e.

$$A^{(i)}_j = \sum_k \mu(j|i, k)M_k.$$  \hspace{1cm} (19)

where $\mu(\cdot|i, k)$ is a probability distribution conditioned on which measurement $A^{(i)}$ we wish to obtain and on each operator $M_k$ of $M$, and therefore satisfies $\mu(j|i, k) \geq 0$ and $\sum_j \mu(j|i, k) = 1$, $\forall i, k$.

This property expresses the fact that for any given quantum system, one can determine an outcome for each $A^{(i)}$ by performing $M$ and, depending on the outcome $k$ obtained, flip a coin $\mu(\cdot|i, k)$. In terms of clean measurements and the pseudo-order of post-processing, this is equivalent to say that there is a measurement that is cleaner than every element of the set [15]. Importantly, the joint measurability of a finite set of $d$-dimensional measurements can be computationally decided in an efficient way by means of semi-definite programming (SDP) [16].

We can now present our main result.

**Theorem 3.3:** Let $\mathcal{A} \in \mathbb{D}_n(\mathcal{H})$ be a DNT. Then $\mathcal{A} = \sum^m_{l=1} \Pi_l \otimes Q_l$, with $Q_l \in \mathbb{P}_n^d$, if and only if the set $R = \{R^{(1)}, \ldots, R^{(m)}\}$ of its row-measurements is jointly measurable.

**Proof:** First, assume that $\mathcal{A} = \sum^m_{l=1} \Pi_l \otimes Q_l$. $\mathcal{A}$ is a DNT for Proposition 3.2, hence we need only to prove that its rows are jointly measurable.

Defining the operators $A_{\alpha\beta}$ as in (17), the row-measurements are given by $R^{(i)} = (A_{ij})_j$. Then the coefficient measurement $Q$ is also a mother measurement for $\mathcal{R}$, since for any $i, j$,

$$A_{ij} = \sum_{l = \Pi_l(j)} Q_l = \sum_l \delta_{i, \Pi_l(j)} Q_l,$$  \hspace{1cm} (20)

so we can recover the $j$-th element of the row-measurement $R^{(i)}$ by post-processing $Q$ with $\mu(j|i, l) = \delta_{i, \Pi_l(j)}$. Therefore the rows of $\mathcal{A}$ are jointly measurable.

Let us now prove the converse. Since the set of jointly measurable POVMs is convex, it suffices to prove that the extremal points of this set admit a decomposition into permutation tensors. Hence, let $R$ be an extremal set of jointly measurable POVMs, and $M$ be a mother measurement for it associated to a deterministic post-processing, namely to a marginalization (which can be done without loss of generality [17]). By Corollary 5 of Ref. [18], this implies that $M$ is an extremal POVM, and, according to Corollary 7 of Ref. [14], its non-vanishing elements are linearly independent. For joint measurability, there exists a
post-processing $\mu$ such that

$$A_{ij} = R_j^{(i)} = \sum_k \mu(j|i, k)M_k. \quad (21)$$

Since $\sum_i A_{ij} = I$, we have that

$$\sum_{k,i} \mu(j|i, k)M_k = I = \sum_k M_k, \quad \forall j. \quad (22)$$

From the linear independence of $\{M_k\}$ (notice we can restrict (21) only to non-vanishing $M_k$), for any $k$ such that $M_k \neq 0$ holds

$$\sum_i \mu(j|i, k) = 1, \quad \forall j. \quad (23)$$

Also, by normalization,

$$\sum_j \mu(j|i, k) = 1, \quad \forall i, k. \quad (24)$$

Hence, for each such $k$ we see that the matrix $(\mu(j|i, k))_{ij} = \sum_{i,j} \mu(j|i, k)E_{ij}$ is doubly stochastic, and according to Birkhoff–von Neumann’s Theorem it has a decomposition into permutation matrices,

$$\sum_{i,j=1}^n \mu(j|i, k)E_{ij} = \sum_l r_l^{(k)} \Pi_l, \quad (25)$$

where $\mathbf{r}^{(k)} = (r_l^{(k)})_l \in \mathbb{S}_n$ is a probability vector, for any $k$ with $M_k \neq 0$.

Therefore, using (21) and (25) we obtain

$$\mathcal{A} = \sum_{ij} E_{ij} \otimes A_{ij} \quad (26a)$$

$$= \sum_{ij} E_{ij} \otimes \left( \sum_k \mu(j|i, k)M_k \right) \quad (26b)$$

$$= \sum_k \left( \sum_{ij} \mu(j|i, k)E_{ij} \right) \otimes M_k \quad (26c)$$

$$= \sum_k \left( \sum_l r_l^{(k)} \Pi_l \right) \otimes M_k \quad (26d)$$

$$= \sum_l \Pi_l \otimes \left( \sum_k r_l^{(k)} M_k \right). \quad (26e)$$

Thus, defining $Q_l := \sum_k r_l^{(k)} M_k$ we see that $Q_l \geq 0$ and $\sum_l Q_l = I$. Therefore, $\mathbf{Q} = (Q_l)$ is the coefficient-measurement that concludes the proof. \(\blacksquare\)
Notice that if the response functions $\mu(j|i,k)$ are deterministic, say, probability measures whose support contains a unique point, then the left-hand side of (25) is already a permutation matrix and the coefficient measurement $Q$ equals the mother measurement $M$. Also, the theorem can be enunciated according to the joint measurability of the column-measurements.

**Corollary 3.1:** The row-measurements of a DNT are jointly measurable if and only if its column-measurements are jointly measurable.

**Proof:** Assuming joint measurability of the rows of $\mathcal{A}$, according to Theorem 3.3 we have $\mathcal{A} = \sum_i \Pi_i \otimes Q_i$, for some POVM $Q$. Considering the transposition $\mathcal{A} = [A_{ij}] \mapsto [A_{ji}] =: A^T$, we see that $A^T = \sum_i \Pi_i^T \otimes Q_i$, and hence $A^T$ has jointly measurable row-measurements, which are the columns of $\mathcal{A}$. Analogously, one can prove the converse. ■

### 4. Incompatible DNTs

In spite of characterizing the set of DNTs that are decomposable into permutation tensors, Theorem 3.3 does not describe general DNTs. For that, we would need to show that all DNTs have jointly measurable row-measurements. Perhaps surprisingly, the next example shows that this is not the case, and not even the strong correlations between the rows of a DNT (given by the normalization of the columns) are sufficient to ensure joint measurability.

**Example 4.1:** For $d = 2$, consider the 3-outcome measurement $\mathcal{A} = (A_1, A_2, A_3)$ whose elements are vertices of an equilateral triangle in the Bloch sphere representation,

\[
A_1 = \frac{I + \sigma_x}{3}, \quad A_2 = \frac{I - (\sigma_x - \sqrt{3}\sigma_z)/2}{3}, \quad A_3 = \frac{I - (\sigma_x + \sqrt{3}\sigma_z)/2}{3},
\]

where $\sigma_x$ and $\sigma_z$ are Pauli matrices. Writing $A'_1 = \sigma_x/3$ and $A''_1 = I/3$, we have $A_1 = A'_1 + A''_1$ and the following DNT,

\[
\mathcal{A} = \begin{bmatrix} A'_1 + A''_1 & A_2 & A_3 \\ A_2 & A'_1 + A_3 & A''_1 \\ A_3 & A''_1 & A'_1 + A_2 \end{bmatrix},
\]

where each entry is positive semi-definite. However, one can check via semi-definite programming that these row-measurements are not jointly measurable [16], and use Theorem 3.3 to conclude that $\mathcal{A}$ is not a combination of permutation tensors.

Nevertheless, we can still write

\[
\mathcal{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes A'_1 + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \otimes A''_1 + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \otimes A_2
+ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \otimes A_3,
\]
which is a combination of permutation matrices where not all coefficient-operators are positive semi-definite, given that $A'_1 \not\succeq 0$.

Despite the fact that the rows of the DNT in the above example are not jointly measurable, they still can be reconstructed by applying a post-processing map to the tuple of coefficients $(A'_1, A'_2, A_2, A_3)$, as the proof of Theorem 3.3 shows. Since $A'_1 \not\succeq 0$, this tuple is not a mother measurement, but it plays the same role as one, and we call it a pseudomother measurement.

**Definition 4.1:** A pseudomother measurement for a set of measurements $\{A^{(1)}, \ldots, A^{(m)}\}$ is a tuple $\tilde{\mathbf{M}} = (\tilde{M}_1, \ldots, \tilde{M}_l)$ of not necessarily positive semi-definite operators, satisfying $\sum_k \tilde{M}_k = I$ and

$$A_j^{(i)} = \sum_k \mu(j|i, k)\tilde{M}_k, \quad \forall i, j. \tag{28}$$

Any set of measurements $\{B^{(1)}, \ldots, B^{(n)}\}$ admits a pseudomother; if we no longer impose positive semi-definitiveness, we can simply consider the products

$$\tilde{M}_{b_1 \ldots b_n} = B_{b_1}^{(1)} B_{b_2}^{(2)} \cdots B_{b_n}^{(n)} \tag{29}$$

and check that $\mu(j|i, b_1 \ldots b_n) = \delta_{b_i, j}$ is an appropriate post-processing map for it, since it satisfies

$$B_j^{(i)} = \sum_{b_1, \ldots, b_n} \tilde{M}_{b_1 \ldots b_n} \delta_{b_i, j}. \tag{30}$$

Notice that $\tilde{\mathbf{M}}$ is still normalized, and many other such pseudomothers can be constructed. For example, the order of the operators in (29) can be arbitrary, as long as it is the same for each element of the pseudomother.

Hence, by dropping positive semi-definiteness from the results in the last section, it is straightforward to obtain the following result.

**Theorem 4.1:** Let $A = [A_{ij}]$ be an $n \times n$ DNT. Then $A = \sum_i \Pi_i \otimes \tilde{Q}_i$, where the operators $\tilde{Q}_i$ are normalized, $\sum_i \tilde{Q}_i = I$, but not necessarily positive semi-definite.

This, however, does not provide a characterization of extremal DNTs. A central point in the proof of Theorem 3.3 is that extremal jointly measurable sets are related to extremal mother measurements. Relaxing positivity prevents such connection between general sets of measurements and pseudomothers, as the latter form an unbounded set which therefore admits no extremal points. Indeed, a perturbation for a pseudomother is a tuple of operators $\tilde{\mathbf{D}}$ such that $\tilde{\mathbf{M}} + \tilde{\mathbf{D}}$ is also a pseudomother, or equivalently that all marginals of $\tilde{\mathbf{D}}$ vanish. But in this case, $-\tilde{\mathbf{D}}$ is also a valid perturbation, and we find the decomposition $\tilde{\mathbf{M}} = [(\tilde{\mathbf{M}} + \tilde{\mathbf{D}}) + (\tilde{\mathbf{M}} - \tilde{\mathbf{D}})]/2$ witnessing its non-extremality on the set of pseudomothers. Hence, characterizing the extremal DNTs remains an open question.
5. DNTs arising from a joint measurability problem

Throughout this work, we presented our motivations to define DNTs and a generalization of Birkhoff–von Neumann’s theorem as purely mathematical, namely to further extend the clear parallel between probability vectors and POVMs. We now show that the set of DNTs can be reached also by a quantum-theoretical path, more specifically in terms of joint measurability.

The property of joint measurability is based on the existence of a mother measurement, but another natural question refers to the uniqueness of such object [18]. Restricting the post-processing map to be deterministic, we can display the mother measurement \( M = (M_{ab}) \) for a pair of POVMs \( A = (A_1, \ldots, A_n), B = (B_1, \ldots, B_n) \) as a table,

\[
\begin{array}{cccc}
M_{11} & \cdots & M_{1n} & A_1 \\
\vdots & \ddots & \vdots & \vdots \\
M_{n1} & \cdots & M_{nn} & A_n \\
B_1 & \cdots & B_n &
\end{array}
\]

(31)

emphasizing the marginalisations \( \sum_b M_{ab} = A_a \) and \( \sum_a M_{ab} = B_b \).

Once we decide to study the plurality of mother measurements for a given pair, it is reasonable to start with the simplest case. Taking \( A = B = (I/n, \ldots, I/n) \), we guarantee that the pair is trivially joint measurable, for being both copies of the same POVM. However, this trivial case allows to see that the general mother measurement for this pair, upon rescaling all the operators by a factor of \( n \) (the number of outcomes), yields a table

\[
\begin{array}{cccc}
nM_{11} & \cdots & nM_{1n} & I \\
\vdots & \ddots & \vdots & \vdots \\
nM_{n1} & \cdots & nM_{nn} & I \\
I & \cdots & I &
\end{array}
\]

(32)

which has exactly the DNT features, i.e. it can be provided with a tensor structure, and its rows and columns form POVMs.

Thus, we see that a generalization of BvN’s theorem emerges not only from an operator-version of the original result but also as the description of the set of mother measurements for perhaps the most trivial pair of POVMs that accepts multiple mothers.

6. Discussion

Following Birkhoff–von Neumann’s Theorem, we define doubly normalized tensors and investigate the extremality of these objects. Our first examples already reveal a complex scenario, as the connection between extremal DNTs and extremal rows/columns is intricate (see Table 1).

Similarly, the richer structure of non-negative operators yields a discrepancy between the set of doubly normalized tensors (that generalize doubly stochastic matrices) and decompositions into permutation tensors (that correspond to convex combinations of permutations). We showed that the latter can be perfectly described as the DNTs composed of jointly measurable POVMs, remarking joint measurability – a quantum-theoretical concept with a strongly operational motivation – as a relevant mathematical property.
by itself. Consequently, the extremal DNTs of this subset are identified by extremal coefficient-measurements.

On the other hand, extremal DNTs in general remain to be characterized. Theorem 4.1 states that DNTs can be written as affine combinations of permutation tensors. This approximates our characterization to another quantum-inspired generalization of BvN’s Theorem presented in Ref. [12], where it was proved that unital quantum channels are affine combinations of unitary channels. In a broader context, the need for quasi-POVMs in our description of incompatible DNTs dialogues with the need for quasi-probability representations in quantum theory [19, 20].

Nevertheless, the investigation that is closest in spirit to ours is Ref. [11]. Despite both works were developed in a completely independent manner, it features an analogue of a DNT extended to general algebras (called quantum magic square) and contains both the notion of decomposability in permutation tensors (called semiclassicity) and our Theorem 3.1. However, these authors are more interested in dilations of these objects and use a stronger notion of matrix convexity to characterize the set of decomposable DNTs as the matrix convex hull of commuting projective DNTs. This is remarkable, since joint measurability restricted to projective measurements reduces to commutation. Hence, the notion of joint measurability – which is central to us – is not directly considered in their work, but appears implicitly. It actually provides alternative proofs for some of their conclusions, as the existence of incompatible DNTs readily shows such matrix convex hull cannot expand the whole DNT set. Overall, both papers complement each other to form a Birkhoff–von Neumann bridge between measurement incompatibility and dilations of sets of quantum measurements. This seems an interesting line for future work.

Although the decomposition of doubly stochastic matrices involves in principle $n!$ terms, it was shown in Ref. [21] that only $(n - 1)^2 + 1$ terms are sufficient. Although many different such decompositions exist, the problem of computing the optimal one (with the minimal number of terms) is NP-complete [22]. Here, Theorem 3.3 establishes that extremal jointly measurable DNTs must correspond to extremal coefficient-measurements. Extremal measurements on dimension $d$ have at most $d^2$ non-null elements, thus by fixing the dimension of the underlying Hilbert space (which is independent of the size $n \times n$ of the tensor), we arrive at the uniform upper bound of $d^2$ for the number of terms in the decomposition of extremal jointly measurable DNTs.

Finally, we recall that doubly stochastic matrices go together with the concept of majorisation of real vectors and Hermitian matrices [23], a celebrated connection with important applications to quantum information theory [24]. It would be interesting to understand whether DNTs are associated to a similar notion of majorisation for measurements, together with its consequences for joint measurability.

**Note**

1. From this point on, we abuse the notation and write $\Pi_l$ both for the permutation map $\Pi_l : \{1, \ldots, n\} \to \{1, \ldots, n\}$ and for its matrix representation.

**Acknowledgments**

The authors are thankful to T. Perche and M. T. Quintino for fruitful discussions and suggestions. AB is partially supported by CNPq.
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

LG was supported by the São Paulo Research Foundation (FAPESP) [grant number 2016/01343-7], [grant number 2018/04208-9].

ORCID

Alexandre Baraviera http://orcid.org/0000-0003-0893-6746

References

[1] Birkhoff G. Tres observaciones sobre el algebra lineal. Univ Nac Tucumán Revista A. 1946;5:147–151.
[2] Cui LB, Li W, Ng MK. Birkhoff-von Neumann theorem for multistochastic tensors. SIAM J Matrix Anal Appl. 2014;35(3):956–973.
[3] Nielsen MA, Chuang IL. Quantum computation and quantum information: 10th anniversary edition. 10th ed. New York (NY): Cambridge University Press; 2011.
[4] Heinosaari T, Miyadera T, Ziman M. An invitation to quantum incompatibility. J Phys A: Math Theor. 2016;49(12):123001.
[5] Brunner N, Cavalcanti D, Pironio S, et al. Bell nonlocality. Rev Mod Phys. 2014;86:419–478.
[6] Busch P, Lahti P, Werner RF. Measurement uncertainty relations. J Math Phys. 2014;55(4):042111.
[7] Musto B, Vicary J. Quantum latin squares and unitary error bases. Quantum Inf Comput. 2016;14:1318–1332.
[8] Musto B, Vicary J. Orthogonality for quantum Latin isometry squares. 15th International Conference on Quantum Physics and Logic, QPL 2018.
[9] Musto B. Constructing mutually unbiased bases from quantum latin squares. Electron Proc Theor Comput Sci. 2017;236:108–126.
[10] Benoist T, Nechita I. On bipartite unitary matrices generating subalgebra-preserving quantum operations. Linear Algebra Appl. 2017;521:70–103.
[11] de las Cuevas G, Drescher T, Netzer T. Quantum magic squares: dilations and their limitations. J Math Phys. 2020;61:111704.
[12] Mendl CB, Wolf MM. Unital quantum channels – convex structure and revivals of Birkhoff’s theorem. Commun Math Phys. 2009;289(3):1057–1086.
[13] Masanes L. Extremal quantum correlations for N parties with two dichotomic observables per site. arxiv: quant-ph/0512100.
[14] D’Ariano GM, Presti PL, Perinotti P. Classical randomness in quantum measurements. J Phys A Math Gen. 2005;38:5979–5991.
[15] Buscemi F, Keyl M, D’Ariano GM, et al. Clean positive operator valued measures. J Math Phys. 2005;46(8):1089–7658.
[16] Wolf MM, Perez-Garcia D, Fernandez C. Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. Phys Rev Lett. 2009;103:230402.
[17] Ali ST, Carmeli C, Heinosaari T, et al. Commutative POVMs and fuzzy observables. Found Phys. 2009;39(6):593–612.
[18] Guerini L, Terra Cunha M. Uniqueness of the joint measurement and the structure of the set of compatible quantum measurements. J Math Phys. 2018;59(4):042106.
[19] Wigner E. On the quantum correction for thermodynamic equilibrium. Phys Rev. 1932;40:749.
[20] Lütkenhaus N, Barnett SM. Nonclassical effects in phase space. Phys Rev A. 1995;51:3340–3342.
[21] Marcus M, Ree R. Diagonals of doubly stochastic matrices. Q J Math. 1959;10(1):296–302.
[22] Dufossé F, Uçar B. Notes on Birkhoff-von Neumann decomposition of doubly stochastic matrices. Linear Algebra Appl. 2016;497:108–115.
[23] Watrous J. https://cs.uwaterloo.ca/watrous/lecturenotes/cs766.fall2011/13.pdf.
[24] Nielsen MA. Conditions for a class of entanglement transformations. Phys Rev Lett. 1999;83:436–439.