IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **46** (2013) 272002 (12pp) doi:10.1088/1751-8113/46/27/272002

FAST TRACK COMMUNICATION

Majorization entropic uncertainty relations*

Zbigniew Puchała[1,2], **Łukasz Rudnicki**[3,4] **and Karol Życzkowski**[2,3]

[1] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice, Poland
[2] Institute of Physics, Jagiellonian University, ul Reymonta 4, 30-059 Kraków, Poland
[3] Center for Theoretical Physics, Polish Academy of Sciences, al Lotników 32/46, 02-668 Warszawa, Poland
[4] Freiburg Insitute for Advanced Studies, Albert-Ludwigs University of Freiburg, Albertstr 19, D-79104 Freiburg, Germany

E-mail: rudnicki@cft.edu.pl

Received 16 May 2013, in final form 5 June 2013
Published 21 June 2013
Online at stacks.iop.org/JPhysA/46/272002

Abstract

Entropic uncertainty relations in a finite-dimensional Hilbert space are investigated. Making use of the majorization technique we derive explicit lower bounds for the sum of Rényi entropies describing probability distributions associated with a given pure state expanded in eigenbases of two observables. Obtained bounds are expressed in terms of the largest singular values of submatrices of the unitary rotation matrix. Numerical simulations show that for a generic unitary matrix of size $N = 5$, our bound is stronger than the well-known result of Maassen and Uffink (MU) with a probability larger than 98%. We also show that the bounds investigated are invariant under the dephasing and permutation operations. Finally, we derive a classical analogue of the MU uncertainty relation, which is formulated for stochastic transition matrices.

PACS number: 03.65.Aa

(Some figures may appear in colour only in the online journal)

1. Introduction

The uncertainty principle is often considered as a key feature of quantum theory, as it explicitly emphasizes the difference with respect to its classical counterpart. The original formulation given by Heisenberg in 1927 [1], which had been devoted to canonically conjugated variables, was further generalized by Robertson in 1929 for arbitrary two observables [2]. If both observables do not commute, it is impossible to specify their precise values simultaneously. In this set-up, uncertainties are characterized by the variances of both variables, and the relation provides a lower bound for the product of these quantities.

* Dedicated to Iwo Białynicki-Birula on the occasion of his 80th birthday.

Another method to describe the uncertainty is to use the continuous entropy of the probability distribution of the measurement outcomes. In 1975, Białynicki-Birula and Mycielski derived the entropic formulation of the uncertainty relation [3], in which the main role is played by the lower bound for the sum of two continuous Shannon entropies calculated for position and momentum probability distributions.

Entropic uncertainty relations, originally introduced for the infinite-dimensional Hilbert space, were later investigated in the case of a finite-dimensional quantum system. Consider a pure state $|\psi\rangle$ belonging to an N-dimensional Hilbert space $\mathcal{H}_N$ and a non-degenerate observable A, the eigenstates $|a_i\rangle$ of which determine an orthonormal basis in $\mathcal{H}_N$. The probability that this observable measured in $|\psi\rangle$ gives the ith outcome is $p_i = |\langle a_i|\psi\rangle|^2$. The non-negative numbers p_i sum up to unity, $\sum_{i=1}^{N} p_i = 1$, so that the properties of the discrete probability distribution $\{p_i\}$ can be described by the Shannon entropy $H(p) = -\sum_i p_i \ln p_i$.

Let $H(q)$ denote the Shannon entropy corresponding to the probability vector $q_j = |\langle b_j|\psi\rangle|^2$ associated with the second observable B. If both observables do not commute the sum of both entropies for any state, $|\psi\rangle$ is bounded from below, and the bound depends only on the unitary rotation matrix $U_{ij} = \langle a_i|b_j\rangle$. The first lower bound

$$H(p) + H(q) \geqslant -2 \ln \frac{1+c}{2} \equiv B_{\mathrm{D}}, \tag{1}$$

where $c = \max_{ij} |U_{ij}|$ was with the help of variational calculus derived by Deutsch in 1983 [4]. Maassen and Uffink (MU) obtained in 1988 a stronger result of the form [5]

$$H(p) + H(q) \geqslant -\ln c^2 \equiv B_{\mathrm{MU}}. \tag{2}$$

Note that for a Fourier matrix of size N, which describes the transition from position to momentum representation, one has $c = 1/\sqrt{N}$, so that $B_{\mathrm{MU}} = \ln N$.

The result of MU, while stronger than the Deutsch lower bound, is known not to be optimal in the general case. The optimal bound is known only for $N = 2$ [6, 7]; however, in higher dimensions the problem remains open (a step going beyond MU was recently performed in [8]). Let us mention that the entropic uncertainty relations were recently formulated in various set-ups [9–12]. To learn more about further developments in that area, the reader is referred to the reviews [13, 14].

In general, it is convenient [9] to work with the Rényi entropy

$$H_\alpha(x) = \frac{1}{1-\alpha} \ln \sum_{i=1}^{N} x_i^\alpha \tag{3}$$

which tends to the Shannon entropy for $\alpha \to 1$, is equal to the min-entropy $-\ln x_{\max}$ in the limit $\alpha \to \infty$ and is a non-increasing function of the parameter α [16]. One may then look for bounds for the sum of two Rényi entropies of order α. For instance, explicit bounds in the case of $\alpha = 1/2$ have been for $N = 2$ recently obtained by Rastegin [17].

The aim of this work is to derive a novel bound, which for a generic unitary U is with high probability stronger than (2). We shall establish lower bounds for the sum of two entropies of an arbitrary order $\alpha > 0$,

$$H_\alpha(p) + H_\alpha(q) \geqslant B_\alpha(U), \tag{4}$$

with $B_\alpha(U)$ depending in general on the whole matrix U.

To improve the approaches of Deutsch and MU (corresponding to the case $\alpha = 1$), we are going to characterize the unitary rotation matrix U by taking into account all its entries. Our approach is based on the concept of majorization.

Consider any two probability vectors x and y of sizes N and M, respectively. Associated vectors of size $\max\{N, M\}$, with coefficients ordered decreasingly and zeros on additional coordinates possibly added to the shorter vector, will be denoted as $\tilde{x}$ and $\tilde{y}$. The vector x is said to be *majorized* by y, written $x \prec y$, if $\tilde{x}$, $\tilde{y}$ satisfy inequalities for all partial sums [18]

$$\sum_{i=1}^{m} \tilde{x}_i \leqslant \sum_{i=1}^{m} \tilde{y}_i, \tag{5}$$

where m runs from 1 to $\max\{N, M\}$. Note that for $m = \max\{N, M\}$, the inequality is trivially saturated as both vectors sum up to 1.

The Rényi entropy is a Schur-concave function for any parameter $\alpha \geqslant 0$, which implies that if $x \prec y$, then $H_\alpha(x) \geqslant H_\alpha(y)$. In general, when a given function F is Schur-concave and two probability vectors satisfy the majorization relation, $x \prec y$, one obtains the inequality $F(y) \leqslant F(x)$.

This paper is organized as follows. Our main result—the explicit uncertainty relations for the sum of Rényi entropies—is derived in section 2. In section 3, we show that the MU bound and the bounds derived in this work are invariant with respect to permutation and dephasing operations. The bounds for some exemplary families of unitary matrices of size $N = 2, 3, 4, 5$ are discussed in section 4. In this section, we also use random unitary matrices to compare the precision of various new and previous bounds. Finally, in section 5, we present a classical analogue of the MU relation for an arbitrary stochastic transition matrix.

2. Main result

In the entropic uncertainty relation (4), we bound the sum of entropies of two probability vectors p and q. However, this sum can be rewritten as the single entropy of the product vector:

$$\begin{aligned} H_\alpha(p) + H_\alpha(q) &= \frac{1}{1-\alpha}\left(\ln \sum_i p_i^\alpha + \ln \sum_j q_j^\alpha\right) \\ &= \frac{1}{1-\alpha} \ln \sum_{ij} (p_i q_j)^\alpha = H_\alpha(r), \end{aligned} \tag{6}$$

where $r = p \otimes q$ is the tensor product of the classical probability vectors.

Assume that p and q are given by the fixed unitary matrix $U \in \mathcal{U}(N)$ and some vector $|\psi\rangle$ as $p_i = |\langle i|\psi\rangle|^2$ and $q_j = |\langle j|U|\psi\rangle|^2$, where the vectors $|i\rangle$ for $i = 1, \ldots, N$ form the orthonormal basis. It was shown by Deutsch [4] that

$$\max_{|\psi\rangle, i, j} p_i q_j = \max_{|\psi\rangle, i, j} \left(\frac{p_i + q_j}{2}\right)^2 = \left(\frac{1+c}{2}\right)^2 \equiv R_1, \tag{7}$$

where as before $c = \max_{ij} |U_{ij}|$. The above result immediately implies that

$$r \prec (R_1, 1 - R_1). \tag{8}$$

Since the Rényi entropies are Schur-concave, we arrive at the first, simple bound

$$H_\alpha(p) + H_\alpha(q) = H_\alpha(r) \geqslant \frac{1}{1-\alpha} \ln\left[R_1^\alpha + (1 - R_1)^\alpha\right]. \tag{9}$$

For any rectangular matrix X, one defines its spectral norm, equal to its largest singular value,

$$\|X\| = \sigma_{\max}(X). \tag{10}$$

By definition, singular values of X are equal to square roots of the eigenvalues of the positive matrix $XX^\dagger$.

$$\overset{(a)}{\begin{bmatrix} 0.19 & 0.50 & \boxed{-0.64} & \boxed{0.55} \\ -0.62 & 0.54 & -0.21 & -0.52 \\ 0.52 & -0.21 & -0.54 & -0.62 \\ -0.55 & \mathbf{-0.64} & -0.50 & 0.19 \end{bmatrix}} = \overset{(b)}{\begin{bmatrix} 0.19 & 0.50 & -0.64 & 0.55 \\ \mathbf{-0.62} & \mathbf{0.54} & -0.21 & -0.52 \\ \mathbf{0.52} & \mathbf{-0.21} & -0.54 & -0.62 \\ \boxed{-0.55} & \boxed{-0.64} & \boxed{-0.50} & 0.19 \end{bmatrix}}$$

Figure 1. Orthogonal matrix U of size 4 with truncated entries. To obtain s_1, we find its entry with the largest modulus, denoted in red boldface in panel (*a*). To obtain s_2, we find vertical and horizontal 2-vectors of the largest norm, marked by the blue ellipses and green boxes, respectively. Calculation of s_3, shown in panel (*b*) requires the comparison of norms of horizontal (green boxes) and vertical (blue ellipses) 3-vectors of U and 2×2 submatrix (bold entries in red) with the largest norm.

Let $\mathcal{A}(m, n)$ denote the set of all $m \times n$ submatrices of $U \in \mathcal{U}(N)$, i.e. the truncations obtained from U by removing arbitrary $N - m$ rows and arbitrary $N - n$ columns. Let $A_{m,n}$ denote the *maximal submatrix*, i.e. the element of $\mathcal{A}(m, n)$ with the largest spectral norm. We shall introduce a set of N coefficients

$$s_k := \max\{||A_{1,k}||,\ ||A_{2,k-1}||, \ldots,\ ||A_{k,1}||\} \tag{11}$$

where the maximum is taken over all submatrices with the same semiperimeter, $m+n = k+1$. By construction, we have $c = s_1 \leqslant s_2 \leqslant \ldots \leqslant s_N = 1$, so that s_1 is equal to the modulus of the largest element of U. Furthermore, s_2 is equal to the maximum of the Euclidean norm of any two-component part of any column or any row of U,

$$s_2 = \max\left\{\max_{i,j_1,j_2}\sqrt{|U_{ij_1}|^2 + |U_{ij_2}|^2}, \max_{i_1,i_2,j}\sqrt{|U_{i_1j}|^2 + |U_{i_2j}|^2}\right\}, \tag{12}$$

so it depends only on the moduli of the matrix entries. In the case of s_3, one needs to find the maximum among Euclidean norms of any 3×1 and 1×3 vectors and spectral norms of any 2×2 submatrix of U belonging to the set $\mathcal{A}(2, 2)$. In the latter case, not only the moduli but also the phases of entries of U become important. In figure 1, we present an exemplary calculation performed for a generic orthogonal matrix of size 4 in which all numbers are truncated up to two decimal digits.

In general, one also has a simple bound

$$s_k \leqslant ||A_{k,k}||. \tag{13}$$

Note that s_N is equal to unity as it is not smaller than the length of any column (or row) of U and is not larger than the spectral norm of the unitary matrix, $\|U\| = 1$.

In the next step, we define

$$R_k = \left(\frac{1+s_k}{2}\right)^2, \tag{14}$$

so that $\left(\frac{1+c}{2}\right)^2 = R_1 \leqslant R_2 \leqslant \ldots \leqslant R_N = 1$. Let us recall that R_1 has been introduced in equation (7). The above notation allows us to formulate key results of this paper.

Theorem 1. *For unitary matrix of size N and any normalized vector* $|\psi\rangle$*, we have*

$$p \otimes q \prec Q, \tag{15}$$

where

$$Q = (R_1, R_2 - R_1, R_3 - R_2, \ldots, R_N - R_{N-1}). \tag{16}$$

Note that from the above theorem we directly obtain the following corollary.

Corollary 1. *For* $Q^{(k)}$ *defined as*

$$Q^{(k)} = (R_1, R_2 - R_1, R_3 - R_2, \ldots, 1 - R_k), \tag{17}$$

we have

$$p \otimes q \prec Q = Q^{(N-1)} \prec Q^{(N-2)} \prec \ldots \prec Q^{(1)}. \tag{18}$$

Corollary 2. *For any unitary matrix U of size N, any normalized vector $|\psi\rangle \in \mathcal{H}_N$ and a Schur-concave function F, we have*

$$F(p \otimes q) \geqslant F(Q) = (Q^{(N-1)}) \\ \geqslant F(Q^{(N-2)}) \geqslant \ldots \geqslant F(Q^{(1)}). \tag{19}$$

Corollary 3. *For a unitary matrix U of size N, any normalized vector $|\psi\rangle \in \mathcal{H}_N$ and every $\alpha \geqslant 0$, we have*

$$H_\alpha(p) + H_\alpha(q) \geqslant H_\alpha(Q), \tag{20}$$

what can be extended to

$$H_\alpha(p) + H_\alpha(q) \geqslant B_\alpha^{N-1} \geqslant B_\alpha^{N-2} \geqslant \cdots \geqslant B_\alpha^1, \tag{21}$$

with $B_\alpha^i = H_\alpha\left(Q^{(i)}\right)$.

Proof of theorem 1. To prove the majorization relation (15), we consider sums of elements of the vector $p \otimes q$, i.e. $\Xi_k = p_{i_1}q_{j_1} + \cdots + p_{i_k}q_{j_k}$ for some indices $i_1, \ldots, i_k$ and $j_1, \ldots, j_k$, such that $(i_l, j_l) \neq (i_{l'}, j_{l'})$ for $l \neq l'$. Assume that the above sum consists of m different elements of the vector q. If we replace them by the m greatest elements of q, i.e. $\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_m$, preserving the order we do not decrease the sum, i.e.

$$\Xi_k \leqslant \tilde{q}_1(p_{i_1^1} + \cdots + p_{i_{k_1}^1}) + \cdots + \tilde{q}_m(p_{i_1^m} + \cdots + p_{i_{k_m}^m}), \tag{22}$$

where $k_1 + k_2 + \cdots + k_m = k$. In each parenthesis above, we shall next replace components of p by the components of the ordered vector $\tilde{p}$

$$\Xi_k \leqslant \tilde{q}_1(\tilde{p}_1 + \cdots + \tilde{p}_{k_1}) + \cdots + \tilde{q}_m(\tilde{p}_1 + \cdots + \tilde{p}_{k_m}). \tag{23}$$

For all values of the index i, we have $k_i \leqslant k - m + 1$ which provides the final estimate

$$\Xi_k \leqslant (\tilde{p}_1 + \cdots + \tilde{p}_{k-m+1})(\tilde{q}_1 + \cdots + \tilde{q}_m). \tag{24}$$

The above reasoning gives us the inequality

$$p_{i_1}q_{j_1} + \cdots + p_{i_k}q_{j_k} \leqslant \max_{1 \leqslant m \leqslant k} \left(\sum_{l=1}^{k-m+1} \tilde{p}_l\right)\left(\sum_{l=1}^{m} \tilde{q}_l\right). \tag{25}$$

Using the fact that the arithmetic mean is not smaller than the geometric mean, we obtain

$$\left(\sum_{l=1}^{k-m+1} \tilde{p}_l\right)\left(\sum_{l=1}^{m} \tilde{q}_l\right) \leqslant \frac{1}{4}\left(\sum_{l=1}^{k-m+1} \tilde{p}_l + \sum_{l=1}^{m} \tilde{q}_l\right)^2. \tag{26}$$

Now we can apply lemma 1 proven in the appendix to bound the inner sums by

$$\sum_{l=1}^{k-m+1} \tilde{p}_l + \sum_{l=1}^{m} \tilde{q}_l \leqslant 1 + \max_{A \in \mathcal{A}_{k-m+1,m}} \sigma_{\max}(A), \tag{27}$$

where for maximum ranges over all sumbatrices of size $(k - m + 1) \times m$. Together with maximization over m, we have

$$\max_{1 \leqslant m \leqslant k} \left(\sum_{l}^{k-m+1} \tilde{p}_l + \sum_{l}^{m} \tilde{q}_l\right) \leqslant 1 + s_k. \tag{28}$$

Thus, we finally obtain the following estimate:

$$p_{i_1}q_{j_1} + \cdots + p_{i_k}q_{j_k} \leqslant R_k, \tag{29}$$

which gives us the desired majorization relation,

$$p \otimes q \prec (R_1, R_2 - R_1, R_3 - R_2, \ldots, R_N - R_{N-1}). \tag{30}$$

In the maximizing case, the inequality (26) can be saturated. This follows from remark 2 in the appendix, as an optimal choice of the vector $|\psi\rangle$ implies that both terms on the right-hand side of (26) are equal, so the geometric and arithmetic means coincide. □

3. Equivalent unitary matrices and entropic uncertainty

We say that two unitary matrices U and V are equivalent $U \sim V$ if there exist permutation matrices P_1, P_2 and diagonal unitary matrices D_1, D_2, such that

$$V = P_1 D_1 U D_2 P_2. \tag{31}$$

It is easy to realize that any unitary matrix is equivalent to another unitary matrix with real elements in the first row and the first column which is often called dephased [15]. Another simple fact is that any 2×2 unitary matrix is equivalent to a real rotation matrix

$$\mathcal{U}(2) \ni U \sim O(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2). \tag{32}$$

The probability distribution induced by a normalized vector $|\psi\rangle$ is invariant with respect to diagonal unitary operations. The permutation matrix changes only the order of coordinates; thus, the above equivalence does not affect the left-hand side of relation (4). This observation suggests that one can restrict his/her attention to functions $B(U)$ which are invariant with respect to the equivalence relation. Note that the functions $B_D(U)$ and $B_{\mathrm{MU}}(U)$, as well as the bounds (21), are invariant with respect to the relation introduced. Therefore, to analyze the entropic uncertainty relations for unitary matrices of order N, one can investigate the $N^2 - 2N - 1$-dimensional set of the dephased matrices.

4. Low-dimensional examples

To demonstrate the new uncertainty relation proven above in action, we consider first the case $N = 2$. As stated in the previous section, it is enough to consider a one-parameter family of rotation matrices, since any 2×2 unitary matrix is similar to a rotation matrix $O(\theta)$ given in (32). In figure 2, we present the bound (20) for different values of the parameter α.

In the case of $N = 3, 4, 5$, we consider a one-parameter family of unitary matrices, given by P^β, where $\beta \in [0, a]$ and P is a circular shift permutation, which in the case of $N = 3$ reads

$$P_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{33}$$

In figures 3, 5 and 6, we have shown the comparison of the bounds (21) with the Deutsch bound (1) and the MU bound (2) for the family of unitary matrices which interpolate between identity and the N-point permutation matrix P_N.

In the case $N = 3$, we analyze a two-dimensional cross-section of the Birkhoff polytope of bistochastic matrices, and select $\mathcal{B}(a, b) = aP_3 + bP_3^2 + (1-a-b)\mathbb{I}$, for $0 \leqslant a, b \leqslant 1, a+b \leqslant 1$. Out of this equilateral triangle of bistochastic matrices of order $N = 3$ formed by the convex hull of permutation matrices P_3, P_3^2 and $P_3^3 = \mathbb{I}$, only a proper subset corresponds to unistochastic matrices, such that there exists a unitary U and $\mathcal{B}_{ij} = |U_{ij}|^2$. For unitary matrices associated with this subset, forming an interior of the 3-hypocycloid [19], we checked whether

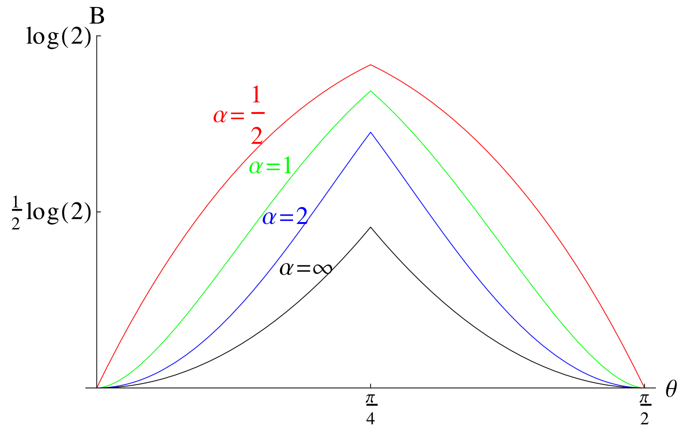

Figure 2. Bound (20) for a rotation matrix of size $N = 2$, obtained for different Rényi parameters $\alpha = 1/2, 1, 2, \infty$ as a function of parameter θ.

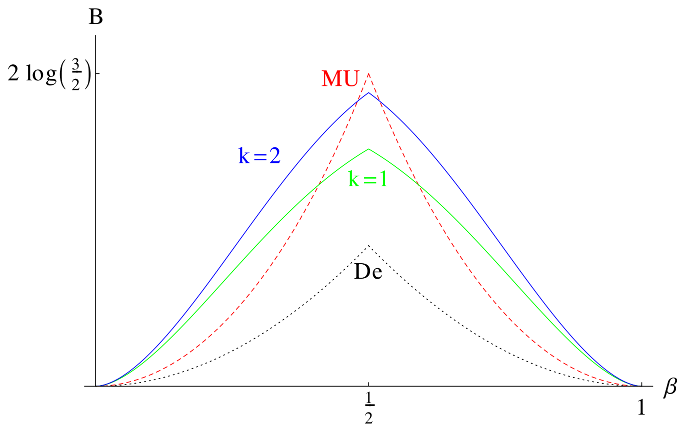

Figure 3. Comparison of bounds (21) for the family of matrices P_3^β of order $N = 3$, where P_3 is a circular shift permutation. The black dotted line represents the Deutsch bound B_D (1), the red dashed line represents the MU bound B_{MU} (2) and the solid lines represent the bounds B_1^k (21) for $k = 1, 2$.

the MU bound B_{MU} is larger than the bound B_1^2. Such a set, denoted in light color in figure 4, contains the center of the figure—the flat bistochastic matrix, $\mathcal{B}_{ij} = 1/3$ associated with the Fourier matrix F_3 for which the MU bound is sharp.

In order to compare the precision of the bound (20) for the standard case of the Shannon entropy $\alpha = 1$, we computed it for random unitary matrices distributed with the Haar measure on $\mathcal{U}(N)$. Probability that for a given unitary matrix U our bound is better than the MU bound, increases with N and reads, $\mathbb{P} = 0.814$ for $N = 2$ and $\mathbb{P} = 0.971, 0.972, 0.984, 0.991$, for $N = 3, 4, 5, 6$, respectively. These numbers are obtained numerically by averaging over samples of 10^7 random unitary matrices.

5. Classical analogues of MU bounds

In the classical case, we discuss an N-point probability vector P and its image with respect to a stochastic matrix, $P' = TP$. Stochasticity conditions, $T_{ij} \geqslant 0$ and $\sum_i T_{ij} = 1$, ensure that P'

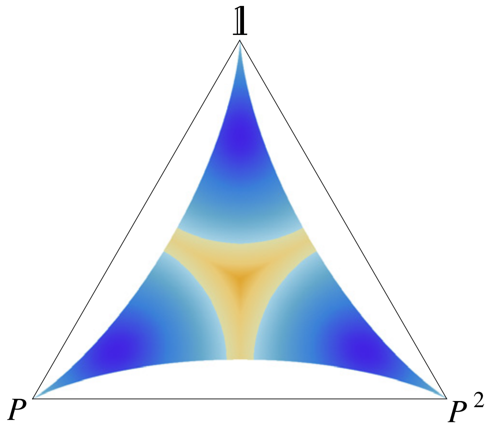

Figure 4. The difference between the MU bound B_{MU} (2) and bound B_1^2 (20) for unitary matrices U of size $N = 3$ corresponding to the cross-section of the set of bistochastic matrices. In the blue (dark) region, the MU bound B_{MU} is lower than the bound B_1^2, while the opposite is true in the yellow (light) region.

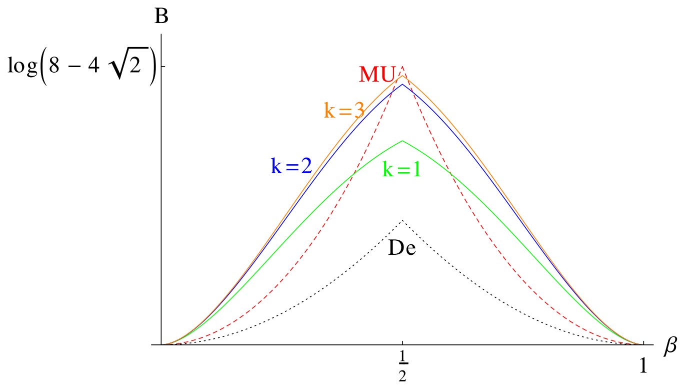

Figure 5. Comparison of bounds B_1^k for a family of matrices P_4^β, where $P \in \mathcal{U}(4)$ is a circular shift permutation. The black dotted line represents the Deutsch bound B_{D} (1), the red dashed line represents the MU bound B_{MU} (2) and the solid lines represent bounds B_1^k (21) for $k = 1, 2, 3$ (3).

is also a normalized probability vector. For a stochastic matrix T and a probability vector P, Słomczyński established [20] the following inequality for the Shannon entropy H:

$$H^{(p)}(T) \leqslant H(TP) \leqslant H^{(p)}(T) + H(P). \tag{34}$$

Here $H^{(p)}(T)$ denotes a statistical mixture of columns of the matrix T with weights p_i, i.e. $H^{(p)}(T) = \sum_i p_i H(\vec{t}_i)$. Using inequality (34), we obtain

$$H(TP) \geqslant H^{(p)}(T) = \sum_i p_i H(\vec{t}_i) \geqslant \min_i H(\vec{t}_i), \tag{35}$$

which gives us

$$H(P) + H(P') \geqslant H(P') \geqslant \min_i H(\vec{t}_i). \tag{36}$$

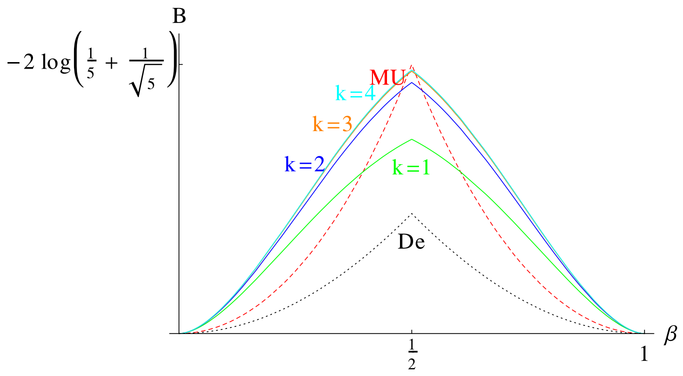

Figure 6. As in figure 5 , comparison of bounds B_1^k for a family of matrices P_5^β, where $P \in \mathcal{U}(5)$ is a circular shift permutation. The black dotted line represents the Deutsch bound B_D (1), the red dashed line represents the MU bound B_MU (2) and the solid lines represent the bounds B_1^k (21) for $k = 1, 2, 3, 4$.

We can continue the estimate and write

$$H(P) + H(P') \geqslant H(P') \geqslant \min_i H_\infty(\vec{t}_i) = \min_i(-\log(\max_j T_{ji})) = -\log\kappa, \tag{37}$$

where $\kappa = \max_{ij} T_{ji}$.

In this way, we obtain an analogue of the MU uncertainty relation for the classical maps represented by stochastic matrices. The sum of the Shannon entropy of any vector P and the entropy of its image $P' = TP$ is bounded from below by the logarithm of the inverse of the largest element of the transformation matrix.

6. Concluding remarks

The problem of establishing optimal entropic uncertainty relations for any two observables, the eigenbases of which are related by a unitary rotation matrix U of size N, remains open for $N \geqslant 3$.

In a recent work of Grudka *et al* [21], the authors analyzed column (or row) v of U, for which the entropy of the probability vector is the largest; $v_i = |U_{ij}|^2$ and $v'_i = |U_{ji}|^2$, where $j = 1, \ldots, N$ and the maximum is taken over i. Observe that in this notation, the MU bound (2) reads $H_1(p) + H_1(q) \geqslant \max[H_\beta(v), H_\beta(v')]$ with $\beta = \infty$, so decreasing the Rényi parameter β would make the bound stronger. Unfortunately, numerical simulations show that an appealing conjecture that the sum of the Shannon entropies is larger than $H_2(v)$ occurs to be true for $N = 2$ and $N = 3$ only.

On the other hand, in this work, we produced a family of inequalities for the sum of the Rényi entropies H_α of an arbitrary order which typically is stronger than the bounds existing in the literature. For instance, in the standard case of $\alpha = 1$, corresponding to the Shannon entropy, our result (25) applied to a random unitary matrix of size $N = 3$ gives a bound stronger than the MU result (2) for a vast majority, 97%, of cases.

It is worth emphasizing that the majorization techniques applied here enable one to obtain explicit bounds for any Schur-concave functions of the probability vector. The explicit formulae for the components of the majorizing vector Q derived in this work are expressed in terms of the spectral norms of the maximal submatrices of the unitary matrix U analyzed. As this norm is equal to the largest singular value of the submatrix [22], our bounds are directly computable. These bounds are shown to be invariant for any unitary matrices equivalent up to permutation and dephasing.

We shall mention that majorization techniques are used in the description of quantum entanglement [23–25]. For instance, the bipartite entanglement criteria by Gühne and Lewenstein [23] rely on the lower bound for the sum of two Rényi entropies. Our results can be immediately incorporated in that framework, providing sharpened entanglement criteria.

As a side remark, we presented a result analogous to the MU bound, but formulated for a classical map described by a stochastic transition matrix T. The sum of the Shannon entropies of an arbitrary initial probability distribution P and its image TP is bounded from below by minus logarithm of the largest entry of T.

Acknowledgments

It is a great pleasure to thank Iwo Białynicki-Birula for the numerous inspiring discussions on quantum theory we have had over the years. We are thankful to P Horodecki, M Horodecki and Ł Pankowski for stimulating discussions and useful correspondence and to V Gheorghiu for drawing our attention to a misprint in the first version of the paper. We also thank R Horodecki and W H Żurek for encouraging remarks. Financial support by NCN grants number DEC-2011/02/A/ST2/00305 (KŻ) and DEC-2012/04/S/ST6/00400 (ZP), and grant number IP2011 046871 of the Polish Ministry of Science and Higher Education (ŁR) are gratefully acknowledged.

Note added in proof. After this work was completed, we learned of the very recent results of Friedland, Gheorghiu and Gour [26]. These authors independently use majorization techniques to establish an entropic uncertainty relation analogous to (20), also valid for any Rényi entropies and arbitrary Schur-concave functions. These powerful bounds can be used to characterize generalized quantum measurements described by an arbitrary number of positive operator-valued measures.

Appendix. A useful lemma

We shall present the lemma which is the main ingredient of the proof of theorem 1.

Lemma 1. *Let* $|1\rangle, |2\rangle, \ldots, |m\rangle \in \mathcal{H}^N$ *and* $|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle \in \mathcal{H}^N$ *be two orthonormal sets of vectors; then*

$$\max_{|\psi\rangle\in\mathcal{H}^N}\left(\sum_{i=1}^{m}|\langle i|\psi\rangle|^2+\sum_{i=1}^{n}|\langle a_i|\psi\rangle|^2\right)=1+\sigma_1(A), \tag{A.1}$$

where $\sigma_1(A)$ *is the leading singular value of a rectangular matrix* $A=\{a_{ij}\}_{i=1,j=1}^{n,m}$ *for* $a_{ij}=\langle a_i|j\rangle$ *and the maximization is performed over normalized vectors* $\langle\psi|\psi\rangle=1$.

Remark 1. Note that in the case $m=1$, the matrix A has the form

$$A=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix}, \tag{A.2}$$

so its norm is equal to the length of the vector, $\sigma_1(A)=\sqrt{\sum_{i=1}^{n}|a_{i1}|^2}$.

Remark 2. One can construct a vector $|\psi_*\rangle$ which maximizes the left-hand side of (A.1). If we denote

$$\{|\xi_0\rangle, |\eta_0\rangle\} = \text{argmax}[\text{Re}\langle\xi|\eta\rangle : |\xi\rangle \in \text{lin}\{|1\rangle, \dots, |m\rangle\}, \\ |\eta\rangle \in \text{lin}\{|a_1\rangle, \dots, |a_n\rangle\}], \tag{A.3}$$

we shall take $|\psi_*\rangle$ as a vector proportional to the sum $|\xi_0\rangle + |\eta_0\rangle$. For this vector, one can show that

$$\sum_{i=1}^{m} |\langle i|\psi_*\rangle|^2 = \sum_{i=1}^{n} |\langle a_i|\psi_*\rangle|^2. \tag{A.4}$$

By $\text{lin}\{|1\rangle, \dots, |m\rangle\}$, we denote the linear space spanned by the unit vectors $|1\rangle, \dots, |m\rangle$.

Proof of lemma. We begin by rewriting the left-hand side of equation (A.1) in terms of matrix multiplication

$$\max_{|\psi\rangle} \left(\sum_{i=1}^{m} |\langle i|\psi\rangle|^2 + \sum_{i=1}^{n} |\langle a_i|\psi\rangle|^2 \right) = \max_{|\psi\rangle} \|C|\psi\rangle\|^2 = \sigma_1^2(C) = \lambda_1(CC^\dagger), \tag{A.5}$$

where the matrix C is defined as

$$C = \begin{pmatrix} \langle 1| \\ \langle 2| \\ \vdots \\ \langle m| \\ \langle a_1| \\ \langle a_2| \\ \vdots \\ \langle a_n| \end{pmatrix}. \tag{A.6}$$

It is easy to calculate that

$$CC^\dagger = \begin{pmatrix} \mathbb{I}_m & A^\dagger \\ A & \mathbb{I}_n \end{pmatrix}. \tag{A.7}$$

Next, we derive the formula for eigenvalues of matrix $CC^\dagger$:

$$\lambda_1(CC^\dagger) = 1 + \lambda_1 \begin{pmatrix} 0 & A^\dagger \\ A & 0 \end{pmatrix} = 1 + \sigma_1(A). \tag{A.8}$$

The last equality follows from Jordan's definition of singular values and can be found e.g. in [22]. □

References

[1] Heisenberg W 1927 Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik *Z. Phys.* **43** 172
[2] Robertson H P 1929 The uncertainty principle *Phys. Rev.* **34** 163
[3] Białynicki-Birula I and Mycielski J 1975 Uncertainty relations for information entropy in wave mechanics *Commun. Math. Phys.* **44** 129
[4] Deutsch D 1983 Uncertainty in quantum measurements *Phys. Rev. Lett.* **50** 631
[5] Maassen H and Uffink J B M 1988 Generalized entropic uncertainty relations *Phys. Rev. Lett.* **60** 1103
[6] Sanches-Ruiz J 1998 Optimal entropic uncertainty in two-dimensional Hilbert *Phys. Lett.* A **244** 189
[7] Ghirardi G, Marinatto L and Romano R 2003 An optimal entropic relation in a two dimensional Hilbert space *Phys. Lett.* A **317** 32

[8] de Vicente J I and Sanchez-Ruiz J 2008 Improved bounds on entropic uncertainty relations *Phys. Rev.* A **77** 042110
[9] Białynicki-Birula I 2006 Formulation of uncertainty relations in terms of the Rényi entropies *Phys. Rev.* A **74** 052101
[10] Tomamichel M and Renner R 2011 Uncertainty relation for smooth entropies *Phys. Rev. Lett.* **106** 110506
[11] Coles P J, Yu L, Gheorghiu V and Griffiths R B 2011 Information-theoretic treatment of tripartite systems and quantum channels *Phys. Rev.* A **83** 062338
[12] Coles P J, Colbeck R, Yu L and Zwolak M 2012 Uncertainty relations from simple entropic properties *Phys. Rev. Lett.* **108** 210405
[13] Wehner S and Winter A 2010 Entropic uncertainty relations—a survey *New J. Phys.* **12** 025009
[14] Białynicki-Birula I and Rudnicki Ł 2011 Entropic uncertainty relations in quantum physics *Statistical Complexity* ed K D Sen (Berlin: Springer) pp 1–34
[15] Tadej W and Życzkowski K 2006 A concise guide to complex Hadamard matrices *Open Syst. Inf. Dyn.* **13** 133–77
[16] Rényi A 1961 On measures of information and entropy *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, CA: University of California Press) pp 547–61
[17] Rastegin A E 2012 Notes on entropic uncertainty relations beyond the scope of Riesz's theorem *Int. J. Theor. Phys.* **51** 1300
[18] Marshall A W and Olkin O 1979 *Inequalities: Theory of Majorization and Its Applications* (New York: Academic)
[19] Dunkl C and Życzkowski K 2009 Volume of the set of unistochastic matrices of order 3 and the mean Jarlskog invariant *J. Math. Phys.* **50** 123521
[20] Słomczyński W 2002 Subadditivity of entropy for stochastic matrices *Open Syst. Inf. Dyn.* **9** 201
[21] Grudka A, Horodecki M, Horodecki P, Horodecki R, Kłobus W and Pankowski Ł 2012 Conjectured strong complementary correlations tradeoff arXiv:1210.8317
[22] Horn R A and Johnson C R 1991 *Topics in Matrix Analysis* (New York: Cambridge University Press)
[23] Gühne O and Lewenstein M 2004 Separability criteria from uncertainty relations *Phys. Rev.* A **70** 022316
[24] Partovi M H 2012 Entanglement detection using majorization uncertainty bounds *Phys. Rev.* A **86** 022309
[25] Huang Y 2010 Entanglement criteria via concave-function uncertainty relations *Phys. Rev.* A **82** 012335
[26] Friedland S, Gheorghiu V and Gour G 2013 Universal uncertainty relations arXiv:1304.6351