RÉNYI ENTROPY FOR SIGNED MEASURES WITH AN APPLICATION TO QUANTUM THEORY

ADAM BRANDENBURGER, PIERFRANCESCO LA MURA, STUART ZOBLE

JUNE 12, 2019

Abstract. We state axioms for Rényi entropy when signed measures are allowed and derive the family of entropy functionals that satisfies these axioms. We use this notion of entropy to provide a characterization of the simplest quantum system, namely, the qubit, in terms of an entropically-stated Uncertainty Principle and an Unbiasedness Principle.

1. Introduction

The maximum entropy method was introduced to physics as a way of deriving the Boltzmann distribution of statistical mechanics (Jaynes [1957]). This method has subsequently been widely used in information theory, statistics, and many applications besides physics. In this paper, we extend the definition of entropy to signed measures, which enables us to provide a maximum-entropy characterization of the simplest quantum system, namely, the qubit. Our approach is to work in quantum phase space (introduced by Wigner [1932]), which, as is well known, involves the use of negative (signed) probabilities. This is what necessitates our extension of entropy.

As for the notion of entropy which we employ, we start with Rényi entropy (Rényi [1961]), which includes Shannnon entropy (Shannon [1948]) as a special case and is used in communication theory, computer science, and quantum information, among other applications. Rényi entropy for ordinary (unsigned) measures was axiomatized in Rényi (1961) and Daróczy (1963). Here, we modify the Rényi axioms so that they retain their intent when signed measures are introduced, and we derive the family of entropy functionals that satisfies our axioms.

For our application to quantum systems of Rényi entropy with signed measures, we take two important features of quantum mechanics and turn them into physically justifiable axioms. The first axiom is an Uncertainty Principle, stated in terms of entropy. The second axiom is an Unbiasedness Principle, which requires that whenever there is complete certainty about the outcome of a measurement of the system in one of three mutually orthogonal directions, there must be maximal uncertainty about the outcomes in each of the two other directions. As we will see, the Unbiasedness Principle sets the value of the lower bound in the Uncertainty Principle. We show that, via Rényi entropy for signed measures, the quantum mechanics of a single qubit is fully characterized by these two axioms.

2. Axioms for Entropy

Rényi (1961) showed that his definition of entropy satisfied a list of axioms which he conjectured gave a characterization. Daróczy (1963) proved the conjecture. The approach followed by Rényi and Daróczy was first to axiomatize entropy for a larger class of measures (non-negative measures with total weight less than or equal to one) and then to specialize the construction to probabilities. We proceed in a similar manner by starting with a set of axioms which characterizes a notion of entropy for signed measures, and then specializing the construction to signed probabilities.

Given a finite set $X = \{x_1, \ldots, x_n\}$, a signed measure Q on X is defined by a tuple $Q = (q_1, \ldots, q_n)$ of real numbers. The quantity $w(Q) = |\Sigma_i q_i|$ will be called the weight of Q. We require $w(Q) \neq 0$ but we do not require $w(Q) = 1$ (except when Q is a signed probability measure).

Given two signed measures $P = (p_1, \ldots, p_m)$ and $Q = (q_1, \ldots, q_n)$, we denote by $P * Q$ the signed measure which is the product $(p_1q_1, \ldots, p_1q_n, \ldots, p_mq_1, \ldots, p_mq_n)$ whenever it is well-defined, i.e., whenever $\Sigma_{i,j} p_i q_j \neq 0$. Also, we denote by $P \cup Q$ the signed measure $(p_1, \ldots, p_m, q_1, \ldots, q_n)$ whenever it is well defined, i.e., whenever $\Sigma_i p_i + \Sigma_j q_j \neq 0$. We write (q) for the signed measure consisting of the scalar q. We impose the following axioms on entropy H:

Axiom 1. *(Real-Valuedness) $H(Q)$ is a non-constant real-valued function of Q.*

Axiom 2. *(Symmetry) $H(Q)$ is a symmetric function of the elements of Q.*

Axiom 3. *(Continuity) $H(Q)$ is a continuous function of each of the elements of Q.*

Axiom 4. *(Calibration) $H((\frac{1}{2})) = 1$.*

Axiom 5. *(Additivity) $H(P * Q) = H(P) + H(Q)$ whenever $H(P * Q)$ is well-defined.*

Axiom 6. *(Mean-Value Property) There is a strictly monotone and continuous function $g : \mathbb{R} \to \mathbb{R}$ such that for any P, Q, whenever $H(P \cup Q)$ is well-defined*

$$H(P \cup Q) = g^{-1}[\frac{w(P)g(H(P)) + w(Q)g(H(Q))}{w(P \cup Q)}].$$

Axiom 7. *(Smoothness) $H((q, 1-q))$ is smooth (C^∞) at $q = 0$.*

Some comments on the axioms. The forms of Axioms 2-6 are carried over without essential change from axioms for Rényi entropy with non-negative measures. (Notice that Axiom 2 is built into the set-up.) Axiom 1 ensures that entropy can be viewed as a measure of the amount or quantity of information in a system, and, to this end, states that entropy must be an ordinary (i.e., real) number. This axiom has bite when applied to signed vs. unsigned measures, because simply extending the domain of ordinary Rényi entropy to negative arguments may yield a complex-valued functional. (In particular, if α is an odd integer, then we may get the log of a negative number.) Concerning Axiom 7, Rényi entropy with non-negative measures is smooth in the interior of its domain. Axiom 7 imposes smoothness at $q = 0$, since this is no longer a boundary value of q.

Theorem 1. *Axioms 1-7 hold if and only if*

$$H(Q) := H_{2k}(Q) = -\frac{1}{2k-1}\log_2\left(\frac{\sum_i \mid q_i \mid^{2k}}{\mid \sum_i q_i \mid}\right), \tag{2.1}$$

where $k = 1, 2, \ldots$ is a free parameter.

Note that when Q is a signed probability measure, i.e., $\Sigma_i q_i = 1$, equation (2.1) reduces to

$$H_{2k}(Q) = -\frac{1}{2k-1}\log_2\left(\sum_i {q_i}^{2k}\right),$$

where we have also omitted the absolute value in the numerator, since $2k$ is an even integer.

The theorem follows from three lemmas.

Lemma 1. *Under Axioms 1, 3, 4, and 5, if $q \neq 0$, then $H((q)) = -\log_2 |q|$.*

Proof. Let $h(q) := H((q))$. Axioms 1 and 3 imply that h is real-valued and continuous. Axiom 5 implies that $h(pq) = h(p) + h(q)$ whenever $p, q \neq 0$. This is a version of Cauchy's logarithmic functional equation (Aczél and Dhombres [1989, Equation (7) and Theorem 3]) with general solution $h(q) = c\log_2 |q|$, where c is a real constant. Axiom 4 fixes $c = -1$. □

Lemma 2. *Under Lemma 1 and Axioms 5 and 6, we have $g(x) = -dx + e$ (linear) or $g(x) = d2^{(1-\alpha)}x + e$ (exponential), where $d \neq 0$, e, and $\alpha \neq 1$ are arbitrary constants.*

Proof. We extend the argument in Daróczy (1963) to signed measures. If Q is a signed measure, then from Lemma 1 and induction on Axiom 6 we obtain

(2.2)
$$H(Q) = H((q_1) \cup \cdots \cup (q_n)) = g^{-1}\left[\frac{\sum_j w((q_j))\, g(H((q_j)))}{w((q_1) \cup \cdots \cup (q_n))}\right] = g^{-1}\left[\frac{\sum_j |q_j|\, g(-\log_2 |q_j|)}{|\sum_j q_j|}\right].$$

From this and Axiom 5, we have for signed measures P and Q, provided $\Sigma_{i,j} p_i q_j \neq 0$

$$g^{-1}\left[\frac{\sum_{i,j} |p_i q_j|\, g(-\log_2 |p_i q_j|)}{|\sum_{i,j} p_i q_j|}\right] = g^{-1}\left[\frac{\sum_i |p_i|\, g(-\log_2 |p_i|)}{|\sum_i p_i|}\right] + g^{-1}\left[\frac{\sum_j |q_j|\, g(-\log_2 |q_j|)}{|\sum_j q_j|}\right].$$

Define $f : \mathbb{R}_{++} \to \mathbb{R}$ by $f(t) = g(-\log_2 t)$. Substituting, we get

$$f^{-1}[\frac{\sum_{i,j} |p_i q_j|\, f(|p_i q_j|)}{|\sum_{i,j} p_i q_j|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}] \times f^{-1}[\frac{\sum_j |q_j|\, f(|q_j|)}{|\sum_j q_j|}].$$

Setting $Q = (q)$ (where $q \neq 0$), this becomes

$$\frac{1}{|q|}\, f^{-1}[\frac{\sum_i |p_i|\, f(|p_i q|)}{|\sum_i p_i|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}].$$

Define $h_q : \mathbb{R}_{++} \to \mathbb{R}$ by $h_q(t) = f(|q|t)$. Then

$$h_q^{-1}[\frac{\sum_i |p_i|\, h_q(|p_i|)}{|\sum_i p_i|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}].$$

This shows that the maps h_q and f generate the same means when restricting the p_i to be non-negative. By a theorem on mean values (Hardy, Littlewood, and Pólya [1952, Theorem 83]), this implies that

$$h_q(t) = a(q)f(t) + b(q),$$

where $a(q)$ and $b(q)$ are independent of t, and $a(q) \neq 0$. Substituting, we get

$$f(|q|t) = a(q)f(t) + b(q).$$

This functional equation (restricting q to be non-negative) has the solution

$$f(t) = d \log_2 t + e,$$

or

$$f(t) = dt^{\alpha-1} + e,$$

where $d \neq 0$, e, and $\alpha \neq 1$ are arbitrary constants (Hardy, Littlewood, and Pólya [1952, Theorem 84]). Recalling the definition of f, we then find that either

$$g(x) = -dx + e, \tag{2.3}$$

or

$$g(x) = d2^{(1-\alpha)x} + e, \tag{2.4}$$

as required. □

Lemma 3. *Under Lemma 2 and Axioms 1 and 7, we have* $g(x) = d2^{(1-2k)x}$*, where* k *is a positive integer.*

Proof. If g is linear as in equation (2.3), then from equation (2.2) we get

$$-d \cdot H(Q) + e = d \cdot \frac{\sum_i |q_i| \log_2 |q_i|}{|\sum_i q_i|} + e \cdot \frac{\sum_i |q_i|}{|\sum_i q_i|}. \tag{2.5}$$

If g is exponential as in equation (2.4), then from equation (2.2) we get

$$d \cdot 2^{(1-\alpha)H(Q)} + e = d \cdot \frac{\sum_i |q_i|^\alpha}{|\sum_i q_i|} + e \cdot \frac{\sum_i |q_i|}{|\sum_i q_i|}. \tag{2.6}$$

Now use Axiom 7. Setting $Q = (q, 1-q)$ in equation (2.5) we find that $H((q, 1-q))$ is not C^1 at $q = 0$. Setting $Q = (q, 1-q)$ in equation (2.6) we find that $H((q, 1-q))$ is C^1 at $q = 0$ only if $e = 0$. If $\alpha < 0$, then $H((0, 1))$ is unbounded (negative), violating real-valuedness in Axiom 1. Thus $\alpha \geq 0$. If $\alpha = 0$, then $H(Q) = 1$ for all Q, violating non-constancy in Axiom 1. Next, suppose α is not an integer and let k be the least integer with $k > \alpha$. Then $\partial H((q, 1-q))/\partial q = \frac{\phi(q)}{\psi(q)}$ where $\phi(0) \neq 0$ and $\psi(0) = 0$. Thus α must be an integer. If α is an odd integer then $H((q, 1-q))$ is eventually not differentiable at 0. It follows that α is an even positive integer. $\square$

The sufficiency direction of Theorem 1 is finished by noting that equation (2.6) reduces to equation (2.1) when $e = 0$ and $\alpha = 2k$. The necessity of Axioms 1-7 is a straightforward calculation.

3. An Application to Quantum Theory

We next apply our Rényi entropy for signed measures to a question in quantum theory. There has been considerable recent interest in finding axioms for quantum theory based on information-theoretic principles. (Examples include using notions such as such as communication complexity (Van Dam [2005]), information causality (Pawlowski et al. [2009]), information capacity (Dakić and Brukner [2011]), and purification (Chiribella et al. [2011]). We provide an entropy-based characterization of the simplest quantum system, namely, a two-level system such as the spin of a particle.

We work in quantum phase space and associate phase-space representations to candidates for quantum states. These representations are, in general, signed probability measures. We use our Rényi entropy for signed measures to assign an entropy to each such representative. We then state our axioms on this structure.

To build phase space for a two-level quantum system, we start with the basis $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ for the space of 2×2 Hermitian matrices, where $\sigma_0 = \mathbf{I}$ is the 2×2 identity matrix and σ_1, σ_2, σ_3 are the Pauli matrices.

Lemma 4. *A 2×2 Hermitian matrix M satisfies $Tr(M) = 1$ if and only if*

$$M = \frac{1}{2}(\mathbf{I} + \sum_{k=1}^{3} r_k \sigma_k)$$

for some vector $r \in \mathbb{R}^3$.

Definition 1. A 2×2 Hermitian matrix M with $Tr(M) = 1$ is called a *potential quantum state*. If, in addition, M is positive semi-definite, then M is a 1-*qubit state*. We also refer to the corresponding r vectors as potential quantum states and 1-qubit states.

The *phase space* $\mathcal{P}$ for a two-level quantum system is the set of all maps which associate an eigenvalue, namely, $+1$ or -1, to each element of the basis excluding the identity matrix, that is

$$\mathcal{P} = \{f \mid f : \{\sigma_1, \sigma_2, \sigma_3\} \to \{+1, -1\}\}.$$

Note that $\mathcal{P}$ may also be viewed abstractly as a map on indices, without mention of Pauli matrices and eigenvalues. This makes our treatment fully axiomatic.

We next define a phase-space representation for each potential quantum state as a signed probability measure over $\mathcal{P}$. Let

$$Q(\mathcal{P}) = \{q \mid q : \mathcal{P} \to \mathbb{R} \text{ and } \sum_{f \in \mathcal{P}} q(f) = 1\}$$

denote the set of all signed probability measures on phase space.

Definition 2. A *phase-space state* is an element q of $Q(P)$.

For a given q and $1 \leq k \leq 3$, let r_k be the expected outcome under q, that is

$$r_k = \sum_{\{f \in \mathcal{P} \mid f(k)=+1\}} q(f) \times (+1) \quad + \sum_{\{f \in \mathcal{P} \mid f(k)=-1\}} q(f) \times (-1). \tag{3.1}$$

This defines a map ϕ from $Q(\mathcal{P})$ to the set of potential quantum states given by

$$\phi(q) = \frac{1}{2}(\mathbf{I} + \sum_{k=1}^{3} r_k \sigma_k).$$

Definition 3. We say q is a *phase-space representation* of $\phi(q)$.

This is a restriction of a linear map from $\mathbb{R}^8$ to $\mathbb{R}^3$ and it will be helpful to fix some notation surrounding a matrix representation.

	σ_1	σ_2	σ_3
f_1	$+1$	$+1$	$+1$
f_2	-1	$+1$	$+1$
f_3	$+1$	-1	$+1$
f_4	-1	-1	$+1$
f_5	$+1$	$+1$	-1
f_6	-1	$+1$	-1
f_7	$+1$	-1	-1
f_8	-1	-1	-1

Enumerate $\mathcal{P}$ as $\{f_j \mid j = 1, ..., 8\}$ as above. Also, let the matrix A be defined by

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

and, for $r \in \mathbb{R}^3$, define $\hat{r} \in \mathbb{R}^4$ by $\hat{r_i} = r_i$ for $i = 1, 2, 3$ and $\hat{r_4} = 1$.

Definition 4. For $q \in \mathbb{R}^8$ and $r \in \mathbb{R}^3$ we say q *represents* r if

$$Aq = \hat{r}.$$

Note we have folded the condition that q is a signed probability measure in as the last equation in the definition of representation. As an example, $q^T = (\frac{1}{4}, 0, \frac{1}{4}, 0, \frac{1}{4}, 0, \frac{1}{4}, 0)$ and $\tilde{q}^T = (\frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}, 0, 0, 0, \frac{1}{2})$ both represent $r = (1, 0, 0)$.

To move to our characterization of the qubit, we look at the largest collection of phase-space states with maximum entropies above a threshold. As for what definition of entropy to use, we refer back to Theorem 1 and use the family of entropy functionals there. That is, in the notation of this section where q is a signed probability measure on phase space, we choose

$$H_{2k}(q) = -\frac{1}{2k-1} \log_2(\sum_{j=1}^{8} (q(f_j))^{2k}).$$

We can now state our Uncertainty Principle, which takes the form of a lower bound on the maximum entropy of a phase-space representation of a potential quantum state.

Uncertainty Principle: A potential quantum state r is allowable if for all k,

$$\max_{\{q \in Q(\mathcal{P}) \mid Aq = \hat{r}\}} H_{2k}(q) \geq 2.$$

The reason for choosing the lower bound to be 2 comes from the unbiasedness property of quantum systems. A set of measurements on a system is called mutually unbiased if complete certainty of the measured value of the outcome of one of them implies maximal uncertainty about the outcomes of the others. As we did with the Uncertainty Principle, we now turn this feature of quantum systems into an axiom. Specifically, we assume that the three measurement directions in our two-level system form a mutually unbiased set, so that if one outcome is +1 (or -1) with probability 1, each of the other two outcomes is +1 with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. Using equation (3.1), we arrive at our second principle.

> Unbiasedness Principle: If a potential state r is allowable and some $r_i = 1$, then $r_j = 0$ for every $j \neq i$.

Lemma 5. *If any of the thresholds in the Uncertainty Principle are reduced, then the Unbiasedness Principle fails.*

Proof. Consider the phase-space probability measure $q = (\frac{1}{4}, +\frac{1}{4}\epsilon, \frac{1}{4}, -\frac{1}{4}\epsilon, \frac{1}{4}, +\frac{1}{4}\epsilon, \frac{1}{4}, -\frac{1}{4}\epsilon)$ where $\epsilon > 0$. Then q represents $r = (1, \epsilon, 0)$ which violates the Unbiasedness Principle, but

$$\begin{aligned} H_{2k}(q) &= -\frac{1}{2k-1}\log_2(4.(\frac{1}{4})^{2k} + 4.(\frac{1}{4}\epsilon)^{2k}) \\ &= -\frac{1}{2k-1}\log_2((\frac{1}{4})^{2k-1}(1+\epsilon^{2k})) \\ &= 2 - \frac{1}{2k-1}\log_2(1+\epsilon^{2k}) \end{aligned}$$

which tends to 2 from below as ϵ tends to 0. □

Lemma 6. *The Unbiasedness Principle holds.*

Proof. Note that

$$H_2(q) \geq 2 \Leftrightarrow \|q\|_2^2 \leq \frac{1}{4}.$$

For a general r, the representation q^* which maximizes 2-entropy is given by

$$q^* = A^T(AA^T)^{-1}\hat{r}.$$

Now use the fact that $AA^T = 8\mathbf{I}$ to write

$$\|q^*\|_2^2 = \hat{r}^T(AA^T)^{-1}\hat{r} = \frac{1}{8}r^Tr + \frac{1}{8} \leq \frac{1}{4} \Leftrightarrow \sum_{k=1}^{3} r_k^2 \leq 1,$$

and the result follows. □

Theorem 2. *The potential quantum states satisfying the Uncertainty Principle at $k = 1$ are precisely the states of the qubit.*

Proof. The matrix $\frac{1}{2}(\mathbf{I} + \sum_{k=1}^{3} r_k\sigma_k)$ is positive semi-definite if and only if $\sum_{k=1}^{3} r_k^2 \leq 1$. □

We conjecture a stronger result:

Conjecture 1. *The potential quantum states satisfying the Uncertainty Principle at all k are precisely the states of the qubit.*

To prove the conjecture, it suffices to show that if a potential state r satisfies the Uncertainty Principle at 2 then it satisfies the Uncertainty Principle at all $2k$.

Lemma 7. *Let $r \in \mathbb{R}^3$. Then*

$$\min\{\|q\|_{2k} \mid q \in \mathbb{R}^8 \text{ and } Aq = \hat{r}\} = \max\{\hat{r}^T x \mid x \in \mathbb{R}^4 \text{ and } \|A^T x\|_{\frac{2k}{2k-1}} \leq 1\},$$

and r satisfies the Uncertainty Principle at $2k$ if and only if

$$\max\{\hat{r}^T x \mid x \in \mathbb{R}^4 \text{ and } \|A^T\| \leq 1\} \leq (\frac{1}{2})^{\frac{2k-1}{k}}.$$

Proof. We can write the problem of maximizing entropy as the norm minimization problem

$$\min_{q \in \mathbb{R}^8} \|q\|_{2k}$$
$$\text{subject to } Aq = \hat{r}.$$

The dual problem is

$$\max_{x \in \mathbb{R}^4} \hat{r}^T x$$
$$\text{subject to} \|A^T x\|_{\frac{2k}{2k-1}} \leq 1.$$

(See Boyd and Vandenberghe, 2004, pp.221-222.) Note that $\|\cdot\|_{\frac{2k}{2k-1}}$ is the dual norm of $\|\cdot\|_{2k}$. Strong Duality holds so the values of the primal and dual problems are equal. For any q we have $H_{2k}(q) \geq 2$ if and only if $\|q\|_{2k} \leq (\frac{1}{2})^{\frac{2k-1}{k}}$ so the result follows. □

Conjecture 2. *If $r \in \mathbb{R}^3$ has a phase-space representation q with $H_{2k}(q) \geq 2$ then it has a phase-space representation q' with $H_{2k+2}(q') \geq 2$.*

To try to prove the conjecture, we assume r satisfies the Uncertainty Principle at $2k$. Let

$$C_k = \{x \in \mathbb{R}^4 \mid \|A^T x\|_{\frac{2k}{2k-1}} \leq 1\}.$$

Thus

$$C_{k+1} = \{x \in \mathbb{R}^4 \mid \|A^T x\|_{\frac{2k+2}{2k+1}} \leq 1\}$$

since $\|\cdot\|_{\frac{2k+2}{2k+1}}$ is the dual norm of $\|\cdot\|_{2k+2}$. Let y^k, y^{k+1} be the maximizers of the dual problems. Let z^k, z^{k+1} be the projections of y^k, y^{k+1} onto the vector $\hat{r}$.

Claim 1. If $\frac{\|z^{k+1}\|_2}{\|z^k\|_2} \leq (\frac{1}{2})^{\frac{1}{k(k+1)}}$ then r satisfies the Uncertainty Principle at $2k+2$.

Proof. We have $z^k = (\hat{r}^T y^k / \|\hat{r}\|)\hat{r}$ and $z^{k+1} = (\hat{r}^T y^{k+1} / \|\hat{r}\|)\hat{r}$. Since the values of the primal and dual problems are equal, this value is positive so $\frac{\|z^{k+1}\|_2}{\|z^k\|_2}$ is equal to the ratio of the value of the $k+1$ problem to the value of the k problem. By assumption

$$\hat{r}^T y^k \leq (\frac{1}{2})^{\frac{2k-1}{k}}$$

so

$$\|z^{k+1}\|_2 \leq \|z^k\|_2 (\frac{1}{2})^{\frac{1}{k(k+1)}} \leq (\frac{1}{2})^{\frac{2k-1}{k}} (\frac{1}{2})^{\frac{1}{k(k+1)}}$$

and thus

$$\|z^{k+1}\|_2 \leq (\frac{1}{2})^{\frac{2k+1}{k+1}}$$

as desired, since

$$\frac{2k-1}{k} + \frac{1}{k(k+1)} = \frac{2k+1}{k+1}.$$

□

Claim 2. Both C_k, C_{k+1} are convex and for every vector w there are unique $\lambda < \nu$ such that

$$\|A^T \nu w\|_{\frac{2k}{2k-1}} = 1$$

and

$$\|A^T \lambda w\|_{\frac{2k+2}{2k+1}} = 1.$$

Proof. Immediate from linearity and continuity. □

Now let

$$f(w) = \frac{\|A^T w\|_{\frac{2k}{2k-1}}}{\|A^T w\|_{\frac{2k+2}{2k+1}}}.$$

By homogeneity of norms, $f(\lambda w) = f(w)$ for any scalar λ. By the claim above, $f(w)$ is the ratio of the distance to the boundary of C_k along the ray through w to the distance to the boundary of C_{k+1}.

Claim 3. If $\max\{f(w) \mid w \in \mathbb{R}^4\} \leq (\frac{1}{2})^{\frac{1}{k(k+1)}}$ then r satisfies the Uncertainty Principle at $2k+2$.

Proof. We claim that $\frac{\|z^{k+1}\|_2}{\|z^k\|_2} \leq \frac{\|y^{k+1}\|_2}{\|w^k\|_2}$ where $w^k = \lambda y^{k+1}$ for the unique λ such that

$$\|A^T \lambda y^{k+1}\|_{\frac{2k}{2k-1}} = 1.$$

Thus $\frac{\|z^{k+1}\|_2}{\|z^k\|_2} \leq \frac{\|y^{k+1}\|_2}{\|w^k\|_2}$ is bounded by a value of f. Note that

$$\hat{r}^T w^k \leq \hat{r}^T y^k$$

so the length of the projection of w^k onto $\hat{r}$ (call this vector v) cannot exceed the length of the projection of y^k onto $\hat{r}$. By similar triangles then

$$\frac{\|y^{k+1}\|}{\|z^{k+1}\|} = \frac{\|w^k\|}{\|v\|} \geq \frac{\|w^k\|}{\|z^k\|}$$

so $\frac{\|z^{k+1}\|_2}{\|z^k\|_2} \leq \frac{\|y^{k+1}\|_2}{\|w^k\|_2} = f(w^k)$. □

To finish the proof, it is enough to show that the maximum value of the function f is less than

$$(\frac{1}{2})^{\frac{1}{k(k+1)}}.$$

Let $\alpha = \frac{2k}{2k-1}, \beta = \frac{2k+2}{2k+1}$, and let $g : \mathbb{R}^8 \to \mathbb{R}$ be given by

$$g(q) = \frac{\|q\|_\alpha}{\|q\|_\beta}.$$

The first-order conditions on g are

$$\frac{\|q\|_\beta}{\|q\|_\alpha^{\alpha-1}} \alpha q_i^{\alpha-1} - \frac{\|q\|_\alpha}{\|q\|_\beta^{\beta-1}} \beta q_i^{\beta-1} = 0$$

for each i. Thus the critical points are of the form $q_i = \pm 1$ for some nonempty set of indices i, and $q_i = 0$ otherwise. So the maximum value of g is 1 at the coordinate axes and the minimum value of g is $8^{\frac{1}{\alpha}-\frac{1}{\beta}}$ at the diagonals. Let q be a critical point with j indices which are ± 1. We have $g(q) = (1/j)^{\frac{1}{2k(k+1)}}$ and so

$$g(q) \leq (\frac{1}{2})^{\frac{1}{k(k+1)}} \Leftrightarrow j \geq 4.$$

Conjecture 3. *The maximum of g over the column space of $A^T$is at $j = 4$.*

Note that $j = 4$ (and the bound) is achieved in the column space of A^T, by settting $(w_1, w_2, w_3, w_4) = (0, 0, 1, 1)$. Computer simulations support this conjecture.

4. Negative Probabilities

Conventionally, negative probabilities are used in phase-space representations of 2-qubit or higher-dimensional systems, where negativity is seen as a witness of entanglement. (By Bell's Theorem (Bell [1964]), there are states of two-qubit systems that cannot be represented by non-negative probabilities on phase space.) We need negative probabilities even in a 1-qubit system because of our requirement that representations have a maximum 2-entropy of at least 2. To see this, consider $(r_1, r_2, r_3) = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. The (unique) maximum 2-entropy representation is

$$q = \frac{1}{8}(1+\sqrt{3}, 1+\frac{1}{\sqrt{3}}, 1+\frac{1}{\sqrt{3}}, 1-\frac{1}{\sqrt{3}}, 1+\frac{1}{\sqrt{3}}, 1-\frac{1}{\sqrt{3}}, 1-\frac{1}{\sqrt{3}}, 1-\sqrt{3}),$$

with negative final component. The 2-entropy of q is 2, so we cannot find another representation with all non-negative components with sufficiently high 2-entropy.

Acknowledgements

We are grateful to Matthew Deutsch for excellent research assistance, to Samson Abramsky, Časlav Brukner, Christopher Fuchs, Leslie Greengard, Michael J.W. Hall, Jerry Keisler, Elliot Lipnowski, David Pine, Rui Soares Barbosa, Yuri Tschinkel, and Noson Yanofsky for valuable comments. Financial support from the NYU Stern School of Business, J.P. Valles, NYU Abu Dhabi, NYU Shanghai, and the HHL - Leipzig Graduate School of Management is gratefully acknowledged.

Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu, http://www.adambrandenburger.com

HHL - Leipzig Graduate School of Management, 04109 Leipzig, Germany, plamura@hhl.de

Signal Fox, Princeton, NJ 08542, U.S.A., szoble@gmail.com, https://signalfox.com

References

Aczél, J., and J. Dhombres, *Functional Equations in Several Variables*, Cambridge University Press, 1989, 26-27.

Bell, J., "On the Einstein-Podolsky-Rosen Paradox," *Physics*, 1, 1964, 195-200.

Bosyk, G., M. Portesi, and A. Plastino, "Collision Entropy and Optimal Uncertainty," *Physical Review A*, 85, 2012, 012108.

Boyd, S., and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

Chiribella, G., G. D'Ariano, and P. Perinotti, "Informational Derivation of Quantum Theory," *Physical Review A*, 84, 2011, 012311.

Dakić, B., and Č. Brukner, "Quantum Theory and Beyond: Is Entanglement Special?" in Halvorson, H. (ed.), *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, Cambridge University Press, 2011, 365-392.

Daróczy, Z., "Über die gemeinsame Charakterisierung der zu den nicht vollständigen Verteilungen gehörigen Entropien von Shannon und von Rényi," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1, 1963, 381-388.

Hardy, G., J. Littlewood, and G. Pólya, *Inequalities*, 2nd edition, Cambridge University Press, 1952.

Pawlowski M., T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, "Information Causality as a Physical Principle," *Nature*, 461, 2009, 1101-1104.

Rényi, A., "On Measures of Information and Entropy," in Neymann, J., (ed.), *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1961, 547-561.

Rényi, A., "On the Foundations of Information Theory," *Review of the International Statistical Institute*, 33, 1, 1965, 1-14.

Shannon, C., "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 1948, 379-423 and 623-656.

Van Dam, W., "Implausible Consequences of Superstrong Nonlocality," 2005, at http://arxiv.org/abs/quant- ph/0501159.

Wigner, E., "On the Quantum Correction For Thermodynamic Equilibrium," *Physical Review*, 40, 1932, 749-759, at at https://doi:10.1103/PhysRev.40.749.