Deriving the Qubit from Entropy Principles: Supplementary Information

Adam Brandenburger[*] Pierfrancesco La Mura[†]

Version 03/14/15

In this supporting material, we provide proofs of the mathematical claims in the main paper. We also include an axiomatization of Rényi entropy[1] with signed probabilities and a discussion of the discrete Wigner function[2].

Appendix A Proof that $\beta_{2k} = 2$ for all k

Consider the phase-space probability distribution:

$$q_{000} = q_{001} = q_{010} = q_{011} = 1/4, \tag{A.1}$$

$$q_{100} = q_{101} = 1/4\epsilon, \tag{A.2}$$

$$q_{110} = q_{111} = -1/4\epsilon. \tag{A.3}$$

By equations (1)-(3) in the main paper, this distribution induces the empirical probabilities $f_x = 1$, $f_y = (1+\epsilon)/2$, $f_z = 1/2$. We also calculate the entropy of the phase-space distribution:

$$H_{2k}(q) = -\frac{1}{2k-1}\log_2[4 \times (\frac{1}{4})^{2k} + 4 \times (\frac{1}{4}\epsilon)^{2k}] \tag{A.4}$$

$$= -\frac{1}{2k-1}\log_2[4^{1-2k}(1+\epsilon^{2k})] \tag{A.5}$$

$$= 2 - \frac{1}{2k-1}\log_2(1+\epsilon^{2k}). \tag{A.6}$$

Suppose $\beta_{2k} = 2 - \delta$ for some $\delta > 0$. Then choose $\epsilon > 0$ sufficiently small that:

$$\frac{1}{2k-1}\log_2(1+\epsilon^{2k}) < \delta. \tag{A.7}$$

Thus $H_{2k}(q) > 2 - \delta$ and the Uncertainty Principle $H_{2k}(q) \geq \beta_{2k}$ is satisfied. But, for any $\epsilon > 0$, the empirical model with $f_x = 1$, $f_y = (1+\epsilon)/2$, $f_z = 1/2$ violates the Unbiasedness Principle.

This establishes that $\beta_{2k} \geq 2$. It is immediate that setting $\epsilon = 0$ in (A.1)-(A.3) gives $H_{2k}(q) = 2$ and that the induced empirical model satisfies the Unbiasedness Principle. This says that $\beta_{2k} \leq 2$, so we conclude that $\beta_{2k} = 2$, as required.

[*] Stern School of Business, Polytechnic School of Engineering, Institute for the Interdisciplinary Study of Decision Making, Center for Data Science, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu, adambrandenburger.com

[†] HHL – Leipzig Graduate School of Management, 04109 Leipzig, Germany, plamura@hhl.de

Appendix B Proof that the Sets $\mathcal{R}_{2k}$ are Increasing

By equation (5) in the main paper, we need to show that if for a point (r_x, r_y, r_z) in the cube $[-1,1]^3$, there is a phase-space probability distribution q that satisfies $H_{2k}(q) \geq 2$ and represents (r_x, r_y, r_z), then there is a phase-space probability distribution q' that satisfies $H_{2k+2}(q') \geq 2$ and represents (r_x, r_y, r_z). We begin by picking q to be the phase-space distribution that maximizes $H_{2k}(q)$ and represents (r_x, r_y, r_z), and then use duality.

We first rewrite Rényi entropy $H_{2k}(q)$ in terms of the $2k$-norm, for $k = 1, 2, 3, \ldots$:

$$H_{2k}(q) = -\frac{2k}{2k-1} \log_2(||q||_{2k}), \tag{B.1}$$

We can then write the problem of maximizing entropy as the norm minimization problem:

$$\min_{q \in \mathbb{R}^8} ||q||_{2k} \tag{B.2}$$

$$\text{subject to} \quad Aq = b, \tag{B.3}$$

where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} (1+r_x)/2 \\ (1+r_y)/2 \\ (1+r_z)/2 \\ 1 \end{pmatrix}. \tag{B.4}$$

The dual problem[3] is:

$$\max_{x \in \mathbb{R}^4} b^{\mathrm{T}} x \tag{B.5}$$

$$\text{subject to} \quad ||A^{\mathrm{T}} x||_{\frac{2k}{2k-1}} \leq 1, \tag{B.6}$$

where the superscript T denotes transpose. Note that $||\cdot||_{\frac{2k}{2k-1}}$ is the dual norm of $||\cdot||_{2k}$. We will also use the fact that $||\cdot||_{\frac{2k+2}{2k+1}}$ is the dual norm of $||\cdot||_{2k+2}$.

The geometry of the dual problem is depicted in Figure B.1. Here, C_k labels the set of points x in $\mathbb{R}^4$ with $||A^{\mathrm{T}}x||_{\frac{2k}{2k-1}} \leq 1$. The point y^k is the maximizer of the dual problem, and the value of the dual problem is $||b||_2 \times ||z^k||_2$ (where $||\cdot||_2$ is the ordinary Euclidean norm). Figure B.1 also depicts the dual problem for $k+1$. Thus, C_{k+1} labels the set of points x in $\mathbb{R}^4$ with $||A^{\mathrm{T}}x||_{\frac{2k+2}{2k+1}} \leq 1$, y^{k+1} is the maximizer of the dual problem for $k+1$, and $||b||_2 \times ||z^{k+1}||_2$ is the corresponding value.

By equation (5) in the main paper and equations (B.1) and (B.2) above, we see that $(r_x, r_y, r_z) \in \mathcal{R}_{2k}$ if and only if the value $||q||_{2k}$ of the primal problem for k is no more than $(1/2)^{\frac{2k-1}{k}}$. Similarly, $(r_x, r_y, r_z) \in \mathcal{R}_{2k+2}$ iff the value of the primal problem for $k+1$ is no more than $(1/2)^{\frac{2k+1}{k+1}}$. Strong duality holds, so the values of each primal and dual are equal. Therefore, the two conditions are, equivalently, that $||b||_2 \times ||z^k||_2 \leq (1/2)^{\frac{2k-1}{k}}$ and $||b||_2 \times ||z^{k+1}||_2 \leq (1/2)^{\frac{2k+1}{k+1}}$.

We want to show that $(r_x, r_y, r_z) \in \mathcal{R}_{2k}$ implies $(r_x, r_y, r_z) \in \mathcal{R}_{2k+2}$. Using

$$\frac{2k-1}{k} + \frac{1}{k(k+1)} = \frac{2k+1}{k+1}, \tag{B.7}$$

it will suffice to show that

$$\frac{||z^{k+1}||_2}{||z^k||_2} \leq (\frac{1}{2})^{\frac{1}{k(k+1)}}. \tag{B.8}$$

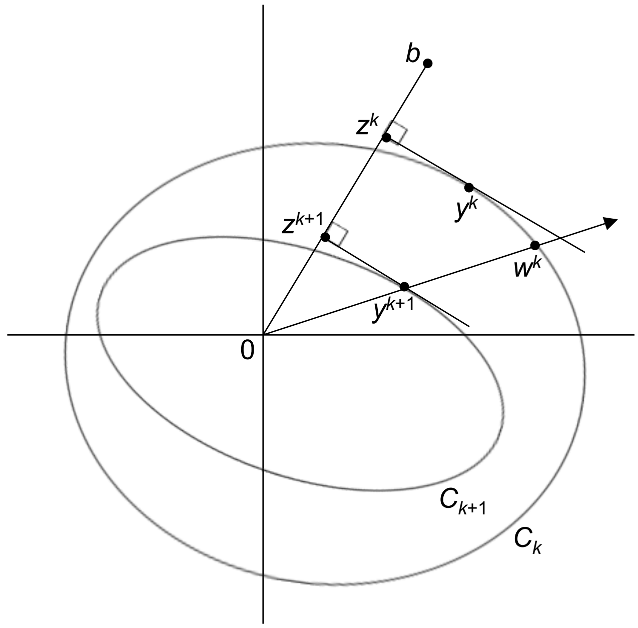

Figure B.1: Dual Problem

By convexity of C_k and similar triangles:

$$\frac{||z^{k+1}||_2}{||z^k||_2} \leq \frac{||y^{k+1}||_2}{||w^k||_2}. \qquad \text{(B.9)}$$

We now find the maximum value of the ratio $||y^{k+1}||_2/||w^k||_2$ as w^k moves around the boundary of C_k. Writing $y^{k+1} = \lambda w^k$, for some λ between 0 and 1, we find:

$$\lambda = \frac{1}{||A^{\mathrm{T}} w^k||_{\frac{2k+2}{2k+1}}}. \qquad \text{(B.10)}$$

In fact, we will write:

$$\lambda = \frac{||A^{\mathrm{T}} w^k||_{\frac{2k}{2k-1}}}{||A^{\mathrm{T}} w^k||_{\frac{2k+2}{2k+1}}} = \frac{||A^{\mathrm{T}} w||_{\frac{2k}{2k-1}}}{||A^{\mathrm{T}} w||_{\frac{2k+2}{2k+1}}} \qquad \text{(B.11)}$$

for any $w \neq 0$ lying on the ray from 0 through w^k. (This uses homogeneity of norms.)

A calculation gives, for any numbers $m, n > 0$:

$$\frac{\partial}{\partial w_4}\left(\frac{||A^{\mathrm{T}} w||_m}{||A^{\mathrm{T}} w||_n} \right) =$$

$$\frac{1}{||A^{\mathrm{T}} w||_n^2} \times \left[||A^{\mathrm{T}} w||_n \times \left(\frac{||A^{\mathrm{T}} w||_{m-1}}{||A^{\mathrm{T}} w||_m} \right)^{m-1} - ||A^{\mathrm{T}} w||_m \times \left(\frac{||A^{\mathrm{T}} w||_{n-1}}{||A^{\mathrm{T}} w||_n} \right)^{n-1} \right]. \qquad \text{(B.12)}$$

Note that $||A^{\mathrm{T}} w||_n \geq ||A^{\mathrm{T}} w||_m$ whenever $m \geq n$, and $||A^{\mathrm{T}} w||_{m-1} \geq ||A^{\mathrm{T}} w||_m$ and $||A^{\mathrm{T}} w||_{n-1} \geq$

$||A^{\mathrm{T}}w||_n$ for all m, n. Setting $m = 2k/(2k-1)$ and $n = (2k+2)/(2k+1)$ we therefore find:

$$\frac{\partial}{\partial w_4}\left(\frac{||A^{\mathrm{T}}w||_{\frac{2k}{2k-1}}}{||A^{\mathrm{T}}w||_{\frac{2k+2}{2k+1}}} \right) \geq 0. \tag{B.13}$$

Again using the fact that the value of the ratio is constant on a given ray, we conclude that the ratio (B.11) is maximized on the ray from 0 through $(0, 0, 0, 1)$. Using $A^{\mathrm{T}}w = (1, 1, 1, 1, 1, 1, 1, 1)^{\mathrm{T}}$ when $w = (0, 0, 0, 1)^{\mathrm{T}}$, we find that this maximum value is $8^{\frac{2k-1}{2k}}/8^{\frac{2k+1}{2k+2}} = 8^{-\frac{1}{2k(k+1)}} = [1/(2\sqrt{2})]^{\frac{1}{k(k+1)}} < (1/2)^{\frac{1}{k(k+1)}}$, as required.

Appendix C Proof that the Set $\mathcal{R}_2$ is the Unit Ball in $\mathbb{R}^3$

By equation (6) in the main paper, we need to show that $r_x^2 + r_y^2 + r_z^2 \leq 1$ if and only if there is a phase-space probability distribution q that satisfies $H_2(q) \geq 2$ and represents (r_x, r_y, r_z). Similar to the previous section, we pick q to be the phase-space distribution that maximizes $H_2(q)$. Since $\log_2(\cdot)$ is strictly increasing, it will suffice to solve the program:

$$\max_{q_{000}\cdots q_{111}} -\sum_{abc} q_{abc}^2 \tag{C.1}$$

$$\text{subject to } \sum_{abc} q_{abc} = 1, \tag{C.2}$$

$$q_{000} + q_{001} + q_{010} + q_{011} = \frac{1}{2}(1 + r_x), \tag{C.3}$$

$$q_{000} + q_{001} + q_{100} + q_{101} = \frac{1}{2}(1 + r_y), \tag{C.4}$$

$$q_{000} + q_{010} + q_{100} + q_{110} = \frac{1}{2}(1 + r_z). \tag{C.5}$$

The objective is strictly concave and the constraints are linear, so there is a unique maximum and the first-order conditions are necessary and sufficient. Introducing Lagrangian multipliers λ, μ, ν, and ξ for the constraints (C.2), (C.3), (C.4), and (C.5), respectively, the first-order conditions for the Lagrangian $\mathcal{L}$ are:

$$\frac{\partial \mathcal{L}}{\partial q_{000}} = 0 \quad \to \quad 2q_{000} = \lambda + \mu + \nu + \xi, \qquad \frac{\partial \mathcal{L}}{\partial q_{001}} = 0 \quad \to \quad 2q_{001} = \lambda + \mu + \nu, \tag{C.6}$$

$$\frac{\partial \mathcal{L}}{\partial q_{010}} = 0 \quad \to \quad 2q_{010} = \lambda + \mu + \xi, \qquad \frac{\partial \mathcal{L}}{\partial q_{011}} = 0 \quad \to \quad 2q_{011} = \lambda + \mu, \tag{C.7}$$

$$\frac{\partial \mathcal{L}}{\partial q_{100}} = 0 \quad \to \quad 2q_{100} = \lambda + \nu + \xi, \qquad \frac{\partial \mathcal{L}}{\partial q_{101}} = 0 \quad \to \quad 2q_{101} = \lambda + \nu, \tag{C.8}$$

$$\frac{\partial \mathcal{L}}{\partial q_{110}} = 0 \quad \to \quad 2q_{110} = \lambda + \xi, \qquad \frac{\partial \mathcal{L}}{\partial q_{111}} = 0 \quad \to \quad 2q_{111} = \lambda. \tag{C.9}$$

Equations (C.2)-(C.9) can be solved in closed form to obtain:

$$q^*_{000} = \frac{1}{8}(1 + r_x + r_y + r_z), \qquad q^*_{001} = \frac{1}{8}(1 + r_x + r_y - r_z), \tag{C.10}$$

$$q^*_{010} = \frac{1}{8}(1 + r_x - r_y + r_z), \qquad q^*_{011} = \frac{1}{8}(1 + r_x - r_y - r_z), \tag{C.11}$$

$$q^*_{100} = \frac{1}{8}(1 - r_x + r_y + r_z), \qquad q^*_{101} = \frac{1}{8}(1 - r_x + r_y - r_z), \tag{C.12}$$

$$q^*_{110} = \frac{1}{8}(1 - r_x - r_y + r_z), \qquad q^*_{111} = \frac{1}{8}(1 - r_x - r_y - r_z). \tag{C.13}$$

From equations (C.10)-(C.13) we find:

$$\begin{aligned} H_2(q^*) = -\log_2 \frac{1}{64}[&(1 + r_x + r_y + r_z)^2 + (1 + r_x + r_y - r_z)^2 + \\ &(1 + r_x - r_y + r_z)^2 + (1 + r_x - r_y - r_z)^2 + (1 - r_x + r_y + r_z)^2 + \\ &(1 - r_x + r_y - r_z)^2 + (1 - r_x - r_y + r_z)^2 + (1 - r_x - r_y - r_z)^2], \end{aligned} \tag{C.14}$$

which after some simplification yields:

$$H_2(q^*) = -\log_2[\frac{1 + r_x^2 + r_y^2 + r_z^2}{8}]. \tag{C.15}$$

It follows that:

$$\begin{aligned} \mathcal{R}_2 &= \{(r_x, r_y, r_z) \in [-1, 1]^3 : -\log_2[\frac{1 + r_x^2 + r_y^2 + r_z^2}{8}] \geq 2\} & \text{(C.16)} \\ &= \{(r_x, r_y, r_z) \in [-1, 1]^3 : r_x^2 + r_y^2 + r_z^2 \leq 1\}, & \text{(C.17)} \end{aligned}$$

as was to be shown.

Appendix D Rényi Entropy with Signed Probabilities

Here we state axioms for Rényi entropy[1] when signed probabilities are allowed and we derive the family of entropies that satisfy these axioms. This is the family of Rényi entropies used in the main text.

Rényi considered ordinary (unsigned) probabilities and showed that his definition satisfied a list of axioms which he conjectured gave a characterization of his definition. Daróczy[4] proved the conjecture. The approach followed by Rényi 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 characterize a notion of entropy for signed measures[5], 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 $(p_1 q_1, \ldots, p_1 q_n, \ldots, p_m q_1, \ldots, p_m q_n)$ 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 D.1 ***(Real-Valuedness)*** *$H(Q)$ is a real-valued function of Q.*

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

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

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

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

Axiom D.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:*

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

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

Some comments on the axioms. The forms of Axioms D.1-D.6 are carried over without change from the corresponding axioms in Rényi[1] for the entropy of non-negative measures. Notice that Axiom D.2 is built into the set-up. The value $H((q))$ could in principle differ according to which of the $x_1, \ldots, x_n$ gets measure q. Under symmetry, only the quantity q matters and so $H(q))$ is well-defined. Axiom D.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 additional bite when applied to signed vs. unsigned measures, because simply extending the domain of ordinary entropy to negative arguments yields a complex-valued functional. (For example, in the case of Shannon entropy, the formula may involve the log of a negative number, in which case real-valuedness is lost.) Axiom D.7 strengthens a continuity requirement in the Rényi formulation. In the interior of its domain, Rényi entropy with non-negative measures is real analytic (and hence approximable by power series). Axiom D.7 imposes smoothness at $q = 0$ (no longer a boundary value of q), which ensures that real analyticity continues to hold with signed measures.

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

$$H(Q) := H_{2k}(Q) = -\frac{1}{2k-1} \log_2(\frac{\sum_i q_i^{2k}}{|\sum_i q_i|}). \quad \text{(D.2)}$$

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

Two lemmas are key to proving the theorem.

Lemma D.2 *Under Axioms D.1, D.3, D.4, and D.5, if $q \neq 0$, then $H((q)) = -\log_2 |q|$.*

Proof. Let $h(q) := H((q))$. Axioms D.1 and D.3 imply that h is real-valued and continuous. Axiom D.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[6]) with general solution $h(q) = c \log_2 |q|$, where c is a real constant. Axiom D.4 fixes $c = -1$. ■

Lemma D.3 *Under Lemma D.2 and Axioms D.5 and D.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[4]. Let Q be a signed measure. From Lemma D.2 and induction on Axiom D.6 we obtain:

$$H(Q) = H((q_1) \cup \cdots \cup (q_n)) = g^{-1}[\frac{\sum_j w((q_j))\, g(H((q_j)))}{w((q_1) \cup \cdots \cup (q_n))}] = g^{-1}[\frac{\sum_j |q_j|\, g(-\log_2 |q_j|)}{|\sum_j q_j|}]. \quad \text{(D.3)}$$

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

$$g^{-1}[\frac{\sum_{i,j} |p_i q_j|\, g(-\log_2 |p_i q_j|)}{|\sum_{i,j} p_i q_j|}] = g^{-1}[\frac{\sum_i |p_i|\, g(-\log_2 |p_i|)}{|\sum_i p_i|}] + g^{-1}[\frac{\sum_j |q_j|\, g(-\log_2 |q_j|)}{|\sum_j q_j|}]. \quad \text{(D.4)}$$

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

$$f^{-1}[\frac{\sum_{i,j} |p_i q_j|\, f(|p_i q_j|)}{|\sum_{i,j} p_i q_j|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}] \times f^{-1}[\frac{\sum_j |q_j|\, f(|q_j|)}{|\sum_j q_j|}]. \quad \text{(D.5)}$$

Setting $Q = (q)$ (so that $q \neq 0$), this becomes:

$$\frac{1}{|q|}\, f^{-1}[\frac{\sum_i |p_i|\, f(|p_i q|)}{|\sum_i p_i|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}]. \quad \text{(D.6)}$$

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

$$h_q^{-1}[\frac{\sum_i |p_i|\, h_q(|p_i|)}{|\sum_i p_i|}] = f^{-1}[\frac{\sum_i |p_i|\, f(|p_i|)}{|\sum_i p_i|}]. \quad \text{(D.7)}$$

This shows (by restricting the p_i to be non-negative) that the maps h_q and f generate the same means. By a theorem on mean values (Hardy, Littlewood, and Pólya[7]), this implies that:

$$h_q(t) = a(q) f(t) + b(q), \quad \text{(D.8)}$$

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

$$f(|q|t) = a(q) f(t) + b(q). \quad \text{(D.9)}$$

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

$$f(t) = d \log_2 t + e, \quad \text{(D.10)}$$

or:

$$f(t) = d t^{\alpha - 1} + e, \quad \text{(D.11)}$$

where $d \neq 0$, e, and $\alpha \neq 1$ are arbitrary constants. (See Hardy, Littlewood, and Pólya[8]; this step also appears in Rényi[1].) Recalling the definition of f, we then find that either:

$$g(x) = -dx + e, \quad \text{(D.12)}$$

or

$$g(x) = d 2^{(1-\alpha)x} + e, \quad \text{(D.13)}$$

as required. ■

Proof of Theorem D.1. To complete the proof of sufficiency of Axioms D.1-D.7, first observe that if, per equation (D.12), g is linear, then from equation (D.3) we find:

$$-d \cdot H(Q) + e = d \cdot \frac{\sum_i |q_i| \log_2 |q_i|}{|\sum_i q_i|} + e \cdot \frac{\sum_i |q_i|}{|\sum_i q_i|}. \tag{D.14}$$

Similarly, if, per equation (D.13), g is exponential, then from equation (D.3) we find:

$$d \cdot 2^{(1-\alpha)H(Q)} + e = d \cdot \frac{\sum_i |q_i|^\alpha}{|\sum_i q_i|} + e \cdot \frac{\sum_i |q_i|}{|\sum_i q_i|}. \tag{D.15}$$

Now use Axiom D.7. Setting $Q = (q, 1-q)$ in equation (D.14) we find that $H(Q)$ is not even C^1 at $q = 0$. Setting $Q = (q, 1-q)$ in equation (D.15) we find that $H(Q)$ is C^1 at $q = 0$ only if $e = 0$ and is C^∞ at $q = 0$ only if α is an even positive integer. In this case, equation (D.15) reduces to equation (D.2), as required.

The proof of necessity of Axioms D.1-D.7 is a straightforward calculation. ■

Finally, equation (D.2) reduces to the family of Rényi entropies used in the main text when Q is a signed probability measure, i.e., when $\Sigma_i q_i = 1$.

Appendix E Discrete Wigner Function

A discrete Wigner function was introduced by Wootters[2] as an analog, for finite-dimensional systems, to the original Wigner[9] function defined for infinite-dimensional systems. The function is defined on the 4-point phase space shown in Figure E.1. This is different from the 8-point phase space which we used in the main paper. It is important that we started with the full (unreduced) state space in the main paper because our goal there was to derive the structure of the qubit from the axioms we imposed. This is not necessary in the case of the discrete Wigner function, which is a representation not a derivation.

To continue the definition, one introduces a set of phase-point operators, i.e., a set of 2×2 matrices $\mathbf{A}_{\delta_1\delta_2}$, for each point $(\delta_1, \delta_2) = (0,0), (1,0), (0,1), (1,1)$:

$$\mathbf{A}_{\delta_1\delta_2} = \frac{1}{2}[\mathbf{I} + (-1)^{\delta_2}\sigma_x + (-1)^{\delta_1+\delta_2}\sigma_y + (-1)^{\delta_1}\sigma_z], \tag{E.1}$$

where $\mathbf{I}$ is the identity matrix and σ_x, σ_y, σ_z are the Pauli matrices. Given a density matrix ρ, the associated (**discrete**) **Wigner function** W (Wootters[2]) is then given by:

$$W_{\delta_1\delta_2} = \frac{1}{2}\mathrm{tr}(\rho \mathbf{A}_{\delta_1\delta_2}), \tag{E.2}$$

for each point (δ_1, δ_2).

One can check that the Wigner function is a signed probability distribution: $W_{00} + W_{10} + W_{01} + W_{11} = 1$. It yields empirical probabilities by marginalization. For example, the probability that a measurement in the x-direction yields outcome 0 is given by the sum $W_{00} + W_{01}$.

The Wigner function can be brought into a more useful form for us by using the well-known fact that any density matrix ρ can be written in the form:

$$\rho = \frac{1}{2}(\mathbf{I} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z), \tag{E.3}$$

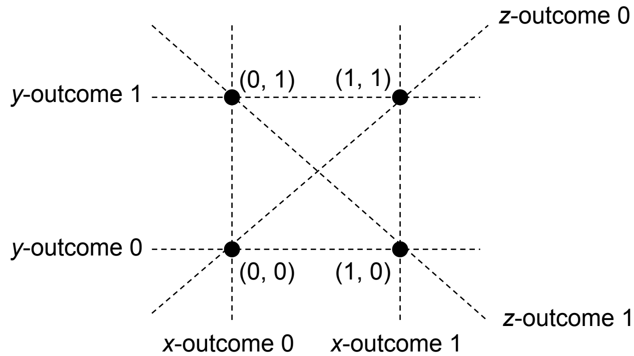

Figure E.1: Wigner Phase Space

where (r_x, r_y, r_z) lies in the Bloch sphere. We now calculate:

$$W_{00} = \frac{1}{2}\mathrm{tr}(\rho \mathbf{A}_{00}) = \frac{1}{8}\mathrm{tr}[(\mathbf{I} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z)(\mathbf{I} + \sigma_x + \sigma_y + \sigma_z)] =$$
$$\frac{1}{8}\mathrm{tr}(\mathbf{I} + r_x\mathrm{I} + r_y\mathrm{I} + r_z\mathrm{I}) = \frac{1}{4}(1 + r_x + r_y + r_z), \quad \text{(E.4)}$$

where we rely on the usual properties of the Pauli matrices. We can similarly calculate:

$$W_{10} = \frac{1}{4}(1 - r_x + r_y - r_z), \quad \text{(E.5)}$$
$$W_{01} = \frac{1}{4}(1 + r_x - r_y - r_z), \quad \text{(E.6)}$$
$$W_{11} = \frac{1}{4}(1 - r_x - r_y + r_z). \quad \text{(E.7)}$$

Comparing equations (E.4)-(E.7) with equations (C.10)-(C.13), we find the following relationships:

$$q^*_{000} + q^*_{001} + q^*_{010} + q^*_{011} = W_{00} + W_{10}, \quad \text{(E.8)}$$
$$q^*_{000} + q^*_{001} + q^*_{100} + q^*_{101} = W_{00} + W_{01}, \quad \text{(E.9)}$$
$$q^*_{000} + q^*_{010} + q^*_{100} + q^*_{110} = W_{00} + W_{11}, \quad \text{(E.10)}$$

from which (using $W_{00} + W_{10} + W_{01} + W_{11} = 1$) we can derive the $W_{\delta_1\delta_2}$ from the q_{abc}.

In this way, we can also view our axiomatization as yielding a representation of the qubit in terms of the discrete Wigner function.

Notes

[1]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.

[2]Wootters, W., "A Wigner-Function Formulation of Finite-State Quantum Mechanics," *Annals of Physics*, 176, 1987, 1-21.

[3]Boyd, S., and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004, pp.221-222.

[4]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, 381388.

[5]Brandenburger, A., and P. La Mura, "The Entropy of a Charge," March 2015, in preparation.

[6]Aczél, J., and J. Dhombres, *Functional Equations in Several Variables*, Cambridge University Press, 1989, pp.26-27, Equation (7) and Theorem 3.

[7]Hardy, G., J. Littlewood, and G. Pólya, *Inequalities*, 2nd edition, Cambridge University Press, 1952, Theorem 83.

[8]Op.cit., Theorem 84.

[9]Wigner, E., "On the Quantum Correction For Thermodynamic Equilibrium," *Physical Review*, 40, 1932, 749-759.