Symmetric Informationally Complete Measurements Identify the Irreducible Difference between Classical and Quantum Systems

John B. DeBrota†, Christopher A. Fuchs†, and Blake C. Stacey

Department of Physics, University of Massachusetts Boston,
100 Morrissey Boulevard, Boston MA 02125, USA

†Stellenbosch Institute for Advanced Study (STIAS),
Wallenberg Research Center at Stellenbosch University, Marais Street, Stellenbosch 7600, South Africa

(Dated: 23 January 2020)

We describe a general procedure for associating a minimal informationally complete quantum measurement (or MIC) with a purely probabilistic representation of the Born Rule. Such representations provide a way to understand the Born Rule as a consistency condition between probabilities assigned to the outcomes of one experiment in terms of the probabilities assigned to the outcomes of other experiments. In this setting, the difference between quantum and classical physics is the way their physical assumptions augment bare probability theory: Classical physics corresponds to a trivial augmentation—one just applies the Law of Total Probability (LTP) between the scenarios—while quantum theory makes use of the Born Rule expressed in one or another of the forms of our general procedure. To mark the irreducible difference between quantum and classical, one should seek the representations that minimize the disparity between the expressions. We prove that the representation of the Born Rule obtained from a symmetric informationally complete measurement (or SIC) minimizes this distinction in at least two senses—the first to do with unitarily invariant distance measures between the rules, and the second to do with available volume in a reference probability simplex (roughly speaking a new kind of uncertainty principle). Both of these arise from a useful result in majorization theory. This work complements recent studies in quantum computation where the deviation of the Born Rule from the LTP is measured in terms of negativity of Wigner functions.

Quantum information theory represents a change of perspective. Rather than regarding quantum physics as a limitation on our abilities—the typical sentiment of older texts—we have learned that it can augment them. In frustrating some ambitions, it enables more subtle ones. Deviation from classicality is a resource, and the idea that this resource can be quantified as a modification of the classical probability calculus dates to the beginning of the field [1]. More recent inquiries have developed this notion precisely: The “negativity” in a Wigner-function representation of quantum states is now understood to be valuable in its own right [2–11]. But what does this line of thinking say about quantum mechanics itself? Can one, following the lead of Carnot, take what might seem a statement of “mere” engineering and find a physical principle? In this paper, we prove some strong results in this regard in the context of finite dimensional Hilbert spaces. In particular, we find the unique form of the quantum mechanical Born Rule that makes it resemble the classical Law of Total Probability (LTP) as closely as possible in at least two senses. Both come from a significant majorization result which may be of general interest for resource theory. This way of tackling the distinction between quantum and classical arises naturally in the quantum interpretive project of QBism [12,43], where the Born Rule is seen as an empirically motivated constraint that one adds to probability theory when using it in the context of alternative (complementary) quantum experiments. We expect the techniques developed here to give an alternative way to explore the paradigm of negativity and to be of use for a range of practical problems.

The standard procedure in quantum theory for generating probabilities starts with an observer, or agent, assigning a quantum state \( \rho \) to a system. When the agent plans to measure the system, she represents the outcomes of her measurement with a positive operator-valued measure (POVM) \( \{ D_j \} \). Assigning \( \rho \) implies that she assigns the Born Rule probabilities \( Q(D_j) = \text{tr} \, \rho D_j \) for the outcomes of her measurement. In this way, any quantum state \( \rho \) may be regarded as a compilation of probability distributions for all possible measurements. However, one does not have to consider all possible measurements to completely specify \( \rho \). In fact, there exist measurements which are informationally complete (IC) in the sense that \( \rho \) is uniquely specified by the agent’s expectations for the outcomes of that single measurement [14]. With respect to an IC measurement, any quantum state, pure or mixed, is equivalent to a single probability distribution. In this paper, we consider minimal informationally-complete POVMs (MICs) for finite dimensional quantum systems. These sets of operators form bases for the vector space of Hermitian operators and lead to probability distributions with the fewest number of entries necessary for reconstructing the quantum state. MICs furnish a convenient way to bypass the language of quantum states, making quantum theory analogous to classical stochastic process theory, in which one puts probabilities in and gets probabilities out.

One can eliminate the need to use the operators \( \rho \) and
$D_j$ in the Born Rule by reexpressing it as a relation between an agent’s expectations for different experiments. Suppose our agent has a preferred reference process consisting of a measurement to which she ascribes the MIC $\{H_i\}$, and, upon obtaining outcome $i$, the preparation of a state $\sigma_i$, drawn from a linearly independent set of post-measurement states $\{\sigma_i\}$. (See Fig. [1]) In her choice of this reference process, she requires linearly independent post-measurement states so that the inner product $D_j \sigma_i$ will uniquely characterize the operators $D_j$. Let $P(H_i)$ be her probabilities for the measurement $\{H_i\}$ and $P(D_j|H_i)$ be her conditional probabilities for a subsequent measurement of $\{D_j\}$. What consistency requirement among $Q(D_j)$, $P(H_i)$, and $P(D_j|H_i)$ does quantum physics entail?

Using the fact that $\{\sigma_i\}$ is a basis, we may write

$$\rho = \sum_j \alpha_j \sigma_j,$$

for some set of real coefficients $\alpha_j$. The probability of outcome $H_i$ is then

$$P(H_i) = \sum_j \alpha_j \text{tr} H_i \sigma_j = \sum_j [\Phi^{-1}]_{ij} \alpha_j,$$

where we have defined the matrix $\Phi$ via its inverse,

$$[\Phi^{-1}]_{ij} := \text{tr} H_i \sigma_j = h_i \text{tr } \rho_i \sigma_j,$$

for $\rho_i := H_i / h_i$ and $h_i := \text{tr } H_i$. The invertibility of $\Phi$ is assured by the linear independence of the MIC and post-measurement sets. This implies that the coefficients of $\rho$ in the $\sigma_i$ basis may be written as an application of the $\Phi$ matrix on the vector of probabilities,

$$\rho = \sum_i \left( \sum_k [\Phi]_{ik} P(H_k) \right) \sigma_i.$$

(4)

The probability of $D_j$ is given by another application of the Born Rule, which becomes

$$Q(D_j) = \sum_{i=1}^d \left( \sum_{k=1}^d [\Phi]_{ik} P(H_k) \right) P(D_j|H_i),$$

where $P(D_j|H_i) = \text{tr } D_j \sigma_i$ is the probability for outcome $D_j$ conditioned on obtaining $H_i$ in the reference measurement. In more compact matrix notation, we can write

$$Q(D) = P(D|H) \Phi P(H),$$

where $P(D|H)$ is a matrix of conditional probabilities.

A SIC \cite{16 25} is a MIC for which all the $H_i$ are rank-1 and

$$\text{tr } H_i H_j = \frac{1}{d^2} \delta_{ij} + \frac{1}{d+1}.$$

(7)

SICs have yet to be proven to exist in all finite dimensions $d$, but they are widely believed to \cite{24} and have even been experimentally demonstrated in some low dimensions \cite{26 29}. The SIC projectors associated with a SIC are the pure states $\rho_i = dH_i$. In dimension 2, a SIC can be represented as a regular tetrahedron inscribed in the Bloch sphere. (States defining a qubit SIC can be extracted from Feynman’s 1987 essay “Negative probabilities” \cite{30}.) In higher dimensions, they are, of course, harder to visualize. When there is no chance of confusion, we will refer to the set of projectors as SICs as well. Prior work has given special attention to the reference procedure where the measurement and post-measurement states are the same SIC \cite{12 31 32}. In this case we denote $\Phi$ by $\Phi_{\text{SIC}}$ and Eq. (5) takes the particularly simple form

$$Q(D_j) = \sum_{i=1}^d \left( (d+1) P(H_i) - \frac{1}{d} \right) P(D_j|H_i).$$

(8)

Recall that the LTP expresses the simple consistency relation between the probabilities one assigns to the second of a sequence of measurements, the probabilities one assigns to the first, and the conditional probabilities for the second given the outcome of the first. Written in vector notation, this is

$$P(D) = P(D|H) P(H).$$

(9)

We write $P(D)$ as opposed to $Q(D)$ to indicate that it is the probability vector for the second of two measurements. $Q(D)$, on the other hand, is the vector of probabilities associated with a single measurement. Aside from the presence of $\Phi$ matrix, Eq. (6) is functionally equivalent to the LTP.
Although $P(H)$, $P(D|H)$, and $Q(D)$ are probabilities, $\Phi P(H)$ often is not. One may see by summing both sides of Eq. (5) over $j$ that the vector is normalized, but in general it may contain negative numbers and values greater than 1. Such a vector is known as a quasi-probability, and matrices like $\Phi$—real-valued matrices with columns summing to 1—which take probabilities to quasiprobabilities are called column-quasistochastic matrices [33].

The subset of column-quasistochastic matrices with non-negative entries are the column-stochastic matrices. The inverse of a column-stochastic matrix is generally a column-quasistochastic matrix; in our case, inspection of Eq. (5) reveals that $\Phi^{-1}$ is column-stochastic.

What would it mean if $\Phi$ could equal $I$? In this case we would have $Q(D) = P(D)$. Then, conceptually, it wouldn’t matter if the intermediate measurement were performed or not. Put another way, we could behave as though measurements simply revealed a preexisting property of the system, as in classical physics where measurements provide information about a system’s coordinates in phase space.

Some amount of what makes quantum theory non-classical resides in the fact that $\Phi$ cannot equal $I$. How close, then, can we make $\Phi$ to $I$ by wisely choosing our MIC and post-measurement states? It turns out that $\Phi_{\text{SIC}}$ is closest to the identity with respect to the distance measure induced by any member of a large family of operator norms called unitarily invariant norms (see section 3.5 in [34]). A unitarily invariant norm is one such that $\|A\| = \|UA\|$ for all unitary matrices $U$ and $V$. These norms include the Schatten $p$-norms (among which are the trace norm, the Frobenius norm, and the operator norm when $p = 1, 2$, and $\infty$ respectively) and the Ky Fan $k$-norms. This result codifies the intuition that Eq. (8) represents the “simplest modification one can imagine to the LTP” [15, p. 1971].

To prove this, we will make use of the theory of majorization [34] [35]. Suppose $x$ and $y$ are vectors of $N$ real numbers and that $x^\downarrow$ and $y^\downarrow$ are $x$ and $y$ sorted in non-increasing order. Then we say that $x$ weakly majorizes $y$ from below, denoted $x \succ_w y$, if

$$\sum_{i=1}^{k} x_i^\downarrow \geq \sum_{i=1}^{k} y_i^\downarrow, \quad \text{for } k = 1, \ldots, N. \quad (10)$$

If the last inequality is an equality, we say $x$ majorizes $y$, denoted $x \succ y$.

Another variant of majorization, called log majorization or multiplicative majorization, is also studied [35]. We say that $x$ weakly log majorizes $y$ from below, denoted $x \succ_w \log y$, if

$$\prod_{i=1}^{k} x_i^\downarrow \geq \prod_{i=1}^{k} y_i^\downarrow, \quad \text{for } k = 1, \ldots, N. \quad (11)$$

If the last inequality is an equality, we say $x$ log majorizes $y$, denoted $x \succ \log y$. Taking the log of both sides of Eq. (11) demonstrates that log majorization is majorization between the vectors after an element-wise application of the log map. Log majorization is strictly stronger than regular majorization; $x \succ_w \log y \Rightarrow x \succ_w y$, but the reverse implication is not true. Majorization is a partial order on vectors of real numbers sorted in non-increasing order.

Throughout this paper we will make use of the standard inequalities between the arithmetic, geometric, and harmonic means for vectors of $n$ positive numbers $x_i$:

$$\frac{1}{n} \sum_{i=1}^{n} x_i \geq \left( \prod_{i=1}^{n} x_i \right)^{1/n} \geq \left( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{x_i} \right)^{-1}. \quad (12)$$

with equality in all cases if and only if $x_i = c$ for all $i$. We now turn to two lemmas.

**Lemma 1.** Let $\Phi_p$ denote the column-quasistochastic matrix associated with a MIC and a proportional post-measurement set. Then $\det \Phi_p \geq \det \Phi_{\text{SIC}}$ with equality iff the MIC is a SIC.

**Proof.** We may write $\Phi_p^{-1} = GA^{-1}$ where $G_{ij} := \text{tr} H_i H_j$ is the Gram matrix of the MIC elements and $A_{ij} := h_i \delta_{ij}$. Note that $\Phi_p^{-1}$ has real, positive eigenvalues because it has the same spectrum as the positive definite matrix $A^{-1/2}GA^{-1/2}$. Also note that

$$\sum_i \frac{1}{\lambda_i(\Phi_p)} = \text{tr} \Phi_p^{-1} = \sum_i h_i \text{tr} \rho_i \sigma_i \leq \sum_i h_i = d. \quad (13)$$

One of the eigenvalues of $\Phi_p$, which we denote $\lambda_d(\Phi_p)$, must equal 1 because an equal-entry row vector is always a left eigenvector with eigenvalue 1 of a matrix with columns summing to unity. Therefore, we may write

$$\sum_{i \leq d} \frac{1}{\lambda_i(\Phi_p)} \leq d - 1. \quad (14)$$

The reciprocal of this expression is proportional to the harmonic mean of the first $d^2-1$ eigenvalues of $\Phi_p$. Thus, because the geometric mean is always greater than or equal to the harmonic mean,

$$\left( \frac{d^2-1}{\sum_{i=1}^{d^2-1} \lambda_i(\Phi_p)} \right)^{d^2-1} \geq \left( \frac{1}{d^2-1} \sum_{i=1}^{d^2-1} \frac{1}{\lambda_i(\Phi_p)} \right)^{-1} \geq d+1, \quad (15)$$

which, noting that $\lambda_d(\Phi_p) = 1$, implies

$$\det \Phi_p \geq (d+1)^{d^2-1} = \det \Phi_{\text{SIC}}. \quad (16)$$

Equality is achieved in this iff all the $\lambda_i(\Phi_p)$ are equal, so Eq. (16) is saturated iff $\lambda(\Phi_p) = \lambda_{\text{SIC}}$. We next show this implies that in fact the MIC is a SIC.

For any $\Phi_p^{-1}$, we may write $\Phi_p^{-1} = P^{-1}D P$ where the rows of $P$ are the left-eigenvectors of $\Phi_p^{-1}$ and $D$ is the diagonal matrix of eigenvalues of $\Phi_p^{-1}$. Since $\Phi_p^{-1}$ is
column-stochastic, the row vector \((1/d, \ldots, 1/d)\) is the (scaled) left-eigenvector of \(\Phi_p^{-1}\) with eigenvalue 1, and so it is the first row of \(P\) when the eigenvalues are in descending order. Left-eigenvectors of a matrix are right-eigenvectors of the transpose of the matrix, so we have

\[
(\Phi_p^{-1})^T|v\rangle = A^{-1}G|v\rangle = A^{-1}GA^{-1}|v\rangle = A^{-1}\Phi_p^{-1}A|v\rangle = \lambda|v\rangle,
\]

where \(|v\rangle\) is an arbitrary left-eigenvector of \(\Phi_p^{-1}\). Combined with our choice of scale for the first row of \(P\), we conclude that the first column of \(P^{-1}\) is 

\[(h_1, h_2, \ldots, h_d)^T.
\]

Now suppose \(\Phi_p\) is such that \(\lambda(\Phi_p) = \lambda(\Phi_{\text{SIC}})\). Then \(G = P^{-1}DA\) where \([D]_{ij} = \frac{1}{d+1}(\delta_{ij} + d\delta_{i1}\delta_{j1})\), and

\[
[G]_{ij} = \sum_{klm} [P^{-1}]_{ik}[D]_{kl}[P]_{lm}[A]_{mj}
\]

\[
= \frac{1}{d+1} \sum_{klm} [P^{-1}]_{ik} \delta_{kl} + d\delta_{k1}\delta_{l1} [P]_{lm} \delta_{mj} h_m
\]

\[
= \frac{1}{d+1} \sum_{kl} [P^{-1}]_{ik} \delta_{kl} + d\delta_{k1}\delta_{l1} [P]_{lj} h_j
\]

\[
= \frac{1}{d+1} (h_j \delta_{ij} + d h_j [P^{-1}]_{i1} [P]_{1j})
\]

\[
= \frac{1}{d+1} (h_j \delta_{ij} + d h_j).
\]

In the last step we used that \([P]_{ij} = 1/d\) and \([P^{-1}]_{i1} = h_i\). If this Gram matrix comes from a MIC, one may use

\[
[G]_{ii} = h_i^2 \text{tr} \rho_i^2 = \frac{1}{d+1}(h_i + h_i^2),
\]

and the fact that \(\text{tr} \rho_i \leq 1\) to show that \(h_i \geq 1/d\). As the average \(h_i\) value must be \(1/d\), this implies that \(h_i = 1/d\) for all \(i\) and furthermore that each \(\rho_i\) is rank-1. Substituting this into Eq. 18 gives

\[
[G]_{ij} = \frac{d \delta_{ij} + 1}{d^2(d+1)},
\]

that is, the MIC is a SIC and \(\Phi_p = \Phi_{\text{SIC}}\).

Let \(s(A)\) denote the vector of singular values of the matrix \(A\) in nonincreasing order. The proof of the following lemma may be found in Appendix A.

**Lemma 2.** Let \(\Phi\) be the column-quasistochastic matrix associated with an arbitrary reference process. Then

\[
|\det \Phi| \geq |\det \Phi_{\text{SIC}}|,
\]

and

\[
s(\Phi) \succ_w s(\Phi_{\text{SIC}}),
\]

with equality iff the MIC and post-measurement states are SICs. In that case,

\[
\Phi^\dagger \Phi = \Phi \Phi^\dagger = \Phi_{\text{SIC}}^2.
\]

We are now poised to prove:

**Theorem 1.** Let \(\Phi\) be the column-quasistochastic matrix associated with an arbitrary reference process. Then for any unitarily invariant norm \(\|\cdot\|\),

\[
||I - \Phi|| \geq ||I - \Phi_{\text{SIC}}||,
\]

with equality iff the MIC and post-measurement states are the same SIC.

**Proof.** By Corollary 3.5.9 in [34], every unitarily invariant norm is monotone with respect to the partial order on matrices induced by weak majorization of the vector of singular values. \(I - \Phi\) is singular with exactly one eigenvalue equal to zero, so one of its singular values is zero as well. Then

\[
s(I - \Phi) \succ_w \left\{ \frac{\sum_i s_i(I - \Phi)}{d^2 - 1}, \ldots, \frac{\sum_i s_i(I - \Phi)}{d^2 - 1} \right\}
\]

\[
\mathbf{s}(I - \Phi) \geq \left( \frac{\sum_i \lambda_i(I - \Phi) - \sum_i \lambda_i(\Phi) + 1}{d^2 - 1} \right)
\]

\[
d(d^2 - 1),
\]

where the first inequality follows Eq. 3.3.13a in [34], the second follows from the triangle inequality, and the last follows from Lemma 2.

It remains to show that the inequality is saturated only if the MIC and post-measurement states are the same SIC. From Lemma 2, we know that \(|\det \Phi| \geq \det \Phi_{\text{SIC}}\) with equality iff both the measurement and post-measurement states are SICs. Then, using the mean orderings [12], we can conclude the inequality

\[
\sum_{i=1}^{d^2} |\lambda_i(\Phi)| = 1 + \sum_{i=1}^{d^2-1} |\lambda_i(\Phi)|
\]

\[
\geq 1 + (d^2 - 1) \left( \prod_{i=1}^{d^2-1} |\lambda_i(\Phi)| \right)^{\frac{1}{d^2-1}}
\]

\[
= 1 + (d^2 - 1)|\det \Phi|^{\frac{1}{d^2-1}}
\]

\[
\geq 1 + (d^2 - 1)|\det \Phi_{\text{SIC}}|^{\frac{1}{d^2-1}} = \sum_{i=1}^{d^2} \lambda_i(\Phi_{\text{SIC}})
\]

(28)

is also saturated iff both the measurement and post-measurement states are SICs which, in turn, implies the same for the last inequality of (27).
If both the measurement and the post-measurement states are SICs, then Lemma 2 implies that Φ commutes with its own adjoint, making it a normal matrix. In turn, this implies that |λ_i(Φ)| = s_i(Φ) = s_i(Φ_{SIC}). The spectrum of Φ is the spectrum of Φ_{SIC}, but with nonzero phases allowed: \((d + 1)e^{i\beta_1}, \ldots, (d + 1)e^{i\beta_{d^2-1}}, 1\). Subject to this constraint, the triangle inequality is saturated if and only if all the eigenvalues are real and positive, which is also when tr Φ and tr Ψ are maximized. Considering the latter, the diagonal elements are \((1/d)tr Π_i Π_i'\) for SIC projectors \(\{Π_i\}\) and \(\{Π_i'\}\) and so the maximum trace is obviously attained only if Π_i = Π_i' for all i. So, in order for the sum of the singular values \(s_i(I - Φ)\) to attain its lower bound, both sets must be the same SIC.

It is known that no quasiprobability representation of quantum theory can be entirely nonnegative [30]. What does this mean in our formalism?

Let \(\mathcal{N}\) be the normalized hyperplane of \(d^2\)-element quasiprobability vectors. Within this is the \((d^2 - 1)\)-simplex of probability vectors, \(Δ\). For any MIC, \(d\)-dimensional quantum state space \(Q_d\) is mapped by the Born Rule to a convex subset of \(Δ\), denoted \(\mathcal{P}\). Note that \(Φ^{-1}(Δ)\) is equal to the convex hull of the \(d^2\) probability vectors tr \(H_j σ_i\), that is, the probabilities for the MIC measurement for each post-measurement state. Consequently, \(Φ^{-1}(Δ) \subset \mathcal{P}\), which implies \(Δ \subset Φ(\mathcal{P})\). These inclusions must be strict, i.e., \(Φ \neq I\): When the MIC and post-measurement states are rank-1, the vertices of the simplex will be among the pure-state probability vectors, but \(\mathcal{P}\) contains more pure states than there are vertices of \(Φ^{-1}(Δ)\). Since the image of some probability vectors consistent with quantum theory must leave the probability simplex under the application of \(Φ\), we have demonstrated that the appearance of negativity is unavoidable in our framework and is in fact characterized by the fact that \(Φ\) cannot equal the identity. Figure 2 illustrates the situation.

The weak log majorization result of Lemma 2 has at least one more important implication for quantifying the quantum deviation from classicality. Instead of looking at the functional form of Eq. (6) and considering how much of a deviation from the LTP it represents, one may approach the problem from a geometric perspective.

Classically one can always imagine assigning probability 1 to an outcome of a putative “maximally informative measurement”—for instance when one knows the system’s exact phase space point. However, in an interpretation of quantum theory without hidden variables, whatever one might mean by “maximally informative,” one cannot mean that the reference measurement’s full probability simplex is available. Indeed, quantum mechanics does not allow probability 1 for the outcome of any MIC measurement [37]. Thus deviation from classicality can also be captured by the fact that the region of probabilities compatible with quantum states is strictly smaller than the full \((d^2 - 1)\)-simplex. In this setting, the irreducible deviation from classicality is defined by the largest possible region for a reference measurement’s probability simplex. The following theorem establishes that a SIC measurement uniquely maximizes the Euclidean volume of this region, thereby answering a question raised by one of us in 2002 [16] pp. 475, 571].

**Theorem 2.** For any MIC in dimension \(d\), let \(\mathcal{P}\) denote the image of \(Q_d\) under the Born Rule and let \(vol_E(\mathcal{P})\) denote its Euclidean volume. Then

\[
vol_E(\mathcal{P}) ≤ vol_E(\mathcal{P}_{SIC}) ,
\]

with equality iff the MIC is a SIC. Furthermore,

\[
vol_E(\mathcal{P}_{SIC}) = \sqrt{\frac{(2π)^d(d-1)}{d^{d^2-1}} \frac{Γ(1)\cdots Γ(d)}{Γ(d^2)}} .
\]

The proof of Theorem 2 involves methods of differential geometry which would be distracting here. We direct the interested reader to Appendix B for details.

The \((d^2 - 1)\)-simplex \(Δ\) has Euclidean volume [38]

\[
vol_E(Δ) = \frac{d}{Γ(d^2)} ,
\]

so we can calculate the ratio of the Euclidean volumes of \(\mathcal{P}_{SIC}\) and the simplex it lies within,

\[
\frac{vol_E(\mathcal{P}_{SIC})}{vol_E(Δ)} = \sqrt{\frac{(2π)^d(d-1)}{d^{d^2-1}} \frac{Γ(1)\cdots Γ(d)}} .
\]

When \(d = 2\), quantum state space is the Bloch ball and \(\mathcal{P}_{SIC}\) is the largest ball which can be inscribed in the regular tetrahedron \(Δ_3\),

\[
\frac{vol_E(\mathcal{P}_{SIC})}{vol_E(Δ_3)} = \frac{π}{6\sqrt{3}} ≈ 0.3023 .
\]
When \(d = 3\),
\[
\frac{\text{vol}(P_{\text{SIC}})}{\text{vol}(\Delta_3)} = \frac{\pi^3}{1296\sqrt{3}} \approx 0.0138.
\] (34)

In general, the ratio is very rapidly decreasing, signifying a greater and greater deviation from classicality with each Hilbert space dimension.

Theorems [1] and [2] show that the SICs provide a way of casting the Born Rule in wholly probabilistic terms, which by two different standards make the difference between classical and quantum as small as possible. Of all the representations deriving from our general procedure, the representation given by Eq. [3] is the essential one for specifying how quantum is quantum.

**Acknowledgements.** CAF thanks K. Kato and M. P. Müller for discussions and A. Khrennikov for accusing us of being “really addicted” to the SICs for no good cause [39]—this paper attempts to show him he was wrong. JBD and CAF were supported in part by the Foundational Questions Institute Fund on the Physics of the Observer (grant FQXi-RFP-1612), a donor advised fund at the Silicon Valley Community Foundation.
Appendix A: Proof of Lemma 2

For a MIC $\{E_i\}$ and a post-measurement set $\{\sigma_j\}$, we have

$$[\Phi^{-1}]_{ij} = \text{tr} E_i \sigma_j .$$

The elements of the MIC may be expanded in the SIC basis

$$E_i = \sum_k [\alpha]_{ik} H_k ,$$

so we may write

$$[\Phi^{-1}]_{ij} = \sum_k [\alpha]_{ik} \text{tr} H_k \sigma_j = \sum_k [\alpha]_{ik} p(k|j) ,$$

where $p(k|j)$ is the probabilistic representation of the state $\sigma_j$ with respect to the SIC $\{H_k\}$. The $\alpha$ matrix must be invertible because it is a transformation between two bases, so the probability vectors can be written

$$p(i|j) = \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} .$$

We know that SIC probability vectors satisfy

$$\sum_i p(i|j)^2 \leq \frac{2}{d(d+1)} \forall j ,$$

so we have

$$\sum_i \left( \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} \right)^2 \leq \frac{2}{d(d+1)} \forall j .$$

Summing over $j$, we then have

$$\sum_{ij} \left( \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} \right)^2 \leq \frac{2d}{d+1} .$$

This expression is the sum of the absolute square entries of a matrix, which is equivalent to the square of the Frobenius norm of the matrix:

$$\|\alpha^{-1}\Phi^{-1}\|^2 = \sum_i s^2(\alpha^{-1}\Phi^{-1}) \leq \frac{2d}{d+1} .$$

From [34] 3.1.11, for any square matrix $A$,

$$\sum_i |\lambda_i(A)|^2 \leq \sum_i \varsigma^2(A) ,$$

so we have a general bound on the absolute squared spectrum:

$$\sum_i |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} .$$

Eq. (38) shows that $\alpha^{-1}\Phi^{-1}$ is column-stochastic and thus that one of its eigenvalues is 1, so we may write:

$$\sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} - 1 = \frac{d-1}{d+1} .$$

Now, using the arithmetic-geometric mean inequality,

$$\frac{d-1}{d+1} \geq \sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \geq (d^2-1) \left( \prod_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \right)^{\frac{1}{2^{d+1}}} \geq (d^2-1) |\det \alpha^{-1}\Phi^{-1}|^{\frac{2}{d+1}} ,$$

which implies

$$|\det \alpha^{-1}\Phi^{-1}| \leq \left( \frac{d-1}{d+1}(d^2-1) \right)^{d^{d+1}} \leq \left( \frac{1}{d+1} \right)^{d^{d+1}} \det \Phi^{-1}_{\text{SIC}} .$$

The elements of the MIC may be expanded in the SIC basis

$$[\Phi^{-1}]_{ij} = \text{tr} E_i \sigma_j .$$

The elements of the MIC may be expanded in the SIC basis

$$E_i = \sum_k [\alpha]_{ik} H_k ,$$

so we may write

$$[\Phi^{-1}]_{ij} = \sum_k [\alpha]_{ik} \text{tr} H_k \sigma_j = \sum_k [\alpha]_{ik} p(k|j) ,$$

where $p(k|j)$ is the probabilistic representation of the state $\sigma_j$ with respect to the SIC $\{H_k\}$. The $\alpha$ matrix must be invertible because it is a transformation between two bases, so the probability vectors can be written

$$p(i|j) = \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} .$$

We know that SIC probability vectors satisfy

$$\sum_i p(i|j)^2 \leq \frac{2}{d(d+1)} \forall j ,$$

so we have

$$\sum_i \left( \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} \right)^2 \leq \frac{2}{d(d+1)} \forall j .$$

Summing over $j$, we then have

$$\sum_{ij} \left( \sum_k [\alpha^{-1}]_{ik} [\Phi^{-1}]_{kj} \right)^2 \leq \frac{2d}{d+1} .$$

This expression is the sum of the absolute square entries of a matrix, which is equivalent to the square of the Frobenius norm of the matrix:

$$\|\alpha^{-1}\Phi^{-1}\|^2 = \sum_i s^2(\alpha^{-1}\Phi^{-1}) \leq \frac{2d}{d+1} .$$

From [34] 3.1.11, for any square matrix $A$,

$$\sum_i |\lambda_i(A)|^2 \leq \sum_i \varsigma^2(A) ,$$

so we have a general bound on the absolute squared spectrum:

$$\sum_i |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} .$$

Eq. (38) shows that $\alpha^{-1}\Phi^{-1}$ is column-stochastic and thus that one of its eigenvalues is 1, so we may write:

$$\sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} - 1 = \frac{d-1}{d+1} .$$

Now, using the arithmetic-geometric mean inequality,

$$\frac{d-1}{d+1} \geq \sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \geq (d^2-1) \left( \prod_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \right)^{\frac{1}{2^{d+1}}} \geq (d^2-1) |\det \alpha^{-1}\Phi^{-1}|^{\frac{2}{d+1}} ,$$

which implies

$$|\det \alpha^{-1}\Phi^{-1}| \leq \left( \frac{d-1}{d+1}(d^2-1) \right)^{d^{d+1}} \leq \left( \frac{1}{d+1} \right)^{d^{d+1}} \det \Phi^{-1}_{\text{SIC}} .$$
From Eq. \[36\], we can write
\[
\text{tr } E_i E_j = \sum_{kl} \alpha_{ik} \alpha_{jl} \text{tr } H_k H_l \iff G = \alpha G_{\text{SIC}} \alpha^T
\]
\[
\iff \det G = (\det \alpha)^2 \det G_{\text{SIC}} ,
\]
where \(G\) is the MIC Gram matrix and \(G_{\text{SIC}}\) is the SIC Gram matrix. Recall the definition of the \(A\) matrix from the proof of Lemma 1. The arithmetic-geometric mean inequality shows \(\text{det } A \leq (1/d)^d\) with equality iff \(h_i = 1/d\). Then, since \(G = \Phi_p^{-1} A\), Lemma 1 shows
\[
\det G = (\det \Phi_p^{-1})(\det A) \leq (\det \Phi_{\text{SIC}}^{-1})(1/d)^d = \det G_{\text{SIC}} ,
\]
with equality iff the MIC is a SIC. This implies \((\det \alpha)^2 \leq 1\), and so \(|\det \alpha| \leq 1\). Since \(|\det \alpha^{-1} \Phi^{-1}| = |\det \alpha^{-1}|| \det \Phi^{-1}|\), we conclude that
\[
|\det \Phi^{-1}| \leq \det \Phi_{\text{SIC}}^{-1} .
\]
Equivalently, \(\det \Phi_{\text{SIC}} \leq |\det \Phi|\). Theorem 3.3.2 in [34] shows \(s(\Phi) \geq \log |\lambda(\Phi)|\) for an arbitrary matrix \(\Phi\). To show the desired weak log majorization result, we wish to prove \(|\lambda(\Phi)| \geq w \log \lambda(\Phi_{\text{SIC}})|\). For this we show weak majorization of the log of the entries.
\[
\log |\lambda(\Phi)| \geq \left(\sum_{i=1}^{d^2} \frac{\log |\lambda_i(\Phi)|}{d^2 - 1}, \ldots, \sum_{i=1}^{d^2} \frac{\log |\lambda_i(\Phi)|}{d^2 - 1}, 0\right)
\]
\[
= \left(\frac{\log |\det \Phi|}{d^2 - 1}, \ldots, \frac{\log |\det \Phi|}{d^2 - 1}, 0\right)
\]
\[
\geq w \left(\frac{\log \det \Phi_{\text{SIC}}}{d^2 - 1}, \ldots, \frac{\log \det \Phi_{\text{SIC}}}{d^2 - 1}, 0\right)
\]
\[
= (\log(d+1), \ldots, \log(d+1), 0) = \lambda(\log \lambda(\Phi_{\text{SIC}})) .
\]
Thus,
\[
s(\Phi) \geq \log |\lambda(\Phi)| \geq w \log \lambda(\Phi_{\text{SIC}}) = s(\Phi_{\text{SIC}}) .
\]
If \(\{H_i\}\) and \(\{\sigma_j\}\) are SICs, \(\Phi_{ij}^{-1} = \frac{1}{d} \text{tr } \Pi_i \Pi_j^\prime\), where \(\{\Pi_i\}\) and \(\{\Pi_j\}\) are SIC projectors in dimension \(d\). Then
\[
[\Phi^{-1} \Phi^{-1}]_{ij} = \frac{1}{d^2} \sum_k (\text{tr } \Pi_i \Pi_k^\prime)(\text{tr } \Pi_j \Pi_k^\prime)
\]
\[
= \frac{1}{d^2} \text{tr } (\Pi_i \otimes \Pi_j^\prime) \left(\sum_k \Pi_k^\prime \otimes \Pi_k^\prime\right)
\]
\[
= \frac{1}{d^2} \text{tr } (\Pi_i \otimes \Pi_j) \left(\frac{2d}{d+1} P_{\text{sym}}\right)
\]
\[
= \frac{1}{d(d+1)} \text{tr } (\Pi_i \otimes \Pi_j) \left[I \otimes I + \sum_{kl} |k\rangle \langle l| \otimes |l\rangle \langle k|\right)
\]
\[
= \frac{1}{d(d+1)} + \text{tr } \Pi_i \Pi_j = \frac{d \delta_{ij} + d + 2}{d(d+1)^2} = [\Phi_{\text{SIC}}^{-2}]_{ij} ,
\]
where \(P_{\text{sym}}\) is the projector onto the symmetric subspace of \(H_d^2\) and in the third step we employed the fact that the SICs form a minimal 2-design [17]. This shows that the modulus of \(\Phi\) is equal to \(\Phi_{\text{SIC}}\) and thus the singular values of \(\Phi\) and \(\Phi_{\text{SIC}}\) coincide. The product \(\Phi^{-1} \Phi^{-1}\) works out to be the same, by interchanging the roles of the two SICs.

On the other hand, suppose \(s(\Phi) = s(\Phi_{\text{SIC}})\). The product of all the singular values is the absolute value of the determinant [33], so \(|\det \Phi^{-1}| = |\det \Phi_{\text{SIC}}^{-1}| \Rightarrow |\det \alpha| = 1 \Rightarrow \det G = \det G_{\text{SIC}} \iff \{E_i\}\) is a SIC. Carrying through the consequences of the MIC being a SIC allows us to see from Eq. [40] that \(\sigma_j\) is rank-1 because the upper bound is saturated for SIC probability vectors. We may expand the \(\{\sigma_j\}\) in the SIC projector basis,
\[
\sigma_j = \sum_k |\beta|_{jk} \Pi_k .
\]
Acting on both sides by a SIC POVM element and computing the trace of both sides, we see
\[
[\Phi^{-1}]_{ij} = \text{tr } E_i \sigma_j = \sum_k |\beta|_{jk} \text{tr } E_i \Pi_k = |\Phi_{\text{SIC}}^{-1}\beta^T|_{ij}
\]
so \(|\det \Phi^{-1}| = |\det \Phi_{\text{SIC}}^{-1}| |\det \beta^T| = |\det \Phi_{\text{SIC}}^{-1}\beta^T|\). Denoting the Gram matrix of states by \(g\), we have, in the same way as before,
\[
det g = (\det \beta)^2 \det g_{\text{SIC}} = \det g_{\text{SIC}} .
\]
We now prove that \(\det g = \det g_{\text{SIC}}\) implies that the basis of projectors forms a SIC. The following lemma is due to Huangjun Zhu [40]. We only use part of Zhu’s conclusion, but the lemma is of enough interest to present in full.

**Lemma 3 (Zhu).** Let \(\lambda\) be the spectrum of the Gram matrix \(g\) of a normalized basis of positive semidefinite operators \(\Pi_j\) sorted in nonincreasing order. Then \(\lambda \geq \lambda_{\text{SIC}}\) with equality iff \(\Pi_j\) forms a SIC.

**Proof.** By assumption \(\text{tr } \Pi_j^2 = 1\) for all \(j\). Since the eigenvalues of \(\Pi_j\) are nonnegative,
\[
1 = \text{tr } \Pi_j^2 = \sum_i \lambda_i^2(\Pi_j) \leq \sum_i \lambda_i(\Pi_j) = \text{tr } \Pi_j .
\]
Define the frame superoperator
\[
F = \sum_j \|\Pi_j\rangle\langle \Pi_j\| ,
\]
where \(\|A\rangle := \sum_{ij} |A|_{ij} |i\rangle |j\rangle\). \(F\) has the same spectrum as the Gram matrix \(g_{ij} = \langle \Pi_i | \Pi_j \rangle = \text{tr } \Pi_i \Pi_j\). To see this, form a projector out of the state \(\sum_k |\Pi_k\rangle\langle \Pi_k|\) where \(|i\rangle\) is an orthonormal basis in \(H_d^2\) and perform partial traces over each subsystem. The results are \(g^T\) and \(F\), and so, by the Schmidt theorem, the spectra of \(F\) and \(g\) are equal: \(\lambda(g) = \lambda(F) = \lambda\).
The expectation value of any operator with respect to an arbitrary normalized state is less than or equal to its maximal eigenvalue. Thus, a lower bound on the maximal eigenvalue \( \lambda_1 \) of \( F \) is given by
\[
\lambda_1 \geq \frac{1}{d} \langle I \| F \| I \rangle = \frac{1}{d} \sum_j (\text{tr} \Pi_j)^2 \geq d. \tag{59}
\]
As our basis is normalized, \( \text{tr} g = d^2 \), so \( \sum_i \lambda_i = d^2 \). With this constraint and our bound on the maximal eigenvalue, we have
\[
\lambda \succ \left( \frac{\lambda_1}{d^2 - 1}, \ldots, \frac{\lambda_1}{d^2 - 1} \right) \succ \left( \frac{d}{d + 1}, \ldots, \frac{d}{d + 1} \right) = \lambda_{\text{SIC}}. \tag{60}
\]

The second majorization becomes an equality when \( \lambda_1 = d \). From Eq. \((59)\), we can see that all \( \Pi_j \) must be rank-1 for this condition to be satisfied. Furthermore, we see that in this case \( \frac{1}{\lambda} \langle I \| I \rangle \) is an eigenvector of \( F \) which achieves the maximal eigenvalue \( d \). When both majorizations are equalities the spectrum \( \lambda_{\text{SIC}} \) tells us that \( F \) takes the form of a weighted sum of a projector and the identity superoperator \( I \), specifically
\[
F = \frac{d}{d + 1} (I + \|I\|\langle I \| I \rangle). \tag{61}
\]

By Cor. 1 in \cite{41}, this implies the \( \Pi_j \) form a SIC. \( \square \)

As in the Lemma, denote by \( \lambda \) the spectrum of \( g \) sorted in nonincreasing order. \( \text{tr} g = d^2 \), so
\[
\sum_{i>1} \lambda_i = d^2 - \lambda_1. \tag{62}
\]

Then because the arithmetic mean is greater than or equal to the geometric mean with equality iff the elements are all equal, we have
\[
\frac{1}{d^2 - 1} \sum_{i>1} \lambda_i = \frac{d^2 - \lambda_1}{d^2 - 1} \geq \left( \prod_{i>1} \lambda_i \right)^{\frac{1}{d^2 - 1}}, \tag{63}
\]
which implies
\[
\det g \leq \lambda_1 \left( \frac{d^2 - \lambda_1}{d^2 - 1} \right)^{\frac{d^2}{d^2 - 1}}, \tag{64}
\]
with equality iff \( \lambda_2 = \cdots = \lambda_{\frac{d^2}{d^2 - 1}} \). When \( \lambda_1 = d \), we then have
\[
\det g = \frac{d^2}{(d + 1)^{d^2 - 1}} = \det g_{\text{SIC}}, \tag{65}
\]
with equality iff \( \lambda = \lambda_{\text{SIC}} \). By Lemma \cite{3} we have equality iff the post-measurement states form a SIC.

Appendix B: Proof of Theorem 2

Equation \((4)\) expanded instead in the \( \rho_i \) basis allows us to relate the differential elements of operator space and probability space for any MIC basis:
\[
d\sigma = \sum_{i,j} [\Phi]_{ij} \rho_i dp^j. \tag{66}
\]
The Hilbert–Schmidt line element is then
\[
ds_{\text{HS}}^2 = \text{tr} (d\sigma)^2 = \sum_{ijkl} [\Phi]_{ij} [\Phi]_{kl} (\text{tr} \rho_i \rho_k) dp^i dp^j. \tag{67}
\]

As in the proof of Lemma \cite{1} we write \( \Phi = AG^{-1} \) where \( [G]_{ij} = \text{tr} H_i H_j \) is the Gram matrix for the MIC and \( [A]_{ij} = h_i \delta_{ij} \). Note further that \( \text{tr} \rho_i \rho_j = [A^{-1} GA^{-1}]_{ij} \).

Then Eq. \((67)\) simplifies to
\[
ds_{\text{HS}}^2 = \sum_{ij} [G^{-1}]_{ij} dp^i dp^j. \tag{68}
\]
The Hilbert–Schmidt volume element on the space of Hermitian operators in \( \mathcal{L}(\mathcal{H}_d) \) may now be related to the Euclidean volume element in \( \mathbb{R}^{d^2} \),
\[
d\Omega_{\text{HS}} = \sqrt{\det G^{-1}} dV_E, \tag{69}
\]
or, equivalently,
\[
dV_E = \sqrt{\det G} d\Omega_{\text{HS}}. \tag{70}
\]
The larger \( \det G \), the larger the corresponding Euclidean volume. Recall Eq. \((49)\) which says
\[
\det G \leq \det G_{\text{SIC}}, \tag{71}
\]
with equality iff the MIC is a SIC. Thus, for any region in operator space, the Euclidean volume is maximal with respect to the SIC basis. In particular, the SIC basis gives the largest volume among positive semidefinite operators \( A \) satisfying \( 1 - \varepsilon \leq \text{tr} A \leq 1 + \varepsilon \) for any \( \varepsilon > 0 \). As \( \varepsilon \to 0 \), we obtain quantum state space \( Q_d \) and the corresponding region in \( \mathbb{R}^{d^2} \) will have the largest hyperarea within \( \Delta \) when computed with the SIC basis.

To calculate this hyperarea, we need to find the metric on \( \Delta \) induced by the Hilbert–Schmidt metric in the SIC basis. We may parameterize \( \Delta \) by
\[
X = \left( p^1, \ldots, p^{d^2-1}, 1 - \sum_{i=1}^{d^2-1} p^i \right), \tag{72}
\]
which has partial derivatives \( \partial_i X^\mu = \delta_i^\mu - \delta_i^{\mu,2} \) where the Latin index runs from 1 to \( d^2 - 1 \) and the Greek index runs from 1 to \( d^2 \). For any MIC, the induced metric \( g \) is given by
\[
[g]_{ij} = \sum_{\mu,\nu=1}^{d^2} \partial_i X^\mu \partial_j X^\nu [G^{-1}]_{\mu \nu}. \tag{73}
\]
It is easily seen that \( G_{\text{SIC}}^{-1} = d(d + 1)I - J \) where \( J \) is the Hadamard identity. One may then calculate \( g_{\text{SIC}} = d(d + 1)(I + J) \) and \( \det g_{\text{SIC}} = d^2(d^2 + d)^{d^2-1} \). The induced volume element on \( \Delta \) is then

\[
d\omega_{\text{HS}} = d\sqrt{(d^2 + d)^{d^2-1}dp^1 \cdots dp^{d^2-1}}. \tag{74}
\]

In a similar way, it may be checked that the Euclidean metric in \( \mathbb{R}^{d^2} \) induces a volume element \( dA_{\text{E}} \) on \( \Delta \) satisfying

\[
\frac{1}{d}dA_{\text{E}} = dp^1 \cdots dp^{d^2-1}, \tag{75}
\]

and so

\[
d\omega_{\text{HS}} = \sqrt{(d^2 + d)^{d^2-1}}dA_{\text{E}}. \tag{76}
\]

We may now integrate over quantum state space to obtain

\[
\text{vol}_{\text{HS}}(Q_d) = \sqrt{(d^2 + d)^{d^2-1}}\text{vol}_{\text{E}}(P_{\text{SIC}}). \tag{77}
\]

Zyczkowski and Sommers \[42\] calculate the Hilbert–Schmidt volume of finite-dimensional quantum state space to be

\[
\text{vol}_{\text{HS}}(Q_d) = \sqrt{\pi^d(d^2-1)/2} \cdot \Gamma(1) \cdot \Gamma(d) / \Gamma(d^2), \tag{78}
\]

from which Eq. \((30)\) follows.