Eur. Phys. J. D (2017) 71: 197
DOI: 10.1140/epjd/e2017-80024-y

THE EUROPEAN PHYSICAL JOURNAL D

Topical Review

Introducing the Qplex: a novel arena for quantum theory

Marcus Appleby[1], Christopher A. Fuchs[2,3,a], Blake C. Stacey[4], and Huangjun Zhu[5]

[1] Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney, Australia
[2] Physics Department, University of Massachusetts Boston, Boston, MA 02125, USA
[3] Max Planck Institute for Quantum Optics, 85748 Garching, Germany
[4] Department of Physics, University of Massachusetts Boston, Boston, MA 02125, USA
[5] Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany

Received 12 January 2017 / Received in final form 18 April 2017
Published online 25 July 2017

Abstract. We reconstruct quantum theory starting from the premise that, as Asher Peres remarked, "Unperformed experiments have no results." The tools of quantum information theory, and in particular the symmetric informationally complete (SIC) measurements, provide a concise expression of how exactly Peres's dictum holds true. That expression is a constraint on how the probability distributions for outcomes of different, hypothetical and mutually exclusive experiments ought to mesh together, a type of constraint not foreseen in classical thinking. Taking this as our foundational principle, we show how to reconstruct the formalism of quantum theory in finite-dimensional Hilbert spaces. The central variety of mathematical entity in our reconstruction is the *qplex,* a very particular type of subset of a probability simplex. Along the way, by closely studying the symmetry properties of qplexes, we derive a condition for the existence of a d-dimensional SIC.

1 Introduction

The arena for standard probability theory is the probability simplex – that is, for a trial of n possible outcomes, the continuous set Δ_n of all n-vectors p with nonnegative entries $p(i)$ satisfying $\sum_i p(i) = 1$. But what is the arena for quantum theory? The answer to this question depends upon how one views quantum theory. If, for instance, one views it as a noncommutative *generalization* of probability theory, then the arena could be the convex sets of density operators and positive-operator-valued measures over a complex Hilbert space. In contrast, references [1–3] have argued that quantum theory is not so much a generalization of probability theory as an *addition* to it. This means that standard probability theory is never invalidated, but that further rules must be added to it when the subject matter concerns measurements on quantum systems. One implication of this is that behind every application of quantum theory is a more basic simplex, which through a not-yet-completely-understood consistency requirement, gets trimmed or cropped to a convex subset isomorphic to the usual space of quantum states [4][1]. In the specific context formalized below, we call an arena of this sort – a suitably cropped simplex as the starting point for a full-fledged derivation of quantum theory – a *qplex*. In a slogan: if the simplex is the starting point for probability theory, the qplex is the starting point for the quantum.

The introduction of a more basic simplex surrounding the qplex, however, should not be construed as a capitulation to the idea of a hidden-variable theory. Rather it is an attempt to bring to the front of the formalism a foundational idea nicely captured by Asher Peres's famous quip "unperformed experiments have no outcomes" [6]. Here the simplex stands for the outcomes of an "unperformed experiment": an experiment that will never be done, but could have been done. How is probability theory all by itself to connect this experiment to any other? It has no tools for the job. But quantum theory does, through the Born rule, when suitably rewritten in the language of the qplex. From this point of view, the meaning of the Born

[a] e-mail: qbism.fuchs@gmail.com

[1] See also (Ref. [5], p. 487) for the historical roots of this idea. In fact, this idea goes back to the very roots of QBism. In a 19 July 1996 note to S.L. Braunstein, one of us (CAF) wrote, "I don't think there's anything interesting to be gained from simply trying to redo [Cox's derivation of probability theory] but with complex numbers. It seems to me that it'll more necessarily be something along the lines of: When you ask me, "Where do all the quantum mechanical outcomes come from?" I must reply, "There is no where there" (with apologies to Gertrude Stein). That is to say, my favorite "happy" thought is that when we know how to properly take into account the piece of prior information that "there is no where there" concerning the origin of quantum mechanical measurement outcomes, then we will be left with "plausibility spaces" that are so restricted as to be isomorphic to Hilbert spaces. But that's just thinking my fantasies out loud."

rule for probabilities in any actual experiment is that "behind" the experiment is a different, hypothetical experiment whose probabilities *must be taken into account* in the calculation.

To be concrete, let us rewrite quantum theory in a language that would make this apparent *were the right mathematical tool available*. Consider the setting of a finite d-level quantum system, and suppose that one of the elusive symmetric informationally complete quantum measurements [7,8] exists for it. We shall call such an object a "SIC" for short. A SIC is a set of d^2 rank-one projection operators $\Pi_i = |\psi_i\rangle\langle\psi_i|$ such that

$$\mathrm{tr}(\Pi_k \Pi_l) = \frac{d\delta_{kl} + 1}{d+1}. \tag{1}$$

For such a set of operators, one can prove that if they exist at all, they must be linearly independent, and rescaling each to $\frac{1}{d}\Pi_i$, they collectively give an informationally complete positive-operator-valued measure (POVM), i.e.,

$$\sum_i \frac{1}{d}\Pi_i = I. \tag{2}$$

Thus, for any quantum state ρ, a SIC can be used to specify a measurement for which the probabilities of outcomes $p(i)$ specify ρ itself. That is, if

$$p(i) = \frac{1}{d}\,\mathrm{tr}(\rho\Pi_i), \tag{3}$$

then

$$\rho = \sum_{i=1}^{d^2}\left[(d+1)p(i) - \frac{1}{d}\right]\Pi_i \tag{4}$$

$$= (d+1)\sum_{i=1}^{d^2} p(i)\Pi_i - I. \tag{5}$$

Is it always possible to write a quantum state like this?[2] Unfortunately, to date, analytic proofs of SIC existence have only been found in dimensions 2–21, 24, 28, 30, 31, 35, 37, 39, 43 and 48[3] [10]. However, very high-precision numerical approximations (many to 8000 and 16 000 digits) have been discovered for all dimensions 2 to 151 without exception, plus some dimensions sporadically beyond that – 168, 172, 195, 199, 228, 259, 323, at last count[4] [11].

[2] To our knowledge the first person to write down this expression was the Cornell University undergraduate Gabriel G. Plunk in an attachment to a 18 June 2002 email to one of us (CAF), though it went undiscovered for many years (see Ref. [9], pp. 472–474).

[3] Exact solutions in dimensions 31, 37, 39 and 45 were found by M. Appleby in 2016; publication forthcoming.

[4] Numerical solutions in dimensions $d = 68$ through $d = 121$, and in $d = 143$, were found by Andrew J. Scott (email communication, May, 2016). Solutions for $d = 122$ through $d = 151$ were found by Michael C. Hoang on the Chimera supercomputer, in collaboration with CAF and BCS, using code developed by Scott.

In general, the mood of the community is that a SIC should exist in every finite dimension d, but we call the SICs "elusive" because in more than 18 years of effort no one has ever proven it. See reference [12] for an extensive bibliography on the subject. For the purpose of the present discussion, let us suppose that at least one SIC can be found in any finite dimension d.

One can now see how to express quantum-state space as a proper subset Q of a probability simplex Δ_{d^2} over d^2 outcomes. That it cannot be the full simplex comes about from the following consideration: for any $p \in \Delta_{d^2}$, equation (4) gives a Hermitian operator ρ with trace 1, but the operator may not be positive-semidefinite as is required of a density operator. Instead, the density operators correspond to a convex subset specified by its extreme points, the pure states $\rho^2 = \rho$. Thanks to an observation by Jones, Flammia and Linden [13,14], we can also characterize pure states as those Hermitian matrices satisfying

$$\mathrm{tr}\,\rho^2 = \mathrm{tr}\,\rho^3 = 1. \tag{6}$$

This expression of purity yields two conditions on the probability distributions p [1,2,15]. First,

$$\sum_{i=1}^{d^2} p(i)^2 = \frac{2}{d(d+1)}, \tag{7}$$

and second,

$$\sum_{ijk} c_{ijk}\, p(i)p(j)p(k) = \frac{d+7}{(d+1)^3}, \tag{8}$$

where we have defined the real-valued, completely symmetric three-index tensor

$$c_{ijk} = \mathrm{Re}\ \mathrm{tr}(\Pi_i\Pi_j\Pi_k). \tag{9}$$

The full state space Q is the convex hull of probability distributions satisfying equations (7) and (8).

So the claim can be made true, but what a strange-looking set the quantum states become when written in these terms! What could account for it except already knowing the full-blown quantum theory as usually formulated?

Nevertheless, every familiar operation in the textbook quantum formalism has its translation into the language of this underlying probability simplex, properly restricted to the subset Q. For example, given a quantum state ρ, one uses the Born rule to calculate the probabilities an experiment will yield its various outcomes with. Using the SIC representation, the description of the measuring apparatus becomes an ordinary set of conditional probabilities, $r(j|i)$. For instance, for a POVM defined by the set of effects

$$\{E_1, \ldots, E_n\},\ \sum_j E_j = I, \tag{10}$$

the Born rule tells us the probabilities $q(j)$ for its outcomes are

$$q(j) = \mathrm{tr}(\rho E_j), \tag{11}$$

but this can be reexpressed as

$$q(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i), \tag{12}$$

where

$$r(j|i) = \mathrm{tr}(E_j \Pi_i), \tag{13}$$

meets the criteria for a conditional probability distribution.

In reference [1], the simple form in equation (12) was considered so evocative of the usual law of total probability from standard probability theory, and seemingly so basic to Peres's "unperformed experiments have no outcomes" considerations, that it was dubbed the *urgleichung* – or German for "primal equation".

Similarly, if we have a quantum state ρ encoding our expectations for the SIC measurement on some system at time $t = 0$, we can evolve that state forward to deduce what we should expect at a later time, $t = \tau$. In textbook language, we relate these two quantum states by a quantum channel – in the simplest case, by a unitary operation:

$$\rho' = U\rho U^\dagger. \tag{14}$$

Let the SIC representation of ρ be $p(i)$, and let the SIC representation of ρ' be $p'(j)$. We translate the unitary U into SIC language by calculating

$$u(j|i) = \frac{1}{d}\,\mathrm{tr}\left(U\Pi_i U^\dagger \Pi_j\right). \tag{15}$$

The object u is a $d^2 \times d^2$ doubly stochastic matrix [16]. But now, something fascinating happens. The two quantum states $p(i)$ and $p'(j)$ are related according to

$$p'(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] u(j|i), \tag{16}$$

an expression identical in form to equation (12).

Formulas (12) and (16) may be compared with *what would have been given* by the standard law of total probability

$$q(j) = \sum_i p(i) r(j|i), \tag{17}$$

and the standard rule for stochastic evolution,

$$p'(j) = \sum_i p(i) u(j|i), \tag{18}$$

were they applicable. This emphasizes again that the quantum laws are different but, in the setting of a SIC-induced simplex, intriguingly similar to their classical counterparts.

This leads one to wonder whether, or to what extent, these very special forms equations (12) and (16) might imply the very arena Q in which they are valid. This is the program laid out in references [1–3] and a key motivation for the geometric studies of references [15,17,18]. Here we will carry the program much further than previously.

Another familiar operation in the standard language of quantum theory is the Hilbert-Schmidt inner product between two quantum states, $\mathrm{tr}(\rho\sigma)$. Using the SIC representations of ρ and σ as probability vectors p and s, it is straightforward to show that

$$\mathrm{tr}(\rho\sigma) = d(d+1)\langle p, s\rangle - 1. \tag{19}$$

Because the inner product of any two quantum states ρ and σ is bounded between 0 and 1, we know that

$$\frac{1}{d(d+1)} \le \langle p, s\rangle \le \frac{2}{d(d+1)}. \tag{20}$$

We designate these the *fundamental inequalities*. The upper bound is simply the quadratic constraint we saw already in equation (7), but the lower bound imposes new and surprisingly intricate conditions on the vectors that can be admissible states.

We will say that two vectors p and s in the probability simplex Δ_{d^2} are *consistent* if their inner product obeys both inequalities in equation (20). If we have a subset of the probability simplex in which every pair of vectors obeys those bounds, we call it a *germ*: It is an entity from which a larger structure can grow. If including one additional vector in a germ could make that set inconsistent, then that germ is said to be a *maximal*. We will see that a maximal germ is one way to define a *qplex*.

Any quantum state space in SIC representation is a qplex. However, the converse is not true: there exist qplexes that are not equivalent to quantum state space. That said, any qplex is already a mathematically rich structure. A primary goal of this paper is to use that richness and identify an extra condition which can be imposed upon a qplex, such that satisfying that constraint will make the qplex into a quantum state space.

In Section 2 we see how quantum physics furnishes a new way that probability assignments can mesh together, a way not foreseen in classical thinking. This will lead us from very general considerations to the specific definition of a qplex. In Section 3 we apply a tool from the theory of polytopes [19,20] to derive a number of basic results about the geometry of an arbitrary qplex. Among other applications, we find a simple, intuitively appealing proof that a polytope embedded in quantum state space cannot contain the in-sphere of quantum state space.

Sections 4 and 5 are the core of the paper. In almost every geometrical problem, a study of the symmetries of the object or objects of interest plays an essential role. However, it turns out that qplexes have the unusual property that the symmetry group, instead of having to be imposed from the outside, is contained internally to the structure. In this they might be compared with elliptic curves [21]. In spite of the extreme simplicity of the defining equation

$$y^2 = x^3 + ax + b, \tag{21}$$

elliptic curves have managed to remain at the cutting edge of mathematics for two millennia, from the work of Diophantus down to the present day. They play an important role in, for example, the recent proof of Fermat's

last theorem [22]. One of the reasons for their high degree of mathematical importance is the fact that they carry within themselves a concealed group. Qplexes have a similar property. In Sections 4 and 5 we describe this property, and examine its implications.

In Section 4 we present our main application. We apply the results established in the previous section to the SIC existence problem and show that SIC existence in dimension d is equivalent to the existence of a certain kind of subgroup of the real orthogonal group in dimension d^2-1. We presented this result in a previous publication [18], where we derived it by more conventional means. In this paper, we describe the way we originally proved it, using the qplex formulation. This is because we believe the method of proof is at least as interesting as the result itself.

In Section 6 we turn to the problem of identifying the "missing assumption" which will serve to pick out quantum state space uniquely from the set of all qplexes. Of course, as is usual in such cases, there is more than one possibility. We identify one such assumption: the requirement that the symmetry group contain a subgroup isomorphic to the projective unitary group. This is a useful result because it means that we have a complete characterization of quantum state space in probabilistic terms. It also has an important corollary: that SIC existence in dimension d is equivalent to the existence of a certain kind of subgroup of the real orthogonal group in dimension d^2-1.

Finally, we wrap up in Section 7 with list of several possible directions for future investigations. If this research program is on the right track, it is imperative that a more basic path from qplex to quantum state space be found. There is plenty of work to do here.

2 The basic scheme

The urgleichung (12) and the inequalities (20) are not independent. In this section, we will start with a generalized form of the urgleichung and, making a few additional assumptions, derive the fundamental inequalities. This is, strictly speaking, not necessary for the mathematical developments in the later sections of the paper. One can assume the fundamental inequalities as a starting point and then proceed from that premise. In fact, we will later see that using that approach, one can derive as consequences the assumptions we will invoke here. Speaking in general terms, we can think of this section as proving the "if" direction, and the following section as proving "only if." One benefit of deriving the fundamental inequalities in this manner is to help compare and contrast our reconstruction of quantum theory with other approaches [23–29]. These other reconstructions are *operational* in character: They take, as fundamental conceptual ingredients, laboratory procedures like "preparations" and "tests." Our language in this section will have a similar tone. However, we will keep Peres' dictum that "unperformed experiments have no results" at the forefront of our considerations.

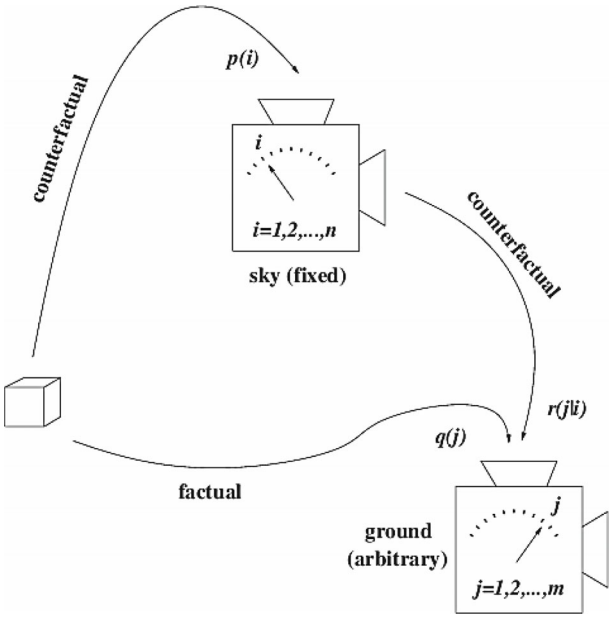

Fig. 1. Analysing one scenario in terms of another: an agent Alice intends to perform an experiment on the ground, whose outcomes she labels with the index j. The other index, i, labels the outcomes of a "Bureau of Standards" measurement which Alice *could* carry out, but which remains unperformed. Classical physics and quantum physics both allow for Bureau of Standards measurements, experiments that are *informationally complete* in the following sense. If Alice has a set of probabilities $p(i)$ for the Bureau of Standards measurement outcomes, she can calculate the proper set of probabilities $q(j)$ for the outcomes of the ground measurement, using the conditional probabilities $r(j|i)$.

Our first step is to understand how the urgleichung is an example of this principle. To do so, we consider the following scenario [3,23].

Fix a dimension $d \geq 2$, and consider a system to which we will ascribe a quantum state in d-dimensional Hilbert space $\mathcal{H}_d$. We will investigate this system by means of two measuring devices, which we model in the standard way by POVMs. One measuring device is a SIC measurement, defined by a set of d^2 rank-1 projection operators $\{\Pi_i\}$. The effects which comprise this POVM are the operators rescaled by the dimension:

$$E_i = \frac{1}{d}\Pi_i. \tag{22}$$

We will refer to this as the "Bureau of Standards" measurement. It is helpful to imagine this measuring device as being located in some comparatively inaccessible place: perhaps inside a vault, or secured in an airship floating through the sky. An agent *can* take her system of interest to the Bureau of Standards device, but she has good reason to want to bypass that step. The other measurement is an arbitrary POVM, whose effects we denote by F_j.

As illustrated in Figure 1, we will consider two experimental scenarios, which we will call the "ground path" and the "sky path." If we follow the ground path, we take

our system of interest directly to the $\{F_j\}$ measuring device, which we will call the measurement on the ground. If we instead follow the sky path, we will take our system to the Bureau of Standards measurement, physically obtain a result by performing that measurement, and then come back down for the second stage, where we conduct the measurement on the ground.

Suppose that Alice follows the sky path in Figure 1. That is, she physically takes her system of interest and performs the Bureau of Standards measurement upon it. Then, she returns the system to the ground and conducts the measurement $\{F_j\}$ there. Before carrying out the Bureau of Standards measurement, she has some expectations for what might happen, which she encodes as a probability distribution $p(i)$. Obtaining an outcome i, she updates her state assignment for the system to the operator Π_i. Her expectations for the outcome of the ground measurement will then be the conditional probabilities $r(j|i)$. Prior to performing the Bureau of Standards measurement, Alice assigns the probability

$$\text{Prob}(j) = \sum_i p(i) r(j|i), \tag{23}$$

to the event of obtaining outcome j when she brings the system back down to the ground and performs the second measurement in the sequence.

Classical intuition suggests that Alice should use the same expression for computing the probability of outcome j on the ground even if she goes directly to the ground experiment and does not perform the measurement in the sky. If $p(i)$ is the probability that she *would* obtain outcome i *were she to perform* the sky measurement, and $r(j|i)$ is the conditional probability for outcome j *if* the event i *were to occur* in the sky, then it is almost instinctive to calculate the probability of j by summing $p(i)r(j|i)$. Mathematically, this is *not necessarily correct,* because the ground path and the sky path are two different physical scenarios. If C_1 and C_2 are two background conditions, then nothing in probability theory forces $\text{Prob}(j|C_1) = \text{Prob}(j|C_2)$. Writing $q(j)$ for the probability of obtaining j by following the ground path, we have that

$$q(j) \text{ is not necessarily equal to } \sum_i p(i) r(j|i). \tag{24}$$

It is merely the assumption that an informationally complete measurement must be measuring some pre-existing physical property of the system that leads Alice to use equation (23) even when she does not physically obtain an outcome in the sky. In other words, using equation (23) to calculate $q(j)$ amounts to assuming that the measurement outcome i is *as good as existing,* even when it remains completely counterfactual.

Probability theory itself does not tell us how to find $q(j)$ in terms of $p(i)$ and $r(j|i)$. Classical intuition suggests one way of augmenting the abstract formalism of probability theory: using equation (23). The crucial point is that *quantum theory gives us an alternative.* It is simply to use the Born rule, in the form of the urgleichung.

The Born-rule probability for obtaining the outcome with index j is

$$q(j) = \text{tr}(\rho F_j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i), \tag{25}$$

where

$$r(j|i) = \text{tr}(\Pi_i F_j). \tag{26}$$

Note that $r(j|i)$ is also the probability that the Born rule would tell us to assign to the outcome j if our quantum state assignment for the system were Π_i.

Probability theory is a way to augment our raw experiences of life: it provides a means to manage our expectations carefully. In turn, quantum theory augments the mathematics of probability, furnishing links between quantities that, considering only the formalism of probability theory, would be unrelated. These new relationships are quantitatively precise, but at variance with classical intuition, reflecting the principle that unperformed experiments have no outcomes.

We now explore the consequences of relating mutually exclusive hypothetical scenarios by the urgleichung. Using seven assumptions, of which the urgleichung is the most radical, we will arrive at the fundamental inequalities (20). Because the constants d^2 and $d+1$ and $1/d$ look rather arbitrary at first glance, we will begin with a more general expression.

Assumption 1. *The Generalized Urgleichung. Given a Bureau of Standards probability distribution* $\{p(i) : i = 1, \ldots, N\}$*, and a matrix of conditional probabilities* $r(j|i)$*, we compute the probabilities for an experiment on the ground by means of*

$$q(j) = \sum_{i=1}^{N} \left[\alpha p(i) - \beta\right] r(j|i). \tag{27}$$

In what follows, this will be our primary means of relating one probability distribution to another. The basic normalization requirements are

$$\sum_i p(i) = 1, \ \sum_j r(j|i) = 1, \ \sum_j q(j) = 1. \tag{28}$$

Normalization relates the constants α, β and N:

$$\alpha = N\beta + 1. \tag{29}$$

We denote the set of valid states p by $\mathcal{P}$, and the set of valid measurements by $\mathcal{R}$. For any $p \in \mathcal{P}$ and any $r(j|i) \in \mathcal{R}$, the vector q calculated using the urgleichung is a proper probability distribution.

If we take any $r \in \mathcal{R}$ and sum over both indices, we find that

$$\sum_{i,j} r(j|i) = \sum_i 1 = N. \tag{30}$$

Assumption 2. *Maximality. The set of all states* $\mathcal{P}$ *and the set of all measurements* $\mathcal{R}$ *together have the property that no element can be added to either without introducing an inconsistency, i.e., a pair* $(p \in \mathcal{P}, r \in \mathcal{R})$ *for which the urgleichung yields an invalid probability.*

It is sometimes helpful to write the urgleichung in vector notation:

$$q = rMp. \tag{31}$$

Here, r is a matrix whose (j, i) entry is given by $r(j|i)$, and M is a linear combination of the identity matrix I and the matrix whose elements all equal 1, the so-called Hadamard identity J:

$$M = \alpha I - \beta J. \tag{32}$$

Assumptions 1 and 2 imply a fair bit about the structure of $\mathcal{P}$ and $\mathcal{R}$.

Lemma 1. *The set $\mathcal{P}$ of all states and the set $\mathcal{R}$ of all measurements are both convex and closed.*

Proof. Let $p_1, p_2 \in \mathcal{P}$, and for any $r \in \mathcal{R}$, define

$$q_1 = rMp_1, \qquad q_2 = rMp_2. \tag{33}$$

By assumption, both q_1 and q_2 are valid probability vectors (i.e., they are normalized, and all their entries are nonnegative). Define

$$p_\lambda = \lambda p_1 + (1 - \lambda)p_2. \tag{34}$$

Then

$$q_\lambda = rMp_\lambda = \lambda q_1 + (1 - \lambda)q_2. \tag{35}$$

This is a convex combination of points in the probability simplex, and as such it also belongs to the probability simplex. By assumption, this holds true for every $r \in \mathcal{R}$, and so by maximality, $p_\lambda \in \mathcal{P}$. The proofs of the convexity of $\mathcal{R}$ and of closure work analogously. □

Consider the case where the ground and sky measurements are the same. In that scenario, we have $q = p$, and so the measurement matrix must be the inverse of M:

$$r_F = M^{-1} = \frac{1}{\alpha} I + \frac{\beta}{\alpha} J. \tag{36}$$

Note that we have to include r_F within $\mathcal{R}$ by the maximality assumption.

The urgleichung is one way that quantum theory builds upon the mathematics of probability, interconnecting our previsions for different experiments, previsions that basic probability theory alone would leave separate. Quantum theory augments the probability formalism in another fashion as well, and it is to that which we now turn.

Our next assumption will establish that the set of measurements $\mathcal{R}$ can be constructed from the set of states $\mathcal{P}$. On a purely mathematical level, we could justify this by saying that we wish to build the most parsimonious theory possible upon the urgleichung, and so we simplify matters by having one fundamental set instead of two. As far as constructing a mathematical theory goes, this is certainly a legitimate way to begin. We can, however, provide a more physical motivation than that.

Probability theory, intrinsically, assumes very little about the structure of event spaces. With it, we can for example discuss rolling a die and recording the side that lands facing up; we say that the realm of possible outcomes for this experiment is the set $\{1, 2, 3, 4, 5, 6\}$. In this experiment, the outcome "1" is no more *like* the outcome "2" than it is *like* the outcome "6". We can ascribe probabilities to these six potential events without imposing a similarity metric upon the realm of outcomes. We use integers as labels, but we care hardly at all about the number-theoretic properties of those integers. When we roll the die, we are indifferent to the fact that 5 is prime and 6 is perfect. Nor is the event of observing a particular integer in this experiment related, necessarily, to the event of observing that same integer in a *different* experiment.

When Alice first learns probability theory, she picks up this habit of tagging events with integers. If Alice considers a long catalogue of experiments that she could perform, she might label the possible outcomes of the first experiment by the integers from 1 to N_1, the outcomes of the second experiment by the integers $\{1, \ldots, N_2\}$ and so on. But, in general, Alice has the freedom to permute these labels as she pleases. She does not have to regard the experience of obtaining $j = 17$ in one experiment as similar to the experience of obtaining $j = 17$ in any other.

But what if Alice wants more structure than this? When Alice contemplates an experiment that she might carry out, she considers a set of possible outcomes for it, i.e., a realm of potential experiences which that action might elicit. She can assign each of those potential experiences a label drawn from whatever mathematical population she desires. Her *index set* for a given experiment can be a subset of whatever population she finds convenient. When Alice adopts the urgleichung as an empirically-motivated addition to the bare fundamentals of probability theory, does she, by that act, also gain a natural collection of mathematical entities from which to build index sets?

In fact, she has just such a collection at hand: She can use the set of valid states, $\mathcal{P}$!

To consider the matter more deeply, we ask the following question: under what conditions would Alice consider two outcomes of two different experiments to be equivalent? For example, Alice contemplates two experiments she might feasibly perform, which she describes by two matrices r and r'. When would Alice treat an outcome j of experiment r to be equivalent to an outcome j' of r' [4]? Generally, the tools she has on hand to make such a judgment are her probability ascriptions for those outcomes. If her overall mesh of beliefs is that her probability of experiencing j upon enacting r is the same as her probability for finding j' when enacting r', no matter what her state assignment p, then she has good grounds to call j and j' equivalent. In order to satisfy $q(j) = q'(j')$ for all $p \in \mathcal{P}$, the measurement matrices r and r' must obey

$$r(j|i) = r'(j'|i), \ \forall i. \tag{37}$$

The simplest way to ensure that this is possible is to build all elements r of the set $\mathcal{R}$ from a common vocabulary. When we construct an element $r \in \mathcal{R}$, we draw each row from a shared pool of ingredients. The natural, parsimonious choice we have on hand for this purpose is the set $\mathcal{P}$.

This means that, up to scaling, measurement outcomes are actually identified with points in the probability simplex.

Let $r \in \mathcal{R}$ be a valid measurement. If each row of the matrix $\{r(j|i)\}$ can also naturally be identified with a vector $s \in \mathcal{P}$, then we are led to consider the vector s sitting inside r in some fashion. The simplest reasonable relation between s, which is a vector with N elements, and the measurement matrix r, whose rows have length N, is to have a row of r be linearly proportional to s.

Assumption 3. *Measurement Matrices are Constructed from States. Given any $r \in \mathcal{R}$, we can write a row $\{r(j|i) \ : \ i = 1, \ldots, N\}$ as a vector $s_j \in \mathcal{P}$, up to a normalization factor:*

$$r(j|i) = N\gamma_j s_j(i). \tag{38}$$

Furthermore, any state in $\mathcal{P}$ can be used in this manner.

For brevity, we will refer to the s_j as "measurement vectors." We will shortly identify the meaning of the constants $\{\gamma_j\}$, which we have written with the prefactor N for later convenience.

Assumption 4. *Possibility of Maximal Ignorance. The state c, defined by*

$$c(i) = \frac{1}{N} \ \forall \ i, \tag{39}$$

belongs to $\mathcal{P}$.

This can be deduced from other postulates, but the state c is a useful tool, and it is helpful to point its existence out explicitly. For example, substituting the state of complete ignorance c into the urgleichung, we obtain

$$q(j) = \frac{1}{N} \sum_i r(j|i). \tag{40}$$

What is the meaning of the factors $\{\gamma_j\}$? To find out, we apply a measurement $r \in \mathcal{R}$ to the state c:

$$q(j) = \frac{1}{N} \sum_i r(j|i) = \gamma_j \sum_i s_j(i) = \gamma_j. \tag{41}$$

The factors $\{\gamma_j\}$ indicate the probability of obtaining the jth outcome on the ground when the agent is completely indifferent to the potential outcomes of the sky experiment.

If the effect of some $r \in \mathcal{R}$, when applied via the urgleichung, is to send c to itself, then we have that

$$c(j) = \frac{1}{N} = \frac{1}{N} \sum_i r(j|i) \Rightarrow \sum_i r(j|i) = 1. \tag{42}$$

Combined with the basic normalization requirement for conditional probabilities, this states that a measurement that preserves c is represented by a *doubly stochastic* matrix.

Lemma 2. *Measurements that send the state c to itself are represented by doubly stochastic matrices.*

When we postulated the urgleichung, we added structure to the bare essentials of probability theory, and the structure we added related one experiment to another in a way above and beyond basic coherence. With Assumption 3, we are also interrelating different experiments. We can appreciate this in another way by considering what it means for a physical system to be usable as a scientific instrument.

What conditions must an object meet in order to qualify as a piece of laboratory apparatus? Classically, a bare minimum requirement is that the object has a set of distinguishable configurations in which it can exist. These might be positions of a pointer needle, heights of a mercury column, patterns of glowing lights and so forth. The essential point is that the system can be in different configurations at different times: a thermometer that always reports the same temperature is useless. We can label these distinguishable configurations by an index j. The *calibration* process for a laboratory instrument is a procedure by which a scientist assigns conditional probabilities $r(j|i)$ to the instrument, relating the readout states j to the inputs i. In order to make progress, we habitually assume that nature is not so perverse that the results of the calibration phase become completely irrelevant when we proceed to the next step and apply the instrument to new systems of unknown character.

But what if nature *is* perverse? Not enough so to forbid the possibility of science, but enough to make life interesting. Quantitatively speaking, what if we must modify the everyday assumption that one can carry the results of a calibration process unchanged from one experimental context to another?

The urgleichung is just such a modification. The $\{r(j|i)\}$ do not become irrelevant when we move from the sky context to the ground, but we do have to use them in a different way.

In quantum physics, we no longer treat "measurement" as a passive reading-off of a specified, pre-existing physical quantity. However, we do still have a counterpart for our classical notion of a system that can qualify as a laboratory apparatus. Instead of asking whether the system can exist in one of multiple possible classical states, we ask whether our overall mesh of beliefs allows us to consistently assign any one of multiple possible catalogues of expectations. That is, if an agent Alice wishes to use a system as a laboratory apparatus, she must be able to say now that she can conceive of ascribing any one of several states to it at a later time. We define a *discrete apparatus* as a physical system with an associated set of states,

$$\{s_1, \ldots, s_m\} \subset \mathcal{P}. \tag{43}$$

The analogue of classical uncertainty about where a pointer might be pointing is the convex combination of the states $\{s_j\}$. Therefore, our basic mental model of a laboratory apparatus is a polytope in $\mathcal{P}$, with the $\{s_j\}$ as its vertices. Assumption 3 says that *Alice can pick up any such apparatus and use it as a "prosthetic hand" to enrich her experience of asking questions of nature.*

We can think of Assumption 3 in another way, if we rewrite equation (38) in the following manner:

$$s_j(i) = \frac{\left(\frac{1}{N}\right) r(j|i)}{\gamma_j}. \tag{44}$$

Earlier, we noted that γ_j is the probability of obtaining the jth outcome on the ground, given complete ignorance about the potential outcomes of the sky experiment. In addition, $1/N$ is the probability assigned to each outcome of the sky experiment by the state of complete ignorance. So,

$$s_j(i) = \frac{\mathrm{PrCI}(i)\, r(j|i)}{\mathrm{PrCI}(j)}, \tag{45}$$

where the notation "PrCI" here indicates a probability assignment given that the state for the sky experiment is c. Note that $\mathrm{PrCI}(j|i) = r(j|i)$. But this means that the expression on the right-hand side above is just the ordinary Bayes formula for inverting conditional probabilities:

$$\mathrm{PrCI}(i|j) = \frac{\mathrm{PrCI}(i)\,\mathrm{PrCI}(j|i)}{\mathrm{PrCI}(j)}. \tag{46}$$

Therefore, we can interpret the mathematical relation established in Assumption 3 as saying that "posteriors from maximal ignorance are priors" [1]. For the remainder of this paper, we will not be considering in detail the rules for changing one's probabilities upon new experiences – a rather intricate subject, all things told [30,31]. So, we will not stress the ideas of "priors" and "posteriors," but it is good to know that this reading of Assumption 3 exists.

Writing the urgleichung in terms of the vector s_j,

$$q(j) = \sum_i \left[\alpha p(i) - \beta\right] N \gamma_j s_j(i) \tag{47}$$

$$= N\alpha\gamma_j \langle p, s_j\rangle - N\beta\gamma_j. \tag{48}$$

The fact that $q(j)$ must be nonnegative for all j implies a lower bound on the scalar product $\langle p, s_j\rangle$:

$$\langle p, s_j\rangle \geq \frac{\beta}{\alpha}. \tag{49}$$

The measurement described by the matrix r_F in equation (36) yields, by construction, equal probabilities for all outcomes given the input state c. That is, it is an experiment with N outcomes, and $\gamma_j = 1/N$ for all of them. Therefore, we can take the rows of r_F as specifying N special vectors within $\mathcal{P}$. We have that

$$r_F(j|i) = e_j(i), \tag{50}$$

where the vector e_j is flat across all but one entries:

$$e_j(i) = \frac{1}{\alpha}(\delta_{ji} + \beta). \tag{51}$$

We will refer to the vectors $\{e_k\}$ as the *basis distributions.*

What happens if we take a measurement $r \in \mathcal{R}$, and act with it via the urgleichung upon a basis distribution e_k? The result is straightforwardly computed to be

$$q(j) = \sum_i \left[\alpha\left(\frac{\beta}{\alpha} + \frac{1}{\alpha}\delta_{ik}\right) - \beta\right] r(j|i) \tag{52}$$

$$= \beta\sum_i r(j|i) + \sum_i \delta_{ik} r(j|i) - \beta\sum_i r(j|i) \tag{53}$$

$$= r(j|k). \tag{54}$$

This will be useful later.

Note that the basis distributions all have magnitude equal to

$$\langle e_k, e_k\rangle = \frac{1 + 2\beta + N\beta^2}{\alpha^2}. \tag{55}$$

This result singles out a *distinguished length scale* in probability space, namely, the radius of the sphere on which all the basis distributions live.

The lower bound (49) suggests the following construction. Let H be the hyperplane of vectors in $\mathbb{R}^N$ that sum to unity:

$$H = \left\{v \in \mathbb{R}^N \colon \langle v, c\rangle = \frac{1}{N}\right\}. \tag{56}$$

This hyperplane includes the probability simplex. For any set A of probability distributions, consider the set

$$A^* = \left\{u \in H \colon \langle u, v\rangle \geq \frac{\beta}{\alpha}\ \forall v \in A\right\}. \tag{57}$$

This set includes all the probability distributions that are consistent with each point in A, with respect to the lower bound we derived from the urgleichung. We will designate the set A^* the *polar* of A, following the terminology for a related concept in geometry [19,20]. Let $\mathcal{P}$ be the set of all valid states. The set of all measurement vectors that are consistent with these states, with respect to the lower bound, is that portion of the polar of $\mathcal{P}$ that lies within the probability simplex:

$$\mathcal{P}^* \cap \Delta = \left\{s : \langle s, p\rangle \geq \frac{\beta}{\alpha} \forall p \in \mathcal{P}\right\} \cap \Delta. \tag{58}$$

If some s in this set is not in the set $\mathcal{P}$, then some measurement vector does not correspond to a state. Likewise, if some $p \in \mathcal{P}$ is not in this set, then that state cannot correspond to a measurement vector. Both of these cases violate the mapping we have advocated on general conceptual grounds. Therefore, our first three assumptions imply that we consider sets $\mathcal{P}$ for which

$$\mathcal{P} = \mathcal{P}^* \cap \Delta. \tag{59}$$

We will see momentarily how to simplify this condition, establishing the condition that a state space $\mathcal{P}$ must be self-polar:

$$\mathcal{P} = \mathcal{P}^*. \tag{60}$$

In order to prove this proposition, we need to know more about the operation of taking the polar. We can derive the

relations we require by adapting some results from the higher-dimensional geometry literature. Grünbaum [19] defines the polar of $A \subseteq \mathbb{R}^{d^2}$ to be the set

$$A^\circ = \{u \in \mathbb{R}^{d^2} : \langle u, v\rangle \leq 1 \ \forall v \in A\}. \tag{61}$$

Our definition of the polar A^* is close enough to this definition of A° that many results about the latter can be carried over with little effort. The properties of the polar A^* are summarized in the following theorem.

Theorem 3. *For all $A \subseteq H$, the polar A^* is a closed, convex set containing c. Since we will frequently be invoking the concept of convex hulls, we introduce the notation $\mathrm{cc}(A)$ for the closed, convex hull of the set A. We have*

$$A^* = \big(\mathrm{cc}(A \cup \{c\})\big)^*, \tag{62}$$
$$A^{**} = \mathrm{cc}(A \cup \{c\}), \tag{63}$$

for all $A \subseteq H$. In particular, A is equal to its double polar A^{**} if and only if it is closed, convex and contains c.

For all A, $B \subseteq H$

$$A \subseteq B \implies B^* \subseteq A^*. \tag{64}$$

If $\mathcal{A}$ is an arbitrary family of subsets of H then

$$\left(\bigcup_{A\in\mathcal{A}} A\right)^* = \bigcap_{A\in\mathcal{A}} A^*. \tag{65}$$

If, in addition, $A^{**} = A$ for all $A \in \mathcal{A}$ then

$$\left(\bigcap_{A\in\mathcal{A}} A\right)^* = \mathrm{cc}\left(\bigcup_{A\in\mathcal{A}} A^*\right). \tag{66}$$

Proof. All these properties follow by relating Grünbaum's definition of the polar with ours. Let $f : \mathbb{R}^N \to \mathbb{R}^N$ be the affine map defined by

$$f(u) = N\alpha(u - c), \tag{67}$$

and let H_0 be the subspace

$$H_0 = \{u \in \mathbb{R}^N : \langle u, c\rangle = 0\}. \tag{68}$$

One then has

$$A^* = f^{-1}\Big(\big(-f(A)\big)^\circ \cap H_0\Big) \tag{69}$$

for all $A \subseteq H$. With this in hand the theorem becomes a straightforward consequence of textbook results. □

Now, consider the relation $\mathcal{P} = \mathcal{P}^* \cap \Delta$, and take the polar of both sides:

$$\mathcal{P}^* = (\mathcal{P}^* \cap \Delta)^* = \mathrm{cc}\left(\mathcal{P}^{**} \cup \Delta_N^*\right). \tag{70}$$

We know that $\mathcal{P}$ is closed and convex, and that it contains the center point c. Therefore,

$$\mathcal{P}^{**} = \mathcal{P}. \tag{71}$$

What is the polar of the probability simplex Δ? In fact, it is the basis simplex Δ_e.

Lemma 4. *The probability simplex and the basis simplex are mutually polar:*

$$\Delta^* = \Delta_\mathrm{e}, \qquad \Delta_\mathrm{e}^* = \Delta. \tag{72}$$

Proof. The probability simplex contains normalized vectors, so it lies in the hyperplane H, and all of its vectors have wholly nonnegative entries. Let v_i be the ith vertex of Δ (so $v_i(j) = \delta_{ij}$). Then the probability simplex is

$$\Delta = \{u \in H : \langle u, v_i\rangle \geq 0 \ \forall i\}. \tag{73}$$

Let $f : H \to H$ be the affine map defined by

$$f(u) = \frac{1}{\alpha}u + \frac{\beta}{\alpha}. \tag{74}$$

Then $\Delta_\mathrm{e} = f(\Delta)$. It follows that

$$\Delta_\mathrm{e} = \left\{u \in H : \langle u, v_i\rangle \geq \frac{\beta}{\alpha} \ \forall i\right\}. \tag{75}$$

Taking account of Theorem 3 we deduce

$$\Delta_\mathrm{e} = \{v_i : i = 1, \ldots, N\}^* = \Delta^*. \tag{76}$$

The fact that $\Delta_\mathrm{e}^* = \Delta$ is an immediate consequence of this and the fact that the double polar of a closed convex set is itself (see Thm. 3). □

Theorem 5. *A state space $\mathcal{P}$ satisfying Assumptions 1–4 is self-polar:*

$$\mathcal{P} = \mathcal{P}^*. \tag{77}$$

Proof. We already know that

$$\mathcal{P}^* = \mathrm{cc}\left(\mathcal{P}^{**} \cup \Delta^*\right), \tag{78}$$

and now we can say that

$$\mathcal{P}^* = \mathrm{cc}(\mathcal{P} \cup \Delta_\mathrm{e}). \tag{79}$$

But we established already that $\mathcal{P}$ always contains the basis distributions, and that $\mathcal{P}$ is closed and convex. Therefore, $\mathcal{P}$ is self-polar. □

The fact that a state space is self-polar implies the existence of two more distinguished length scales. To see why, it is helpful to work in barycentric coordinates, shifting all our vectors so that the origin lies at the barycenter point of the simplex, the point c:

$$p \to p' = p - c. \tag{80}$$

In these coordinates, our lower bound (49) becomes

$$\langle p', s'\rangle \geq -\frac{1}{N\alpha}. \tag{81}$$

Any basis distribution e_j satisfies

$$\langle e_j', e_j'\rangle = \frac{N-1}{N\alpha^2}. \tag{82}$$

We define the *out-sphere* S_o to be the sphere centered on the barycenter with radius

$$r_{\mathrm{o}}^2 = \frac{N-1}{N\alpha^2}. \tag{83}$$

The ball bounded by S_o is the *out-ball* B_{o}. We will see shortly that the polar of the out-ball is a ball centered at the barycenter and having radius

$$r_i^2 = \frac{1}{N(N-1)}. \tag{84}$$

We designate this ball the *in-ball* B_i, and its surface is the *in-sphere* S_i. Finally, note that if we take

$$r_m^2 = \frac{1}{N\alpha}, \tag{85}$$

any two points both lying within r_m of the barycenter will be consistent with respect to the bound (49). This defines the *mid-ball* B_m and its surface, the *mid-sphere* S_m. It follows that

$$r_i r_{\mathrm{o}} = r_m^2. \tag{86}$$

We now prove the fact we stated a moment ago.

Lemma 6. *The out- and in-balls are mutually polar:*

$$B_{\mathrm{o}}^* = B_i, \qquad B_i^* = B_{\mathrm{o}}. \tag{87}$$

Proof. Let $f\colon H \to H$ be the affine map defined by

$$f(u) = c + \frac{r_{\mathrm{o}}}{r_i}(u-c). \tag{88}$$

Then $f(B_i) = B_{\mathrm{o}}$. Consequently, given arbitrary $u \in H$,

$$u \in B_{\mathrm{o}}^* \tag{89}$$

$$\iff \quad \langle u, f(v)\rangle \geq \frac{\beta}{\alpha} \qquad \forall v \in B_i \tag{90}$$

$$\iff \quad \langle u-c, f(v)-c\rangle \geq -r_{\mathrm{o}} r_i \qquad \forall v \in B_i \tag{91}$$

$$\iff \quad \langle u-c, v-c\rangle \geq -r_i^2 \qquad \forall v \in B_i \tag{92}$$

$$\iff \quad u \in B_i \tag{93}$$

So $B_{\mathrm{o}}^* = B_i$. The fact that $B_i^* = B_{\mathrm{o}}$ is an immediate consequence of this and the fact that the double polar of a closed convex set is itself. □

These distinguished length scales suggest another assumption we ought to make about our state space. Earlier, we stated that the barycenter c must belong to our set of admissible probability distributions. It is natural to ask how far away from complete ignorance we can go before we encounter complications. Can our state space $\mathcal{P}$ contain all the points in a little ball around c? Intuitively, it is hard to see why not. How big can we make that ball around the center point c before we run into trouble? The simplest assumption, in this context, is to postulate that the first complication we encounter is the edge of the probability simplex itself. Where does a sphere centered at c touch the faces of the simplex? The center of a face of the probability simplex is found by taking the average of $N-1$ of its vertices:

$$\bar{v}_k(i) = \frac{1}{N-1}(1-\delta_{ik}). \tag{94}$$

The sphere centered on c that just touches these points has a radius given by

$$(\bar{v}_k - c)^2 = \frac{1}{N(N-1)}. \tag{95}$$

The in-sphere S_i is just the *inscribed* sphere of the probability simplex.

Assumption 5. *Every state space $\mathcal{P}$ contains the in-ball.*

Because the polar of the in-ball is the out-ball, and polarity reverses inclusion, it follows that every self-consistent state space is bounded by the out-sphere. This result has the form of an "uncertainty principle": it means that our probability distributions can never become too narrowly focused. For any two points p and s within our state space $\mathcal{P}$, we have

$$L \leq \langle p, s\rangle \leq U, \tag{96}$$

where the lower and upper bounds are given by

$$L = -\frac{1}{N\alpha} + \frac{1}{N}, \tag{97}$$

$$U = \frac{N-1}{N\alpha^2} + \frac{1}{N}. \tag{98}$$

Recall from Lemma 4 that the polar of the probability simplex is the simplex defined by the basis distributions e_k, which in barycentric coordinates is seen to be the probability simplex rescaled:

$$e_k'(i) = e_k(i) - c(i) = \frac{1}{\alpha}\left(\delta_{ik} - c(i)\right). \tag{99}$$

Call two extremal states p and s in a state space *maximally distant* if they saturate the lower bound:

$$\langle p', s'\rangle = -\frac{1}{N\alpha}. \tag{100}$$

Let

$$\{p_k' : k = 1, \ldots, m\}, \tag{101}$$

be a set of Mutually Maximally Distant (MMD) states. That is, for all k,

$$\langle p_k', p_k'\rangle = r_{\mathrm{o}}^2, \tag{102}$$

and for $k \neq l$,

$$\langle p_k', p_l'\rangle = -r_m^2. \tag{103}$$

Construct the vector quantity

$$V = \sum_k p_k'. \tag{104}$$

From the fact that the magnitude $\langle V, V\rangle \geq 0$, it follows that

$$m \leq 1 + \frac{r_{\mathrm{o}}^2}{r_m^2}. \tag{105}$$

Substituting in the definitions of the radii, we arrive at the relation

$$m \leq 1 + \frac{N-1}{\alpha}. \tag{106}$$

Let us now make an assumption: we want this bound to be attainable.

Assumption 6. *A state space $\mathcal{P}$ contains an MMD set of size*

$$m_{\max} = 1 + \frac{N-1}{\alpha}. \tag{107}$$

Note that both N and $m_{\max}$ are positive integers by assumption. This means that α must divide $N-1$ neatly.

To set the context for our next assumption, switch back to the original frame. Recall that any two points p and s within our state space $\mathcal{P}$ satisfy

$$L \leq \langle p, s \rangle \leq U, \tag{108}$$

where the lower and upper bounds are given by

$$L = -\frac{1}{N\alpha} + \frac{1}{N}, \tag{109}$$

$$U = \frac{N-1}{N\alpha^2} + \frac{1}{N}. \tag{110}$$

Comparing these two quantities, and using equation (107) to simplify, we obtain

$$\frac{U}{L} = 1 + \frac{m_{\max}}{\alpha - 1}, \tag{111}$$

where $m_{\max}$ is a positive integer. This expression makes it inviting to set the ratio on the right-hand side to unity by fixing

$$m_{\max} = \alpha - 1, \tag{112}$$

and thus $U/L = 2$ is, in a sense, the natural first option to explore.

Assumption 7. *The upper and lower bounds in the fundamental inequalities are related by*

$$U = 2L. \tag{113}$$

This lets us solve for N in terms of α:

$$N = (\alpha - 1)^2. \tag{114}$$

Thanks to our two latest assumptions, we can fix all three parameters in the generalized urgleichung (27) in terms of the maximal size of an MMD set:

$$N = m_{\max}^2, \;\; \alpha = m_{\max} + 1, \;\; \beta = \frac{1}{m_{\max}}. \tag{115}$$

Relabeling $m_{\max}$ by d for brevity, we recover the formulas familiar from the SIC representation of quantum state space. Here, the generalized urgleichung takes the specific form

$$q(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i), \tag{116}$$

and we arrive at the following pair of inequalities:

$$\frac{1}{d(d+1)} \leq \langle p, s \rangle \leq \frac{2}{d(d+1)}. \tag{117}$$

Consequently, the polar of a set A is

$$A^* = \left\{ u \in H \colon \langle u, v \rangle \geq \frac{1}{d(d+1)} \;\forall v \in A \right\}. \tag{118}$$

We now arrive at the definition upon which the rest of our theory will stand.

Definition 1. A *qplex* is a self-polar subset of the out-ball in the probability simplex Δ_{d^2}, with the parameters in the generalized urgleichung set to $\alpha = (d+1)$ and $\beta = 1/d$.

3 Fundamental geometry of qplexes

In the previous section, we began with the urgleichung and, making a few assumptions of an operational character, arrived at the double inequality

$$\frac{1}{d(d+1)} \leq \langle p, s \rangle \leq \frac{2}{d(d+1)}. \tag{119}$$

Here, we will take this as established, and we will demonstrate several important geometrical properties of the sets that maximally satisfy it – the qplexes.

A qplex is a subset of Δ, the probability simplex in $\mathbb{R}^{d^2}$ (i.e. the space of probability distributions with d^2 outcomes). Δ is, in turn, a subset of the hyperplane

$$H = \left\{ u \in \mathbb{R}^{d^2} \colon \langle u, c \rangle = \frac{1}{d^2} \right\}, \tag{120}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product on $\mathbb{R}^{d^2}$ and

$$c = \left(\tfrac{1}{d^2} \;\ldots\; \tfrac{1}{d^2}\right)^{\mathrm{T}} \tag{121}$$

is the barycenter of Δ.

It is important to appreciate the geometrical relationships between the four sets $\Delta, \Delta_{\mathrm{e}}, B_{\mathrm{o}}, B_i$. Specializing our results from the previous section, we have

$$e_i - c = \frac{1}{d+1}(v_i - c), \tag{122}$$

$$r_i = \frac{1}{d-1} r_{\mathrm{o}}. \tag{123}$$

So the basis simplex is obtained from the probability simplex by scaling by a factor $1/(d+1)$, while the in-ball is obtained from the out-ball by scaling by a factor $1/(d-1)$. In particular $B_i = B_{\mathrm{o}}$ when $d = 2$, but is otherwise strictly smaller. We have

$$\langle e_j, e_k \rangle = \frac{d\delta_{jk} + d + 2}{d(d+1)^2}. \tag{124}$$

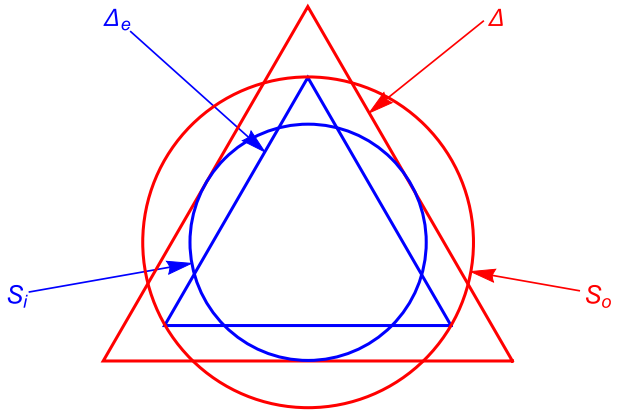

Fig. 2. Relations between the sets Δ, Δ_e, S_o, S_i when $d > 2$. The diagram is schematic only. It shows the inclusion relations, and points of contact, but does not reproduce the metric relations, which are impossible to depict accurately in a 2-dimensional diagram. In particular, the basis simplex Δ_e is much smaller in relation to the probability simplex Δ than is shown here. This is also true of the in-sphere S_i and the out-sphere S_o.

If $d = 2$ then

$$\Delta_e \subseteq B_i = B_o \subseteq \Delta. \tag{125}$$

If $d > 2$ then one still has

$$\Delta_e \cup B_i \subseteq \Delta \cap B_o, \tag{126}$$

but

$$\Delta_e \nsubseteq B_i, \qquad B_o \nsubseteq \Delta. \tag{127}$$

The first of these statements is an immediate consequence of the foregoing. To prove the second observe that $e_i \in \Delta_e$ but $\notin B_i$, while $c + (r_o/r_i)(\bar{e}_i - c) \in B_o$ but $\notin \Delta$.

These facts are perhaps most easily appreciated by examining the diagram in Figure 2. Observe, however, that the metric relations are impossible to reproduce in a 2-dimensional diagram. So, although Figure 2 reproduces the inclusion relations, and points of contact, it badly misrepresents the sizes of the sets Δ_e, B_i in comparison to the sets Δ, B_o.

General properties of qplexes include the following:

- Any qplex is convex and closed, and is thus the convex hull of its extremal points.
- Because a qplex is self-polar, it can be thought of as the intersection of half-spaces. Each half-space is defined, per equation (118), by a hyperplane that is composed of points all maximally distant from an extreme point of the qplex.
- For every extreme point of a qplex, there exists at least one point that is maximally distant to it, in the sense of saturating the lower bound in equation (119).
- Call a vector $p \in Q$ a *pure* vector if $\langle p, p\rangle = 2/(d(d+1))$. Any set of pure vectors that pairwise saturate the lower bound of the consistency condition (119) contains no more than d elements.
- Suppose we have a qplex Q that is a polytope, i.e., the convex hull of a finite set of vertices. Because all qplexes contain the basis distributions, this polytope must have at least d^2 vertices. The polar of each extreme point is a half-space bounded by a hyperplane, all of the points on which are maximally distant from that extreme point. The intersection of the half-spaces defined by all these hyperplanes forms a polytope. By self-polarity, this polytope is identical to Q. It follows that each extreme point of Q must lie on at least $d^2 - 1$ such hyperplanes. Therefore, each vertex of Q is maximally distant from at least $d^2 - 1$ other extreme points.
- It follows from the above that a qplex cannot be a simplex. Consequently, any point in the interior of a qplex can be written in more than one way as a convex combination of points on the boundary. This is a generalization of the result that any mixed quantum state has multiple convex decompositions into different sets of pure states, a theorem that has historically been of some significance in interpreting the quantum formalism [32–35]. Also, a result of Plávala implies that any qplex admits incompatible measurements [36].
- If Q is a qplex, then no vector $p \in Q$ can have an element whose value exceeds $1/d$.
- The total number of zero-valued entries in any vector belonging to a qplex is bounded above by $d(d-1)/2$.

A SIC representation of a quantum state space is a qplex with a continuous set of pure points. All qplexes with this property enjoy an interesting geometrical relation with the polytopes that can be inscribed within them.

Theorem 7. *If Q is a qplex that contains an infinite number of pure points, then any polytope inscribed in Q cannot contain the in-sphere S_i.*

Proof. Suppose that P is a polytope inscribed in Q that contains the in-sphere S_i. Recall that the polarity operation reverses inclusion (Thm. 3), so the polar polytope P^* of P must contain the polar Q^* of Q. But all qplexes are self-polar, so $Q \subset P^*$. Likewise, because the polar of the in-ball B_i is the out-ball B_o, it follows that P^* is contained within the out-sphere S_o. Consequently, Q can have only a finite number of pure points. □

Let us consider the two-outcome measurement r_s defined by rescaling a state $s \in Q$:

$$r_s(0|i) = d^2\gamma_0 s(i), \;\; \gamma_0 = \frac{1}{d}. \tag{128}$$

We fix the other row of the matrix $r_s(j|i)$ by normalization:

$$r(0|i) + r(1|i) = 1. \tag{129}$$

Does this actually define a legitimate measurement? Because $\langle p, s\rangle$ is always bounded above and below for any vector $p \in \mathcal{P}$, then applying r_s to any $p \in \mathcal{P}$ via the urgleichung will yield a valid probability vector q. Therefore, r_s defined in this way is indeed a member of $\mathcal{R}$.

What's more, if we apply r_s to the state s itself, then we can be *certain* about the outcome, if s lies on the same sphere as the basis distributions. In such a case, we have $q(0) = 1$. If Alice ascribes a state having this magnitude to a system, she is asserting her confidence that performing a particular experiment will have a specific result. But certainty about *one* experiment does not, and indeed cannot, imply certainty about *all*. Even when Alice is certain about what would happen should she perform the experiment r_s, she is necessarily uncertain about what would happen if she brought the Bureau of Standards measurement down to the ground and applied it.

Note that when we apply r_s to a state p, we compute

$$q(0) = d(d+1)\langle p, s\rangle - 1. \tag{130}$$

The bound established by Assumption 5 implies that we can associate the factor d just as well with s or with p. That is, both r_s and r_p are valid measurements within $\mathcal{R}$, and we obtain the same probability $q(0)$ when we apply r_s to p as we would if we applied r_p to the state s.

This is a point worth considering in depth. With Assumption 3, we introduced a relation between the set of all states and the set of all measurements. Now, thanks to the additional assumptions we have invoked since then, we have a more specific correspondence between the two sets: for every pure state, there is a binary measurement for which that state, and no other state, implies certainty. This result depends upon our assumption that departures from complete ignorance are minimally constrained, or equivalently, that the basis distributions are extremal. As a consequence, we know that we can take any valid state s and scale by a factor d to create a row in a measurement matrix. In the language of Asher Peres, the fact that we can interpret equation (130) as r_s applied to p or as r_p applied to s, for any states p and s, is the reciprocity of "preparations" and "tests" [37].

This reciprocity is an important concept for many mathematical treatments of quantum physics. For example, it is one of the primary axioms in Haag's formulation [38,39]. To those who apply category theory to quantum mechanics, it is the reason why they construct "dagger-categories," and how the basic idea of an inner product is introduced into their diagrammatic language [24].

Next, we consider sets which are related to qplexes.

Definition 2. A subset A of the probability simplex Δ is a *germ* if it satisfies the fundamental inequalities (117) for all p, $s \in A$.

Definition 3. A germ is *maximal* if no point can be added to it without violating the fundamental inequalities (117).

We start by proving two results about germs that follow from the Cauchy-Schwarz inequality. Originally, these theorems were proved for qplexes [1,2,15], but they apply more broadly.

Theorem 8. *If G is a germ, then no vector $p \in G$ can have an element whose value exceeds $1/d$.*

Proof. Let $p \in G$ be a point on the out-sphere. Assume without loss of generality that $p(0) \geq p(i)$. Then

$$\frac{2}{d(d+1)} = p(0)^2 + \sum_{i=1}^{d^2-1} p(i)^2, \tag{131}$$

and using the Cauchy-Schwarz inequality,

$$\frac{2}{d(d+1)} \geq p(0)^2 + \frac{1}{d^2-1}\left(\sum_{i=1}^{d^2-1} p(i)\right)^2. \tag{132}$$

By normalization, we can simplify the sum in the last term, yielding

$$\frac{2}{d(d+1)} \geq p(0)^2 + \frac{1}{d^2-1}\left(1 - p(0)\right)^2. \tag{133}$$

Thus,

$$p(0) \leq \frac{1}{d}, \tag{134}$$

with equality if and only if all the other $p(i)$ are equal, in which case, normalization forces them to take the value $1/(d(d+1))$. □

Remark 1. If the germ G contains the basis distributions, this result also follows from

$$\langle p, e_k\rangle = \frac{1}{d(d+1)} + \frac{p_k}{d+1} \leq \frac{2}{d(d+1)}. \tag{135}$$

Theorem 9. *The total number of zero-valued entries in any vector belonging to a germ is bounded above by $d(d-1)/2$.*

Proof. Let G be a germ and choose $p \in G$. Square the basic normalization condition to find

$$\left(\sum_i p(i)\right)^2 = 1. \tag{136}$$

Apply the Cauchy-Schwarz inequality to show, writing n_0 for the number of zero-valued elements in p,

$$(d^2 - n_0)\sum_{\{i:p(i)>0\}} p(i)^2 \geq \left(\sum_{\{i:p(i)>0\}} p(i)\right)^2 = 1. \tag{137}$$

Consequently,

$$n_0 \leq d^2 - \frac{d(d+1)}{2} = \frac{d(d-1)}{2}. \tag{138}$$

□

It follows from Zorn's lemma [40] that every germ is contained in at least one maximal germ. In other words, we can extend any germ in at least one way to form a set that is also a germ, but which admits no further consistent extension. Adding any new point to a maximal germ

implies that some pair of points will violate the inequalities (117). Every qplex is a germ, but the converse is not true. Using the theory of polarity, we will show that any maximal germ is a self-polar subset of the out-ball. That is, a maximal germ is a qplex, and in fact, any qplex is also a maximal germ.

It is an immediate consequence of the definition that if G is an arbitrary germ then

$$G \subseteq \Delta \cap B_{\mathrm{o}}, \tag{139}$$

where B_{o} is the out-ball:

$$B_{\mathrm{o}} = \left\{ u \in H \colon \langle u, u \rangle \leq \frac{2}{d(d+1)} \right\} \tag{140}$$

$$= \{ u \in H \colon \|u - c\| \leq r_{\mathrm{o}} \}. \tag{141}$$

Taking polars on both sides of equation (139) and taking account of what polarity does to inclusion and intersection (Thm. 3), we find

$$\mathrm{cc}(\Delta^* \cup B_{\mathrm{o}}^*) \subseteq G^*, \tag{142}$$

for every germ G. Recall from Lemma 4 that the polar of Δ is the basis simplex Δ_{e}, and by Lemma 6 we know that the polar of the out-ball B_{o} is the in-ball B_i. Therefore,

$$\mathrm{cc}(\Delta_{\mathrm{e}} \cup B_i) \subseteq G^*. \tag{143}$$

We are now able to prove

Theorem 10. *Let A be a subset of $\Delta \cap B_{\mathrm{o}}$. Then*

1. *A is a germ if and only if $A \subseteq A^*$.*
2. *A is a maximal germ if and only if $A = A^*$.*

Therefore, the terms "maximal germ" and "qplex" are equivalent.

Proof. The first statement is an immediate consequence of the definition. To prove the second statement we need to do a little work. This is because it is not immediately apparent that if A is a maximal germ then $A^* \subseteq B_{\mathrm{o}}$.

Suppose that A is a maximal germ. We know from the first part of the theorem that $A \subseteq A^*$. To prove the reverse inclusion let $u \in A^*$ be arbitrary. In order to show that $u \in A$ first consider the vector

$$\tilde{u} = c - \frac{r_i}{\|u - c\|}(u - c). \tag{144}$$

We have, for all $v \in A$,

$$-r_i r_{\mathrm{o}} \leq \langle \tilde{u} - c, v - c \rangle \leq r_i r_{\mathrm{o}}, \tag{145}$$

implying

$$\frac{1}{d(d+1)} \leq \langle \tilde{u}, v \rangle \leq \frac{2}{d(d+1)}. \tag{146}$$

Also

$$\frac{1}{d(d+1)} \leq \langle \tilde{u}, \tilde{u} \rangle = \frac{1}{d^2 - 1} \leq \frac{2}{d(d+1)}. \tag{147}$$

So $\tilde{u} \in A$. We now use this to show that $u \in A$. In fact

$$\begin{aligned} -r_i \|u - c\| &= \langle u - c, \tilde{u} - c \rangle \\ &= \langle u, \tilde{u} \rangle - \frac{1}{d^2} \\ &\geq -r_i r_{\mathrm{o}}, \end{aligned} \tag{148}$$

implying $u \in B_{\mathrm{o}}$. Consequently

$$\begin{aligned} \langle u, v \rangle &= \langle u - c, v - c \rangle + \frac{1}{d^2} \\ &\leq r_{\mathrm{o}}^2 + \frac{1}{d^2} \\ &= \frac{2}{d(d+1)}, \end{aligned} \tag{149}$$

for all $v \in A$. The fact that $u \in A^*$ means

$$\langle u, v \rangle \geq \frac{1}{d(d+1)}, \tag{150}$$

for all $v \in A$. Finally

$$\frac{1}{d(d+1)} \leq \|u - c\|^2 + \frac{1}{d^2} = \langle u, u \rangle \leq \frac{2}{d(d+1)}. \tag{151}$$

So $u \in A$. This completes the proof that if A is a maximal germ then $A = A^*$. The converse statement, that if $A = A^*$ then A is a maximal germ, is an immediate consequence of the definition. □

Let us note that this theorem means, in particular, that every maximal germ contains the basis simplex. If we start with the fundamental inequalities and assume that the set of points satisfying them is maximal, then that set turns out to be self-polar. Because the state space is contained within the probability simplex, the state space must *contain* the *polar of* the probability simplex, which by Lemma 4 is the basis simplex. In earlier papers on germs [1,2,15], the existence of the basis distributions was an extra assumption in addition to maximality; here, using the concept of polarity, we have been able to derive it.

Let us also note, as another consequence of this theorem, that if Q is a maximal germ, and if q is any element of Q, then there exists a measurement r and index a such that $q = s_a$, where s_a is the distribution

$$s_a(j) = \frac{r(a|j)}{\sum_k r(a|k)}. \tag{152}$$

This too was something that was assumed in older work [1,2,15], but which we are now in a position to derive. To see that it is true observe that the statement is trivial if $q = c$ (simply take r to be the one-outcome measurement). If, on the other hand, $q \neq c$ we can define

$$q' = c - \frac{r_i}{\|q - c\|}(q - c). \tag{153}$$

By construction $q' \in S_i$. So it follows from the theorem that $q' \in Q$. Consequently, if we define

$$r(a|i) = \begin{cases} \frac{d^2 r_i}{\|q-c\| + r_i} q(i) & a = 1, \\ \frac{d^2 \|q-c\|}{\|q-c\| + r_i} q'(i) & a = 2, \end{cases} \tag{154}$$

then r describes a two-outcome measurement such that

$$\frac{r(1|i)}{\sum_k r(1|k)} = q(i), \qquad \frac{r(2|i)}{\sum_k r(2|k)} = q'(i). \tag{155}$$

At this stage, we turn to the question of what germs can have in common, and how they can differ. In order to develop this topic, we introduce some more definitions. Given an arbitrary germ G, let $\mathcal{Q}_G$ denote the set of all qplexes containing G (necessarily nonempty, as we noted above).

Definition 4. The *stem* of a germ G is the set

$$\mathcal{S}(G) = \bigcap_{Q\in\mathcal{Q}_G} Q, \tag{156}$$

and the *envelope* of G is the set

$$\mathcal{E}(G) = \bigcup_{Q\in\mathcal{Q}_G} Q. \tag{157}$$

When G is the empty set $\mathcal{Q}_\emptyset$ is the set of all qplexes, without restriction. In that case we omit the subscript and simply denote it $\mathcal{Q}$. Similarly we write $\mathcal{S}(\emptyset) = \mathcal{S}$ and $\mathcal{E}(\emptyset) = \mathcal{E}$. We will refer to $\mathcal{S}$ and $\mathcal{E}$ as the principal stem and envelope.

Theorem 11. *Let G be a germ. Then*

$$\mathcal{S}(G) = \mathrm{cc}(\Delta_\mathrm{e} \cup B_i \cup G), \tag{158}$$
$$\mathcal{E}(G) = \Delta \cap B_\mathrm{o} \cap G^*. \tag{159}$$

In particular, $\mathcal{S}(G)$ and $\mathcal{E}(G)$ are mutually polar.

Proof. If Q is a qplex containing G we must have $Q = Q^* \subseteq \Delta \cap B_\mathrm{o} \cap G^*$. So

$$\mathcal{E}(G) \subseteq \Delta \cap B_\mathrm{o} \cap G^*. \tag{160}$$

On the other hand if p is any point in $\Delta \cap B_\mathrm{o} \cap G^*$ then $G \cup \{p\}$ is a germ, and so must be contained in some $Q \in \mathcal{Q}_G$. The second statement now follows.

To prove the first statement we take duals on both sides of

$$\bigcup_{Q\in\mathcal{Q}_G} Q = \Delta \cap B_\mathrm{o} \cap G^*. \tag{161}$$

We find

$$\begin{aligned}\mathcal{S}(G) &= \mathrm{cc}\Big(\Delta_\mathrm{e} \cup B_i \cup \mathrm{cc}\big(G \cup \{c\}\big)\Big)\\ &= \mathrm{cc}(\Delta_\mathrm{e} \cup B_i \cup G).\end{aligned} \tag{162}$$

□

Corollary 1. *The principal stem and envelope are given by*

$$\mathcal{S} = \mathrm{cc}(\Delta_\mathrm{e} \cup B_i), \tag{163}$$
$$\mathcal{E} = \Delta \cap B_\mathrm{o}. \tag{164}$$

Proof. Immediate. □

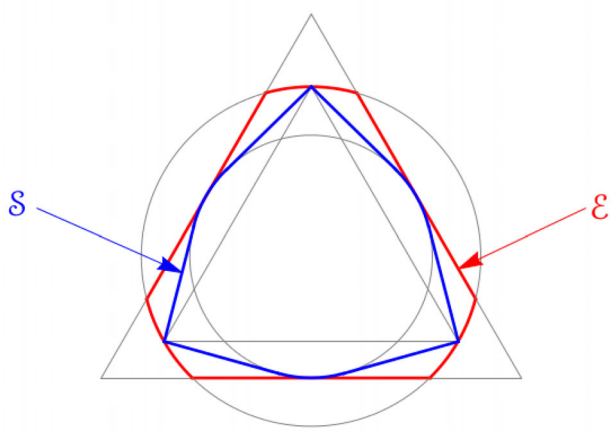

Fig. 3. The principal stem and envelope when $d > 2$. The sets Δ, Δ_e, S_o, S_i are shown in gray. The surface of every qplex lies between the blue and red surfaces. As with Figure 2 the diagram is schematic only. In particular, the total measure of the set $\mathcal{S}$ is much smaller in comparison with the total measure of the set $\mathcal{E}$ than the diagram suggests.

This result is illustrated schematically in Figure 3.

Corollary 2. *Let G be a closed, convex germ containing $\mathcal{S}$. Then*

$$\mathcal{S}(G) = G \qquad \mathcal{E}(G) = G^*. \tag{165}$$

Moreover, given arbitrary $p \in G^*$ such that $p \notin G$ there exist qplexes Q_1, Q_2 containing G such that

$$p \notin Q_1 \qquad p \in Q_2. \tag{166}$$

Proof. Immediate. □

Every germ can be extended to a qplex. It is natural to ask how many ways there are of performing the extension. The following theorem provides a partial answer to that question.

Theorem 12. *Let G be a closed, convex germ containing $\mathcal{S}$. If G is not already a qplex, then there are uncountably many qplexes containing G.*

Proof. It will be convenient to begin by introducing some notation. Given any two points p_1, $p_2 \in H$ we define

$$[p_1, p_2] = \{\lambda p_1 + (1-\lambda)p_2 : 0 \le \lambda \le 1\} \tag{167}$$
$$(p_1, p_2] = \{\lambda p_1 + (1-\lambda)p_2 : 0 < \lambda \le 1\} \tag{168}$$
$$[p_1, p_2) = \{\lambda p_1 + (1-\lambda)p_2 : 0 \le \lambda < 1\} \tag{169}$$
$$(p_1, p_2) = \{\lambda p_1 + (1-\lambda)p_2 : 0 < \lambda < 1\}. \tag{170}$$

Turning to the proof, suppose that G is not a qplex. Then we can choose $p \in G^*$ such that $p \notin G$. Let q be the point where $[c, p]$ meets the boundary of G. We will show that, for each $s \in (q, p)$ there exists a qplex Q_s such that $G \cup [c, s] \subseteq Q_s$ and $(s, p] \cap Q_s = \emptyset$. The result will then follow since $Q_s \neq Q_{s'}$ if $s \neq s'$.

To construct the qplex Q_s for given $s \in (q, p)$ observe that it follows from the basic theory of convex sets [19]

that there exists a hyperplane through s and not intersecting G. This means we can choose $u \in H$ such that

$$\langle u, v\rangle < \langle u, s\rangle = 1, \tag{171}$$

for all $v \in G$. Observe that for all $t \in (s, p)$ we have

$$t = \lambda s + (1-\lambda)c, \tag{172}$$

for some $\lambda > 1$ and, consequently,

$$\langle u, t\rangle = \frac{(d^2-1)\lambda + 1}{d^2} > 1. \tag{173}$$

Let

$$u' = \left(1 + \frac{1}{(d+1)(d^2-1)}\right)c - \frac{1}{(d+1)(d^2-1)}u, \tag{174}$$

and let $A = \mathrm{cc}\big(G \cup \{s\}\big)$. Then it is easily seen that $u' \in A^*$ while

$$\langle u', t\rangle < \frac{1}{d(d+1)}, \tag{175}$$

for all $t \in (s, p)$. A is a closed, convex germ containing $\mathcal{S}$, so it follows from Corollary 2 that there exists a qplex Q_s containing A and u'. By construction $t \notin Q_s$ for all $t \in (s, p)$, so Q_s has the required properties. □

The result just proved shows that there exist uncountably many qplexes. However, we would like to know a little more: namely, how many qplexes there are which are geometrically distinct. We now prove a series of results leading to Theorem 16, which states that there are uncountably many qplexes which are not isomorphic to each other, or to quantum state space.

Definition 5. Let $s \in H$ be arbitrary. We define the polar point of s to be the point

$$s^* = c - \frac{r_o r_i}{\|s-c\|^2}(s-c), \tag{176}$$

and the polar hyperplane of s to be the set

$$H_s = \left\{u \in H : \langle u, s\rangle = \frac{1}{d(d+1)}\right\}. \tag{177}$$

Observe that $s^{**} = s$ for all s, and

$$\langle s^*, s\rangle = \frac{1}{d(d+1)}, \tag{178}$$

(so $s^* \in H_s$). The relations between the polar $\{s\}^*$, the polar point s^* and the polar hyperplane H_s are depicted in Figure 4. It follows from these definitions that if s is any point on S_o (respectively S_i), then s^* is on S_i (respectively S_o).

Theorem 13. *Let G be a closed germ, and let s be any point on S_o. Then $s \in G$ if and only if s^* is on the boundary of G^*.*

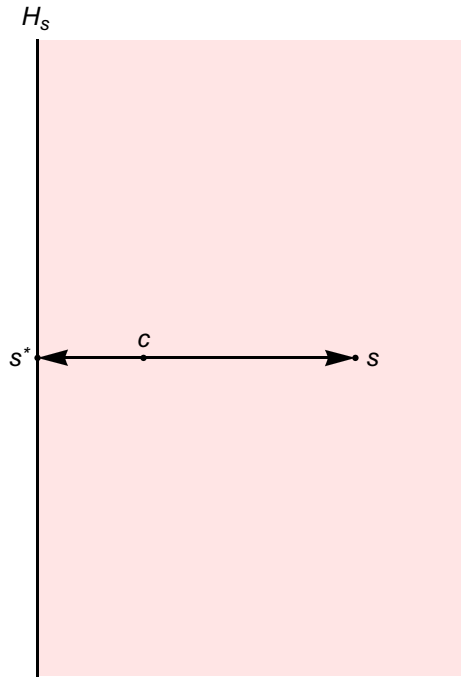

Fig. 4. Diagram to illustrate the relationships between the polar $\{s\}^*$, the polar point s^* and the polar hyperplane H_s. $\{s\}^*$ is the pink-shaded region to the right of H_s.

Remark 2. Specializing to the case when G is a qplex, the theorem says that the points where the boundary of G touches the out-sphere are antipodal to the points where it touches the in-sphere. This is a subtle property of quantum state space [41,42].

Proof. Suppose $s \in G$. Then it follows that $s^* \in S_i$. The fact that G is a germ means $G \subseteq \mathcal{E}$, implying $\mathcal{S} \subseteq G^*$. So $s^* \in G^*$. Moreover, if we define

$$t_n = \frac{n+1}{n}s^* - \frac{1}{n}c, \tag{179}$$

then

$$\langle t_n, s\rangle = \frac{1}{d(d+1)} - \frac{1}{nd^2(d+1)} < \frac{1}{d(d+1)}, \tag{180}$$

for all n. So t_n is a sequence outside G^* converging to s^*. We conclude that s^* is on the boundary of G^*.

Conversely, suppose s^* is on the boundary of G^*. Then we can choose a sequence $t_n \notin G^*$ such that $t_n \to s^*$. For each n there must exist $p_n \in G$ such that

$$\langle t_n, p_n\rangle < \frac{1}{d(d+1)}. \tag{181}$$

Since G is closed and bounded it is compact. A theorem of point set topology has it that in a compact set, every sequence contains a convergent subsequence. Therefore, we can choose a convergent subsequence $p_{n_j} \to p \in G$. Also, the fact that $t_{n_j} \to s^*$ means $t^*_{n_j} \to s$. So

$$\|p - s\|^2 = \lim_j \left(\|p_{n_j} - t^*_{n_j}\|^2\right). \tag{182}$$

We can expand the quantity inside the limit as

$$\|p_{n_j}-t^*_{n_j}\|^2 = \|p_{n_j}-c\|^2+\|t^*_{n_j}-c\|^2-2\left\langle t^*_{n_j}-c, p_{n_j}-c\right\rangle. \tag{183}$$

In turn, we have that

$$\lim_j \left(\|p_{n_j}-t^*_{n_j}\|^2\right) \leq \|p-c\|^2 - r_{\mathrm{o}}^2. \tag{184}$$

Because p is contained in the out-ball, its distance from c has to be less than r_{o}, meaning that

$$\|p-s\|^2 \leq \|p-c\|^2 - r_{\mathrm{o}}^2 \leq 0. \tag{185}$$

So p coincides with s, which consequently belongs to G. □

We next prove two results which show that we can restrict our attention to the out-sphere when trying to establish the existence of non-isomorphic qplexes. The first of these, Lemma 14, is a technical result which will also be used in Section 5.

Lemma 14. *Let G be a closed germ containing $\mathcal{S}$, and let*

$$C = G \cup (G^* \cap B_m), \tag{186}$$

where B_m is the mid-ball. Then C is a closed germ such that $C' \cap S_o = G \cap S_o$ for every germ C' containing C.

Proof. It follows from Theorem 3 and Lemma 6 that

$$C^* = G^* \cap \mathrm{cc}(G \cup B_m), \tag{187}$$

from which one sees that $C \subseteq C^*$. Moreover, the fact that $\mathcal{S} \subseteq G$ means $G^* \subseteq \Delta \cap B_{\mathrm{o}}$, implying $C \subseteq \Delta \cap B_{\mathrm{o}}$. So C is a closed germ containing G. Let C' be any germ containing C. It is immediate that $G \cap S_o = C \cap S_o \subseteq C' \cap S_o$. Suppose, on the other hand, that s is a point on S_o not belonging to G. Let G^*_{b}, C_{b} be the boundaries of G^*, C respectively. The fact that $C \cap B_m = G^* \cap B_m$ is easily seen to imply $G^*_{\mathrm{b}} \cap S_i = C_{\mathrm{b}} \cap S_i$. So it follows from Theorem 13 that $s^* \notin C_{\mathrm{b}}$. Since $S_i \subseteq C$ this means s^* must lie in the interior of C. So there exists $\lambda > 1$ such that

$$t = \lambda s^* + (1-\lambda)s, \tag{188}$$

is in C. Since

$$\langle t, s\rangle = \frac{2-\lambda}{d(d+1)} < \frac{1}{d(d+1)}, \tag{189}$$

it follows that $s \notin C^*$. Consequently $s \notin C'$. □

Theorem 15. *Let G be a closed germ containing the vertices of the basis simplex. Then there exists a qplex Q such that $Q \cap S_o = G \cap S_o$.*

Proof. Let

$$\tilde{G} = \mathrm{cc}(G \cup \Delta_{\mathrm{e}} \cup B_i). \tag{190}$$

Then $\tilde{G}$ is a closed germ containing $\mathcal{S}$. Moreover $\tilde{G} \cap S_o = G \cap S_o$. Let

$$C = \tilde{G} \cup \left(\tilde{G}^* \cap B_m\right). \tag{191}$$

Then it follows from Lemma 14 that C is a germ and that $Q \cap S_o = G \cap S_o$ for any qplex Q containing C. □

Before proving Theorem 16 we need to give a sharp definition of what it means for two qplexes to be isomorphic.

Definition 6. We say that two qplexes Q and Q' are isomorphic if and only if there exists a linear bijection $f\colon \mathbb{R}^{d^2} \to \mathbb{R}^{d^2}$ such that

1. $Q' = f(Q)$.
2. For all q_1, $q_2 \in Q$

$$\langle f(q_1), f(q_2)\rangle = \langle q_1, q_2\rangle. \tag{192}$$

We are now ready to prove the final result of this section.

Theorem 16. *There exist uncountably many qplexes which are not isomorphic to each other.*

Proof. We have

$$\langle e_i, e_j\rangle = \frac{d\delta_{ij}+d+2}{d(d+1)^2}, \tag{193}$$

for all i, j (c.f. Eq. (124)). So if we define

$$p_\theta = c + \cos\theta(e_1 - c) + \sin\theta(e_2 - c), \tag{194}$$

$$G_\theta = \{p_\theta, e_1, \ldots, e_{d^2}\}, \tag{195}$$

then, for sufficiently small ϵ, the set G_θ is a germ for all $\theta \in (0,\epsilon)$. It follows from Theorem 15 that we can choose qplexes Q_θ such that $Q_\theta \cap S_o = G_\theta \cap S_o = G_\theta$. By construction, the scalar products $\langle p_\theta, e_j\rangle$ are different for different choices of θ, and so qplexes Q_θ corresponding to different values of θ are non-isomorphic. Moreover, the fact that the intersection with S_o is finite means that Q_θ is non-isomorphic to quantum state space for all θ. □

So far we have been focussing on qplexes in general. However, it seems to us that the method of analysis employed is a potentially insightful way of thinking about the geometry of quantum state space.

4 Type-preserving measurements

We now come to the central result of this paper. We will show that the symmetry group of a qplex can be identified with a set of measurements, which in turn can be identified with a set of regular simplices within the qplex whose vertices all lie on the out-sphere.

Let Q be a qplex and r a measurement with n outcomes. For each $q \in Q$, let q_r be the distribution given by the urgleichung, equation (12). Then the map $q \to q_r$ takes Q to

$$Q_r = \{q_r\colon q \in Q\} \subseteq \Delta_n, \tag{196}$$

(where Δ_n is the $n-1$ dimensional probability simplex). We refer to Q_r as the measurement set, and the map $q \to q_r$ as the measurement map.

We are interested in measurements having d^2 outcomes for which the measurement set is another qplex. We will

refer to such measurements as type-preserving. We are particularly interested in the case when the measurement set is Q itself, in which case we will say that the measurement is Q-preserving.

Let r be an arbitrary measurement. Then it is easily seen that the urgleichung can be written in the alternative form

$$q_r(i) = \sum_j R_{ij} q(j), \tag{197}$$

where

$$R_{ij} = (d+1) r(i|j) - \frac{1}{d} \sum_k r(i|k). \tag{198}$$

We refer to R as the stretched measurement matrix. Note that equation (198) can be inverted:

$$r(i|j) = \frac{1}{d+1} \left(R_{ij} + \frac{1}{d} \sum_k R_{ik} \right). \tag{199}$$

So the stretched measurement matrix uniquely specifies the measurement.

Now specialize to the case of a type-preserving measurement. In that case it turns out that R must be an orthogonal matrix. To see this we begin by observing that, since the basis simplex belongs to both Q and Q_r, there must exist $s_i \in Q$, $s'_i \in Q_r$ such that

$$Rs_i = e_i, \qquad Re_i = s'_i. \tag{200}$$

We then have

Lemma 17. *Let R, s_i, s'_i be as above. Then*

1. $\det R = \pm 1$.
2. *s_i, $s'_i \in S_o$ for all i.*
3. *$\mathrm{cc}(\{s_i\})$ and $\mathrm{cc}(\{s'_i\})$ are regular simplices.*

Proof. The proof is based on the fact [43] that the simplices of maximal volume within a ball are precisely the regular simplices with vertices on the sphere that bounds the ball. The desired result follows from considering the simplex formed by the s_i and the origin (and the corresponding simplex formed by the s'_i and the origin). □

To complete the proof that R is an orthogonal matrix, we observe that maps from regular simplices to regular simplices are orthogonal. From this, we can derive the following theorem.

Theorem 18. *Let Q be a qplex, and let R be the stretched measurement matrix of a type-preserving measurement. Then R is an orthogonal matrix such that $Rc = c$. Moreover there exists a regular simplex with vertices $s_i \in Q \cap S_o$ such that*

$$R_{ij} = (d+1) s_i(j) - \frac{1}{d}, \tag{201}$$

$$Rs_i = e_i, \tag{202}$$

$$(Re_i)(j) = s_j(i). \tag{203}$$

Remark 3. We will refer to $\mathrm{cc}(\{s_i\})$ as the measurement simplex.

For a given qplex Q define

1. $\mathcal{T}_Q$ to be the class of type-preserving measurements.
2. $\mathcal{S}_Q$ to be the class of regular simplices with vertices in $Q \cap S_o$.
3. $\mathcal{O}_Q$ to be the class of orthogonal matrices R such that RQ is a qplex.

The previous theorem states that to each element of $\mathcal{T}_Q$ there corresponds an element of $\mathcal{S}_Q$ and an element of $\mathcal{O}_Q$. The next theorem we prove states that the correspondences are in fact bijective, so that we can identify the three classes $\mathcal{T}_Q$, $\mathcal{S}_Q$ and $\mathcal{O}_Q$.

Theorem 19. *Let Q be a qplex and let $s_i \in Q \cap S_o$ be the vertices of a regular simplex Δ_s. Then Δ_s is the measurement simplex of a type-preserving measurement. Likewise, if R is an orthogonal matrix such that RQ is also a qplex, then R is the stretched measurement matrix for a type-preserving measurement.*

Proof. Define

$$r(i|j) = s_i(j). \tag{204}$$

It is immediate that the $r(i|j)$ are the conditional probabilities defining a measurement with stretched measurement matrix

$$R_{ij} = (d+1) s_i(j) - \frac{1}{d}. \tag{205}$$

We need to show that the measurement is type-preserving. In other words, we need to show that the set RQ is a qplex. For all $q \in Q$

$$(Rq)(i) = (d+1) \langle s_i, q \rangle - \frac{1}{d}, \tag{206}$$

from which it follows

$$(Rq)(i) \geq 0 \qquad \sum_i (Rq)(i) = 1. \tag{207}$$

So $RQ \subseteq \Delta$. Also, it follows from the same considerations that led to Theorem 18 that R is orthogonal. The defining condition of a germ, equation (117), is invariant under orthogonal transformations. Therefore, RQ is a qplex.

We now prove the other direction of the correspondence. Let R be an orthogonal matrix such that RQ is a qplex. We know that the basis distributions e_i must belong to RQ. So, there exist $s_i \in Q$ such that

$$e_i = Rs_i. \tag{208}$$

Since $\det R = \pm 1$ we have, by the same argument used to prove Lemma 17, that the $s_i \in S_o$ and are the vertices of a regular simplex. It now follows from the considerations above that $\mathrm{cc}(\{s_i\})$ is the measurement simplex of a type-preserving measurement, with stretched measurement matrix

$$R'_{ij} = (d+1) s_i(j) - \frac{1}{d}. \tag{209}$$

By multiplying both sides of equation (208) by R^{T}, we find that

$$s_i(j) = \sum_k R^{\mathrm{T}}_{jk} e_i(k) = \frac{1}{d+1} R_{ij} + \frac{1}{d(d+1)} \sum_k R_{kj}. \tag{210}$$

Summing over i on both sides of this equation we find $\sum_k R_{kj} = 1$ for all j and, consequently, $R = R'$. □

At this stage, we recall our definition of an isomorphism between qplexes: Two qplexes Q and Q' are isomorphic if and only if there exists an inner-product-preserving map $f\colon \mathbb{R}^{d^2} \to \mathbb{R}^{d^2}$ that sends Q to Q'.

Theorem 20. *Let Q, Q' be qplexes. Then Q and Q' are isomorphic if and only if there is a type-preserving measurement on Q such that $Q' = RQ$, where R is the stretched measurement matrix.*

Proof. Sufficiency is immediate. To prove necessity suppose that $f\colon Q \to Q'$ is an isomorphism. The fact that f preserves scalar products on a set which spans $\mathbb{R}^{d^2}$ means that it must be represented by an orthogonal matrix. The claim now follows from Theorem 19. □

So far we have been looking at type-preserving measurements in general. Let us now focus on the special case of Q-preserving measurements. Suppose that we have two such measurements, with measurement matrices R, R'. Then RR' is also an orthogonal matrix, with the property that $RR'Q = Q$. So it follows from Theorem 19 that RR' is the stretched measurement matrix for a Q-preserving measurement. Similarly with R^{T}, the inverse. In short, the Q-preserving measurement maps form a group. For ease of reference let us give it a name:

Definition 7. Let Q be a qplex. The preservation group of Q, denoted G_Q, is the group of type-preserving measurement maps between Q and itself.

The elements of G_Q are symmetries of Q. The question naturally arises, whether they comprise *all* the symmetries. The above considerations are not sufficient to answer that question because they leave open the possibility that Q is invariant under orthogonal transformations which do not fix the origin of $\mathbb{R}^{d^2}$. The following theorem eliminates that possibility.

Theorem 21. *Let Q be a qplex. Then the preservation group is the symmetry group of Q.*

Proof. The symmetry group of a subset of a normed vector space is defined to be the group of isometries of the set. It has been shown above that every Q-preserving measurement map is an isometry of Q. We need to show the converse. Let f be an isometry of Q. It follows from Theorem 18 that $f(c) = c$.

Now define a map $\tilde{f}\colon Q - c \to Q - c$ by

$$\tilde{f}(u) = f(u + c) - c. \tag{211}$$

One easily sees that $\|\tilde{f}(u)\| = \|u\|$ for all $u \in Q - c$. Consequently

$$\tilde{f}(u) = Tu, \tag{212}$$

for some orthogonal transformation T of the subspace $H - c$. We may extend T to an orthogonal transformation R of the whole space $\mathbb{R}^{d^2}$ by defining $Rc = c$. It is then immediate that $RQ = Q$. The result now follows by Theorem 19. □

5 From preservation group to qplex

In this section we ask what conditions a subgroup of $\mathrm{O}(d^2)$ must satisfy in order to be the preservation group of some qplex. This will lead us to the question of when symmetries are powerful enough to determine a qplex essentially uniquely. Let Q be a qplex and $\mathcal{G}$ be its preservation group. Under what conditions can Q be maximally symmetric, in the sense that $\mathcal{G}$ is not a proper subgroup of the symmetry group of any qplex? The answer will turn out to depend upon how the group $\mathcal{G}$ acts on the basis simplex.

Quantum state space has the property that any pure state can be mapped to any other pure state by some unitary operation, that is, by some symmetry of the state space. Indeed, given any pure state, the set of all pure states is the orbit of the original state under the action of the symmetry group. This leads us to consider the general question of qplexes whose extremal points form a single orbit under the action of the qplex's symmetries. One can prove that if Q is such a qplex, then the symmetry group of Q is maximal, and furthermore, any other qplex Q' with the same symmetry group is identical to Q.

Given a group $\mathcal{G} \subseteq \mathrm{O}(d^2)$, can $\mathcal{G}$ be the preservation group of a qplex? It is easy to find a necessary condition. Following our previous paper [18], we introduce the concept of a stochastic subgroup:

Definition 8. A subgroup $\mathcal{G} \subseteq \mathrm{O}(d^2)$ is stochastic if, for all $R \in \mathcal{G}$,

$$R_{ij} \geq -\frac{1}{d} \quad \forall i, j \qquad \text{and} \qquad Rc = c. \tag{213}$$

Equivalently, we may say that a subgroup $\mathcal{G} \subseteq \mathrm{O}(d^2)$ is stochastic if every matrix in $\mathcal{G}$ is of the form

$$R_{ij} = (d+1)S_{ij} - \frac{1}{d}, \tag{214}$$

where $S_{ij} = s_i(j)$ is a doubly-stochastic matrix (hence the name). It can then be seen from Theorem 18 that every preservation group is a stochastic subgroup of $\mathrm{O}(d^2)$.

It is natural to ask whether the condition is sufficient as well as necessary, so that every stochastic subgroup of $\mathrm{O}(d^2)$ is the preservation group of some qplex. We have not been able to answer this question in full generality. However, we have obtained some partial results. We can show that any stochastic subgroup $\mathcal{G} \subseteq \mathrm{O}(d^2)$ is at least contained in the preservation group of some qplex. To see why, we start with a preliminary result.

Lemma 22. *Let $\mathcal{G}$ be a stochastic subgroup of $\mathrm{O}(d^2)$. For each $R \in \mathcal{G}$ define the vectors s_i^R by applying R to the basis distributions:*

$$s_i^R(j) = \frac{1}{d(d+1)}\left(dR_{ij} + 1\right). \tag{215}$$

Then $s_i^R \in \Delta \cap S_o$ for all i and $\mathrm{cc}(\{s_i^R\})$ is a regular simplex. Moreover

$$G = \{s_i^R \colon R \in \mathcal{G},\ i = 1, \ldots, d^2\} \tag{216}$$

is a germ.

Proof. Straightforward consequence of the definitions. □

Definition 9. Let $\mathcal{G}$ be a stochastic subgroup of $\mathrm{O}(d^2)$. The orbital germ is the orbit of the basis distributions under the action of $\mathcal{G}$, that is, the set G specified in the statement of Lemma 22.

Theorem 23. *Let $\mathcal{G}$ be a stochastic subgroup of $\mathrm{O}(d^2)$. Then there exists a qplex Q such that $\mathcal{G} \subseteq G_Q$.*

Proof. Let G be the orbital germ of $\mathcal{G}$, and let $\mathcal{A}_G$ be the set of all germs P such that

1. P contains G.
2. $RP = P$ for all $R \in \mathcal{G}$.

It follows from Zorn's lemma that $\mathcal{A}_G$ contains at least one maximal element. Let Q be such a maximal element. Observe that if P is in $\mathcal{A}_G$ then its convex closure is also in $\mathcal{A}_G$; consequently Q must be convex and closed. Observe, also, that if R is any element of $\mathcal{G}$, then c is in the interior of the simplex $\mathrm{cc}(\{s_i^R\})$; consequently c is in the interior of Q.

We claim that Q is in fact a qplex. For suppose it were not. Then we could choose $p \in \Delta \cap B_\mathrm{o} \cap Q^*$ such that $p \notin Q$. For each λ in the closed interval $[0, 1]$ define $p_\lambda = \lambda p + (1 - \lambda)c$. The fact that Q is closed, convex together with the fact that c is in the interior of Q means that there exists $\lambda_0 \in (0, 1)$ such that $p_\lambda \in Q$ if and only if $\lambda \in [0, \lambda_0]$. We have

$$\langle p, Rp_\lambda\rangle \geq \frac{1}{d(d+1)}, \tag{217}$$

for all $R \in \mathcal{G}$, $\lambda \in [0, \lambda_0]$. Consequently

$$\langle p_\lambda, Rp_\lambda\rangle \geq \frac{1}{d(d+1)} + \frac{1-\lambda}{d^2(d+1)}, \tag{218}$$

for all $\lambda \in [0, \lambda_0]$, $R \in \mathcal{G}$. By continuity this inequality must hold for all $\lambda \in [0, \lambda_0]$, $R \in \bar{\mathcal{G}}$, where $\bar{\mathcal{G}}$ is the closure of $\mathcal{G}$ in $\mathrm{O}(d^2)$. It follows that there must exist a fixed number $\mu \in (\lambda_0, 1]$ such that

$$\langle p_\mu, Rp_\mu\rangle \geq \frac{1}{d(d+1)}, \tag{219}$$

for all $R \in \mathcal{G}$. For suppose that were not the case. Then we could choose a sequence $\nu_n \downarrow \lambda_0$, and a sequence $R_n \in \mathcal{G}$, such that

$$\langle p_{\nu_n}, R_n p_{\nu_n}\rangle < \frac{1}{d(d+1)}, \tag{220}$$

for all n. The group $\bar{\mathcal{G}}$ is compact (because $\mathrm{O}(d^2)$ is compact[5]) as is the closed interval $\left[0, \frac{1}{d(d+1)}\right]$. Consequently we can choose a subsequence n_j such that $R_{n_j} \to \bar{R} \in \bar{\mathcal{G}}$ and

$$\langle p_{\nu_{n_j}}, R_{n_j} p_{\nu_{n_j}}\rangle \to a, \tag{221}$$

for some $a \in \left[0, \frac{1}{d(d+1)}\right]$. But this would imply that

$$\langle p_{\lambda_0}, \bar{R}p_{\lambda_0}\rangle = a \leq \frac{1}{d(d+1)}, \tag{222}$$

which is a contradiction.

Now consider the set

$$Q' = Q \cup \{Rp_\mu \colon R \in \mathcal{G}\}. \tag{223}$$

Observe that

$$(Rp_\mu)(i) = (d+1)\langle s_i^R, p_\mu\rangle - \frac{1}{d} \geq 0, \tag{224}$$

for all i and all $R \in \mathcal{G}$ (because $p_\mu \in Q^* \subseteq G^*$). So $Q' \subseteq \Delta$. It is immediate that $Q' \subseteq B_\mathrm{o}$ and $Q' \subseteq Q'^*$. So Q' is a germ such that $RQ' = Q'$ for all $R \in \mathcal{G}$, and which is strictly larger than Q – which is a contradiction.

It is now immediate that $\mathcal{G}$ is a subgroup of G_Q. □

We can make stronger statements if we introduce some new concepts.

Definition 10. A stochastic subgroup $\mathcal{G} \subseteq \mathrm{O}(d^2)$ is maximal if it is not contained in any larger stochastic subgroup.

Definition 11. A stochastic subgroup $\mathcal{G} \subseteq \mathrm{O}(d^2)$ is strongly maximal if it is maximal and if, in addition, the closed convex hull of the orbital germ is a qplex.

We then have the following results.

Corollary 3. *Let $\mathcal{G}$ be a maximal stochastic subgroup of $\mathrm{O}(d^2)$. Then there exists a qplex Q such that $\mathcal{G} = G_Q$.*

Proof. Immediate consequence of Theorem 23. □

Theorem 24. *Let $\mathcal{G}$ be a strongly maximal stochastic subgroup of $\mathrm{O}(d^2)$ and let G be the orbital germ. Then $\mathrm{cc}(G)$ is the unique qplex Q such that $\mathcal{G} = G_Q$.*

[5] Given any $M \in \mathrm{O}(n)$, columns of M form an orthonormal basis, meaning that any column of any M is a unit vector. M therefore lies within a topology that is the product of n copies of the sphere S^{n-1}. Because the topology of $\mathrm{O}(n)$ is compact, it is a compact group.

Proof. We know from Corollary 3 that there exists at least one qplex Q such that $\mathcal{G} = G_Q$. If Q, Q' are qplexes such that $\mathcal{G} = G_Q = G_{Q'}$ then Q, Q' must both contain $\mathrm{cc}(G)$, where G is the orbital germ. Since $\mathrm{cc}(G)$ is a qplex we must have $Q = \mathrm{cc}(G) = Q'$. □

This brings us back to the claim we made at the beginning of this section.

Corollary 4. *If Q is a qplex whose extreme points form a single orbit under the action of the preservation group, then the preservation group of Q is strongly maximal.*

Proof. Let Q be a qplex and $\mathcal{G}$ be its preservation group. Assume that the extremal points form a single orbit under the action of $\mathcal{G}$. The basis distributions are among the extremal points, so all extremal points are on the same orbit as any basis distribution. In other words, the orbital germ is the set of extreme points. Suppose that Q' is a qplex whose preservation group contains $\mathcal{G}$. Then Q' contains all the extremal points of Q, and thus, Q' contains Q. But a qplex is a maximal germ, so we must have $Q' = Q$. □

6 Characterizing qplexes isomorphic to quantum state space

We are, of course, most interested in qplexes corresponding to SIC measurements. In this section, we will define what it means for a qplex to be isomorphic to quantum state space. We will prove that if Q is a qplex isomorphic to quantum state space, then its preservation group is isomorphic to the projective extended unitary group, essentially the group of all unitaries and anti-unitaries with phase factors quotiented out. Then, we will establish the converse: If the preservation group of a qplex is isomorphic to the projective extended unitary group, then that qplex is isomorphic to quantum state space. This result indicates one way of recovering quantum theory from the urgleichung.

Definition 12. Let B_{H} be the space of Hermitian operators on d-dimensional Hilbert space and let S be the space of density matrices. We will say that a qplex Q is isomorphic to quantum state space if there exists an $\mathbb{R}$-linear bijection $f\colon B_{\mathrm{H}} \to \mathbb{R}^{d^2}$ such that

1. $Q = f(S)$.
2. For all ρ, $\rho' \in S$

$$\big\langle f(\rho), f(\rho')\big\rangle = \frac{\mathrm{tr}(\rho\rho') + 1}{d(d+1)}. \tag{225}$$

A qplex that is isomorphic to quantum state space will be designated a Hilbert qplex.

It is straightforward to verify that Definitions 6 and 12 are consistent, in the sense that if Q is a Hilbert qplex, and if Q' is any other qplex, then Q' is a Hilbert qplex if and only if it is isomorphic to Q in the sense of Definition 6.

Theorem 25. *Let Q be a qplex. Then a map $f\colon S \to Q$ is an isomorphism of quantum state space onto Q if and only if there is a SIC Π_j such that*

$$(f(\rho))(j) = \frac{1}{d}\,\mathrm{tr}(\rho\Pi_j), \tag{226}$$

for all j and all $\rho \in S$.

Remark 4. Thus, to each isomorphism of quantum state space onto Q, there corresponds a unique SIC. In particular a SIC exists in dimension d if and only if a Hilbert qplex exists in dimension d.

Proof. Suppose $f\colon S \to Q$ is an isomorphism. Define

$$\Pi_j = f^{-1}(e_j). \tag{227}$$

Then

$$\mathrm{tr}(\Pi_j\Pi_k) = d(d+1)\langle e_j, e_k\rangle - 1 = \frac{d\delta_{jk}+1}{d+1}. \tag{228}$$

So Π_j is a SIC. Moreover, for all $\rho \in S$, and all j,

$$\frac{1}{d}\,\mathrm{tr}(\rho\Pi_j) = (d+1)\langle f(\rho), e_j\rangle - \frac{1}{d} = (f(\rho))(j). \tag{229}$$

Suppose, on the other hand, $f\colon S \to Q$ is a map for which equation (226) is satisfied for some SIC Π_j. Then we can extend f to a linear bijection of B_{H} onto $\mathbb{R}^{d^2}$. We know from prior work [1,2,15] that $f(S)$ is a qplex. Since it is contained in Q we must have $f(S) = Q$. Moreover, since

$$\rho = \sum_j \left((d+1)(f(\rho))(j) - \frac{1}{d}\right)\Pi_j, \tag{230}$$

with a similar expression for ρ', we have

$$\mathrm{tr}(\rho\rho') = d(d+1)\langle f(\rho), f(\rho')\rangle - 1, \tag{231}$$

from which equation (225) follows. □

One might wonder if other qplexes, not isomorphic to Q (and we know that these exist, per Theorem 16 and Appendix A), correspond to other informationally complete POVMs. This is not the case. It follows from the foregoing that there is no measurement which will take us from a qplex of one kind to a qplex of a different, nonisomorphic kind.

Knowing this, let us characterize the preservation group of a Hilbert qplex Q. We define the extended unitary group, denoted $\mathrm{EU}(d)$, to be the group consisting of all unitary and anti-unitary operators, and the projective extended unitary group, denoted $\mathrm{PEU}(d)$, to be the quotient $\mathrm{EU}(d)/\mathrm{M}(d)$, where $\mathrm{M}(d)$ is the sub-group consisting of all unitaries of the form $e^{i\theta}I$, for some phase $e^{i\theta}$.

Theorem 26. *Let Q be a Hilbert qplex. Then G_Q is isomorphic to* $\mathrm{PEU}(d)$.

Proof. Straightforward consequence of Wigner's theorem [44]. □

We showed in Theorem 26 that if Q is a Hilbert qplex then G_Q is isomorphic to PEU(d). Now, we will prove the converse: if G_Q is isomorphic to PEU(d), then Q is a Hilbert qplex. It turns out, in fact, that a weaker statement is true: if G_Q contains a subgroup isomorphic to PU(d), then Q is a Hilbert qplex.

In the Introduction we remarked on the need for an extra assumption, additional to the basic definition of a qplex, which will serve to uniquely pick out those qplexes which correspond to quantum state space. The theorem we will prove momentarily supplies us with one possible choice for this assumption. As we remarked in the introduction, there may be others.

As a by-product of this result we obtain a criterion for SIC existence: namely, a SIC exists in dimension d if and only if PU(d) is isomorphic to a stochastic subgroup of O(d^2). We proved this result by another method in a previous paper [18], but this is the route by which we were originally led to it. Indeed, it is hard to see why it should occur to anyone that stochastic subgroups of O(d^2) might be relevant to SIC existence if they were not aware of the role that such subgroups play in the theory of qplexes.

The result depends on the following method for embedding a qplex in operator space. The question of whether a SIC exists in every dimension is very hard, and, indeed, is still unsolved. But if one simply asks for a set of operators $\Pi_1, \ldots, \Pi_{d^2}$ satisfying the equations

$$\mathrm{tr}(\Pi_j) = 1, \tag{232}$$

$$\mathrm{tr}(\Pi_j \Pi_k) = \frac{d\delta_{jk} + 1}{d+1}, \tag{233}$$

without imposing any further constraint – in particular, without requiring that the Π_j be positive semi-definite – then the problem becomes almost trivial. To see this consider the real Lie algebra $\mathfrak{su}(d)$ (i.e. the space of trace-zero Hermitian operators). Equipped with the Hilbert-Schmidt inner product

$$\langle B, B' \rangle = \mathrm{tr}(BB'), \tag{234}$$

this becomes a (d^2-1)-Euclidean space, so the existence of operators $B_1, \ldots, B_{d^2}$, each of length 1, and forming the vertices of a regular simplex, is guaranteed. These operators satisfy

$$\mathrm{tr}(B_j B_k) = \begin{cases} 1 & j = k; \\ -\frac{1}{d^2-1} & j \neq k. \end{cases} \tag{235}$$

If we now define

$$\Pi_j = \sqrt{\frac{d-1}{d}} B_j + \frac{1}{d} I, \tag{236}$$

then the Π_j satisfy equations (232) and (233). We will refer to them as a quasi-SIC.

Now let Q be an arbitrary qplex, and for each $q \in Q$ define, by analogy with equation (4)

$$\rho_q = \sum_j \left((d+1) q(j) - \frac{1}{d} \right) \Pi_j. \tag{237}$$

If Π_j really were a SIC, and if the $q(j)$ really were the outcome probabilities for a measurement with that SIC, then ρ_q would be a density matrix. In general, however, neither of those conditions need hold true. So, ρ_q will typically not be positive semi-definite (though it will be trace-1). We will refer to it as a quasi-density matrix. It will also be convenient to define

$$S_Q = \{\rho_q \colon q \in Q\}. \tag{238}$$

We will refer to S_Q as quasi-state space. It is easily verified that

$$0 \leq \langle \rho, \rho' \rangle \leq 1, \tag{239}$$

for all $\rho, \rho' \in S_Q$, just as is the case for genuine density matrices.

We are now in a position to prove

Theorem 27. *Let Q be a qplex. Then the following statements are equivalent:*

1. *G_Q contains a subgroup isomorphic to* PU(d).
2. *Q is a Hilbert qplex.*

Proof. The implication (2) $\implies$ (1) is an immediate consequence of Theorem 26. It remains to prove the implication (1) $\implies$ (2).

Let Π_j be a quasi-SIC, and use this quasi-SIC to map the qplex Q into operator space, creating the quasi-state space S_Q. The fact that the qplex Q contains a subgroup isomorphic to the projective unitary group PU(d) implies that the quasi-state space S_Q is invariant under unitary transformations. That is, the projective unitary symmetry of one set carries over to the other. This result is fairly natural; for completeness, we provide an explicit proof in Appendix B.

Suppose $q \in Q \in S_o$. Then

$$\mathrm{tr}(\rho_q) = \mathrm{tr}(\rho_q^2) = 1. \tag{240}$$

Also, it follows from equation (239) and unitary invariance of the quasi-state space that

$$0 \leq \mathrm{tr}(\rho_q U \rho_q U^\dagger) \leq 1. \tag{241}$$

for every unitary U. By choosing U to give the appropriate permutation of the eigenvalues we deduce that

$$0 \leq \sum_i \lambda_i^\uparrow \lambda_i^\downarrow \leq 1, \tag{242}$$

where $\lambda_i^\uparrow$ (respectively $\lambda_i^\downarrow$) are the eigenvalues of ρ_q arranged in increasing (respectively decreasing) order.

We now invoke a lemma proven in [18]. If λ is a vector in $\mathbb{R}^d$ such that

$$\sum_{j=0}^{d-1} \lambda_j = \sum_{j=0}^{d-1} \lambda_j^2 = 1, \tag{243}$$

then

$$\langle \lambda^\uparrow, \lambda^\downarrow \rangle \leq 0. \tag{244}$$

The inequality is saturated if and only if $d-1$ entries in λ are equal. This can occur when

$$\lambda^{\downarrow} = (1, 0, \ldots, 0), \tag{245}$$

or when

$$\lambda^{\downarrow} = \left(\frac{2}{d}, \ldots, \frac{2}{d}, \frac{2}{d} - 1\right). \tag{246}$$

So we must have

$$\sum_i \lambda_i^{\uparrow} \lambda_i^{\downarrow} = 0. \tag{247}$$

Moreover, the possible solutions for the eigenvalue spectrum $\lambda^{\downarrow}$ imply that either $\rho_q = P$ or $\rho_q = (2/d)I - P$ for some rank-1 projector P. If $d = 2$, then ρ_q is a rank-1 projector either way. Otherwise, if $d > 2$, suppose q, $q' \in Q \in S_o$ were such that $\rho_q = P$ and $\rho_{q'} = (2/d)I - P'$ where P and P' are rank-1 projectors. In that case there would be a unitary U such that $UP'U^{\dagger} = P$, which would mean, by unitary invariance, that the quasi-state space contained both P and $(2/d)I - P$. But

$$\mathrm{tr}\left(P\left(\frac{2}{d}I - P\right)\right) = \frac{2}{d} - 1 < 1, \tag{248}$$

which contradicts equation (239). We conclude that if $d > 2$ then, either ρ_q is a rank-1 projector for all $q \in Q$, or else $(2/d)I - \rho_q$ is a rank-1 projector for all $q \in Q$. In the latter case we may define a new quasi-SIC

$$\tilde{\Pi}'_j = \frac{2}{d}I - \tilde{\Pi}_j. \tag{249}$$

One easily verifies that the new quasi-state space is also unitarily invariant. Moreover, if we define

$$\rho'_q = \sum_j \left((d+1)q(j) - \frac{1}{d}\right)\tilde{\Pi}'_j, \tag{250}$$

then

$$\rho'_q = \frac{2}{d}I - \rho_q, \tag{251}$$

implying that ρ'_q is a rank-1 projector for all $q \in Q \in S_o$. There is therefore no loss of generality in assuming that our original quasi-state space is such that ρ_q is a rank-1 projector for all $q \in Q \in S_o$. Since

$$\rho_{e_i} = \tilde{\Pi}_i, \tag{252}$$

this means in particular that the Π_i are rank-1 projectors, and therefore constitute a genuine SIC.

Let us note that unitary invariance means that the set $\{\rho_q \colon q \in Q \in S_o\}$ does not merely consist of rank-1 projectors; it actually comprises all the rank-1 projectors. It follows, that if ρ is an arbitrary density matrix, and if $q(j) = (1/d)\,\mathrm{tr}(\rho\tilde{\Pi}_j)$, then q is a convex combination of points in $Q \in S_o$, and therefore $q \in Q$. Since the SIC probabilities are a qplex, it follows that Q does not contain any other points than these, and is therefore isomorphic to quantum state space as claimed. □

Let us observe that in proving this theorem we have incidentally shown that if there is a qplex Q which contains an isomorphic copy of $\mathrm{PU}(d)$, then a SIC exists in dimension d. So the theorem has the following corollary:

Corollary 5. *The following statements are equivalent:*

1. $\mathrm{PU}(d)$ *is isomorphic to a stochastic subgroup of* $\mathrm{O}(d^2)$.
2. *A SIC exists in dimension* d.

Proof. The implication (2) $\implies$ (1) is an immediate consequence of Theorem 27. To prove the implication (1) $\implies$ (2), let $\mathcal{G}$ be a stochastic subgroup of $\mathrm{O}(d^2)$ which is isomorphic to $\mathrm{PU}(d)$. It follows from Theorem 23 that there exists a qplex Q such that $\mathcal{G} \subseteq G_Q$. In view of Theorem 27 this implies Q is the set of outcome probabilities for a SIC measurement, which means, in particular, that a SIC must exist in dimension d. □

7 Discussion

Our investigation of qplexes exists in the context of many years' effort toward the goal of reconstructing quantum theory. Early pioneers of the subject, like Birkhoff and von Neumann, sought a broader mathematical environment in which quantum theory could be seen to dwell. This led to the subjects of quantum logic and Jordan algebras [45]. However, despite the mathematical developments, the influence on physics – and, indeed, on the philosophy thereof – was rather subdued. The intensely mathematical character of the work may have played a role in this. Moreover, this work predated the invention and integration into physics of information theory, which turned out to be a boon to the reconstruction enterprise. It also predated the theorems of Bell, Kochen and Specker [46,47], and thus it could not benefit from their insight into what is robustly strange about quantum physics.

One might say that the "modern age" of quantum reconstructions was inaugurated by Rovelli in 1996. He advocated a research program of deriving quantum theory from physical principles, in a manner analogous to the derivation of special relativity's mathematical formalism [48]. During the same time period, one of the authors (CAF) also began advocating this project [4,5,23]. An early success was Hardy's "Quantum theory from five reasonable axioms" [27,28], which pointed out the importance of what we call a Bureau of Standards measurement (Ref. [9], p. 368).

Looking over the papers produced in this "modern age," one technical commonality worth remarking upon is the idea of building up the unitary (or projective unitary) group from a universal gate set [25,26]. This is an idea from the field of quantum computation. For example, it is known that any unitary operator can be broken down into a sequence of two-level unitaries, applied in succession (Ref. [49], p. 188). Also, given a collection of N qubits, all the projective unitaries acting on their joint state space – that is, the group $\mathrm{PU}(2^N)$ – can be synthesized using single-qubit unitaries and an entangling gate, like a Controlled NOT operation, that can be applied

to any pair of qubits [50]. This suggests one way of making progress in the theory of qplexes, by replacing the unitarity assumption.

Recall that in any qplex, a set of mutually maximally distant points can have at most d elements [15,23]. Thus, although a qplex is originally defined as living within a d^2-dimensional space, in a sense it has an "underlying dimensionality" [23] equal to d. Consider a qplex Q, equipped with a set of d mutually maximally distant pure states. What if we require that any $d-1$ of those states defines a structure isomorphic to a smaller qplex? Applying this recursively, we arrive eventually at the condition that any two maximally distant points define a set of probability distributions isomorphic to a qplex with $d = 2$, which is automatically a Hilbert qplex. This is a strong condition, although it makes no direct mention of a particular symmetry group. At the moment, we see no way to satisfy this condition other than having Q be a Hilbert qplex.

Alternatively, one can try to make progress by relaxing the unitarity assumption. For example, instead of imposing a particular symmetry group, what if we seek the qplexes of maximal allowed symmetry? Assuming that a SIC exists in dimension d, then a qplex in Δ_{d^2} can be at least as symmetric as a Hilbert qplex. We conjecture that no qplex can be more symmetric than a Hilbert qplex, where we quantify the degree of symmetry by, for example, the dimension of the Lie group of qplex-preserving maps. This conjecture leads to another: we suspect that of all the qplexes of a given dimension, the Hilbert qplexes have maximal Euclidean volume.

Another outstanding question is, out of all the conceivable additions one could make to probability theory in order to relate expectations for different hypotheticals, why pick the urgleichung? To our knowledge, no one considered such a relation before quantum mechanics and the SIC representation. And yet, it is a comparatively mild modification of the classical relationship. This is particularly evident when the measurement on the ground is modeled by a set of d orthogonal projectors, i.e., when it is a von Neumann measurement. In that case,

$$q(j) = (d+1)\sum_i p(i)r(j|i) - 1. \tag{253}$$

This is just a rescaling and shifting of the classical formula [23].

In Section 2, we began with a general affine relationship between Bureau of Standards probabilities and the probabilities for other experiments. By invoking a series of assumptions, we narrowed the parameter values in the generalized urgleichung down to those that occur in quantum theory. (Our last assumption, which fixed the upper bound at twice the lower bound, may be related to the choice of complex numbers over real numbers and quaternions for Hilbert-space coordinates [2]. For an unexpected connection between SICs and the normed division algebras, see [51,52].) This has the appealing feature that a linear stretching is just about the simplest deformation of the classical Law of Total Probability that one can imagine. However, this area is still, to a great extent, unknown territory: Why linearity? Are qualitatively greater departures from classicality mathematically possible?

Many of the quantum reconstruction efforts to date share the feature that they make quantum physics as unremarkable as possible: while the technical steps from axioms to theorems are unassailable, the choice of axioms gives little insight into what is truly strange about quantum phenomena. To borrow a phrase from David Mermin, these re-expressions tend to make quantum theory sound "benignly humdrum" [53].

For example, should one aim to derive quantum theory from the fact that quantum states cannot be cloned? Arguably not: even classical distributions over phase space are uncloneable [54]. What about quantum teleportation? At root, teleportation is a protocol for making information about one system instead relevant to another, and it has exact analogues in classical statistical theories [55–57]. In 2003, Clifton, Bub and Halvorson [58] proposed a derivation of quantum theory that started with C^* algebras and then added, as postulates, some results of quantum information science, such as the no-broadcasting theorem [59]. However, the no-broadcasting theorem – despite its original motivation (Ref. [9], p. 2235) – also applies in classical statistical theories [55–57], and thus seems a poor foundation to build the quantum upon. Overall, it seems that choosing C^* algebras for a starting point implicitly does a great deal of the work already (Ref. [9], p. 1125).

Similarly, a more recent derivation by Chiribella, D'Ariano and Perinotti [60] invokes, at a key juncture, the postulate that any mixed state can be treated as a marginal of a pure state ascribed to a larger system. This postulate, the purifiability of mixed states, is an essential ingredient in their recovery of quantum theory. As with the examples above, however, it is also true in classical statistical theories [55–57,61]. From that perspective, it is consequently a less than fully compelling candidate for the essence of quantumness.

By contrast, we have chosen as our starting point what we consider to be the "jugular vein" of quantum strangeness: Theories of intrinsic hidden variables do so remarkably badly at expressing the vitality of quantum physics. The urgleichung is our way of stating this physical characteristic of the natural world in the language of probability. Quantum states, it avers, are catalogues of expectations – but *not* expectations about hidden variables. This view is in line with "participatory realist" interpretations of quantum mechanics [62,63], like QBism [1,3,64] and related approaches [65–67].

C.A.F. thanks Ben Schumacher for helpful discussions and the Max Planck Institute for Quantum Optics for a safe haven in which to work out some of these radical ideas. M.A. acknowledges support by the Australian Research Council via EQuS project number CE11001013. We thank John DeBrota for comments. Open access funding provided by Max Planck Society.

Author contribution statement

The content of this paper arose from daily chalkboard interactions over a period of time between all the authors.

CAF is predominantly responsible for formulating the technical questions addressed in the paper; BCS worked out the detailed conceptual aspects and wrote much of the manuscript. MA and HZ provided many of the mathematical proofs.

Appendix A: A qplex which does not correspond to a SIC measurement

By definition, a qplex is a subset of the probability simplex Δ_{d^2} such that each pair of points within it satisfy the fundamental inequalities,

$$\frac{1}{d(d+1)} \leq \sum_i p(i)s(i) \leq \frac{2}{d(d+1)}. \tag{A.1}$$

We can construct a qplex which is not isomorphic to quantum state space in the following way. Begin with a set A defined by the intersection of the probability simplex with the ball

$$\sum_i p(i)^2 \leq \frac{2}{d(d+1)}. \tag{A.2}$$

Our plan is to trim this set down until it becomes a qplex. First, we break A into $d^2!$ regions, which we label F_k, for $k = 1, \ldots, d^2!$. We define the region F_1 to be all probability vectors in the set A whose entries appear in decreasing magnitude. That is,

$$F_1 = \left\{p : p \in A \text{ and } p(1) \geq p(2) \geq \cdots \geq p(d^2)\right\}. \tag{A.3}$$

The region F_1 is consistent with the fundamental inequalities, because for every $p \in F_1$,

$$\langle p, p\rangle \geq \langle p, c\rangle \geq \frac{1}{d^2} > \frac{1}{d(d+1)}. \tag{A.4}$$

We define the other regions F_k analogously. Because k runs from 1 to $(d^2)!$, it labels the permutations in the symmetric group on d^2 elements. Each F_k consists of the vectors obtained by taking the vectors in F_1 and permuting the components according to permutation k. All of the regions F_k so defined will be internally consistent.

To obtain a qplex Q, start with F_1 and include all the points from F_1 in Q. Then, take all the points from F_2 that are consistent with all the points in F_1, and include them in Q. Continue in this manner, adding the points in each F_k that are consistent with every point added to Q so far. The end result will be a qplex that is surely not isomorphic to quantum state space.

Appendix B: Unitary symmetry of quasi-state spaces

Let Π_j be a quasi-SIC, as defined in equations (232) and (233) of the main text. For each $U \in \mathrm{PU}(d)$ we have a matrix S^U_{jk} such that

$$U\Pi_j U^\dagger = \sum_k S^U_{jk}\Pi_k. \tag{B.1}$$

The matrix is given explicitly by

$$S^U_{jk} = \frac{d+1}{d}\,\mathrm{tr}\left(\Pi_k U \Pi_j U^\dagger\right) - \frac{1}{d}, \tag{B.2}$$

from which one sees

$$\sum_j S^U_{jk} = 1, \qquad \sum_k S^U_{jk} = 1, \tag{B.3}$$

and

$$\sum_k S^U_{ik}S^U_{jk} = \frac{d+1}{d}\,\mathrm{tr}\left(\left(\sum_k S^U_{ik}\Pi_k\right)U\Pi_j U^\dagger\right) - \frac{1}{d} = \delta_{ij}. \tag{B.4}$$

So S^U_{ij} is an orthogonal matrix.

We now appeal to the assumption that G_Q contains a subgroup isomorphic to $\mathrm{PU}(d)$. So for each $U \in \mathrm{PU}(d)$ there exists an orthogonal matrix $R^U_{jk} \in G_Q$. It can be proven that, up to equivalence, the adjoint representation of $\mathrm{PU}(d)$ is the only nontrivial irreducible representation of $\mathrm{PU}(d)$ having degree d^2-1 or smaller, when $d \geq 2$ [18]. Thus, the two representations here must be equivalent, so that

$$R^U = TS^UT^{-1}, \tag{B.5}$$

for all U and some fixed orthogonal matrix T. Summing over k on both sides of

$$\sum_j R^U_{ij}T_{jk} = \sum_j T_{ij}S^U_{jk}, \tag{B.6}$$

and appealing to the fact that the representations are irreducible on the subspace orthogonal to c we deduce that

$$\sum_j T_{ij} = t, \tag{B.7}$$

for some constant t, independent of i. Similarly

$$\sum_i T_{ij} = s, \tag{B.8}$$

for some constant s, independent of i. Since

$$d^2t = \sum_{ij} T_{ij} = d^2 s, \tag{B.9}$$

we must in fact have $s = t$. Multiplying both sides of

$$\sum_j T_{ij} = t, \tag{B.10}$$

by T_{ik} and summing over i we find

$$1 = t \sum_i T_{ik} = t^2. \tag{B.11}$$

So $t^2 = \pm 1$. If $t = -1$ we can make the replacement $T \to -T$ without changing equation (B.5). We may therefore assume, without loss of generality,

$$\sum_j T_{ij} = \sum_j T_{ji} = 1, \tag{B.12}$$

for all i. It follows that, if we define

$$\tilde{\Pi}_i = \sum_j T_{ij} \Pi_j, \tag{B.13}$$

then the $\tilde{\Pi}_i$ are also a quasi-SIC. Moreover

$$\begin{aligned} U\tilde{\Pi}_i U^\dagger &= \sum_{j,k} T_{ij} S^U_{jk} T^{\mathrm{T}}_{kl} \tilde{\Pi}_l \\ &= \sum_l R^U_{il} \tilde{\Pi}_l. \end{aligned} \tag{B.14}$$

Suppose we now use the $\tilde{\Pi}_i$ to map Q into operator space by defining

$$\rho_q = \sum_j \left((d+1) q(j) - \frac{1}{d} \right) \tilde{\Pi}_j, \tag{B.15}$$

for all $q \in Q$. It follows from the foregoing that, for all $q \in Q$

$$U \rho_q U^\dagger = \rho_{q'}, \tag{B.16}$$

where

$$q'(j) = \sum_k R^U_{kj} q(k) \tag{B.17}$$

is also in Q. It follows that the quasi-state space $\{\rho_q \colon q \in Q\}$ is invariant under unitary transformations.

Appendix C: An alternate route to the fundamental inequalities

In the main text, we began with the urgleichung and eventually arrived at the fundamental inequalities

$$\frac{1}{d(d+1)} \leq \langle p, s \rangle \leq \frac{2}{d(d+1)}, \tag{C.1}$$

proving in Theorem 10 that a self-polar subset of the outball B_{o} is a maximal germ. Because Theorem 10 is an if-and-only-if result, it is natural to wonder if one could argue for the fundamental inequalities from a different premise, in which case self-polarity would be a consequence of assuming maximality.

One counterintuitive feature of quantum theory is that two quantum states can be perfectly distinguishable by a von Neumann measurement, yet less distinguishable by an informationally complete measurement [52,68,69]. This runs counter to experience with classical probability and stochastic processes, which leads one to think of a non-IC measurement as a coarse-graining (or a convolution by some kernel) of an IC measurement. If hypothesis A is that the system is in region A of phase space, and hypothesis B is that the system is in region B, classical intuition says that hypothesis A and B being perfectly distinguishable means that their regions have no overlap. Therefore, if we measure where the system is in phase space – the fundamental classical image of what an IC experiment can be – then some outcomes would be consistent with hypothesis A, some with hypothesis B, and none with both.

In quantum physics, two pure states being orthogonal means that the overlap of their SIC representations is minimal, but minimal is not zero. If we regard two orthogonal states $|0\rangle$ and $|1\rangle$ as two hypotheses that Alice can entertain about how a system will behave, then there exists some measurement with the property that no outcome is compatible with both hypotheses. Whatever the outcome of that experiment, one hypothesis or the other will be excluded [68]. But the two hypotheses $|0\rangle$ and $|1\rangle$ have SIC representations p_0 and p_1, and $\langle p_0, p_1 \rangle = 1/(d(d+1))$. The measurement that defines the SIC representation, although informationally complete, does not itself automatically exclude either hypothesis, because some possible outcomes of it are consistent with both.

With this motivation, we derive quantum state space in the following way. We again postulate a Bureau of Standards measurement, but we assume as little as possible about the meshing of probability distributions. Instead of the urgleichung (27), we merely postulate some functional relation [70],

$$q(j) = F\left(\{ p(i), r(j|i) : i = 1, \ldots, N \} \right), \tag{C.2}$$

with the property that state vectors with nonzero overlap are incompatible hypotheses with respect to some measurement. We assume, then, that the inner product of two state-space vectors is bounded below, and take this as an aspect of quantum strangeness. Then, we assume that certainty is bounded. This is less strange, since even classically, we can imagine a constraint that probability distributions can never get too focused. These two postulates tell us that the inner product of two state vectors lies in the interval $[L, U]$. Note that U, being an upper bound on $\langle p, p \rangle$, has an interpretation as an upper bound on an *index of coincidence,* which is inversely related to the *effective population size* [31,69,71]. Imagine an urn filled with marbles in N different colors. We draw a marble at random from the urn, note its color, replace it and draw at random again. If all colors are equally probable, then the probability of obtaining the same color twice in succession is $1/N$. More generally, if the colors are weighted by

some probability vector p, then the probability of obtaining the same color twice – i.e., a "coincidence" of colors – is $\langle p, p\rangle$. So, we can take the reciprocal of this quantity as the effective number of colors present. Regarding the probability vector p as a hypothesis about a system, the effective population size

$$N_{\text{eff}}(p) = \frac{1}{\langle p, p\rangle} \tag{C.3}$$

is the effective number of experiment outcomes that are compatible with that hypothesis. Given two probability vectors p and s, we can take

$$N_{\text{eff}}(p, s) = N_{\text{eff}}(p) N_{\text{eff}}(s) \langle p, s\rangle = \frac{\langle p, s\rangle}{p^2 s^2} \tag{C.4}$$

as the effective number of outcomes compatible with both hypotheses p and s.

By following the logic in Section 2, we can get an upper bound on the size of a Mutually Maximally Distant set. If we postulate that this bound is saturated, we can relate L, U and N to the effective dimensionality:

$$d = 1 + \frac{U - 1/N}{1/N - L}. \tag{C.5}$$

If we take $L = 0$ in the above expression, which we can heuristically regard as going to the "classical limit," then we end up with $d = NU$. This says that the total number of MMD states is the total size of the sample space (N), divided by the area per state, i.e., the effective population size $1/U$.

Instead of taking $L = 0$, if we choose – for whatever reason – that $L = U/2$ and that $N = d^2$, we get the familiar upper and lower bounds that define a germ. Postulating that our state space is maximal then implies that it is self-polar. Because the state space is contained within the probability simplex, it contains the polar of the probability simplex, which is the basis simplex. By Theorem 21, all the isometries of this set are specified by the regular simplices whose vertices are valid states lying on the out-sphere.

Suppose that p and s are two *pure* states. Then

$$N_{\text{eff}}(p) = N_{\text{eff}}(s) = \frac{d(d+1)}{2}, \tag{C.6}$$

and

$$N_{\text{eff}}(p, s) = \frac{d^2(d+1)^2}{4} \langle p, s\rangle \geq \frac{d(d+1)}{4}. \tag{C.7}$$

Thus, the fundamental inequalities imply that two hypotheses of maximal certainty can only disagree by so much that their overlap is *half* the effective number of outcomes consistent with either hypothesis alone.

We note that Wootters [72], Hardy [27] and others [34] have used various premises to argue for a relation of the form $N = d^2$. It bears something of the flavor of a classical state space whose points are labeled by discretized position and momentum [57,73] (and this resonates sympathetically with the fact that the Weyl-Heisenberg group, which is projectively equivalent to $\mathbb{Z}_d \times \mathbb{Z}_d$, is the canonical way to generate SICs [10,74]). However, at the moment we find it neither an obvious choice nor a consequence of a uniquely compelling assumption.

References

1. C.A. Fuchs, R. Schack, Rev. Mod. Phys. **85**, 1693 (2013)
2. C.A. Fuchs, R. Schack, Found. Phys. **41**, 345 (2011)
3. C.A. Fuchs, arXiv:1003.5209 [quant-ph] (2010)
4. C.A. Fuchs, in *Quantum Theory: Reconsideration of Foundations*, edited by A. Khrennikov (Växjö University Press, 2002), pp. 463–543
5. C.A. Fuchs, *Coming of Age with Quantum Information: Notes on a Paulian Idea* (Cambridge University Press, Cambridge, UK, 2010).
6. A. Peres, Am. J. Phys. **46**, 745 (1978)
7. G. Zauner, Int. J. Quantum Inf. **9**, 445 (2011)
8. J.M. Renes, R. Blume-Kohout, A.J. Scott, C.M. Caves, J. Math. Phys. **45**, 2171 (2004)
9. C.A. Fuchs, *My Struggles with the Block Universe (2014); foreword by M. Schlosshauer*, edited by B.C. Stacey, arXiv:1405.2390 [quant-ph]
10. M. Appleby, S. Flammia, G. McConnell, J. Yard, Generating Ray Class Fields of Real Quadratic Fields via Complex Equiangular Lines, arXiv:1604.06098 [quant-ph] (2016)
11. A.J. Scott, M. Grassl, J. Math. Phys. **51**, 042203 (2010)
12. D.M. Appleby, H. Yadsan-Appleby, G. Zauner, Quantum Inf. Comput. **13**, 672 (2013)
13. S.T. Flammia, Unpublished (2004)
14. N.S. Jones, N. Linden, Phys. Rev. A **71**, 012324 (2005)
15. D.M. Appleby, Å. Ericsson, C.A. Fuchs, Found. Phys. **41**, 564 (2011)
16. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985)
17. D.M. Appleby, S.T. Flammia, C.A. Fuchs, J. Math. Phys. **52**, 022202 (2011)
18. D.M. Appleby, C.A. Fuchs, H. Zhu, Quantum Inf. Comput. **15**, 61 (2015)
19. B. Grünbaum, *Convex Polytopes* (Springer, 2003)
20. G.M. Ziegler, *Lectures on Polytopes* (Springer, 1998)
21. N. Koblitz, *Introduction to Elliptic Curves and Modular Forms* (Springer, 1993)
22. A. Wiles, Ann. Math. **141**, 443 (1995)
23. C.A. Fuchs, B.C. Stacey, in *Quantum Theory: informational Foundations and Foils* (Springer, 2016)
24. B. Coecke, R. Duncan, A. Kissinger, Q. Wang, in *Quantum Theory: Informational Foundations and Foils* (Springer, 2016)
25. P.A. Hoehn, C. Wever, Phys. Rev. A **95**, 012102 (2017)
26. L. Masanes, M.P. Müller, New J. Phys. **13**, 063001 (2011)
27. L. Hardy, Quantum Theory From Five Reasonable Axioms, arXiv:quant-ph/0101012v4 (2001)
28. R. Schack, Found. Phys. **33**, 1461 (2003)
29. H. Barnum, M.P. Müller, C. Ududec, New J. Phys. **16**, 123029 (2014)
30. C.A. Fuchs, R. Schack, in *Probability and Physics*, edited by Y. Ben-Menahem, B. Hemmo (Springer, 2012), pp. 233–47

31. B.C. Stacey, Multiscale Structure in Eco-Evolutionary Dynamics, Ph.D. thesis, Brandeis University (2015)
32. E.T. Jaynes, Phys. Rev. **108**, 171 (1957)
33. W. Ochs, Erkenntnis **16**, 339 (1981)
34. C.M. Caves, C.A. Fuchs, R. Schack, J. Math. Phys. **43**, 4537 (2002)
35. B.C. Stacey, Philos. Trans. R. Soc. A **374**, 2068 (2016)
36. M. Plávala, Phys. Rev. A **94**, 042108 (2016)
37. A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, 1993)
38. R. Haag, Lect. Notes Phys. **153**, 168 (1982)
39. H. Araki, Commun. Math. Phys. **75**, 1 (1980)
40. M. Zorn, Bull. Am. Math. Soc. **41**, 667 (1935)
41. G. Kimura, A. Kossakowski, Open Syst. Inf. Dyn. **12**, 207 (2005)
42. D.M. Appleby, Opt. Spectrosc. **103**, 416 (2007)
43. M.S. Klamkin, G.A. Tsintsifas, Math. Mag. **52**, 20 (1979)
44. E.P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quanten mechanik der Atomspektren* (Friedrich Vieweg und Sohn, 1931), pp. 251–254. Translation by J.J. Griffin in *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, 1959), pp. 233–236
45. K. Mccrimmon, Bull. Am. Math. Soc. **84**, 612 (1978)
46. N.D. Mermin, Rev. Mod. Phys. **65**, 803 (1993)
47. N.D. Mermin, Rev. Mod. Phys. **88**, 039902 (2016)
48. C. Rovelli, Int. J. Theor. Phys. **35**, 1637 (1996)
49. M.A. Nielsen, I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2011)
50. A. Harrow, Quantum Inf. Comput. **8**, 715 (2008)
51. B.C. Stacey, Sporadic SICs and the Normed Division Algebras, arXiv:1605.01426 [quant-ph] (2016)
52. B.C. Stacey, Geometric and Information-Theoretic Properties of the Hoggar Lines, arXiv:1609.03075 [quant-ph] (2016)
53. N.D. Mermin, Phys. Today **42**, 9 (1989)
54. C.M. Caves, C.A. Fuchs, Ann. Israel Phys. Soc. **12**, 226 (1996)
55. R.W. Spekkens, Phys. Rev. A **75**, 032110 (2007)
56. S.D. Bartlett, T. Rudolph, R.W. Spekkens, Phys. Rev. A **86**, 012103 (2012)
57. R.W. Spekkens, Quasi-quantization: classical statistical theories with an epistemic restriction, arXiv:1409.5041 [quant-ph] (2014)
58. R. Clifton, J. Bub, H. Halvorson, Found. Phys. **33**, 1561 (2003)
59. H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa, Phys. Rev. Lett. **76**, 2818 (1996)
60. G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A **84**, 012311 (2011)
61. L. Disilvestro, D. Markham, Quantum protocols within Spekkens' toy model, arXiv:1608.09012 [quant-ph] (2016)
62. A. Cabello, Interpretations of quantum theory: a map of madness, arXiv:1509.04711 [quant-ph] (2015)
63. C.A. Fuchs, in *Information & Interaction: Eddington, Wheeler, and the Limits of Knowledge*, edited by I.T. Durham, D. Rickles (2017)
64. C.A. Fuchs, N.D. Mermin, R. Schack, Am. J. Phys. **82**, 749 (2014)
65. A. Zeilinger, Nature **438**, 743 (2005)
66. J. Kofler, A. Zeilinger, Eur. Rev. **18**, 469 (2010)
67. D.M. Appleby, Mind and Matter, arXiv:1305.7381 [physics.hist-ph] (2013)
68. C.M. Caves, C.A. Fuchs, R. Schack, Phys. Rev. A **66**, 062111 (2002)
69. B.C. Stacey, Mathematics **4**, 36 (2016)
70. C.A. Fuchs, R. Schack, Phys. Scripta **90**, 015104 (2014)
71. T. Leinster, C.A. Cobbold, Ecology **93**, 477 (2012)
72. W.K. Wootters, Found. Phys. **16**, 391 (1986)
73. H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, 1950), translated from the German by H.P. Robertson
74. H. Zhu, J. Phys. A **43**, 305305 (2010)
75. D.M. Appleby, H.B. Dang, C.A. Fuchs, Entropy **16**, 1484 (2014)