Classification of all alternatives to the Born rule in terms of informational properties

Thomas D. Galley and Lluis Masanes

Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
June 14, 2017

The standard postulates of quantum theory can be divided into two groups: the first one characterizes the structure and dynamics of pure states, while the second one specifies the structure of measurements and the corresponding probabilities. In this work we keep the first group of postulates and characterize all alternatives to the second group that give rise to finite-dimensional sets of mixed states. We prove a correspondence between all these alternatives and a class of representations of the unitary group. Some features of these probabilistic theories are identical to quantum theory, but there are important differences in others. For example, some theories have three perfectly distinguishable states in a two-dimensional Hilbert space. Others have exotic properties such as lack of "bit symmetry", the violation of "no simultaneous encoding" (a property similar to information causality) and the existence of maximal measurements without phase groups. We also analyze which of these properties single out the Born rule.

1 Introduction

There is a long-standing debate on whether the structure and dynamics of pure states already encodes the structure of measurements and probabilities. This discussion arises within the dynamical description of a quantum measurement, decoherence theory, and the many-worlds interpretation. There have been attempts to derive the Born rule [1–3], however these are often deemed controversial. In this work we suggest a more neutral approach where we consider *all* possible alternatives to the Born rule and explore their physical and informational properties.

More precisely, we provide a complete classification of all the alternatives to the structure of measurements and the formula for outcome probabilities with the following property: in any system with a finite-dimensional Hilbert space, the number of parameters that is required to specify a mixed state is finite. This property is necessary if state estimation can be performed with a finite number of measurements and no additional assumptions. We also show that, among all these alternative theories, those which have no restriction on the allowed effects violate a physical principle called "bit symmetry" [4], which is satisfied by quantum theory. Bit symmetry states that any pair of perfectly distinguishable pure states can be reversibly mapped to any other. Thus the Born rule is the only possible probability assignment to measurement outcomes under these requirements. We hope that these results may help to settle the above mentioned debate.

Most other known non-quantum general probabilistic theories seem somewhat pathological [5–7] with some exceptions such as quantum theory over the field of real numbers [8] or quaternions [9] and theories based on euclidean Jordan algebras [10].

For example they often violate most of the axioms used to reconstruct quantum theory, rather than just one or two per reconstruction. In particular, modifying the Born rule without creating inconsistencies or absurdities has been argued to be difficult (see, e.g. [11]).

For example, if we keep the usual association of the outcomes of a measurement with the elements of an orthonormal basis $\{|k\rangle\}$, a straightforward alternative for the probability of $|k\rangle$ given sate $|\psi\rangle$ is

$$P(k|\psi) = \frac{|\langle k|\psi\rangle|^\alpha}{\sum_{k'} |\langle k'|\psi\rangle|^\alpha} , \qquad (1)$$

Thomas D. Galley: thomas.galley.14@ucl.ac.uk

with $\alpha \neq 2$. However, this violates the finiteness assumption stated above.

In this work, we provide many non-quantum probabilistic theories with a modified Born rule which still inherit many of quantum theory's properties. These theories can be used as foils to understand quantum theory "from the outside". In particular, we analyse in full detail all finite-dimensional state spaces where pure states are rays in a 2-dimensional Hilbert space. We discuss properties such as the number of distinguishable states, bit symmetry, no-simultaneous encoding and phase groups for maximal measurements. We also analyze a class of alternative theories which have a restriction on the allowed effects such that they share all the above properties with quantum theory.

In section 2 we introduce the necessary framework and outline the main result showing the equivalence between alternative sets of measurement postulates and certain representations of the dynamical group. We then find these representations. In section 3 we study these alternative theories starting with a classification of theories with pure states given by rays on a 2-dimensional Hilbert space. We then show that all unrestricted non-quantum theories with pure states given by rays on a d-dimensional Hilbert space violate bit symmetry. In section 4 we discuss the results and their relation to Gleason's theorem. Proofs of all the results are in the appendix.

2 All alternatives to the Measurement Postulates

2.1 Outcome probability functions

The standard formulation of quantum theory can be divided in two parts. The first part postulates that the pure states of a system are the rays of a complex Hilbert space[1] $\psi \in \mathrm{P}\mathbb{C}^d$, and that the evolution of an isolated quantum system is unitary $\psi \mapsto U\psi$, where $U \in \mathrm{PU}(d)$ is an element of the projective unitary group[2]. The second part postulates that each measurement is

---

[1] The set of rays of a Hilbert space $\mathcal{H}$ is called the projective Hilbert space, denoted $\mathrm{P}\mathcal{H}$.

[2] $\mathrm{PU}(d)$ is defined by taking $\mathrm{SU}(d)$ and identifying the matrices that differ by a global phase $U \cong e^{i\theta}U$. For example, the two matrices $\pm\mathbb{1} \in \mathrm{SU}(2)$ correspond to the same element in $\mathrm{PU}(2)$.

characterized by a list of positive semi-definite operators $(Q_1, Q_2, \ldots, Q_k)$ adding up the identity $\sum_i Q_i = \mathbb{1}$, and that the probability of outcome Q_i when the system is in state ψ is given by $P(Q_i|\psi) = \langle\psi|Q_i|\psi\rangle$ (Born's rule).

In this article we consider *all* possible alternatives to this second part. For this, we denote by $P(F|\psi)$ the probability of outcome F given the pure state ψ. And we recall that, from an operational perspective, the mathematical characterization of an outcome F is given by the probability of its occurrence for all states. Hence, each outcome F is represented by the function $F : \mathrm{P}\mathbb{C}^d \rightarrow [0,1]$ defined as $F(\psi) = P(F|\psi)$. Note that, unlike in [12], we do not require outcomes F to correspond to those of quantum mechanics (POVM elements). And even more, a priori, these outcome probability functions (OPFs) F are arbitrary; for example, they can be non-linear in $|\psi\rangle\langle\psi|$ or even discontinuous. However, we show below that, if mixed states are required to have finitely-many parameters then OPFs must be polynomial in $|\psi\rangle\langle\psi|$.

In this generalized setup, each measurement is characterized by a list of OPFs $(F_1, F_2, \ldots, F_k)$ satisfying the normalization condition $\sum_i F_i(\psi) = 1$ for all $\psi \in \mathrm{P}\mathbb{C}^d$. Therefore, each alternative to the Measurement Postulates can be characterized by the list of all "allowed" measurements in each dimension d, e.g. $\{(F_1, \ldots, F_k), (F'_1, \ldots, F'_{k'}), \ldots\}$. For each alternative, we denote the set of OPFs on $\mathrm{P}\mathbb{C}^d$ by $\mathcal{F}_d = \{F_1, F_2, \ldots, F'_1, F'_2, \ldots\}$. From an operational perspective, the set $\mathcal{F}_d$ is not completely arbitrary: the composition of a measurement and a unitary generates another measurement. That is, if $F \in \mathcal{F}_d$ and $U \in \mathrm{PU}(d)$ then $F \circ U \in \mathcal{F}_d$.

In this generalized setup, a *mixed state* is an equivalence class of indistinguishable ensembles. Let (p_i, ψ_i) be the ensemble where state $\psi_i \in \mathrm{P}\mathbb{C}^d$ is prepared with probability p_i. Two ensembles (p_i, ψ_i) and (p'_i, ψ'_i) are indistinguishable in $\mathcal{F}_d$ if

$$\sum_i p_i F(\psi_i) = \sum_i p'_i F(\psi'_i) \qquad (2)$$

for all $F \in \mathcal{F}_d$. This equation defines the equivalence relation with which the equivalence classes of ensembles (mixed states) are defined.

Figure 1: This diagram expresses the commutation of equation (6).

2.2 The convex representation

In what follows we introduce a different representation for pure states, unitaries and OPFs, which naturally incorporates mixed states. This representation is the one used in general probabilistic theories [13–16], which is sometimes called the convex operational framework [17–20]. Before starting let us introduce some notation. Given a real vector space V, the general linear group $\mathrm{GL}(V)$ is the set of invertible linear maps $T : V \to V$, and $\mathcal{E}(V)$ is the set of affine[3] functions $E : V \to \mathbb{R}$.

Result 1. Given a set $\mathcal{F}_d$ of OPFs for $\mathrm{P}\mathbb{C}^d$ (encoding an alternative to the measurement postulates) there is a (possibly infinite-dimensional) real vector space V and the maps

$$\Omega : \mathrm{P}\mathbb{C}^d \to V \,, \tag{3}$$
$$\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(V) \,, \tag{4}$$
$$\Lambda : \mathcal{F}_d \to \mathcal{E}(V) \,, \tag{5}$$

satisfying the following properties:

- *Preservation of dynamical structure (see Figure 1):*

$$\Gamma_U \Omega_\psi = \Omega_{U\psi} \,, \tag{6}$$
$$\Gamma_{U_1}\Gamma_{U_2} = \Gamma_{U_1 U_2} \,. \tag{7}$$

- *Preservation of probabilistic structure:*

$$\Lambda_F(\Omega_\psi) = F(\psi) \,. \tag{8}$$

[3]A function $E : V \to \mathbb{R}$ is affine if it satisfies $E\left(x\omega_1 + (1-x)\omega_2\right) = xE(\omega_1) + (1-x)E(\omega_2)$ for all $x \in \mathbb{R}$ and $\omega_1, \omega_2 \in V$. When V is finite-dimensional, each $E \in \mathcal{E}(V)$ can be written in terms of a scalar product as $E(\omega) = \mathbf{e} \cdot \omega + c$, where $\mathbf{e} \in V$ and $c \in \mathbb{R}$.

- *Minimality of V[4]:*

$$\mathrm{Aff}\left(\Omega_{\mathrm{P}\mathbb{C}^d}\right) = V \,. \tag{9}$$

- *Uniqueness:* for any other maps $\Omega', \Gamma', \Lambda'$ satisfying all of the above, there is an invertible linear map $L : V \to V$ such that

$$\Omega'_\psi = L(\Omega_\psi) \,, \tag{10}$$
$$\Gamma'_U = L\Gamma_U L^{-1}, \tag{11}$$
$$\Lambda'_F = \Lambda_F L^{-1}. \tag{12}$$

In this new representation, the affinity of the OPFs Λ_F implies the following. All the predictions for an ensemble (p_i, ψ_i) can be computed from the vector $\omega = \sum_i p_i \Omega_{\psi_i} \in V$. In more detail, if a source prepares state ψ_i with probability p_i then the probability of outcome F is

$$\sum_i p_i P(F|\psi_i) = \sum_i p_i \Lambda_F(\Omega_{\psi_i}) = \Lambda_F(\omega) \,.$$

Hence, the "mixed state" ω contains all the physical information of the ensemble, in the same sense as the density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ does in quantum theory. More precisely, ω is the analogue of the traceless part of the density matrix $\rho - \mathbb{1}\frac{\mathrm{tr}\rho}{d}$, which in the $d = 2$ case is the Bloch vector. So V is the analogue of the space of traceless Hermitian matrices or Bloch vectors, not the full space of Hermitian matrices. This is why probabilities are affine instead of linear functions. For example, in this representation, the maximally mixed state is always the zero vector $\int d\psi\, \Omega_\psi = 0 \in V$. Also, the linearity of the action $\Gamma_U : V \to V$ implies that, transforming the ensemble as $(p_i, \psi_i) \to (p_i, U\psi_i)$, is equivalent to transforming the mixed state as

$$\sum_i p_i \Omega_{U\psi_i} = \sum_i p_i \Gamma_U \Omega_{\psi_i} = \Gamma_U \omega \,.$$

When we include all these mixed states the (full) state space $\mathcal{S}$ becomes convex:

$$\mathcal{S} = \mathrm{conv}\, \Omega_{\mathrm{P}\mathbb{C}^d} \subset V \,. \tag{13}$$

An *effect* on $\mathcal{S}$ is an affine function $E : \mathcal{S} \to [0,1]$. In the representation introduced in Result 1, OPFs Λ_F are effects. We say that a given set

[4]Here $\mathrm{Aff}(S)$ refers to the affine span of vectors in S. The affine span is defined as all linear combinations of vectors in S where the coefficients add to 1.

of OPFs $\mathcal{F}_d$ is *unrestricted* [21] if it contains all effects of the corresponding state space $\mathcal{S}$. That is, for every effect E on $\mathcal{S}$ there is an OPF $F \in \mathcal{F}_d$ such that $\Lambda_F = E$. When $\mathcal{F}_d$ is unrestricted the map Λ is redundant, and the theory is characterized by Ω and Γ, which provide the state space $\mathcal{S}$. When $\mathcal{F}_d$ is not unrestricted, the theory can be understood as an unrestricted one with an extra restriction on the allowed effects. Hence, unrestricted theories play a special role.

The dimension of $\mathcal{S}$ (equal to that of V) corresponds to the number of parameters that are necessary to specify a general mixed state. In quantum theory this number equal to $d^2 - 1$. Despite the finite dimensionality of $\mathrm{P}\mathbb{C}^d$ the state space $\mathcal{S} = \mathrm{conv}\,\Omega_{\mathrm{P}\mathbb{C}^d}$ corresponding to a general $\mathcal{F}_d$, can be infinite-dimensional. When this is the case, general state estimation requires collecting an infinite amount of experimental data, which makes this task impossible. Recall that, when performing state estimation in infinite-dimensional quantum mechanics, additional assumptions are required; like, for example, an upper bound on the energy of the state. For this reason, in the rest of the paper, we assume that when the set of pure states $\mathrm{P}\mathbb{C}^d$ is finite-dimensional then so is the set of mixed states $\mathcal{S}$.

Let us also assume that any mathematical transformation $T : \mathcal{S} \to \mathcal{S}$ which can be arbitrarily-well approximated by physical transformations $\Gamma_{\mathrm{PU}(d)}$ should be considered a physical transformation $T \in \Gamma_{\mathrm{PU}(d)}$. This means that the group of matrices $\Gamma_{\mathrm{PU}(d)}$ is topologically closed. Together with the homomorphism property (7) and the finite-dimensionality of V this implies that $\Gamma : \mathrm{PU}(d) \to GL(V)$ is a *continuous* group homomorphism (this is shown in appendix A.1.1). In other words, Γ is a representation of the Lie group $\mathrm{PU}(d)$.

2.3 Classification results

In quantum mechanics all pure states ψ can be related to a fixed state ψ_0 via a unitary $\psi = U\psi_0$. This and equation (6) imply that, given the representation Γ and the image of the fixed pure state $\psi_0 \to \Omega_{\psi_0}$, we can obtain the image of any pure state as

$$\Omega_\psi = \Omega_{U\psi_0} = \Gamma_U \Omega_{\psi_0} . \qquad (14)$$

Now note that, for any OPF F, we can write

$$F(\psi) = F(U\psi_0) = \Lambda_F(\Gamma_U \Omega_{\psi_0}) . \qquad (15)$$

This, the affinity of Λ_F and the continuity of Γ imply that $F : \mathrm{P}\mathbb{C}^d \to [0,1]$ is continuous.

We also see that, an unrestricted theory is characterized by a representation Γ and a reference vector $\Omega_{\psi_0} \in V$. But not all pairs $(\Gamma, \Omega_{\psi_0})$ satisfy (6). Below we prove that only a small class of representations Γ allow for (6), and for each of these, there is a unique (in the sense of (10)-(12)) vector Ω_{ψ_0} satisfying (6).

Any representation of $\mathrm{PU}(d)$ is also a representation of $\mathrm{SU}(d)$, and these are all well classified in the finite-dimensional case. We now consider a family of irreducible representations of $\mathrm{SU}(d)$ which we call $\mathcal{D}_j^d$. There are many other representations of $\mathrm{SU}(d)$ but these are not compatible with the dynamical structure of quantum theory (6). For any positive integer j, let $\mathcal{D}_j^d : \mathrm{SU}(d) \to \mathrm{GL}(\mathbb{R}^{D_j^d})$ be the highest-dimensional irreducible representation inside the reducible one $\mathrm{Sym}^j U \otimes \mathrm{Sym}^j U^*$, where $\mathrm{Sym}^j U$ is the projection of $U^{\otimes j}$ into the symmetric subspace [22, Appendix 2]. Here U is the fundamental representation of $\mathrm{SU}(d)$ acting on $\mathbb{C}^d$. Note that any global phase $e^{i\theta}U$ disappears in the product $\mathrm{Sym}^j U \otimes \mathrm{Sym}^j U^*$, hence $\mathcal{D}_j^d$ is also an irreducible representation of $\mathrm{PU}(d)$. We recall that these irreducible representations are real, in the sense that there exists a basis in which all matrix elements of $\mathcal{D}_j^d(U)$ are real, for all U. The dimension D_j^d of the real vector space acted upon by $\mathcal{D}_j^d(U)$ is

$$D_j^d = \left(\frac{2j}{d-1} + 1 \right) \prod_{k=1}^{d-2} \left(1 + \frac{j}{k} \right)^2 , \qquad (16)$$

(see [22, p.224]).

Note that quantum theory corresponds to $j = 1$. In figure 2 the weight diagrams of the quantum ($\mathcal{D}_1^3$) and lowest dimensional non-quantum ($\mathcal{D}_2^3$) representations of $\mathfrak{su}(3)$ (the Lie algebra of $\mathrm{SU}(3)$) are shown. For $d = 2$, $\mathcal{D}_j^2$ are the $\mathrm{SU}(2)$ irreducible representations with integer spin j, which are also irreducible representations of $\mathrm{SO}(3) \cong \mathrm{PU}(2)$. For $d \geq 3$, these irreducible representations are also denoted with the Dynkin label $(j, \underbrace{0, \ldots, 0}_{d-3}, j)$.

Result 2. Let $\omega_j^d \in \mathbb{R}^{D_j^d}$ be the unique (up to proportionality) invariant vector $\mathcal{D}_j^d(U)\omega_j^d = \omega_j^d$

for all elements of the subgroup

$$U = \begin{pmatrix} e^{i\alpha} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & e^{-i\alpha/(d-1)}u & \\ 0 & & & \end{pmatrix}, \quad u \in \mathrm{SU}(d-1) .$$
(17)

Each finite-dimensional representation Ω : $\mathrm{P}\mathbb{C}^d \to \mathbb{R}^n$ and $\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(\mathbb{R}^n)$ satisfying (6), (7), (9) is of the form

$$\Gamma_U = \bigoplus_{j \in \mathcal{J}} \mathcal{D}_j^d(U) ,$$
(18)

$$\Omega_{\psi_0} = \bigoplus_{j \in \mathcal{J}} \omega_j^d ,$$
(19)

where $\mathcal{J}$ is any finite set of positive integers.

Each set $\mathcal{J}$ corresponds to an unrestricted theory, and clearly the dimension of the state space $n = \sum_{j \in \mathcal{J}} D_j^d$. Quantum theory corresponds to $\mathcal{J} = \{1\}$ and $n = d^2 - 1$. Now, let us analyze a simple example: quantum theory with $d = 2$. In this case, the subgroup $\mathrm{SU}(d-1) = \{\mathbb{1}\}$ is trivial, and the subgroup (17) is e^{Zt} where

$$Z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$
(20)

The action of this subgroup on the Bloch vector has two invariant states $(0, 0, \pm 1)$. But, as the result says, both are the same vector up to a proportionality factor. Also note that the two state spaces generated by the two reference vectors $\omega_1^2 = (0, 0, \pm 1)$ are related as in (10)-(12).

A consequence of Result 2 is that all finite-dimensional group actions Γ are polynomials of the fundamental action U. Alternative (1) is not polynomial, and hence needs an infinite-dimensional representation V. This shows our previous claim that an infinite number of parameters is necessary to specify a mixed state in this alternative theory.

A desirable feature of any alternative to the measurement postulates $\mathcal{F}_d$ is *faithfulness*: for any pair pure states $\psi, \psi' \in \mathrm{P}\mathbb{C}^d$ there is a measurement $F \in \mathcal{F}_d$ which distinguish them $F(\psi) \neq F(\psi')$. When this does not happen, the two states ψ, ψ' become operationally equivalent. Faithfulness translates to the injectivity of the map $\Omega : \mathrm{P}\mathbb{C}^d \to V$. The following result tells us which of the representations (18)-(19) are faithful.

Result 3 (Faithfulness). When $d \geq 3$ the map Ω is always injective. When $d = 2$ the map Ω is injective if and only if $\mathcal{J}$ contains at least one odd number.

In this work we are specially interested in faithful state spaces. Because if two pure states in $\mathrm{P}\mathbb{C}^d$ are indistinguishable, then, from an operational point of view, they become the same state.

2.4 A simpler characterization

In this section we present a different representation for all these alternative theories that makes them look closer to quantum theory. For this, we have to relax the "minimality of V" property (given by equation (9)).

For any $\psi \in \mathrm{P}\mathbb{C}^d$ and $U \in \mathrm{PU}(d)$ let

$$\Omega_\psi = |\psi\rangle\langle\psi|^{\otimes N} ,$$
(21)

$$\Gamma_U \omega = U^{\otimes N} \omega \, U^{\dagger \otimes N} ,$$
(22)

where N is a positive integer and ω is a Hermitian matrix acting on the symmetric subspace of $(\mathbb{C}^d)^{\otimes N}$. It can be seen that this representation decomposes as

$$\Gamma_U = \bigoplus_{j=0}^{N} P_{N,j}^d \otimes \mathcal{D}_j^d(U) ,$$
(23)

where $P_{N,j}^d$ is a projector whose rank depends on N, j, d. The rank of projector $P_{N,j}^d$ counts the number of copies of representation $\mathcal{D}_j^d$. For example $\mathrm{rank} P_{N,0}^d = 1$ and $\mathrm{rank} P_{N,N}^d = 1$. Also, we define $\mathcal{D}_0^d$ to be the trivial irreducible representation.

The fact that representation (21)-(22) contains the trivial irreducible representation, allows for effects to be linear instead of affine function. Hence, all effects can be written as

$$P(E|\psi) = \mathrm{tr}(E|\psi\rangle\langle\psi|^{\otimes N}) ,$$
(24)

where E is a Hermitian matrix. However, this matrix E does not need to be positive semidefinite. For example, any N-party entanglement witness W has negative eigenvalues but gives a positive value $\mathrm{tr}(W|\psi\rangle\langle\psi|^{\otimes N}) \geq 0$ for all ψ. An example of this is given in Section 3.1.

If we want to use notation (21)-(22) to represent the alternative theory characterized by $\mathcal{J}$, then we have to set $N \geq \max \mathcal{J}$ and constrain
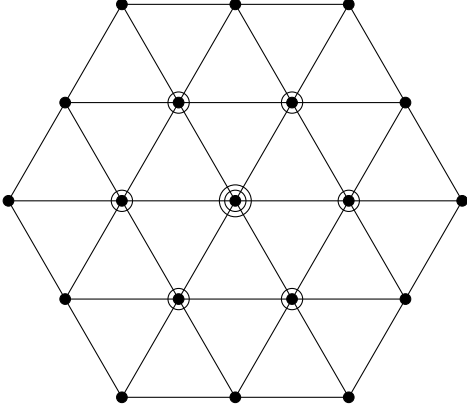
(a)
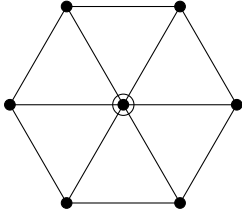
(b)

Figure 2: (a) The weight diagram of the irreducible representation of $\mathfrak{su}(3)$ with $j = 1$. This is the adjoint representation and corresponds to quantum theory. (b) The weight diagram of the irreducible representation of $\mathfrak{su}(3)$ with $j = 2$. This corresponds to the simplest non-quantum state space for $\mathrm{P}\mathbb{C}^3$.

the allowed effects to only have support on the subspaces of (23) with $j \in \mathcal{J}$. The fact that for some $j \in \mathcal{J}$ the representations $\mathcal{D}_j^d$ may be repeated is irrelevant. We can restrict the effects to act on a single copy or not. In the second case the action of the effect becomes the average of its action in each copy.

3   Phenomenology of the alternative theories

In this section we explore the properties of the alternative theories classified in the previous section. We first consider the case $d = 2$ for both irreducible and reducible representations, and later we generalize to $d \geq 3$. Before considering these families of theories we provide an example of a specific theory of $\mathrm{P}\mathbb{C}^2$ which differs significantly from quantum theory.

3.1   Three distinguishable states in $\mathbb{C}^2$

Result 4. The $d = 2$ state space $\mathcal{J} = \{1, 2\}$ has at least three perfectly distinguishable states when all effects are allowed.

(Result independently obtained in [23].) The irreducible representations $\mathcal{D}_j^2$ are given by the symmetric product $\mathcal{D}_j^2(U) = \mathrm{Sym}^{2j}U$ [22, p.150], where U is the fundamental representation on $\mathbb{C}^2$. According to Result 3, these are faithful when j is odd (Ω injective). To prove Result 4 we first show that the unfaithful $\mathrm{P}\mathbb{C}^2$ theory $\mathcal{J}' = \{2\}$ corresponds to a quantum $\mathrm{P}\mathbb{C}^3$ system restricted to the real numbers (i.e. $\mathrm{P}\mathbb{R}^3$), and hence has 3 distinguishable states. By including the trivial representation, the representation $\mathcal{J}'$ can be expressed as: $\mathrm{Sym}^4U \oplus \mathrm{Sym}^0U = \mathrm{Sym}^2(\mathrm{Sym}^2U)$ [22, p.152]. Using the fact that Sym^2U is the quantum representation, the states generated by applying $\mathcal{J}'$ to the reference state can be expressed as the rank-one projectors

$$|\phi\rangle\langle\phi| = \mathcal{D}_1^2(U)\omega_1^2 \, [\mathcal{D}_1^2(U)\omega_1^2]^T \,, \qquad (25)$$

where $\omega_1^2 \in \mathbb{R}^3$ is a unit Bloch vector and hence $|\phi\rangle \in \mathbb{R}^3$ with $\langle\phi|\phi\rangle = 1$. This is the state space of a 3-dimensional quantum system restricted to the real numbers $\mathrm{P}\mathbb{R}^3$, and hence, has three distinguishable states.

Coming back to $\mathcal{J} = \{1, 2\}$, this theory has transformations given by the representation $\mathrm{Sym}^4U \oplus \mathrm{Sym}^2U \oplus \mathrm{Sym}^0U$ and hence it is faithful. It has at least three distinguishable states since one can make measurements with support on the blocks $\mathrm{Sym}^4U \oplus \mathrm{Sym}^0U$ to distinguish the three states. It may be the case that by making measurements with support in both blocks one could distinguish more states. We now consider families of alternative theories with dynamics SU(2) and find various properties which distinguish them from the qubit.

3.2   Irreducible $d = 2$ theories

In this section we define a class of $d = 2$ theories and explore which properties they have in common with the qubit, and which properties distinguish them from quantum theory. We denote by $\mathcal{T}_2^I$ the class of non-quantum $d = 2$ theories which are faithful, unrestricted (all effects allowed) and have irreducible Γ. According to Result 3, each

of these theories is characterized by an odd integer $j \geq 3$, such that $\mathcal{J} = \{j\}$, or equivalently $\Gamma = \mathcal{D}_j^2$.

3.2.1 Number of distinguishable states

An important property of a system is the maximal number of perfectly distinguishable states it has. This quantity determines the amount of information that can be reliably encoded in one system. For instance a classical bit has two distinguishable states, as does a qubit. Distinguishable states of a qubit are orthogonal rays in the Hilbert space, or equivalently, antipodal states on the Bloch sphere.

The following result tells us about the similarities of the theories in $\mathcal{T}_2^I$ with respect to quantum mechanics. It also applies to other theories, because it does not require the assumption of unrestricted effects.

Result 5. Any $d = 2$ theory with irreducible Γ has a maximum of two perfectly distinguishable states.

However all these theories have an important difference with quantum theory.

Result 6. All theories in $\mathcal{T}_2^I$ have pairs of non-orthogonal (in the underlying Hilbert space) rays which are perfectly distinguishable.

We also prove that two rays $\psi, \phi \in \mathrm{P}\mathbb{C}^2$ are orthogonal if and only if they are represented by antipodal states $\Omega_\psi = -\Omega_\phi$. The existence of distinguishable non-antipodal states entails that theories in $\mathcal{T}_2^I$ have exotic properties not shared by qubits. We discuss a few of these in what follows.

3.2.2 Bit symmetry

Bit symmetry, as defined in [4], is a property of theories whereby any pair of pure distinguishable states (ω_1, ω_2) can be mapped to any other pair pure of distinguishable states (ω_1', ω_2') with a reversible transformation U belonging to the dynamical group, i.e. $\Gamma_U \omega_1 = \omega_1'$ and $\Gamma_U \omega_2 = \omega_2'$. The qubit is bit symmetric since distinguishable states are orthogonal rays, and any pair of orthogonal rays can be mapped to any other pair of orthogonal rays via a unitary transformation.

Result 7. All theories in $\mathcal{T}_2^I$ violate bit symmetry.

This result follows directly from Result 6 and from the fact that there does not exist any reversible transformation which maps a pair of antipodal states (representations of orthogonal rays) to a pair of non-antipodal states (representations of non-orthogonal rays).

3.2.3 Phase invariance of measurements

We now consider the concept of phase groups of measurements following [24, 25]. The phase group T of a measurement $(E_1, ..., E_n)$ is the maximal subgroup of the transformation group of the state space $\mathcal{S}$ which leaves all outcome probabilities unchanged:

$$E_i \circ \Gamma_U = E_i \quad \forall U \in T . \qquad (26)$$

For example, in the case of the qubit and a measurement in the Z basis, the phase group associated to this measurement is $T = \{e^{Zt} : t \in \mathbb{R}\}$, where (20). We maintain the historical name "phase group", although in general, it need not be abelian.

A measurement which distinguishes the maximum number of pure states in a state space is often called a *maximal measurement* [24]. For a qubit, these are the projective measurements (since they perfectly distinguish two states) which have a phase group U(1).

Result 8. All theories in $\mathcal{T}_2^I$ have maximal measurements with trivial phase groups $T = \{\mathbb{1}\}$.

These maximal measurements with trivial phase groups are those used to distinguish non-antipodal states. We note that the antipodal states can be distinguished by measurements with U(1) phase groups (as in quantum theory). We observe that contrary to [24] we find maximal measurements with trivial phase groups in a non-classical theory. This is because unlike [24] we do not consider all allowed transformations which map states to states, but only the transformations Γ_U.

3.2.4 No simultaneous encoding

No simultaneous encoding [26] is an information-theoretic principle which states that if a system is used to perfectly encode a bit it cannot simultaneously encode any other information (similarly for a trit and higher dimensions). More precisely,

consider a communication task involving two distant parties, Alice and Bob. Similarly as in the scenario for information causality [27], suppose that Alice is given two bits $a, a' \in \{0, 1\}$, and Bob is asked to guess only one of them. He will base his guess on information sent to him by Alice, encoded in one $\mathcal{T}_2^I$ system. Alice prepares the system with no knowledge of which of the two bits, a or a', Bob will try to guess. No simultaneous encoding imposes that, in a coding/decoding strategy in which Bob can guess a with probability one, he knows nothing about a'. That is, if b, b' are Bob's guesses for a, a' then

$$P(b|a, a') = \delta_b^a \Rightarrow P(b'|a, a' = 0) = P(b'|a, a' = 1)$$

where δ_b^a is the Kronecker tensor.

As an example consider a qubit. Alice decides to perfectly encode bit a, which she can only do by encoding $a = 0$ and $a = 1$ in two perfectly distinguishable states. Without loss of generality she can choose to encode $a = 0$ in $|0\rangle$ and $a = 1$ in $|1\rangle$, with $\langle 0|1 \rangle = 0$. She now also needs to encode a' whilst keeping a perfectly encoded. Since, $|0\rangle$ is the only state which is perfectly distinguishable from $|1\rangle$, we have that both $a = 0, a' = 0$ and $a = 0, a' = 1$ combinations must be assigned to $|0\rangle$. Similarly $a = 1, a' = 0$ and $a = 1, a' = 1$ combinations must be assigned to $|1\rangle$. In this case we see that whilst Bob can perfectly guess the value of a if he chooses to, if he chooses to guess the value of a' he cannot do so. Hence this property is met by qubits, however

Result 9. All theories in $\mathcal{T}_2^I$ violate no simultaneous encoding.

We see that the above three properties: bit symmetry, existence of phase groups for maximal measurements and no-simultaneous encoding single out the qubit amongst all $\mathcal{T}_2^I$ theories.

3.2.5 Restriction of effects

The study of theories in $\mathcal{T}_2^I$ has shown that they differ from quantum theory in many ways. We now consider a new family of theories $\tilde{\mathcal{T}}_2^I$, which is constructed by restricting the effects of the theories in $\mathcal{T}_2^I$. These theories turn out to be closer to quantum theory in that they obey all the above properties. This approach is similar to the self-dualization procedure outlined in [21] (which also recovers bit symmetry).

	Qubit	$\mathcal{T}_2^I$	$\tilde{\mathcal{T}}_2^I$	$\mathcal{T}_2^R$
Distinguishable states	2	2	2	≥ 2
Bit symmetry	✓	X	✓	X
Phase invariant group	✓	X	✓	?
No simultaneous encoding	✓	X	✓	?

Figure 3: Summary of results for $d = 2$ theories.

We call a theory *pure-state dual* if the only allowed effects are "proportional" to pure states. That is, for every allowed effect, there is a pure state ψ and a pair of normalization constants α, β such that

$$E(\omega) = \alpha(\Omega_\psi^T \cdot \omega) + \beta \ . \tag{27}$$

All theories in $\tilde{\mathcal{T}}_2^I$ have a maximum of two distinguishable states, and all pairs of distinguishable states are antipodal.

Result 10. All theories in $\tilde{\mathcal{T}}_2^I$ are bit symmetric, have phase groups for all maximal measurements and obey no-simultaneous encoding.

3.3 Reducible $d = 2$ theories

We consider the set of all unrestricted faithful non-quantum state spaces generated by reducible representations of SU(2). These representations are given by equation (18) with $d = 2$ and $|\mathcal{J}| > 1$ containing at least one odd number. We denote the set of all these theories $\mathcal{T}_2^R$.

For theories in $\mathcal{T}_2^R$ the number of distinguishable states can be more than two (as shown for $\mathcal{J} = \{1, 2\}$ in Section 3.1). It is in general a difficult task to find the number of distinguishable states. However we find:

Result 11. All theories in $\mathcal{T}_2^R$ violate bit symmetry.

That is, there are pairs of distinguishable pure states (ψ_1, ψ_2) which cannot be mapped to another pair of distinguishable states (ψ_1', ψ_2') with a reversible transformation U belonging to the dynamical group. This implies that

Bit symmetry singles out quantum theory amongst all $d = 2$ unrestricted faithful state spaces (both irreducible and reducible).

3.4 Arbitrary d

Having studied in detail different families of $d = 2$ theories, we now consider properties of arbitrary-d unrestricted theories. We also show that the property of bit symmetry singles out quantum theory in all dimensions. This is proven by showing that all faithful P$\mathbb{C}^d$ state spaces have embedded within them faithful (reducible) P$\mathbb{C}^{d-1}$ state spaces. We then show that if a state space P$\mathbb{C}^{d-1}$ violates bit symmetry, then any state space P$\mathbb{C}^d$ it is embedded in also does. Using a proof by induction we find:

Result 12. All unrestricted non-quantum theories with pure states P$\mathbb{C}^d$ and transformations $\mathcal{D}_j^d$ are not bit symmetric.

This result shows that the Born rule can be singled out amongst all possible probability assignments from the following set of assumptions.

The Born rule is the unique probability assignment satisfying: (i) no restriction on the allowed effects, and (ii) bit symmetry.

Although many works in GPT's only consider unrestricted theories no-restriction remains an assumption of convenience with a less direct physical or operational meaning. Bit symmetry however is a property which has computational and physical significance as discussed in [4]. It is generally linked to the possibility of reversible computation, since bit symmetric theories allow any logical bit (pair of distinguishable states) of the theory to be reversibly transformed into any other logical bit.

4 Conclusion

In this work we have considered theories with the same structure and dynamics for pure states as quantum theory but different sets of OPF's (not necessarily following the Born rule). These alternative theories were shown to be in correspondence with certain class of representations of the unitary group. Moreover these theories (assuming no restriction on the effects) were shown to differ from quantum theory in that they are not

bit symmetric. In the case of $d = 2$, the number of distinguishable states was bounded for irreducible representations, and further properties which single out quantum theory were found.

In order to fully develop these alternative theories it is necessary to determine how the state spaces studied compose. This would then allow us to study important properties related to multipartite systems.

Gleason's theorem [28] can be understood as a derivation of the Born rule. Gleason assumes that measurements are associated with orthogonal bases and shows that states must be given by density matrices and measurement outcome probabilities by the Born rule. This is a different approach to the one presented in this paper. In this work we show that, assuming that pure states are rays on a Hilbert space and time evolution is unitary, the Born rule can be derived (or singled out) from the requirement of bit symmetry. We allow for measurements to take any form, not just being associated with orthogonal bases in the corresponding Hilbert space. Unlike Gleason's theorem this applies in the case $d = 2$ and may arguably provide a more physical justification for the Born rule. There exists a derivation of Gleason's theorem for POVM's [29] which is closer in spirit to our approach than the original. However we emphasise that in our work we begin from the structure of pure states and dynamics to derive that of measurements. In Gleason's theorem (and its extension) it is the structure of measurements which are assumed and that of states derived from it.

Future work will be focused on multipartite systems and post-measurement state update rules for the alternative theories analysed in this work. This will allow us to consider more general information processing tasks for these novel probabilistic theories.

5 Acknowledgements

References

[1] David Deutsch. Quantum theory of probability and decisions. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences.* 455, 3129-3137. 1999.

[2] Wojciech Hubert Zurek. Probabilities from entanglement, born's rule $p_k = \mid \psi_k \mid^2$ from envariance. *Phys. Rev. A.* 71, 052105. May 2005.

[3] David Wallace. How to Prove the Born Rule. In Adrian Kent Simon Saunders, Jon Barrett and David Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality.* Oxford University Press. 2010.

[4] Markus P. Müller and Cozmin Ududec. Structure of reversible computation determines the self-duality of quantum theory. *Phys. Rev. Lett..* 108, 130401. Mar 2012.

[5] David Gross, Markus Müller, Roger Colbeck, and Oscar C. O. Dahlsten. All reversible dynamics in maximally nonlocal theories are trivial. *Phys. Rev. Lett..* 104, 080402. Feb 2010.

[6] Sabri W Al-Safi and Anthony J Short. Reversible dynamics in strongly non-local box-world systems. *Journal of Physics A: Mathematical and Theoretical.* 47, 325303. 2014.

[7] Sabri W Al-Safi and Jonathan Richens. Reversibility and the structure of the local state space. *New Journal of Physics.* 17, 123001. 2015.

[8] L. Hardy and W. K. Wootters. Limited Holism and Real-Vector-Space Quantum Theory. *Foundations of Physics.* 42, 454-473. March 2012.

[9] D. Finkelstein, J.M. Jauch, and D. Speiser. *Notes on Quaternion Quantum Mechanics.* CERN publications: European Organization for Nuclear Research. CERN. 1959.

[10] H. Barnum, M. Graydon, and A. Wilce. Composites and Categories of Euclidean Jordan Algebras. *eprint ArXiv:1606.09331 quant-ph.* June 2016.

[11] Scott Aaronson. Is Quantum Mechanics An Island In Theoryspace? *eprint arXiv:quant-ph/0401062.* 2004.

[12] Y. D. Han and T. Choi. Quantum Probability assignment limited by relativistic causality. *Sci. Rep..* 6, 22986. July 2016.

[13] L. Hardy. Quantum theory from five reasonable axioms. *eprint arXiv:quant-ph/0101012.* January 2001.

[14] Jonathan Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A.* 75, 032304. March 2007.

[15] Lluís Masanes and Markus P. Muller. A derivation of quantum theory from physical requirements. *New Journal of Physics.* 13, 063001. 2011.

[16] C. Brukner B. Dakic. Quantum Theory and Beyond: Is Entanglement Special? In H. Halvorson, editor, *Deep Beauty - Understanding the Quantum World through Mathematical Innovation.* pages 365–392. Cambridge University Press. 2011.

[17] Robin Giles. Foundations for Quantum Mechanics. *Journal of Mathematical Physics.* 11, 2139. 1970.

[18] Bogdan Mielnik. Generalized quantum mechanics. *Communications in Mathematical Physics.* 37, 221–256. 1974.

[19] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Probabilistic theories with purification. *Physical Review A.* 81, 062348. jun 2010.

[20] H. Barnum, J. Barrett, L. Orloff Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke. Entropy and information causality in general probabilistic theories. *New Journal of Physics.* 12, 033024. mar 2010.

[21] Peter Janotta and Raymond Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Physical Review A.* 87, 052131. 2013.

[22] William Fulton and Joe Harris. *Representation theory : a first course.* Graduate texts in mathematics. Springer-Verlag. New York, Berlin, Paris. 1991.

[23] Jonathan Barrett. private communication.

[24] Andrew J P Garner, Oscar C O Dahlsten, Yoshifumi Nakata, Mio Murao, and Vlatko Vedral. A framework for phase and interference in generalized probabilistic theories. *New Journal of Physics.* 15, 093044. 2013.

[25] Ciarán M Lee and John H Selby. Generalised phase kick-back: the structure of computational algorithms from physical principles. *New Journal of Physics.* 18, 033023. 2016.

[26] L. Masanes, M. P. Muller, R. Augusiak, and D. Perez-Garcia. Existence of an information unit as a postulate of quantum theory. *Proceedings of the National Academy of Sciences.* 110, 16373-16377. 2013.

[27] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski. Information causality as a physical principle. *nature.* 461, 1101-1104. October 2009.

[28] Andrew M. Gleason. *Measures on the Closed Subspaces of a Hilbert Space.* pages 123–133. Springer Netherlands. Dordrecht. 1975.

[29] Carlton M Caves, Christopher A Fuchs, Kiran K Manne, and Joseph M Renes. Gleason-type derivations of the quantum probability rule for generalized measurements. *Foundations of Physics.* 34, 193–209. 2004.

[30] Karl Heinrich. Hofmann and Sidney A. Morris. *The structure of compact groups : a primer for students, a handbook for the expert.* Berlin, Boston: De Gruyter. third edition, revised and augmented edition. 2013.

[31] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States.* Cambridge University Press. 2006. Cambridge Books Online.

[32] M. L. Whippman. Branching Rules for Simple Lie Groups. *Journal of Mathematical Physics.* 6, 1534-1539. October 1965.

[33] W.J. Holman and L.C. Biedenharm. The Representations and Tensor Operators of the Unitary Groups U(n). In Ernest M. Loebl, editor, *Group Theory and its Applications.* pages 1 – 73. Academic Press. 1971.

[34] Roe Goodman and Nolan R. Wallachs. *Symmetry, Representations, and Invariants.* Springer-Verlag New York. New York. 2009.

[35] Fernando Antoneli, Michael Forger, and Paola Gaviria. Maximal subgroups of compact Lie groups. *eprint ArXiv:math/0605784.* 2006.

A Proofs

A.1 The convex representation (proof of Result 1)

Let us fix the value of d and the set of OPFs $\mathcal{F}_d$. Also, we assume that all OPFs of the form $F \circ U$ are contained in $\mathcal{F}_d$, for all $U \in \mathrm{PU}(d)$ and $F \in \mathcal{F}_d$. And recall that, if this is not the case, we can always re-define $\mathcal{F}_d$ to include them. Before defining the representation $(\Omega, \Gamma, \Lambda)$ of Result 1, it is convenient to define a temporary one $(\tilde{\Omega}, \tilde{\Gamma}, \tilde{\Lambda})$ in which states $\tilde{\Omega}_\psi$ are represented by a list of outcome probabilities. For this, we take a minimal subset $\{F^{(1)}, F^{(2)}, F^{(3)}, \ldots\} \subseteq \mathcal{F}_d$ that affinely generates $\mathcal{F}_d$. That is, for any $F \in \mathcal{F}_d$ there are constants $c^{(0)}, c^{(1)}, c^{(2)}, \ldots \in \mathbb{R}$ such that $F = c^{(0)} + \sum_r c^{(r)} F^{(r)}$. This construction requires the minimal generating subset $\{F^{(1)}, F^{(2)}, \ldots\}$ to be countable, but it can still be infinite. This does not pose any loss of generality, because, from an operational point of view, the characterization of arbitrary states would be impossible.

If we represent the pure state ψ by the list of probabilities

$$\tilde{\Omega}_\psi = \begin{pmatrix} P(F^{(1)}|\psi) \\ P(F^{(2)}|\psi) \\ P(F^{(3)}|\psi) \\ \vdots \end{pmatrix} = \begin{pmatrix} F^{(1)}(\psi) \\ F^{(2)}(\psi) \\ F^{(3)}(\psi) \\ \vdots \end{pmatrix} \in V ,$$

(28)

then we can use the rules of probability to extend this definition to mixed states. For example, the probability of outcome F when the system is prepared in the ensemble (p_i, ψ_i) is

$$P(F|(p_i, \psi_i)_i) = \sum_i p_i P(F|\psi_i) = \sum_i p_i F(\psi_i) .$$

(29)

Therefore, the probabilistic representation of ensemble (p_i, ψ_i) is $\sum_i p_i \tilde{\Omega}_{\psi_i}$. And the set of all states, pure and mixed, is

$$\tilde{\mathcal{S}} = \mathrm{conv}\, \tilde{\Omega}_{\mathbb{PC}^d} .$$

(30)

In the representation $\tilde{\Omega}$, any probability $F(\psi)$ is either a component of the vector $\tilde{\Omega}_\psi$ or an affine function of its components. Hence, for any F, there is an affine function $\tilde{\Lambda}_F : V \to \mathbb{R}$ such that

$$\tilde{\Lambda}_F(\tilde{\Omega}_\psi) = F(\psi) \quad \forall \psi .$$

(31)

Any U can be seen as a relabeling of the pure states $\mathbb{PC}^d$, hence, $\{(F^{(1)} \circ U), (F^{(2)} \circ U), \ldots\} \subseteq$

$\mathcal{F}_d$ is also a minimal generating subset. There-fore, there is an affine map $\tilde{\Gamma}_U : V \to V$ such that

$$\tilde{\Omega}_{U\psi} = \tilde{\Gamma}_U(\tilde{\Omega}_\psi) \quad \forall U, \psi . \tag{32}$$

Next, we define a new representation Ω, Γ, Λ in which $\Gamma_U : V \to V$ is linear. This new representation is that of Result 1. For this, we need to define the maximally mixed state

$$\tilde{\omega}_{\mathrm{mm}} = \int_{\mathrm{PU}(d)} dU \, \tilde{\Gamma}_U(\tilde{\Omega}_\psi) , \tag{33}$$

where dU is the Haar measure and ψ is any pure state. We also note that the maximally mixed state is invariant under any unitary: $\tilde{\Gamma}_U(\tilde{\omega}_{\mathrm{mm}}) = \tilde{\omega}_{\mathrm{mm}}$ for any U. Now, we define the new representation as

$$\Omega_\psi = \tilde{\Omega}_\psi - \tilde{\omega}_{\mathrm{mm}} , \tag{34}$$

$$\Gamma_U(\omega) = \tilde{\Gamma}_U(\tilde{\omega}) - \tilde{\omega}_{\mathrm{mm}} , \tag{35}$$

$$\Lambda_F(\omega) = \tilde{\Lambda}_F(\tilde{\omega}) , \tag{36}$$

which extends to general mixed states as $\omega = \tilde{\omega} - \tilde{\omega}_{\mathrm{mm}}$. Using (31) and (32) we obtain (6), (7). In this representation, the maximally mixed state is the zero vector $\omega_{\mathrm{mm}} = 0 \in V$. Now, recalling that ω_{mm} is invariant under unitaries, we have that $\Gamma_U(0) = 0$, which together with the affinity of Γ_U implies that $\Gamma_U : V \to V$ is linear, for all U.

To continue with the proof of Result 1, we note that the affinity of $\tilde{\Lambda}_F$ implies that of Λ_F; and that (30) implies (9).

Now, it only remains to prove uniqueness. Let us suppose that there are other maps $\Omega', \Gamma', \Lambda'$ with the properties stated in Result 1. These properties imply the existence of an affine function $\tilde{L}$ such that

$$\tilde{\Omega}_\psi = \begin{pmatrix} F^{(1)}(\psi) \\ F^{(2)}(\psi) \\ \vdots \end{pmatrix} = \begin{pmatrix} \Lambda'_{F^{(1)}}(\Omega'_\psi) \\ \Lambda'_{F^{(2)}}(\Omega'_\psi) \\ \vdots \end{pmatrix} = \tilde{L}(\Omega'_\psi) . \tag{37}$$

Which in turn implies the existence of another affine function L such that $\Omega_\psi = L(\Omega'_\psi)$. According to the properties stated in Result 1, both Γ_U and Γ'_U are linear; which implies that the maximally mixed state in both representations is the zero vector $\omega_{\mathrm{mm}} = \omega'_{\mathrm{mm}} = 0 \in V$. The affine map L also changes the representation of the maximally mixed state $\omega_{\mathrm{mm}} = L(\omega'_{\mathrm{mm}})$. But $0 = L(0)$ is only possible if L is not just affine, but linear.

All the results of this section are valid when V is finite- and infinite-dimensional. In the rest of the appendices, only the finite-dimensional case is considered.

A.1.1 Continuity of homomorphisms Γ

The number of parameters that define a mixed state in $\mathcal{S}$ is equal to the dimension of $V = \mathrm{conv}\,\mathcal{S}$. Hence, the requirement that general mixed states in $\mathcal{S}$ can be estimated with finite means implies that V is finite-dimensional. Now, we recall that in the probabilistic representation $\tilde{\Omega}$, the entries of states $\tilde{\omega} \in \tilde{\mathcal{S}}$ are bounded. Hence, due to (34), the same is true in $\mathcal{S}$. This implies that the absolute value of all matrix elements of the group $\Gamma_{\mathrm{PU}(d)}$ are also bounded.

Now, we argue that, from a physical point of view, the group of transformations $\Gamma_{\mathrm{PU}(d)}$ must be topologically closed. This follows from the fact that any mathematical transformation that can be approximated arbitrarily well by physical transformations should be a physical transformation too. In summary, the fact that the set $\Gamma_{\mathrm{PU}(d)} \subseteq \mathbb{R}^n$ is bounded and closed implies that it is compact. And this is the premise of the following theorem.

It is proven in Theorem 5.64 of [30, p.167] that any group homomorphism $\Gamma : \mathrm{PU}(d) \to \mathcal{G}$ in which $\mathcal{G}$ is compact must be continuous.

A.2 Preservation of dynamical structure (Result 2)

The stabilizer subgroup $\mathcal{G}_\psi \leq \mathrm{PU}(d)$ of a pure state $\psi \in \mathrm{P}\mathbb{C}^d$ is the set of unitaries U leaving the state invariant $U\psi = \psi$. If the state ψ' is related to ψ via $\psi' = U\psi$, then their corresponding stabilizer subgroups are related via $\mathcal{G}_{\psi'} = U\mathcal{G}_\psi U^\dagger$. Therefore, all stabilizers are isomorphic.

Equation (7) implies that Γ is a representation of $\mathrm{SU}(d)$. Following the premises of Result 2, this representation is finite-dimensional. Hence, we can decompose Γ into real irreducible representations

$$\Gamma = \bigoplus_r \Gamma^r . \tag{38}$$

Using the same partition into real linear subspaces, we also decompose the map

$$\Omega = \bigoplus_r \Omega^r . \tag{39}$$

Equation (6) independently holds for each summand $\Gamma_U^r \Omega_\psi^r = \Omega_{U\psi}^r$. In particular $\Gamma_U^r \Omega_\psi^r = \Omega_\psi^r$ for all $U \in \mathcal{G}_\psi$, which implies that each Γ^r has a $\mathcal{G}_\psi$-invariant subspace.

Now, let us concentrate on the $d \geq 3$ case. To figure out the structure of the stabilizer subgroup, we consider the state $\psi_0 = (1, 0, \ldots, 0)$, and note that the group $\mathcal{G}_{\psi_0}$ is the set of unitaries of the form (17). Hence, when $d \geq 3$, all stabilizers are isomorphic to $U(1) \times SU(d-1)$.

Lemma 1. The finite-dimensional irreducible representations of $SU(d)$ that have $U(1) \times SU(d-1)$-invariant vectors are the $\mathcal{D}_j^d$ introduced above. Additionally, the vector is always unique (up to a constant).

This lemma, proven in appendix A.12, tells us that most of the irreducible representations of $SU(d)$ do not have a vector which is invariant under $U(1) \times SU(d-1)$, and hence violate (6). As mentioned above, the $SU(d)$ representations $\mathcal{D}_j^d$ are also $PU(d)$ representations, which is what is required (see Result 1). A technical point (fully addressed in appendix A.12.6) arises. The representations we are looking for act on real vector spaces, however the representation theory tools used to find the $\mathcal{D}_j^d$ deal with representations acting on complex vector spaces. A representation which is irreducible on a real space is not necessarily irreducible on a complex space. We have found all complex irreducible representations (which are also real irreducible) with the correct properties, however it may be the case that there are real irreducible representations which are complex reducible which were not found with the methods used. It is shown in appendix A.12.6 that the $\mathcal{D}_j^d$ are the only real irreducible representations with the required properties.

A very convenient fact is that there is a single one-dimensional invariant subspace. Hence, the maps Ω^r are completely determined up to a factor. However, this factor does not play any role, since it can be modified via an equivalence transformation L of the type (10)-(12). If there were multiple trivial invariant subspace, one could construct different alternatives to the measurement postulates having the same representation Γ.

In the case $d = 2$, the stabilizer subgroup $\mathcal{G}_\psi$ is isomorphic to $U(1)$. This subgroup is generated by a single Lie algebra element (e.g. Z). The irreducible representations $\mathcal{D}$ in which, the

subgroup $\mathcal{G}_\psi$ has an invariant vector, are those in which $\mathcal{D}(Z)$ has at least one zero eigenvalue. These are the integer spin representations. Also, the multiplicity of this eigenvalue is always one, implying that, as in the $d \geq 3$ case, the map Ω is completely determined up to equivalence transformations of the type (10)-(12).

A.3 Faithfulness (Result 3)

Let us prove that when $d \geq 3$ the map Ω specified in Result 2 is injective. We start by assuming the opposite: there are two different pure states $\psi \neq \psi'$ which are mapped to the same vector $\Omega_\psi = \Omega_{\psi'}$. This vector must be invariant under the action of the two stabilizer subgroups: $\Gamma_U \Omega_\psi = \Omega_\psi$ for all $U \in \mathcal{G}_\psi$ and all $U \in \mathcal{G}_{\psi'}$. Now, note that if $\Gamma_U \Omega_\psi = \Gamma_{U'} \Omega_\psi = \Omega_\psi$ then also $\Gamma_{UU'} \Omega_\psi = \Omega_\psi$. Hence, the stabilizer group of the vector Ω_ψ contains the group $\{UU' | \forall U \in \mathcal{G}_\psi, U' \in \mathcal{G}_{\psi'}\}$.

Lemma 2. When $d \geq 3$ the group $SU(d)$ can be generated by two stabilizer subgroups $SU(d-1) \times U(1)$ of two distinct rays on $P\mathbb{C}^d$.

This lemma, proven in Appendix A.13, shows that the group generated by any two stabilizer subgroups is the full group $SU(d)$. Therefore, the vector Ω_ψ is invariant under any transformation Γ_U. This implies that all states ψ' are mapped to the same vector $\Omega_{\psi'} = \Omega_\psi$ and have exactly the same outcome probabilities. Equivalently, all functions $F \in \mathcal{F}_d$ are constant.

Now, let us analyze the $d = 2$ case. In what follows we prove that, the map Ω^r associated to an irreducible representation $\Gamma^r = \mathcal{D}_j^2$ is injective when j is odd. Hence, the global map Ω is injective if it contains at least one summand Ω^r that is injective. An irreducible representation of $SU(2)$ can be expressed as a symmetric power of the fundamental representation [22, p.150]:

$$\mathcal{D}_j^2 = \text{Sym}^{(2j)} \mathcal{D}_{\frac{1}{2}}^2 \ , \qquad (40)$$

where $D_{\frac{1}{2}}^2$ is the fundamental representation of $SU(2)$ acting on $\mathbb{C}^2$, with basis $\{\psi_0, \psi_1\}$. The action of the Lie algebra element Z on this basis is:

$$Z\psi_0 = i\psi_0 \ , \qquad (41)$$
$$Z\psi_1 = -i\psi_1 \ . \qquad (42)$$

Given $\mathcal{D}_j^2$, we take as reference state $\Omega_{\psi_0}^j$, the 0 eigenstate of $\mathcal{D}_j^2(Z)$ (since ψ_0 is invariant under all transformations generated by Z).

$$\mathrm{e}^{\mathcal{D}_j^2(Z)t}\Omega_{\psi_0}^j = \Omega_{\psi_0}^j \ . \tag{43}$$

The 0 eigenstate is given by $\Omega_{\psi_0}^j = \psi_0^{\otimes j} \otimes \psi_1^{\otimes j}$ (where the product is the symmetric product) [22, p.150]. All states can be obtained by applying a unitary to the reference state: $\Omega_\psi^j = \mathcal{D}_j^2(U)\Omega_{\psi_0}^j$. We call U_Z the set of transformations generated by Z. All states $\psi \neq \psi_0$ are of the form $\psi = U\psi_0$, $U \notin U_Z$.

We show that for j odd there are no states $U\psi_0$, $U \notin U_Z$ such that $\mathcal{D}_j^2(U)\Omega_{\psi_0}^j = \Omega_{\psi_0}^j$ hence ψ_0 and ψ are mapped to distinct states. For j even we show that there is a $U \notin U_Z$ such that $\Omega_\psi^j = D_j^2(U)\Omega_{\psi_0}^j = \Omega_{\psi_0}^j$ and hence the representation is not faithful. A generic $U \in \mathrm{SU}(2)$ acting on $\mathbb{C}^2$ has the following action:

$$\begin{aligned} \psi_0 &\to \alpha\psi_0 + \beta\psi_1 \ , \\ \psi_1 &\to \alpha^*\psi_1 - \beta^*\psi_0 \ , \end{aligned} \tag{44}$$

where $|\alpha|^2 + |\beta|^2 = 1$. The action $\mathcal{D}_j^2(U)$ on $\Omega_{\psi_0}^j$ is the same as that of $U^{\otimes 2j}$ since $\Omega_{\psi_0}^j$ belongs to the symmetric subspace.

$$\begin{aligned} \psi_0^{\otimes j} &\to (\alpha\psi_0 + \beta\psi_1)^{\otimes j} \ , \\ \psi_1^{\otimes j} &\to (\alpha^*\psi_1 - \beta^*\psi_0)^{\otimes j} \ . \end{aligned} \tag{45}$$

We now determine which unitaries U preserve the state:

$$\psi_0^{\otimes j}\psi_1^{\otimes j} = (\alpha\psi_0 + \beta\psi_1)^{\otimes j}(\alpha^*\psi_1 - \beta^*\psi_0)^{\otimes j} \ . \tag{46}$$

This only holds when either α or β is 0. When $\beta = 0$ this corresponds to a $U \in U_Z$. When $\alpha = 0$ we have:

$$\psi_0^{\otimes j}\psi_1^{\otimes j} = (-1)^j \psi_0^{\otimes j}\psi_1^{\otimes j} \ , \tag{47}$$

since by unitarity requirement $|\beta| = 1$. For even j there is a unitary $U \notin U_Z$, such that $\mathcal{D}_j^2(U)\Omega_{\psi_0}^j = \Omega_{\psi_0}^j$ hence the map Ω is not injective. It maps orthogonal rays to the same state. For odd j any unitary $U \notin U_Z$ maps $\Omega_{\psi_0}^j$ to a different state and so the map Ω is injective. Moreover we see that $U\psi_0 = \psi_1$ and hence orthogonal rays ψ_0 and ψ_1 are mapped to antipodal states Ω_{ψ_0} and $\Omega_{\psi_1} = -\Omega_{\psi_0}$ for odd j and to the same state $\Omega_{\psi_0} = \Omega_{\psi_1}$ for even j.

A.4 Distinguishable states in $\mathcal{T}_2^I$ (Result 5)

The proof is by contradiction: we assume the existence of 3 distinguishable states, and show that any distinguishing measurement has an outcome with a negative probability for certain states. Hence the maximum number of perfectly distinguishable states is two. Since effects are affine functions we can write an effect E as a pair $(\mathbf{e}, c)$.

$$E(\omega) = \mathbf{e} \cdot \omega + c \ . \tag{48}$$

A measurement is a set of effects $\{E_i\}$ such that $\sum E_i(\omega) = 1$ for all states ω; this entails $\sum \mathbf{e}_i = \mathbf{0}$ and $\sum c_i = 1$. If three states ω_1, ω_2 and ω_3 are distinguishable then there exists a measurement which distinguishes them of the following form: $\{E_1 = (c_1, \mathbf{e}_1), E_2 = (c_2, \mathbf{e}_2), E_3 = (c_3, \mathbf{e}_3)\}$, with $E_i(\omega_j) = \delta_{ij}$. This entails

$$\mathbf{e}_i \cdot \omega_i = 1 - c_i \ . \tag{49}$$

From Result 3 there exists an antipodal state $-\omega_i$ for every ω_i in theories in $\mathcal{T}_2^I$. We compute the outcome probabilities for these antipodal states:

$$E_i(-\omega_i) = \mathbf{e}_i \cdot (-\omega_i) + c_i = -1 + 2c_i \tag{50}$$

The sum of these three measurement outcomes probabilities is:

$$\sum_i E_i(-\omega_i) = -3 + 2(c_1 + c_2 + c_3) = -1 \tag{51}$$

Therefore at least one of the outcome probabilities is negative and therefore not legitimate. By contradiction this proves there is no measurement which perfectly distinguishes three states in these theories.

A.5 Distinguishable non-antipodal states in $\mathcal{T}_2^I$ (Result 6)

In this section we show that for each non-quantum theory in $\mathcal{T}_2^I$ there exists a measurement of the form $M = \{(\mathbf{e}, c), (-\mathbf{e}, 1 - c)\}$ with $\mathbf{e} \propto \omega_0^\dagger \cdot \mathcal{D}_j^2(X)^\dagger$ which perfectly distinguishes pairs of non antipodal states lying on the $\mathcal{D}(U_X(t))\omega_0$ orbit.

Theories in $\mathcal{T}_2^I$ are generated by representations $\mathcal{D}_j^2$ with j odd. The dimension of such representations is $n = 2j + 1$. We write $U_K(t) = e^{Kt/2}$ for an arbitrary element of the Lie Algebra K. We will use the following elements of $\mathfrak{su}(2)$:

$$Z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \ , \tag{52}$$

$$X = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \ , \qquad (53)$$

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} i & i \\ i & -i \end{pmatrix} \ , \qquad (54)$$

We note that $\mathcal{D}_j^2(e^{Kt/2}) = e^{\mathcal{D}_j^2(K)t}$. The reference state is chosen to be invariant under all $U_Z(t)$:

$$\mathcal{D}_j^2(U_Z(t))\omega_j^2 = \omega_j^2 \ , \qquad (55)$$

which implies $\mathcal{D}_j^2(Z)\omega_j^2 = 0$. The 0-weight subspace is one dimensional, the other possible (nor-

malised) state is $-\omega_j^2$. This will be the antipodal state. The manifold of pure states can be generated in the following manner:

$$\omega(s,t) = \mathcal{D}_j^2(U_Z(s))\mathcal{D}_j^2(U_X(t))\omega_j^2 \ . \qquad (56)$$

This is just a 2-sphere (embedded in a space of dimension n) parametrised by polar angle s and azimuthal angle t. The generators of the irreducible (unitary) representations of SU(2) can be written as [31, p.387]:

$$\mathcal{D}_j^2(X) = i \begin{pmatrix} 0 & \frac{\sqrt{2j}}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{\sqrt{2j}}{2} & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & \frac{\sqrt{2j}}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{2j}}{2} & 0 \end{pmatrix} \qquad (57)$$

$$\mathcal{D}_j^2(Z) = i \begin{pmatrix} j & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & j-1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -j+1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -j \end{pmatrix} \ . \qquad (58)$$

We define $k = j + 1$ where $\mathcal{D}(K)_{(k,k)}$ is the central element of the matrices $\mathcal{D}(K)$ for some Lie algebra element K. It is clear that in this basis the reference state (0 eigenstate of $\mathcal{D}_j^2(Z)$) is:

$$(\omega_j^2)_i = \delta_{ik} \ . \qquad (59)$$

In the following we drop the explicit reference to the representation, and use $\mathcal{D}(U)$ for $\mathcal{D}_j^2(U)$. We now use the notation ω_{k+i}^K for the i eigenstate of the representation matrix $\mathcal{D}(K)$ of the Lie Algebra element K. Here i runs from $-j$ to j. In the case of the diagonal operator $\mathcal{D}(Z)$ we observe that an eigenvector with eigenvalue i has a single entry at position $k + i$. The zero eigenstate $(\omega_k^Z)_m = \delta_{km}$ which is just a vector with a single entry at position k (the central entry). The 0 eigenstate of the operator $\mathcal{D}(X)$ for example

is ω_k^X. Hence we use ω_k^Z for the reference state ω_j^2 (since it is the 0 eigenstate of the Z operator). The 0 eigenstate of $\mathcal{D}(X)$ corresponds to the state ω_k^X which is equal to $\mathcal{D}(U_Y(-\frac{\pi}{2}))\omega_k^Z$ (and is invariant under $\mathcal{D}(U_X(t))$). These 0 eigenstates of Lie Algebra matrices $\mathcal{D}(K)$ correspond to states since they have the correct invariance properties, that is, invariant under $e^{i\mathcal{D}(K)t}$.

As shown in the proof of Result 3 these state spaces have antipodal states for all states. Consider an effect which has as maximum ω_m will have as minimum $-\omega_m$. This effect may be noisy (i.e. not have values 0 and 1 for the minimum and maximum) but will be proportional to an effect which is tight (gives values 0 and 1 on the minimum and maximum). Hence any tight effect which has as maximum ω_m will distinguish

it from $-\omega_m$. In the following we show that there exist effects which have 2 global maxima and minima when applied to the manifold of pure states. These effects are proportional to tight effects which distinguish ω_0 and $-\omega_0$ and also distinguish ω_0 and ω_1 (for $\omega_1 \neq -\omega_0$).

Since effects $(\mathbf{e}, c)$ are of the form $\mathbf{e} \cdot \omega(s,t) + c$, we need only consider $\mathbf{e} \cdot \omega(s,t)$ to establish the number of global extrema and their locations. We now prove the effect $(\mathbf{e}, c)$ with $\mathbf{e} = (\omega_k^Z)^\dagger \mathcal{D}(X)^\dagger$ (up to normalisation) perfectly distinguishes non-antipodal states in all representations. Here we use the $\dagger$ since the specific representation of the states have complex entries (even though the vector space spanned by the states is a real vector space). We first show that $\mathbf{e}$ gives real values when applied to the state space (whose states span a real vector space). There is a basis for the representation where all entries are real which is related to this representation by $A\mathcal{D}(U)A^{-1}$. It acts on states as $A\omega$ and effects as $\mathbf{e}A^{-1}$. Hence $(\omega_j^Z)^\dagger A^{-1} \cdot A\mathcal{D}(U)A^{-1}A\omega$ is real for all ω (since it is an inner product of vectors with real entries). This is equal to $\mathbf{e} \cdot \omega$ for all omega. Hence $\mathbf{e}$ applied to the state space is real valued.

We study the maxima and minima of $f(s,t) = \mathbf{e} \cdot \omega(s,t)$ which corresponds to applying $\mathbf{e}$ to the entire manifold of pure states:

$$f(s,t) = \mathbf{e} \cdot \mathcal{D}(U_Z(s))\mathcal{D}(U_X(t))\omega_k^Z . \quad (60)$$

We consider the first part (a row vector)

$$\mathbf{e}_s = \mathbf{e}\, e^{\mathcal{D}(Z)s} = (\omega_k^z)^\dagger \mathcal{D}(X)^\dagger e^{\mathcal{D}(Z)s} , \quad (61)$$

for which entries are 0 apart from entries $k \pm 1$:

$$(\mathbf{e}_s)_{k\pm 1} = i\frac{\sqrt{j(j+1)}}{2}\, e^{\mp is} . \quad (62)$$

We can write this as:

$$\mathbf{e}_s =$$
$$(0,...,i\frac{\sqrt{j(j+1)}}{2}e^{is}, 0, i\frac{\sqrt{j(j+1)}}{2}e^{-is}, 0,...,0) . \quad (63)$$

We now determine the second part of the function $f(s,t)$: $\mathcal{D}(U_X(t))\omega_k^Z$. Since ω_k^Z has a single entry equal to 1 at position k this just selects the kth column of $\mathcal{D}(U_X(t))$, we call these states $\omega_X(t)$; they form an orbit. The inner product

$\mathbf{e}_s \cdot \omega_X(t)$ will only have two terms corresponding to the $k-1$ and $k+1$ entries of each vector.

$$\mathbf{e}_s \cdot \omega_t =$$
$$i\frac{\sqrt{j(j+1)}}{2}(e^{is}\omega_X(t)_{k-1} + e^{-is}\omega_X(t)_{k+1}) . \quad (64)$$

This is real valued for all values of s, hence it must be the case that $\omega_X(t)_{k-1} = \omega_X(t)_{k+1}$ and both are imaginary (or both real and $\omega_X(t)_{k-1} = -\omega_X(t)_{k+1}$). However since the ω_k^Y state belongs to the orbit $\omega_X(t)$ and has the property $(\omega_k^Y)_{k-1} = (\omega_k^Y)_{k+1}$ it must be the former.

$$\mathbf{e}_s \cdot \omega_t = \sqrt{j(j+1)}\omega_X(t)_{k+1}\cos(s) . \quad (65)$$

The maxima and minima of this function occur for $s = 0$ (or π but this corresponds to the same orbit). Hence we need to find the maxima and minima of the function:

$$\begin{aligned}
f(s, t=0) = g(t) &= \mathbf{e} \cdot \mathcal{D}(U_X(t))\omega_k^Z \\
&= (\omega_k^Z)^\dagger \mathcal{D}(X)^\dagger \cdot \mathcal{D}(U_X(t))\omega_k^Z \\
&= (\omega_k^X)^\dagger \mathcal{D}(Z)^\dagger \cdot \mathcal{D}(U_Z(t))\omega_k^X .
\end{aligned} \quad (66)$$

Where we have used H the Hadamard transformation defined above which changes the basis from Z to X: $HZH^\dagger = X$. Moreover $HU_Z(t)H^\dagger = U_X(t)$. Therefore:

$$\mathcal{D}(U_X(t))\omega_k^Z = \mathcal{D}(H)\mathcal{D}(U_Z(t))\mathcal{D}(H)^\dagger\omega_k^Z , \quad (67)$$

we note that $\mathcal{D}(H)^\dagger\omega_k^Z$ is a 0 eigenstate of $\mathcal{D}(X)$ and is therefore equal to ω_k^X.

We now evaluate the second part of $g(t)$: $\mathcal{D}(U_Z(t))\omega_k^X$ where ω_k^X is a normalised zero eigenvector of $\mathcal{D}(X)$ and has the following form:

$$\omega_k^X = \begin{pmatrix} a_j \\ 0 \\ a_{(j-2)} \\ 0 \\ \vdots \\ 0 \\ -a_{(j-2)} \\ 0 \\ -a_j \end{pmatrix} , \quad (68)$$

where $|a_n| > |a_m|$ for $n > m$.

$$\mathcal{D}(e^{Zt})\omega_k^X = \begin{pmatrix} a_j e^{ijt} \\ 0 \\ a_{(j-2)}\, e^{i(j-2)t} \\ 0 \\ \vdots \\ 0 \\ -a_{(j-2)}e^{-i(j+2)t} \\ 0 \\ -a_j e^{-ijt} \end{pmatrix} . \qquad (69)$$

We can also determine $(\omega_k^X)^\dagger \mathcal{D}(Z)^\dagger$:

$$(\omega_k^X)^\dagger \mathcal{D}(Z)^\dagger =$$
$$(ja_j^*, 0, (j-2)a_{j-2}^*, 0, ..., 0, (j-2)a_{j-2}^*, 0, ja_j^*) .$$
$$(70)$$

We can now compute $g(t)$:

$$g(t) = (\omega_k^X)^\dagger \mathcal{D}(Z)^\dagger \cdot \mathcal{D}(U_Z(t))\omega_k^X$$
$$= 2 \sum_{l=1,\ l\ \text{odd}}^{j} l\ |a_l|^2 \sin(kt) . \qquad (71)$$

$|a_n| > |a_m|$ for $n > m$ hence $|a_n|^2 > |a_m|^2$ for $n > m$. We restrict ourselves to the interval $[0, 2\pi)$. We have been considering non-quantum theories with $j > 1$, however we briefly describe what this effect would correspond to for the quantum case. When $j = 1$ (i.e $k = 2$) there is a single maximum occurring for $t_m = \pi/2$. Hence the global minimum occurs for $t = 3\pi/2$. The two states distinguished by this effect are antipodal and correspond to the $\mathcal{D}(X)$ eigenstates. Indeed we observe that this effect is just given by the tangent to ω_2^Z (image of the ray $|0\rangle$) in the $\mathcal{D}(X)$ direction which is proportional to ω_2^X. We know that this effect gives the outcome probability of being in the ω_2^X state (image of the $|+\rangle$ ray) which is maximal for ω_2^X and minimal for $-\omega_2^X$ (image of the $|-\rangle$ ray).

We now show that $g(t)$ has two global maxima and two global minima for $j > 1$. $g(t) = g(\pi - t)$, therefore given a maximum/minimum we can find another (unless $t_m = \pi/2$). Moreover since $g(t) = -g(t + \pi)$ given a maximum/minimum we can find a minimum/maximum. To prove our claim we need to find a global maximum/minimum in the interval $[0, \pi)$. If this extremum occurs for a value which is not $\pi/2$ then we can find the other maximum/minimum in the same interval and the

two minima/maxima in $[\pi, 2\pi)$. We now show that for $j > 1$ the global maximum/minimum does not occur for $t_m = \pi/2$.

We first note that $g(0) = 0$, hence a global maximum is positive and a global minimum is negative. We first compute $g(\pi/2)$ and show that: $g(\pi/2) > 0 \to g(\pi/(2j)) > g(\pi/2)$. This implies that $g(\pi/2)$ cannot be a global maximum, since if it is positive there is a $\tau = \pi/(2j)$ such that $g(\tau) > g(\pi/2)$. Similarly we show that $g(\pi/2) < 0 \to g(\pi/(2j)) < g(\pi/2)$ which entails that when $g(\pi/2)$ is negative it cannot be a global minimum. Hence the global extrema of $g(t)$ do not occur for $t = \pi/2$ when $j > 1$.

$$g(\pi/2) = \sum_{l=1,l\ \text{odd}}^{j} (-1)^{\frac{l-1}{2}} l\ |a_l|^2 . \qquad (72)$$

We note that:

$$l|a_l|^2 > (l-2)|a_{l-2}|^2, \quad l\ \text{odd},\ l > 1 . \qquad (73)$$

Since each term in $g(\pi/2)$ alternates sign and the absolute value of each term increases for each l the sign of $g(\pi/2)$ is determined by that of its highest term $(-1)^{\frac{j-1}{2}} l\ |a_j|^2$. This implies $g(\pi/2) > 0$ for $j = 5, 9, 13, ...$. We consider these cases. The last term (which is the largest) $(-1)^{\frac{j-1}{2}} j\ |a_j|^2$ is always positive. We observe that the sum of the remaining terms is negative. This implies:

$$g(\pi/2) - j|a_j|^2 < 0 , \qquad (74)$$

We now determine

$$g(\frac{\pi}{2j}) = \sum_{l=1,l\ \text{odd}}^{j} l\ |a_l|^2 \sin(\frac{k\pi}{2j}) . \qquad (75)$$

The arguments in each of the sin functions in $g(\frac{\pi}{2j})$ are always between 0 and $\frac{\pi}{2}$, hence each term is positive. Therefore:

$$g(\frac{\pi}{2j}) - j|a_l|^2 > 0 . \qquad (76)$$

Combining equations 74 and 76 we obtain $g(\frac{\pi}{2j}) > g(\frac{\pi}{2})$ when $g(\frac{\pi}{2}) > 0$. This shows that the maximum of the function does not occur for $t = \pi/2$. The function therefore has at least two maxima in the interval $[0, \pi]$.

For $j = 3, 7, 11, ...$ the same argument can be applied to show that the function has two minima within the interval $[0, \pi]$ and hence two maxima within $[\pi, 2\pi]$.

This entails that there exists a global maximum which is separated from a global minimum by a value which is not π. Hence there exist states which are distinguishable but are not antipodal.

A.6 Bit symmetry (Result 7)

In theories belonging to $\mathcal{T}_2^I$ images of the orthogonal rays ψ_0 and ψ_1 are antipodal states Ω_{ψ_0} and Ω_{ψ_1} with $\Omega_{\psi_0} = -\Omega_{\psi_1}$. These states can be perfectly distinguished using the measurement $(\mathbf{e}, \frac{1}{2}), (-\mathbf{e}^T, \frac{1}{2})$ where $\mathbf{e} = \Omega_{\psi_0}^T/2$. From Result 6 there exists a state Ω_{ψ_2} which is distinguishable from Ω_{ψ_0} and not antipodal to it. Due to the faithfulness of Ω for j odd we have $\psi_2 \neq \psi_1$. Since ψ_1 is the unique ray orthogonal to ψ_0 in $\mathrm{P}\mathbb{C}^2$ ψ_2 is not orthogonal to ψ_0. There is no unitary which maps the pair of orthogonal rays (ψ_0, ψ_1) to the pair of non-orthogonal rays (ψ_0, ψ_2). Hence there exist pairs of distinguishable states which are not related by a reversible transformation belonging to the dynamical group.

A.7 Phase groups in $\mathcal{T}_2^I$ (Result 8)

We show that the maximal measurement (from the proof of Result 6) which distinguishes two non-antipodal states Ω_ψ and Ω_ϕ in theories $\mathcal{T}_2^I$ cannot have a phase group. Without loss of generality we choose ψ to be invariant under transformations $U_Z(t)$. The phase group of a measurement is a subgroup of the transformation group which leaves all outcome probabilities of the measurement unchanged. This measurement has exactly two maxima and two minima on the state space (as shown in the proof of Result 6). We call E one of the effects which compose this measurement:

$$E(\Omega_\psi) = E(-\Omega_\phi) = 1 ,$$
$$E(\Omega_\phi) = E(-\Omega_\psi) = 0 .$$

The states Ω_ψ and Ω_ϕ are the images of non-orthogonal rays on $\mathbb{C}^2$. We now show that there is no subgroup of the transformation group which leaves the outcome probabilities of the 4 states $\Omega_\psi, \Omega_\phi, -\Omega_\psi$ and $-\Omega_\phi$ invariant (which are the images of $\psi, \phi, \psi^\perp$ and $\phi^\perp$). This is sufficient that there is no phase group for this measurement. Such a phase group must either preserve the states, or map them to states which give the same outcome probabilities. For example it must either preserve ψ or map it to $\phi^\perp$. First consider

transformations which map Ω_ϕ to itself. These correspond to the U(1) subgroup $U_Z(t)$ (or a discrete subgroup of this subgroup). These transformations do not map Ω_ϕ to itself (nor to $-\Omega_\psi$) and hence do not preserve the outcome probabilities of the effect. The only other possible transformations which preserve the outcome probabilities for these 4 states are of the form:

$$U\psi = \phi^\perp ,$$
$$U\phi = \psi^\perp .$$

In order for a phase group to exist $U^\dagger$ must also preserve the outcome probabilities:

$$U^\dagger \psi = \phi^\perp ,$$
$$U^\dagger \phi = \psi^\perp ,$$

entailing that $U^\dagger = U$. The only $U \in \mathrm{PU}(2)$ which has this property is the identity. Therefore the only phase group of this maximal measurement is the trivial one.

A.8 No-simultaneous encoding (Result 9)

All theories in $\mathcal{T}_2^I$ have pairs of non-antipodal states ω_0 and ω_1 which are perfectly distinguishable. The measurement which distinguishes them perfectly, also distinguishes $-\omega_0$ and $-\omega_1$.

$$E_i(\omega_j) = \delta_{ij} , \tag{77}$$
$$E_i(-\omega_j) = -\delta_{ij} + 1 . \tag{78}$$

A first bit a can be perfectly encoded as follows:

$$a = 0 \rightarrow \omega_0 \text{ or } -\omega_1 ,$$
$$a = 1 \rightarrow \omega_1 \text{ or } -\omega_0 .$$

The second bit a' can be encoded as:

$$a' = 0 \rightarrow \omega_0 \text{ or } \omega_1 ,$$
$$a' = 1 \rightarrow -\omega_1 \text{ or } -\omega_0 .$$

For example if Alice needed to encode the bits $a = 0$, $a' = 0$ she would choose the state ω_0. According to the scenario Alice encodes her bits a and a' in a single system and sends it to Bob. He then tries to guess one of the bits. If he chooses to guess the value of bit a he can do so with certainty (using the measurement which perfectly distinguishes ω_0 and $-\omega_1$ from $-\omega_0$ and ω_1). If Bob chooses to guess the value of bit a' he can use another effect to obtain partial information about

whether the state is ω_0 or ω_1 or whether the state is $-\omega_1$ or $-\omega_0$. This could be any effect which partially distinguishes $\frac{\omega_0+\omega_1}{2}$ from $-\frac{\omega_0+\omega_1}{2}$. Since neither of these is the maximally mixed state such an effect exists.

A.9 Pure-state dual theories (Result 10)

A two outcome measurement is a pair of effects $E_1 = (\mathbf{e}, c)$ and $E_2 = (-\mathbf{e}, 1-c)$. In theories with irreducible Γ (for $d = 2$) for each pure state $\Omega_\psi \in \mathcal{S}$ (where $\mathcal{S}$ is the state space) the state $-\Omega_\psi$ also exists. For a *pure state dual* theory we impose that effects are proportional to states and tight. Hence the effects are given by:

$$\left\{ \mathbf{e} = \frac{\Omega_\psi^T}{2} | \psi \in \mathrm{P}\mathbb{C}^2 \right\} . \tag{79}$$

The set of effects is such that for every $\mathbf{e}$ the linear functional $-\mathbf{e}$ also exists. Two outcome measurements are of the form $M = \{(\mathbf{e}, \frac{1}{2}), (-\mathbf{e}, \frac{1}{2})\}$. The linear functional $\mathbf{e}$ has a single maximum for state Ω_ψ^T and a single minimum for $-\Omega_\psi^T$. Hence the only states which are distinguishable are antipodal. This entails that the state spaces are bit symmetric. The fact that for all effects there is a single maximum/minimum entails that no-simultaneous encoding holds. Given a functional $\mathbf{e}$ there is a U(1) subgroup which leaves Ω_ψ and hence $\mathbf{e}$ invariant. Hence every measurement $M = \{(\mathbf{e}, \frac{1}{2}), (-\mathbf{e}, \frac{1}{2})\}$ has a U(1) phase group.

A.10 Bit symmetry in $\mathcal{T}_2^R$ (Result 11)

Consider a theory in $\mathcal{T}_2^R$; its state space $\mathcal{S}$ is a direct sum of state spaces of theories in $\mathcal{T}_2^I$:

$$\mathcal{S} = \bigoplus_{j \in \mathcal{J}} \mathcal{S}_j . \tag{80}$$

Moreover $\mathcal{S}$ has at least one faithful block in the decomposition. In the case where this block is non-quantum we denote it by k ($k > 1$, k odd). We assume all effects are allowed. Now consider effects with support solely on subspace k. The state space restricted to this subspace is just a state space corresponding to a theory with $j = k$ (and hence in $\mathcal{T}_2^I$). There exists a measurement (with support just in this subspace) which can distinguish pairs of non-antipodal states (by Result 6). Moreover there also exists a measurement (with support only in this subspace) which distinguishes pairs of antipodal states. Hence the

entire state space $\mathcal{S}$ has pairs of distinguishable states which are images of orthogonal rays and pairs which are not images of orthogonal rays.

If the generators contain a single faithful block $j = 1$ (with all others unfaithful) then a measurement on that block can distinguish a pair of states (which are images of orthogonal rays). The unfaithful blocks have pairs of distinguishable states which are not images of orthogonal rays. This follows from the fact that orthogonal rays are mapped to the same state in unfaithful representations and that the unfaithful block are also state spaces when considered alone. Since all state spaces have at least two distinguishable states it follows that pairs of states which are not images of orthogonal rays can be distinguished using effects with support in the unfaithful blocks.

Therefore all theories in $\mathcal{T}_2^R$ have pairs of distinguishable states which are images of orthogonal rays and pairs which are not; these theories violate bit symmetry.

A.11 Proof of Result 12

In order to show that all unrestricted state spaces are not bit symmetric we first need to establish how a $\mathrm{P}\mathbb{C}^{d-1}$ state space is embedded in a $\mathrm{P}\mathbb{C}^d$ state space. In the following we assume that all state spaces are unrestricted.

Lemma 3. Given a theory of $\mathrm{P}\mathbb{C}^d$ corresponding to maps Ω and Γ, where the highest weight representation in Γ is $\mathcal{D}_j^d$, the image of a $\mathrm{P}\mathbb{C}^{d-1}$ subspace of $\mathrm{P}\mathbb{C}^d$ under Ω is equivalent to a state space of a PU$(d-1)$ theory with representation Γ', whose highest weight component is $\mathcal{D}_j^{d-1}$.

This lemma is proven in appendix A.14. It follows from the lemma that any non-quantum $\mathrm{P}\mathbb{C}^3$ (i.e. which has a block $\mathcal{D}_j^3$ with $j > 1$) state space when restricted to a $\mathrm{P}\mathbb{C}^2$ subspace is equivalent to a non-quantum state space with a representation which has a block $\mathcal{D}_j^2$ (if j is even it must also be the case that the representation has a block with j odd since the representation of states is faithful). Hence the restricted state space belong to $\mathcal{T}_2^R$. By Result 11 this $\mathrm{P}\mathbb{C}^2$ state space is not bit symmetric, that is to say it has pairs of distinguishable states which have different (Hilbert space) inner products. Moreover the $\mathrm{P}\mathbb{C}^3$ state space is not bit symmetric since there are no transformations

in PU(3) mapping pairs of states with different Hilbert space inner products.

Any non-quantum $P\mathbb{C}^d$ state space (i.e. which has a block $\mathcal{D}_j^d$ with $j > 1$) has a $P\mathbb{C}^{d-1}$ subspace which is equivalent to a state space associated to a reducible representation which has a block $\mathcal{D}_j^{d-1}$. If the $P\mathbb{C}^{d-1}$ subspace is not bit symmetric (i.e. the Hilbert space inner product between two pairs of distinguishable state is not the same) then the $P\mathbb{C}^d$ state space is not bit symmetric either (since even considering the whole PU(d) transformation group its elements will still only map between pairs of states with the same inner product).

Hence by induction any non-quantum $P\mathbb{C}^d$ ($d \geq 2$) unrestricted state space is not bit symmetric.

A.12   Proof of Lemma 1

A.12.1   Partitions, U(d) and SU(d)

A partition λ is a set of d integers such that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{d-1} \geq \lambda_d \geq 0$. We define the size of the partition $|\lambda| = \sum_{i=1}^n \lambda_i = k$ and say that λ is a partition of k into d parts. We use square brackets to denote $\lambda = [\lambda_1, ..., \lambda_d]$.

Each partition λ is associated to an irreducible representation π_d^λ of U(d) [32]. Moreover any irreducible representation π_d^λ of U(d) is also an irreducible representation of SU(d) [32] . By this we mean that if we consider a subgroup $\pi_d^\lambda(U)$, $U \in$ SU(d) this representation is not reducible. Two irreducible representations $\lambda^{(1)}$ and $\lambda^{(2)}$ of U(d), when restricted to SU(d), are equivalent representations of SU(d) if and only if $\lambda_i^{(1)} - \lambda_i^{(2)}$ is constant for all i [32].

A representation of SU(d) corresponding to a partition λ and a representation of SU(d) corresponding to $\lambda + c$ are therefore equivalent. Hence an irreducible representation of SU(d) is characterised by the differences between elements of a partition λ. We can map λ to $(\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, ..., \lambda_{d-1} - \lambda_d)$ which is a vector of positive integers of size $d - 1$ which is the same for λ and $\lambda + c$ for all c . This vector $(\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, ..., \lambda_{d-1} - \lambda_d)$ is the previously introduced Dynkin index (which is in one-to-one correspondence with irreducible representations of SU(d)). We use square brackets to differentiate a partition from a Dynkin index. In the following it will be useful to describe representations of SU(d) using their partition (even if the map from partitions to representations of SU(d) is many to one). Since partitions differing by a constant give the same representation we fix $\lambda_d = 0$. This requirement fixes a representation of SU(d) for each λ.

A.12.2   U(1) × SU($d-1$) invariant subspaces

For a representation λ of SU(d) acting on V to have an U(1)×SU($d-1$)-invariant vector v means the following. If we take the representation of a U(1) × SU($d - 1$) subgroup (which is reducible) and consider its action on v, it leaves v invariant. The reducible representation of U(1) × SU($d-1$) is composed of irreducible blocks. If we decompose v according to these blocks, we see that the components of v in these subspaces (for non-trivial representations of U(1) × SU($d - 1$)) will not be left invariant by all of the transformations U(1)×SU($d-1$) (since each block is irreducible, so cannot have invariant subspaces within it). However if one of the blocks corresponds to a trivial representation of U(1) × SU($d - 1$) then it will leave that component of v invariant. Thus we can choose v to have support just in those subspaces which correspond to trivial representations in the decomposition of λ when restricting to a U(1) × SU($d - 1$) subgroup.

We now find the representations λ such that there exists at least one trivial representation of U(1)×SU($d-1$) when λ is restricted to this U(1)× SU($d - 1$) subgroup.

A.12.3   Branching rule for U(d) → U($d-1$)

Restricting representations of groups to subgroups leads to a reducible representation of the subgroup given by a branching rule. We introduce the necessary tools.

We consider a partition $\lambda = [\lambda_1, ..., \lambda_n]$. A partition $\mu = [\mu_1, ..., \mu_{n-1}]$ is said to interlace λ when:

$$\lambda_1 \geq \mu_1 \geq ... \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n \qquad (81)$$

The branching rule from U(d) to U($d - 1$) is as follows. Let U(d) have irreducible representation π_d^λ acting on a space V_λ. Then there is a unique decomposition of V_λ into subspaces under the action of U($d-1$) [33, p.19]:

$$V_\lambda = \bigoplus_\mu V_\mu \,, \qquad (82)$$

where the sum is over every μ which interlaces λ and V_μ is a carrier space for an irreducible representation of $U(d-1)$ labelled by μ.

A.12.4   Branching rule for $SU(d) \to SU(d-1) \times U(1)$

The restriction of π_d^λ (a representation of $U(d)$) to an $SU(d)$ subgroup is an irreducible representation of $SU(d)$ with partition λ (which for $SU(d)$ is defined up to a constant). This representation acts irreducibly on V_λ. The subgroup $U(d-1)$ is given by:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{pmatrix} , \quad U_{(d-1)}U_{(d-1)}^\dagger = \mathbb{1} . \quad (83)$$

It's action on V_λ decomposes as a direct sum of V_μ. We wish to consider the action of $SU(d-1)$ on V_λ. This subgroup is given by:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{pmatrix} , \quad U_{(d-1)\times(d-1)} \in SU(d-1) .$$
$$(84)$$

$U(d-1)$ acts on each V_μ irreducibly and restricting an irreducible representation of $U(d-1)$ to $SU(d-1)$ gives an irreducible representation with the same partition. Hence the $SU(d-1)$ acts irreducibly on each V_μ in the decomposition. The branching rule $SU(d) \to SU(d-1)$ is the same as $U(d) \to U(d-1)$ [32].

We now establish the branching $U(d) \to SU(d-1) \times U(1)$ subgroup. The $U(1)$ part corresponds to all matrices C of the following form:

$$C = \begin{pmatrix} e^{it} & & & \\ & e^{\frac{-i}{d-1}t} & & \\ & & \ddots & \\ & & & e^{\frac{-i}{d-1}t} \end{pmatrix} . \quad (85)$$

We note that this $U(1)$ subgroup commutes with the $SU(d-1)$ subgroup hence its action will leave the subspaces V_μ invariant. Therefore the space V_λ decomposes into V_μ invariant subspaces under the action of $SU(d-1) \times U(1)$. We now establish the action of the $U(1)$ subgroup on these subspaces. We use a similar technique as that found

in proof of theorem 8.1.2 of [34, p.364]. We consider a matrix $A \in U(d)$:

$$A = \begin{pmatrix} e^{it} & & & \\ & e^{it} & & \\ & & \ddots & \\ & & & e^{it} \end{pmatrix} , \quad (86)$$

And a matrix $B \in U(d-1)$:

$$B = \begin{pmatrix} 1 & & & \\ & e^{-i\frac{d}{d-1}t} & & \\ & & \ddots & \\ & & & e^{-i\frac{d}{d-1}t} \end{pmatrix} . \quad (87)$$

We note that $AB = C$ for the above defined $U(1)$ subgroup. We determine the action of A and B on V_λ which will allow us to know that of C. The action of A on the whole carrier space V_λ is multiplication by a scalar $e^{it|\lambda|}$. Similarly the action of B (which belongs to the $U(d-1)$ subgroup and commutes with all elements of $U(d-1)$) on each subspace V_μ is multiplication by $e^{-i\frac{d}{d-1}t|\mu|}$ [34, theorem 8.1.2. (proof) , p.364]. Hence we can compute the action of the $U(1)$ subgroup on each subspace which is multiplication by :

$$e^{it|\lambda|}e^{-i\frac{d}{d-1}t|\mu|} = e^{it(|\lambda|-\frac{d}{d-1}|\mu|)} , \quad (88)$$

Hence the action of this $U(1)$ subgroup on the carrier space V_λ acts by scalar multiplication on each V_μ (where the scalar can be the same for different μ). We can now summarise: given a representation π_d^λ of $SU(d)$ acting on V_λ there exists a decomposition (into invariant subspaces) under the action of $SU(d-1) \times U(1)$ given by:

$$V_\lambda = \bigoplus_\mu V_\mu , \quad (89)$$

where the sum is over every μ which intertwines λ. The μ determine irreducible representations of $SU(d-1)$ acting on each subspace. The $U(1)$ parts acts like $e^{it(|\lambda|-\frac{d}{d-1}|\mu|)}$ on each subspace.

A.12.5   Representation of $SU(d)$ with trivial representations under the action of $SU(d-1) \times U(1)$

We can now address the problem of finding which representations of $SU(d)$ are such that their restriction to $SU(d-1) \times U(1)$ contains a trivial representation of $SU(d-1) \times U(1)$.

A trivial representation of SU($d-1$) has Dynkin coefficients $(0,...0)$ and hence partition $\mu = [\mu_1,...\mu_{d-1}]$ where $\mu_1 = \mu_2 = ... = \mu_{d-1}$. Hence representations of SU(d) which contain trivial representations of SU($d-1$) in this decomposition will have partitions λ where $\lambda_2 = \lambda_3 = ... = \lambda_{n-1} = \mu_1$, following the requirement that μ intertwine λ.

We now consider the requirement that the U(1) action is trivial. Since the action of U(1) on V_μ is given by $e^{it(|\lambda|-\frac{d}{d-1}|\mu|)}$, it is trivial when $|\lambda| - \frac{d}{d-1}|\mu| = 0$. We note that some authors multiply the U(1) charge by a constant so that it is an integer value. Since we are considering only 0 U(1) charge this will not concern us.

We can now add this requirement on λ to the preceding one: $\lambda_2 = \lambda_3 = ... = \lambda_{d-1} = \mu_1$. As stated above we are considering $\lambda_d = 0$ in order to identify one representation of SU($d-1$) with each partition λ. We have

$$|\mu| = (d-1)\mu_1 \, , \qquad (90)$$

and

$$|\lambda| = \lambda_1 + (d-2)\mu_1 \, , \qquad (91)$$

We substitute this into the requirement $|\lambda| - \frac{d}{d-1}|\mu| = 0$:

$$\begin{aligned} \lambda_1 + (d-2)\mu_1 - d\mu_1 &= 0 \\ \lambda_1 &= 2\mu_1 \end{aligned} \qquad (92)$$

Hence the representation π_n^λ of SU(d) with partition $[2\mu_1, \mu_1, ..., \mu_1, 0]$ which corresponds to Dynkin label $(\mu_1, 0, ..., 0, \mu_1)$ meets the requirements. This shows that any representation of SU(d) with Dynkin label $(j, 0, ..., 0, j)$ (for any positive integer j) has a trivial subspace under the action of SU(d) $\times$ U(1). We now show that these representations have a single trivial representation in the decomposition. These are representations with partition $\lambda = [2j, j, ..., j, 0]$. Under the branching rule into SU($d-1$) $\times$ U(1) trivial representations correspond to partitions μ have components μ_i which are all equal (in order for the SU($d-1$) component to be trivial) and these must be equal to j. There is a single such μ of this form. This shows that there is a unique trivial representation in the decomposition.

The representations of SU(d) of the form $(j, 0, ..., 0, j)$ correspond to the representations $\mathcal{D}_j^d$. These are the only representations of SU(d) (and PU(d)) which leave a vector invariant under SU(d) $\times$ U(1) (moreover this vector is unique up to normalisation).

A.12.6 A comment on real and complex irreducibility

In this work we are interested in classifying certain representations of the dynamical group SU(d). These representations act on state spaces which span real vector spaces.

All the representation theory results used above concern representations acting on complex vectors spaces. A representation which acts on a complex vector space is real if it can be expressed in a basis where all the matrix elements are real. This is the case of the irreducible representations $\mathcal{D}_j^d$. These representations are irreducible when acting on both complex and real vector spaces. However not all real irreducible representations correspond to complex irreducible representations. There are reducible representations acting on a complex vector space whose action is irreducible when acting on a real vector space. We have found all real irreducible representations which are also complex irreducible which meet our criteria. However it may be the case that there are representations which are real irreducible but complex reducible which meet our criteria. If this is the case our approach would not have found them. We now show that there are no such representations satisfying the invariance properties we require.

A real representation Γ which is irreducible on a real vector space but reducible on a complex one can be block diagonalised in the following form $\Gamma = \rho \oplus \bar{\rho}$ [22, Exercise 3.39, p. 41]. For example consider the following representation of SO(2) $\cong$ U(1):

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} . \qquad (93)$$

This is irreducible when acting on $\mathbb{R}^2$. However its action on $\mathbb{C}^2$ can be diagonalised as follows:

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} . \qquad (94)$$

Hence there may be representations which are real irreducible which have SU($d-1$) $\times$ U(1) invariant vectors but are complex reducible. All real irreducible representations are either: complex irreducible or complex reducible of the form $\rho + \bar{\rho}$.

Let us consider a real irreducible representation Γ which is complex reducible. There exists a transformation L such that $\Gamma = L(\rho \oplus \bar\rho)L^{-1}$. We now ask the same question: does there exist a vector which is invariant under the subgroup $H = \mathrm{SU}(d-1) \times \mathrm{U}(1)$. Let us call this vector v. We have:

$$\Gamma_{|H} v = v . \tag{95}$$

We observe that if such a vector exists then the (complex) vector $L^{-1}v$ is invariant under $L^{-1}\Gamma_{|H}L = \rho_{|H} \oplus \bar\rho_{|H}$. Moreover if there exists a vector w which is left invariant by $\rho_{|H} \oplus \bar\rho_{|H}$ then the vector Lw is left invariant under $\Gamma_{|H}$. Hence if we show that there are no vectors left invariant under $\rho_{|H} \oplus \bar\rho_{|H}$ then there are no vectors left invariant under $\Gamma_{|H}$.

There exists a vector w invariant under $\rho_{|H} \oplus \bar\rho_{|H}$ if and only if this representation (obtained by restricting $\rho \oplus \bar\rho$) has at least one trivial component. This is only possible if at least one of the representations $\rho_{|H}$ or $\bar\rho_{|H}$ has one or more trivial components.

Here ρ and $\bar\rho$ are complex irreducible representations of $\mathrm{SU}(d)$. The only such representations with a $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ invariant vector are of the form $\mathcal{D}_j^d$ which is real (and we observe $\mathcal{D}_j^d = \bar{\mathcal{D}}_j^d$). Hence there are no representations which are complex reducible but real irreducible which have these invariant vectors.

A.13   Proof of Lemma 2

We consider two $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ stabilizer subgroups of two distinct rays on $\mathrm{P}\mathbb{C}^d$ ($d \geq 3$). Consider a ray ψ with stabilizer group G_ψ. The Lie algebra $\mathfrak{g}_\psi$ which generates this group has elements X_ψ (corresponding to $\mathrm{U}(1)$)

$$X_\psi = \left(\begin{array}{c|ccc} i & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & -\frac{i}{d-1}\, \mathbb{I}_{d-1} & \\ 0 & & & \end{array} \right), \tag{96}$$

and elements Y_ψ (corresponding to the $\mathrm{SU}(d-1)$ group)

$$Y_\psi = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A & \\ 0 & & & \end{array} \right), \quad A = -A^\dagger,\ \mathrm{Tr}(A) = 0 . \tag{97}$$

As can be seen by the generators ψ is the unique ray stabilized by this subgroup. Hence distinct rays ψ and ψ' have distinct stabilizer groups (which are equivalent up to conjugation). This is not the case for $\mathrm{SU}(2)$ for example; the $\mathrm{U}(1)$ stabilizer of $|0\rangle$ also stabilizes $|1\rangle$.

The two subgroups G_ψ and $G_{\psi'}$ (which stabilizes ψ') are *maximal* subgroups of $\mathrm{SU}(d)$ [35]. A maximal subgroup H of G is a proper subgroup (i.e. $H \neq G$) such that if $H \leq K \leq G$ then $H = K$ or $K = G$. The group H generated by these two groups G_ψ and $G_{\psi'}$ is not equal to either G_ψ and $G_{\psi'}$. Hence it is equal to the full group $\mathrm{SU}(d)$.

A.14   Proof of Lemma 3

A.14.1   Embeddedness of simple theories

In this section we look at theories which do not obey the constraint given by (9) and study how $\mathrm{P}\mathbb{C}^{d-1}$ subspaces are embedded in $\mathrm{P}\mathbb{C}^d$ state space. We consider a $\mathrm{P}\mathbb{C}^d$ state space, which can be generated by applying all transformations Γ_U to a reference state $\Omega_0 = |0\rangle\langle 0|^{\otimes N}$:

$$\begin{aligned} \Gamma_U \Omega_0 &= U^{\otimes N} |0\rangle\langle 0|^{\otimes N} U^{\dagger \otimes N} \\ &= \mathrm{Sym}^N U |0\rangle\langle 0|^{\otimes N} \mathrm{Sym}^N U^\dagger, \end{aligned} \tag{98}$$

The action of the product $\mathrm{Sym}^N U \otimes \mathrm{Sym}^N U^\dagger$ can be decomposed (using known rules for Young tableaux) as:

$$\begin{aligned} &(N, 0, ..., 0) \otimes (0, ..., 0, N) \\ &= (0, ..., 0) \oplus (1, 0, ..., 0, 1) \oplus ... \oplus (N, 0, ..., 0, N) . \end{aligned} \tag{99}$$

In our previous notation this is just:

$$\Gamma_U = \bigoplus_{i=0}^{N} \mathcal{D}_i^d . \tag{100}$$

The state space is given by Hermitian matrices acting on the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$. However it is not immediately clear that every summand in (100) acts on the space of Hermitian matrices (which is the state space). For example when we generate the state space by acting with a reducible representation on a reference state, if the reference state does not have support in every block of the representation then there are certain components of the representation which are superfluous (since they do not act on the state

space). The dimension of the space of Hermitian matrices is:

$$D_\omega = \binom{d+n-1}{n}^2 \qquad (101)$$

Using the dimension formula for the representations $\mathcal{D}_j^d$ given by (16) and requiring that the representation of the transformations acts on a space of dimension D_ω we see that every summand in (100) acts on the state space. Hence we see that these theories are equivalent to reducible theories of $\mathbb{PC}^d$ with $\mathcal{J} = \{1, 2, ..., N\}$. We consider a $\mathbb{PC}^{d-1}$ subspace of the Hilbert space and see which subspace of the state space it corresponds to. It can be generated from the state $|1\rangle\langle 1|^{\otimes N}$ by applying an SU($d-1$) subgroup:

$$U^{\otimes N} |1\rangle\langle 1|^{\otimes N} U^{\dagger \otimes N} , \qquad (102)$$

With:

$$U = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{pmatrix} \qquad (103)$$

We notice that this state space is just a $\mathbb{PC}^{d-1}$ state space of the type described above with representation:

$$\Gamma' = \bigoplus_{i=0}^{N} \mathcal{D}_i^{d-1} . \qquad (104)$$

A.14.2   Embeddedness of irreducible theories

We consider a $\mathbb{PC}^d$ theory with representation $\mathcal{D}_j^d$. We want to know the state space corresponding to a $\mathbb{PC}^{d-1}$ subspace. We consider a basis $\{|0\rangle, |1\rangle, ..., |d-1\rangle\}$ and determine the image of all states of the form $\alpha_1 |1\rangle + ... + \alpha_{d-1} |d-1\rangle$. These are all the states (apart from $|0\rangle$) which are invariant under the U(1) group of the SU($d-1$) × U(1) subgroup specified in the proof of Lemma 2. From Lemma 4 (proven in appendix A.15) we know that restricting the representation $\mathcal{D}_j^d$ to a SU($d-1$) × U(1) subgroup gives a reducible representation of SU($d-1$) × U(1). Moreover the U(1) action is trivial for the following blocks:

$$\bigoplus_{i=0}^{j} \mathcal{D}_i^{d-1} . \qquad (105)$$

The image of $\alpha_1 |1\rangle + ... + \alpha_{d-1} |d-1\rangle$ therefore lies in this subspace (since it is invariant under

this U(1) action) which we call the $\alpha = 0$ subspace. Moreover the image of $|0\rangle$ is uniquely determined as the trivial subspace $\mathcal{D}_0^{d-1}$ (which is invariant under the whole subgroup). However we do not know if the image of the $\mathbb{PC}^{d-1}$ subspace spans the whole $\alpha = 0$ subspace. We show that it must have support in the subspace $\mathcal{D}_j^{d-1}$ (which is part of the $\alpha = 0$ subspace).

We first consider a reducible theory described with representation:

$$\Gamma = \bigoplus_{i=0}^{j} \mathcal{D}_i^d . \qquad (106)$$

This corresponds to a reducible theory as characterised above. We call its state space $\mathcal{S}$. Considering a $\mathbb{PC}^{d-1}$ subspace gives a state space $\mathcal{S}'$ with representation:

$$\Gamma' = \bigoplus_{i=0}^{j} \mathcal{D}_i^{d-1} . \qquad (107)$$

We observe that the state space with representation Γ is equivalent to a direct sum of state spaces with representation $\mathcal{D}_k^d$ for $k = 1, ..., j$. We label these state spaces $\mathcal{S}_k$. We write:

$$\mathcal{S} = \bigoplus_{k=0}^{j} \mathcal{S}_k . \qquad (108)$$

Moreover reducing a state space $\mathcal{S}_k$ gives a $\mathbb{PC}^{d-1}$ state space with a representation inside the direct sum (i.e. which may or may not have support in each representation):

$$\bigoplus_{i=0}^{k} \mathcal{D}_i^{d-1} . \qquad (109)$$

We have seen that the restriction of $\mathcal{S}$ gives a state space $\mathcal{S}'$ with support in block $\mathcal{D}_j^{d-1}$. Moreover since $\mathcal{S}$ is a direct sum of state spaces $\mathcal{S}_k$ it must be the case that (at least) one of these $\mathcal{S}_k$ reduces to a state space with support in $\mathcal{D}_j^{d-1}$. The state space $\mathcal{S}_j$ is the only state space in the direct sum in (108) which can have support in $\mathcal{D}_j^{d-1}$ when restricted. That is to say the reduction of $\mathcal{S}_j$ (with representation $\mathcal{D}_j^d$) gives a $\mathbb{PC}^{d-1}$ state space with support in block $\mathcal{D}_j^{d-1}$.

A.14.3   Embeddedness of arbitrary reducible theories

From the above considerations we see that an irreducible theory $\mathcal{J} = j$ of $\mathbb{PC}^d$ when restricted to

a $P\mathbb{C}^{d-1}$ subspace must give a state space which has support in the subspace $\mathcal{D}_j^{d-1}$, and hence corresponds to a theory of $P\mathbb{C}^{d-1}$ which has a representation which contains at least a block $\mathcal{D}_j^{d-1}$. A reducible theory of $P\mathbb{C}^d$ $\mathcal{J} = j_1, ..., j_n$ has a state space $\mathcal{S}$ which is a direct sum of state spaces $\mathcal{S}_{j_i}$. The restriction of $\mathcal{S}$ to a $P\mathbb{C}^{d-1}$ subspace is equivalent to restricting each of the $\mathcal{S}_{j_i}$ state spaces. Hence the restriction of $\mathcal{S}$ to $P\mathbb{C}^{d-1}$ will have support in at least subspaces $\mathcal{D}_k^{d-1}$ for $k = j_1, ..., j_n$.

The Lemma follows directly from this.

A.15 Proof of Lemma 4

Lemma 4. Consider a representation $\mathcal{D}_j^d$ and its decomposition under a $SU(d-1)\times U(1)$ subgroup. The reducible representation of $SU(d-1) \times U(1)$ has blocks with various U(1) charges. The subspace with 0 U(1) charge is acted upon by the representation of $SU(d-1)$:

$$\bigoplus_{i=0}^{j} \mathcal{D}_{d-1}^i \ . \tag{110}$$

Similarly to Lemma 1 we label $\mathcal{D}_j^d$ with partition $\lambda = [2j, j, j...., j, 0]$. The restriction to a $SU(d-1) \times U(1)$ subgroup acts reducibly on the carrier space V_λ as:

$$V_\lambda = \bigoplus_\mu V_\mu \ , \tag{111}$$

where the sum is over every μ which intertwines λ and V_μ is a carrier space for an irreducible representation of $SU(d-1)$ with partition μ . The μ which intertwine λ are of the form:

$$\mu = [a, j, j, ..., b], \ 2j \geq a \geq j \ , j \geq b \geq 0 \ . \tag{112}$$

We have $|\lambda| = dj$ and $|\mu| = (d-3)j + a + b$. From Lemma 1 we have the condition that the U(1) charge is 0 when $|\lambda| - \frac{d}{d-1}|\mu| = 0$. We now substitute in the relevant expressions:

$$|\lambda| - \frac{d}{d-1}|\mu| = 0$$
$$dj - \frac{d}{d-1}((d-3)j + a + b) = 0 \tag{113}$$
$$(d-1)j - (d-3)j - a - b = 0$$
$$a = 2j - b \ .$$

This entails that when $b = 0$, $a = 2j$, when $b = 1, a = 2j - 1...$ and when $b = j, a = j$. Every possible μ (intertwining λ) is in the direct sum

and hence every μ which meets the above condition for having 0 U(1) charge is in the decomposition. Hence there are $j + 1$ subspaces V_μ where the action of U(1) is trivial. If we express the representations acting on these in terms of Dynkin notation we observe that $\mu = [2j-b, j, ..., j, 2j-b]$ becomes $(j - b, 0, 0..., j - b)$ for $b = 0, ..., j$. The terms in the direct sum are therefore just the representations $\mathcal{D}_i^{d-1}$ for $i = 0, ..., j$.