How dynamics constrains probabilities in general probabilistic theories

Thomas D. Galley[1] and Lluís Masanes[2]

[1]Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada

[2]Department of Computer Science, University College London, Gower Street, London WC1E 6BT, United Kingdom

We introduce a general framework for analysing general probabilistic theories, which emphasises the distinction between the dynamical and probabilistic structures of a system. The dynamical structure is the set of pure states together with the action of the reversible dynamics, whilst the probabilistic structure determines the measurements and the outcome probabilities. For transitive dynamical structures whose dynamical group and stabiliser subgroup form a Gelfand pair we show that all probabilistic structures are rigid (cannot be infinitesimally deformed) and are in one-to-one correspondence with the spherical representations of the dynamical group. We apply our methods to classify all probabilistic structures when the dynamical structure is that of complex Grassmann manifolds acted on by the unitary group. This is a generalisation of quantum theory where the pure states, instead of being represented by one-dimensional subspaces of a complex vector space, are represented by subspaces of a fixed dimension larger than one. We also show that systems with compact two-point homogeneous dynamical structures (i.e. every pair of pure states with a given distance can be reversibly transformed to any other pair of pure states with the same distance), which include systems corresponding to Euclidean Jordan Algebras, all have rigid probabilistic structures.

1 Introduction

General probabilistic theories (GPTs) provide a framework for the study of operational theories beyond quantum theory. Within this framework quantum theory appears as one non-classical theory amongst many. This field has its origin in the work of Segal [1], Mackey [2] and Ludwig [3–5] with other notable contributions at the time including [6–13] amongst others. Contemporary interest in GPTs was kickstarted by Hardy's seminal work [14] followed by a detailed exposition of the framework by Barrett [15]. Important applications of the framework include the operational derivations of quantum theory of [16–18]. Current treatments have tended to emphasise finite dimensional systems and system composition. Using this framework (or related frameworks such as convex operational theories [19, 20] and operational probabilistic theories [21]) many physical and informational features of general probabilistic theories have been studied, such as interference phenomena [22–25], computation [26, 27], thermodynamics [28–30] and others [31–38].

Examples of GPTs (excepting classical and quantum theory) include Boxworld [15, 39–42], quantum theory over the field of real numbers [43–45] or quaternions [46], theories based on Euclidean Jordan algebras [35], quartic quantum theory [47], d-balls [16, 48, 49], density cubes [50] and quantum systems with modified measurements [51]. Amongst these, only Boxworld, quantum theory over real or quaternionic fields and theories based on Euclidean Jordan algebras are full theories, in that they have non-trivial composites.

The aim of this paper is to provide tools to systematically explore the space of non-classical systems. Rather than generating examples of non-classical systems we can give full classifications of families of non-classical systems which share a common dynamical structure (pure states and reversible dynamics) but different probabilistic structures (measurements and measurement outcome

---

Thomas D. Galley: tgalley1@perimeterinstitute.ca

probabilities); as done in [51] for systems which share the dynamical structure of quantum systems. We can thus obtain a richer picture of the space of non-classical systems, of which quantum systems are just one example.

We provide a general framework for convex systems and use it to study *transitive* systems, that is to say systems for which any two pure states are related by a reversible transformation. This is a generalisation of the OPF (outcome probability function) framework of [51–53], where the pure states and dynamical group no longer have to be those of quantum theory. We restrict ourselves to systems with reversible dynamics given by finite and compact groups, noting that all the examples of GPT systems mentioned previously are transitive systems with finite or compact dynamical groups. The assumption of transitivity has played an important role in derivations of quantum theory from operational/information theoretic principles, for instance in Hardy's original derivation [14] as well as subsequent derivations by other authors [16–18, 54]. It is worth mentioning that many derivations of the second law of thermodynamics from more fundamental principles (see for example [55–59]) use as the central premise the reversibility of the underlying dynamics (both in the classical and quantum frameworks). Also, when all the transformations that can be implemented on a system are generated by reversible dynamics, all the achievable states of the system form a transitive space. Hence, there is a connection between transitivity and the second law of thermodynamics.

We show that for a given dynamical structure (pure states and dynamical group) every possible probabilistic structure (measurements and outcome probabilities) is in correspondence with a representation of the dynamical group. Moreover we find necessary and sufficient conditions on the dynamical structure (the dynamical group and subgroup form a Gelfand pair) which make this correspondence one-to-one. We find that certain probabilistic structures cannot be infinitesimally deformed and call these rigid. We show that all dynamical structures which are Gelfand pairs do not have any probabilistic structures which can be infinitesimally deformed. We apply the methods developed to classify generalisations of quantum systems, with pure states given by Grassmann manifolds and unitary dynamics. We introduce the family of systems with compact two point homogeneous dynamical structures and show that they all have rigid probabilistic structures.

1.1 Structure of the paper

In Section 2 we introduce the OPF framework used for studying transitive systems and present relevant known results (or slight generalisations thereof). In Section 3 we give the main theorem of this work (the classification theorem), establishing a correspondence between probabilistic structures of transitive systems and group representations, as well as the conditions under which this correspondence is one-to-one. In Section 4 we introduce the notion of deformation of probabilistic structures, and show that the only dynamical structures which admit probabilistic structures which can be infinitesimally deformed are those corresponding to non-Gelfand pairs. We also give an explicit example of deformations of a non-rigid probabilistic structure. In Section 5 we introduce the family of compact dynamical structures which are two point homogeneous and show that they are all rigid. In Section 6 we apply the classification theorem to systems with dynamical structures given by complex Grassmann manifolds (a generalisation of complex projective space). In Section 7 we discuss the results of this paper in light of existing work as well as comment on the implications of new concepts and results of the present work. Lastly we close with some concluding remarks in Section 8. A glossary of notation is given in Appendix A and an introduction to some of the representation theory used in this paper can be found in Appendices B and C.

2 Single system state spaces

We provide a characterisation of single systems within the GPT framework which emphasises the pure states and reversible dynamics. This will allow us to consider families of systems with the same pure states and reversible dynamics, but different measurements. This is a generalisation of [51] where all systems with the same pure states and reversible dynamics as quantum theory were classified and their informational properties studied. We first describe quantum systems in this framework.

2.1 A characterisation of finite dimensional quantum systems

Quantum systems are often characterised directly in terms of mixed states, that is to say their convex representation. States are positive semi-definite operators on a finite dimensional complex space $\mathbb{C}^d$ equipped with a sesquilinear inner product, transformations are CPTP maps and measurements are associated to POVMs, with the probability of an outcome occuring being given by the usual trace rule. Here we provide a characterisation of quantum systems which separates their dynamical structure from the probabilistic structure. In this characterisation the mixed state representation described above is derived, rather than postulated. Moreover this distinction between dynamical and probabilistic structures will provide us with a way of classifying families of more general systems which share a common dynamical structure. A quantum system $\mathcal{S}_d^{\texttt{Quant}}$ associated to the space $\mathbb{C}^d$ is given by the following:

    I. Pure states: $\psi \in \mathrm{P}\mathbb{C}^d$.

    II. Reversible dynamics: $\psi \mapsto U\psi$, $U \in \mathrm{PU}(d)$.

    III. Outcome probabilities: $\mathcal{F}_{\mathrm{P}\mathbb{C}^d}^{\texttt{Quant}} = \{Q : \mathrm{P}\mathbb{C}^d \to [0,1] \,|\, Q(\psi) = \langle\psi|\hat{Q}|\psi\rangle \,, \, \forall \hat{Q} : 0 \leq \hat{Q} \leq \mathbb{I}\}$.

$\mathrm{P}\mathbb{C}^d$ is the complex projective space with elements corresponding to one dimensional subspaces of $\mathbb{C}^d$. $\mathrm{PU}(d)$ is the projective unitary group constructed from $\mathrm{U}(d)$ by taking equivalence classes of unitaries under multiplication by a complex phase: $U_0 = e^{i\theta}U_1 \iff U_0 \cong U_1$ where $U_0, U_1 \in \mathrm{U}(d)$ and $\theta \in \mathbb{R}$.

We assume that any subset $\{Q_i\}_{i=1}^n$ such that $\sum_i Q_i(\psi) = 1$ for all $\psi \in \mathrm{P}\mathbb{C}^d$ forms a valid measurement. This implies that a measurement consists of positive semi-definite operators $\hat{Q}_i$ such that $\sum_i \hat{Q}_i = 1$. Here I. and II. are the dynamical structure, whilst III. is the probabilistic structure. The mixed state representation (density operators) is derived from the dynamical and probabilistic structures. We observe that the probability assignment (Born rule) is not given in terms of the trace, since this already presumes the structure of mixed states. We now define general non-classical systems in terms of dynamical and probabilistic structures and show how to derive the convex representation.

2.2 Dynamical structure

The pure states of a system $\mathcal{S}$ form a set X, and the reversible dynamics a group G. The action of G on X is given by a group action $\varphi : G \times X \to X$. This gives X the structure of a G-space.

Definition 1 (Dynamical structure). A dynamical structure $\mathcal{D}$ is a triplet

$$\mathcal{D} = (X, G, \varphi), \tag{1}$$

where X is a set, G is a group and φ a group action.

In the following we leave φ implicit and write gx for $\varphi(g, x)$. An important family of dynamical structures are transitive. A dynamical structure is *transitive* when for any two pure states $x, x' \in X$ there exists a transformation $g \in G$ such that $x' = gx$. In other words X is the orbit of G acting on an arbitrary $x \in X$.

A central notion to the approach used in this work is that of a *stabilizer subgroup* (also known as isotropy group) of an element $x \in X$, which is just the subgroup of all transformations in G which leave a point x invariant. We write $H_x := \{g \in G : gx = x\}$ for the stabilizer subgroup of a point $x \in X$. For a transitive group action, the stabilizer groups for different points are isomorphic, hence we write H as the stabilizer group.

Given a group G and a subgroup $H \subseteq G$, we denote by $\phi_{G,H}$ the action of G on the set of left cosets G/H. Given a transitive dynamical structure $\mathcal{D} = (X, G, \varphi)$ with stabilizer subgroup H we have the following isomorphism of dynamical structures $(X, G, \varphi) \cong (G/H, G, \phi_{G,H})$.

Typically dynamical structures (X, G, φ) also have topological (and sometimes differentiable) structure. A topological group G acts on a topological space X when the action φ is a continuous function: $\varphi : G \times X \to X$ (where $G \times X$ has the product topology). If H is a subgroup of G then

the space G/H of left cosets is a topological space with respect to the quotient topology, which is the finest topology making the quotient map $q : G \rightarrow G/H$, $q : g \mapsto gH$ continuous.

For transitive dynamical structures (X, G, φ) where X is Hausdorff and G is compact the isomorphism $(X, G, \varphi) \cong (G/H, G, \phi_{G,H})$ also involves the topological structure of each component of the triplet [60, Proposition 1.10]. When G is a Lie group the isomorphism $(X, G, \varphi) \cong (G/H, G, \phi_{G,H})$ involves the differentiable structure of each component of the triplet [61, Theorem 20.12].

In the following we restrict our attention to compact dynamical structures (which includes finite dynamical structures as a special case) implying that the isomorphism $(X, G, \varphi) \cong (G/H, G, \phi_{G,H})$ includes the topological and differentiable structures of the dynamical structures considered.

For this reason we use the abbreviation

$$\mathcal{D} = (X, G, \varphi) = (G, H) . \tag{2}$$

for transitive dynamical structures with stabilizer subgroup H.

The case of non-compact dynamical structures is discussed in Section 2.5.

2.3   Probabilistic structure

A system is determined by its pure states, dynamics and measurements. Given the dynamical structure we need to specify its probabilistic structure, which characterises the measurements which can be performed on the system.

Definition 2 (Outcome probability function (OPF))**.** An outcome of a measurement on a system with pure states X is given by a function $\mathbf{f} : X \rightarrow [0, 1]$, where the probability of the associated outcome $\mathbf{f}$ occurring is $P(\mathbf{f}|x) = \mathbf{f}(x)$.

Definition 3 (Measurement)**.** A measurement $\mathcal{M}_j$ with countable outcomes $i = 1, ..., n, ..$ is specified by the list $\{\mathbf{f}_1^j, ..., \mathbf{f}_n^j, ...\}$. The elements of this list obey the condition:

$$\sum_i \mathbf{f}_i^j(x) = 1, \ \forall x \in X . \tag{3}$$

Definition 4 (Unit OPF)**.** The unit OPF $\mathbf{u}$ is $\mathbf{u}(x) = 1, \forall x \in X$.

Definition 5 (Probabilistic structure)**.** The probabilistic structure of a system is the set $\mathcal{F}_X$ of all outcome probability functions $\mathbf{f}$.

Typically we assume that any set $\{\mathbf{f}_1, ..., \mathbf{f}_n, ...\}$ such that $\sum_i \mathbf{f}_i = \mathbf{u}$ forms a valid measurement, however this assumption is not necessary. When this assumption does not hold, one needs to supplement the set $\mathcal{F}_X$ with a specification of which OPFs form a valid measurement. One example of such a specification is the 'finite measurement outcomes' assumption:

Definition 6 (Finite measurement outcome assumption)**.** Only finite sets of OPFs $\{\mathbf{f}_1, ..., \mathbf{f}_n\}$ such that $\sum_i \mathbf{f}_i = \mathbf{u}$ form valid measurements.

The above assumption is sometimes viewed as part of the definition of measurements in an operational framework, since we can never carry out measurements with infinitely many outcomes. We will be making this assumption in the present work.

Operational considerations impose the following constraints on $\mathcal{F}_X$:

i . $\mathcal{F}_X$ is closed under taking mixtures: for all $\mathbf{f}_1, \mathbf{f}_2 \in \mathcal{F}_X$ and all $\lambda \in [0, 1]$ we have that $\lambda \mathbf{f}_1 + (1 - \lambda)\mathbf{f}_2 \in \mathcal{F}_X$.

ii . $\mathcal{F}_X$ is closed under composition with group transformations: for all $\mathbf{f} \in \mathcal{F}_X$ and $g \in G$ we have that $\mathbf{f} \circ g \in \mathcal{F}_X$, where $(\mathbf{f} \circ g)(x) = \mathbf{f}(gx)$.

iii . $\mathcal{F}_X$ is closed under coarse graining of measurement outcomes: for any pair of outcomes $\mathbf{f}_i^k, \mathbf{f}_j^k \in \mathcal{M}_k$ of a given measurement $\mathcal{M}_k$ we have that $\mathbf{f}_i^k + \mathbf{f}_j^k \in \mathcal{F}_X$.

iv .   For every $\mathbf{f} \in \mathcal{F}_X$, the complement OPF $\mathbf{f}_c = \mathbf{u} - \mathbf{f}$ is also in $\mathcal{F}_X$

The first constraint implies that $\mathcal{F}_X$ is a convex set, hence it can be extended to a vector space $\mathbb{R}[\mathcal{F}_X]$ with addition $(\mathbf{f}_1 + \mathbf{f}_2)(x) = \mathbf{f}_1(x) + \mathbf{f}_2(x)$ and scalar multiplication $(\alpha \mathbf{f})(x) = \alpha \mathbf{f}(x)$ for any $\alpha \in \mathbb{R}$. Explicitly $\mathbb{R}[\mathcal{F}_X] := \{\tilde{\mathbf{f}}(x) | \tilde{\mathbf{f}}(x) = \alpha \mathbf{f}_1(x) + \beta \mathbf{f}_2(x), \ \alpha, \beta \in \mathbb{R}, \ \mathbf{f}_1, \mathbf{f}_2 \in \mathcal{F}\}$. We have that $\mathcal{F}$ is a convex subset of $\mathbb{R}[\mathcal{F}_X]$ by construction and that $\text{span}_{\mathbb{R}}(\mathcal{F}_X) = \mathbb{R}[\mathcal{F}_X]$.

Closedness under composition with group transformations implies that $\mathcal{F}_X$ is a G-space. This and the fact that the group action commutes with taking mixtures implies that $\mathbb{R}[\mathcal{F}_X]$ is a linear representation of G. Closedness under coarse graining of measurement outcomes implies that every $\mathcal{F}_X$ contains the unit OPF and the existence of the complement guarantees the existence of the $\mathbf{0}$ OPF. We introduce the following property, though we will not always require it in the present treatment.

Definition 7 (Separability of pure states)**.** A probabilistic structure $\mathcal{F}_X$ separates pure states when for any two pure states $x_1, x_2 \in X$ there exists an OPF $\mathbf{f} \in \mathcal{F}_X$ such that $\mathbf{f}(x_1) \neq \mathbf{f}(x_2)$.

If one does not have this requirement, the probabilistic structure $\mathcal{F}_X = \{\mathbf{u}\}$ leading to a trivial system for all dynamical structures is valid for example.

2.4 Systems, state spaces and associated group representations

The above definitions allow us to formally define a system $\mathcal{S}_X$.

Definition 8 (System)**.** A system $\mathcal{S}_X$ is a triple $\mathcal{S}_X = (X, G, \mathcal{F}_X)$, where (X, G) is a dynamical structure and $\mathcal{F}_X$ is a probabilistic structure.

In the following we briefly outline how the general state space (including mixed states) of a system is derived, both from an operational starting point and directly from the mathematical starting point $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$.

2.4.1 Operational derivation of the state space

Operationally for a single system one has access to a preparation device which is wired up sequentially with a transformation and measurement devices. These devices have classical settings (for instance which transformation to apply) and classical readouts (for instance which measurement outcome occurred). In an experiment one collects the statistics for different outcomes given choices of settings. Typically one assumes that statistics are gathered for all possible setting choices, and that the relative frequencies obtained become probabilities as the number of runs tends to infinity. Using these probabilities (which are directly given by the set $\mathcal{F}$ in the OPF framework) one derives the convex state space (and effect space) of the system. We refer the reader to [62] about how one can in practice derive a state and effect space from experimental data.

2.4.2 Mathematical derivation of the state space

In this work we will make the assumption of the possibility of state estimation using a finite outcome set (known as 'Possibility of state estimation' in [53]).

Definition 9 (Possibility of state estimation using a finite outcome set)**.** The system $\mathcal{S} = \{X, G, \mathcal{F}_X\}$ is such that the value of a finite number of outcomes $\mathbf{f}_1, ..., \mathbf{f}_n \in \mathcal{F}_X$ on any ensemble $\{(p_i, x_i)\}_i$ determines the value of any OPF $\mathbf{f} \in \mathcal{F}_X$ on the ensemble $\{(p_i, x_i)\}_i$.

It is shown in Lemma 2 of [53] that this implies that $\mathbb{R}[\mathcal{F}_X]$ is finite dimensional. Equivalently the convex set of mixed stated is embeddable in a finite dimensional real vector space. We now briefly outline the derivation of the space of mixed state for a system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ under the assumption "Possibility of state estimation using a finite outcome set". First the probability of an outcome $\mathbf{f}$ (defined on X) occurring for an ensemble $\{(p_i, x_i)\}_i$ is $P(\mathbf{f}|\{(p_i, x_i)\}_i) = \sum_i p_i \mathbf{f}(x_i)$. This allows us to define *equivalent* ensembles.

Definition 10 (Equivalent ensembles)**.** Two ensembles $\{(p_i, x_i)\}_i$ and $\{(p'_j, x'_j)\}_j$ are equivalent if $P(\mathbf{f}|\{(p_i, x_i)\}_i) = P(\mathbf{f}|\{(p'_j, x'_j)\}_j) \ \forall \mathbf{f} \in \mathcal{F}_X$, and we write $\{(p_i, x_i)\}_i \sim \{(p'_j, x'_j)\}_j$.

The mixed states are defined as equivalence classes of ensembles under this equivalence relation. For each state $x \in X$ we define the linear functional $\Omega_x : \mathbb{R}[\mathcal{F}_X] \to \mathbb{R}$ as $\Omega_x(\mathbf{f}) = \mathbf{f}(x)$. The probability of outcome $\mathbf{f}$ on ensemble $\{(p_i, x_i)\}_i$ can be written as

$$P(\mathbf{f}|\{(p_i, x_i)\}_i) = \sum_i p_i \mathbf{f}(x_i) = \sum_i p_i \Omega_x(\mathbf{f}) = \omega(\mathbf{f}) \,, \tag{4}$$

where we define the functional associated to ensemble $\{(p_i, x_i)\}_i$ as $\omega = \sum_i p_i \Omega_x$. Therefore, two ensembles $\{(p_i, x_i)\}_i$ and $\{(p'_j, x'_j)\}_j$ are equivalent if and only if, their corresponding functionals are identical $\sum_i p_i \Omega_{x_i} = \sum_j p'_j \Omega_{x'_j}$. The outcome probabilities $P(\mathbf{f}|\{(p_i, x_i)\}_i)$ on the space of ensembles uniquely define linear functionals $\Lambda_\mathbf{f}$ on the space of mixed states, such that $\Lambda_\mathbf{f} \cdot \omega = \omega(\mathbf{f})$ for all mixed states ω.

The group action $\varphi : X \times G \to X$, naturally extends to the space of mixed states (embedded linearly in $\mathbb{R}[\mathcal{F}_X]^*$) as $\Omega_x \xrightarrow{g} \Omega_{gx}$. This is a linear action which is such that $\Omega_x \xrightarrow{g} \Omega_{gx} \xrightarrow{g'} \Omega_{g'gx}$ is the same as $\Omega_x \xrightarrow{g'g} \Omega_{g'gx}$; hence there exists a homomorphism $\Gamma : G \mapsto \mathrm{GL}(\mathbb{R}[\mathcal{F}_X]^*)$. We call this the group representation *associated* to the system $\mathcal{S}$. This naturally induces a representation $\Gamma^* : G \mapsto \mathrm{GL}(\mathbb{R}[\mathcal{F}_X])$, which is isomorphic to Γ since the representations are unitary and real.

We can summarise the above in the following theorem (fully proven in Appendix D), which is a straightforward generalisation of Result 1 of [51] to arbitrary dynamical structures:

Theorem 1 (Result 1 of [51]). For every system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ obeying 'Possibility of state estimation using a finite outcome set' there exists an embedding of $\mathcal{S}_X$ into a finite dimensional real vector space $V \cong \mathbb{R}[\mathcal{F}_X]^*$ and its dual V^* given by the following maps:

$$\Omega : X \to V \tag{5}$$
$$\Gamma : G \to \mathrm{GL}(V) \tag{6}$$
$$\Lambda : \mathcal{F}_X \to V^* \tag{7}$$

satisfying the following properties:

1. Preservation of dynamical structure:

$$\Gamma_g \Omega_x = \Omega_{gx} \tag{8}$$
$$\Gamma_{g_1} \Gamma_{g_2} = \Gamma_{g_1 g_2} \tag{9}$$

2. Preservation of probabilistic structure:

$$\Lambda_\mathbf{f} \cdot \Omega_x = \mathbf{f}(x) \tag{10}$$

3. Uniqueness: The embedding of $\mathcal{S}_X$ into (V, V^*) given by the maps Ω, Γ, Λ (satisfying all of the above) is unique up to equivalence.

   Two embeddings of $\mathcal{S}_X$ into (V, V^*) given by the maps Ω, Γ, Λ and $\Omega', \Gamma', \Lambda'$ are equivalent if there exists an invertible linear map $L : V \to V$ such that:

$$\Omega'_x = L\Omega_x, \ \forall x \in X, \tag{11}$$
$$\Gamma'_g = L\Gamma_g L^{-1}, \ \forall g \in G, \tag{12}$$
$$\Lambda'_\mathbf{f} = \Lambda_\mathbf{f} L^{-1}, \forall \mathbf{f} \in \mathcal{F}_X \,. \tag{13}$$

We call the representation Γ of Equation (6) the representation of G *associated* to the system $\mathcal{S}_X$. $\mathrm{conv}(\Omega_X)$ is the convex hull of the extremal states, which we call *state space* [1]. A standard representation of the states Ω_X is given as a vector of fiducial outcome probabilities.

---

[1]The object $\mathrm{conv}(\Omega_X) = \mathrm{conv}(\Gamma_G \Omega_x)$ is the convex hull of an orbit of the group G in the vector space V and is known as an *orbitope*, see [63] for a detailed study of the theory of orbitopes.

Remark 1. For a system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ with maps Ω, Γ, Λ the vectors Ω_X admit of a standard representation in terms of a fiducial outcome set (which is non-unique). A fiducial outcome set $\{\mathbf{f}_i\}_{i=1}^d$ is a linearly independent basis $\{\mathbf{f}_i\}_{i=1}^d$ of $\mathbb{R}[\mathcal{F}_X]$ (here $\dim(\mathbb{R}[\mathcal{F}_X]) = d$), i.e. it is a set of OPFs $\{\mathbf{f}_i\}_{i=1}^d$ such that every other $\mathbf{f} \in \mathcal{F}_X$ can be uniquely expressed as $\mathbf{f}(x) = \sum_{i=1}^d c_i \mathbf{f}_i(x)$ for all $x \in X$ and c_i coefficients in $\mathbb{R}$. In this representation a state Ω_x is written as:

$$\Omega_x = \begin{pmatrix} \mathbf{f}_1(x) \\ \vdots \\ \mathbf{f}_d(x) \end{pmatrix} , \tag{14}$$

and an effect $\Lambda_{\mathbf{f}}$ (where $\mathbf{f}(x) = \sum_{i=1}^d c_i \mathbf{f}_i(x)$) is a dual vector:

$$\Lambda_{\mathbf{f}} = (c_1, ..., c_d) \ .$$

One can immediately verify that $\Lambda_{\mathbf{f}} \cdot \Omega_x = \mathbf{f}(x)$.

In general the convex hull of a set of points P will not have that set of points as extremal points, since generically some points in P might lie in the convex hull of other points of P. As such it is not immediate that $\mathrm{conv}(\Omega_X)$ has extremal points Ω_X. The following lemma tells us that this is the case. Let us denote by $\delta_e(C)$ the extremal points of some convex set C.

Lemma 1. The embedding of a system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ into (V, V^*) with maps Ω, Γ, Λ is such that $\delta_e(\mathrm{conv}(\Omega_X)) = \Omega_X$.

This lemma is proven in Appendix D.2, and makes use of the transitivity of the action of G on the pure states X. It follows from the fact that Ω_X is a subset of a hypersphere in the affine span of Ω_X centred on the maximally mixed state. See also [63, Proposition 2.2] for an equivalent statement of the lemma and proof.

Definition 11 (Tomographically equivalent probabilistic structures). Two probabilistic structures $\mathcal{F}$ and $\mathcal{F}'$ are tomographically equivalent if they yield the same equivalence classes of ensembles (i.e. mixed states).

We note that two systems $\mathcal{S}_X = (X, G, \mathcal{F}_X)$ and $\mathcal{S}'_X = (X, G, \mathcal{F}'_X)$ with embeddings Ω, Γ, Λ and $\Omega', \Gamma', \Lambda'$ are tomographically equivalent if and only if Ω_X and Ω'_X are affinely isomorphic (i.e. equivalent as convex set).

For a given system the asymptotic limit consists of the scenario where all preparation procedures are of n copies of the same state and n tends to infinity. In this case all states (including mixed) become perfectly distinguishable (though this does not lift the degeneracy of equivalent ensembles). We denote $\bar{\mathcal{F}}_X$ the equivalence class of all tomographically equivalent probabilistic structures, hence $(X, \bar{\mathcal{F}}_X)$ can be identified with the state space (convex set) $\mathrm{conv}(\Omega_X)$ which is the same for all systems $(X, \mathcal{F}_X)$ with $\mathcal{F}_X \in \bar{\mathcal{F}}_X$. A representative element is the probabilistic structure corresponding to the (effect) unrestricted system.

Remark 2 (On the link between tomographically equivalent probabilistic structures and restriction of effects). The notion of tomographically equivalent probabilistic structures can be cast in terms of *restriction of effects*. A state space is *effect unrestricted* when all linear functionals $\mathrm{GL}(\mathbb{R}[\mathcal{F}_{G/H}]^*) \to [0,1]$ correspond to allowed measurement outcomes. A system is restricted when some of the mathematically allowed functionals do not represent any measurement outcomes of the theory. However when a system has restricted effects, it is always the case that the allowed effects span the dual space V^* of the state space embedded in V. In other words both the restricted and unrestricted systems have the same mixed states (the restricted effects are always such that they separate the initial state space). A system with restricted effects has a tomographically equivalent probabilistic structure to the unrestricted system. Two tomographically equivalent probabilistic structures can be obtained by restriction of a common probabilistic structure.

2.4.3 Topological and differentiable features of the probabilistic representation

In the case where the dynamical structure (X, G) has topological/differentiable features, we assume that the maps $\Omega : X \to V$ and $\Gamma : G \to \mathrm{GL}(V)$ are continuous/smooth, implying that the action $\Omega_x \to \Omega_{gx}$ is also continuous/smooth. In the cases where $\mathcal{F}_X$ separates X and G (implying that Ω and Γ are injections) the inverse maps Ω^{-1} and Γ^{-1} are also assumed to be continuous, implying that Ω_X and Γ_G are homeomorphic/diffeomorphic to X and G respectively.

For topological groups a group representation is a continuous homomorphism $\Gamma : G \to \mathrm{GL}(V)$ (smooth for Lie groups), hence the continuity/smoothness of Γ will allow us to make use of the representation theory of topological/Lie groups in the rest of the work. Since $\mathbb{R}[\mathcal{F}_X]$ is assumed to be finite dimensional continuity of these maps entails that the functions $\mathbf{f}$ are continuous on X.

Remark 3. The assumption that $\Omega : X \to V$ and $\Gamma : G \to \mathrm{GL}(V)$ are continuous/smooth is justified by the following. First note that continuity of these maps implies that Ω_X and Γ_G are isomorphic to X and G not just as set/groups, but as topological spaces/topological groups (differentiable manifolds/Lie groups). Consider the case where the dynamical structure has topological/differentiable structure but the maps Ω and Γ are not continuous/smooth, i.e. Ω_X and Γ_G are not homeomorphic/diffeomorphic to X and G respectively. Then given access only to the operational system, with state space $\mathrm{conv}(\Omega_X)$ and transformation space $\mathrm{conv}(\Gamma_G)$ and asked to reconstruct the dynamical structure, we would not assign it a dynamical structure with X a topological space acted on continuously by the group G (or differentiable manifold acted on smoothly by a Lie group). Rather we would assign it the set of pure states X without any topological structure. To summarise: the operational perspective begins from some experimental data, then constructs the convex state, transformation and effect spaces and only then can one infer the pure states and reversible transformations of those systems (i.e. dynamical structure). From this perspective the dynamical structure has all the structural properties of the experimentally determined Ω_X and Γ_G, implying that the maps Ω and Γ must preserve these structures.

2.5 Non-compactness and "Possibility of state estimation using a finite outcome set"

Although we have restricted our attention to probabilistic systems with X and G compact, there are well defined probabilistic systems which are non-compact such as infinite dimensional quantum systems. These systems violate "Possibility of state estimation using a finite outcome set" and a natural question to ask is whether "Possibility of state estimation using a finite outcome set" rules out non-compact dynamical structures in general. In the following we no longer assume X and G compact, and briefly explore this question.

Lemma 2. Under the assumption of "Possibility of state estimation using a finite outcome set" the sets $\mathrm{conv}(\Omega_X)$ and $\mathrm{conv}(\Gamma_G)$ are compact.

Observe however that compactness of $\mathrm{conv}(\Omega_X)$ and $\mathrm{conv}(\Gamma_G)$ does not imply compactness of Ω_X and Γ_G. There exists compact subset of $\mathbb{R}^n$ whose extremal points are not closed (for an example of such a set see [64, proof of Lemma 0.22]). As such one cannot use compactness of a convex set to infer compactness of its extremal points. Although there exist compact convex sets in $\mathbb{R}^n$ with a non-compact set of extremal points, it is not known to the authors whether any such sets where the extremal points are transitive under some group action exist. As such it may be the case that for transitive dynamical structures "Possibility of state estimation using a finite outcome set" imposes that X and G are compact.

In the case of transitive non-compact dynamical structures with a non-compact group one can make use of group representation theory to rule out the existence of probabilistic structure compatible with the assumption of "Possibility of state estimation using a finite outcome set". For many non-compact groups G (such as non-compact simple Lie groups) it is known that there are no non-trivial finite dimensional continuous unitary representations, which rules out probabilistic structures which violate "Possibility of state estimation using a finite outcome set" for dynamical structures with those groups. An open question is whether are there any transitive dynamical structures, where X and G are non-compact, which are consistent with "Possibility of state estimation using a finite outcome set".

8

3 Classification theorem

Before stating the main theorem of this section we will need to define the notion of a *Gelfand pair*. Here $(\Gamma, V, \mathbb{F})$ refers to a representation $\Gamma : G \mapsto \mathrm{GL}(V)$ over a field $\mathbb{F}$.

Definition 12 (Gelfand pair). A pair (G, H) with G a group and H a subgroup of G form a Gelfand pair when for all irreducible representations $(\Gamma, V, \mathbb{C})$ of G, the restriction $\Gamma_{|H}$ has at most one trivial sub-representation.

In other words, for a Gelfand pair (G, H), every irreducible representation $(\Gamma, V, \mathbb{C})$ of G is such that all the vectors $v \in V$ which are invariant under H span a subspace of dimension at most 1. A vector $v \in V$ is invariant under H when $\Gamma_h v = v$ for all $h \in H$.

This definition applies to complex irreducible representations. For irreducible representations over the field $\mathbb{R}$ the restriction $\Gamma_{|H}$ may contain two trivial sub-representations, however all H-invariant vectors are related by invertible transformations which commute with the group action (this does not contradict Schur's Lemma, which applies to irreducible representation over the complex field). More details and proofs can be found in Appendix C.

A representation $(\Gamma, V, \mathbb{C})$ of a group G which has a non-zero H-invariant vector (i.e. for which $\Gamma_{|H}$ contains a trivial sub-representation) is called a *spherical representation* of (G, H).

Theorem 2 (Classification theorem). Let $\mathcal{D} = (G, H)$ be a transitive dynamical structure, and let us consider probabilistic structures $\mathcal{F}_{G/H}$ such that $\mathbb{R}[\mathcal{F}_{G/H}]$ is finite-dimensional. By Theorem 1 every system $\mathcal{S}_{G/H} = (G, H, \mathcal{F}_{G/H})$ has an associated representation $\Gamma : G \to \mathrm{GL}(\mathbb{R}[\mathcal{F}_{G/H}]^*)$.

  i. Every probabilistic structure $\mathcal{F}_{G/H}$ (up to tomographic equivalence) has an associated representation Γ of the form:

$$\Gamma = \bigoplus_j \Gamma_j, \tag{15}$$

  where each term $(\Gamma_j, V_j, \mathbb{R})$ is a real-irreducible representation with at least one trivial sub-representation when restricted to H.

 ii. Conversely every representation of the form (15) (where each irreducible representation in the decomposition has at leat one trivial subrepresentation when restricted to H) is associated to at least one probabilistic structure $\mathcal{F}_{G/H}$.

iii. When (G, H) forms a Gelfand pair the correspondence between representations $(\Gamma, V, \mathbb{R})$ of the form (15) and probabilistic structures (up to tomographic equivalence) $\mathcal{F}_{G/H}$ is one-to-one.

 iv. When (G, H) does not form a Gelfand pair then some representations $(\Gamma, V, \mathbb{R})$ of the form (15) have infinitely-many tomographically inequivalent probabilistic structures $\mathcal{F}_{G/H}$ associated to them.

This theorem is proven in Appendix D.4.

Parts i. and iii. entail that for a dynamical structure (G, H) which form a Gelfand pair one can classify all possible probabilistic structures $\mathcal{F}$ (up to equivalence) by finding the irreducible representations Γ of G such that $\Gamma_{G|H}$ has a trivial representation.

Parts iii. and iv. tell us that for Gelfand pairs all inequivalent probabilistic structures are characterised by different representations of G. Therefore for Gelfand pairs all probabilistic structures are in one-to-one correspondence with representations of the dynamical group, up to restriction of effects. For non-Gelfand pairs there are inequivalent probabilistic structures which are associated to the same representation of G.

The one-to-one correspondence between probabilistic structures and representations for Gelfand pairs is a direct consequence of the existence of an invertible transformation which commutes with group action for all invariant H-vectors (see Corollary 5 in Appendix C). For real irreducible representations which are also complex irreducible this is just the identity (by Schur's lemma), however for real irreducible of complex type (i.e. which are complex reducible) the linear space of transformations which commutes with all H-invariant vectors is two dimensional. As shown

in Lemma 10 there are no real irreducible representations of quaternionic type which have an H-invariant vector when (G, H) Gelfand.

We observe that this theorem does not guarantee that for a given representation Γ of the form (15) the associated OPF set $\mathcal{F}$ separates the pure states. For instance the trivial representation $\Gamma : G \to \mathrm{GL}(\mathbb{R})$, $\Gamma(g) = \mathbb{I}_{\mathbb{R}}$ for all $g \in G$ is such that any vector $v \in \mathbb{R}$ is H-invariant, and the state space obtained for any choice of non-zero reference vector v is trivial: $\Omega_x = v$ for all $x \in X$.

4 Rigidity of dynamical structures

In this section we analyse which probabilistic structures can be continuously deformed [2]. We first study which dynamical structures (G, H) have probabilistic structures which are arbitrarily close. Following this we show that these probabilistic structures can be continuously deformed. In order to do so, we define an operational distance between probabilistic structures in terms of how difficult is to discriminate them.

Obviously, one can always smoothly deform a probabilistic structure by restricting the set of OPFs; for example, by adding noise to the measurements [3]. However, all these variants have the same set of mixed states, or in other words, the same equivalence classes of ensembles of pure states $\{(p_i, x_i)\}_i$. We call all these probabilistic structures *tomographically equivalent* because, in estimation processes with multiple measurements, they agree on the set of mixed states. In each tomographically-equivalent class of probabilistic structures there is a privileged element: the unrestricted probabilistic structure. This $\mathcal{F}$ includes all linear maps $\Lambda : V \to \mathbb{R}$ that map pure states to probabilities $\Lambda : \Omega(X) \to [0, 1]$. In order to avoid considering trivial deformations (i.e. those which leave the space of mixed states unchanged), in this section, we only consider unrestricted probabilistic structures.

A probabilistic structure $\mathcal{F}_0$ for which every other probabilistic structure $\mathcal{F}_1$ of the same linear dimension is at a finite bounded distance is called *rigid*. In other words, once the dimension of the space of mixed states is fixed, there is a finite bound on the minimal error when discriminating between probabilistic structures compatible with that dimension.

Theorem 2 tells us that if a dynamical structure (G, H) is a Gelfand pair then the set of unrestricted probabilistic structures is countable. We prove that each finite-dimensional probabilistic structure of a Gelfand pair (G, H) is rigid. We show that for non-Gelfand pairs there exists probabilistic structures $\mathcal{F}_0$ which are not rigid, and which can be continuously deformed to other probabilistic structures of the same linear dimension.

4.1 Distance between inequivalent probabilistic structures

For a given dynamical structure (G, H) (with $X \cong G/H$) there is a natural notion of distance between probabilistic structures $\mathcal{F}_X$. The distance between two OPFs $\mathbf{f}^0 \in \mathcal{F}_X^0$ and $\mathbf{f}^1 \in \mathcal{F}_X^1$ is given by:

$$\mathrm{dist}(\mathbf{f}^0, \mathbf{f}^1) = \max_{x \in X} |\mathbf{f}^0(x) - \mathbf{f}^1(x)| . \tag{16}$$

This distance is directly related to the minimal error made when discriminating between $\mathbf{f}^0$ and $\mathbf{f}^1$. We define the distance between two probabilistic structures $\mathcal{F}_X^0$ and $\mathcal{F}_X^1$ as:

$$D(\mathcal{F}^0, \mathcal{F}^1) = \max_{\mathbf{f}^0 \in \mathcal{F}_X^0} \min_{\mathbf{f}^1 \in \mathcal{F}_X^1} \mathrm{dist}(\mathbf{f}^0, \mathbf{f}^1) , \tag{17}$$

which informs us about the error that we make when certifying that a system behaves according to $\mathcal{F}_X^0$ and not $\mathcal{F}_X^1$ in the optimal experimental setting $\mathbf{f}^0 \in \mathcal{F}_X^0$. Note that D is not symmetric and hence it is not a metric distance. We introduce the symmetrised distance:

$$D_{\mathrm{sym}}(\mathcal{F}^0, \mathcal{F}^1) = \max\{D(\mathcal{F}^0, \mathcal{F}^1), D(\mathcal{F}^1, \mathcal{F}^0)\} , \tag{18}$$

---

[2]Here by continuity we do not mean in the sense of a continuous map between topological spaces, but rather in the sense that there exists a connected path between the two probabilistic structures being deformed.

[3]Adding noise to a measurement $\{\Lambda_{\mathbf{f}_i}\}_i$ consists of replacing each effect $\Lambda_{\mathbf{f}_i}$ by $(1 - \lambda)\Lambda_{\mathbf{f}_i} + \lambda\Lambda_{\mathbf{u}}$ for some noise parameter $\lambda \in [0, 1]$

which is a metric distance. The following theorem (proven in Appendix E.1) provides us with a lower bound on the distance between certain pairs of probabilistic structures $\mathcal{F}^0$ and $\mathcal{F}^1$.

Theorem 3. Let the dynamical structure (G, H) be a Gelfand pair, and let $\mathcal{F}^0$ and $\mathcal{F}^1$ be two unrestricted probabilistic structures of (G, H). If $\mathcal{F}^0$ has an irreducible representation of dimension d^0 which does not appear in $\mathcal{F}^1$ then

$$D(\mathcal{F}^0, \mathcal{F}^1) \geq \frac{1}{4d^0} \ . \tag{19}$$

Now we recall that for Gelfand systems, two unrestricted probabilistic structures are equal if and only if they have the same irreps in their decomposition. Hence, the above theorem implies that, for Gelfand systems, each pair of unrestricted probabilistic structure can be discriminated by finite means.

4.2 Rigid and non-rigid probabilistic structures

Theorem 3 tells us that, if we fix a dynamical structure consisting of a Gelfand pair (G, H), then the hypothesis that "the observed data is generated by a particular probabilistic structure $\mathcal{F}_0$ of (G,H)", in opposition to "the observed data is not generated by $\mathcal{F}_0$", can be tested with finite means. We now look at the property of rigidity of probabilistic structures, i.e. which probabilistic structures are such that every other probabilistic structure of the same linear dimension is at a finitely bounded distance.

Theorem 4. Let $\mathcal{F}^0$ of dimension $\dim\mathbb{R}[\mathcal{F}_0] = d_0$ be an unrestricted probabilistic structure of (G, H) with associated representation Γ_G.

i. If every pair of H-invariant vectors in $\mathbb{R}[\mathcal{F}^0]$ is related by an invertible transformation which commutes with Γ_G then $\mathcal{F}^0$ is rigid and any other inequivalent probabilistic structure $\mathcal{F}_1$ such that $\dim\mathbb{R}[\mathcal{F}_1] = d_0$ is at distance:

$$D_{\mathrm{sym}}(\mathcal{F}^0, \mathcal{F}^1) \geq \frac{1}{4(d_0 - 1)} \ , \tag{20}$$

ii. If there is a pair of H-invariant vectors in $\mathbb{R}[\mathcal{F}^0]$ that are not related by any invertible transformation which commutes with Γ_G then $\mathcal{F}^0$ is non-rigid and for any $\epsilon > 0$ ($\epsilon \ll 1$) there is an inequivalent probabilistic structure $\mathcal{F}^1$ with $\dim\mathbb{R}[\mathcal{F}_1] = d_0$ at distance

$$D_{\mathrm{sym}}(\mathcal{F}^0, \mathcal{F}^1) \leq 2\epsilon \ . \tag{21}$$

This theorem is proven in Appendix E.2.

For Gelfand pairs all irreducible spherical representations Γ_G are such that all pairs of H-invariant vectors are related by invertible transformations which commute with Γ_G, hence all probabilistic structures for Gelfand pairs are rigid. For non-Gelfand pairs there exist probabilistic structures $\mathcal{F}^0$ which have associated representations such that for all pairs of H-invariant vectors there is no invertible transformation relating them which commutes with Γ_G. Hence we have the following corollary:

Corollary 1. Let $\mathcal{D} = (G, H)$ be a dynamical structure.

1. If (G, H) is a Gelfand pair, then every unrestricted probabilistic structure $\mathcal{F}_{G/H}$ is rigid.

2. If (G, H) is not a Gelfand pair, then there exist probabilistic structures $\mathcal{F}_{G/H}$ which are not rigid, which are those with associated representations Γ_G which admit H-invariant vectors related by invertible transformations which do not commute with Γ_G.

In Lemma 4 we show that these probabilistic structures can be *continuously deformed* to other probabilistic structures of the same linear dimension. Deformation of probabilistic structure is defined in Section 4.4 where deformation maps between different probabilistic structures are explicitly characterised.

Before studying the general case we provide an example of a non-rigid probabilistic structure and how to continuously deform it.

4.3 Continuous deformation of probabilistic structures: an example

In this section we analyse a dynamical system (G, H) that is not a Gelfand pair. Hence, some of its probabilistic structures can be continuously deformed, giving rise to varying statistical properties. This is an interesting feature of GPTs that has not been explored in the literature.

Definition 13 (The deformable state space). Consider the set of pure states $\Omega_X = \{U\Omega_{x_0}U^\dagger : \forall U \in \mathrm{SU}(3)\}$ generated by the adjoint action of $G = \mathrm{SU}(3)$ on the reference state

$$\Omega_{x_0} = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}, \tag{22}$$

where the three real coefficients α_i are different ($\alpha_i \neq \alpha_j$ for all i, j) and add up to one $\sum_i \alpha_i = 1$. This state space has stabiliser subgroup

$$H = \left\{ \begin{pmatrix} e^{i\phi_1} & 0 & 0 \\ 0 & e^{i\phi_2} & 0 \\ 0 & 0 & e^{-i(\phi_1+\phi_2)} \end{pmatrix} : \forall \phi_1, \phi_2 \in \mathbb{R} \right\} \cong \mathrm{S}\left(\mathrm{U}(1) \times \mathrm{U}(1) \times \mathrm{U}(1)\right). \tag{23}$$

We observe that the embedding of the pure states is given $\Omega : X \to V$ where V is the real space of 3×3 Hermitian matrices. In this representation outcome probabilities are given by effects (i.e. linear functionals of the states Ω_x), hence they are given by the trace inner product: $\Lambda_{\mathbf{f}}(\Omega_x) = \mathrm{tr}\left(H_{\mathbf{f}}\Omega_x\right)$ where $H_{\mathbf{f}}$ a Hermitian matrix. The unit effect $\Lambda_{\mathbf{u}}$ is linear functional evaluating to 1 on all Ω_x, hence $\Lambda_{\mathbf{u}}(\Omega_x) = \mathrm{tr}(\Omega_x)$.

The case with two equal coefficients is equivalent to the familiar three-level quantum system, $(\alpha_1, \alpha_2, \alpha_3) = (1, 0, 0)$, which has stabiliser subgroup $H \cong \mathrm{S}\left(\mathrm{U}(2) \times \mathrm{U}(1)\right)$ [51] different than our example (23). This illustrates how we cannot deform quantum theory without changing its stabiliser subgroup and hence changing its dynamical structure (G, H). By contrast, the above-defined family of state spaces can be deformed without changing the stabiliser nor the dynamical structure.

The pair $(\mathrm{SU}(3), \mathrm{S}\left(\mathrm{U}(2) \times \mathrm{U}(1)\right))$ is a Gelfand pair [51], whereas $(\mathrm{SU}(3), \mathrm{S}\left(\mathrm{U}(1) \times \mathrm{U}(1) \times \mathrm{U}(1)\right))$ is not. To show this we need to find a single irreducible representation Γ of G such that the restriction $\Gamma_{|H}$ to the subgroup H has more than one trivial subrepresentation. Take the adjoint action of G (see Definition 13) acting on the full complex space of complex matrices; this representation decomposes into a trivial representation of G (acting on the subspace spanned by the identity matrix) and the adjoint representation acting on the complementary subspace. This subspace is spanned by the trace 0 complex matrices and carries an irreducible representation of $G \cong \mathrm{SU}(3)$ (the adjoint representation). We observe that all diagonal matrices are invariant under the adjoint action of the subgroup H defined in Equation (23), and as such the space of H-invariant vectors is spanned by the trace 0 diagonal matrices. This is a 2 dimensional subspace of the full space of trace 0 matrices carrying the irreducible representation of G implying that $(\mathrm{SU}(3), \mathrm{S}\left(\mathrm{U}(1) \times \mathrm{U}(1) \times \mathrm{U}(1)\right))$ is not a Gelfand pair.

We know that the three-level quantum system has three perfectly distinguishable states. The following theorem tells us that this is not the case for above-defined state spaces.

Theorem 5. All state spaces introduced in Definition 13 have two perfectly distinguishable states and no more.

Proof. In the following proof we write x instead of Ω_x (similarly X instead of Ω_X).

Let us start by assuming the existence of three perfectly distinguishable states $x_1, x_2, x_3 \in V$. This implies the existence of a three-outcome measurement $A_1, A_2, A_3 \in V$ such that $\mathrm{tr}(A_i x_j) = \delta_{ij}$. Without loss of generality we can take the three states to be pure $x_i \in X \subseteq V$. In the following analysis we use a V-basis where A_1 is diagonal

$$A_1 = \begin{pmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{pmatrix}. \tag{24}$$

The probability of A_1 with any state x only depends on the diagonal of the state (in this basis). Therefore, in what follows, we characterise the projection of $\text{conv}(X) \subseteq V$ into the diagonal. A general state $U x_0 U^\dagger$ has diagonal projection

$$[U x_0 U^\dagger]_{ii} = \sum_j U_{ij} \alpha_j [U^\dagger]_{ji} = \sum_j |U_{ij}|^2 \alpha_j \ . \tag{25}$$

The unitarity of U implies that $|U_{ij}|^2$ is a doubly stochastic matrix, and Birkhoff's theorem tells us that $|U_{ij}|^2$ is a mixture of the six permutation-matrices of three elements. Conversely, the six permutation matrices can be written as $|U_{ij}|^2$. This, together with the convexity of the state space, implies that the projection of $\text{conv}\,X$ into the diagonal is the convex set generated by the six extreme points

$$y_\sigma = \begin{pmatrix} \alpha_{\sigma(1)} \\ \alpha_{\sigma(2)} \\ \alpha_{\sigma(3)} \end{pmatrix}, \ \sigma \in \mathcal{S}_3 \ , \tag{26}$$

where $\mathcal{S}_3$ is the group of permutations of 3 elements. These six points (also denoted $y_1, y_2, \ldots, y_6$) are depicted in Figure 1.

Let us show that each of the three pairs of opposite lines in Figure 1 are indeed parallel. For example:

$$y_1 - y_2 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} - \begin{pmatrix} \alpha_1 \\ \alpha_3 \\ \alpha_2 \end{pmatrix} \propto \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \tag{27}$$

$$y_6 - y_3 = \begin{pmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_1 \end{pmatrix} - \begin{pmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_3 \end{pmatrix} \propto \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \tag{28}$$

$$y_5 - y_4 = \begin{pmatrix} \alpha_3 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} - \begin{pmatrix} \alpha_3 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} \propto \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{29}$$

Condition $\text{tr}(A_1 x_j) = \delta_{1j}$ implies that the scalar product $(\gamma_1, \gamma_2, \gamma_3) \cdot y_\sigma$ takes the value zero for two permutations σ and the value 1 on at least one permutation. However, as shown in Figure 1, the only outcome that tells apart states y_1 from y_4, y_5 is A_1, which gives probability one for states y_1, y_2 and zero for y_4, y_5. It is worth mentioning that the vectors y_1, y_2, y_4, y_5 correspond to four pure states with zero off-diagonal components. Hence, the outcome probabilities of these pure states can be calculated by only looking at the diagonal projection (i.e. the figure).

The figure also shows that, no matter how we choose the direction A_2, the states y_4, y_5 (or y_1, y_2) cannot be perfectly distinguished. This proves the non-existence of three perfectly distinguishable states in this family of state spaces. Of course, this argument breaks down for the three-level quantum system, when the projection becomes a triangle instead of an hexagon. $\qquad\square$

The state spaces under consideration (Definition 13) have a remarkable property that is not present in quantum theory. This property is sometimes called "violation of no-simultaneous encoding" [65] and it is very similar to "information causality" [66]. This property allows to perfectly encode one bit of information (e.g. y_1, y_2 versus y_4, y_5) and simultaneously imperfectly encode another bit (y_1, y_5 versus y_2, y_4). Although only one of the two bits can be retrieved, there is a sense in which this system encodes more than one bit of information despite having only two perfectly-distinguishable states. Different choices of $(\alpha_1, \alpha_2, \alpha_3)$ will give different success probability when optimally guessing the second bit. This is a statistical feature that distinguishes inequivalent values of $(\alpha_1, \alpha_2, \alpha_3)$ within the family of state spaces of Definition 13.

4.4 Continuous deformation of probabilistic structures: the general case

In the following we only consider systems up to tomographic equivalence, where two systems are tomographically equivalent if and and only if they have the same equivalence classes of ensembles (i.e. the same mixed states). For a given set of pure states X each equivalence class of probabilistic
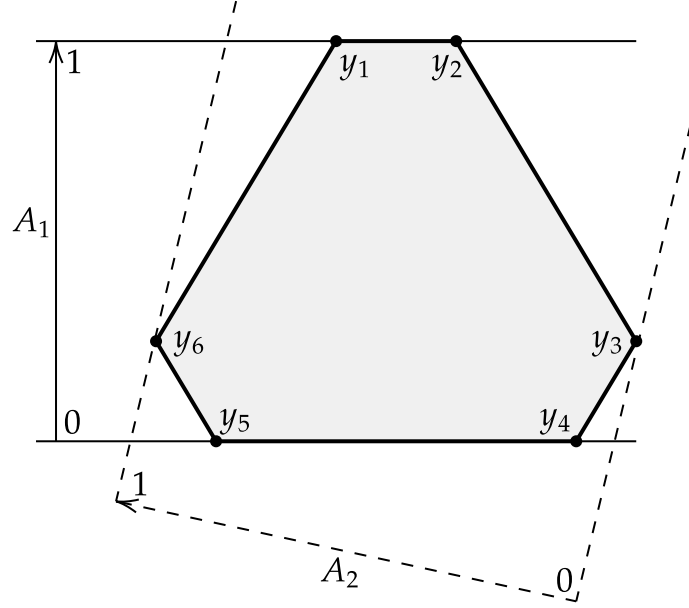
Figure 1: This six-sided figure is the projection onto the diagonal of a generic state space of the family introduced in Definition 13. The extreme points $y_1, \ldots, y_6$ are the six permutations of the vector $(\alpha_1, \alpha_2, \alpha_3)$, which can be represented in a two-dimensional plane because of the normalisation condition $\sum_i \alpha_i = 1$. In the generic case $(\alpha_1, \alpha_2, \alpha_3)$ the figure has two types of sides with alternating length, and for certain values of $(\alpha_1, \alpha_2, \alpha_3)$ the six sides have equal length. In the quantum case $(\alpha_1, \alpha_2, \alpha_3) = (1, 0, 0)$ the short sides have zero length zero and the figure becomes a triangle. The outcome A_1 (A_2) has probability 0 or 1 in the solid (dotted) lines.

structures $\bar{\mathcal{F}}_X$ induces a map $\Omega : X \to V$ (where $V \cong \mathbb{R}^n$) by Theorem 1. We will define a deformation of a probabilistic structure for X as a map between different images Ω_X of Ω-maps for X (since there is a one-to-one correspondence between equivalence classes of probabilistic structures $\bar{\mathcal{F}}_X$ and images Ω_X of Ω-maps) satisfying certain conditions which we will formalize with the aid of some further definitions. We only consider equivalence classes of probabilistic structures which separate X.

For a given dynamical structure (X, G) let us call $\mathfrak{F}$ the space of all images Ω_X (equivalently the space of equivalence classes of probabilistic structures $\bar{\mathcal{F}}_X$). For a given $d \in \mathbb{R}_{\mathbf{0}}^+$ let us call $\mathfrak{F}_d$ the space of all Ω_X whose linear span is isomorphic to $\mathbb{R}^d$ (equivalently the space of equivalence classes of probabilistic structures $\bar{\mathcal{F}}_X$ such that $\mathbb{R}[\bar{\mathcal{F}}_X] \cong \mathbb{R}^d$). We observe that $\mathfrak{F}_d$ might be the empty set (if there are no representations of G of the form of Equation (15) acting on $\mathbb{R}^d$). For a Gelfand dynamical structure all $\mathfrak{F}_d$ are countable sets, whilst for non-Gelfand pairs there exist $\mathfrak{F}_d$ which are uncountable.

The symmetrised distance defined in Equation (18) turns $\mathfrak{F}_d$ into a metric space. This metric induces a topology on $\mathfrak{F}_d$.

We are interested in continuously deforming probabilistic structures, namely how to continuously vary one set of mixed states into another (i.e. one probabilistic structure into another) whilst keeping the linear dimension fixed. First we define a general transformation between sets of embedded pure states Ω_X.

Definition 14 (Trivial Ω-transformation map). For two systems $\mathcal{S}_X^0 = (X, \bar{\mathcal{F}}_X^0)$ and $\mathcal{S}_X^1 = (X, \bar{\mathcal{F}}_X^1)$ with associated maps $\Omega^0 : X \to V^0$ and $\Omega^1 : X \to V^1$ we define the trivial Ω-transformation map $M_{0 \to 1} : \mathfrak{F} \to \mathfrak{F}$, $M_{0 \to 1} : \Omega_X^0 \mapsto \Omega_X^1$ (where it acts as $M_{0 \to 1} : \Omega_x^0 \mapsto \Omega_x^1$, $\forall x \in X$).

These trivial Ω-transformation maps allow us to map between any two probabilistic structures in $\mathfrak{F}$. A deformation of probabilistic structure is a specific map between probabilistic structures such that the linear dimension of the space of mixed states is unchanged.

Definition 15 (Deformation map). For two systems $\mathcal{S}_X^0 = (X, \bar{\mathcal{F}}_X^0)$ and $\mathcal{S}_X^1 = (X, \bar{\mathcal{F}}_X^1)$ with

associated maps $\Omega^0 : X \to V^0$ and $\Omega^1 : X \to V^1$ where $V^0 \cong V^1 \cong \mathbb{R}_d$ a trivial Ω-transformation map $M_{0\to1} : \mathfrak{F}_d \to \mathfrak{F}_d$, $M_{0\to1} : \Omega_X^0 \mapsto \Omega_X^1$ is a deformation map.

Lemma 3. A non-trivial deformation map $M_{0\to1} : \Omega_X^0 \mapsto \Omega_X^1$ ($\Omega_X^0 \neq \Omega_X^1$) cannot be extended to a linear transformation on $\text{span}_{\mathbb{R}}(\Omega_X^0)$.

Proof. Since $\bar{\mathcal{F}}_X^0$ and $\bar{\mathcal{F}}_X^1$ separate X, Ω_X^0 and Ω_X^1 are homeomorphic. Let us prove by contradiction and assume that $M_{0\to1}$ is linear when extended to $\text{span}_{\mathbb{R}}(\Omega_X^0)$. Let us call $\tilde{M}_{0\to1} : \text{span}_{\mathbb{R}}(\Omega_X^0) \to \text{span}_{\mathbb{R}}(\Omega_X^1)$ this linear extension of $M_{0\to1}$. We observe that $\tilde{M}_{0\to1}$ has a trivial kernel since no element of Ω_X^0 is in the kernel, hence no element of $\text{span}_{\mathbb{R}}(\Omega_X^0)$ is either. Moreover the image of $\tilde{M}_{0\to1}$ is $\text{span}_{\mathbb{R}}(\Omega_X^1)$: any element in $\text{span}_{\mathbb{R}}(\Omega_X^1)$ is a linear combination $\sum_i \alpha_i \Omega_{x_i}^1$ of elements in Ω_X^1, and hence by linearity of $\tilde{M}_{0\to1}$ $\sum_i \alpha_i \Omega_{x_i}^1$ is the image under $\tilde{M}_{0\to1}$ of the linear combination $\sum_i \alpha_i \Omega_{x_i}^0$ in the domain $\text{span}_{\mathbb{R}}(\Omega_X^0)$. Hence $\tilde{M}_{0\to1}$ is an invertible linear transformation which takes $\text{conv}(\Omega_X^0)$ to $\text{conv}(\Omega_X^1)$ and $\text{conv}(\Omega_X^0) \cong \text{conv}(\Omega_X^1)$. This contradicts the assumption that Ω^0 and Ω^1 are induced by tomographically inequivalent probabilistic structures. $\square$

A non-trivial deformation map $M_{0\to1}$ cannot be extended to the spaces of mixed states $\text{conv}(\Omega_X^0)$ and $\text{conv}(\Omega_X^1)$, which are not isomorphic as convex sets. For such systems one can use the map $M_{0\to1}$ to deform the probabilistic structure $\mathcal{F}_X^0$ to $\mathcal{F}_X^1$ to change the space of mixed states from $\text{conv}(\Omega_X^0)$ to $\text{conv}(\Omega_X^1)$ whilst preserving the linear dimension.

We observe that if we consider the general trivial Ω-transformation maps (i.e. the linear dimension of the two state spaces is no longer required to be the same) then some of these maps do linearly extend to $\text{span}_{\mathbb{R}}(\Omega_X^0)$ when $\Omega_X^0 \neq \Omega_X^1$. An example of such a map is the projection from a system $\mathcal{S}_X^0$ with an associated representation Γ^0 which is reducible to a system $\mathcal{S}_X^1$ with an associated representation Γ^1 which is a sub-representation of Γ^0, i.e. $\Gamma^0 \cong \Gamma^1 \oplus \Gamma^2$ for some representation Γ^3 of G.

Definition 16 (Continuous deformation of probabilistic structure). A probabilistic structure $\bar{\mathcal{F}}_X^0$ can be continuously deformed to another probabilistic structure $\bar{\mathcal{F}}_X^1$ if there is a continuous function $\gamma : [0,1] \to \mathfrak{F}_d$, $\gamma(t) = \Omega_X^t$ in $\mathfrak{F}_d$ with $t \in [0,1]$ such that $\gamma(0) = \Omega_X^0$ and $\gamma(1) = \Omega_X^1$. Here continuity is defined with respect to the usual topology on $[0,1] \subset \mathbb{R}$. γ defines a connected path in $\mathfrak{F}_d$.

We observe that the connected path $\gamma(t)$ defines a family of deformation maps $M_{0\to t} : \Omega_X^0 \mapsto \Omega_X^t$ for every $t \in [0,1]$.

In other words Ω_X^0 can be continuously deformed to Ω_X^1 if there exists a connected path between the two passing through a continuum of different Ω_X^t (each spanning a linear space of the same dimension). If there is no such path between the two then Ω_X^0 cannot be continuously deformed to Ω_X^1.

In the following theorem we show the relation between the rigidity of a probabilistic structure and the possibility of continuously deforming a probabilistic structure.

Lemma 4. Any non-rigid probabilistic structure $\bar{\mathcal{F}}_{G/H}^0$ can be continuously deformed into another (non-rigid) probabilistic structure $\bar{\mathcal{F}}_1$. Any rigid probabilistic structure $\bar{\mathcal{F}}_{G/H}^0$ cannot be continuously deformed to another probabilistic structure.

This lemma is proven in Appendix E.3. Together with Corollary 1 this lemma implies the following corollary.

Corollary 2. If (G,H) is a Gelfand pair, then no probabilistic structure can be continuously deformed. If (G,H) is a non-Gelfand pair then it has probabilistic structures which can be continuously deformed.

We now provide an explicit construction of a continuous deformation of a probabilistic structure whose associated representation has inequivalent H-invariant vectors; for an arbitrary non-Gelfand pair (G,H).

Example 1 (Continuous deformation). Consider a non-Gelfand pair (G,H), and an irreducible spherical representation Γ acting on $V \cong \mathbb{R}^d$ with inequivalent H-invariant vectors v_H^0 and v_H^1 (which always exists for a given non-Gelfand pair). Any vector in $\text{span}(v_H^0, v_H^1)$ is H-invariant,

moreover any two vectors in $\text{span}(v_H^0, v_H^1)$ which are not proportional will generate inequivalent state spaces under Γ_G. Let us take two inequivalent normalised H-invariant vectors v_H^0 and v_H^1. Let us call $\Omega_X^0 = \{\Gamma_g v_H^0 | g \in G\}$ and $\Omega_X^1 = \{\Gamma_g v_H^1 | g \in G\}$ the two embedded sets of pure states generated from these two reference vectors. Consider a rotation $R \in \text{GL}(\text{span}(v_H^0, v_H^1))$ such that $R v_H^0 = v_H^1$. For a given state $v_{gH}^0 = \Gamma_g v_H^0 \in \Omega_X^0$ and $v_{gH}^1 = \Gamma_g v_H^1 \in \Omega_X^1$ we have the relation:

$$v_{gH}^1 = \Gamma_g R v_H^0 = \Gamma_g R \Gamma_{g^{-1}} v_{gH}^0 \ . \tag{30}$$

Hence there is a map $m(g) = \Gamma_g R \Gamma_{g^{-1}}$ between the two points $v_{gH}^0 \in \Omega_X^0$ and $v_{gH}^1 \in \Omega_X^1$, for all points $gH \in X \cong G/H$. The deformation map is then:

$$M_{0 \to 1} : \Omega_X^0 \to \Omega_X^1 \ , \tag{31}$$

$$M_{0 \to 1} : v_{gH}^0 \mapsto m(g) v_{gH}^0 = v_{gH}^1 \ . \tag{32}$$

Since this map is g-dependent (i.e. its action depends on each extremal point v_{gH}^0) it is not linear and cannot be extended to the mixed states.

Let us define Ω_X^t to be the state space generated by the H-invariant reference vector $v_H^t = R(t) v_H^0$ for $t \in [0, 1]$ and $m^t(g) = \Gamma_g R(t) \Gamma_{g^{-1}}$. Then the deformation map from Ω_X^0 to Ω_X^t is given by:

$$M_{0 \to t} : \Omega_X^0 \to \Omega_X^t \ , \tag{33}$$

$$M_{0 \to t} : v_{gH}^0 \mapsto m^t(g) v_{gH}^0 = v_{gH}^t \ . \tag{34}$$

The family of deformation maps $M_{0 \to t}$ (for $t \in [0, 1]$) defines a connected path $\gamma(t) = M_{0 \to t} \Omega_X^0 = \Omega_X^t$ in $\mathfrak{F}_d$ (see the proof of Lemma 4 in Appendix E.3). Hence the probabilistic structure associated to Ω_X^0 can be continuously deformed to the probabilistic structure associated to Ω_X^1.

Remark 4. The possibility of continuously deforming a probabilistic structure without altering its dynamical structure (and without restricting effects) is a very peculiar feature that is not found in any of the known GPTs (such as boxworld and quantum theory over the field of reals, complex or quaternions) to the best of our knowledge. Moreover we posit that this is a typical feature of GPT systems, in that, most dynamical structures (G, H) are not Gelfand pairs. If a probabilistic structure can be continuously deformed then the probabilities can be fine-tuned to suitably describe the observed statistics, and hence, make the theory more difficult to falsify. Hence, the fact that the probabilistic structure of a theory cannot be smoothly deformed makes the falsifiability of the theory more straightforward. We believe that this is a desirable property of a theory. If we consider a dynamical structure (G, H) being a Gelfand pair then we can be sure that any of its probabilistic structures will be straightforwardly falsifiable. Finally, it is important to mention that a dynamical structure (G, H) cannot be continuously deformed due to the group and sub-group structures of G and H. That is, adding a single element to G or H will generate lots of new elements via products and inverses. And hence, the probabilistic structure is the only part of a theory that, a priori, could be continuously deformed.

5 Gelfand pairs and two point homogeneity

In Theorem 2 we have singled out dynamical structures corresponding to Gelfand pairs as being of interest, namely for the convenient property that their probabilistic structures can be classified via the associated group representations. This implies that there are countably many of them, and that they cannot be continuously deformed. This rigidity is a highly desirable property for a fundamental theory of physics, because it does not allow for *ad hoc* parameter adjustment, and is thereby easier to falsify. Apart from this, one may also ask whether there are other informational/physical motivations for considering Gelfand pairs. One such reason may be the following.

Definition 17 (Two-point homogeneous action [67])**.** A group G acts two-point homogeneously on a metric space (X, dist) if for every pair of points (x_1, x_2) and (x_1', x_2') in X with $\text{dist}(x_1, x_2) = \text{dist}(x_1', x_2')$ there is an element $g \in G$ such that $g x_1 = x_1'$ and $g x_2 = x_2'$.

X	G/H
S^d	$\mathrm{SO}(d)/\mathrm{SO}(d-1)$
$\mathrm{P}\mathbb{R}^d = S^d/\{\pm I\}$	$\mathrm{O}(d)/(\mathrm{O}(d-1) \times \mathrm{O}(1)) \cong \mathrm{SO}(d)/\mathrm{S}(\mathrm{O}(d-1) \times \mathrm{O}(1))$
$\mathrm{P}\mathbb{C}^d$	$\mathrm{SU}(d)/\mathrm{S}(\mathrm{U}(d-1) \times \mathrm{U}(1)) \cong \mathrm{U}(d)/(\mathrm{U}(d-1) \times \mathrm{U}(1)$
$\mathrm{P}\mathbb{H}^d$	$\mathrm{Sp}(d)/(\mathrm{Sp}(d-1) \times \mathrm{Sp}(1))$
$\mathrm{P}\mathbb{O}^3$	$\mathrm{F}(4)/\mathrm{Sp}(9)$

Table 1: List of all two-point homogeneous spaces which are connected and compact. Here $\mathrm{P}\mathbb{O}^3$ is the projective space of octonionic planes known as the Cayley plane.

Two-point homogeneity implies transitivity, since for any points x_1 and x_2 we have $\mathrm{dist}(x_1, x_1) = \mathrm{dist}(x_2, x_2)$ and hence there exists an element such that $gx_1 = x_2$. The following is a very remarkable result.

Lemma 5 (Prop 2.2 [68])**.** If G acts two-point homogeneously on a metric space X and H is the stabilizer of a point, then (G, H) is a Gelfand pair.

The requirement of two-point homogeneity restricts us to dynamical structures corresponding to Gelfand pairs. We observe that this requirement requires an additional metric structure to be imposed on the dynamical structure. A natural metric on GPT state spaces is the following.

Lemma 6. (Pure-state metric distance.) The distance between any pair of pure states $x, x' \in X$ defined by

$$\mathrm{dist}(x, x') = \sup_{\mathbf{f} \in \mathcal{F}_X} [\mathbf{f}(x) - \mathbf{f}(x')] \tag{35}$$

is bounded $\mathrm{dist}(x, x') \leq 1$, it satisfies the metric axioms:

1. $\mathrm{dist}(x, x') = 0$ if and only if $x = x'$,

2. $\mathrm{dist}(x, x') = \mathrm{dist}(x', x)$,

3. $\mathrm{dist}(x, x'') \leq \mathrm{dist}(x, x') + \mathrm{dist}(x', x'')$,

and it is G-invariant

4. $\mathrm{dist}(gx, gx') = \mathrm{dist}(x, x')$ for all $g \in G$.

Therefore we conclude that two-point homogeneous state spaces are Gelfand pairs. It is remarkable that the purely dynamical property of two-point homogeneity implies that all probabilistic structures are rigid.

However we note that not all Gelfand pairs (G, H) give rise to a homogeneous space $X \cong G/H$ which is two-point homogeneous. Indeed the classification of all the compact and connected two point homogeneous symmetric spaces was given in [67]. These are listed in Table 1.

The full classification of all finite dimensional probabilistic structures for the compact connected two point homogeneous spaces G/H (where all pairs (G, H) corresponding to such spaces are given in Table 1) directly follows from the classification of all irreducible spherical representations. Equivalently these are the irreducible subspaces of the function space $C(G/H, \mathbb{C})$ (continuous functions from G/H to $\mathbb{C}$) under the action of G, where a specific basis for an irreducible subspace is given by *spherical harmonics*. This is a generalisation of the well known spherical harmonics for $L^2(S^2)$, where the irreducible representation labelled by l has a basis $Y_{lm}(\theta, \phi)$ spanning a $2l + 1$ dimensional subspace.

The (G, H) spherical irreducible representations for these pairs are characterised by a condition on the highest weights given by the *Cartan-Helgason Theorem* [69, 70] (see [71, Theorem 11.4.10.]). Explicit characterisations of these (G, H) spherical irreducible representations (either in terms of the highest weights or other methods) for the pairs in Table 1 can be found in the literature.

6   Grassmannian systems

In this section we introduce a family of non-classical systems which generalise the dynamical structure of quantum systems, and make use of Theorem 2 to provide a full classification of these systems.

The pure states of finite dimensional quantum systems are given by $\mathrm{P}\mathbb{C}^d$. This is the set of all one dimensional subspaces of $\mathbb{C}^d$. We now consider systems with pure states given by the set of all k-dimensional subspaces $W \subseteq \mathbb{C}^d$. This set is known as a Grassmann manifold $\mathrm{Gr}(k, \mathbb{C}^d)$:

$$\mathrm{Gr}(k, \mathbb{C}^d) = \{W \subseteq \mathbb{C}^d, \dim(W) = k\} \ . \tag{36}$$

Hence $\mathrm{P}\mathbb{C}^d \cong \mathrm{Gr}(1, \mathbb{C}^d)$. Since $\mathrm{SU}(d)$ acts transitively on $\mathrm{Gr}(k, \mathbb{C}^d)$ it can also be expressed as follows (re-parametrising $k = m$ and $d = m + n$):

$$\mathrm{Gr}(m, \mathbb{C}^{m+n}) \cong \mathrm{SU}(m+n)/\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)) \tag{37}$$

Here the embedding of $\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n))$ into $\mathrm{SU}(m + n)$ is the direct sum embedding:

$$\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)) = \left\{ \begin{pmatrix} e^{i\theta} A & 0 \\ 0 & B \end{pmatrix} : A \in \mathrm{U}(m), \ B \in \mathrm{U}(n), \ e^{i\theta m} \det A \det B = 1 \right\} \ . \tag{38}$$

Similarly one can define Grassmann manifolds over $\mathbb{R}$ and $\mathbb{H}$, generalising the dynamical structures of quantum theory over $\mathbb{R}$ and $\mathbb{H}$. These are:

$$\mathrm{Gr}(m, \mathbb{R}^{m+n}) \cong \mathrm{SO}(m+n)/\mathrm{S}(\mathrm{O}(m) \times \mathrm{O}(n)) \ , \tag{39}$$

$$\mathrm{Gr}(m, \mathbb{H}^{m+n}) \cong \mathrm{Sp}(m+n)/(\mathrm{Sp}(m) \times \mathrm{Sp}(n)) \ . \tag{40}$$

In the next section we will make use of Theorem 2 to classify all possible probabilistic structures for each dynamical structure which is a complex Grassmann manifold.

6.1   Full classification of all probabilistic structures for complex Grassmann manifolds

Theorem 2 states that for a dynamical structure (G, H) corresponding to a Gelfand pair, every probabilistic structure is in one-to-one correspondence with a spherical representation of (G, H). Hence the first step in classifying probabilistic structures for the Grassmann dynamical structure $\mathrm{Gr}(m, \mathbb{C}^{m+n}) \cong \mathrm{SU}(m+n)/\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n))$ is to determine whether $(\mathrm{SU}(m + n), \mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)))$ form a Gelfand pair.

Lemma 7.

1. $(\mathrm{SU}(m + n), \mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)))$ form a Gelfand pair.

2. The spherical representations $(\mathrm{SU}(m+n), \mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)))$ have a real structure.

The first part of the lemma is found in [72, Corollary 3] and the second part is proven in Appendix F. This lemma entails (using Theorem 2) that all probabilistic structures $\mathcal{F}_X$ where $X \cong \mathrm{SU}(n + m)/\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n))$ are in one-to-one correspondance with the spherical representations $(\mathrm{SU}(m + n), \mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)))$. Irreducible spherical representations are typically defined over $\mathbb{C}$, and in general the irreducible representations of a group G over $\mathbb{C}$ are not in one-to-one correspondence with those over $\mathbb{R}$. Part 2. of the lemma allows us to classify the real irreducible spherical representations of $(\mathrm{SU}(m + n), \mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n)))$ by studying the irreducible spherical representations over $\mathbb{C}$.

The restriction of representations of $\mathrm{SU}(m + n)$ to $\mathrm{S}(\mathrm{U}(m) \times \mathrm{U}(n))$ has been studied in [72]. We summarise the result below. Representations of $\mathrm{SU}(m + n)$ are labelled by a partition λ of an integer k in $m + n - 1$ parts (often represented as a *Young diagram*). One can construct the associated irreducible representation by applying the *Schur functor* $\mathbb{S}_\lambda$ to $(\mathbb{C}^{m+n})$ [73, Proposition 15.15]

Lemma 8. The representations of $\mathrm{SU}(n + m)$ which have a trivial representation when restricted to $\mathrm{S}(\mathrm{U}(n) \times \mathrm{U}(m))$ are of the following form: When $m = n$:

$$\lambda = (2b_1, b_1 + b_2, ..., b_1 + b_m, b_1 - b_m, ..., b_1 - b_2, 0) \ . \tag{41}$$

When $n \geq m + 1$:

$$\lambda = (2b_1, b_1 + b_2, ..., b_1 + b_m, \underbrace{b_1, ..., b_1}_{n-m \text{ times}}, b_1 - b_m, ..., b_1 - b_2, 0) \,. \tag{42}$$

In each case $b_1 \geq b_2 \geq ... \geq b_m \geq 0$. We have added a redundant 0 entry; with it λ has length $m + n$.

6.2 Quartic quantum theory over $\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$

Quartic quantum theory over $\mathbb{C}$, introduced in [47], is a theory which contains some of the systems classified above. In this theory systems $\mathcal{S}_{k,\mathbb{C}}^{\texttt{Quart}}$ ($k \in \mathbb{Z}, k > 2$) have pure states given by the Grassman manifold $\text{Gr}(k, \mathbb{C}^{k^2})$ and a probabilistic structure $\mathcal{F}_{k,\mathbb{C}}^{\texttt{Quart}}$ given by the adjoint representation. For example the state space for the system $k = 2$ can be generated by taking reference state:

$$\rho = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \,, \tag{43}$$

applying the SU(4) dynamical group in the adjoint representation:

$$U \rho U^\dagger \,, \tag{44}$$

and taking the convex hull of the $\text{Gr}(2, \mathbb{C}^4)$ manifold embedded in $V = \text{Herm}\left(\mathbb{C}^4\right)$ (the real linear space of Hermitian matrices on $\mathbb{C}^4$).

One problematic feature of quartic quantum theory is that it does not have well defined composition [25, 47], and as such is just a collection of systems rather than a full theory.

We can similarly introduce two theories (without composition): *real quartic quantum theory* and *quaternionic quartic quantum theory* where systems are given by $\mathcal{S}_{k,\mathbb{R}}^{\texttt{Quart}} := \left(\text{Gr}(k, \mathbb{R}^{k^2}), \mathcal{F}_{k,\mathbb{R}}^{\texttt{Quart}}\right)$ and $\mathcal{S}_{k,\mathbb{H}}^{\texttt{Quart}} = \left(\text{Gr}(k, \mathbb{H}^{k^2}), \mathcal{F}_{k,\mathbb{H}}^{\texttt{Quart}}\right)$, where $\mathbb{R}[\mathcal{F}_{k,\mathbb{R}}^{\texttt{Quart}}]$ and $\mathbb{R}[\mathcal{F}_{k,\mathbb{H}}^{\texttt{Quart}}]$ are acted on by the adjoint representation of $\text{SO}(k^2)$ and $\text{Sp}(k^2)$ respectively. In both cases the states space for the system associated to $k = 2$ can be generated by taking the reference state ρ above acting with the adjoint representation of the dynamical group and taking the convex hull.

7 Discussion

7.1 Comparison to previous work

7.1.1 The generalised quantum mechanics of Mielnik

The OPF framework presented in this work is similar to the 'Group theoretical model' of [10]. The novel aspects of this work include Theorem 2 which, building on the framework, establishes a correspondence between probabilistic structures and group representations. We find specific conditions on transitive dynamical structures which make this correspondence one-to-one (namely that the dynamical group and stabilizer subgroup form a Gelfand pair). Moreover Mielnik studies examples with the same pure states as quantum theory, but different dynamical groups. We study (and classify) systems which have different pure states and dynamics.

7.1.2 Classification of all alternatives to the measurement postulates of quantum system

Theorem 2 is a generalisation of the classification theorem of [51], where the dynamical structure is no longer constrained to be that of quantum systems. We also find the necessary and sufficient conditions for which dynamical structures have probabilistic structures which are in one-to-one correspondence with group representations.
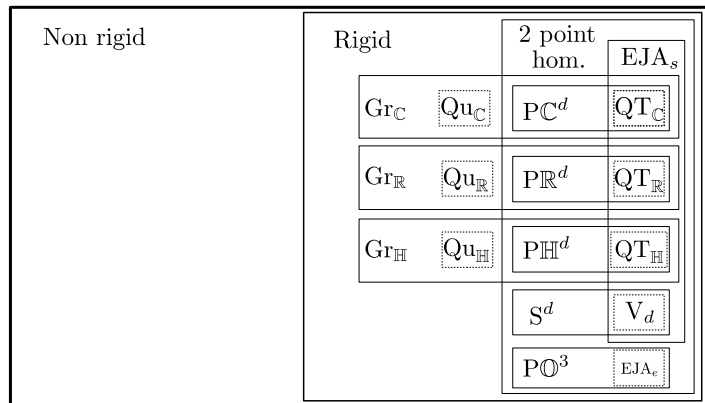
Transitive systems with compact pure states

Figure 2: Map of the space of transitive non-classical systems with compact pure states $X = G/H$. 'Rigid' and 'Non rigid' are notions defined in this paper. '2 point hom.' stands for two point homogeneous. For a field $\mathbb{F}$, $\mathrm{Gr}_{\mathbb{F}}$ is the family of systems with pure states given by the Grassmann manifold $\mathrm{Gr}(\mathbb{F}^d, \mathbb{F}^k)$ for all $2 < d < \infty$, $k < d$. $\mathrm{P}\mathbb{F}^d$ is the family of systems with pure states given by projective space over $\mathbb{F}^d$ for all $1 < d < \infty$, hence $\mathrm{P}\mathbb{F}^d := \mathrm{Gr}(\mathbb{F}^d, \mathbb{F}^1)$. $\mathrm{QT}_{\mathbb{F}}$ is quantum theory over $\mathbb{F}$ whilst $\mathrm{Qu}_{\mathbb{F}}$ is quartic quantum theory over $\mathbb{F}$. 'EJA$_s$' labels special Euclidean Jordan Algebras (EJA) and 'EJA$_e$' the exceptional EJA. V_d is the $d-$sphere in the standard embedding in $\mathbb{R}^{d+1}$ whilst S^d is the family of systems with pure states given by S^d (hence embeddings of S^d in $\mathbb{R}^k$ where k not necessarily equal to $d+1$). This map does not capture all the relations, namely there are 'coincidences' like the qubit being both in $\mathrm{QT}_{\mathbb{C}}$ and V_d.

7.2 Mapping the space of GPT systems

In the GPT formalism a *theory* is considered to be a set of systems together with some composition rules. Quantum theory for example is the set of systems $\mathrm{QT}_{\mathbb{C}} := \{\mathcal{S}_d^{\mathtt{Quant}}\}_{d=2}^{\inf}$ together with the standard tensor product composition rule and partial trace. We note that $\mathrm{QT}_{\mathbb{C}}$ alone is not a theory, just a set of systems.

In this work we also consider sets of systems which are not expected to form theories; these are sets of systems which share a common dynamical structure. For example the set of systems with shared dynamical structure $\left(\mathrm{P}\mathbb{C}^2, \mathrm{SU}(2)\right)$ form a *sub-family* of systems, and the set of systems which contain all systems with dynamical structure $\left(\mathrm{P}\mathbb{C}^d, \mathrm{SU}(d)\right)$ $(d > 2)$ form a *family* of systems.

In this work we have introduced new families of systems (see Sections 5 and 6) which generalise previously known systems. In Figure 2 we map out the space of transitive systems with compact pure states including the new families of systems introduced in this work.

The advantage of the methods introduced in this work are two-fold: firstly we can generate examples of non-classical systems and secondly we can systematically classify non-classical systems, thus providing us with a fuller picture of non-classical systems lying beyond quantum theory.

7.3 The search for alternative theories and the issue of composition

The tools presented in this work allow us to systematically search for non-classical systems. However it is not certain that these systems compose in a non-trivial manner (existence of entangled states and measurements). For example it is shown in [53] that the only full theory with systems having the same dynamical structure as quantum theory is quantum theory itself. In [48] it is shown (under certain additional assumptions such as local tomography) that the only systems corresponding to d-balls which compose non-trivially are for $d = 3$. However, removing the requirement of local tomography allows for d-balls to compose non-trivially in at least some cases, including those of real and quaternionic bits [35] (where a category in which arbitrary d-balls combine via a monoidal product is also described, though the acceptability of these composites is less clear). Out of the family of systems classified in section 6 it is known that one of them (quartic quantum theory) does not compose under the assumption of local tomography and the requirement that the theory remain quartic [25, 47]. The question remains open as to whether

any of the systems with pure states given by Grassmann manifolds compose non-trivially, with or without the assumption of local tomography

8 Conclusion

In this work we have introduced the *OPF framework* which is used to characterise systems in GPTs. By separating the dynamical and probabilistic components of systems this framework provides new insight into non-classical systems. It allows us to consider *families* of systems which share a common dynamical structure. We introduce the notion of a *rigid* dynamical structure and show that for such structures one can classify all probabilistic structures using representations of the dynamical group. A key feature of rigid dynamical structures is that they do not admit continuous deformation of probabilistic structure.

Moreover we introduced multiple new families of non-classical systems, such as the *complex Grassmann* systems. Many of these families contain known non-classical systems, as well as providing infinitely many examples of non-classical systems which were not known. As well as exploring the space of non-classical systems by finding new examples, we mapped out this space in a more systematic manner by introducing families of systems which share a dynamical structure.

The present work has limited itself to single systems. In general it is not a given that these systems can be made to compose in a non-trivial way (i.e. with entangled states). Future work will involve extending the OPF framework to include composition (along the lines of [52, 53]) and determining whether any of the new families of systems presented in this work can compose in a non-trivial manner. Finding such examples would provide new non-classical theories whose informational properties could be characterised and contrasted to quantum theory. Alternatively a proof that none of the Grassmann systems can compose would lend credence to the proposition that there are few fully compositional general probabilistic theories.

9 Acknowledgements

References

[1] I. E. Segal, Postulates for general quantum mechanics, Annals of Mathematics **48**, 930–948 (1947).

[2] G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (A. Benjamin, Inc., New York, 1963).

[3] G. Ludwig, Versuch einer axiomatischen Grundlegung der Quantenmechanik und allgemeinerer physikalischer Theorien, Zeitschrift für Physik **181**, 233–260 (1964).

[4] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories. II, Comm. Math. Phys. **4**, 331–348 (1967).

[5] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories. III, Communications in Mathematical Physics **9**, 1–12 (1968).

[6] G. Dähn, Attempt of an axiomatic foundation of quantum mechanics and more general theories. IV, Communications in Mathematical Physics **9**, 192–211 (1968).

[7] P. Stolz, Attempt of an axiomatic foundation of quantum mechanics and more general theories. V, Communications in Mathematical Physics **11**, 303–313 (1969).

[8] P. Stolz, Attempt of an axiomatic foundation of quantum mechanics and more general theories VI, Communications in Mathematical Physics **23**, 117–126 (1971).

[9] B. Mielnik, Theory of filters, Communications in Mathematical Physics **15**, 1–46 (1969).

[10] B. Mielnik, Generalized quantum mechanics, Communications in Mathematical Physics **37**, 221–256 (1974).

[11] R. Giles, Foundations for Quantum Mechanics, Journal of Mathematical Physics **11**, 2139 (1970).

[12] S. Gudder, Convex structures and operational quantum mechanics, Communications in mathematical Physics **29**, 249–264 (1973).

[13] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Communications in Mathematical Physics **17**, 239–260 (1970).

[14] L. Hardy, Quantum theory from five reasonable axioms, eprint arXiv:quant-ph/0101012 (2001), quant-ph/0101012 .

[15] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

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

[17] L. Masanes and M. P. Muller, A derivation of quantum theory from physical requirements, New Journal of Physics **13**, 063001 (2011).

[18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Informational derivation of quantum theory, Physical Review A **84**, 012311 (2011).

[19] H. Barnum and A. Wilce, Information processing in convex operational theories, Electronic Notes in Theoretical Computer Science **270**, 3 – 15 (2011).

[20] H. Barnum, R. Duncan, and A. Wilce, Symmetry, compact closure and dagger compactness for categories of convex operational models, Journal of Philosophical Logic **42**, 501–523 (2013).

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

[22] A. J. P. Garner, O. C. O. Dahlsten, Y. Nakata, M. Murao, and V. Vedral, A framework for phase and interference in generalized probabilistic theories, New Journal of Physics **15**, 093044 (2013).

[23] H. Barnum, M. P. Muller, and C. Ududec, Higher-order interference and single-system postulates characterizing quantum theory, New Journal of Physics **16**, 123029 (2014).

[24] H. Barnum, C. Lee, C. Scandolo, and J. Selby, Ruling out Higher-Order Interference from Purity Principles, Entropy **19**, 253 (2017), arXiv:1704.05106 [quant-ph] .

[25] C. M. Lee and J. H. Selby, Higher-order interference in extensions of quantum theory, Foundations of Physics **47**, 89–112 (2017).

[26] M. P. Muller and C. Ududec, Structure of reversible computation determines the self-duality of quantum theory, Phys. Rev. Lett. **108**, 130401 (2012).

[27] C. M. Lee and J. Barrett, Computation in generalised probabilisitic theories, New Journal of Physics **17**, 083001 (2015), arXiv:1412.8671 [quant-ph] .

[28] N. Brunner, M. Kaplan, A. Leverrier, and P. Skrzypczyk, Dimension of physical systems, information processing, and thermodynamics, New Journal of Physics **16**, 123050 (2014).

[29] G. Chiribella and C. M. Scandolo, Operational axioms for diagonalizing states, EPTCS **2**, 96–115 (2015).

[30] G. Chiribella and C. M. Scandolo, Entanglement and thermodynamics in general probabilistic theories, New Journal of Physics **17**, 103027 (2015).

[31] 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 (2010), arXiv:0909.5075 [quant-ph] .

[32] G. Kimura, K. Nuida, and H. Imai, Distinguishability measures and entropies for general probabilistic theories, Reports on Mathematical Physics **66**, 175 – 206 (2010).

[33] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, Limits on nonlocal correlations from the structure of the local state space, New Journal of Physics **13**, 063024 (2011).

[34] P. Janotta and R. Lal, Generalized probabilistic theories without the no-restriction hypothesis, Physical Review A **87**, 052131 (2013).

Accepted in 〉〈uantum 2021-05-11, click title to verify. Published under CC-BY 4.0.

22

[35] H. Barnum, M. A. Graydon, and A. Wilce, Composites and Categories of Euclidean Jordan Algebras, Quantum **4**, 359 (2020).

[36] J. Bae, D.-G. Kim, and L.-C. Kwek, Structure of optimal state discrimination in generalized probabilistic theories, Entropy **18**, 39 (2016).

[37] J. H. Selby and J. Sikora, How to make unforgeable money in generalised probabilistic theories, Quantum **2**, 103 (2018).

[38] T. Heinosaari, L. Leppäjärvi, and M. Plávala, No-free-information principle in general probabilistic theories, Quantum **3**, 157 (2019).

[39] E. Davies, Example Related to the Foundations of Quantum Theory, Journal of Mathematical Physics **13**, 39–41 (1972).

[40] A. J. Short and J. Barrett, Strong nonlocality: a trade-off between states and measurements, New Journal of Physics **12**, 033034 (2010).

[41] D. Gross, M. Muller, R. Colbeck, and O. C. O. Dahlsten, All reversible dynamics in maximally nonlocal theories are trivial, Phys. Rev. Lett. **104**, 080402 (2010).

[42] S. W. Al-Safi and A. J. Short, Reversible dynamics in strongly non-local boxworld systems, Journal of Physics A: Mathematical and Theoretical **47**, 325303 (2014).

[43] L. Hardy and W. K. Wootters, Limited Holism and Real-Vector-Space Quantum Theory, Foundations of Physics **42**, 454–473 (2012), arXiv:1005.4870 [quant-ph] .

[44] A. Aleksandrova, V. Borish, and W. K. Wootters, Real-vector-space quantum theory with a universal quantum bit, Physical Review A **87**, 052106 (2013).

[45] W. K. Wootters, Optimal Information Transfer and Real-Vector-Space Quantum Theory, arXiv eprints:quant-ph/1301.2018 (2013).

[46] D. N. Finkelstein, *Notes on Quaternion Quantum Mechanics*, CERN publications: European Organization for Nuclear Research (CERN, 1959).

[47] K. Życzkowski, Quartic quantum theory: an extension of the standard quantum mechanics, Journal of Physics A: Mathematical and Theoretical **41**, 355302 (2008).

[48] L. Masanes, M. P. Mueller, D. Perez-Garcia, and R. Augusiak, Entanglement and the three-dimensionality of the Bloch ball, Journal of Mathematical Physics **55**, 122203 (2014).

[49] M. Krumm and M. P. Müller, Quantum computation is the unique reversible circuit model for which bits are balls, npj Quantum Information **5**, 7 (2019), arXiv:1804.05736 [quant-ph] .

[50] B. Dakić, T. Paterek, and C. Brukner, Density cubes and higher-order interference theories, New Journal of Physics **16**, 023028 (2014).

[51] T. D. Galley and L. Masanes, Classification of all alternatives to the Born rule in terms of informational properties, Quantum **1**, 15 (2017).

[52] T. D. Galley and L. Masanes, Any modification of the Born rule leads to a violation of the purification and local tomography principles, Quantum **2**, 104 (2018).

[53] L. Masanes, T. Galley, and M. P. Muller, The measurement postulates of quantum mechanics are operationally redundant, Nature Communications **10**, 1361 (2019).

[54] A. Wilce, Four and a half axioms for finite dimensional quantum mechanics (2009), arXiv:0912.5530 [quant-ph] .

[55] L. d. Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy, Nature **474**, 61–63 (2011).

[56] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nature Communications **4**, 2059 (2013).

[57] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource theory of quantum states out of thermal equilibrium, Physical Review Letters **111**, 250404 (2013).

[58] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nature Communications **5**, 4185 (2014).

[59] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, The second laws of quantum thermodynamics, Proceedings of the National Academy of Sciences **112**, 3275–3279 (2015).

[60] K. H. Hofmann and S. A. Morris, *The structure of compact groups : a primer for students, a handbook for the expert*, third edition, revised and augmented ed. (Berlin, Boston: De Gruyter, 2013).

[61] J. H. Gallier, Advanced geometric methods in computer science, `https://www.cis.upenn.edu/~cis610/cis610-18-sl18.pdf` (2018), chapter 20: Manifolds arising from group actions.

[62] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories, PRX Quantum **2**, 020302 (2021).

[63] R. Sanyal, F. Sottile, and B. Sturmfels, Orbitopes, Mathematika **57**, 275–314 (2011).

[64] L. Vesely, Extreme points of compact convex sets, `http://www.mat.unimi.it/users/libor/AnConvessa/ext.pdf` (2010), NoStop

[65] 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).

[66] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature **461**, 1101–1104 (2009).

[67] H.-C. Wang, Two-point homogeneous spaces, Annals of Mathematics **55**, 177–191 (1952).

[68] B. H. Gross, Some applications of gelfand pairs to number theory, Bull. Amer. Math. Soc. (N.S.) **24**, 277–301 (1991).

[69] E. Cartan, Sur la détermination d'un système orthogonal complet dans un espace de Riemann symétrique clos, Rendiconti del Circolo Matematico di Palermo (1884-1940) **53**, 217–252 (1929).

[70] S. Helgason, *Groups and geometric analysis: integral geometry, invariant differential operators, and spherical functions*, Pure and applied mathematics No. 113 (Academic Press, Orlando, 1984).

[71] J. Wolf, *Harmonic Analysis on Commutative Spaces*, Mathematical Surveys and Monographs, Vol. 142 (American Mathematical Society, Providence, Rhode Island, 2007).

[72] M. B. Halima, Branching rules for unitary groups and spectra of invariant differential operators on complex Grassmannians, Journal of Algebra **318**, 520–552 (2007).

[73] W. Fulton and J. Harris, *Representation theory : a first course*, Graduate texts in mathematics (Springer-Verlag, New York, Berlin, Paris, 1991).

[74] M. R. Sepanski, *Compact Lie Groups*, 1st ed. (Springer-Verlag New York, 2007).

[75] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, 2nd ed., Quaderni Monographs No. 1 (Edizioni della normale, Pisa, 2011).

A   Notation

$\mathcal{S}_X$	System with pure states X
$\mathrm{Sym}(X)$	Symmetric group on set X
$\mathrm{Diff}(X)$	Group of diffeomorphisms on manifold X
Ω_x	Image of Ω map for en element x
Ω_X	The set $\Omega_x, \forall x \in X$
$\mathrm{conv}\,(\Omega_X)$	Convex hull of Ω_X
$\{(p_i, x_i)\}_i$	Ensemble of pure states
$\{(p_i, x_i)\}_i \mapsto \sum_i p_i \Omega_{x_i}$	Mixed state representation of an ensemble
$\mathcal{F}_X$	set of OPFs (outcome probability functions) for system with pure states X
$\mathbb{R}[\mathcal{F}_X]$	$\mathbb{R}[\mathcal{F}_X] := \{\tilde{\mathbf{f}}(x) \| \tilde{\mathbf{f}}(x) = \alpha \mathbf{f}_1(x) + \beta \mathbf{f}_2(x), \ \alpha, \beta \in \mathbb{R}, \ \mathbf{f}_1, \mathbf{f}_2 \in F\}$
$\mathbf{f} \mapsto \Lambda_{\mathbf{f}}$	Extension of an OPF to ensembles/mixed states
$(\Gamma, V, \mathbb{F})$	Representation $\Gamma : G \to \mathrm{GL}(V)$ of a group G on linear space V over $\mathbb{F}$
$m(V, V_i)$	Multiplicity of the sub-representation V_i in the representation V
$\mathrm{Hom}_G(V, W)$	Linear space of G-equivariant maps from V to W
$W_{\mathbb{C}}$	Complexification of a real representation W
$V_{\mathbb{R}}$	Restriction to reals of a complex representation V
$C(X, \mathbb{F})$	Continuous functions $X \to \mathbb{F}$ for a topological space X and a field $\mathbb{F}$
$\mathbb{F}[G/H]$	Vector space of functions $G/H \to \mathbb{F}$ for G, H finite
V_G	G-module (carrier space of a representation of G)
$\mathbb{I}_V$	Identity operator on V
$\bar{V}$	Complex conjugate vector space
$\bar{v}$	Complex conjugate of a vector v

B   Background group theory and group representation theory

We briefly outline a few concepts from group representation theory which will be needed for the proofs.

B.1   Representation theory basics

Definition 18 (Group representation)**.** A representation of a group G is a homomorphism $\rho : G \to \mathrm{GL}(V)$, where V is a finite dimensional vector space. When G is a Lie group this is a continuous map. V is called the carrier space of the representation, or a G-module.

Let us denote by $\mathrm{Hom}(V, W)$ the space of linear maps $V \to W$.

Definition 19 (Intertwining operators)**.** Let (ρ, V) and (σ, W) be representations of G. An element $T \in \mathrm{Hom}(V, W)$ is called an intertwining operator if $T \circ \rho = \sigma \circ T$. We denote the linear space of all such maps $\mathrm{Hom}_G(V, W)$.

$\mathrm{Hom}_G(V, W)$ is a subspace of $\mathrm{Hom}(V, W)$.

Theorem 6 (Schur's Lemma)**.** Let $(\rho, V, \mathbb{C})$ and $(\sigma, W, \mathbb{C})$ be irreducible representations of G over $\mathbb{C}$. Then $\dim \mathrm{Hom}_G(V, W) = 1$ when $\rho \cong \sigma$ and $\dim \mathrm{Hom}_G(V, W) = 0$ otherwise.

This implies that $\mathrm{Hom}_G(V, V) = \mathbb{C} \mathbb{I}_V$ for irreducible V. Schur's Lemma also has important consequences for reducible representations. For instance consider the case where $\dim(\mathrm{Hom}_G(V, W)) = 1$ where V is irreducible and W is reducible. Then Schur's Lemma entails that W must contain the irreducible representation V exactly once, and that $\mathrm{Hom}_G(V, W) = \mathbb{C} \mathbb{I}_V$.

B.2   Left regular and $C(G)$ representations

Definition 20 (Left regular representation (finite group))**.** The left regular representation of a finite group G is given by:

$$\rho : G \to \mathrm{GL}(\mathbb{C}[G]) \,, \tag{45}$$

$$\rho_g |g'\rangle = |gg'\rangle \,, \tag{46}$$

where $\mathbb{C}[G]$ is a complex linear space spanned with orthonormal basis $\{|g\rangle\}_{g \in G}$.

Here the action of G on $\mathbb{C}[G]$ is just permutation of the basis vectors.

Definition 21 (Left regular representation (compact group))**.** The left regular representation of a compact topological group G is given by:

$$\rho_g f(g') = f(g^{-1}g') , \tag{47}$$

where $f \in C(G, \mathbb{C})$.

ρ is a continuous homomorphism: $\rho_g \rho_h f(g') = \rho_h f(g^{-1}g') = f(h^{-1}g^{-1}g') = f((gh)^{-1}g') = \rho_{gh} f(g')$.

B.3 Restricted and induced representations

B.3.1 Restricted representation

Definition 22 (Restricted representation)**.** The restriction of a representation (ρ, V) of a group G to a subgroup H is a representation of H on V with operators $\rho(h) : V \to V$ for all $h \in H$. It is denoted $\rho_{|H}$.

In general the restriction of an irreducible representation G will give a reducible representation H.

Example 2. Consider the fundamental representation of SO(3) on $\mathbb{R}^3$ and restrict to the subgroup SO(2) with matrices:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} . \tag{48}$$

This is a reducible representation of SO(2) containing the trivial and fundamental representation.

B.3.2 Induced representation

Definition 23 (Induced representation: G and H finite)**.** A representation $\rho : H \to \mathrm{GL}(V)$ of a subgroup H of G induces a representation of G given by the right regular representation of G on:

$$\mathrm{Ind}_H^G(V) = \{\phi : G \to V | \phi(hg) = \rho(h)\phi(g), \forall h \in H, g \in G\} . \tag{49}$$

The carrier space $\mathrm{Ind}_H^G(V)$ is the space of functions $G \to V$ which are H-equivariant. We can write it as $\mathrm{Ind}_H^G(V) \cong \mathrm{Hom}_H(G, V)$

Definition 24 (Induced representation: G locally compact and H closed)**.** A representation $\rho : H \to \mathrm{GL}(V)$ of a subgroup H of G induces a representation $(\sigma, \mathrm{Ind}_H^G(V))$ of G:

$$\mathrm{Ind}_H^G(V) = \{\phi : G \to V | \phi(gh^{-1}) = \sigma(h)\phi(g), \forall h \in H, g \in G, \ \phi \text{ continuous}\} . \tag{50}$$

The action of G on $\mathrm{Ind}_H^G(V)$ is given by $\sigma(g)\phi(g') = \phi(g^{-1}g')$ where $\sigma : G \to \mathrm{GL}(\mathrm{Ind}_H^G(V))$.

B.4 Frobenius reciprocity

Theorem 7 (Frobenius reciprocity)**.** [74, Theorem 7.47] Take G a finite group and H a subgroup or G a Lie group and H a closed subgroup of G. Given V a representation of H and W a representation of G, then as vector spaces

$$\mathrm{Hom}_G(W, \mathrm{Ind}_H^G(V)) \cong \mathrm{Hom}_H(W_{|H}, V) . \tag{51}$$

We define $\mathbb{C}[G/H]$ (with G, H finite) to be the space of functions $G/H \to \mathbb{C}$ which forms a vector space under pointwise addition and scalar multiplication. It carries a representation of G, inherited from the action of G on G/H. For G/H a topological space, $C(G/H, \mathbb{C})$ is the vector space of continuous functions $G/H \to \mathbb{C}$, also carrying a representation of G.

Corollary 3. For (G, H) finite the irreducible representation W occurs in $\mathbb{C}[G/H]$ with multiplicity equal to the multiplicity of the trivial representation in $W_{|H}$. For (G, H) compact the irreducible representation W occurs in $C(G/H, \mathbb{C})$ with multiplicity equal to the multiplicity of the trivial representation in $W_{|H}$.

Proof. First observe for (G, H) finite $\mathrm{Ind}_H^G(\mathbb{C}) = \mathbb{C}[G/H]$ and for (G, H) compact $\mathrm{Ind}_H^G(\mathbb{C}) = C(G/H, \mathbb{C})$.

Let us consider the case where V is the trivial representation: $V \cong \mathbb{C}$. Then $W_{|H}$ contains a trivial representation with multiplicity $m(W_{|H}, 1_H)$. Using Schur's lemma the only H-equivariant map from $W_{|H}$ to V are identities on the trivial representations. In other words $\mathrm{Hom}_H(W_{|H}, \mathbb{C}) = \mathbb{C}^{m(W_{|H}, 1_H)}$.

Similarly we have that $\mathrm{Hom}_G(W, \mathrm{Ind}_H^G(\mathbb{C})) = \mathbb{C}^{m(\mathrm{Ind}_H^G(\mathbb{C}), W)}$ where $m(\mathrm{Ind}_H^G(\mathbb{C}), W)$ is the multiplicity of W in $\mathrm{Ind}_H^G(\mathbb{C})$. By Theorem 7 we obtain:

$$m(\mathrm{Ind}_H^G(\mathbb{C}), W) = m(W_{|H}, 1) . \tag{52}$$

$\square$

C Representations over $\mathbb{R}$

A representation Γ is a homomorphism $G \to \mathrm{GL}(V)$ for some vector space V. Typically V is assumed to be a vector space over $\mathbb{C}$, and many tools in representation theory apply to this case. Schur's lemma for instance holds for representations over the complex field, but not always for those over $\mathbb{R}$. For a Gelfand pair (G, H) the property of having a trivial irreducible representation when restricted to H which is of multiplicity 0 or 1 holds for representations over $\mathbb{C}$, and not necessarily $\mathbb{R}$.

In this section we explore some of the subtleties involved in dealing with representations over $\mathbb{R}$ and prove some lemmas which will be needed for Theorem 2. First we present an example to introduce some of the relevant concepts.

Example 3 (Fundamental representation of SO(2))**.** Consider the representation of SO(2) over $\mathbb{R}^2$:

$$\Gamma(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} . \tag{53}$$

This representation is irreducible over $\mathbb{R}^2$. However consider this representation acting on $\mathbb{C}^2$ (obtained from $\mathbb{R}^2$ by allowing complex linear combinations of the basis elements). Then there exist the following matrices S and S^{-1}:

$$S = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} , \ S^{-1} = \begin{pmatrix} \frac{i}{2} & \frac{1}{2} \\ -\frac{i}{2} & \frac{1}{2} \end{pmatrix} , \tag{54}$$

such that $\Gamma'(\theta) = S^{-1}\Gamma(\theta)S$, with

$$\Gamma'(\theta) = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} . \tag{55}$$

So the irreducible representation over $\mathbb{R}^2$ is reducible over $\mathbb{C}^2$. Consider once again the irreducible representation Γ over $\mathbb{R}$. Then this commutes with all matrices proportional to the identity, but also matrices proportional to J, where J is:

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{56}$$

Moreover one can show that only matrices which are linear combinations of J and $\mathbb{I}$ commute with the whole group.

C.1 Definitions

Definition 25 (Real, complex and quaternionic structure)**.** Consider an irreducible representation ρ, V with V a complex vector space. Then V has a real structure if there exists an equivariant anti-linear map $j : V \to V$ such that $j^2 = \mathbb{I}$, V has a quaternionic structure if there exists an equivariant anti-linear map $j : V \to V$ such that $j^2 = -\mathbb{I}$. Otherwise V has a complex structure.

Lemma 9 (Descent map)**.** Given a representation $(\Gamma, V, \mathbb{C})$ of a group G over a complex space V equipped with a real structure j the projection $V \to V^j$ with kernel the $j = -1$ eigenspace of V is a descent map. $V^j : \{v \in V : j(v) = v\}$ carries a real representation of G.

Proof. Consider $V^j : \{v \in V : j(v) = v\}$. Observe that this set is closed under real linear combinations, and not complex linear combinations. As such it has the structure of a real vector space. Moreover $V^j \otimes_{\mathbb{R}} \mathbb{C} \cong V$.

Since $j(gv) = gj(v)$, V^j carries an action of G. Moreover for $v_1, v_2 \in V^j$ we have $agv_1 + bgv_2 = agj(v_1) + bgj(v_2) = gj(av_1 + bv_2) = g(av_1 + bv_2)$ for $a, b \in \mathbb{R}$. Therefore V^j carries a real representation of G.

$\square$

Typically j is the conjugation map: $j(v) = \bar{v}$. We see that it is equivariant when $j(gv) = \bar{g}\bar{v}$, i.e. the representation g and $\bar{g}$ are isomorphic.

Definition 26 (Complexification)**.** A real vector space V has complexification $V_{\mathbb{C}} := V \otimes \mathbb{C}$ where $V_{\mathbb{C}}$ is a complex vector space. A basis for this complex vector space is $v \otimes 1$. $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(V_{\mathbb{C}})$.

We sometimes write $V_{\mathbb{C}} \cong V \oplus iV$.

Definition 27 (Restriction of scalars)**.** A complex vector space V is isomorphic (as a real vector space) to the real vector space $V_{\mathbb{R}}$. $\dim_{\mathbb{R}}(V_{\mathbb{R}}) = 2 \dim_{\mathbb{C}}(V)$. If $\{e_i\}$ is a basis for V, then $\{\{e_i\}, \{ie_j\}\}$ is a basis for $V_{\mathbb{R}}$.

Consider V a complex vector space and W a real vector space.

$$(V_{\mathbb{R}})_{\mathbb{C}} \cong_{\mathbb{C}} V \oplus \bar{V} \ , \tag{57}$$

$$(W_{\mathbb{C}})_{\mathbb{R}} \cong_{\mathbb{R}} W \oplus W \ , \tag{58}$$

where $\bar{V}$ is the complex conjugate vector space of V. $\bar{V}$ has the same elements and addition rule as V, but scalar multiplication is given by $\lambda \cdot v = \bar{\lambda}V$.

Theorem 8. [73, Theorem 3.37, p.41] For an irreducible representation $(\rho, W, \mathbb{R})$ the action of G extends naturally to $W_{\mathbb{C}}$. It has one of the following decompositions:

1. $W_{\mathbb{C}} \cong V$ for a complex irreducible representation V if and only if V has a real structure.

2. $W_{\mathbb{C}} \cong V \oplus \bar{V}$ for a complex irreducible representation V if and only if V has a complex structure.

3. $W_{\mathbb{C}} \cong V \oplus V$ for a complex irreducible representation V if and only if V has a quaternionic structure.

Every real irreducible representation $(\rho, W, \mathbb{R})$ can be obtained from irreducible complex representations of one of the above three forms, by taking the descent map $W_{\mathbb{C}}^j$.

Corollary 4. Consider the set of all irreducible complex representations of a group G, these are the representations $\{V_i^r\}$ of real type, the representation $\{V_i^c\}$ complex type and the representations $\{V_j^q\}$ of quaternionic type. The set of all real irreducible representations has a unique representative element in $\{V_i^r\} \bigcup \{V_i^c/*\} \bigcup \{V_j^q\}$, where $\{V_i^c/*\}$ is the set of complex representations where conjugate representations are equivalent.

C.2   Lemmas

In the following section G is assumed finite or compact.

Lemma 10. For a Gelfand pair (G, H) all irreducible representations $(\Gamma, V, \mathbb{C})$ with an H-invariant subspace have either a real or complex structure.

Proof. For (G, H) finite every representation $(\Gamma, V, \mathbb{C})$ with an H-invariant subspace occurs exactly once in the decomposition of $\mathbb{C}[G/H]$ by Corollary 3. $\mathbb{C}[G/H]$ has a real structure (compatible with the group action) and decomposes as $\mathbb{R}[G/H] \oplus i\mathbb{R}[G/H] \cong \mathbb{R}[G/H]_{\mathbb{C}}$.

By Theorem 8 the complexification map $\mathbb{R}[G/H] \to \mathbb{R}[G/H]_{\mathbb{C}} \cong \mathbb{C}[G/H]$ sends every real irreducible representation W in $\mathbb{R}[G/H]$ to either:

1. An irreducible representation $W_{\mathbb{C}} \cong V$ in $\mathbb{C}[G/H]$ if W is real.

2. A reducible representation $W_{\mathbb{C}} \cong V \oplus \bar{V}$ in $\mathbb{C}[G/H]$ if W is complex.

3. A reducible representation $W_{\mathbb{C}} \cong V \oplus V$ in $\mathbb{C}[G/H]$ if W is quaternionic.

However since $\mathbb{C}[G/H]$ only contains single copies of irreps it cannot contain $V \oplus V$ and hence no irrep W can be of quaternionic form.

For (G, H) compact every representation $(\Gamma, V, \mathbb{C})$ with an H-invariant subspace occurs exactly once in the decomposition of $C(G/H, \mathbb{C})$ by Corollary 3. $C(G/H, \mathbb{C})$ is the complexification of $C(G/H, \mathbb{R})$ and the same proof as for the finite case follows through. $\square$

Lemma 11. For (G, H) a Gelfand pair and $(\Gamma, V, \mathbb{R})$ an irreducible representation over $\mathbb{R}$, then the following holds:

1. If Γ is irreducible over $V_{\mathbb{C}}$ then the (real) dimension of the H-invariant subspace in V is the same as the (complex) dimension of the H-invariant subspace in $V_{\mathbb{C}}$: $\dim_{\mathbb{R}}(\mathrm{Hom}_H(\Gamma, 1_H)) = \dim_{\mathbb{C}}(\mathrm{Hom}_H(\Gamma, 1_H))$. It is either 0 or 1.

2. If Γ is reducible over $V_{\mathbb{C}}$ then the (real) dimension of the H-invariant subspace in V is twice that of the (complex) dimension of the H-invariant subspace in $V_{\mathbb{C}}$: $\dim_{\mathbb{R}}(\mathrm{Hom}_H(\Gamma, 1_H)) = 2\dim_{\mathbb{C}}(\mathrm{Hom}_H(\Gamma, 1_H))$. It is either 0 or 2.

Proof.   1. $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$. The H-invariant complex subspace V'_H of $V_{\mathbb{C}}$ is 0 or 1 dimensional. Consider the 1 dimensional case, then V'_H is spanned by a vector $v'_H = v_H \otimes_{\mathbb{R}} \gamma$. Now since $V \cong V_{\mathbb{C}}^j$ for the map $j(v') = \bar{v}'$, we have that under the descent map the subspace V'_H is mapped to the (real) one dimensional subspace V_H of V, spanned by $v_H \otimes_{\mathbb{R}} \gamma$ with $\gamma \in \mathbb{R}$. Any vector v such that $j(v') = v'$ which is not in V'_H will be mapped to a vector in V which is not H-invariant (since j is equivariant). As such the subspace of all H-invariant vectors in V is just V_H and is one dimensional. The case where V'_H is of dimension 0 is immediate.

2. By Lemma 10 if Γ is irreducible over $\mathbb{R}$ but reducible over $\mathbb{C}$ then it has a complex structure and is of the form $\rho \oplus \rho^*$ where ρ and ρ^* are irreducible. Since (G, H) is a Gelfand pair either both ρ and ρ^* contain a one dimensional H-invariant complex subspace or neither do. Let us consider the case where they both do.

$$\Gamma' = \begin{pmatrix} \rho & 0 \\ 0 & \rho^* \end{pmatrix} . \tag{59}$$

We use a matrix representation where the entries of ρ^* are the complex conjugates of those of ρ. Let v_i be a basis for V, then

$$v_i^+ = \begin{pmatrix} v_i \\ \bar{v}_i \end{pmatrix} , v_i^- = \begin{pmatrix} v_i \\ -\bar{v}_i \end{pmatrix} , \tag{60}$$

form a basis for $V \oplus \bar{V}$. By also considering the elements:

$$iv_i^+ = \begin{pmatrix} iv_i \\ i\bar{v}_i \end{pmatrix} , iv_i^- = \begin{pmatrix} iv_i \\ -i\bar{v}_i \end{pmatrix} , \tag{61}$$

we obtain a real basis for $V \oplus \bar{V}$, i.e. any $w \in V \oplus \bar{V}$ is a real linear combination of the above vectors.

Let

$$\begin{pmatrix} v \\ \mathbf{0} \end{pmatrix} , \tag{62}$$

be an H invariant vector, then so is:

$$\begin{pmatrix} \mathbf{0} \\ \bar{v} \end{pmatrix} . \tag{63}$$

All H-invariant vectors are of the form:

$$v_H = \begin{pmatrix} \alpha v \\ \bar{\beta}\bar{v} \end{pmatrix} , \tag{64}$$

since the H invariant subspaces in V and $\bar{V}$ are one dimensional.

Now we consider the change of basis given by matrices S and S^{-1}:

$$S = \begin{pmatrix} -i\mathbb{I} & i\mathbb{I} \\ \mathbb{I} & \mathbb{I} \end{pmatrix} , \ S^{-1} = \begin{pmatrix} \frac{i\mathbb{I}}{2} & \frac{\mathbb{I}}{2} \\ -\frac{i\mathbb{I}}{2} & \frac{\mathbb{I}}{2} \end{pmatrix} , \tag{65}$$

and apply $S\Gamma'S^{-1}$ to obtain

$$\Gamma = \begin{pmatrix} \frac{\rho+\rho^*}{2} & \frac{i(\rho^*-\rho)}{2} \\ \frac{i(\rho-\rho^*)}{2} & \frac{\rho+\rho^*}{2} \end{pmatrix} , \tag{66}$$

has real valued entries. The action of S on the basis is:

$$Sv_i^+ = w_i^+ = \begin{pmatrix} i(\bar{v}_i - v_i) \\ v_i + \bar{v}_i \end{pmatrix} \quad Sv_i^- = w_i^- = \begin{pmatrix} -i(v_i + i\bar{v}_i) \\ v_i - \bar{v}_i \end{pmatrix} , \tag{67}$$

Considering the real basis elements we also obtain:

$$iw_i^+ = \begin{pmatrix} v_i - \bar{v}_i \\ i(v_i + \bar{v}_i) \end{pmatrix} , iw_i^- = \begin{pmatrix} v_i + \bar{v}_i \\ i(v_i - \bar{v}_i) \end{pmatrix} . \tag{68}$$

Under the real structure $j(v) = \bar{v}$ the basis vectors w_i^+ and iw_i^- are $+1$ eigenvectors of j, and form a basis for $V^j \cong \mathbb{R}^n$. The other basis vectors are -1 eigenvectors and form a basis for $i\mathbb{R}$. Since $\Gamma = \bar{\Gamma}$ we have that $j(v) = v \implies j(gv) = gv$ and so the subspace V^j is closed under G.

The image of the H-invariant vectors under S is:

$$Sv_H(\alpha, \beta) = w_H(\alpha, \beta) = \begin{pmatrix} -i\alpha v + i\bar{\beta}\bar{v} \\ \alpha v + \bar{\beta}\bar{v} \end{pmatrix} . \tag{69}$$

These are invariant under j for $\beta = \alpha$. Any such j invariant vector is a real linear combination of $w_H^1 = w_H(\alpha = 1, \beta = 1)$ and $w_H^2 = w_H(\alpha = i, \beta = i)$. As such the H-invariant subspace in V^j is two dimensional with basis vectors written out explicitly as:

$$w_H^1 = \begin{pmatrix} i(\bar{v} - v) \\ v + \bar{v} \end{pmatrix} , \tag{70}$$

$$w_H^2 = \begin{pmatrix} v + \bar{v} \\ i(v - \bar{v}) \end{pmatrix} . \tag{71}$$

$\square$

Corollary 5. For (G, H) a Gelfand pair and Γ an irreducible representation over $\mathbb{R}^n$, then all H-invariant vectors are related by an invertible transformation which commutes with Γ.

Proof. In the case where Γ is complex irreducible this is immediate from Schur's lemma. In the case where Γ is complex reducible, then its complexification is reducible:

$$\Gamma' = \begin{pmatrix} \rho & 0 \\ 0 & \rho^* \end{pmatrix} . \tag{72}$$

By Schur's lemma all matrices which commute with the whole group are of the form:

$$M'(\alpha, \beta) = \begin{pmatrix} \alpha\mathbb{I} & 0 \\ 0 & \beta\mathbb{I} \end{pmatrix} . \tag{73}$$

Using the transformation S of the previous proof one obtains:

$$M(\alpha, \beta) = SM'(\alpha, \beta)S^{-1} = \begin{pmatrix} \frac{\alpha+\beta}{2}\mathbb{I} & \frac{i(\beta-\alpha)}{2}\mathbb{I} \\ \frac{i(\alpha-\beta)}{2}\mathbb{I} & \frac{\alpha+\beta}{2}\mathbb{I} \end{pmatrix} , \tag{74}$$

which all commute with Γ. This is real valued for $\alpha = \beta = 1$, $\alpha = -\beta = i$ and all real linear combinations of these matrices. $M(1,1) = \mathbb{I}$ and $M(i,-i) = J$ where

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{75}$$

Consider the two basis H invariant vectors w_H^1 and w_H^2 of Equation (70). These are related by $-M(i,-i)$. Therefore all linear combinations of these are related by real linear combinations of $M(1,1)$ and $M(i,-i)$. $\qquad\square$

Corollary 6. For Γ, W an irreducible representation over $\mathbb{R}^n$, then the following holds:

1. If $W_\mathbb{C}$ is irreducible then $\mathrm{Hom}_G(\mathbb{R}^n, \mathbb{R}^n) = \mathbb{R}$. The only equivariant homomorphisms are scalar multiples of the identity.

2. If $W_\mathbb{C}$ is reducible into irreducible representations with complex structure then $\mathrm{Hom}_G(\mathbb{R}^n, \mathbb{R}^n) \cong \mathbb{R}^2$.

Lemma 12. For a Gelfand pair of finite groups (G, H), the representation of G acting on $\mathbb{R}[G/H]$ contains exactly once every irreducible representation of G which has a trivial sub-representation when restricted to H and does not contain any other representations.

Proof. By Corollary 3 $\mathbb{C}[G/H]$ contains every irreducible representation W with an H invariant subspace with multiplicity 1, if an irreducible is of complex type then its complex conjugate is necessarily also in $\mathbb{C}[G/H]$, since if W has a H-invariant vector so does W^*.

We have $\mathbb{C}[G/H] = \mathbb{R}[G/H]_\mathbb{C}$ which contains only irreducible representations with complex or real structures by Lemma 10. Therefore we have:

$$\mathbb{C}[G/H] = \bigoplus V_i \oplus \bigoplus (U_j \oplus \bar{U}_j) , \tag{76}$$

where V_j are the irreducible representations of real type, and U_j are irreducible representations of complex type. There are no degeneracies.

$\mathbb{R}[G/H]$ decomposes into real irreducible representations W_i.

$$\mathbb{R}[G/H] = \bigoplus_i W_i , \tag{77}$$

$$\mathbb{C}[G/H] = \mathbb{R}[G/H]_\mathbb{C} = \mathbb{R}[G/H] \otimes \mathbb{C} = \bigoplus_i (W_i \otimes \mathbb{C}) . \tag{78}$$

Hence every irreducible representation W_i is sent to $(W_i)_\mathbb{C}$ in $\mathbb{C}[G/H]$. By Theorem 8 these are of the form V for V irreducible of real type, $V \oplus \bar{V}$ for V of complex type and $V \oplus V$ for V of quaternionic type, where we know by Lemma 10 that this latter case does not occur.

Combining the above:

$$\bigoplus_i (W_i \otimes \mathbb{C}) \cong \bigoplus V_j \oplus \bigoplus_k (U_k \oplus \bar{U}_k) , \tag{79}$$

where every real irreducible subspace $(W_i \otimes \mathbb{C})$ corresponds to either a subspace V_j or $(U_k \oplus \bar{U}_k)$. Since these occur with multiplicity 1, every $(W_i \otimes \mathbb{C})$ occurs with multiplicity one in $\mathbb{R}[G/H]_{\mathbb{C}}$. This implies that every W_i occurs with multplicity 1 in $\mathbb{R}[G/H]$. $\square$

Lemma 13. For a Gelfand pair (G, H), where G compact and H a closed subgroup, the representation of G acting on $C(G/H, \mathbb{R})$ contains exactly once every irreducible representation of G which has a trivial sub-representation when restricted to H and does not contain any other representations.

Proof. Direct from proof of Lemma 12 replacing $\mathbb{C}[G/H]$ with $C(G/H, \mathbb{C})$ and $\mathbb{R}[G/H]$ with $C(G/H, \mathbb{R})$ $\square$

D Convex linear state space

In this appendix we derive the convex linear state space associated to a system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$.

Definition 28 (Preparation space)**.** The preparation space of $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ consists of the space of all ensembles $\{(p_i, x_i)\}_i$ for X finite, which is affinely equivalent to the simplex over the set X (i.e. with the simplex with $|X|$ extremal points, corresponding to the Dirac measures on X). When X is a topological space, the preparation space consists of all positive normalised measures on X, which is the generalised simplex $M_1^+(X)$

In some treatments [75] a mixture space is considered, where the mixing operation applies also to ensembles. For example one can take j ensembles $\{(p_i^j, x_i^j)\}_i$ to prepare an ensemble $\{(p_j, \{(p_i^j, x_i^j)\}_i)\}_j$. In our case we can consider ensembles of ensembles by setting $\{(p_j, \{(p_i^j, x_i^j)\}_i)\}_j = \{(p_j p_i^j, x_i^j)\}_{ij}$ which is an ensemble of pure states and hence in the preparation space. We note that in a mixture space $\{(p_j, \{(p_i^j, x_i^j)\}_i)\}_j$ and $\{(p_j p_i^j, x_i^j)\}_{ij}$ are two distinct elements, even if there can be no affine functional which separates them, by definition. In the "maximally separated" classical state space, these two ensembles are always equivalent. Therefore there is no loss of generality when considering a preparation space, as opposed to a mixture space.

D.1 Proof of Theorem 1

Proof. We consider a system $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ where $\mathbb{R}[\mathcal{F}_X]$ is finite dimensional as a linear space. We consider only a preparation space of finite ensembles for this proof (one can extend to continuous ensembles, but for finite dimensional systems this is not necessary). One can extend the OPFs to the preparation space:

$$P(\mathbf{f}|\{(p_i, x_i)\}_i) = \sum_i p_i \mathbf{f}(x_i) . \tag{80}$$

Two ensembles $\{(p_i, x_i)\}_i$ and $\{(p_j', x_j')\}_j$ where $P(\mathbf{f}|\{(p_i, x_i)\}_i) = P(\mathbf{f}|\{(p_j', x_j')\}_j)$ for all $\mathbf{f} \in \mathcal{F}$ are equivalent ensembles. We can define the linear form $\Omega_x \in \mathbb{R}[\mathcal{F}_X]^*$: $\Omega_x(\mathbf{f}) = \mathbf{f}(x)$ $\forall \mathbf{f} \in \mathcal{F}_X$.

$$P(\mathbf{f}|\{(p_i, x_i)\}_i) = \sum_i p_i \mathbf{f}(x_i) = (\sum_i p_i \Omega_{x_i})(\mathbf{f}) . \tag{81}$$

Hence to the ensemble $\{(p_i, x_i)\}_i$ we associate the mixed state $\omega = \sum_i p_i \Omega_{x_i}$, which is naturally an element of $\mathbb{R}[\mathcal{F}_X]^*$. Any two indistinguishable ensembles are mapped to the same mixed state. Now we can map the OPFs to linear functionals $\Lambda_\mathbf{f}$ in $(\mathbb{R}[\mathcal{F}_X]^*)^*$:

$$\Lambda_\mathbf{f}(\omega) = P(\mathbf{f}|\omega), \ \forall \omega \in \text{conv}(\Omega_X) . \tag{82}$$

By assumption $\Omega_x \to \Omega_{gx}$ is a continuous group action when G is a continuous group. Moreover it extends to $\mathrm{conv}(\Omega_X)$ as follows

$$\sum_i p_i \Omega_{x_i} \xrightarrow{g} \sum_i p_i \Omega_{gx_i} \ . \tag{83}$$

This uniquely extends to $\mathrm{span}(\Omega_X) \cong \mathbb{R}[\mathcal{F}_X]^*$ and hence $\Gamma : G \to \mathrm{GL}(\mathbb{R}[\mathcal{F}_X]^*)$ is a group representation.

$\square$

D.2 Proof of Lemma 1

Proof. Let Ω, Γ, Λ be an embedding of $\mathcal{S}_X = \{X, G, \mathcal{F}_X\}$ into (V, V^*) with $V \cong \mathbb{R}^d$.

Ω_X can be generated by applying Γ_g to a reference vector Ω_{x_0} for all $g \in G$. Since G is compact the matrices Γ_g can be made orthogonal. Hence every $\Omega_x = \Gamma_g \Omega_{x_0}$ for some $g \in G$ is such that $\Omega_x^T \Omega_x = \Omega_{x_0}^T \Omega_{x_0}$. All Ω_x are such that $\mathbf{u} \cdot \Omega_x = 1$ with $\mathbf{u}$ the unit effect, therefore Ω_X lies in the affine span of Ω_X which is a $d-1$ dimensional hyperplane. This is the hyperplane composed of all $v \in V$ such that $\mathbf{u} \cdot v = 1$. Every Ω_x can be written as:

$$\begin{pmatrix} 1 \\ \tilde{\Omega}_x \end{pmatrix} \tag{84}$$

in a basis where the first element corresponds to the subspace spanned by the unit effect. Observe that Γ_G acts trivially on the subspace spanned by the unit effect, and hence decomposes as $\Gamma = 1 + \tilde{\Gamma}$ where 1 is the trivial representation. Observe that the maximally mixed state $\omega = \int_{g \in G} \Gamma_g \Omega_x dg$ for an arbitrary Ω_x is invariant under G. Hence it lies fully in the subspace spanned by the unit effect. Therefore ω can be written as:

$$\begin{pmatrix} 1 \\ \tilde{\omega} \end{pmatrix} \tag{85}$$

where $\tilde{\omega}$ is the zero vector.

Restricting our attention to the affine span of Ω, i.e. ignoring the unit effect subspace we observe that the matrices $\tilde{\Gamma}$ are orthogonal, hence $\tilde{\Gamma}_g \tilde{\Omega}_{x_0}$ lies on a $d-1$ sphere centred on ω.

Since Ω_X lies on a hypersphere, this implies that no point in Ω_X lies inside $\mathrm{conv}(\Omega_X)$, proving the lemma.

$\square$

D.3 Proof of Lemma 2

Proof. In the representation in terms of fiducial OPFs the vectors $\mathrm{conv}(\Omega_X)$ are bounded. Moreover they are closed, since any limit of physically realisable extremal preparations is indistinguishable from an extremal physical preparation. Hence $\mathrm{conv}(\Omega_X)$ is a closed and bounded subset of $\mathbb{R}[\mathcal{F}_X]^*$ (isomorphic to $\mathbb{R}^n$ by the assumption of "Possibility of state estimation using a finite outcome set"), and by the Heine-Borel theorem it is compact in the induced (subspace) topology.

Since $\mathrm{conv}(\Omega_X)$ is bounded this implies that the absolute value of the matrix entries of $\mathrm{conv}(\Gamma_G)$ are also bounded. Moreover we assume that from a physical point of view the space of transformations $\mathrm{conv}(\Gamma_G) \subset \mathbb{R}^n$ must be topologically closed (in the vector space topology of the space $\mathbb{R}^{n^2}$ of linear transformations $\mathbb{R}^n \to \mathbb{R}^n$), since any mathematical transformation which can be arbitrarily well approximated by physical transformations is indistinguishable from a physical transformation. This implies that the set $\mathrm{conv}(\Gamma_G) \subset \mathbb{R}^{n^2}$ is bounded and closed, and hence compact by the Heine-Borel theorem. $\square$

D.4 Proof of Theorem 2

Proof. i. Take $\mathcal{S} = \{X \cong G/H, \mathcal{F}\}$ with (G, H) a Gelfand pair and $\mathbb{R}[\mathcal{F}]$ finite. To every $x \in X$ we associated a linear functional $\Omega_x \in \mathbb{R}[\mathcal{F}]^*$. Let us call Γ the representation of G acting on $\mathbb{R}[\mathcal{F}]^*$.

In general Γ may be reducible:

$$\Gamma = \bigoplus_i \Gamma^i \ . \tag{86}$$

We can also decompose states:

$$\Omega_x = \bigoplus_i \Omega_x^i \ . \tag{87}$$

We have the following equality:

$$\Gamma_g \Omega_x = \Omega_{gx} \ . \tag{88}$$

Since x is stabilized by H_x (where H_x is stabilizer of x: $hx = x$ for all $h \in H_x$) we have

$$\Gamma_h \Omega_x = \Omega_x, \ \forall h \in H \ . \tag{89}$$

This implies

$$\Gamma_h^i \Omega_x^i = \Omega_x^i, \ \forall h \in H \ . \tag{90}$$

This implies that each $\Gamma_{|H}^i$ has at least one trivial sub-representation.

ii. Let us consider a representation of G acting on a real vector space V:

$$\Gamma = \bigoplus_i \Gamma_i \tag{91}$$

such that each Γ_i has at least one H-invariant subspace. Take a reference vector $v \in V$ which has support only in the H-invariant subspaces, and has support in each subspace V_i. By applying Γ_G to v we obtain $\Omega_{G/H} \in V$, where we observe that conv$\left(\Omega_{G/H}\right)$ has $\Omega_{G/H}$ as extremal points, since Γ_G can be expressed in orthogonal matrices and hence $\Omega_{G/H} \subset S^n$ (a hyper-sphere in the affine span of the normalised states, centred on the maximally mixed state) as proven in Lemma 1. By taking the convex set of all linear functionals which give values in $[0,1]$ we obtain $\Lambda_\mathcal{F}$ a probabilistic structure.

iii. Since (G, H) is a Gelfand pair, all real irreducible representations V_i are such that any pair of H-invariant vectors are related by an invertible linear transformation L_i. For a representation:

$$\Gamma = \bigoplus_i \Gamma^i \ , \tag{92}$$

take two H-invariant vectors v and v' which have support in every irreducible subspace. These are related by a transformation L: $Lv = v'$ which commutes with the group action. We call their orbits under Γ_G Ω_X and Ω_X'. Since L commutes with the group action $L\Omega_X = \Omega_X'$. Let us consider the unrestricted effect spaces for both: $\Lambda_\mathcal{F}$ and $\Lambda_{\mathcal{F}'}'$. $\Lambda_{F'}'(\Omega_X') = \Lambda_{\mathcal{F}'}'(L\Omega_X)$. The set of all effects on $L\Omega_X$ is just $\Lambda_F L^{-1}$, hence $\Lambda_{\mathcal{F}'}' = \Lambda_\mathcal{F} L^{-1}$.

iv. Let us take the case where (G, H) is not a Gelfand pair. There exist (complex) irreducible representations W such that $W_{|H}$ contains more than one trivial sub-representation.

One can obtain a real irreducible representation from W by one of the three following methods:

(a) Restriction to $\mathbb{R}$ of W if W is of real type.
(b) Restriction to $\mathbb{R}$ of $W \oplus \bar{W}$ if W is of complex type.
(c) Restriction to $\mathbb{R}$ of $W \oplus W$ if W if of quaternionic type

In each case the real irreducible representation V obtained is such that it has invariant H-vectors which are not related by a transformation which commutes with the group representation. Let us fix a basis and consider the matrices Γ_g acting on V. Let us pick two H-invariant vectors

v and v' such that there is no transformation which commutes with the group action such that $Lv = v'$.

We now show that v and v' can be used (along with Γ_G) to define two tomographically inequivalent probabilistic structures, associated with the same carrier space V.

Let us call x_0 an element in X stabilised by H and define $\Omega_{x_0} = v$ and $\Omega'_{x_0} = v'$. Furthermore let us define $\Omega_{gx_0} = \Gamma_g \Omega_{x_0}$ for every $g \in G$ (and similarly for Ω'_{gx_0}). Observe that for any g_1, g_2 such that $g_1 x_0 = g_2 x_0$ we have $\Omega_{g_1 x_0} = \Omega_{g_2 x_0}$ (and similarly for $\Omega'_{g_1 x_0} = \Omega'_{g_2 x_0}$). This defines the maps $\Omega : X \to V$ and $\Omega' : X \to V$. We call Ω_X and Ω'_X the orbits of Ω_{x_0} and Ω'_{x_0} under Γ_G.

We call $\Lambda_{\mathcal{F}} \in V^*$ the set of all linear functionals giving values in $[0, 1]$ on $\mathrm{conv}\,(\Omega_X)$ and $\Lambda'_{\mathcal{F}'} \in V^*$ the set of all linear functionals giving values in $[0, 1]$ on $\mathrm{conv}\,(\Omega'_X)$. Necessarily $\mathrm{span}(\Lambda'_{\mathcal{F}'}) = \mathrm{span}(\Lambda'_{\mathcal{F}'}) = V^*$. $\Lambda_{\mathcal{F}}$ and $\Lambda'_{\mathcal{F}'}$ define two OPF sets $\mathcal{F}_X$ and $\mathcal{F}'_X$ respectively.

We have constructed two systems $(X, G, \mathcal{F}_X)$ and $(X, G, \mathcal{F}'_X)$ with associated maps $(\Omega, \Gamma, \Lambda)$ and $(\Omega', \Gamma, \Lambda')$ respectively, associated to the same representation Γ. We now show that these two systems are tomographically inequivalent, i.e. the mixed states $\mathrm{conv}\,(\Omega_X)$ and $\mathrm{conv}\,(\Omega'_X)$ are not isomorphic. In other words there is no linear invertible transformation mapping $\mathrm{conv}\,(\Omega_X)$ to $\mathrm{conv}\,(\Omega'_X)$.

Consider an invertible linear transformation such that $L\Omega_X = \Omega'_X$ (if such a transformation exists, then it is well defined on the convex hulls and maps $\mathrm{conv}\,(\Omega_X)$ to $\mathrm{conv}\,(\Omega'_X)$). This implies that $L\Omega_x = \Omega'_{gx}$ for some fixed $g \in G$ and for all pure states $x \in X$. Let us redefine $L \mapsto \Gamma_{g^{-1}}$ such that the equality $L\Omega_x = \Omega'_x$ for all $x \in X$ holds.

Since the state spaces are transitive we have $\Omega_x = \Omega_{gx_0} = \Gamma_g \Omega_{x_0}$ for reference state x_0, and $\Omega'_x = \Omega'_{gx_0} = \Gamma_g \Omega'_{x_0}$. Using $L\Omega_x = \Omega'_x\ \forall x \in X$ we obtain:

$$L\Gamma_g \Omega_{x_0} = \Gamma_g \Omega'_{x_0}, \forall g \in G \ . \tag{93}$$

Moreover from $L\Omega_{x_0} = \Omega'_{x_0}$ we obtain

$$L\Gamma_g \Omega_{x_0} = \Gamma_g L\Omega_{x_0}, \forall g \in G \ . \tag{94}$$

Hence there is a transformation L which commutes with the group action such that $L\Omega_{x_0} = \Omega'_{x_0}$. This is in contradiction with the assumption that the two reference states were not related by such a transformation.

$\square$

E Deformation of probabilistic structure

E.1 Proof of Theorem 3

Proof. Let $\Gamma^i : G \to \mathrm{GL}(V^i)$ be the representation of G associated to $\mathcal{F}^i$, let $\Omega^i : X \to V^i$ be the representation of pure states, and let $\Lambda^i : \mathcal{F}^i \to (V^i)^*$ be the representation of OPFs, for $i = 0, 1$. We can decompose the group action into (real) irreducible representations as $\Gamma^i = \bigoplus_j \Gamma^i_j$, $V^i = \bigoplus_j V^i_j$, $\Gamma^i_j : G \to \mathrm{GL}(V^i_j)$, where $j = 0, 1, \dots$ Recall that there must be one (and only one) trivial irrep, which we label by $j = 0$. Also, we can decompose the representation of pure states $\Omega^i = \bigoplus_j \Omega^i_j$ and $\Omega^i_j : X \to V^i_j$, and the representation of OPFs $\Lambda^i = \bigoplus_j \Lambda^i_j$ and $\Lambda^i_j : \mathcal{F}_j \to (V^i_j)^*$. Within this proof we follow the convention that

$$\langle \Omega^i_j(x), \Omega^i_j(x) \rangle_j = 1 \ , \tag{95}$$

for all i, j, x, where $\langle \cdot, \cdot \rangle_j$ is a group-invariant scalar product. This scalar product provides an isomorphism $(V^i_j)^* \cong V^i_j$. Using this decomposition we can write any $\mathbf{f}^i \in \mathcal{F}^i$ as

$$\mathbf{f}^i(x) = c^i_0(\mathbf{f}^i) + \sum_{j \geq 1} \langle \Lambda^i_j(\mathbf{f}^i), \Omega^i_j(x) \rangle_j \ , \tag{96}$$

where we define

$$c^i_j(\mathbf{f}^i) = \langle \Lambda^i_j(\mathbf{f}^i), \Lambda^i_j(\mathbf{f}^i) \rangle^{1/2}_j \ . \tag{97}$$

Schur's Lemma tells us that any non-trivial irreducible representation $\Gamma : G \to \mathrm{GL}(V)$ satisfies $\int_G dg\, \Gamma(g) = 0$. Therefore, only the trivial rep survives the following average

$$\int_G dg\, \mathbf{f}^i(gx) = \int_G dg \left[c_0^i(\mathbf{f}^i) + \sum_{j \geq 1} \langle \Lambda_j^i(\mathbf{f}^i), \Gamma_j^i(g)\, \Omega_j^i(x) \rangle_j \right] = c_0^i(\mathbf{f}^i) \ . \tag{98}$$

Next we use another consequence of Schur's Lemma, namely that for any pair of (real) inequivalent irreducible representations $\Gamma : G \to \mathrm{GL}(V)$ and $\Gamma' : G \to \mathrm{GL}(V')$ we have

$$\int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle \langle \Lambda', \Gamma'(g)\Omega' \rangle = 0 \ , \tag{99}$$

for any vectors Λ, Ω in V and any vectors Λ', Ω' in V'. Also, if the real irreducible representation $\Gamma : G \to \mathrm{GL}(V)$ has complexification with real structure then

$$\int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle^2 = \int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle \langle \Omega, \Gamma(g)^\dagger \Lambda \rangle \tag{100}$$

$$= \int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle \langle \Omega, \Gamma(g^{-1})\Lambda \rangle \tag{101}$$

$$= \frac{1}{d}\langle \Lambda, \Lambda \rangle \langle \Omega, \Omega \rangle \ , \tag{102}$$

where d is the dimension of V and Λ, Ω are any vectors in V.

The case where the real irreducible representation $\Gamma : G \to \mathrm{GL}(V)$ acts on the complexified space $\mathbb{C}V \cong W \oplus \bar{W}$ as the direct sum of two complex irreducible representations $\gamma \oplus \bar{\gamma}$, with complex structure and one being the dual of the other. For what follows we are interested in elements of $W \oplus \bar{W}$ of the form $\Omega = \omega \oplus \bar{\omega}$ where $\bar{\omega}$ is the complex conjugate of ω, and analogously for $\Lambda = \lambda \oplus \bar{\lambda}$. Proceeding as above and using Schur's Lemma we obtain

$$\int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle^2 = \int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle \langle \Omega, \Gamma(g)^\dagger \Lambda \rangle \tag{103}$$

$$= \int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle \langle \Omega, \Gamma(g^{-1})\Lambda \rangle \tag{104}$$

$$= \int_G dg\, \left(\langle \lambda, \gamma(g)\omega \rangle + \langle \bar{\lambda}, \bar{\gamma}(g)\bar{\omega} \rangle \right) \left(\langle \omega, \gamma(g^{-1})\lambda \rangle + \langle \bar{\omega}, \bar{\gamma}(g^{-1})\bar{\lambda} \rangle \right)$$

$$= \frac{2}{d}\langle \lambda, \lambda \rangle \langle \omega, \omega \rangle + \frac{2}{d}\langle \bar{\lambda}, \bar{\lambda} \rangle \langle \bar{\omega}, \bar{\omega} \rangle \tag{105}$$

where $d/2$ is the dimension of W (and $\bar{W}$). The structure of the vector Ω implies

$$\langle \Omega, \Omega \rangle = \langle \omega, \omega \rangle + \langle \bar{\omega}, \bar{\omega} \rangle = 2\langle \omega, \omega \rangle = 2\langle \bar{\omega}, \bar{\omega} \rangle \ , \tag{106}$$

and analogously for Λ. Which in turn implies

$$\int_G dg\, \langle \Lambda, \Gamma(g)\Omega \rangle^2 = \frac{1}{d}\langle \Lambda, \Lambda \rangle \langle \Omega, \Omega \rangle \ . \tag{107}$$

In summary, the above average is independent of whether the complexification of representation Γ is reducible or not.

The remaining case, where W has quaternionic structure, need not be considered, because Lemma 7 proves that it never appears in Gelfand systems. Next we use the identities (99) and (107), the normalisation convention (95) and definition (97) to calculate the following average

$$\int_G dg\, [\mathbf{f}^i(gx)]^2 = c_0^i(\mathbf{f}^i)^2 + \sum_{j \geq 1} \frac{1}{d_j^i} \langle \Lambda_j^i(\mathbf{f}^i), \Lambda_j^i(\mathbf{f}^i) \rangle \langle \Omega_j^i(x), \Omega_j^i(x) \rangle_j$$

$$= \sum_j \frac{c_j^i(\mathbf{f}^i)^2}{d_j^i} \ , \tag{108}$$

where $d_j^i = \dim V_j^i$.

Now, suppose that $\mathcal{F}^0$ and $\mathcal{F}^1$ are unrestricted and tomographically-inequivalent probabilistic structures of (G, X). Then, without loss of generality, we assume that irrep Γ_1^0 is inequivalent to irrep Γ_j^1 for all j. Also, let us choose $\mathbf{f}^0 \in \mathcal{F}^0$ so that its corresponding vector $\Lambda^0(\mathbf{f}^0)$ only has support on the subspaces $j = 0, 1$, that is

$$\mathbf{f}^0(x) = c_0^0(\mathbf{f}^i) + \langle \Lambda_1^0(\mathbf{f}^0), \Omega_1^0(x) \rangle_1 \ . \tag{109}$$

This choice is possible because the probabilistic structure $\mathcal{F}^0$ is unrestricted. Also, due to the fact that we are obtaining a lower bound for the distance $\mathcal{D}(\mathcal{F}^0, \mathcal{F}^1)$, we do not need to maximise over OPFs $\mathbf{f}^0 \in \mathcal{F}^0$, it is enough to pick any one. Next, let us lower-bound the distance between the chosen OPF $\mathbf{f}^0 \in \mathcal{F}^0$ and any given OPF $\mathbf{f}^1 \in \mathcal{F}^1$ as

$$\text{dist}(\mathbf{f}^0, \mathbf{f}^1) = \max_{x \in X} |\mathbf{f}^0(x) - \mathbf{f}^1(x)| \geq \int_G dg \ |\mathbf{f}^0(gx_0) - \mathbf{f}^1(gx_0)|$$
$$\geq \int_G dg \ [\mathbf{f}^0(gx_0) - \mathbf{f}^1(gx_0)]^2 \ , \tag{110}$$

for any $x_0 \in X$. This bound is independent of whether the probabilistic structures $\mathcal{F}^0$ and $\mathcal{F}^1$ are equivalent or not. Now, using Schur's lemma as in (98) and (108) we obtain

$$\text{dist}(\mathbf{f}^0, \mathbf{f}^1) \geq \sum_{j=0}^{1} \frac{c_j^0(\mathbf{f}^0)^2}{d_j^0} + \sum_j \frac{c_j^1(\mathbf{f}^1)^2}{d_j^1} - 2 \, c_0^0(\mathbf{f}^0) \, c_0^1(\mathbf{f}^1)$$
$$= \frac{c_1^0(\mathbf{f}^0)^2}{d_1^0} + \sum_{j \geq 1} \frac{c_j^1(\mathbf{f}^1)^2}{d_j^1} + [c_0^0(\mathbf{f}^0) - c_0^1(\mathbf{f}^1)]^2$$
$$\geq \frac{c_1^0(\mathbf{f}^0)^2}{d_1^0} + \sum_{j \geq 1} \frac{c_j^1(\mathbf{f}^1)^2}{d_j^1} \ , \tag{111}$$

for any $\mathbf{f}^1 \in \mathcal{F}^1$, therefore

$$\mathcal{D}(\mathcal{F}^0, \mathcal{F}^1) \geq \min_{\mathbf{f}^1 \in \mathcal{F}^1} \text{dist}(\mathbf{f}^0, \mathbf{f}^1) \geq \frac{c_1^0(\mathbf{f}^0)^2}{d_1^0} \ . \tag{112}$$

Again, using the fact that we can pick any $\mathbf{f}^0 \in \mathcal{F}^0$, we further restrict (109) to

$$\mathbf{f}^0(x) = \frac{1}{2} + \frac{1}{2} \langle \Omega_1(x_0), \Omega_1(x) \rangle \in [0, 1] \ , \tag{113}$$

for any fixed state $x_0 \in X$. The above $\mathbf{f}^0$ is an OPF because the normalisation condition (95) implies that it takes values within $[0, 1]$. This chosen OPF (113) has $c_1^0(\mathbf{f}^0) = 1/2$, which when substituted in (112) gives

$$\mathcal{D}(\mathcal{F}^0, \mathcal{F}^1) \geq \frac{1}{4d_1^0} \ , \tag{114}$$

which is the statement of Theorem 3. $\qquad \square$

E.2 Proof of Theorem 4

E.2.1 Proof of Theorem 4 i.

Proof. Let $\mathcal{F}^0$ of dimension $\dim \mathbb{R}[\mathcal{F}_0] = d_0$ be an unrestricted probabilistic structure of (G, H) with associated representation Γ_G^0 such that all H-invariant vectors are related by invertible transformations which commute with Γ_G^0.

Consider two such vectors v_H and v_H' where $v_H' = L v_H$, with $L \in \text{GL}(\mathbb{R}[\mathcal{F}_0])$. We can generate state spaces with extremal points G/H by applying Γ_G to the reference states:

$$\Omega_{G/H} = \Gamma_G v_H \ , \tag{115}$$
$$\Omega'_{G/H} = \Gamma_G v_H' = \Gamma_G L v_H = L \Omega_{G/H} \ . \tag{116}$$

Both the state spaces $\operatorname{conv}\left(\Omega_{G/H}\right)$ and $\operatorname{conv}\left(\Omega'_{G/H}\right)$ are related by an invertible transformation and hence are equivalent as convex sets. In other words they correspond to tomographically equivalent probabilistic structures. Since we are only considering unrestricted probabilistic structures they both generate the same state space corresponding to $\mathcal{F}_0$ (the linear transformation L is just a change of co-ordinates, including rescalings). For probabilistic structures whose H-invariant vectors are all related to each other by invertible transformations which commute with Γ_G there is a unique unrestricted probabilistic structure associated to Γ_G.

Therefore the only other probabilistic structures $\mathcal{F}_1$ such that $\dim\mathbb{R}[\mathcal{F}_1] = d_0$ (if they exist) are associated to different representations Γ_G^1. The largest dimensional irreducible representation they can differ by is of dimension $d_0 - 1$ (corresponding to the case where either Γ^0 or Γ^1 consists of a $d_0 - 1$ dimensional irreducible representation and the trivial and the other representation consists of the trivial representation and an inequivalent $d_0 - 1$ irreducible representation or some reducible representation of total dimension $d_0 - 1$). Therefore by Theorem 3 the distance between the two probabilistic structures is lower bounded by $\frac{1}{4(d_0-1)}$. Any other possibility (i.e. both consist of a trivial representation and different reducible ones would give a lower bound which is higher). $\qquad\square$

E.2.2 Proof of Theorem 4 *ii.*

Let $\mathcal{F}^0$ of dimension $\dim\mathbb{R}[\mathcal{F}_0] = d_0$ be an unrestricted probabilistic structure of (G, H) with associated representation Γ_G^0 such that there are pairs of H-invariant vectors not related by invertible transformations which commute with Γ_G^0. These H-invariant vectors belong to an H-invariant subspace V_H.

Take $\Omega_H \in V_H$ and $\Omega_H^t = L^t \Omega_H$ where $L^t = e^{iRt}$ with $R \in \mathfrak{gl}(V_H)$. Take Ω_H^ϵ for small $\epsilon > 0$. We then expand:

$$\Omega_H^\epsilon = (\mathbb{I} + \epsilon R + ...)\Omega_H = \Omega_H + \epsilon R\Omega_H + O(\epsilon^2) . \tag{117}$$

We consider a first order approximation and ignore terms of order ϵ^2 and higher.

The image of a point $x = gH$ is $\Omega_{gH} = \Gamma_g \Omega_H$ and $\Omega_{gH}^\epsilon = \Gamma_g \Omega_H^\epsilon$:

$$\Omega_{gH}^\epsilon = \Omega_{gH} + \epsilon\Gamma_g R\Omega_H . \tag{118}$$

Let us take an OPF $\mathbf{f}^0 \in \mathcal{F}_0$ with associated effect $\Lambda_{\mathbf{f}^0}$ and an OPF $\mathbf{f}^1 \in \mathcal{F}_\epsilon$ with associated effect $\Lambda_{\mathbf{f}^1}^\epsilon$:

$$\mathbf{f}^0(gH) = \Lambda_{\mathbf{f}^0} \cdot \Omega_{gH} ,$$
$$\mathbf{f}^1(gH) = \Lambda_{\mathbf{f}^1}^\epsilon \cdot \Omega_{gH}^\epsilon .$$

$$\operatorname{dist}(\mathbf{f}^0, \mathbf{f}^1) = \max_{x \in X} |\mathbf{f}^0(x) - \mathbf{f}^1(x)| = \max_{x \in X} |\Lambda_{\mathbf{f}^0} \cdot \Omega_x - \Lambda_{\mathbf{f}^1}^\epsilon \cdot \Omega_x^\epsilon| \tag{119}$$

$$= \max_{g \in G} |(\Lambda_{\mathbf{f}^0} - \Lambda_{\mathbf{f}^1}^\epsilon) \cdot \Omega_{gH} - \epsilon\Lambda_{\mathbf{f}^1}^\epsilon \cdot (\Gamma_g R\Omega_H)| . \tag{120}$$

For a fixed $\mathbf{f}_0$ we want to find an $\mathbf{f}^1$ which minimizes the distance.

$$\min_{\mathbf{f}^1 \in \mathcal{F}_\epsilon} \operatorname{dist}(\mathbf{f}^0, \mathbf{f}^1) = \min_{\mathbf{f}^1 \in \mathcal{F}_\epsilon} \max_{g \in G} |(\Lambda_{\mathbf{f}^0} - \Lambda_{\mathbf{f}^1}^\epsilon) \cdot \Omega_{gH} - \epsilon\Lambda_{\mathbf{f}^1}^\epsilon \cdot (\Gamma_g R\Omega_H)| . \tag{121}$$

We observe that that the transformations Γ_g and R leave the normalisation degree of freedom (component in the normalisation subspace V_0) unchanged, therefore if we take $\mathbf{f}^0 = \mathbf{u}^0$ then we can choose $\mathbf{f}^1 = \mathbf{u}^1$ to obtain $\operatorname{dist}(\mathbf{u}^0, \mathbf{u}^1) = 0$.

For any effect with support outside the normalisation subspace V_0 the expression (121) is not linear in Ω_{gH} (due to the last term) and as such there is no choice of $\Lambda_{\mathbf{f}^1}$ which will make $(\Lambda_{\mathbf{f}^0} - \Lambda_{\mathbf{f}^1}^\epsilon)$ cancel out with $\epsilon\Lambda_{\mathbf{f}^1}^\epsilon\Gamma_g R\Omega_H$ for all $g \in G$, which is needed to make the distance go to 0. Hence for $\mathbf{f}^0$ not proportional to $\mathbf{u}^0$ we have the bound:

$$\min_{\mathbf{f}^1 \in \mathcal{F}_\epsilon} \operatorname{dist}(\mathbf{f}^0, \mathbf{f}^1) > 0 .$$

We observe that every Ω_x^ϵ is ϵ close to Ω_x, therefore any $\Lambda_{\mathbf{f}^0}$ which is not $\epsilon-$tight on Ω_X (i.e. which gives values in $[0+\epsilon, 1-\epsilon]$) will be valid on Ω_X^ϵ. Let us minimize the distance $\min_{\mathbf{f}^1 \in \mathcal{F}^\epsilon} \text{dist}(\mathbf{f}^0, \mathbf{f}^1)$ for the choice of $\mathbf{f}^0$ not $\epsilon-$tight. We observe that by choosing $\Lambda_{\mathbf{f}^1}^\epsilon = \Lambda_{\mathbf{f}^0}$ we obtain:

$$\max_{g \in G} |(\Lambda_{\mathbf{f}^0} - \Lambda_{\mathbf{f}^1}^\epsilon) \cdot \Omega_{gH} - \epsilon \Lambda_{\mathbf{f}^1}^\epsilon \cdot (\Gamma_g R v_H)| = |\epsilon \Lambda_{\mathbf{f}^1}^\epsilon \cdot (\Gamma_g R \Omega_H)| \leq \epsilon , \tag{122}$$

which implies that

$$\min_{\mathbf{f}^1 \in \mathcal{F}^\epsilon} \text{dist}(\mathbf{f}^0, \mathbf{f}^1) \leq \epsilon . \tag{123}$$

Finally if $\mathbf{f}^0$ is $\epsilon-$tight (i.e. has some values in $[0, \epsilon]$ or $[1 - \epsilon, 1]$), one can take $\Lambda_{\mathbf{f}^0}$ and add ϵ noise (i.e. take $\Lambda_{\mathbf{f}^0}' = (1 - \epsilon)\Lambda_{\mathbf{f}^0} + \epsilon \Lambda_{\mathbf{u}^0}$) in order to obtain an effect $\Lambda_{\mathbf{f}^0}'$ which is valid on Ω_X^ϵ:

$$\max_{g \in G} |(\Lambda_{\mathbf{f}^0} - \Lambda_{\mathbf{f}^0}') \cdot \Omega_{gH} - \epsilon \Lambda_{\mathbf{f}^0}' \cdot (\Gamma_g R \Omega_H)| \tag{124}$$

$$= \min_{\mathbf{f}^1 \in \mathcal{F}^\epsilon} \max_{g \in G} |\epsilon - \epsilon \Lambda_{\mathbf{f}^1}^\epsilon \cdot (\Gamma_g R \Omega_H)| \leq 2\epsilon . \tag{125}$$

Therefore

$$\min_{\mathbf{f}^1 \in \mathcal{F}^\epsilon} \text{dist}(\mathbf{f}^0, \mathbf{f}^1) \leq 2\epsilon . \tag{126}$$

This, together with Equation (123), implies that

$$\max_{\mathbf{f}^0 \in \mathcal{F}^0} \min_{\mathbf{f}^1 \in \mathcal{F}^\epsilon} \text{dist}(\mathbf{f}^0, \mathbf{f}^1) \leq 2\epsilon . \tag{127}$$

E.3 Proof of Lemma 4

Proof. Two convex sets S_0 and S_1 with pure states X linearly embedded in isomorphic spaces $V_0 \cong V_1 \cong \mathbb{R}^d$ (i.e. for which there exist injective maps $\Omega^0 : X \to \mathbb{R}^d$ and $\Omega^1 : X \to \mathbb{R}^d$ and $S_i = \text{conv}(\Omega^i(X))$) can be continuously deformed one into the other if there exists a *path connected map* $M_{0 \to t} : \Omega_X^0 \to \Omega_X^t$ in the space $\mathfrak{F}_d$ (of Ω maps with codomain $\mathbb{R}^d$) parametrised by $t \in [0, 1]$. $M_{0 \to t}(\Omega_X^1) = \Omega_X^t$ is an embedding of X in $\mathbb{R}^n$ for every $t \in [0, 1]$. We observe that $M_{0 \to t}$ is defined just on the extremal points of Ω_X^0 (and is invertible), but cannot be extended to $\text{conv}(\Omega_X^0)$ by Lemma 3. A connected path on $\mathfrak{F}_d$ (the space of Ω maps with codomain $\mathbb{R}^d$) is a continuous map from $[0, 1] \subset \mathbb{R}$ (with the usual topology) to $\mathfrak{F}_d$ (with the topology inherited from the metric D_{sym}).

We now show that for any non-rigid probabilistic structure one can find probabilistic structures which can be continuously deformed into one another, whilst rigid probabilistic structures cannot be continuously deformed.

In the following we make use of the isomorphism $X \cong G/H$ and label points in X by their stabilizer subgroup. The point x with stabilizer subgroup H is written as H, and any point $x' = gx$ is written as gH.

A probabilistic structure $\mathcal{F}_X$ is non-rigid if the associated representation Γ contains an irreducible spherical representation Γ with inequivalent H-invariant vectors v_H^0 and v_H^1. Two H-invariant vectors v_H^0 and v_H^1 are inequivalent if they are not related by an invertible transformation which commutes with Γ_G.

We show that a non-rigid probabilistic structure $\mathcal{F}_X^0$ associated to an irreducible representation Γ can be continuously deformed. The case of a non-rigid probabilistic structure $\mathcal{F}_X^0$ associate to a reducible representation follows since one can deform the full reducible probabilistic structure by deforming the subspace of $\mathcal{F}_X^0$ acted on by the irreducible representation with inequivalent H-invariant vectors.

The embedded pure state Ω_X^0 associated to $\mathcal{F}_X^0$ can be generated as $\Omega_X^0 = \{\Gamma_g v_H^0 | g \in G\}$, where v_H^0 is an H-invariant vector. Let v_H^1 be an inequivalent H-invariant vector. Any vector in $\text{span}(v_H^0, v_H^1)$ is H-invariant, moreover any two vectors in $\text{span}(v_H^0, v_H^1)$ which are not proportional will generate inequivalent state spaces under Γ_G. Let us call $\Omega_X^1 = \{\Gamma_g v_H^1 | g \in G\}$ the embedded

sets of pure states generated from the reference state v_H^1. We now show that Ω_X^0 can be continuously deformed to Ω_X^1.

Consider a rotation $R \in \mathrm{GL}(\mathrm{span}(v_H^0, v_H^1))$ such that $R v_H^0 = v_H^1$.

Let us write $R(t) = e^{iKt}$ with $R(1) = R$ for some non-unique generator K.

We now show the existence of a connected path in $\mathfrak{F}_d$ from Ω_X^0 to Ω_X^1, the two sets pure states generated by two non-equivalent H-invariant vectors v_H^0 and v_H^1.

Let us define Ω_X^t the state space generated by the H-invariant reference vector $v_H^t = R(t) v_H^0$.

We now define $\gamma : [0, 1] \to \mathfrak{F}_d$, where $\gamma(t) = \Omega_X^t$. We observe that $\gamma(0) = \Omega_X^0$ and $\gamma(1) = \Omega_X^1$. We now show that γ is a continuous function (where $\mathfrak{F}_d$ has the topology induced by the metric $\mathcal{D}_{\mathrm{sym}}$ of Equation (18))

γ is continuous at a point $t_0 \in [0, 1]$ if for every $\epsilon > 0$ it is the case that if $\mathcal{D}_{\mathrm{sym}}(\Omega_X^{t_0}, \Omega_X^t) < \epsilon$ (for $t \in [0, 1]$) then we can find a δ such that $|t - t_0| < \delta$ implies $\mathcal{D}_{\mathrm{sym}}(\Omega_X^{t_0}, \Omega_X^t) < \epsilon$.

For ϵ large (but ≤ 1) this holds since one can just take δ to be small, in which case $\mathcal{D}_{\mathrm{sym}}(\Omega_X^{t_0}, \Omega_X^{t_0}) \leq 2\delta$ by the proof of Theorem 4 *ii.* which is less than or equal to ϵ.

If $\epsilon << 1$ (and greater than 0) then setting $\delta = \frac{\epsilon}{2}$ gives $|t - t_0| < \frac{\epsilon}{2}$ which implies $\mathcal{D}_{\mathrm{sym}}\left(\Omega_X^{t_0}, \Omega_X^{t_0}\right) < \epsilon$ using the proof of Theorem 4 *ii.*.

This holds for all $t_0 \in [0, 1]$ hence γ is a continuous function from $[0, 1]$ to $\mathfrak{F}_d$ with extremal points Ω_X^0 and Ω_X^1. Hence there is a connected path from Ω_X^0 and Ω_X^1 in the space $\mathfrak{F}_D$ of deformation maps, meaning that Ω_X^0 can be continuously deformed to Ω_X^1.

In the case where $\mathcal{F}_X^0$ is a rigid probabilistic structure it is at a finite bounded distance from every other probabilistic structure in $\mathfrak{F}_d$ by definition. Consider the associated set of embedded pure states Ω_X^0 and take an arbitrary path $\gamma(t)$ to an arbitrary $\Omega_X^1 \in \mathfrak{F}_d$. $\gamma(t)$ is not continuous at $t = 0$, since for ϵ lower than the finite bounded distance from $\mathcal{F}_X^0$ to any other probabilistic structure, there is no δ such that $|t| < \delta$ implies $\mathcal{D}_{\mathrm{sym}}\left(\Omega_X^0, \Omega_X^t\right) < \epsilon$.

$\square$

F Proof of Lemma 7 part 2

Proof. We make use of Proposition 26.24 of [73] which we translate slightly:

Lemma 14. Any irreducible representation of SU(d) with Dynkin indices $j = (j_1, ..., j_{d-1})$ is complex if $j_i \neq j_d - i$ for any i, real if $j_i = j_{d-i}$ for all i and n is odd, or $n = 4k$, or $n = 4k + 2$ and j_{2k+1} is even, and quaternionic if $j_i = j_{d-i}$ for all i and j_{2k+1} is odd.

We translate the partitions (41) and (42) into Dynkin indices:
When $m = n$:
$$j = (b_1 - b_2, b_2 - b_3, ..., 2b_m, b_{m-1} - b_m, ..., b_1 - b_2) \tag{128}$$

where this has $m + n - 1$ entries.

For all i $j_i = j_{m+n-i}$, hence for $m+n$ odd and $m+n = 4k$ these are all real. For $m+n = 4k+2$ we observe that j_{2k+1} (the central entry) is $2b_m$ and hence even. This implies the representation is real.

When $n \geq m + 1$:
$$\lambda = (b_1 - b_2, b_2 - b_3, ..., b_m, \underbrace{0, ..., 0}_{\text{times } n-m-1}, b_m, ..., b_2 - b_3, b_1 - b_2) \tag{129}$$

where this has $m + n - 1$ entries.

For all i $j_i = j_{m+n-i}$, hence for $m+n$ odd and $m+n = 4k$ these are all real. For $m+n = 4k+2$ we observe that j_{2k+1} (the central entry) is 0 and hence even. This implies the representation is real.

$\square$