Geometry of Quantum States from Symmetric Informationally Complete Probabilities

by

Gelo Noel M. Tabia

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics - Quantum Information

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

It is usually taken for granted that the natural mathematical framework for quantum mechanics is the theory of Hilbert spaces, where pure states of a quantum system correspond to complex vectors of unit length. These vectors can be combined to create more general states expressed in terms of positive semidefinite matrices of unit trace called density operators. A density operator tells us everything we know about a quantum system. In particular, it specifies a unique probability for any measurement outcome. Thus, to fully appreciate quantum mechanics as a statistical model for physical phenomena, it is necessary to understand the basic properties of its set of states. Studying the convex geometry of quantum states provides important clues as to why the theory is expressed most naturally in terms of complex amplitudes. At the very least, it gives us a new perspective into thinking about structure of quantum mechanics.

This thesis is concerned with the structure of quantum state space obtained from the geometry of the convex set of probability distributions for a special class of measurements called symmetric informationally complete (SIC) measurements. In this context, quantum mechanics is seen as a particular restriction of a regular simplex, where the state space is postulated to carry a symmetric set of states called SICs, which are associated with equiangular lines in a complex vector space. The analysis applies specifically to 3-dimensional quantum systems or qutrits, which is the simplest nontrivial case to consider according to Gleason's theorem. It includes a full characterization of qutrit SICs and includes specific proposals for implementing them using linear optics. The infinitely many qutrit SICs are classified into inequivalent families according to the Clifford group, where equivalence is defined by geometrically invariant numbers called triple products. The multiplication of SIC projectors is also used to define structure coefficients, which are convenient for elucidating some additional structure possessed by SICs, such as the Lie algebra associated with the operator basis defined by SICs, and a linear dependency structure inherited from the Weyl-Heisenberg symmetry. After describing the general one-to-one correspondence between density operators and SIC probabilities, many interesting features of the set of qutrits are described, including an elegant formula for its pure states, which reveals a permutation symmetry related to the structure of a finite affine plane, the exact rotational equivalence of different SIC probability spaces, the shape of qutrit state space defined by the radial distance of the boundary from the maximally mixed state, and a comparison of the 2-dimensional cross-sections of SIC probabilities to known results. Towards the end, the representation of quantum states in terms of SICs is used to develop a method for reconstructing quantum theory from the postulate of maximal consistency, and a procedure for building up qutrit state space from a finite set of points corresponding to a Hesse configuration in Hilbert space is sketched briefly.

Acknowledgements

> There is no one way the world is because the world is still in creation, still being hammered out.
>
> – Christopher A. Fuchs, letter to Howard Barnum and Tony Sudbury, August 18, 2003, quoted in *Scientific American*, Sept. 2004

I wish to express my sincerest gratitude to my advisor Chris Fuchs, for having provided me the opportunity and the privilege of working with him in my pursuit of a doctoral degree. I will always feel greatly indebted to him, since he was willing to take a chance on me four years ago when I had been looking for a supervisor for almost a year to no avail. Chris is an excellent advisor, who cares not just about how much work his students get done but also in their general well-being. I have personally benefitted from his valuable insights and sage advice on what problems to look at or what next steps to consider. His knowledge, guidance, patience and support have been invaluable to my research and studies. But I think just as important is his overall influence in my way of thinking and approach to physics and life. I feel some of the important lessons I have learned from him, which are worth mentioning here:

(i) Do not be afraid to ask silly questions. True, they are usually a waste of other people's time but occasionally they turn out not to be silly at all.

(ii) Be constructive, not just productive. Do not fall into the trap of busywork.

(iii) When you obtain a result that is simpler than you expected, find out why.

(iv) Do not be satisfied with vague claims based on intuition. Always demand precision and some rigorous mathematics.

(v) But while mathematics is important, remember that often the most profound ideas in physics can be expressed meaningfully with a few, well-chosen words.

It is my sincere hope that these lessons, taken to heart, bring me a few steps closer to the kind of physicist he is.

I would like to extend my special thanks to Marcus Appleby, Åsa Ericsson, Hoan Dang, and Matthew Graydon, friends and comrade-in-arms in research topics related to SICs. To Marcus and Åsa, I am grateful for their willingness to discuss many of my ideas, most of which did not work, and for sharing their valuable insight into various problems that

I have struggled with in my research. To Hoan and Matthew, I am grateful for all the time we spent as students of Chris, for all the enlightening and enjoyable conversations on both physics and non-physics matters alike, and for the tradition of attending APS March Meetings, which I hope continues for many more years.

I would like to thank Norbert Lütkenhaus, Kevin Resch, Joseph Emerson, and Andrew Childs for serving in my PhD committee and for always making sure I was always on track in fulfilling the requirements of my program. I would also like to thank Ingemar Bengtsson for acting as the external examiner for my thesis defense. His work on the geometry of quantum states has been a source of inspiration to my research and it is an honor and a privilege to have him evaluate my contributions to the field.

Each of the following have in one way or another, influenced this thesis, sometimes directly in its content but even if only in little ways that affected its emphasis on certain ideas: Berge Englert, Zach Medendorp, Christoph Schaeff, Bill Wootters, Rob Spekkens, Lucien Hardy, Jeff Bub, John Watrous, Ben Reichardt, Frank Wilhelm, and Huangjun Zhu.

I am grateful to the Perimeter Institute for providing a stimulating environment for research, and its staff in particular, for the numerous times they have assisted me on administrative matters. I would like to mention in particular my officemates Jeff Hnybida and Katja Reid, for many delightful and stimulating conversations about their own work and other things, which provide a short reprieve from the mostly tedious part of research that happens on my desk.

Special thanks to Alfonso Cesar Albason, a companion from distance in my journey through my graduate studies. Getting a PhD is hard when you have many concerns and distractions beyond research and having a friend to talk to who understands the experience is both helpful and encouraging.

My continuing education would not be possible without the unwavering support and understanding of my dearest family and closest friends, for the many times they have encouraged me to carry on when things became difficult, and for relieving some of the pressures of those hard and trying moments. Many thanks to them.

To the loving memory of my father, Angel Garrucha Tabia, Jr.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Why does quantum mechanics have the structure that it has? In particular, why are quantum systems most convenient to describe in terms of linear operators and complex-valued probability amplitudes? John Wheeler tells us that it follows from two distinctive features of quantum phenomena: a complementary choice of questions to ask each quantum system, and answers to those questions that must be deduced from probabilities [129].

As Wheeler points out, Ronald Fisher has taught us that the correct number to determine from experiments is not the probability of the desired answer but its square root [38]. The need to adequately distinguish between nearly identical sets of measurement statistics is what then leads to the notion of statistical distance. In Hilbert space, this translates into the angle θ between two probability amplitude vectors $|\psi\rangle$ and $|\phi\rangle$ given by

$$\theta = \arccos |\langle\psi|\phi\rangle|. \tag{1.1}$$

Ernst Stückelberg showed that probability amplitudes that are exclusively real are incompatible with Heisenberg's uncertainty relations [123]. Equivalently, Wheeler has argued that this means that they are incompatible with Bohr's complementarity principle.

Yet since the only truly observable quantities in the theory are the outcome probabilities of measurements and not the amplitudes, it must be possible to understand the Hilbert space structure solely in terms of probabilities. William Wootters gives such a description in terms of the probability tables of mutually unbiased bases, where the role of complex amplitudes in transition probabilities are replaced by "coincidence sums" between probability tables [133]. In this thesis, we formulate quantum theory using probability distributions obtained from symmetric, informationally complete (SIC) measurements and highlight the geometric properties of quantum states exhibited in this framework.

1.1 Historical background and motivation

In a 1926 paper, Max Born proposed the statistical interpretation of quantum mechanics that is with us to this day [20]. In his discussion of a scattering problem, he mentions in a footnote that the probability of a specific outcome is obtained from the square of the Schrodinger wave function. In modern parlance, we would write

$$p(x) = |\langle x|\psi\rangle|^2 \tag{1.2}$$

for the probability of getting outcome x given that the system being measured is described by the state $|\psi\rangle$. Put differently, $p(x)$ gives the transition probability of finding the system in state $|x\rangle$ when it was in state $|\psi\rangle$ before measurement. Once probabilities enter into the picture, we know that we are dealing with a statistical theory, where the outcomes of certain measurements are nondeterministic.

Even in the formative years of quantum mechanics, physicists wondered about the physical meaning behind this underlying indeterminism. For his part, Werner Heisenberg argued that quantum randomness is a consequence of having to ascribe a conjugate pair of variables to each physical degree of freedom, where the simultaneous values of the two variables can only be determined with some irreducible uncertainty [62]. In his matrix mechanics, this follows from the canonical commutation relation

$$[X, P] = i\hbar, \tag{1.3}$$

where X and P refer to any such conjugate pair. Here we clearly see that complex numbers play a necessary and prominent role in quantum theory.

Around the same time as this particular assertion by Heisenberg, Niels Bohr formulated the principle of complementarity [18], a concept concerning wave and particle aspects of quantum systems. To Bohr, the uncertainty relations were manifestations of the deeper principle of complementarity, a claim duly noted by Heisenberg in an addendum to his paper on the uncertainty principle. Complementarity can be made formally precise in which-way experiments in quantum optics [120], where the complementary observables for path distinguishability D and interference pattern visibility V have been shown to obey the duality relation [35]

$$D^2 + V^2 \leq 1. \tag{1.4}$$

However, despite its profound significance, the complementarity principle does not really help us understand why the probabilities are obtained from complex amplitudes. To do that, we need a more general and systematic way to explore the function of complex numbers in the theory, which we gain when we use Hilbert spaces.

Hilbert spaces were first introduced into quantum mechanics by David Hilbert, Lothar Nordheim, and John von Neumann [63] in 1927 in a collaborative effort to make the transformation theory of Pascual Jordan and Paul Dirac more mathematically rigorous. The familiar form we learn today was first established by von Neumann in a follow-up paper [127].

In finite dimensions, Hilbert space is just a complex vector space supplied with an inner product. Hilbert space represents a significant and useful advance in quantum theory because much of the geometric intuition learned from Euclidean geometry still applies to it. For example, an exact analog of the Pythagorean theorem holds for Hilbert spaces. Also, vectors in a Hilbert space can be uniquely specified with respect to a coordinate system in exactly the same way Cartesian coordinates can be assigned to points on a plane. Linear operators on Hilbert spaces are likewise easy to visualize as certain transformations that stretch or shrink the space in different directions.

To connect the abstract notions of Hilbert space to concrete facts about measurements and observations in quantum theory, Garrett Birkhoff and John von Neumann developed the idea of a logic of questions about properties of a quantum system that can be checked with yes-no measurements [17]. They demonstrated that quantum properties exhibit the algebraic structure of an orthocomplemented lattice, which translates into a representation of states in terms of projection operators on a Hilbert space.

While the discovery of quantum logic is a nice and important result, it would be hardly enlightening or satisfying for anyone to claim that the laws of quantum mechanics originate from the distinct, non-Boolean flavor of its logical propositions. What we really desire is a more intuitive way to understand the role of the complex amplitudes in the indeterminism implied by quantum laws. In particular, what sort of statistical theory is implied by the Hilbert space structure of quantum mechanics?

Bogdan Mielnik taught us that the true lesson from quantum logic is that there is a more general way to describe the statistical properties of quantum ensembles using geometrical concepts. In particular, he showed that if we interpret logical propositions in terms of filters that transmit beams of particles of a specific type, the Born rule becomes a purely geometric constraint on the absorption coefficient, or transition probability, between pairs of filters [99]. The resulting space of transition probabilities is equivalent to Hilbert space if the structure of its 2-dimensional subspaces lead to the usual linear superposition principle.[1] This allows us to consider state spaces whose geometries do not admit a Hilbert space representation, something that might be relevant in attempts to develop a working theory of quantum gravity.

[1] Details on this can be found in Appendix E.

In general, we can assign to any statistical theory a geometric description where states are represented by points of a convex set S, which we consider to be a subset of the real Euclidean space $\mathbb{R}^n$. If we take a pair of states $\vec{x}_1, \vec{x}_2 \in S$ then the state $x = p_1\vec{x}_1 + p_2\vec{x}_2$ denotes a mixture of the states $\vec{x}_1$ and $\vec{x}_2$ in proportions given by the probabilities p_1 and p_2, respectively. Because the set is convex, any point lying on the line segment joining any pair of points in S also belongs in S. On the boundary of S, there are special points that can not be expressed as a nontrivial combination of other states and are called extreme points. Because these points do not represent mixtures of different states, we call them pure states. We assume that the set S is compact, which means that every state corresponding to some point in S can be written as a finite convex combination of pure states.

In classical mechanics, the convex set describing the state space of a system with N perfectly distinguishable states is given by the set of probability vectors on a regular simplex Δ^{N-1}. The vertices of the simplex are associated with classical pure states, and any mixture has a unique decomposition in terms of these pure states. In a typical setting of classical dynamics where states refer to continuous degrees of freedom, such as position and momentum, the simplices are not only infinite-dimensional, they also possess a kind of symplectic structure that accommodates canonical transformations by Hamiltonian functions.

In quantum mechanics, the simplest convex set is a 3-dimensional ball called the Bloch ball; it is the state space for 2-dimensional quantum systems, or qubits. The points on the surface of the Bloch ball represent pure states while the interior corresponds to mixtures. Relations among qubit density operators on a Hilbert space translate into specific geometric properties of the Bloch ball. For instance, orthogonal states correspond to antipodal points on the sphere. Also, any mixture can be obtained from mixing any pair of pure states lying on the line segment passing through the point representing that mixture. Thus, even in the simplest case, the geometry of the convex set of quantum states expresses the indistinguishability of quantum mixtures, a phenomenon shared by any statistical theory with a state space that is not equivalent to a probability simplex. That is, given any mixture, it is impossible to physically distinguish between two different ways of combining states that lead to the same mixed state.

Although the geometry of the Bloch ball is elegant, it is actually too simple. In fact, there is a specific sense in which it is not representative of the convex geometry we expect from quantum theory. This is embodied in two important foundational results concerning the structure of probabilities in quantum mechanics: Gleason's theorem [51] and the Bell-Kochen-Specker theorem [9, 82]. Gleason's theorem states that the Born rule specifying the probabilities for the outcomes of quantum measurements is the only possible way to obtain them from density operators. In particular, Andrew Gleason proved that the

theorem holds in all finite dimensions if and only if it holds in $d = 3$. On the other hand, the Bell-Kochen-Specker theorem states that assigning simultaneous non-contextual values to all quantum observables is impossible. It can be proven as a corollary to Gleason's theorem [22]. Together the theorems suggest that an accurate picture of the complex structure of quantum states can only be gleaned from Hilbert spaces that are at least 3-dimensional. Thus, in formulating a geometric framework for understanding quantum probabilities, we will employ qutrits as our case study for explicit calculations and analysis.

Several authors have examined cross-sections of qutrit state space in order to gain some insights into the rich, intricate geometry of quantum states, most notably the work by Sandeep Goyal, et. al, which gives a comprehensive analysis of the 2- and 3-dimensional cross-sections of the generalized Bloch ball [56]. An important part of their results clarifies the connection between a cross-section forming an obese tetrahedron and the tetrahedral symmetry represented by a group of signed permutation matrices. Gniewomir Saribicki and Ingemar Bengtsson supplement Goyal, et al. by classifying all possible 2-dimensional cross-sections of a qutrit in terms of the plane cubic curve that defines the boundary of every such cross-section [115].

Despite the great strides made in recent years in our understanding of the convex geometry of qutrits, much of what we have determined involves specific observations that do not necessarily clarify why, for instance, the space of quantum states have a positivity structure and unitary symmetry, and what are the implications of these features for the corresponding statistical theory. Particularly, we are still searching for a simple and intuitive geometric characterization of shape of qutrit state space that generalizes to all finite dimensions, similar to how we know that all classical state spaces are regular simplices.

It may be worth mentioning here that despite the somewhat limited scope of examining just qutrits, a better understanding of qutrits is important even beyond their relevance in foundational studies of quantum mechanics. In fact, qutrits play essential roles in various applications of quantum information theory. This may come as a bit of a surprise since the vast majority of quantum information processing is designed for and achieved by manipulating qubits but there are several instances when qutrits are necessary in principle. Some examples are

(a) the optimal solution to the Byzantine agreement problem requires entangled qutrits [39, 48],

(b) a quantum computer with a finite number of distinguishable states has a maximal Hilbert space if it is partitioned into qutrits [58],

(c) threshold schemes for quantum secret sharing work only if the quantum system used for the secret is at least 3-dimensional [28],

(d) quantum key distribution with qutrits provides increased coding density and a higher security margin for errors [59],

(e) recently it has been shown that the smallest possible heat engines involves a qutrit with spatially separated basis states [86], and

(f) qutrits are more robust under certain forms of decoherence, making them suitable for storing quantum information [96].

We end this section by noting that it is possible to explore the geometry of quantum states directly if one takes the set of rank-1 projection operators. In this context, one can then examine the projective geometry of quantum mechanics analogous to the classical phase space geometry of Hamiltonian dynamics. This is, in fact, the subject matter of what is known in the literature as *geometric quantum mechanics*. Tom Kibble is usually credited for pioneering this geometric approach to quantum theory, where in a series of papers in the 1970s, he showed that quantum mechanics can be formulated in terms of a Hamiltonian theory where quantum phase space is the symplectic manifold of pure state trajectories [76].

Inspired by Kibble, Dorje Brody and Lane Hughston provide a detailed study of the interplay between aspects of quantum theory such as linear superposition, entanglement, and uncertainty relations and the geometrical properties of the complex projective space $\mathbb{C}\mathrm{P}^{d-1}$ of d-dimensional pure states with distances measured by the Fubini-Study metric [21]. Some key insights gained from the rich geometry of the Fubini-Study manifold are

(a) the Fubini-Study metric determines the transition probability $|\langle\phi|\psi\rangle|^2$ between any 2 pure states $|\psi\rangle$ and $|\phi\rangle$,

(b) the complex projective space is rigid: any bijective map from the space of pure states to itself that preserves the Fubini-Study distance describes a unitary or anti-unitary transformation on quantum states,

(c) curves in $\mathbb{C}\mathrm{P}^{d-1}$ are just solutions to the Schrödinger equation and the gauge-invariant Berry phase is a consequence of the parallel transport condition on these curves

(d) the geodesic distance of any point in the manifold to the nearest disentangled state with respect to a Hopf mapping provides a measure of entanglement[2],

[2]We provide some details of this in Appendix F.

(e) if we think of $\mathbb{C}\mathrm{P}^{d-1}$ as a real manifold Γ with twice the dimension, then the structure of Γ is such that geometric inequalities for vector fields defined on Γ are equivalent to Heisenberg uncertainty relations for quantum observables, and

(f) geometric notions on the manifold of pure states can be extended to mixed states when one considers purifications on a larger Hilbert space; for example, the Bures-Uhlmann metric generalizes the notion of distances to all quantum states [126].

Although these are interesting observations, they do not help explain the why projective geometry applies to quantum mechanics, since here it is assumed at the onset. What the quantum phase space approach is most useful for is when we consider nonlinear modifications to the Schrödinger equation. Most of the general features of the complex manifold carry over or can be adapted quite naturally to many aspects of the nonlinear regime.

1.2 Research scope and objectives

This thesis aims to establish some of the defining features of the of convex set of quantum states by examining the structures and symmetries that come into view when quantum states are expressed in terms of the probabilities for the outcomes of a symmetric, informationally complete (SIC) measurement, a special class of measurement whose elements are rank-1 projectors that form a set of d^2 equiangular lines in Hilbert space. By representing quantum states in terms of SICs, the full machinery of probability theory becomes immediately applicable to the investigation of the properties of quantum state space.

Note that SICs also appear in various applications of quantum information, where they typically play an important role in quantum state tomography and quantum cryptography. Thus, there is a strong interconnection among the foundational aspects surrounding the nature of quantum states, the mathematical aspects regarding the geometry of quantum state space, and the practical aspects concerning the manipulation of quantum information, and SICs seem to play a vital, and potentially central, role in linking them together.

The nontrivial character of the geometry of quantum states is reflected most simply in 3-dimensional quantum systems and so this research is devoted primarily to understanding the structure of qutrits in all their intricate details, with some emphasis on the geometrical aspects. Particularly, the thesis attempts to address the following issues:

(i) *Understand basic properties of SICs:* Part of this research is concerned with unraveling the full symmetry of SICs by studying their connection with different areas

in mathematics such as projective geometry, finite field theory, and design theory. Such results may be of vital importance in efforts to construct a rigorous proof for the existence of SICs in all dimensions. However, in this thesis, we focus our attention on the geometric aspects of qutrit SICs with Weyl-Heisenberg symmetry, the details of which are most relevant in our subsequent analysis of the structure of the corresponding state space.

(ii) *Explore qutrits as a test bed for identifying general features of quantum state space:* The structure of quantum probabilities follows from Gleason's theorem, a result for dimensions $d \geq 3$ that assigns a unique probability measure to quantum states through density operators. This suggests that the 3-dimensional case already contains the essential ingredients of the complex structure of quantum theory, making it appropriate for initial studies. Some of our key results in this regard include an elegant formula for describing pure qutrits, a full listing of geometric invariants for qutrit SICs and the rotational transformation between the probability vectors of distinct qutrit SICs, and a characterization of the shape of qutrit state space in terms of the radial distance of boundary states. Several results provide natural stepping stones for extended work in higher dimensions.

(iii) *Develop experimental schemes for direct implementation of SIC measurements:* The vast majority of experiments in the lab involve projective measurements, which is why they are considered standard quantum measurements. However, advances in quantum information technology have shown that more ingenious measurements are possible and it has only been recently that technology is up to the task in performing them. Devising a practical implementation of SICs should provide a big boost for motivating them as canonical measurements. More generally, methods for realizing SICs in practice are potentially important for advancing experimental studies on higher-dimensional quantum systems.

(iv) *Reconstruct quantum theory from simple postulates:* A better knowledge of the convex geometry of quantum states would highlight which basic features exhibit the probabilistic structure of quantum theory, allowing us to deduce a set of physically motivated assumptions sufficient for recovering quantum mechanics from some modification of classical probability theory. The specific approach we use takes the Born rule as a normative postulate that formally leads to the idea of quantum state space as a maximal consistent set.

1.3 Roadmap for the thesis

The thesis provides a detailed analysis of the geometric structure of the convex body of quantum states obtained from a representation of states in terms of SIC probabilities.

Chapter 2 presents an overview of quantum mechanics. It focuses on establishing the quantum formalism in finite dimensions and discusses some general properties of the space of density operators for qubits and higher dimensions, with emphasis on geometric features such as its convex unitary structure and the Hilbert-Schmidt distance between two operators. It includes a section on how quantum probabilities are obtained from an algebraic perspective.

Chapter 3 introduces symmetric informationally complete measurements in general and Weyl-Heisenberg covariant qutrit SICs in particular. We define SICs as they appear in different contexts and characterize the group symmetry of SICs that are generated from fiducial vectors. We classify qutrit SICs into eight families according to the independent orbits generated by the extended Clifford group and identify a SIC with the nicest properties as the most exceptional SIC. We discuss various geometric aspects of SICs such as the triple products that characterize unitary equivalence between SICs, the Lie algebraic properties of the imaginary part of the triple products, and a linear dependency structure associated with the Weyl-Heisenberg symmetry of qutrit SICs.

Chapter 4 describes various practical methods for implementing SICs. It includes the earlier polarimetry experiments of Thomas Durt, et al. on Wigner qubit tomography and a more recent experiment by Zach Medendorp, et al. using an optical storage loop with weak projections for simulating a qutrit SIC-POVM. We also discuss proposed methods for future experiments, one by Amir Kalev, et al. that describes a 2-step procedure with a fuzzy measurement followed by a projection onto the Fourier basis; the other one is my own proposed scheme using multiport devices designed to perform the Naimark extension, with specific designs for qubit and qutrit SICs.

Chapter 5 describes the SIC representation of quantum states, which we first motivate in the framework of Quantum Bayesianism or QBism, where quantum states are characterized as objects representing an agent's degrees of belief about the future behavior of a quantum system. We then proceed to define pure quantum states in terms of SIC probabilities and observe how this leads to a definition of pure states in terms of fixed and stationary points of certain linear maps. We also see how other physical quantities appear when expressed in the SIC language, such as the fidelity function used for distinguishing quantum states and purification for mixed states. We also include a short discussion on entanglement in SIC terms, although limited to the partial positivity criterion for 2-qubit systems. We

end with a description of maps between density operators in terms of affine maps on the corresponding SIC probability vectors.

Chapter 6 uses the SIC representation of qutrits to examine the geometric features of qutrit state space. We see how the pure state conditions for the most exceptional SIC lead to an elegant formula for qutrit pure states. We construct the orthogonal transformation between the probability vectors of distinct SICs for the same quantum state in terms of a 3-dimensional circulant rotation matrix. We describe a permutation symmetry for qutrit SIC probabilities that provides some insight into the finite affine plane symmetry of qutrit pure states. We determine the boundary states of qutrits in terms of radial distance from the uniform distribution. We end with a discussion of the plane sections of SIC probabilities, which are shown to be equivalent to those obtained in the generalized Bloch representation.

Chapter 7 introduces the notion of a maximal consistent set in the context of postulating a "quantum law of total probability" for extending classical probability theory to quantum theory. We define a general notion of maximal consistent sets and show that classical and quantum state spaces are important examples. We also describe general properties of these maximal consistent sets concerning convexity, boundary features, symmetry groups and the notion of duality. We describe a method for reconstructing qutrit state space from a finite consistent set, where we show that a symmetry related to the Hesse configuration arises quite naturally.

Chapter 8 provides a summary of our key results, a list of problems recommended for future investigations, and some concluding remarks.

1.4 List of specific contributions

The results presented in this thesis representing my own specific contributions are contained in the following sections:

(i) Section 4.4 is based on GNM Tabia, *Phys. Rev. A*, **86** (2012) 062107.

(ii) Sections 3.3, 3.4, 5.5, 6.1, 6.2, and 6.4 are based on GNM Tabia and DM Appleby, *Phys. Rev. A*, **88** (2013) 012131.

(iii) Sections 3.5, 3.6, 3.7, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 7.4, and 7.5 are based on unpublished research notes.

Chapter 2

Overview of quantum mechanics

Quantum mechanics is considered to be one of humanity's most impressive intellectual achievements. It certainly seems to be as mystifying as it is remarkable. To date, it is our most accurate scientific theory, capable of making predictions on the behavior of atomic scale objects with incredible precision. The mystery lies in what sort of picture quantum theory paints about physical reality and a means of understanding in clear, unambiguous terms why Nature behaves the way it does.

Part of the reason why the interpretation remains unresolved is because the mathematical structure of the theory, as established in the pioneering works of Erwin Schrödinger, Werner Heisenberg, Paul Dirac, and John von Neumann, does not naturally lead to a world view that is fully consistent with intuitions developed in classical mechanics. This thesis is part of an attempt to better understand the implications of quantum theory as a form of probability calculus that applies to physical systems which can manifest nonclassical phenomena such as interference and entanglement.

But in order to gain some appreciation for both the elegant and puzzling aspects of quantum theory, it is necessary to get acquainted with the mathematics it employs, the framework of Hilbert spaces. Because we will not be concerned with systems with continuous degrees of freedom, we focus our attention on Hilbert spaces of finite dimensions, which is easily understood using methods of matrix algebra with complex numbers. Here we present a brief summary of the mathematical formalism in Dirac notation, some geometric aspects of the set of quantum states, and the introduction of probabilities in quantum mechanics from an algebraic perspective.

2.1 Outline of the quantum formalism

The traditional backdrop of quantum mechanics is the vectors of a Hilbert space $\mathcal{H}$, which represent the set of pure states for a quantum system. In Dirac notation, a vector is denoted by ket $|\psi\rangle$. To each ket there corresponds a dual vector, denoted by a bra $\langle\psi|$. In finite dimensions, $|\psi\rangle$ can be thought of as a column vector with d complex entries,

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix}, \tag{2.1}$$

where the entries refer to the components of the vector with respect to some orthonormal basis, usually just the standard basis, while $\langle\psi|$ is the row vector whose entries are complex conjugate to that of $|\psi\rangle$, that is,

$$\langle\psi| = \begin{pmatrix} \alpha_1^*, & \alpha_2^*, & \ldots, & \alpha_d^* \end{pmatrix}. \tag{2.2}$$

Thus, a d-dimensional Hilbert space $\mathcal{H}^d$ is effectively the complex vector space $\mathbb{C}^d$ equipped with a scalar product. More precisely, a pure state corresponds to an equivalence class of vectors $e^{i\phi}|\psi\rangle, \phi \in \mathbb{R}$, all physically equivalent. In effect, the true space of pure quantum states is represented by the complex projective space $\mathbb{C}\mathrm{P}^{d-1}$.

The Dirac notation is such that the inner product is naturally denoted by a bracket $\langle\psi|\phi\rangle$, e.g., if we have

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \text{ and } |\phi\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \qquad \text{then } \langle\psi|\phi\rangle = \alpha_1^*\beta_1 + \alpha_2^*\beta_2. \tag{2.3}$$

It is also quite natural to represent the projection operator that projects onto the subspace spanned by $|\psi\rangle$ by

$$|\psi\rangle\langle\psi| = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} \alpha_1^*, & \alpha_2^* \end{pmatrix} = \begin{pmatrix} |\alpha_1|^2 & \alpha_1\alpha_2^* \\ \alpha_1^*\alpha_2 & |\alpha_2|^2 \end{pmatrix}. \tag{2.4}$$

Physical observables are described using Hermitian operators $A = A^\dagger$, which can be

written as

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{d1} & \dots & \dots & a_{dd} \end{pmatrix} = \sum_i a_{ij} |i\rangle\langle j|, \tag{2.5}$$

where $|i\rangle$ represent standard basis column vectors, i.e.,

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad |d\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{2.6}$$

while $a_{ii} \in \mathbb{R}$ and $a_{ij}^* = a_{ji}$, that is, diagonal entries are real and off-diagonal entries that are transposes of each other are complex conjugates.

Let a_i denote the eigenvalues of A with corresponding eigenvectors $|\psi_i\rangle$. We can always choose the set of eigenvectors to be an orthogonal set so that

$$A\,|\psi_j\rangle = \left(\sum_i a_i |\psi_i\rangle\langle\psi_i| \right) |\psi_j\rangle = \sum_i a_i \delta_{ij}\,|\psi_i\rangle = a_j\,|\psi_j\rangle\,, \tag{2.7}$$

where we used the so-called spectral decomposition of A,

$$A = \sum_i a_i |\psi_i\rangle\langle\psi_i|. \tag{2.8}$$

According to the spectral theorem, any normal operator $A^\dagger A = AA^\dagger$ can be diagonalized by some unitary operator and thus may be decomposed in this way.

If we have a quantum system described by state $|\phi\rangle$ and we wish to know the value of observable A for that system, we obtain the mean or expectation value $\langle A\rangle$ of A by sandwiching A by a bra and ket for the given state,

$$\langle A\rangle = \langle\phi|A\,|\phi\rangle = \sum_i a_i |\langle\psi_i|\phi\rangle|^2 = \mathrm{Tr}\left(A|\psi\rangle\langle\psi|\right). \tag{2.9}$$

Probabilities are expectation values themselves since, for instance, if we want to calculate the probability of value a_i for state $|\phi\rangle$ then

$$\Pr\left[A = a_i|\,|\phi\rangle\right] = |\langle\psi_i|\phi\rangle|^2 = \langle\phi|\left(|\psi_i\rangle\langle\psi_i|\right)|\phi\rangle\,. \tag{2.10}$$

Now, if instead the system was prepared as an equal mixture of $|\phi_1\rangle$ and $|\phi_2\rangle$, we would compute $\langle A\rangle$ to be

$$\begin{aligned}\frac{1}{2}\langle\phi_1|A\,|\phi_1\rangle + \frac{1}{2}\langle\phi_2|A\,|\phi_2\rangle &= \mathrm{Tr}\left[A\left(\frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2|\right)\right] \\ &= \mathrm{Tr}\,(A\rho) \end{aligned} \tag{2.11}$$

where

$$\rho = \frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2| \tag{2.12}$$

is the mathematical object for describing mixtures of pure states and is called the density or statistical operator. The density operator ρ summarizes all that we know about the preparation of the quantum system and the outcome probabilities for any measurement on it.

Formally, a density operator ρ corresponds to a quantum state if it satisfies the following conditions:

$$\rho^\dagger = \rho, \qquad \rho \geq 0, \qquad \mathrm{Tr}\,(\rho) = 1. \tag{2.13}$$

If $\rho^2 = \rho$, then ρ is a projection operator and it corresponds to a pure quantum state, which we may write as

$$\rho = |\psi\rangle\langle\psi| \tag{2.14}$$

In general, a density operator can be expressed as a convex combination of pure states in a non-unique way. Nonetheless, for any particular decomposition, ρ can be written as

$$\rho = \sum_i p(i)|\psi_i\rangle\langle\psi_i| \tag{2.15}$$

where $p(i)$ denotes the weight of each state $|\psi_i\rangle$ in the convex combination.

In practice, the value of any observable is determined by performing a measurement. In quantum mechanics, a general measurement is represented by a set of positive operators $E_i = M_i^\dagger M_i$ such that

$$\sum_i E_i = I. \tag{2.16}$$

This is called a positive operator valued measure, or POVM. When a system is measured using the set $\{E_i\}_{i=1}^N$, the post-measurement state $\rho^{(j)}$ when you get outcome j in an ideal measurement is given by

$$\rho^{(j)} = \frac{M_j \rho M_j^\dagger}{p(j)}, \tag{2.17}$$

where $p(j) = \mathrm{Tr}\,(\rho E_j)$ is the probability for outcome j.

In the special case that $E_i E_j = E_j \delta_{ij}$, the measurement consists of an orthogonal set of projections. This is known as a projective or von Neumann measurement. In this case, we may write

$$E_i = |e_i\rangle\langle e_i|, \qquad \langle e_i|e_j\rangle = \delta_{ij}, \qquad E_i^\dagger E_i = E_i \tag{2.18}$$

so that the post-measurement state corresponding to outcome j is just

$$\rho^{(j)} = \frac{E_j \rho E_j^\dagger}{p(j)} = |e_j\rangle\langle e_j| = E_j. \tag{2.19}$$

Density operators assign probabilities to measurement outcomes. However, even for pure states, probabilities not equal to 0 or 1 appear as soon as the observable A being measured has

$$[A, \rho] = A\rho - \rho A \neq 0, \tag{2.20}$$

where $[A, \rho]$ is called the commutator of A and ρ. A pure state $|\psi\rangle$ therefore behaves like a pure classical state only with respect to observables that commute with $|\psi\rangle\langle\psi|$.

Finally we consider how quantum states evolve in time. The dynamics of pure states obeys the Schrodinger equation

$$i\hbar \frac{\partial\,|\psi\rangle}{\partial t} = H\,|\psi\rangle \tag{2.21}$$

where H is called the Hamiltonian; typically, it is the observable that describes the total energy of the system. Because H is a Hermitian matrix, the Schrodinger equation implies unitary time evolution for pure states.

By linearity, the time evolution of density operators is given by

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]. \tag{2.22}$$

Observe the similarity of Eq. (2.22) with the Liouville equation for the phase space distribution function $\rho^{(C)}$:

$$\frac{\partial \rho^{(C)}}{\partial t} = \left\{H, \rho^{(C)}\right\} \tag{2.23}$$

where we have a Poisson bracket instead of a commutator. In fact, the process of deriving quantum laws via the replacement of classical dynamical variables with linear operators and Poisson brackets by commutators is called canonical quantization, which was first introduced by Paul Dirac in his doctoral thesis as a classical analogy for quantum mechanics. Of course, it must be mentioned that the particular connection being made here for the dynamics holds in the infinite-dimensional case but not in the finite-dimensional ones, where there is no obvious way to define a Poisson bracket.

2.2 Qubits and the Bloch ball

It is one thing to say that the space of d-dimensional pure states is the complex projective space $\mathbb{C}\mathrm{P}^{d-1}$; it is another thing to point out the one-to-one correspondence between elements of that space and an actual configuration of the quantum system modeled with it. To illustrate how physical statements can be made from the quantum formalism, we consider the specific example of a spin-1/2 particle.

Spin is a quantum property associated with an intrinsic angular momentum possessed by particles such as electrons, atoms, etc. A Stern-Gerlach experiment reveals the spin of a beam of particles as a deflection of the beam as it is sent through an inhomogeneous magnetic field. For spin-1/2 systems, every pure state corresponds to a unique direction in physical space. Hence, it can be described by a point on a unit sphere, with the surface representing all possible directions in 3 dimensions. Formally, we say that

$$S^2 \cong \mathbb{C}\mathrm{P}^1. \tag{2.24}$$

In quantum mechanics, this sphere is called the Bloch sphere. Opposite points on the Bloch sphere are associated with orthogonal states and convex mixtures of pure states are represented by points inside it. Thus, the Bloch ball, i.e., the spherical surface and its interior, corresponds to the density operator space for spin-1/2 systems.

Because orthogonal states occur in pairs, a spin-1/2 system is an example of a 2-level quantum system. Any quantum system with two degrees of freedom, which correspond to a pair of fully distinguishable states, are mathematically equivalent. Such a system will have a state space represented by a Bloch ball and are called qubits, in contrast with 2-level classical systems, which are called bits.

Density operators for qubit states are often written in the form

$$\rho = \frac{1}{2}\left(I + \vec{n}\cdot\vec{\sigma}\right) \tag{2.25}$$

where $\vec{n}$ is a vector associated with a point on the Bloch ball and $\vec{\sigma}$ is a vector of Pauli matrices,

$$\vec{\sigma} = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix}, \tag{2.26}$$

which are typically chosen to be

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.27}$$

The eigenvalues of ρ are given by

$$\lambda = \frac{1}{2}\left(1 \pm \|\vec{n}\|\right) \tag{2.28}$$

which tells us that $\|\vec{n}\| \leq 1$ for positivity, with equality for pure states.

The explicit connection between pure states and the surface of the sphere can be made if we consider

$$|\psi\rangle = e^{i\alpha}\begin{pmatrix}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}\end{pmatrix} \tag{2.29}$$

where $\alpha \in \mathbb{R}$ is an overall phase which has no physical significance. In the Bloch sphere, this pure state is represented by the unit vector

$$\vec{n} = \begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix}. \tag{2.30}$$

The simplicity of the Bloch sphere is actually misleading when we look at quantum state space in higher dimensions. The primary difference is that the positivity constraints on ρ lead to more complicated convex bodies for Hilbert spaces with $d \geq 3$. It is not difficult to write down the equations but they are not very illuminating. So instead we will consider more general properties of these higher dimensional quantum state space.

2.3 The set of density operators

Recall that a complex $d \times d$ matrix ρ is a density operator if and only if

$$\rho^\dagger = \rho, \qquad \rho \geq 0, \qquad \mathrm{Tr}\,(\rho) = 1. \tag{2.31}$$

We denote the set of density operators by $\mathcal{D}^d$; it is a convex subset in the vector space of Hermitian matrices, with pure states $\rho^2 = \rho$ that form a projective Hilbert space. Geometrically, $\mathcal{D}^d$ is the intersection of the cone of positive operator of unit trace with the hyperplane of unit trace Hermitian matrices [14].

Consider a d-dimensional Hilbert space $\mathcal{H}^d$. There is a dual space $\mathcal{H}^{d*}$ defined by the space of linear mappings from $\mathcal{H}^d$ to $\mathbb{C}$. In finite dimensions, these 2 spaces are isomorphic to one another.

There is also an associated space of operators on $\mathcal{H}^d$, which is a Hilbert space itself when equipped with the inner product

$$\langle A, B\rangle = \kappa \mathrm{Tr}\left(A^\dagger B\right) \qquad \text{for } A, B \in \mathcal{B}\left(\mathcal{H}^d\right), \tag{2.32}$$

where $\kappa \in \mathbb{R}$ is just a possibly convenient scaling factor and $\mathcal{B}\left(\mathcal{H}^d\right)$ denotes the space of bounded linear operators on $\mathcal{H}^d$. Eq. (2.32) is known as the Hilbert-Schmidt inner product and it naturally gives rise to the distance measure

$$D_{\mathrm{HS}}(A, B) = \sqrt{\mathrm{Tr}\left[(A-B)(A^\dagger - B^\dagger)\right]} = \|A - B\|_{\mathrm{HS}} \tag{2.33}$$

called the Hilbert-Schmidt metric. $D_{\mathrm{HS}}(\rho, \sigma)$ measures the distance between density operators $\rho, \sigma \in \mathcal{D}^d$ analogous to the way the Euclidean metric

$$D(\vec{a}, \vec{b}) = \sqrt{\sum_{i=1}^{N} (a_i - b_i)^2}. \tag{2.34}$$

is used to define the distance between points $\vec{a}$ and $\vec{b}$, with Cartesian coordinates a_i and b_i, respectively.

The Hilbert-Schmidt metric induces a flat geometry on the space of density operators, from which we can identify a few basic properties [16]:

(i) The set $\mathcal{D}^d$ is a compact convex set in d^2-1 dimensions that is topologically equivalent to a ball.

(ii) The set of pure states is a connected set of $2d-2$ dimensions, which has zero measure with respect to the (d^2-2)-dimensional boundary.

(iii) For any 2 pure states, there exists a continuous path of transformations joining one pure state to another.

(iv) The 2-dimensional projections of $\mathcal{D}^d$ are polyhedral, that is, they can be obtained from the intersection of 2 boundary segments.

(v) The volume V_d and surface area A_d of $\mathcal{D}^d$ with respect to the flat metric are given by [138]

$$\begin{aligned} V_d &= (2\pi)^{d(d-1)/2} \sqrt{d} \left(\frac{\Gamma(1)\cdots\Gamma(d)}{\Gamma(d^2)} \right), \\ A_d &= (2\pi)^{d(d-1)/2} \sqrt{d-1} \left(\frac{\Gamma(1)\cdots\Gamma(d+1)}{\Gamma(d)\Gamma(d^2-1)} \right). \end{aligned} \tag{2.35}$$

If we consider the radius r of the largest sphere inscribed in $\mathcal{D}^d$, then

$$r = \frac{1}{\sqrt{d(d-1)}} \tag{2.36}$$

and so

$$\frac{rA_d}{V_d} = d^2 - 1, \tag{2.37}$$

meaning $\mathcal{D}^d$ can be thought of as a union of disjoint pyramids of constant height [138].

A generic way to furnish vector coordinates for $\mathcal{D}^d$ is to express density operators as

$$\rho = \frac{1}{N} I + \sum_{i=1}^{d^2-1} r_i \sigma_i \tag{2.38}$$

where $\{\sigma_i\}_{i=1}^{d^2-1}$ is the set of generators of $SU(d)$

$$\sigma_i \sigma_j = \frac{2}{N}\delta_{ij} + i \sum_k f_{ijk}\sigma_k + \sum_k g_{ijk}\sigma_k \tag{2.39}$$

where f_{ijk} are the structure constants, which are antisymmetric in the indices, and g_{ijk} are components of a totally symmetric tensor. The operators σ_i form the standard generalization of Pauli matrices for qubits so Eq. (2.38) is usually referred to as the generalized Bloch representation and r_i are the components of the generalized Bloch vector.

One convenient thing about the Bloch coordinates is that the origin or zero vector corresponds to the maximally mixed state

$$\rho_\star = \frac{1}{d} I, \tag{2.40}$$

which allows us to identify $\mathcal{D}^d$ with a subset of the Lie algebra of $SU(d)$.

Using Eq. (2.39), it is fairly straightforward to deduce that the pure state condition $\rho^2 = \rho$ becomes

$$\begin{aligned} \vec{r}^2 &= \frac{d-1}{2d}, \\ (\vec{r} \star \vec{r})_i \equiv \sum_{j,k} g_{ijk} r_j r_k &= \left(\frac{d-2}{d}\right) r_i. \end{aligned} \tag{2.41}$$

The first condition states that the Bloch vector is confined to a (d^2-2)-dimensional sphere while the second condition restricts pure states to a particular subspace of this out-sphere, a surface of dimension $2(d-1)$, which is considerably smaller than the dimension of the out-sphere for $d \geq 3$. Later in this thesis we shall see that for our approach that uses probability distributions obtained from a special measurement called a SIC-POVM as quantum states, the pure-state conditions look more symmetric both in its algebraic appearance and geometric implications.

An important property we know about the set of quantum states is that it forms a convex set. If we think of interior points of the set as convex combinations of extreme points, we can introduce probabilities as a measure of mixtures on the manifold of pure states. Also, if we imagine the space of density operators $\mathcal{D}^d$ as some rigid, convex body lying in $\mathbb{R}^{d^2-1}$, we can ask about any symmetries it might possess. For one, we know it has a global symmetry associated with $SU(d)$, since global unitary transformations leave the space invariant. In this convex picture, unitary invariance corresponds explicitly to a proper subgroup of the rotation group $SO(d^2-1)$. Not all possible rotations are included since quantum state space is never a sphere, except in the case of qubits, in which case we do have the symmetry of $SO(3)$.

The precise statement about the unitary symmetry of quantum state space is contained in a result due to Richard Kadison, which is concerned with operations that preserve the convex structure of $\mathcal{D}^d$ [66].

Theorem 2.3.1 (Kadison)**.** Let Φ be a bijective mapping from $\mathcal{D}^d \to \mathcal{D}^d$, i.e., from the space of density operators to itself, such that

$$\Phi(\alpha\rho_1 + (1-\alpha)\rho_2) = \alpha\Phi(\rho_1) + (1-\alpha)\Phi(\rho_2). \tag{2.42}$$

Then the mapping takes the form

$$\Phi(\rho) = U\rho U^\dagger \tag{2.43}$$

where U is either a unitary or anti-unitary operator.

For density operators, U must be a unitary operator since we can consider the infinitesimal form of the mapping and see that it must hold for the maximally mixed state $\rho_\star$, which is a multiple of the identity. In that case, only unitary operations maintain the convex structure globally.

One way to characterize the unitary operators is to say that they define a rotational symmetry for the set of quantum states that preserves the spectrum of ρ, that is, it leaves

the trace powers $\mathrm{Tr}\left(\rho^k\right), 1 \leq k \leq d$ unchanged. It is worth noting that rotations generally preserve only $\mathrm{Tr}\left(\rho^2\right)$, which corresponds geometrically to leaving the distance from any state ρ to the maximally mixed state $\rho_\star$ invariant. In even dimensions, a rotation typically has a single fixed point; a rotation that maintains the trace powers of ρ will have more. For example, if U has distinct eigenvalues then the set of fixed points consists of the convex hull of states that make U diagonal, which is also known as the eigenvalue simplex of U. The set of fixed points is larger when U is degenerate.

We end this section by remarking on how probabilities in quantum theory are obtained from density operators. The notion of state in quantum mechanics as we usually know it is actually determined by the algebraic structure of the closed linear subspaces of a Hilbert space. In the language of density operators, this is reflected by the fact that any state can be written as a convex combination of projection operators. What is not readily apparent from this is that the way in which probabilities arise from the formalism is actually severely limited. This is the main content of Gleason's theorem [51], a central result in quantum foundations.

If we think of probabilities as functions on the rays in a Hilbert space, we can make two basic assumptions:

(i) Elements $|i\rangle$ of any orthonormal basis are assigned probabilities $p(i)$ such that

$$\sum_{i=1}^{d} p(i)|i\rangle\langle i| = I. \tag{2.44}$$

(ii) The vector $|i\rangle$ may be an element of many different bases and the probability assigned to $|i\rangle$ is independent of the basis for it. For example, in $d = 3$, we can consider the orthonormal bases

$$B_1 = \{|0\rangle, |1\rangle, |2\rangle\}, \qquad B_2 = \left\{|0\rangle, \frac{|1\rangle + |2\rangle}{\sqrt{2}}, \frac{|1\rangle - |2\rangle}{\sqrt{2}}\right\} \tag{2.45}$$

and $\Pr[|0\rangle]$ should be the same whether $|0\rangle$ belongs to B_1 or B_2.

We can then prove that for a real or complex Hilbert space of dimension $d \geq 3$, there exists a density operator ρ such that

$$p(i) = \mathrm{Tr}\left(\rho|i\rangle\langle i|\right). \tag{2.46}$$

As long as we believe that quantum mechanics can be represented using Hilbert spaces then this is how probabilities are uniquely obtained from density operators. In a sense, the Born rule for calculating outcome probabilities is forced upon us by the algebraic structure of complex projective spaces.

Chapter 3

SIC-POVMs and Weyl-Heisenberg qutrit SICs

In quantum mechanics, the value of a physical observable is determined using a measurement, with outcomes represented by a set of positive semidefinite operators E_i that add up to the identity,

$$\sum_{i=1}^{N} E_i = I. \tag{3.1}$$

The set $\{E_i\}_{i=1}^{N}$ is formally known as a positive operator-valued measure (POVM).

Generally, measurement results correspond to relative frequencies for its various outcomes, which can be used to specify the probability for each outcome. When the probability distribution determined from the measurement statistics is sufficient for assigning a unique quantum state to the measured system, we say that the measurement is informationally complete. Since a density operator is specified by $d^2 - 1$ real parameters, any informationally complete POVM must have at least d^2 distinct elements satisfying Eq. (3.1).

If a POVM consists of exactly d^2 elements $E_i^2 = [\mathrm{Tr}\,(E_i)]^2\, E_i$ such that pairwise distinct operators have a constant trace overlap, i.e.,

$$\mathrm{Tr}\,(E_i E_j) = \text{const.} \qquad \text{for all } i \neq j, \tag{3.2}$$

then the measurement is an example of a symmetric informationally complete POVM or SIC-POVM. A SIC-POVM for qubits is depicted in Fig. 3.1.

In this chapter, we describe the standard method for constructing SIC-POVMs and explore several geometric properties associated with qutrit SIC-POVMs.

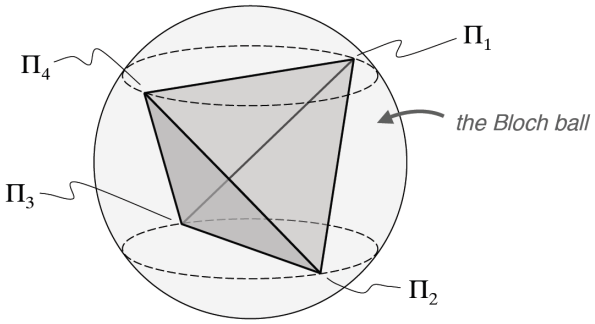


Figure 3.1: Qubit SIC-POVM on the Bloch sphere is represented by the vertices of an inscribed tetrahedron. Reproduced from [8].

3.1 Characterizing SIC-POVMs

SIC-POVMs have been studied extensively in the literature [3–7, 12, 24, 29, 40, 42, 52, 57, 75, 77, 80, 113, 118, 119, 128, 132, 134, 135], with some emphasis on the question of their existence in finite dimensions. Despite strong numerical evidence [118] that they exist for dimensions $d \leq 67$, the proof that they exist in all finite dimensions remains elusive.

SIC-POVMs are known by different names depending on context. In the study of projective spaces, they are known as equiangular lines. In the field of combinatorial designs, they correspond to minimal 2-designs. In frame theory, they are called maximally equiangular tight frames. Because each definition provides some useful insights into properties of SIC-POVMs, we briefly go through each of them here.

3.1.1 Equiangular lines

The most common way of defining SIC-POVMs is in terms of d^2 lines in a complex vector space $\mathbb{C}^d$. To every line $L_\psi \subset \mathbb{C}^d$, there is a unit vector $|\psi\rangle \in \mathbb{C}^d$ specifying its direction. In fact, every unit vector spans a subspace corresponding to some line in $\mathbb{C}^d$.

A set S of lines in $\mathbb{C}^d$ are said to be equiangular if for any two lines $L_\psi, L_\phi \in S$ associated with unit vectors $|\psi\rangle$ and $|\phi\rangle$, the angle θ between them is given by

$$\cos\theta = |\langle\phi|\psi\rangle|. \tag{3.3}$$

Infinitely many vectors specify the same line because $|\psi\rangle$ and $e^{i\alpha}|\psi\rangle, \alpha \in \mathbb{R}$ span the same subspace. However, the projection operator describing each line is unique since

$$P_\psi = \left(e^{i\alpha}|\psi\rangle\right)\left(e^{-i\alpha}\langle\psi|\right) = |\psi\rangle\langle\psi|. \tag{3.4}$$

In $\mathbb{C}^d$, the maximal number of lines that are equiangular is obtained by assigning projection operators to each line and demonstrating that those unit vectors are linearly independent. Since the lines in $\mathbb{C}^d$ are determined from the corresponding projection operators, then the space of interest is really the space of $d \times d$ Hermitian matrices to which the projections belong to, viewed here as a real vector space of dimension d^2.

Let Π_i be the projection operator associated with line L_{ψ_i}. Since $\Pi_i = |\psi_i\rangle\langle\psi_i|$, we know that $\mathrm{Tr}\,(\Pi_i) = 1$ and $\mathrm{Tr}\,(\Pi_i\Pi_j) = k \leq 1$. Now suppose there is a set $\{c_i\}$ such that

$$\sum_i c_i \Pi_i = 0. \tag{3.5}$$

If we take the trace on both sides, we get $\sum_i c_i = 0$. We can also multiply both sides of Eq. (3.5) by Π_j and take the trace to obtain

$$\sum_i c_i \mathrm{Tr}\,(\Pi_i\Pi_j) = c_j + \sum_{i\neq j} k c_i \tag{3.6}$$

for any choice of j. Thus, $c_i = 0$ for all i and the projectors Π_i must be linearly independent. Since the projections belong to a vector space of dimension d^2, this puts a bound on the size of a set of equiangular lines.

Theorem 3.1.1. A set of equiangular lines in $\mathbb{C}^d$ has at most d^2 elements.

Let L be a set of d^2 equiangular lines in $\mathbb{C}^d$. If we consider

$$I = \sum_i c_i \Pi_i, \tag{3.7}$$

taking the trace of both sides yields

$$\mathrm{Tr}\,(I) = d = \sum_i c_i. \tag{3.8}$$

Likewise, we can multiply Eq. (3.7) by Π_j throughout before taking the trace to obtain

$$\mathrm{Tr}\,(\Pi_j) = 1 = c_j + \sum_{i\neq j} k c_i. \tag{3.9}$$

Multiplying Eq. (3.8) by k and subtracting Eq. (3.9) from it gives $c_j(k-1) = dk - 1$ for all j, which then implies that

$$c_j = \frac{1}{d} \qquad \text{and} \qquad k = \frac{1}{d+1}. \tag{3.10}$$

3.1.2 Complex projective design

An alternative way to define SIC-POVMs is to consider a particular structure in combinatorial design theory called a t-design. A finite set X of points on a d-dimensional unit hypersphere S^d is called a spherical t-design if the mean value of any polynomial of degree t or less on the discrete set X equals the mean value of the polynomial over the entire sphere. In other words,

$$\frac{1}{|X|} \sum_{x \in X} f(x) = \frac{1}{\mathcal{V}} \int_{S^d} dy \; f(y) \tag{3.11}$$

where $\mathcal{V}$ denotes the measure or volume of S^d. In practice, this means that we can approximate an integral over a sphere numerically by a suitable choice of equally weighted $|X|$ points.

The application to quantum mechanics takes the average over pure quantum states represented by points in a complex projective space. Thus, a set $S = \{|\psi_i\rangle \in \mathcal{H}\}_{i=1}^n$ is called a complex projective t-design if

$$\frac{1}{|S|} \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|^{\otimes t} = \int_{\mathbb{C}\mathrm{P}^{d-1}} d\mu(\phi) \; |\phi\rangle\langle\phi|^{\otimes t} \tag{3.12}$$

where μ is the unitarily invariant Haar measure.

Let $\mathcal{H}_t = \otimes_t \mathcal{H}$ be the t-fold tensor product of Hilbert spaces $\mathcal{H}$ and consider the operator

$$S_t = \sum_{i=1}^{n} |\Psi_i^{(t)}\rangle\langle\Psi_i^{(t)}|, \qquad \left|\Psi_i^{(t)}\right\rangle = |\psi_i\rangle^{\otimes t}. \tag{3.13}$$

The set S forms a complex projective t-design if and only if

$$\mathrm{Tr}\left(S_t^2\right) = \sum_{i,j} |\langle\psi_i|\psi_j\rangle|^{2t} = \frac{n^2 t!(d-1)!}{(t+d-1)!} \tag{3.14}$$

with $n \geq {}^{t+d-1}\mathrm{C}_{d-1}$, the lower bound being the dimension of the symmetric subspace of $\mathcal{H}_t$. Thus, SIC-POVMs are called minimal 2-designs since for S corresponding to elements of a SIC-POVM,

$$\mathrm{Tr}\left(S_2^2\right) = \sum_{i,j} |\langle\psi_i|\psi_j\rangle|^4 = \frac{2d^3}{d+1} \tag{3.15}$$

where minimal refers to the fact that they achieve the lower bound on the size of S for 2-designs, $|S| \geq d^2$. All 2-designs with $|S| = d^2$ are necessarily equivalent to sets of equiangular lines and therefore correspond to SIC-POVMs.

3.1.3 Maximal equiangular tight frame

A third way to characterize SIC-POVMs is through the notion of a frame, which generalizes the notion of a basis. We know that any linearly independent set of vectors that span the entire vector space can serve as a basis for that space. In particular, the vectors in the basis do not have to be orthogonal. Generally, this leads to a biorthogonal system for basis E and F: for every vector $|e_i\rangle \in E$, there is a dual vector $|f_i\rangle \in F$ such that

$$\langle f_i|e_j\rangle = \delta_{ij}. \tag{3.16}$$

A frame extends this idea to representations that expand over a linearly dependent set.

Here we are interested in the case of frames for the space of Hermitian operators $\mathrm{H}(\mathcal{H})$ on a Hilbert space of dimension d. Defining the inner product to be

$$\langle A, B\rangle = \mathrm{Tr}\,(AB) \tag{3.17}$$

for any $A, B \in \mathrm{H}(\mathcal{H})$, we see that $\mathrm{H}(\mathcal{H})$ is itself a real Hilbert space of dimension d^2.

A frame $\mathcal{F}$ for $\mathrm{H}(\mathcal{H})$ is a set of operators $F(\lambda) \in \mathcal{F}$ that satisfy

$$a\|A\|^2 \leq \int_\Lambda d\lambda \; |\,\langle f(\lambda), A\rangle\,|^2 \leq b\|B\|^2 \tag{3.18}$$

for all $A \in \mathrm{H}(\mathcal{H})$ and some $0 < a \leq b$. Observe that for an orthogonal basis $\{B_k\}$,

$$\sum_{k=1}^{d^2} |\,\langle B_k, A\rangle\,|^2 = \|A\|^2 \tag{3.19}$$

for all $A \in \mathrm{H}(\mathcal{H})$. A frame $\mathcal{D} = \{D(\lambda)\}$ is said to be dual to $\mathcal{F}$ if

$$A = \int_\Lambda d\lambda \;\; \langle F(\lambda), A\rangle \, D(\lambda) \tag{3.20}$$

for all $A \in \mathrm{H}(\mathcal{H})$.

Define the frame operator

$$S(A) = \int_\Lambda d\lambda \, \langle F(\lambda), A\rangle \, F(\lambda), \tag{3.21}$$

which is a superoperator acting on Hermitian operators. If $S(A) = kA$ for some constant k, we say that $\mathcal{F}$ is a tight frame. From the frame operator, we can construct a dual frame $S^{-1}(\mathcal{F})$ called the canonical dual frame of $\mathcal{F}$, which is defined by

$$A = S^{-1}[S(A)] = \int_\Lambda d\lambda \, \langle F(\lambda), A\rangle \, S^{-1}[F(\lambda)]. \tag{3.22}$$

This means that canonical dual frame operators consists of operators proportional to the frame operators in $\mathcal{F}$.

A SIC-POVM forms a maximally equiangular tight frame with frame operators

$$F_i = \frac{1}{d}|\psi_i\rangle\langle\psi_i| \tag{3.23}$$

and it has a unique canonical dual frame given by operators

$$D_i = S^{-1}(F_i) \equiv d(d+1)F_i - I. \tag{3.24}$$

It is straightforward to check that $\langle D_i, F_j\rangle = \delta_{ij}$. The equiangular condition means that the frame operators satisfy

$$\langle F_i, F_j\rangle = \frac{d\delta_{ij} + 1}{d^2(d+1)}. \tag{3.25}$$

Chris Ferrie and Joseph Emerson [37] note that a SIC-POVM yields a frame representation for finite-dimensional Hilbert spaces where states are represented by the frame operators via the density operator reconstruction formula

$$\rho = \sum_i \langle F_i, \rho\rangle \, D_i = \sum_i \langle F_i, \rho\rangle \, [d(d+1)F_i - 1] \tag{3.26}$$

where $p(i) = \langle F_i, \rho\rangle$ is a true probability and measurements are represented by dual frame operators via the total probability law

$$\Pr(k) = \sum_j \langle F_j, \rho\rangle \, \langle D_j, E_k\rangle \tag{3.27}$$

where we think of $\langle D_j, E_k\rangle$ as something like a conditional probability. It is however not a true conditional probability since it is sometimes negative.

3.2 Weyl-Heisenberg SICs and the Clifford group

The conventional way to make a SIC-POVM is to find a set of d^2 vectors $|\psi_i\rangle \in \mathcal{H}^d$ such that

$$|\langle\psi_i|\psi_j\rangle|^2 = \frac{d\delta_{ij}+1}{d+1} \tag{3.28}$$

in which case the SIC-POVM elements are obtained from rescaling the rank-1 projectors

$$\Pi_i = |\psi_i\rangle\langle\psi_i|, \qquad i = 1, 2, \ldots, d^2. \tag{3.29}$$

It is generally more convenient to think of a SIC-POVM in terms of these rank-1 projectors because $\mathrm{Tr}\,(\Pi_i) = 1$. Thus, we distinguish the set $\{\Pi_i\}_{i=1}^{d^2}$ from a SIC-POVM by calling it a SIC set, or just SIC for short.

All SICs constructed to date have a group covariance property. Let G be a finite group with d^2 elements and $g \mapsto G$ be a projective representation of G on the Hilbert space $\mathcal{H}^d$, i.e., we consider an injective map $g \mapsto U_g$ where U_g is a d-dimensional unitary operator such that for all $g, h \in G$,

$$U_g U_h = e^{i\phi_{gh}} U_{gh}, \tag{3.30}$$

where $e^{i\phi_{gh}}$ is just some overall phase.

Take some $|\psi\rangle \in \mathcal{H}^d$. If the set of vectors $U_g\,|\psi\rangle$ generates a SIC, we say that the corresponding SIC is covariant with respect to the group G. In almost all known examples, there is just a single group used to construct SICs and is called the Weyl-Heisenberg group.

To describe the group, let $\{|j\rangle\}_{j=1}^{d}$ be an orthonormal basis for $\mathbb{C}^d$. Introduce a pair of operators called the shift operator X and phase operator Z, defined as follows:

$$\begin{aligned} X\,|j\rangle &= |j \oplus 1\rangle\,, \\ Z\,|j\rangle &= \omega^j\,|j\rangle\,, \end{aligned} \tag{3.31}$$

where $\oplus$ indicates addition modulo d and $\omega = e^{\frac{2\pi i}{d}}$. Now consider $\mathbf{p} = (p_1, p_2) \in \mathbb{Z}_d^2$, that is, $p_1, p_2 = 0, 1, ..., d-1$. It may be worth mentioning here that $\mathbf{p}$ corresponds to a point on the discrete phase space of a finite-dimensional quantum system as described by William Wootters [131]. The vectors $\mathbf{p}$ allow us to define the displacement operator

$$D_{\mathbf{p}} = \tau^{p_1 p_2} X^{p_1} Z^{p_2} \tag{3.32}$$

where $\tau = -e^{i\frac{\pi}{d}}$. If we can find a suitable seed vector $|\psi\rangle$ such that $|\langle\psi|D_{\mathbf{p}}\,|\psi\rangle\,|^2 = \frac{1}{d+1}$ for $\vec{p} \neq (0,0)$, then the projectors

$$\Pi_{\mathbf{p}} = D_{\mathbf{p}}|\psi\rangle\langle\psi|D_{\mathbf{p}}^{\dagger} \tag{3.33}$$

form a SIC of dimension d. The seed vector $|\psi\rangle$ is known as a *fiducial vector*.

Two key properties of displacement operators are:

$$\begin{aligned} D_{\mathbf{p}}^{\dagger} &= D_{-\mathbf{p}}, \\ D_{\mathbf{p}} D_{\mathbf{q}} &= \tau^{\langle \mathbf{p},\mathbf{q} \rangle} D_{\mathbf{p}+\mathbf{q}}, \end{aligned} \tag{3.34}$$

where $\langle \mathbf{p}, \mathbf{q} \rangle = p_2 q_1 - p_1 q_2$ is called the symplectic form.

The set of operators

$$W(d) = \left\{ \tau^r D_{\mathbf{p}} \mid r = 0, 1, ..., d-1, \mathbf{p} \in \mathbb{Z}_d^2 \right\} \tag{3.35}$$

is called the Weyl-Heisenberg group. Observe that while $W(d)$ contains d^3 elements, when we act its elements on a SIC fiducial, we only get d^2 unique projectors, since vectors that differ by an overall phase factor correspond to the same projection. The SIC generated by $W(d)$ is called a Weyl-Heisenberg SIC.

We are also interested in the normalizer of the Weyl-Heisenberg group, called the Clifford group $C(d)$ [3, 5, 53–55, 108], which is a subgroup of the set of unitary operators in dimension d. If $U \in C(d)$ then

$$U W(d) U^{\dagger} = W(d). \tag{3.36}$$

If the set of anti-unitary operators that map $W(d)$ to itself are added to the unitaries, we get a set called the extended Clifford group.

Marcus Appleby introduced a convenient way to parameterize the Clifford unitaries in terms of symplectic matrices [3]. If we just consider the simpler case when d is odd, let $\mathrm{SL}(2, \mathbb{Z}_d)$ be the group of 2×2 symplectic matrices with elements

$$F = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{3.37}$$

such that $\alpha, \beta, \gamma, \delta \in \mathbb{Z}_d$ and $\det(F) = 1 \mod d$. If the multiplicative inverse β^{-1} of β exists in $\mathbb{Z}_d$ then we can construct the Clifford unitary

$$U_F = \frac{1}{\sqrt{d}} \sum_{r,s=0}^{d-1} \tau^{\beta^{-1}\left(r^2\delta - 2rs + s^2\alpha\right)} |r\rangle \langle s|. \tag{3.38}$$

such that

$$U_F D_p U_F^{\dagger} = D_{Fp}. \tag{3.39}$$

If β^{-1} does not exist, one can look for some integer x such that $\alpha + x\beta$ is nonzero and $\mathrm{GCD}(\delta + x\beta, d) = 1$, then $F = F_1 F_2$ where [3]

$$F_1 = \begin{pmatrix} 0 & -1 \\ 1 & x \end{pmatrix}, \qquad F_2 = \begin{pmatrix} \gamma + x\alpha & \delta + x\beta \\ -\alpha & -\beta \end{pmatrix} \tag{3.40}$$

and U_{F_1} and U_{F_2} can be constructed as before. U_F and $U_{F_1} U_{F_2}$ give the same Clifford unitary up to a phase factor.

A particularly important symplectic matrix for finding SICs is given by

$$F_Z = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \tag{3.41}$$

and its corresponding Clifford unitary U_{F_Z}. We shall call F_Z the Zauner matrix. Note that the Zauner matrix can be defined in any finite dimension. According to a conjecture by Gerhard Zauner [134] and Marcus Appleby [3], there exists a SIC fiducial vector in every dimension which is an eigenvector of U_{F_Z}, a unproven claim that is supported by all available numerical results to date. In Sec. 3.5, we will see that the Zauner matrix in $d = 3$ plays a prominent role in the linear dependencies of qutrit SICs.

3.3 Classifying qutrit SICs into inequivalent families

Since we are ultimately interested in the geometric features of quantum state space in dimension three, we begin by examining the properties of qutrit SICs. All known SICs in $d = 3$ are Weyl-Heisenberg SICs, of which there are infinitely many.

Using the displacement operators in Eq. (3.32), a fiducial vector $|\psi\rangle$ generates a SIC with elements

$$\Pi_{\mathbf{p}} = D_{\mathbf{p}} |\psi\rangle\langle\psi| D_{\mathbf{p}}^{\dagger}. \tag{3.42}$$

where the index $\mathbf{p} = (a, b)$ specifies the powers for the X and Z operators, respectively. In this case, it is useful to adopt the following convention for labeling the SIC elements:

$$\begin{array}{lll} 1 : |\psi\rangle & 2 : Z|\psi\rangle & 3 : Z^2|\psi\rangle \\ 4 : X|\psi\rangle & 5 : XZ|\psi\rangle & 6 : XZ^2|\psi\rangle \\ 7 : X^2|\psi\rangle & 8 : X^2Z|\psi\rangle & 9 : X^2Z^2|\psi\rangle \end{array} \tag{3.43}$$

where the SIC elements are given by projectors associated with each unit vector. For example, $\Pi_5 = XZ|\psi\rangle\langle\psi|Z^\dagger X^\dagger$.

Consider the continuous 1-parameter family of qutrit SIC fiducial vectors given by

$$|\psi_t\rangle = \begin{pmatrix} 0 \\ 1 \\ e^{2it} \end{pmatrix}. \tag{3.44}$$

Marcus Appleby [3] has shown that there are 3 types of orbits of the extended Clifford group in $d = 3$ for which $|\psi(t)\rangle$ is in the orbit: a generic type corresponding to $t \in (0, \frac{\pi}{6})$ and two exceptional ones corresponding to the endpoints $t = 0, \frac{\pi}{6}$. Each orbit of the Clifford group determines a set of fiducial vectors from which a number of distinct SICs can be identified. For example, in the generic case, the orbit consists of 72 fiducials, forming 8 distinct SICs for any fixed choice of $0 < t < \frac{\pi}{6}$.

An important property of SICs on different orbits of the Clifford group is that generally, they are not related to one another by a unitary transformation. This makes it convenient for us to classify qutrit SICs into 8 inequivalent SIC-families represented by the following fiducial vectors:

$$\left|\psi_t^{(0\pm)}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ e^{\mp it} \\ -e^{\pm it} \end{pmatrix}, \qquad \left|\psi_t^{(1\pm)}\right\rangle = \sqrt{\frac{2}{3}} \begin{pmatrix} \omega \sin t \\ \sin\left(t \pm \frac{2\pi}{3}\right) \\ \sin\left(t \mp \frac{2\pi}{3}\right) \end{pmatrix},$$

$$\left|\psi_t^{(2\pm)}\right\rangle = \sqrt{\frac{2}{3}} \begin{pmatrix} \omega^2 \sin t \\ \sin\left(t \pm \frac{2\pi}{3}\right) \\ \sin\left(t \mp \frac{2\pi}{3}\right) \end{pmatrix}, \qquad \left|\psi_t^{(3\pm)}\right\rangle = \sqrt{\frac{2}{3}} \begin{pmatrix} \sin t \\ \sin\left(t \pm \frac{2\pi}{3}\right) \\ \sin\left(t \mp \frac{2\pi}{3}\right) \end{pmatrix}, \tag{3.45}$$

where $\omega = e^{i\frac{2\pi}{3}}$ and $0 \leq t \leq \frac{\pi}{6}$. In this parametrization, the exceptional types correspond to $t = 0$, where all 8 fiducials correspond to the same SIC, and $t = \frac{\pi}{6}$, where there are 4 distinct SICs.

The unique SIC for $t = 0$ is generated by the SIC fiducial

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{3.46}$$

We will see that this particular SIC possesses nice properties that make it the simplest one to consider when studying the probabilities obtained from qutrit SICs. In anticipation of its important role for the remainder of this thesis, we refer to it as the *most exceptional SIC*.

The only other exceptional case for which we have less than 8 SICs is when $t = \frac{\pi}{6}$, where there are 4 distinct SICs, whose fiducial vectors may be chosen as

$$\left|\psi^{(0)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix}, \qquad \left|\psi^{(1)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}\omega\\1\\-2\end{pmatrix},$$
$$\left|\psi^{(2)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}\omega^2\\1\\-2\end{pmatrix}, \qquad \left|\psi^{(3)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}. \tag{3.47}$$

It is worth mentioning here that the SIC-families of Eq. (3.45) are inequivalent with respect to Clifford unitaries; however, some of the SIC-families for different values of the parameter t are still related by a unitary operator that does not belong to the Clifford group. Specifically, Huangjun Zhu [135] has shown that the SIC-families for t, $\frac{\pi}{9} - t$ and $\frac{\pi}{9} + t$ are, in fact, unitarily equivalent to each other, with the unitary transformation relating them being

$$U = \mathrm{diag}(1, s, s^2), \quad s = e^{-i\frac{2\pi}{9}}. \tag{3.48}$$

Moreover, there are no other unitary equivalences. This means that every pair of SICs on any two different orbits corresponding to $t \in \left[0, \frac{\pi}{18}\right]$ are not equivalent.

3.4 Geometric invariants of qutrit SICs

Every SIC is fully characterized by two sets of geometrically invariant quantities: the first set consist of numbers T_{ijk} that give the trace of the product of every 3 SIC elements and are called *triple products*, i.e.,

$$T_{ijk} = \mathrm{Tr}\left(\Pi_i \Pi_j \Pi_k\right). \tag{3.49}$$

The other set is composed of the expansion coefficients S_{ijk} of operator multiplication between SIC elements when expressed in the basis defined by SIC projectors and are called

structure coefficients, i.e.,

$$\Pi_i \Pi_j = \sum_{k=1}^{d^2} S_{ijk} \Pi_k. \tag{3.50}$$

Actually, triple products and structure coefficients are essentially equivalent quantities since they are directly related to each other:

$$S_{ijk} = \frac{1}{d}\left[(d+1)T_{ijk} - \frac{d\delta_{ij}+1}{d+1}\right]. \tag{3.51}$$

Triple products are significant because they demonstrate the unitary equivalence of SICs: It is shown in Ref. [7] that two SICs are unitarily equivalent if and only if the triple products are the same, up to permutation. In Chapter 5, we will show that it is the real part of T_{ijk} that plays an important role in characterizing geometric properties of SIC probability spaces. Thus, we write

$$\tilde{T}_{ijk} = \mathrm{Re}\left[T_{ijk}\right] \tag{3.52}$$

to denote the real part of Eq. (3.49). Similarly, we write

$$\tilde{S}_{ijk} = \mathrm{Re}\left[S_{ijk}\right] \tag{3.53}$$

to denote the real part of the structure coefficients.

Substituting Eq. (3.42) into Eq. (3.49), the possible values for T_{ijk} can be obtained from

$$\mathrm{Tr}\left(\Pi_{\mathbf{p}}\Pi_{\mathbf{q}}\Pi_{\mathbf{r}}\right) = \langle Z^{-r_2}X^{p_1-r_1}Z^{p_2}\rangle\langle Z^{-p_2}X^{q_1-p_1}Z^{q_2}\rangle\langle Z^{-q_2}X^{r_1-q_1}Z^{r_2}\rangle, \tag{3.54}$$

where $\langle A\rangle = \langle\psi|A\,|\psi\rangle$ is the expected value of A given by the SIC fiducial $|\psi\rangle$ and $\mathbf{p} = (p_1, p_2)$, $\mathbf{q} = (q_1, q_2)$, and $\mathbf{r} = (r_1, r_2)$ corresponds to some index triple (ijk) if we consider Eq. (3.43). It is fairly straightforward to try out all combinations of $\mathbf{p}, \mathbf{q}$, and $\mathbf{r}$ consistent with Eq. (3.54) and consequently obtain all distinct values of T_{ijk}, which are

$$T_{ijk} \in \left\{1, \frac{1}{4}, -\frac{1}{8}\omega^r, -\frac{1}{8}\omega^r e^{\pm 6it}\right\}, \qquad \text{for} \quad r = 0, 1, 2. \tag{3.55}$$

Using Eq. (3.51) we can also calculate the structure coefficients associated with every triple product. Thus, we can use Eq. (3.55) to determine the real parts $\tilde{T}_{ijk}$, which ultimately are the values most relevant to determining properties of quantum state space.

Of particular interest in Eq. (3.55) are the 4 possible values of $\tilde{T}_{ijk}$ we get when i, j, k are all distinct. In these cases, the corresponding $\tilde{S}_{ijk}$ are given by

$$\tilde{S}_{ijk} = \frac{1}{3}\left[4\tilde{T}_{ijk} - \frac{1}{4}\right], \qquad (i \neq j \neq k). \tag{3.56}$$

One possibility corresponds to

$$\tilde{T}_{ijk} = -\frac{1}{8}\mathrm{Re}\left[\omega^r\right] = \frac{1}{16} \tag{3.57}$$

which implies that $\tilde{S}_{ijk} = 0$. The remaining 3 possibilities yield $\tilde{S}_{ijk} = x_t, y_t, z_t$ where

$$\begin{aligned} x_t &= -\frac{1}{6}\left(\cos 6t + \frac{1}{2}\right), \\ y_t &= -\frac{1}{6}\left[\cos\left(6t - \frac{2\pi}{3}\right) + \frac{1}{2}\right] \equiv x_{\frac{\pi}{9}+t}, \\ z_t &= -\frac{1}{6}\left[\cos\left(6t + \frac{2\pi}{3}\right) + \frac{1}{2}\right] \equiv x_{\frac{\pi}{9}-t}. \end{aligned} \tag{3.58}$$

For the most exceptional SIC given by Eq. (3.46),

$$x_0 = -\frac{1}{4}, \quad y_0 = 0, \quad z_0 = 0, \tag{3.59}$$

in which case we have

$$\tilde{S}_{ijk}|_{t=0} \in \left\{0, 1, \frac{1}{4}, -\frac{1}{4}\right\}. \tag{3.60}$$

Table 3.1 shows the distinct values for the structure coefficients $\tilde{S}_{ijk}$ for $0 \leq t \leq \pi/6$ and their frequencies of occurrence, which add up to the total number of index triples (ijk), $d^3 = 729$.

Although triple products and structure coefficients carry the same information about the structure of SICs, we will find that the structure coefficients are typically more convenient for numerical calculations since they vanish for some index triples (ijk), whereas all triple products are nonzero. In fact, according to Table 3.1 more than half of the index triples (ijk) yield $\tilde{S}_{ijk} = 0$. Furthermore, it tells us that it is sufficient to know the real part of the structure coefficients for just one of the 8 SIC-families, since the same set of $\tilde{S}_{ijk}$ are obtained for all other fiducial vectors in Eq. (3.45) once we permute the indices. In fact, this follows from the fact that the 8 SIC-families are inequivalent in terms of the triple products T_{ijk} only because the signs of the imaginary parts are different.

Since we already know the possible values of $\tilde{S}_{ijk}$, the only thing left to do is to assign index triples (ijk) to each possible value. To illustrate how the indices are assigned,

Structure coefficient	Frequency
1	9
$\frac{1}{4}$	144
$-\frac{1}{4}$	18
$-\frac{1}{3}\left(\frac{1}{2}\cos 6t + \frac{1}{4}\right)$	54
$\frac{1}{12}\left(\cos 6t + \sqrt{3}\sin 6t - 1\right)$	54
$\frac{1}{12}\left(\cos 6t - \sqrt{3}\sin 6t - 1\right)$	54
0	396

Table 3.1: Distinct values for $\tilde{S}_{ijk}$ given parameter t.

consider the structure coefficients for the $\left|\psi_t^{(0+)}\right\rangle$ in Eq. (3.45). We remind ourselves that the possible values for $\tilde{S}_{ijk}$ are given by

$$\tilde{S}_{ijk} \in \left\{0, 1, \pm\frac{1}{4}, x_t, y_t, z_t\right\}, \tag{3.61}$$

with x_t, y_t, z_t defined by Eq. (3.58). Which index triples they correspond to is specified by the following rules:

(a) For $\tilde{S}_{ijk} = 1$, all indices have to be the same, that is $i = j = k$. Obviously, there are nine ways to achieve this.

(b) For $\tilde{S}_{ijk} = \frac{1}{4}$, two indices must be the same. In this case, it is convenient to think of writing $\tilde{S}_{ijk}$ as the elements of 9 matrices $\tilde{S}^{(k)}_{ij} = \tilde{S}_{ijk}$, for which we can say that we get $\tilde{S}_{ijk} = \frac{1}{4}$ whenever $i = k$ or $j = k$ but not on the diagonal of $S^{(k)}$, i.e, $i \neq j$. Thus, for each matrix $\tilde{S}^{(k)}$ there are 16 such entries, corresponding to the elements in the kth row and kth column excluding the diagonal. It follows that there are a total of $16 \times 9 = 144$ index triples (ijk) with this value.

(c) For $\tilde{S}_{ijk} = -\frac{1}{4}$, the indices i, j, k have to be all different. There is a simple pattern for determining which index triples (ijk) give this value which can be easily stated in terms of a table of indices called an *index generator*, which are given in Table 3.2. In

terms of the index generator, $\tilde{S}_{ijk} = -\frac{1}{4}$ when the index triple (ijk) is on the same row. For instance, for the SIC $\left|\psi_t^{(2-)}\right\rangle$, the relevant index triples (ijk) belong to the set

$$\{(186), (294), 375)\}, \tag{3.62}$$

as well as all permutations of indices in each index triple, i.e., $(168), (681)$, and so on.

(d) For $\tilde{S}_{ijk} = x_t$, we take the index triples that lie on the same column in Table 3.2 or those that lie on entirely different rows and columns. That is, the desired index triples(ijk) belong to the set

$$\{(123), (897), (645), (195), (843), (627), (147), (825), (693)\} \tag{3.63}$$

and all permutations of indices in each index triple.

(e) For $\tilde{S}_{ijk} = y_t$, we take the index triples such that first 2 indices are on the same column and the last one belongs to the succeeding column when counting in a cyclic manner. By succeeding we mean that column 2 is after column 1, column 3 is after column 2, and column 1 is after column 3. Thus, the relevant (ijk) are in the set

$$\{(127), (895), (643), (238), (976), (451), (319), (784), (562)\} \tag{3.64}$$

and all permutations of indices in each index triple.

(f) For $\tilde{S}_{ijk} = z_t$, we have a similar rule as in y_t above but we take the last index from the preceding row. This yields the set of index triples

$$\{(125), (893), (647), (236), (971), (458), (314), (782), (569)\} \tag{3.65}$$

and all permutations of indices in each index triple.

(g) It is straightforward to check that the above rules account for the 54 index triples for $\tilde{S}_{ijk} = x_t, y_t, z_t$, respectively, as listed in Table 3.1. Any other index triple not specified in the above rules has $\tilde{S}_{ijk} = 0$.

Observe that once the structure coefficients for, say the SIC $\left|\psi_t^{(0+)}\right\rangle$ are known, $\tilde{S}_{ijk}$ for any other qutrit SIC with the same value of parameter t are also completely determined since all we need to do is relabel the indices according to the permutation implied by the

$$G_{0+} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad G_{0-} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 5 \\ 7 & 9 & 8 \end{bmatrix}$$

$$G_{1+} = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 7 \\ 3 & 4 & 8 \end{bmatrix} \quad G_{1-} = \begin{bmatrix} 1 & 9 & 5 \\ 2 & 7 & 6 \\ 3 & 8 & 4 \end{bmatrix}$$

$$G_{2+} = \begin{bmatrix} 1 & 6 & 8 \\ 2 & 4 & 9 \\ 3 & 5 & 7 \end{bmatrix} \quad G_{2-} = \begin{bmatrix} 1 & 8 & 6 \\ 2 & 9 & 4 \\ 3 & 7 & 5 \end{bmatrix}$$

$$G_{3+} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \quad G_{3-} = \begin{bmatrix} 1 & 7 & 4 \\ 2 & 8 & 5 \\ 3 & 9 & 6 \end{bmatrix}$$

Table 3.2: Structure coefficient index generators for qutrit SICs. The index triples (ijk) for each value of $\tilde{S}_{ijk}$ are obtained from rules that make use of these index generators.

index generators. For example, comparing G_{0+} to G_{2-}, after doing Eq. (3.70), we apply the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 8 & 6 & 2 & 9 & 4 & 3 & 7 & 5 \end{pmatrix} \tag{3.66}$$

on the indices to obtain the corresponding J_{ijk} for the SIC $\left|\psi_t^{(2-)}\right\rangle$. It might also be worth noting here that the permutation of indices is such that it always preserves the affine lines of G_{0+}, that is, the rows, columns, and broken diagonals of G_{0+} are mapped onto rows, columns, or broken diagonals for other index generators.

We can supplement the rules above with ones for getting the imaginary parts of S_{ijk}, which is a bit more involved. Consider the matrix

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu & -\mu & -\mu & \mu & \mu & \mu \\ 0 & 0 & 0 & \mu & \mu & \mu & -\mu & -\mu & -\mu \\ 0 & \mu & -\mu & 0 & \mu & -\mu & x'_t & y'_t & z'_t \\ 0 & \mu & -\mu & -\mu & 0 & \mu & y'_t & z'_t & x'_t \\ 0 & \mu & -\mu & \mu & -\mu & 0 & z'_t & x'_t & y'_t \\ 0 & -\mu & \mu & -x'_t & -y'_t & -z'_t & 0 & -\mu & \mu \\ 0 & -\mu & \mu & -y'_t & -z'_t & -x'_t & \mu & 0 & -\mu \\ 0 & -\mu & \mu & -z'_t & -x'_t & -y'_t & -\mu & \mu & 0 \end{pmatrix}, \tag{3.67}$$

where the components $J_{1jk} = \mathrm{Im}[S_{1jk}]$ are the imaginary parts for the structure coefficients of the SIC-family $\left|\psi_t^{(0+)}\right\rangle$, $\mu = \frac{\sqrt{3}}{12}$, and

$$\begin{aligned} x'_t &= \frac{1}{6}\sin 6t, \\ y'_t &= \frac{1}{6}\sin\left(6t + \frac{2\pi}{3}\right), \\ z'_t &= \frac{1}{6}\sin\left(6t - \frac{2\pi}{3}\right), \end{aligned} \tag{3.68}$$

are the corresponding imaginary parts of Eq. (3.58). For instance, we have $S_{147} = x_t + ix'_t$.

The rules for determining (ijk) are as follows:

(i) First, the index triples for $\pm\mu$ for the SIC-family $\left|\psi_t^{(0+)}\right\rangle$ are specified by the components of J_1 and the permutation matrices

$$P_{1j} = \begin{pmatrix} V_j & 0 & 0 \\ 0 & V_j & 0 \\ 0 & 0 & V_j \end{pmatrix}, \quad P_{2j} = \begin{pmatrix} 0 & 0 & V_j \\ V_j & 0 & 0 \\ 0 & V_j & 0 \end{pmatrix}, \quad P_{3j} = \begin{pmatrix} 0 & V_j & 0 \\ 0 & 0 & V_j \\ V_j & 0 & 0 \end{pmatrix}, \tag{3.69}$$

where $V_1 = I$, $V_2 = X$ and $V_3 = X^2$. The permutations are used to generate all other matrices J_i:

$$\begin{aligned} J_2 &= P_{12} J_1 P_{12}^T, & J_6 &= P_{23} J_1 P_{23}^T, \\ J_3 &= P_{13} J_1 P_{13}^T, & J_7 &= P_{31} J_1 P_{31}^T, \\ J_4 &= P_{21} J_1 P_{21}^T, & J_8 &= P_{32} J_1 P_{32}^T, \\ J_5 &= P_{22} J_1 P_{22}^T, & J_9 &= P_{33} J_1 P_{33}^T. \end{aligned}$$

(ii) Next, to obtain the set of J_i matrices for the other SIC-families $\left|\psi_t^{(r+)}\right\rangle$ with $r = 1, 2, 3$ and $\left|\psi_t^{(0-)}\right\rangle$, first we perform the replacements

$$\begin{aligned} \mu &\mapsto -\mu, & x'_t &\mapsto -x'_t, \\ y'_t &\mapsto -z'_t, & z'_t &\mapsto -y'_t. \end{aligned} \tag{3.70}$$

then we permute the indices of J_{ijk} for $\left|\psi_t^{(0+)}\right\rangle$ according to the index generators in Table 3.2. For example, comparing G_{0+} to G_{2+}, after doing Eq. (3.70), we apply the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 6 & 8 & 2 & 4 & 9 & 3 & 5 & 7 \end{pmatrix}$ on the indices to obtain the corresponding J_{ijk} for the SIC $\left|\psi_t^{(2+)}\right\rangle$.

(iii) Finally, for the SIC-families $\left|\psi_t^{(r-)}\right\rangle$ with $r = 1, 2, 3$, we simply permute the indices of J_{ijk} for $\left|\psi_t^{(0+)}\right\rangle$ according to the index generators.

3.5 Linear dependency structure of qutrit SICs

We have already seen that by thinking in terms of a set of equiangular lines in a complex linear space, a SIC can always be described with a set of vectors

$$S = \left\{ |\psi_j\rangle \in \mathcal{H}^d \,\middle|\, j = 1, 2, \ldots, d^2 \right\} \tag{3.71}$$

that define the 1-dimensional subspaces on which SIC elements project. Because the number of vectors in such a set is much larger than the number of dimensions, they are certainly linearly dependent. However, we may ask if choosing any d vectors from the set S spans the entire Hilbert space. That is, does every subset of d vectors from S form a linearly independent set and a basis for $\mathcal{H}^d$? It turns out that the answer to this question is negative for $d = 3$. In fact, according to Hoan Dang, et al., such linearly dependent subsets can always be found for SICs in dimensions divisible by 3 [30]. In this section, we describe such a linear dependency structure for qutrit SICs.

We begin with some definitions:

Definition 3.5.1. A p-vector is any 2-element vector

$$\mathbf{p} = \begin{pmatrix} a \\ b \end{pmatrix} \tag{3.72}$$

where $a, b = 0, 1, \ldots, d-1$ so that each $\mathbf{p}$ corresponds to a displacement operator $D_{\mathbf{p}}$.

Definition 3.5.2. A p-set is any set of d p-vectors

$$\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_d\}. \tag{3.73}$$

If $\mathbf{p}_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$ for $i = 1, 2, ..., d$, we will write the corresponding p-set in shorthand notation as

$$(a_1 b_1, a_2 b_2, ..., a_d b_d). \tag{3.74}$$

Definition 3.5.3. In dimension 3, the set of three p-vectors $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ is called a good p-set whenever the corresponding set of vectors

$$\{D_{\mathbf{p}_1} |\psi\rangle, D_{\mathbf{p}_2} |\psi\rangle, D_{\mathbf{p}_3} |\psi\rangle\} \tag{3.75}$$

form a linearly dependent set for some fiducial vector $|\psi\rangle$.

For the 8 SIC-families listed by Eq. (3.45), we obtain three good p-sets for each SIC-family if we take the following p-sets:

$$\begin{aligned}
\left|\psi_t^{(0\pm)}\right\rangle &: (00,01,02),(10,11,12),(20,21,22),\\
\left|\psi_t^{(1\pm)}\right\rangle &: (00,12,21),(01,10,22),(02,11,20),\\
\left|\psi_t^{(2\pm)}\right\rangle &: (00,11,22),(01,12,20),(02,10,21),\\
\left|\psi_t^{(3\pm)}\right\rangle &: (00,10,20),(01,11,21),(02,12,22).
\end{aligned} \tag{3.76}$$

where for instance, each p-set for $\left|\psi_t^{(0\pm)}\right\rangle$ corresponds to the rows of

$$G_{0+} = \begin{bmatrix} 00 & 01 & 02 \\ 10 & 11 & 12 \\ 20 & 21 & 22 \end{bmatrix} \tag{3.77}$$

which can be identified with the index generator G_{0+} in Table 3.2 if we apply our labeling conventions in Eq. (3.43). In fact, one can check from the rules for specifying the index triples (ijk) of the structure coefficients described in Sec. 3.4 that the good p-sets of Eq. (3.76) for each SIC-family correspond to index triples (ijk) such that $\tilde{S}_{ijk} = -\frac{1}{4}$, that is, the rows of the index generators in Table 3.2.

Something more interesting happens when we consider a few special cases, specifically the linear dependencies for the most exceptional SIC given by $t = 0$ and the 8 SICs corresponding to $t = \frac{\pi}{9}$ (which are equivalent to each other and related by the unitary matrix given by Eq. (3.48).) Again, in terms of p-vectors, we obtain linearly dependent sets of SIC-vectors with the fiducials in Eq. (3.45) for the following 12 good p-sets:

$$\begin{array}{llll}
(00,01,02), & (00,11,22), & (00,10,20), & (00,12,21),\\
(10,11,12), & (01,12,20), & (01,11,21), & (01,10,22),\\
(20,21,22), & (02,10,21), & (02,12,22), & (02,11,20),
\end{array} \tag{3.78}$$

which coincidentally is the union of all the p-sets in Eq. (3.76).

The linear dependency structure of the most exceptional SIC is described in detail in Table 3.3. The linear dependency condition specifies how the SIC-vectors corresponding to the p-sets sum to zero and the orthogonal subspace is specified by a unit vector that spans the 1-dimensional space. It is worth noting that every 3 consecutive orthogonal subspace

Linear Dependency	Orthogonal Subspace
$\|1\rangle + \|4\rangle + \|7\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$
$\|2\rangle + \|5\rangle + \|8\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega^2\\\omega\end{pmatrix}$
$\|3\rangle + \|6\rangle + \|9\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega\\\omega^2\end{pmatrix}$
$\|1\rangle + \|2\rangle + \|3\rangle = 0$	$\begin{pmatrix}1\\0\\0\end{pmatrix}$
$\|4\rangle + \omega^2\|5\rangle + \omega\|6\rangle = 0$	$\begin{pmatrix}0\\0\\1\end{pmatrix}$
$\|7\rangle + \omega\|8\rangle + \omega^2\|9\rangle = 0$	$\begin{pmatrix}0\\1\\0\end{pmatrix}$
$\omega^2\|1\rangle + \|5\rangle + \|9\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}\omega^2\\1\\1\end{pmatrix}$
$\omega^2\|2\rangle + \|6\rangle + \|7\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\\omega^2\end{pmatrix}$
$\omega^2\|3\rangle + \|4\rangle + \|8\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega^2\\1\end{pmatrix}$
$\omega\|1\rangle + \|6\rangle + \|8\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}\omega\\1\\1\end{pmatrix}$
$\omega\|2\rangle + \|4\rangle + \|9\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\\omega\\1\end{pmatrix}$
$\omega\|3\rangle + \|5\rangle + \|7\rangle = 0$	$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\\omega\end{pmatrix}$

Table 3.3: Linear dependency structures for the most exceptional SIC in $d = 3$. The vectors $|j\rangle$ correspond to SIC vectors and are labeled according to Eq. (3.43).

vector as listed in Table 3.3 (1st to 3rd, 4th to 6th, etc.) form a basis for the 3-dimensional Hilbert space. Moreover, the four bases constitute a complete set of mutually unbiased bases (MUBs) in $d = 3$.

Aside from the most exceptional SIC, there are 8 other SICs which also have 12 good p-sets, those given by $t = \frac{\pi}{9}$. The good p-sets for these SICs are listed in Table 3.4. A key difference between good p-sets of the most exceptional SIC and that of the 8 SICs with $t = \frac{\pi}{9}$ is that for the most exceptional SIC, every good p-set has the property

$$\sum_{i=1}^{3} \mathbf{p}_i = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \mod 3. \tag{3.79}$$

This, however, is not the case for the 8 SICs with $t = \frac{\pi}{9}$ since a quick glance at Table 3.4 will show that only p-sets in the first column have this property.

Lane Hughston [65] showed that the linear dependency structure for the 9 SICs with 12 good p-sets corresponds in projective geometry to structure known as a *Hesse configuration*. A Hesse configuration consists of a set of 12 lines and 9 points in a complex projective plane where the 9 points are the inflection points of a non-singular cubic function

$$P(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3 + t x_1 x_2 x_3, \qquad t \in \mathbb{C} \tag{3.80}$$

such that every 3 inflection points fall on a line on the projective plane [64]. The inflection points are given by points on the curve for which the determinant of the Hessian matrix of P vanishes, that is, if we take the matrix H with elements

$$H_{ij} = \frac{\partial^2 P}{\partial x_i \partial x_j} \tag{3.81}$$

then we have

$$\det(H) = (2t^3 + 216) x_1 x_2 x_3 - 6t^2 (x_1^3 + x_2^3 + x_3^3) = 0. \tag{3.82}$$

The inflection points coincide for all values of t and correspond to the most exceptional SIC vectors. Moreover, the family of cubic polynomials defined by $P(x_1, x_2, x_3)$ has 4 singular members corresponding to $t = 0$ and the cube roots of $t^3 + 81 = 0$, the resulting projective lines form triangles whose vertices are given by the orthogonal subspace vectors in Table 3.3 [15]. Because of this, we refer to the most exceptional SIC and the 8 SICs with $t = \frac{\pi}{9}$ as *Hesse SICs*.

For $\left\vert\psi_{\frac{\pi}{9}}^{(0+)}\right\rangle$			
(00,01,02)	(00,10,22)	(00,11,21)	(00,12,20)
(10,11,12)	(01,11,20)	(01,12,22)	(01,10,21)
(20,21,22)	(02,12,21)	(02,10,20)	(02,11,22)

For $\left\vert\psi_{\frac{\pi}{9}}^{(2+)}\right\rangle$			
(00,12,21)	(00,10,11)	(00,02,22)	(00,01,20)
(01,10,22)	(12,22,20)	(12,01,11)	(12,10,02)
(02,11,20)	(21,01,02)	(21,10,20)	(21,21,11)

For $\left\vert\psi_{\frac{\pi}{9}}^{(0-)}\right\rangle$			
(00,01,02)	(00,12,22)	(00,10,21)	(00,11,20)
(10,11,12)	(01,10,20)	(01,11,22)	(01,12,21)
(20,21,22)	(02,11,21)	(02,12,20)	(02,10,22)

For $\left\vert\psi_{\frac{\pi}{9}}^{(2-)}\right\rangle$			
(00,12,21)	(00,22,20)	(00,10,02)	(00,01,11)
(01,10,22)	(12,01,02)	(12,22,11)	(12,10,20)
(02,11,20)	(21,10,11)	(21,01,20)	(21,22,02)

For $\left\vert\psi_{\frac{\pi}{9}}^{(1+)}\right\rangle$			
(00,11,22)	(00,12,10)	(00,20,02)	(00,01,21)
(01,12,20)	(11,20,21)	(11,01,10)	(11,12,02)
(02,10,21)	(22,01,02)	(22,12,21)	(22,20,10)

For $\left\vert\psi_{\frac{\pi}{9}}^{(3+)}\right\rangle$			
(00,10,20)	(00,11,12)	(00,21,02)	(00,01,22)
(01,11,21)	(10,21,22)	(10,01,12)	(10,11,02)
(02,12,22)	(20,01,02)	(20,11,22)	(20,21,12)

For $\left\vert\psi_{\frac{\pi}{9}}^{(1-)}\right\rangle$			
(00,11,22)	(00,20,21)	(00,12,02)	(00,01,10)
(01,12,20)	(11,01,02)	(11,20,10)	(11,12,21)
(02,10,21)	(22,12,10)	(22,01,21)	(22,20,02)

For $\left\vert\psi_{\frac{\pi}{9}}^{(3-)}\right\rangle$			
(00,10,20)	(00,21,22)	(00,11,02)	(00,01,12)
(01,11,21)	(10,01,02)	(10,21,12)	(10,11,22)
(02,12,22)	(20,11,12)	(20,01,22)	(20,21,02)

Table 3.4: Good p-sets yielding linear dependencies for the 8 SICs with $t = \frac{\pi}{9}$.

3.6 The Zauner subspace for qutrit SICs

Recall that in Sec. 3.2, we mentioned that in the Clifford group for displacement operators, there is a symplectic unitary U_{F_Z} given by Eq. (3.38) that is generated from the Zauner matrix F_Z defined by Eq. (3.41). The significance of U_{F_Z} is that according to a conjecture by Gerhard Zauner, one of its eigenvectors is always a SIC fiducial [134], and existing numerical evidence strongly supports this claim. In $d = 3$, the Zauner matrix is given by

$$F_Z = \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} \tag{3.83}$$

because $-1 \mod 3 = 2$ and the corresponding Zauner unitary U_{F_Z} takes the form

$$U_{F_Z} = \frac{1}{\sqrt{3}} e^{i\frac{\pi}{6}} \sum_{r,s=0}^{2} \tau^{2(2r^2-2rs)} |r\rangle \langle s| = \frac{1}{\sqrt{3}} e^{i\frac{\pi}{6}} \sum_{r,s=0}^{2} \omega^{2r(r-s)} |r\rangle \langle s|, \tag{3.84}$$

where $\tau = -e^{i\frac{\pi}{3}}$ and $\tau^2 = \omega$. In the standard basis, we can write U_{F_Z} as

$$U_{F_Z} = \frac{e^{i\frac{\pi}{6}}}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ \omega^2 & 1 & \omega \\ \omega^2 & \omega & 1 \end{pmatrix}. \tag{3.85}$$

In Zauner's conjecture [3, 134], the eigenspace of interest is the largest subspace where the overall phase can be chosen so that it has a unit eigenvalue. We denote the Zauner subspace by $\mathcal{Z}$, which in $d = 3$ is given by

$$\mathcal{Z} = \text{span} \left\{ \begin{pmatrix} -\omega^2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\omega^2 \\ 1 \\ 0 \end{pmatrix} \right\}, \tag{3.86}$$

where the span is over the complex field $\mathbb{C}$.

In Sec 3.5, we saw that there are 9 qutrit SICs for which the number of good p-sets is 12: the most exceptional SIC and the 8 SICs generated with $t = \frac{\pi}{9}$. For any other SIC-family, the number of good p-sets is 3. In this section, we want to answer the following questions:

(a) Are there any vectors in the Zauner subspace other than SIC fiducials that give rise to linear dependencies?

(b) Do the Hesse SICs exhaust all SICs with a maximal number of linear dependencies in $d = 3$?

We will show that the answer is affirmative for both questions in $d = 3$.

Observe that any Zauner eigenvector $|\psi\rangle \in \mathcal{Z}$ can be written as

$$|\psi\rangle = \begin{pmatrix} -\omega^2(x+y) \\ y \\ x \end{pmatrix}, \tag{3.87}$$

where $x, y \in \mathbb{C}$. For each $|\psi\rangle \in \mathcal{Z}$, we associate a set P

$$P = \{D_{\mathbf{p}} |\psi\rangle \mid \mathbf{p} \in \mathbb{Z}_3^2\}, \tag{3.88}$$

where $D_{\mathbf{p}}$ are just the usual displacement operators defined by Eq. (3.32). As we have done previously, because there is a one-to-one correspondence between P and all p-vectors in $\mathbb{Z}_3^2$ (except when $y = \omega x$), we can refer to vectors in P by their corresponding p-vectors.

It was mentioned in Sec. 3.5 that the p-sets

$$(00, 12, 21), (01, 10, 22), (02, 11, 20) \tag{3.89}$$

are always good p-sets. This may be seen if we explicitly list down the set P generated by Eq. (3.87):

$$P = \left\{ \begin{array}{lll} \begin{pmatrix} -\omega^2(x+y) \\ y \\ x \end{pmatrix}, & \begin{pmatrix} x \\ -\omega^2(x+y) \\ y \end{pmatrix}, & \begin{pmatrix} y \\ x \\ -\omega^2(x+y) \end{pmatrix}, \\ \begin{pmatrix} -\omega^2(x+y) \\ \omega y \\ \omega^2 x \end{pmatrix}, & \begin{pmatrix} x \\ -(x+y) \\ \omega^2 y \end{pmatrix}, & \begin{pmatrix} y \\ \omega x \\ -\omega(x+y) \end{pmatrix}, \\ \begin{pmatrix} -\omega^2(x+y) \\ \omega^2 y \\ \omega x \end{pmatrix}, & \begin{pmatrix} x \\ \omega(x+y) \\ \omega y \end{pmatrix}, & \begin{pmatrix} y \\ \omega^2 x \\ -(x+y) \end{pmatrix} \end{array} \right\}. \tag{3.90}$$

The vectors of P are displayed so that each row matches exactly with the p-sets in Eq. (3.89).

The Hesse SICs are the only ones which maximize the number of good p-sets. This follows from considering any of 3 SIC vectors that do not correspond to any p-set in Eq. (3.89) and setting the value of their determinant to zero [30]. Doing so gives

$$\begin{vmatrix} 0 & -e^{2it}\omega^{y} & 1 \\ 1 & 0 & e^{2it}\omega^{z} \\ -e^{2it}\omega^{x} & 1 & 0 \end{vmatrix} = 1 + e^{6it}\omega^{x+y+z} = 0 \tag{3.91}$$

which implies that

$$\omega^{r} = e^{6it}, \qquad r = 0, 1, 2. \tag{3.92}$$

which holds if and only if $t = 0$ and $t = \frac{\pi}{9}$, which are the Hesse SICs. Any other SIC with 12 good p-sets must then be unitarily equivalent to the Hesse SICs.

3.7 Lie algebraic properties of qutrit SICs

As we have mentioned before, our interest in SICs stems primarily from the fact that an arbitrary density operator can be expanded in terms of a SIC. However, Marcus Appleby, Steve Flammia, and Chris Fuchs [7] have shown that SICs can be used not just as a basis for the space of Hermitian matrices, they also can serve as a basis for the space of all complex matrices of dimension d, that is, the Lie algebra $\mathfrak{gl}_d(\mathbb{C})$.

If we take any linear operator B acting on some Hilbert space $\mathcal{H}^d$, we can find a unique set of expansion coefficients b_k such that

$$B = \sum_{k=1}^{d^2} b_k \Pi_k. \tag{3.93}$$

For the special case $B = \Pi_i \Pi_j$ the expansion coefficients b_k correspond to the structure coefficients S_{ijk}. If we define the quantity

$$J_{rst} = \frac{d+1}{d}\left(T_{rst} - T^{*}_{rst}\right), \tag{3.94}$$

then using Eq. (3.51), we may write

$$[\Pi_r, \Pi_s] = \sum_{t=1}^{d^2} J_{rst}\Pi_t \tag{3.95}$$

so that the J_{rst} are, in fact, the structure constants for the Lie algebra $\mathfrak{gl}_d(\mathbb{C})$. Note that the J_{rst} are proportional to the imaginary part of the triple products.

It is not difficult to verify that the structure constants J_{rst} satisfy the properties of a Lie algebra. In particular, they obey the Jacobi identity

$$\sum_{q=1}^{d^2} (J_{rsq}J_{tqu} + J_{stq}J_{rqu} + J_{trq}J_{squ}) = 0, \tag{3.96}$$

for all $r, s, t, u = 1, ..., d^2$. It may be worth noting that the matrix J_r with elements

$$[J_r]_{st} = J_{rst} \tag{3.97}$$

is the *adjoint representative* of Π_r in the basis formed by the SIC projectors, i.e.,

$$\mathrm{ad}_{\Pi_r}\Pi_s = [\Pi_r, \Pi_s] = \sum_{t=1}^{d^2} J_{rst}\Pi_t. \tag{3.98}$$

An important property of the adjoint matrix J_r is that it has a spectral decomposition of the form

$$J_r = Q_r - Q_r^T \tag{3.99}$$

where Q_r is a rank-$(d-1)$ projector with the special property that

$$Q_r Q_r^T = 0. \tag{3.100}$$

We refer to this as the Q-Q^T property [7].

The relationship between the triple products and the projectors Q_r is given by

$$T_r = \frac{d}{d+1}Q_r + \frac{2d}{d+1}|e_r\rangle\langle e_r|, \tag{3.101}$$

where the matrix T_r has elements $(T_r)_{st} = T_{rst}$ and

$$|e_r\rangle = \sqrt{\frac{d+1}{2d}} \sum_{s=1}^{d^2} K_{rs}^2 \, |s\rangle \,, \qquad K_{rs} = \sqrt{\frac{d\delta_{rs}+1}{d+1}}, \tag{3.102}$$

and the set $\{|s\rangle\}_{s=1}^{d^2}$ is the standard basis in dimension d^2. Observe that the quantity K_{rs} is just the absolute value of the inner product between SIC vectors, that is, the elements

$$G_{rs} = \langle\psi_r|\psi_s\rangle = K_{rs}e^{i\theta_{rs}} \tag{3.103}$$

of the Gram matrix of the SIC with vectors $|\psi_r\rangle$.

The Gram matrix satisfies an identity

$$\sum_r G_{s_1 r} G_{s_2 r} G_{r t_1} G_{r t_2} = \frac{d}{d+1} (G_{s_1 t_1} G_{s_2 t_2} G_{s_1 t_2} G_{s_2 t_1}) \tag{3.104}$$

which follows from the 2-design property of SICs:

$$\sum_r \Pi_r \otimes \Pi_r = \frac{2d}{d+1} P_{\text{sym}} \tag{3.105}$$

where P_{sym} denotes the projector onto the symmetric subspace of $\mathcal{C}_d \otimes \mathcal{C}_d$. The triple products are related to the Gram matrix elements by

$$T_{rst} = G_{rs} G_{st} G_{tr} = K_{rs} K_{st} K_{tr} e^{i\theta_{rst}} \tag{3.106}$$

where

$$\theta_{rst} = \theta_{rs} + \theta_{st} + \theta_{tr}. \tag{3.107}$$

This shows that the most interesting features of the Gram matrix and the triple products are already contained in the angle tensors θ_{rs}.

Using geometrical considerations of the Q-Q^T property, Appleby, Flammia, and Fuchs showed that the existence of a SIC in dimension d is equivalent to the construction of a highly symmetric basis for the Lie algebra $\mathfrak{gl}_d(\mathbb{C})$ [7]. More specifically, let L_r be any set of d^2 Hermitian operators which form a basis for $\mathfrak{gl}_d(\mathbb{C})$ and C_r be the adjoint representative of L_r. Then the necessary and sufficient condition for C_r to be decomposable into

$$C_r = Q_r - Q_r^T \tag{3.108}$$

where Q_r is a rank $d-1$ projector such that $Q_r Q_r^T = 0$ is that there exists a SIC with elements Π_r such that

$$L_r = \epsilon_r \left(\Pi_r + \alpha I\right) \tag{3.109}$$

for some fixed $\alpha \in \mathbb{R}$ and $\epsilon_r = \pm$. Moreover, the existence of a Hermitian basis for $\mathfrak{gl}_d(\mathbb{C})$ with a Q-Q^T property is necessary and sufficient for the existence of a SIC in dimension d.

To illustrate this, we can use Eq.(3.101) to calculate Q_r from the triple products T_{ijk} for qutrit SICs. In particular, we are interested in the Q-projectors for the most exceptional

SIC given by Eq. (3.46). Suppose we take the projector Q_1, we find that

$$Q_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & -\frac{1}{4} & \mu & \mu & \mu & -\mu & -\mu & -\mu \\ 0 & -\frac{1}{4} & \frac{1}{4} & -\mu & -\mu & -\mu & \mu & \mu & \mu \\ 0 & -\mu & \mu & \frac{1}{4} & -\mu & \mu & -\frac{1}{4} & -\mu & \mu \\ 0 & -\mu & \mu & \mu & \frac{1}{4} & -\mu & -\mu & \mu & -\frac{1}{4} \\ 0 & -\mu & \mu & -\mu & \mu & \frac{1}{4} & \mu & -\frac{1}{4} & -\mu \\ 0 & \mu & -\mu & -\frac{1}{4} & \mu & -\mu & \frac{1}{4} & \mu & -\mu \\ 0 & \mu & -\mu & \mu & -\mu & -\frac{1}{4} & -\mu & \frac{1}{4} & \mu \\ 0 & \mu & -\mu & -\mu & -\frac{1}{4} & \mu & \mu & -\mu & \frac{1}{4} \end{pmatrix}, \tag{3.110}$$

where $\mu = \frac{\sqrt{3}}{12}i$. Note that Q_1 has a very similar structure to J_1 in Eq. (3.67), which is not a terrible surprise since they are related by Eq. (3.99). By inspection, it is easy to see that $Q_1^T = Q_1^*$. Furthermore, it can be verified that Q_1 is a rank-2 projector ($Q_1^2 = Q_1$) and that $Q_1 Q_1^T = 0$.

Actually, the properties above follow directly from Eq. (3.110). For instance, to show that Q_1 is rank-2, observe that there are 6 independent ways to add up rows R_j of Q_1 that vanish, one possible combination being

$$\begin{aligned} R_2 + R_3 &= 0, \\ R_4 + R_7 &= 0, \\ R_5 + R_9 &= 0, \\ R_6 + R_8 &= 0, \\ \left(\frac{1}{4} - \mu\right)(R_4 - R_5) + \left(\frac{1}{4} + \mu\right)(R_5 - R_6) &= 0, \\ \left(\frac{1}{4} + \mu\right)(R_2 + R_5) + \left(\frac{1}{4} - \mu\right)(R_2 - R_4) &= 0. \end{aligned} \tag{3.111}$$

Similarly, we can check that Q_1 is a projector possessing the Q-Q^T property from Eq. (3.110). To show this, it may be useful to know that

$$2\mu\left(\frac{1}{4} + \mu\right) + \left(\frac{1}{4} - \mu\right)^2 = 0. \tag{3.112}$$

Observe that it is sufficient to know any one of the Q_r matrices since the others can be obtained by some rearrangement of elements in Q_1, as dictated by the corresponding S_{ijk}. Thus, let us introduce the matrices

$$V_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad V_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.113}$$

and use them to define the permutation matrices

$$U_{1j} = \begin{pmatrix} V_j & 0 & 0 \\ 0 & V_j & 0 \\ 0 & 0 & V_j \end{pmatrix}, \quad U_{2j} = \begin{pmatrix} 0 & 0 & V_j \\ V_j & 0 & 0 \\ 0 & V_j & 0 \end{pmatrix}, \quad U_{3j} = \begin{pmatrix} 0 & V_j & 0 \\ 0 & 0 & V_j \\ V_j & 0 & 0 \end{pmatrix}. \tag{3.114}$$

Therefore, if we know Q_1 then we can generate all other Q-projectors as follows:

$$\begin{aligned} Q_2 &= U_{12} Q_1 U_{12}^T, & Q_3 &= U_{13} Q_1 U_{13}^T, \\ Q_4 &= U_{21} Q_1 U_{21}^T, & Q_5 &= U_{22} Q_1 U_{22}^T, \\ Q_6 &= U_{23} Q_1 U_{23}^T, & Q_7 &= U_{31} Q_1 U_{31}^T, \\ Q_8 &= U_{32} Q_1 U_{32}^T, & Q_9 &= U_{33} Q_1 U_{33}^T. \end{aligned}$$

In fact, once we know all structure coefficients S_{rst}, then we already know all the projectors Q_r, too. To demonstrate this, define the matrix S_r with elements

$$[S_r]_{st} = S_{rst}. \tag{3.115}$$

The elements Q_{rst} of the Q-projectors are then given by

$$Q_{rst} = [Q_r]_{st} = [S_r]_{st} \left(\frac{d\delta_{st} + 1}{d+1} \right) \delta_{rt}. \tag{3.116}$$

The matrix S_r has d unit eigenvalues and $d^2 - d$ zero eigenvalues so Eq. (3.116) is consistent with the fact that the matrices Q_r have rank $d-1$: the term $(1-\delta_{rt})$ essentially turns a column of S_r into zeroes, which means the rank of Q_r must be one less than S_r. We can show that Eq.(3.116) holds in general by considering

$$\begin{aligned} S_{rst} - S_{rts} &= \frac{1}{d} \left[(d+1)(T_{rst} - T_{rts}) - \frac{d}{d+1} (\delta_{rs} - \delta_{rt}) \right] \\ &= (Q_{rst} - Q_{rts}) - \frac{1}{d+1} (\delta_{rs} - \delta_{rt}), \end{aligned} \tag{3.117}$$

where we used

$$T_{rts} = \mathrm{Tr}\left(\Pi_r \Pi_t \Pi_s\right) = \mathrm{Tr}\left(\Pi_t \Pi_s \Pi_r\right) = \mathrm{Tr}\left(\left(\Pi_r \Pi_s \Pi_t\right)^\dagger\right) = T^*_{rst}, \tag{3.118}$$

and Eq.(3.99).

Looking at Eq. (3.117), a reasonable guess would be

$$Q_{rst} = S_{rst} - \left(\frac{1}{d+1}\right)\delta_{rt}. \tag{3.119}$$

However, we also know that $r = t \neq s, S_{rst} = \frac{1}{d+1}$ so it seems that we should be able to just write

$$Q_{rst} = S_{rst}(1 - \delta_{rt}). \tag{3.120}$$

To include the case $r = s = t$, we must use the Q-Q^T property in Eq. (3.100); from it we conclude that

$$Q_{rrr} = 0, \tag{3.121}$$

which also follows from Eq.(3.119) only if it is true that

$$Q_{rsr} = Q_{rrs} = 0, \qquad \forall r \neq s. \tag{3.122}$$

Thus, to fully incorporate the Q-Q^T property, we should rewrite Eq.(3.119) as

$$Q_{rst} = S_{rst} - \left(\frac{d\delta_{st} + 1}{d+1}\right)\delta_{rt}. \tag{3.123}$$

Note that this is not directly obtainable from Eq. (3.117) because in that case we were only looking at differences $Q_{rst} - Q_{rts}$. The Q-Q^T property is necessary for us to determine all components such that $Q^2 = Q$. Eq. (3.123) says that knowing the set of Q-projectors is equivalent to knowing the structure coefficients of a SIC, which can be used to reconstruct the SIC.

Chapter 4

Practical implementations of SIC-POVMs

Quantum state tomography describes various methods for estimating the state of a quantum system from the statistics of a sufficiently informative set of measurements [67,103]. It continues to be one of the primary challenges in quantum information and quantum computing since many quantum information processing tasks requires the ability to determine the quantum state in a robust and efficient manner.

The state of an ensemble of d-dimensional quantum systems is typically described by a density operator, which in general is an abstract list of d^2 numbers. This means that to perform tomography involves carrying out at least d^2 linearly independent measurements and then using one of a number of mathematical techniques to determine the state, such as maximum likelihood estimation.

In practical quantum state tomography, SIC-POVMs are particularly important because they are optimal according to the following fairly standard measures of efficiency:

(i) they use the minimal number of expectation values, leading to less redundancy and faster convergence in state reconstruction, and

(ii) they achieve the smallest error probability with respect to Hilbert-Schmidt distance between the estimated and "true" density operator for the measured system [119].

Here we describe various methods for experimental SIC-POVMs, including two new proposals, one involving a storage loop with weak projections, while another involving Naimark unitaries implemented with integrated optics.

4.1 Experimental SICs using optimal polarimetry

In a quantum polarimetry experiment, Alexander Ling et al. [87] estimated the polarization of an identically prepared ensemble of single photons by an optimal measurement equivalent to a SIC-POVM. We provide a brief summary of their experiment in this section.

In classical optics, the polarization state of light is often described in terms of 4 Stokes parameters S_0, S_1, S_2, S_3 where S_0 is the total intensity. They are typically written in vector form called the Stokes vector $\vec{S}$. A SIC-POVM can be identified from $\vec{S}$ if we take the normalized Stokes vector

$$\vec{S} = \left(1, \frac{S_1}{S_0}, \frac{S_2}{S_0}, \frac{S_3}{S_0}\right) \tag{4.1}$$

and identify the last 3 components with points on a Poincaré sphere. A minimal scheme for estimating the Stokes vector requires a measurement that corresponds to the vertices of a tetrahedron inscribed by the Poincaré sphere. Each measurement operator E_j can be written as

$$E_j = \frac{1}{4}\vec{e_j} \cdot \vec{\sigma} \tag{4.2}$$

where $\vec{e_j} = (1, v_{j1}, v_{2j}, v_{j3})$ are the normalized Stokes vectors associated with the tetrahedron whose vertices are given by the 3-dimensional vectors $\vec{v_j}$ and $\vec{\sigma}$ is a vector consisting of the identity matrix σ_0 and the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$. Intuitively, the regular tetrahedron minimizes any redundancies in the measurement results because it samples the state space in the most uniform way possible for 4 outcomes.

The measurement operators E_j are related to the detected intensities I_j by

$$\frac{I_j}{I} = \langle E_j \rangle = \frac{1}{4}\vec{e_j} \cdot \vec{S}. \tag{4.3}$$

We can solve for the Stokes vector from

$$\vec{S} = M^{-1}\vec{I} \tag{4.4}$$

where $\vec{I} = (I_1, I_2, I_3, I_4)$ and M is the so-called instrument matrix

$$M = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}} & 0 \\ 1 & \sqrt{\frac{1}{3}} & -\sqrt{\frac{2}{3}} & 0 \\ 1 & -\sqrt{\frac{1}{3}} & 0 & -\sqrt{\frac{2}{3}} \\ 1 & -\sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{2}{3}} \end{pmatrix} \tag{4.5}$$

with rows that can be constructed from the vectors e_j.

Writing the entries of the qubit density matrix ρ as a column vector $\vec{\rho}$, we obtain the Stokes vector from the relation

$$\vec{\rho} = \frac{1}{2}\Sigma\vec{S} \tag{4.6}$$

where Σ is a matrix whose columns are the Pauli matrices,

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}. \tag{4.7}$$

Combining Σ and M into the tomography matrix $T = \Sigma M^{-1}/2$, we get

$$\vec{\rho} = T\vec{I}, \tag{4.8}$$

that is, by choosing the measurement corresponding to T, we can estimate the quantum state from the intensity readings I_j of the tetrahedron operators.

As a figure of merit, the determinant of the matrix that maps the POVM histograms onto the coefficients of the density operator is a measure of optimality. The tetrahedron measurement gives the smallest determinant that can be obtained for any POVM and is regarded as optimal in this sense.

The tetrahedron measurement is also identified as a way to realize a discrete Wigner distribution for qubits since

$$A(q,p) = \frac{1}{2}\sum_{j,k=0}^{1}(-1)^{pj-qk+jk/2}X^jZ^k \tag{4.9}$$

corresponds to a qubit phase operator on phase point (q,p), where X and Z are the usual Pauli operators, and

$$W_\rho(q,p) = \frac{1}{2}\mathrm{Tr}\left(\rho A(q,p)\right) \tag{4.10}$$

is Wootters' discrete Wigner function. In a subsequent experiment by Thomas Durt, et al. [34], they used a 2-qubit POVM consisting of the above tetrahedron measurement and its counterpart anti-tetrahedron measurement obtained by taking the points opposite to the vertices of the original one to perform tomography on Bell states, as a component of a quantum key distribution scheme with entangled qubits.

4.2 Implementing SICs by successive measurements

In a recent paper, Amir Kalev, et al. propose a method for performing a SIC-POVM with linear optics in a two-step measurement process [71]. They describe the specific decomposition of known SIC-POVMs in dimensions 2, 3, 4, and 8. Here we provide a brief summary of their two-step procedure as it applies to the general case.

Recall that a general quantum measurement is composed of a set of N outcomes represented by positive operators E_i, where the probability of each outcome is given by

$$p(j) = \mathrm{Tr}\left(\rho E_j\right). \tag{4.11}$$

If the j^{th} outcome is found, the post-measurement state of the system is given by

$$\rho_j = \frac{1}{p(j)} A_j \rho A_j^\dagger \tag{4.12}$$

where A_j is the Kraus operator for the j^{th} outcome,

$$E_j = A_j^\dagger A_j \tag{4.13}$$

The Kraus decomposition of POVM $\{E_j\}_{j=1}^N$ is not unique since

$$\left(U_j A_j\right)^\dagger \left(U_j A_j\right) = A_j^\dagger A_j = E_j \tag{4.14}$$

for any unitary transformation U_j. The unitaries correspond in practice to different ways of implementing the measurement.

Consider a system subject to a sequence of measurements $\mathcal{E}$ and $\mathcal{F}$, both with d outcomes,

$$\mathcal{E} = \left\{E_i = A_i^\dagger A_i\right\}_{i=1}^d, \qquad \mathcal{F} = \left\{F_j^{(i)}\right\}_{j=1}^d \tag{4.15}$$

where the subscript (i) suggests that the outcomes of the second measurement generally depends on the actual outcome obtained in the first measurement.

Using the Born rule,

$$\Pr\left[E_i \text{ and } F_k\right] = \mathrm{Tr}\left(\rho A_i^\dagger F_j^{(i)} A_i\right). \tag{4.16}$$

Thus, measuring $\mathcal{E}$ and $\mathcal{F}$ in sequence is equivalent to a single POVM with d^2 outcomes,

$$\mathcal{M} = \left\{M_{ij} = A_i^\dagger F_j^{(i)} A_i\right\}_{i,j=1}^d. \tag{4.17}$$

In particular, we are interested in finding $\mathcal{E}$ and $\mathcal{F}$ such that $\mathcal{M}$ corresponds to a SIC-POVM.

Let X, Z be the generators of the Weyl-Heisenberg group on $\mathcal{H}_d$ defined by their action on some orthonormal basis vectors $|j\rangle$,

$$X\,|j\rangle = |j \oplus 1\rangle\,, \qquad Z\,|j\rangle = \omega^j\,|j\rangle\,, \tag{4.18}$$

where $\oplus$ signifies addition modulo d and $\omega = e^{2\pi i/d}$. A SIC-POVM with Weyl-Heisenberg symmetry is generated from a fiducial projector $|\psi\rangle\langle\psi|$ under the action of the group elements

$$\Pi_{ij} = X^i Z^j |\psi\rangle\langle\psi| Z^{\dagger j} X^{\dagger i} \tag{4.19}$$

with the fiducial state $|\psi\rangle$ chosen such that

$$\mathrm{Tr}\,(\Pi_{ij}\Pi_{kl}) = \frac{d\delta_{ik}\delta_{jl} + 1}{d^2(d+1)}. \tag{4.20}$$

Expressing the fiducial state as

$$|\psi\rangle = \sum_{j=0}^{d-1} \alpha_j\,|j\rangle \tag{4.21}$$

we obtain

$$\Pi_{ij} = \sum_{k,l=0}^{d-1} \alpha_i \alpha_j^* \omega^{j(k-l)}\,|i \oplus k\rangle\,\langle l \oplus i| \tag{4.22}$$

which can be recognized as a two-step measurement where

(i) the first measurement has Kraus operators

$$A_i = \sum_{j=0}^{d-1} \alpha_j |i \oplus j\rangle\langle i \oplus j| \tag{4.23}$$

(ii) and the second measurement consists of projection operators onto the Fourier basis

$$F_j = \frac{1}{d} \sum_{k,l=0}^{d-1} \omega^{j(k-l)}\,|k\rangle\,\langle l|. \tag{4.24}$$

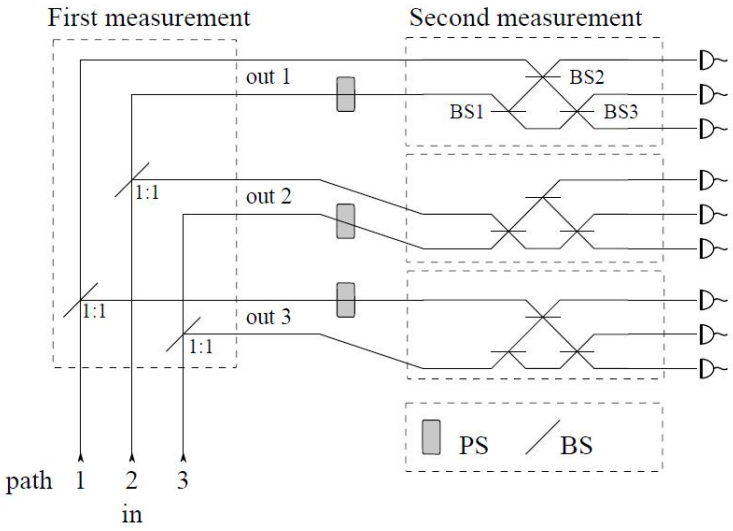


Figure 4.1: Implementing a continuous family of SICs for a path-encoded qutrit using successive measurements. Image reproduced from Ref. [71]

It is straightforward to verify that Eqs. (4.23) and (4.24) yield

$$\Pi_{ij} = A_i^\dagger F_j A_i \tag{4.25}$$

Note that the first measurement corresponds to a fuzzy measurement with full-rank elements consisting of mixtures of basis projectors with the identity operator, i.e.,

$$E_i = \frac{1}{d}(1-\mu)I + \mu|i\rangle\langle i| \tag{4.26}$$

where $-\frac{1}{d-1} \leq \mu \leq 1$ is required for positivity.

In $d = 3$, there is a one-parameter family of SICs which are given by the fiducial vectors

$$\left|\phi_t^{jk}\right\rangle = \frac{1}{\sqrt{2}}\left(|k\rangle - \omega^j e^{2it}\,|k\oplus 1\rangle\right) \tag{4.27}$$

This continuous SIC-family can be realized in a sequence consisting of a fuzzy measurement with operators

$$A_k^\dagger = \frac{1}{\sqrt{2}}\left(|k\oplus 1\rangle\langle k\oplus 1| + |k\oplus 2\rangle\langle k\oplus 2|\right), \qquad k = 1,2,3, \tag{4.28}$$

followed by the application of a diagonal unitary transformation

$$U_k^\dagger = |k\rangle\langle k| - e^{2it}|k\oplus 1\rangle\langle k\oplus 1| + |k\oplus 2\rangle\langle k\oplus 2| \tag{4.29}$$

whenever the outcome k is obtained from the fuzzy measurement, and then ending with a projection onto the Fourier basis. In an optical setting with path-encoded qutrits, this sequence correspond to a set of equal beam splitters for the fuzzy measurement, path-dependent phase shifters for implementing the unitary $U_k^\dagger$ and a Fourier transformation realized with three beam splitters BS1, BS2, and BS3, which perform the unitaries

$$U_{\rm BS1} = \frac{\sigma_X + \sigma_Z}{\sqrt{2}}, \quad U_{\rm BS2} = \frac{\sqrt{2}\sigma_X + \sigma_Z}{\sqrt{3}}, \quad U_{\rm BS3} = \frac{\sigma_Y + \sigma_Z}{\sqrt{2}}, \tag{4.30}$$

respectively, where σ_i is a Pauli operator. The general optical circuit is shown in Fig. 4.1.

4.3 Storage loop SIC- POVM experiment with weak projections

In this section, we describe the first experiment to characterize the state of a qutrit using a SIC-POVM performed by Medendorp et al. in Toronto [95]. A schematic of the experiment in shown in Fig. 4.2. To implement the SIC-POVM, consider a sequence of d^2 partial projections involving low-reflectivity partially polarizing beam splitters (PPBS). In the ideal case, the signal consisting of identically prepared photons are cycled into an optical storage loop where the PPBS projects the photons onto the d^2 SIC-POVM elements. Since the projections are weak, the state of the photons is nearly unaltered as they propagate through the loop. In the limit of no coupling and for a loop with no losses, the signal goes around many times until it eventually leaks out of one of the ports. The time-integral of each port constitutes the measurement of a single SIC-POVM projector. Of course, in practice, the optical components have losses that affect the symmetry of the measurement setup, leading to deviations from ideal SIC-POVM outcomes.

For a signal photon, the qutrit is encoded in the path and polarization information. The photon has some probability amplitude to traverse each path, where the first path has both vertical and horizontal polarization and the second path allows only vertical polarization. Thus, the basis states for a qutrit are given by

$$|0\rangle = |V_1\rangle\,, \qquad |1\rangle = |H_1\rangle\,, \qquad |2\rangle = |V_2\rangle\,. \tag{4.31}$$

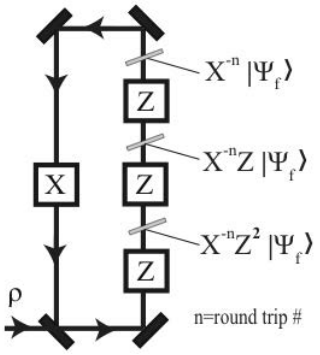


Figure 4.2: SIC-POVM storage loop scheme implemented by partial projections over three round trips. The PPBS labels indicate how the detector click is interpreted.

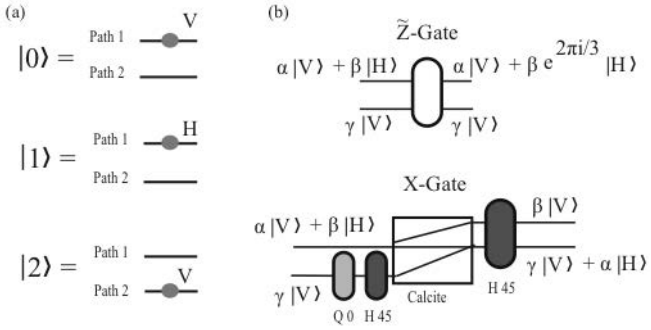


Figure 4.3: (a) Qutrit state encoding. (b) Implementing the X and $\tilde{Z}$ gates.

The qutrit state encoding is shown in Fig. 4.3(a).

In the actual experiment, the SIC-POVM chosen was the one associated with the fiducial vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{4.32}$$

which is the one of the exceptional SICs for $t = \frac{\pi}{6}$. Part of the reason for choosing this SIC is that it leads to SIC vectors which can be measured by projection onto the first path, since the zero component can always be put in the second path.

A schematic of the experimental apparatus is displayed in Fig. 4.4. The SIC-POVM measurement is obtained from partial projections and permuting through the SIC elements with X and Z gates. The storage loop is constructed with 3 Z gates and 1 X gate which couple with the signal in 3 round trips. If n is the round trip number, the weak

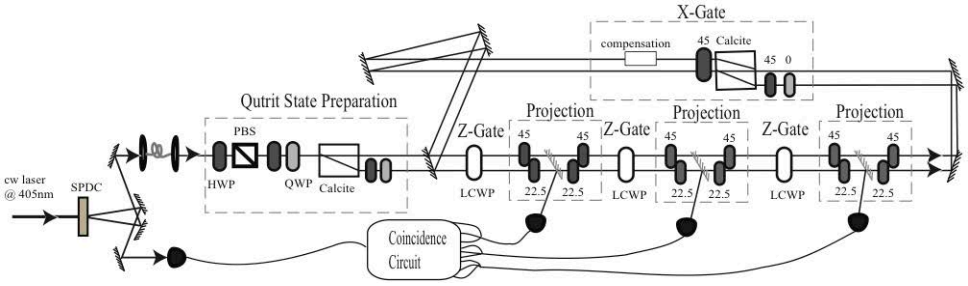


Figure 4.4: Schematic for the storage loop SIC experiment.

projections performed after each Z gate correspond to measuring the SIC-POVM element corresponding to $X^{-n}Z^2, X^{-n}Z$, andX^n, respectively, as shown in Fig. 4.3(b).

A liquid crystal wave plate (LCWP) applies a $\frac{2\pi}{3}$ phase shift between the horizontal and vertical polarization, performing the $\tilde{Z}$ gate. This is different from the Z gate in that the second path experiences no phase shift. But because of the particular SIC-POVM chosen, the second path is never directly measured, ensuring that $Z^n |\psi\rangle = \tilde{Z}^n |\psi\rangle$, and as long as $\tilde{Z}^3 = I$, the loop functions correctly. A calcite beam displacer and 3 HWP allows us to switch between basis states and thus acts as the X gate.

The weak projections are performed by a collection of HWPs and PPBSs after each Z gate. Since the measurements are performed on the first path only, the coherence between paths is significant only for the X gate. With a compensating crystal at one point in the loop, coherence is maintained for the first two rounds. Custom thickness pieces of BK7 are used for coherence in the last round.

The experiment was designed to test the quantum law of total probability (QLTP), which is the Born rule expressed in terms of a SIC-POVM

$$p(B_j) = (d+1)\sum_{i=1}^{d^2} p(A_i)p(B_j|A_i) - 1 \tag{4.33}$$

where $p(A_i)$ are the outcome probabilities of a SIC-POVM, $p(B_j|A_i)$ are conditional probabilities for the outcome B_j of some other measurement given that the SIC-POVM outcome A_i was obtained first, and $p(B_j)$ are the probabilities for the other measurement being performed directly on the system. This particular form of the Born rule looks almost like the classical law of total probability and so is called the QLTP.

Using the experimentally corrected SIC-POVM outcomes, we tested the QLTP for several input states. For example, for the input state $\frac{1}{\sqrt{2}}(|0\rangle + |2\rangle)$, the QLTP predicted

the probability of measuring $\frac{1}{\sqrt{2}}(|0\rangle + i|2\rangle)$ to be $0.526 \pm 0.009(\text{stat}) \pm 0.03(\text{sys})$, while in a direct measurement we found $0.506 \pm 0.004 \pm 0.002$. For the input state $|0\rangle$, the QLTP prediction for the probability of measuring $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$ was $0.01 \pm 0.0002(\text{stat}) \pm 0.009(\text{sys})$, while in a direct measurement we found $0.005 \pm 0.0004 \pm 0.001$. These results are consistent with the QLTP.

4.4 Multiport scheme from Naimark dilation of SIC-POVMs

In this section, we describe an experimental scheme for performing SIC-POVMs with integrated optics. An integrated optical circuit is a device that combines multiple functions for manipulating photons analogous to a microchip for electronic systems, the essential difference being that the former works in optical wavelengths. Photonic integrated circuits allows for compact, efficient, and scalable performance, which is important for optical systems used to implement a quantum computation. Physical realizations of integrated optics include silica-based planar lightwave circuit interferometers fabricated on a semiconductor chip [73, 74, 94, 107], periodically poled lithium niobate waveguides in standard fiber optics technology [19, 93, 109, 116] and the direct-write femtosecond laser technique for manufacturing optical waveguides in dielectric materials [32, 92, 102], which allows one to realize three-dimensional photonic circuits, in contrast with the flat architecture of conventional lithography.

The flexibility afforded by integrated optics makes it the natural candidate for realizing something called a *multiport device*, which is a blackbox for performing any discrete unitary operation through a sequence of qubit operations acting on various optical modes, in the case of photonic systems [112]. An optical version of a multiport devices features a sequence of beam splitters and phase shifters with parameters that can be adjusted to perform any quantum gate operation, i.e., unitary transformation. Note that a qubit unitary acting on optical modes transforms two input modes (k, l) into two output modes (k', l') according to

$$\begin{pmatrix} k' \\ l' \end{pmatrix} = \begin{pmatrix} e^{i\phi}\sin\theta & e^{i\phi}\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \begin{pmatrix} k \\ l \end{pmatrix} \tag{4.34}$$

and allows measurements to be done in any other basis. With a multiport device, unitary operations, and therefore arbitrary projective measurements, can also be realized in higher-dimensional quantum systems, providing a means of performing full state tomography with a quantum circuit, in particular, a means for realizing a SIC-POVM.

But how does one realize a measurement in terms of a unitary operation? Here it is important to recall a particular result called Naimark's theorem [83]. In most common version of the theorem states that there is no loss in generality in considering only projective measurements because any general measurement or POVM on a quantum system can be achieved by a joint projective measurement on the system and an ancilla, provided the ancilla has large enough a dimension and that it is initialized to a known pure quantum state [106].

In the finite dimensional case, there is a simple procedure for determining the projective measurement that corresponds to a POVM. Suppose we have a POVM $\{E_i : i = 1, 2, ..., N\}$ and each POVM operator E_i is proportional to a rank-1 projection, which is described by the subspace spanned by $e_i \in \mathcal{H}_M$. Given $N > M$ and the POVM elements are not mutually orthogonal, we can construct the $M \times N$ matrix

$$V = (e_1 \dots e_N) \tag{4.35}$$

and look for an $(N - M) \times N$ matrix W such that

$$U = \begin{pmatrix} V \\ W \end{pmatrix} \tag{4.36}$$

is an $N \times N$ unitary matrix. In this case, the projectors associated with the columns of U correspond to the desired projective measurement. It is by using the Naimark extension of SIC-POVMs that we can design experiments for realizing SIC-POVMs with optical multiport devices.

The method here is completely general in that the Naimark extension can be found for SIC-POVMs of any dimension but for concreteness, we describe experiments for qubit and qutrit SIC-POVMs [124]. Experiments that feature higher-dimensional quantum systems are far from trivial, making the qutrit case more interesting but the qubit case is useful for illustrating the basic idea behind the experimental scheme with multiport devices and so we consider it first.

4.4.1 Multiport Qubit SIC-POVM

Any qubit SIC-POVM is associated with the vertices of a regular tetrahedron circumscribed by the Bloch sphere. There are infinitely many ways to choose this tetrahedron but we may choose the projectors corresponding to the set of vectors

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} e^{i\frac{\pi}{3}} \\ e^{-i\frac{\pi}{3}}\sqrt{2} \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} e^{-i\frac{\pi}{3}} \\ e^{i\frac{\pi}{3}}\sqrt{2} \end{pmatrix} \right\}. \tag{4.37}$$

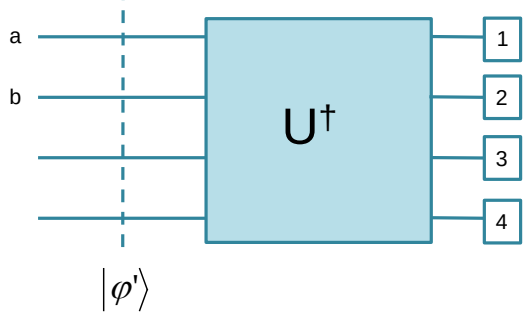


Figure 4.5: Schematic for a multiport qubit SIC-POVM experiment.

This qubit SIC-POVM can be realized using an optical setup with path-encoded qubits, where information is encoded in the photon amplitude to be found in each mode.

Writing the qubit SIC-POVM vectors as a 2×4 matrix, one particularly simple Naimark extension is given by

$$U = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & 1 & e^{i\frac{\pi}{3}} & e^{-i\frac{\pi}{3}} \\ 0 & \sqrt{2} & \sqrt{2}e^{-i\frac{\pi}{3}} & \sqrt{2}e^{i\frac{\pi}{3}} \\ \sqrt{3} & -1 & -e^{i\frac{\pi}{3}} & -e^{-i\frac{\pi}{3}} \\ 0 & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \end{pmatrix}. \tag{4.38}$$

The significance of U for the multiport experiment is as follows. Denote the columns of U by $|U_j\rangle$ and the vectors in Eq. (4.37) by $|u_j\rangle$. If $\{|j\rangle : j = 1, 2, \ldots, 4\}$ is the standard basis in 4 dimensions then

$$U^\dagger |U_j\rangle = |j\rangle . \tag{4.39}$$

Now consider the path-encoded qubit $|\varphi\rangle = (a, b)^T$ depicted in Fig. 4.5. Treating this qubit as the 4-level state $|\varphi'\rangle = (a, b, 0, 0)^T$, it is easy to check that

$$|\langle\varphi'|U_j\rangle|^2 = \frac{1}{2} |\langle\varphi|u_j\rangle|^2 . \tag{4.40}$$

This means that the outcomes corresponding to the qubit SIC-POVM in Eq. (4.37) are obtained by a projective measurement in the standard basis after applying $U^\dagger$ to $|\varphi'\rangle$.

In practice, $U^\dagger$ can be implemented using a 4×4 multiport device. To derive an optical implementation for such a multiport, we need to decompose U into a sequence of unitaries composed of 2×2 blocks, each of which can then be performed by some combination of phase shifters and beam splitters on the appropriate optical modes. One such decomposition is

achieved using the following matrices:

$$U_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \qquad U_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{3} & 0 & 0 & 0 \\ 0 & 1 & \sqrt{2} & 0 \\ 0 & -\sqrt{2} & 1 & 0 \\ 0 & 0 & 0 & \sqrt{3} \end{pmatrix},$$

$$U_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\frac{\pi}{3}} & 0 \\ 0 & 0 & 0 & e^{-i\frac{\pi}{6}} \end{pmatrix},$$

$$U_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U_6 = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

$$U_7 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{-i\frac{2\pi}{3}} \end{pmatrix}. \tag{4.41}$$

It is straightforward to check that

$$U = U_1 U_2 U_3 U_4 U_5 U_6 U_7. \tag{4.42}$$

The optical multiport circuit that implements the qubit SIC-POVM according to Eq. (4.42) is shown in Fig. 4.6.

4.4.2 Multiport Qutrit SIC-POVM

Next we do a similar construction for the qutrit case. Here the qutrit SIC-POVM that we wish to implement is the most exceptional SIC given by the fiducial vector

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{4.43}$$

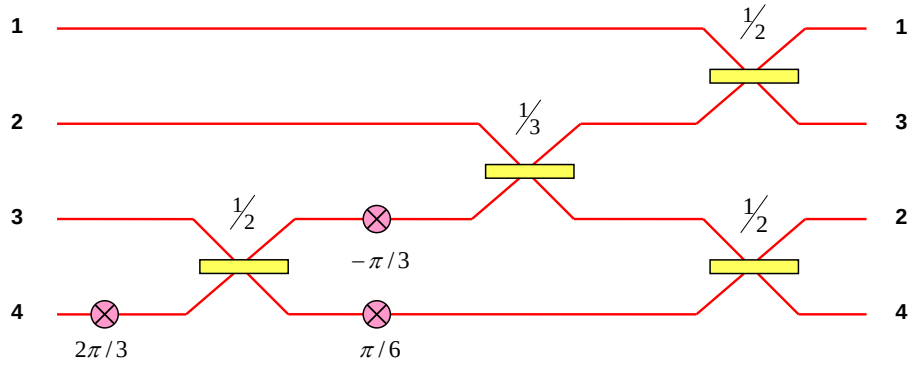


Figure 4.6: Qubit SIC-POVM implemented with an optical multiport device.

The multiport experiment for implementing the SIC-POVM corresponding to $|\psi\rangle$ is shown in Fig. 4.7.

To explain how the circuit works, consider the input state $|\phi\rangle = (a, b, c)^T$. Analogous to the qubit case, we think of $|\phi\rangle$ as a 9-level state $|\phi'\rangle = (a, 0, 0, b, 0, 0, c, 0, 0)^T$. We are looking for a projective measurement on $|\phi'\rangle$ that gives the same statistics as measuring the most exceptional SIC on $|\phi\rangle$. The desired measurement basis correspond to the columns of

$$V = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & 0 & 0 & -1 & -\omega^2 & -\omega & 1 & \omega & \omega^2 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \omega^2 & \omega & 1 & \omega & \omega^2 & 0 & 0 & 0 \\ 1 & \omega & \omega^2 & 0 & 0 & 0 & -1 & -\omega^2 & -\omega \\ 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \omega^2 & \omega & 1 & \omega & \omega^2 \\ -1 & -\omega^2 & -\omega & 1 & \omega & \omega^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 1 & \omega & \omega^2 & 0 & 0 & 0 & 1 & \omega^2 & \omega \end{pmatrix}. \tag{4.44}$$

Observe that the 9 column vectors formed by rows 1, 4, and 7 are exactly the SIC vectors for the most exceptional SIC. Thus, V for the qutrit case plays the role of U in the qubit case.

Denote the columns of V by $|V_j\rangle$ and the normalized column vectors formed by taking rows 1, 4, and 7 of V by $|v_j\rangle$. If $\{|j\rangle : j = 1, \ldots, 9\}$ is the standard basis in 9 dimensions

then $V^\dagger |V_j\rangle = |j\rangle$. Moreover, one can check that

$$|\langle\phi'|V_j\rangle|^2 = \frac{1}{3}|\langle\phi|v_j\rangle|^2 \tag{4.45}$$

so the outcomes of the standard basis measurement after applying $V^\dagger$ yield statistics that is expected from the qutrit SIC-POVM.

The unitary $V^\dagger$ can be realized with a 9×9 multiport device. However, the particular structure of V actually allows us to perform the SIC-POVM in terms of 3×3 unitary operations. The decomposition of V into a sequence of 3×3 unitaries is possible because V exhibits the block-circulant structure

$$V = \begin{pmatrix} A & B & C \\ C & A & B \\ B & C & A \end{pmatrix} \tag{4.46}$$

where A, B, C are the 3×3 blocks in Eq. (4.44). V can be transformed into block diagonal form by

$$S = P \otimes I_3 \equiv \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ \omega^2 & 0 & 0 & \omega & 0 & 0 & 1 & 0 & 0 \\ 0 & \omega^2 & 0 & 0 & \omega & 0 & 0 & 1 & 0 \\ 0 & 0 & \omega^2 & 0 & 0 & \omega & 0 & 0 & 1 \\ \omega & 0 & 0 & \omega^2 & 0 & 0 & 1 & 0 & 0 \\ 0 & \omega & 0 & 0 & \omega^2 & 0 & 0 & 1 & 0 \\ 0 & 0 & \omega & 0 & 0 & \omega^2 & 0 & 0 & 1 \end{pmatrix}, \qquad P = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ \omega^2 & \omega & 1 \\ \omega & \omega^2 & 1 \end{pmatrix}. \tag{4.47}$$

It is worth noting that the generalization of P to all higher dimensions is known as the Bell multiport [85], and there is an efficient system for realizing the Fourier transform with a small number of optical elements [137].

Thus, we have

$$Q = SVS^\dagger \equiv \begin{pmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & Q_3 \end{pmatrix} \tag{4.48}$$

where

$$Q_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & \omega - \omega^2 & \omega^2 - \omega \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 2 & -1 & -1 \end{pmatrix},$$
$$Q_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} \omega^2 - \omega & 0 & \omega - \omega^2 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \\ -\omega^2 & 2\omega^2 & -\omega^2 \end{pmatrix}, \tag{4.49}$$
$$Q_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} \omega - \omega^2 & \omega^2 - \omega & 0 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \\ -\omega & -\omega & 2\omega \end{pmatrix}.$$

Since S and Q are both composed of blocks of 3×3 unitaries, this gives us the decomposition of V,

$$V = S^\dagger Q S, \tag{4.50}$$

where Q has blocks acting on modes (1,2,3), (4,5,6), and (7,8,9) while S has blocks acting on modes (1,4,7), (2,5,8), and (3,6,9) of the input state $|\phi'\rangle$. The qutrit SIC-POVM requires $V^\dagger$, which translates into the scheme displayed in Fig. 4.7. Because each $P^\dagger$ acts on independent sets of modes, it is more economical in practice to include them in the input state preparation, reducing the depth of the measurement circuit to just 2 parallel layers corresponding to the Q_k and P. Since all components are 3×3 unitaries, a very compact realization of the qutrit SIC-POVM is possible using integrated 3×3 multimode interference devices.

To complete the construction, we determine optical networks for the unitaries $P^\dagger$ and $Q_k^\dagger$, $k = 1, 2, 3$. A circuit for $Q_k^\dagger$ can be constructed from the Euler decomposition of a three-dimensional rotation, permutation matrices, and diagonal unitaries. We can write

$$Q_k^\dagger = G_k R D_k, \tag{4.51}$$

where D_k are the diagonal matrices

$$D_1 = \begin{pmatrix} -i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} -i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \omega \end{pmatrix}, \quad D_3 = \begin{pmatrix} -i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \tag{4.52}$$

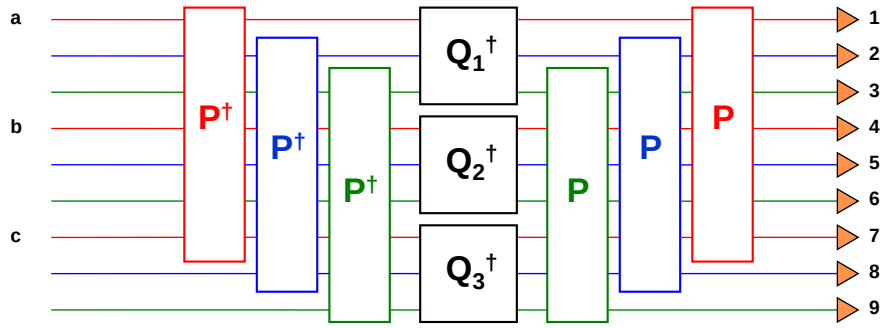


Figure 4.7: Schematic for a qutrit SIC-POVM generated by Eq.(4.43) using multiport devices.

which can be performed using phase shifters alone, R is the rotation matrix

$$R = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & \sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & -1 \\ -\sqrt{3} & \sqrt{2} & -1 \end{pmatrix}, \tag{4.53}$$

and G_k are matrices that permutes the rows of $Q_k^\dagger$,

$$G_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{4.54}$$

Using the Euler decomposition

$$R = R_1(x)R_2(y)R_3(z) \tag{4.55}$$

where

$$R_1(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos x & -\sin x \\ 0 & \sin x & \cos x \end{pmatrix}, \quad R_2(y) = \begin{pmatrix} \cos y & 0 & -\sin y \\ 0 & 1 & 0 \\ \sin y & 0 & \cos y \end{pmatrix}, \quad R_3(z) = \begin{pmatrix} \cos z & -\sin z & 0 \\ \sin z & \cos z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.56}$$

the optical elements for our quantum circuit can be determined using

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \sqrt{\epsilon} & \sqrt{1-\epsilon} \\ \sqrt{1-\epsilon} & -\sqrt{\epsilon} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.57}$$

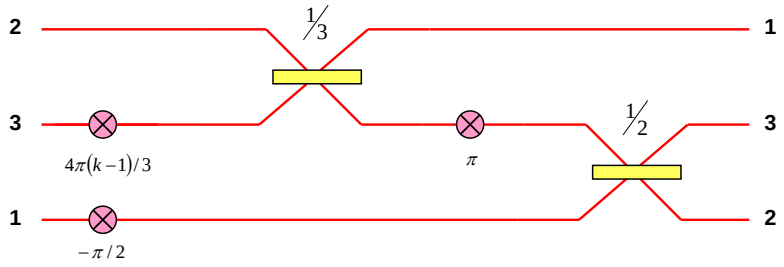


Figure 4.8: Multiport circuit for the unitary operators $Q_k^\dagger$. The circuits for $Q_k^\dagger$ differ only by a phase shift and a relabeling of output modes according to G_k. The output modes shown are for $Q_1^\dagger$.

which shows how a 2-dimensional rotation can be achieved by a beam splitter with reflectivity $\epsilon = \cos^2\alpha$ and a π-phase shifter. For the particular decomposition in Eq. (4.55), we have

$$x = -\frac{\pi}{4}, \quad y = -\arccos\left(-\frac{1}{\sqrt{3}}\right), \quad z = \frac{\pi}{2}. \tag{4.58}$$

The unitaries $Q_k^\dagger$ have almost identical optical circuits as prescribed by Eq. (4.51) since they differ only by a permutation and relative phase in one row. The optical network for $Q_k^\dagger$ is shown in Fig. 4.8. With top-to-bottom path labeling, the circuit shown realizes $Q_1^\dagger$. Only two phase shifters are needed because $R_3(\pi/2)$ is simply a swap operator sandwiched by two π-phase shifts acting on different paths, which collapses to a mere relabeling of paths.

The unitary P transforms the modes into the Fourier basis and can be implemented using the same circuit as $Q_k^\dagger$ since

$$P = Q_1^\dagger \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & 1 & 0 \end{pmatrix}, \tag{4.59}$$

that is, we just combine an equal beam splitter with $Q_1^\dagger$.

4.4.3 Remarks on improvements and feasibility

We have described an experimental scheme for directly implementing SIC-POVMs with integrated optics. It is worthwhile to comment on how the optical multiport scheme serves

as an improvement over previous techniques and what possible issues there are regarding its practical feasibility.

A key component for realizing the scheme is the SIC-POVM decomposition into optical networks. Designing the scheme for linear optics is practical given that various experimental groups are already currently working on multimode interference experiments with path encoded qudits but in principle the scheme can be realized with any setup that can perform the Naimark unitaries correspond to SIC-POVMs. As to the prescribed decomposition itself, we would like to emphasize that our decomposition shows a remarkable improvement over the general procedure for discrete unitary operators as described by Reck, et al. [112]. Reck's method can be sketched in following way. Let d be the Hilbert space dimension and $N = d^2$. Then any N-dimensional unitary operator can be realized by an optical circuit with $(N-1)$ beam splitters between modes 1 and 2, $(N-2)$ beam splitters between modes 2 and 3, and so on until one beam splitter between modes $(N-1)$ and N, a phase shifter accompanying each beam splitter, and finally $(N-1)$ phase shifters before measuring the output. Counting all the optical elements needed, we obtain (N^2-1) elements. Thus, by Reck's method no more than (d^4-1) elements are required to realize the unitary for a d-dimensional SIC-POVM, but for particular unitary operators, it is possible that considerably fewer elements are needed. For example, Reck's method suggests that the 4×4 unitary for the qubit SIC-POVM needs 15 elements but our optical multiport scheme here uses only 7 elements. Similarly, the unitary corresponding to a qutrit SIC-POVM would generally require 80 elements but the one we present here uses only 44 elements.

We also remark on how the optical multiport scheme for qutrit SIC-POVMs compares to the storage loop experiment [95] of Sec 4.3. In that scheme, photons are sent into an optical storage loop and the SIC-POVM is performed by weakly coupled PPBS, with a full measurement completed after the signal makes 3 rounds. Because of the loop design, it uses only 22 elements for the measurement: 3 liquid crystal wave plates, 1 beam displacer, 3 PPBS, and the rest are half wave plates. Ideally, each PPBS extracts an infinitesimal portion of the signal that circles the loop infinitely many times. In the actual experiment, each PPBS reflected about 3% of the horizontal polarization and 13% of the vertical polarization. This led to a significant drop in visibilities even before we finish performing the SIC-POVM. Even with compensation wave plates that help tune the visibility of interference between modes, the visibilities achieved were 74% and 36% for the first and second rounds, respectively. Errors due to beam deformation and other losses in the loop were also corrected during the post-measurement analysis, but some of the results were still off by as much as 20% from the theoretically expected values. An accurate assessment of the sensitivity of the multiport scheme to losses in the optical elements will depend on the actual apparatus used to realized it but because there is no weak coupling demand,

we expect it to be more robust to errors and show a vast improvement in performance compared to the Medendorp results.

To properly assess the feasibility of the multiport experiment, it is important to know how sensitive the scheme is to fabrication errors or losses in the optical elements. The specifics will depend on the particular apparatus used for the multiport scheme. For instance, one might employ commercially available micro-optical chips used in some telecom devices, which have built-in optical elements that can then be set to the appropriate values. Or in more standard cases, one is likely to fabricate the integrated optics circuit himself. In any case, the biggest issue will be maintaining the interferometric phase stability among the optical modes. For example, in the qutrit case, the fact that the Q_k and P act on different sets of modes means one has to actively check and correct for possibly unstable relative phases, which is quite a nontrivial task at present. A way to possibly avoid the phase stability issue is to fabricate the 9×9 unitary for the qutrit case as a single nine-mode integrated device, but since our decomposition involves 44 elements that might create its own challenges.

Chapter 5

The SIC representation of quantum states

Quantum foundations is the branch of physics that seeks to elucidate what the laws of quantum mechanics imply about the nature of the physical phenomena it describes. In particular, it aims to identify a correct interpretation of the quantum formalism, which will hopefully provide a clear picture of reality in a quantum world.

So far we have studied properties of SICs mainly in the context of quantum information, where it is a special case of a generalized measurement appropriate for various tomographic purposes. Here we shift gears and examine the role SICs play in understanding quantum theory itself.

The conceptual significance of SICs follows from the fact that the set $\{\Pi_i\}_{i=1}^{d^2}$ is linearly independent and thus forms a basis for the space of d-dimensional density operators. This allows us to define the bijective mapping $\vec{p} \mapsto \rho$,

$$\rho = \sum_{i=1}^{d^2} \left[(d+1)p(i) - \frac{1}{d} \right] \Pi_i, \tag{5.1}$$

which establishes a one-to-one correspondence between probability distributions for a SIC and quantum states represented by density operators. Therefore, the probabilities $p(i)$ provide an equivalent description of quantum states, which is called the *SIC representation* and is the subject of this chapter.

5.1 Quantum states as degrees of belief

Progress in quantum foundations has generally meant taking the mathematical structure of quantum mechanics as given but establishing an account that identifies the real, physical objects implied by the theory, which is independent of any external observer. Numerous attempts to furnish such an account have led to a variety of interpretations that can be broadly categorized into two camps, ontic and epistemic.

The ontic camp regards the quantum state as a real entity with an objective existence akin to the physical properties possessed by a classical system. The standard approach is to identify mathematical symbols in the theory literally as some actual physical object. It suffices to consider 3 examples to understand how this works. One is the de Broglie-Bohm pilot wave theory, where particles exhibit quantum behavior by obeying essentially classical equation but accompanied by a real wave function in configuration space that dictates the dynamics in a completely nonlocal fashion. Another is the class of spontaneous collapse theories, of which the Ghirardi-Rimini-Weber (GRW) variant is most popular, which accepts the usual unitary evolution of quantum systems but interrupts this occasionally with some random nonlinear effect, removing the need for the process of measurement. The last one we mention is the many-worlds interpretation that introduces the idea of a multiverse, obtained from a deterministic quantum state evolving into multiple branches with independent realities, and where the probabilistic nature of quantum systems is merely an apparent effect of residing in a single branch of reality.

While the ontic approaches have their merits, the main issue with them is that they appear mostly to be arbitrary quick fixes that are in no way made necessary by the quantum formalism. Moreover, to attribute reality to quantum states introduces a host of paradoxes that one either accepts as necessary, like a wave function that connects everything nonlocally, or is explained away as being unobservable, like the unpredictable nonlinear collapses or other branches of the multiverse.

The alternative is to adopt the epistemic view that quantum states are states of incomplete knowledge about the system [61,122]. The epistemic approaches can be distinguished by how they characterize the information encoded in quantum states. There are the objectivists who ascribe this information to some hidden property of the quantum system, which is objective but not fully accessible. This lack of complete information leads to the probabilities obtained from quantum states. On other hand, the subjectivists regard the information as representing an agent's credible beliefs, reflecting his state of knowledge about the system.

Those who subscribe to a subjective view of quantum states are sometimes called

quantum Bayesians. The term "Bayesian" refers to the mathematician Thomas Bayes, who provided the first rigorous analysis of the problem of inference. To evaluate the probability of a hypothesis, a Bayesian starts with some prior probability, which is updated when new information is available. This Bayesian notion of probability treats probabilities as an extension of logical reasoning and manipulates them according to a set of criteria for testing consistency, in much the same way as logic tells us how to manipulate propositions. A probability assignment is a tool an agent uses to make decisions based on his beliefs. The rules of probability theory ensures that the agent accounts for those beliefs in a coherent way.

Rudolph Peierls [104] believed that the only legitimate way to understand quantum mechanics is to consider the density operator as representing "our knowledge of the system we are trying to describe," thereby advocating an epistemic view of quantum states. As once emphasized by Chris Fuchs and Asher Peres [44], "Quantum theory does not describe physical reality. What it does is provide an algorithm for computing probabilities for the macroscopic events that are the consequence of our experimental interventions."

Chris Fuchs and Rüdiger Schack push the Bayesian viewpoint further by stating that the information embodied by a quantum state represents an agent's degrees of belief about the future behavior of a quantum system. To distinguish their approach from other versions of quantum Bayesianism, they have dubbed it as QBism. The general idea of a Bayesian approach to quantum probabilities can actually be traced back to Lane Hughston in 1984, although he mentioned this in the context of introducing a nonlinear modification to the Schrodinger equation. By contrast, QBism does not make any radical changes to the quantum formalism. What QBism proposes is a personalist interpretation to the probabilities that appear in quantum theory. The personalist Bayesian considers probability theory as a calculus of consistency about one's beliefs, irrespective of the source of those beliefs. Hence, a probability is a mere calculational tool for a consistent accounting of information; it does not exist as an objective property of the system.

One important implication of subjective quantum probabilities is that there is no such thing as a correct quantum state for a system, since everyone assigns a state according to his own beliefs. There are circumstances, however, where two agents are logically led to the same state. For instance, when we speak of determining an unknown quantum state from a tomographic measurement, the de Finetti theorem says that what we really do is update a prior probability assignment assuming the systems being measured are exchangeable [26]. As long as an agent does not assign a zero probability to some state, the theorem shows that independent observers with minimal agreement in beliefs will converge to some product state assignment for the exchangeable systems.

But if the proper way of thinking about quantum states just a matter of interpreting probabilities correctly, why is the quantum formalism not directly written in the language of probabilities, instead of using concepts in Hilbert space? Actually, it is always possible to formulate quantum theory in terms of probabilities. We only need the sort of measurement that can be used to determine a quantum state uniquely in a tomography experiment. For QBism, this key ingredient is a SIC-POVM, which establishes a one-to-one mapping between density operators and probabilities via Eq. (5.1).

In quantum theory, probabilities are computed from the Born rule. This means that we can also think of these probabilities as subject to the agent who assigns them. The only difference in the quantum case is that in order to be consistent with the rules of quantum mechanics, the probability assignment must also be constrained by the Born rule, which when written in terms of a SIC measurement $\{\Pi_i\}_{i=1}^{d^2}$ gives

$$q(j) = (d+1)\sum_{i=1}^{d^2} p(i)p(j|i) - 1, \tag{5.2}$$

where $p(i)$ is the probability of outcome π_i measured directly, $q(j)$ is the probability of outcome E_j of any other measurement $\{E_j\}_{j=1}^{N}$ measured directly, and $p(j|i) = \mathrm{Tr}\left(\Pi_i E_j\right)$ is the conditional probability of E_j given Π_i. Because of its similarity to the classical law of total probability, Eq. (5.2) is also called the quantum law of total probability. Fuchs and Schack develop an axiomatic framework for reconstructing quantum theory by treating Eq. (5.2) as an empirical addition to Bayesian probability theory [47]. In the next two chapters, we explore the structure of quantum states in the QBist framework.

5.2 Representing quantum states with SICs

Our first task is to find the conditions for a SIC probability distribution to be a pure state. Recall that a density operator ρ corresponds to a pure quantum state if and only if it is a rank-1 projector, i.e., $\rho^2 = \rho$. The SIC representation for ρ is given by Eq. (5.1). As long as $\vec{p}$ is a probability distribution, ρ in Eq. (5.1) is guaranteed to be a Hermitian operator. In this case, a remarkable theorem independently shown by Nick Jones and Noah Linden [70], and by Steve Flammia [41] states that an equivalent condition for pure states is given by

$$\mathrm{Tr}\left(\rho^2\right) = \mathrm{Tr}\left(\rho^3\right) = 1. \tag{5.3}$$

Translating Eq. (5.3) in terms of SICs, we obtain constraints on SIC probabilities associated with pure states. From $\mathrm{Tr}\,(\rho^2) = 1$ we obtain

$$\sum_i p(i)^2 = \frac{2}{d(d+1)}. \tag{5.4}$$

From $\mathrm{Tr}\,(\rho^3) = 1$, we can work out that

$$\sum_{i,j,k} T_{ijk} p(i)p(j)p(k) = \frac{d+7}{(d+1)^3} \tag{5.5}$$

where

$$T_{ijk} = \mathrm{Tr}\,(\Pi_i \Pi_j \Pi_k) \tag{5.6}$$

are the triple products we first encountered in Chapter 3.

Geometrically, we can think of the SIC probabilities as points on a simplex embedded in real vector space. Then Eq. (5.4) says that the probabilities corresponding to pure states lie on a sphere while Eq. (5.5) picks out a manifold of points on that sphere that actually corresponds to quantum states. Since quantum state space is a compact, convex set, the convex hull of probabilities satisfying Eq. (5.4) and Eq. (5.5) corresponds to the set of all quantum states. Moreover, because the extreme points lie on a sphere, it means the length of any vector belonging to the set is less than or equal to the radius of the sphere. This implies that the scalar product of two probability distributions is bounded. The upper bound is given by

$$\vec{p} \cdot \vec{q} \leq \frac{2}{d(d+1)}, \tag{5.7}$$

which follows from the radius of the sphere, while a lower bound can be obtained from considering $\mathrm{Tr}\,(\rho\sigma) \geq 0$ for density operators, which translates into

$$d(d+1)\vec{p} \cdot \vec{q} - 1 \geq 0 \quad \Longrightarrow \quad \vec{p} \cdot \vec{q} \geq \frac{1}{d(d+1)}. \tag{5.8}$$

An important observation that we will exploit throughout the rest of our discussion is the fact that the real part of T_{ijk} is sufficient in Eq. (5.5) since

$$\begin{aligned}
\sum_{i,j,k} T_{ijk} p(i)p(j)p(k) &= \frac{1}{2} \sum_{i,j,k} \mathrm{Tr}\,(\Pi_k \Pi_j \Pi_i)\, p(k)p(j)p(i) + \frac{1}{2} \sum_{i,j,k} \mathrm{Tr}\,(\Pi_i \Pi_j \Pi_k)\, p(i)p(j)p(k) \\
&= \sum_{i,j,k} \frac{1}{2} \Big[\mathrm{Tr}\left((\Pi_i \Pi_j \Pi_k)^\dagger \right) + \mathrm{Tr}\,(\Pi_i \Pi_j \Pi_k) \Big] p(i)p(j)p(k), \\
&= \sum_{i,j,k} \tilde{T}_{ijk} p(i)p(j)p(k)
\end{aligned} \tag{5.9}$$

where $\tilde{T}_{ijk} = \mathrm{Re}\,[T_{ijk}]$. This also follows from the fact that the imaginary parts of T_{ijk} are completely antisymmetric.

For $\mathrm{Tr}\,(\rho^3) = 1$, we get another expression in terms of the structure coefficients S_{ijk} by using

$$\sum_k S_{ijk} = \frac{d\delta_{ij} + 1}{d+1}, \tag{5.10}$$

which is obtained by taking the trace of

$$\Pi_i \Pi_j = \sum_{k=1}^{d^2} S_{ijk} \Pi_k. \tag{5.11}$$

Performing the algebra, we get

$$\sum_{i,j,k} S_{ijk} p(i) p(j) p(k) = \frac{4}{d(d+1)^2} \tag{5.12}$$

where we may also replace S_{ijk} with just the real parts $\tilde{S}_{ijk}$.

Actually, we can get a condition for pure state probabilities directly from the idempotence condition $\rho = \rho^2$ if we plug in ρ from Eq. (5.1) and take the trace of both sides, yielding

$$p(k) = \frac{(d+1)^2}{3d} \sum_{i,j} T_{ijk} p(i) p(j) - \frac{1}{3d} \tag{5.13}$$

for $k = 1, 2, ..., d^2$. The same condition but with structure coefficients is given by

$$p(k) = \frac{(d+1)}{3} \sum_{i,j} S_{ijk} p(i) p(j) + \frac{2}{3d(d+1)}. \tag{5.14}$$

Because we will always be concerned with probability distributions obtained from SICs, any probability distribution that satisfies Eq. (5.4) and Eq. (5.5) or Eq. (5.12) will also be called a pure state, which is implicitly identified with a rank-1 projector via Eq. (5.1).

We obtain a slight simplification if we note that

$$T_{iii} = 1, \quad T_{ijj} = T_{jij} = T_{jji} = \frac{1}{d+1}, \tag{5.15}$$

from which it follows that

$$S_{iii} = 1, \quad S_{ijj} = S_{jij} = \frac{1}{d+1}, \quad S_{jji} = 0. \tag{5.16}$$

Substituting these values into Eq.(5.12) gives

$$\sum_i p(i)^3 + 2\left(\frac{1}{d+1}\sum_i\sum_{j\neq i} p(i)^2 p(j)\right) + \sum_{i\neq j\neq k} S_{ijk}p(i)p(j)p(k) \tag{5.17}$$

where using Eq.(5.4) and arranging terms, we get

$$\frac{d-1}{d+1}\sum_i p(i)^3 + \sum_{i\neq j\neq k} S_{ijk}p(i)p(j)p(k) = 0, \tag{5.18}$$

5.3 Pure states as fixed points of a map

If we look at Eq. (5.18), we see that the left-hand side can be regarded of as a real-valued scalar function

$$F(\vec{p}) = \left(\frac{d-1}{d+1}\right)\sum_i p(i)^3 + \sum_{i\neq j\neq k} S_{ijk}p(i)p(j)p(k) \tag{5.19}$$

acting on probability vectors. In this case, we get the condition that for $\vec{p}$ satisfying

$$\vec{p}\cdot\vec{p} = \frac{2}{d(d+1)}, \tag{5.20}$$

then $F(\vec{p}) = 0$ for pure states. So in the SIC representation, the intricate geometry of the convex set of quantum states that is a consequence of the positivity conditions on density operators appears to be encoded primarily in this cubic constraint. Thus, it is useful to know some of the properties of $F(\vec{p})$.

One of our attempts to find a some simple property of $F(\vec{p})$ was to take derivatives of it with respect to the probabilities $p(i)$. This led us to an alternative way to define pure states in terms of the fixed points of a certain mapping from the simplex to the real vector space in which the simplex is embedded. Fixed points are interesting because in many different contexts, they are related to the concepts of stability and equilibrium. For instance, in areas of study where the function describes an iterative process, the important fixed points describe values for which the function tends to converge, in which case the fixed points are said to be attractive.

To construct the fixed point mapping for pure states, we start by computing the gradient of $F(\vec{p})$, which is obtained from taking its partial derivatives with respect to $p(k)$:

$$\frac{\partial F(\vec{p})}{\partial p(k)} = 3\left(\frac{d-1}{d+1}\right)p(k)^2 + 3\sum_{i\neq j(\neq k)} S_{ijk}p(i)p(j) \tag{5.21}$$

where the factor of 3 in the second term on the right-hand side is because

$$S_{ijk} = S_{jki} = S_{kij}, \qquad \text{for } i \neq j \neq k, \tag{5.22}$$

each of which should be counted separately. The notation $(\neq k)$ is meant to indicate that the indices i, j, k are all distinct but there is no sum over k.

We note that

$$\begin{aligned}\sum_{i,j} S_{ijk} p(i) p(j) = &\sum_{j(\neq k)} S_{kjk} p(k) p(j) + \sum_{i \neq k} S_{ikk} p(i) p(k) \\ &+ \sum_{i(\neq k)} S_{iik} p(i)^2 + p(k)^2 + \sum_{i \neq j (\neq k)} S_{ijk} p(i) p(j).\end{aligned} \tag{5.23}$$

Substituting the known values of S_{ijk} in Eq. (5.16) into Eq. (5.23) we have

$$\sum_{i,j} S_{ijk} p(i) p(j) = \frac{2}{d+1} p(k) + \left(\frac{d-1}{d+1}\right) p(k)^2 + \sum_{i \neq j \neq k} S_{ijk} p(i) p(j), \tag{5.24}$$

Plugging this into Eq. (5.13), we get

$$p(k) = (d+1) \left[\sum_{i \neq j (\neq k)} S_{ijk} p(i) p(j) + \left(\frac{d-1}{d+1}\right) p(k)^2 \right] + \frac{2}{d(d+1)}. \tag{5.25}$$

Comparing Eq. (5.21) with Eq. (5.25), we finally obtain

$$p(k) = \frac{(d+1)}{3} \frac{\partial F}{\partial p(k)} + \frac{2}{d(d+1)}. \tag{5.26}$$

We can summarize the main result as follows: consider the mapping

$$G : \Delta_{d^2} \to \mathbb{R}^{d^2}; \quad \vec{p} \mapsto \vec{q} \tag{5.27}$$

from probability distributions $\vec{p} \in \Delta^{d^2-1}$ to vectors $\vec{q} \in \mathbb{R}^{d^2}$ such that

$$\vec{q} = G(\vec{p}) = \frac{(d+1)}{3} \vec{\nabla} F(\vec{p}) + \frac{2}{d(d+1)} \vec{J} \tag{5.28}$$

where $\vec{\nabla}F$ is the gradient of Eq. (5.19) and $\vec{J}$ is the vector of all ones,

$$\vec{J} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \tag{5.29}$$

that is, it is a column vector with all d^2 entries equal to 1. The pure states are then given by the fixed points of G,

$$G(\vec{p}) = \vec{p} \tag{5.30}$$

provided that

$$\sum_{i=1}^{d^2} p(i)^2 = \frac{2}{d(d+1)}. \tag{5.31}$$

5.4 Pure states as stationary points of a map

In constructing the fixed point mapping in Eq. (5.28), we used derivatives of $F(\vec{p})$. Once we have derivatives of a function, we can ask for the its stationary points, which can be a minimum, maximum, or a saddle point. Stationary points are important for smooth real-valued functions since minima and maxima tend to signify optimal values in many contexts. In this section, we show the probability vectors representing pure states correspond to the stationary points of a certain map, and maybe knowing whether they are minimum or maximum points of this map will tell us something new about pure states.

The form of Eq. (5.28) suggests that one can find a function $M(\vec{p})$ such that $\vec{\nabla}M(\vec{p}) = \vec{0}$ for SIC probabilities $\vec{p}$ representing pure states. To construct the function M, we use a lemma by Kimura [78] that expresses the positivity condition for density operators in terms of trace powers of ρ. Define

$$\begin{aligned} a_1 &= \mathrm{Tr}\left(\rho\right), \\ a_2 &= \frac{\mathrm{Tr}\left(\rho\right)^2 - \mathrm{Tr}\left(\rho^2\right)}{2}, \\ a_3 &= \frac{1}{6}\left[\mathrm{Tr}\left(\rho\right)^3 - 3\mathrm{Tr}\left(\rho\right)\mathrm{Tr}\left(\rho^2\right) + 2\mathrm{Tr}\left(\rho^3\right)\right], \end{aligned} \tag{5.32}$$

where $a_0 = 1, a_1, a_2$, and a_3 are first few coefficients of the characteristic polynomial of ρ,

$$\det(\lambda I - \rho) = \sum_{i=0}^{d} a_i x^{d-i} = 0. \tag{5.33}$$

If $a_1 = 1$ and $a_2 = a_3 = 0$, then

$$\mathrm{Tr}\left(\rho\right) = \mathrm{Tr}\left(\rho^2\right) = \mathrm{Tr}\left(\rho^3\right) = 1, \tag{5.34}$$

so ρ must be a pure state, according to Eq. (5.3).

To derive M, we will first relax some of the usual conditions on ρ. In particular, we do not assume that $\mathrm{Tr}\left(\rho\right) = 1$. We will then show that

$$M(\vec{p}) = \frac{1}{6}\left[\mathrm{Tr}\left(\rho(\vec{p})\right)^3 - 3\mathrm{Tr}\left(\rho(\vec{p})\right)\mathrm{Tr}\left(\rho(\vec{p})^2\right) + 2\mathrm{Tr}\left(\rho(\vec{p})^3\right)\right], \tag{5.35}$$

where $\rho(\vec{p})$ is the density operator expressed in the SIC representation given by Eq. (5.1).

Calculating the required traces:

$$\mathrm{Tr}\left(\rho\right) = (d+1)\sum_i p(i) - d, \tag{5.36}$$

$$\mathrm{Tr}\left(\rho^2\right) = d\left(d+1\right)\sum_i p(i)^2 + (d+1)\left(\sum_i p(i)\right)^2 - 2\left(d+1\right)\sum_i p(i) + d,$$

$$\begin{aligned}\mathrm{Tr}\left(\rho^3\right) = {} & d\left(d+1\right)^2 \sum_{i\neq j\neq k} S_{ijk}p(i)p(j)p(k) - d + d\left(d+1\right)\left(d-1\right)\sum_i p(i)^3 \\ & + 3\left(d+1\right)\sum_i p(i) + 2d\left(d+1\right)\sum_i p(i)^2 \sum_i p(i) - 3d\left(d+1\right)\sum_i p(i)^2 \\ & - 3\left(d+1\right)\left(\sum_i p(i)\right)^2 + (d+1)\left(\sum_i p(i)\right)^3 + d\left(d+1\right)\sum_i p(i)^2 \sum_i p(i).\end{aligned}$$

The components of $\vec{\nabla}M$ are given by

$$\frac{\partial M}{\partial p(k)} = \frac{1}{6}\left[3\mathrm{Tr}\left(\rho\right)^2 \frac{\partial \mathrm{Tr}\left(\rho\right)}{\partial p(k)} - 3\mathrm{Tr}\left(\rho\right)\frac{\partial \mathrm{Tr}\left(\rho^2\right)}{\partial p(k)} - 3\frac{\partial \mathrm{Tr}\left(\rho\right)}{\partial p(k)}\mathrm{Tr}\left(\rho^2\right) + 2\frac{\partial \mathrm{Tr}\left(\rho^3\right)}{\partial p(k)}\right]. \tag{5.37}$$

Differentiating the trace powers in Eq. (5.36), we get

$$\begin{aligned}
\frac{\partial \mathrm{Tr}\left(\rho\right)}{\partial p(k)} &= d+1, \\
\frac{\partial \mathrm{Tr}\left(\rho^2\right)}{\partial p(k)} &= 2d\left(d+1\right)p(k) + 2\left(d+1\right)\sum_i p(i) - 2\left(d+1\right), \\
\frac{\partial \mathrm{Tr}\left(\rho^3\right)}{\partial p(k)} &= 3d\left(d+1\right)^2 \sum_{i\neq j(\neq k)} S_{ijk}p(i)p(j) + 3d\left(d^2-1\right)p(k)^2 + 2d\left(d+1\right)\sum_i p(i)^2 \\
&\quad + 4d\left(d+1\right)p(k)\sum_i p(i) + 3\left(d+1\right) - 6d\left(d+1\right)p(k) - 6\left(d+1\right)\sum_i p(i) \\
&\quad + 3\left(d+1\right)\left(\sum_i p(i)\right)^2 + 2d\left(d+1\right)p(k)\sum_i p(i) + d\left(d+1\right)\sum_i p(i)^2.
\end{aligned} \tag{5.38}$$

Next we substitute Eq. (5.38) into Eq. (5.39) and demand that $\vec{p}$ be a probability distribution. After some algebra, we arrive at

$$\begin{aligned}
\left.\frac{\partial M(\vec{p})}{\partial p(k)}\right|_{\sum_i p(i)=1} &= \frac{1}{6}\Bigg\{6d(d+1)\left[\left(\frac{d+1}{3}\right)\frac{\partial F}{\partial p(k)} - p(k)\right] \\
&\qquad + \left[6d(d+1) - 3d(d+1)^2\right]\sum_i p(i)^2 + 6(d+1)\Bigg\}.
\end{aligned} \tag{5.39}$$

where we used Eq. (5.21) in the previous section. To show that Eq. (5.39) vanishes for pure states, we can rewrite it as

$$\begin{aligned}
\frac{\partial M(\vec{p})}{\partial p(k)} &= \frac{1}{6}\Bigg\{6d(d+1)\left[\left(\frac{d+1}{3}\right)\frac{\partial F}{\partial p(k)} - p(k) + \sum_i p(i)^2\right] \\
&\quad + \left[-3d(d+1)^2\sum_i p(i)^2 + 6(d+1)\right]\Bigg\}, \\
&= d(d+1)\left[\left(\frac{d+1}{3}\right)\frac{\partial F}{\partial p(k)} - p(k) + \frac{2}{d(d+1)}\right] - \frac{d(d+1)^2}{2}\left[\sum_i p(i)^2 - \frac{2}{d(d+1)}\right] \\
&= 0.
\end{aligned} \tag{5.40}$$

where the first term is zero because it just Eq. (5.28) while the second term is zero if the

probability distribution obeys lies on the sphere containing pure states,

$$\vec{p} \cdot \vec{p} = \frac{2}{d(d+1)}. \tag{5.41}$$

We summarize the result here: consider the function $M : \mathbb{R}^{d^2} \to \mathbb{R}$ that maps real vectors $\vec{p}$ to real numbers $M(\vec{p})$,

$$M(\vec{p}) = \frac{1}{6}\left[\mathrm{Tr}\left(\rho\right)^3 - 3\mathrm{Tr}\left(\rho\right)\mathrm{Tr}\left(\rho^2\right) + 2\mathrm{Tr}\left(\rho^3\right)\right], \tag{5.42}$$

where

$$\rho = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right]\Pi_i. \tag{5.43}$$

If $\vec{p}$ corresponds to a probability distribution then $\vec{p}$ determines a pure quantum state if and only if

$$M(\vec{p}) = 0 \quad \text{and} \quad \vec{\nabla} M(\vec{p}) = \vec{0} \tag{5.44}$$

provided that

$$\vec{p} \cdot \vec{p} = \frac{2}{d(d+1)}. \tag{5.45}$$

Therefore, a probability distribution $\vec{p}$ obeying Eq. (5.45) is a pure state if and only if it is a stationary point of M.

As you may already suspect, since the function $M(\vec{p})$ depends on $\vec{p}$ only through the SIC representation, it seems that its stationary points are pure states independent of the representation used for ρ. This is indeed the case. Consider

$$\rho(\vec{a}) = \sum_j a_j A_j \tag{5.46}$$

for some basis $\{A_j\}$ for the space of density operators, or more generally, for the space of Hermitian operators. Supposing that $\mathrm{Tr}\left(\rho\right) = 1$,

$$M(\rho) = \frac{1}{6}\left[1 - 3\mathrm{Tr}\left(\rho(\vec{a})^2\right) + 2\mathrm{Tr}\left(\rho(\vec{a})^3\right)\right]. \tag{5.47}$$

We then have

$$\frac{\partial M(\vec{a})}{\partial a_k} = -3\frac{\partial \mathrm{Tr}\left(\rho^2\right)}{\partial a_k} + 2\frac{\partial \mathrm{Tr}\left(\rho^3\right)}{\partial a_k} = -6\mathrm{Tr}\left[\left(\rho - \rho^2\right)\frac{\partial \rho}{\partial a_k}\right] \tag{5.48}$$

Since

$$\frac{\partial \rho}{\partial a_k} = A_k \neq 0 \tag{5.49}$$

then Eq. (5.48) is equal to zero for any k if and only if $\rho = \rho^2$, that is, ρ is a rank-1 projector. Thus, $M = 0$ and $\vec{\nabla} M = 0$ define pure states for any representation $\rho(\vec{a})$ assuming only that $\mathrm{Tr}\,(\rho) = 1$.

The next logical question is, what sort of stationary points are pure states? The conventional way to determine this is by computing the eigenvalues of the Hessian matrix $H(M)$, which has elements

$$[H(M)]_{ij} = \frac{\partial^2 M(\vec{p})}{\partial p(j) \partial p(i)} \tag{5.50}$$

evaluated at the stationary points $\vec{p}$. According to the second partial derivative test, if the eigenvalues of $H(M)$ are all positive (negative) then the stationary point is a local minimum (maximum). If $H(M)$ has a mix of positive and negative eigenvalues then $\vec{p}$ is a saddle point. If there are any zero eigenvalues then the test is deemed inconclusive.

For the Hessian matrix $H(M)$, we need the second partial derivatives of the traces:

$$\begin{aligned}
\frac{\partial^2 \mathrm{Tr}\,(\rho)}{\partial p(j) \partial p(k)} &= 0, \\
\frac{\partial^2 \mathrm{Tr}\,(\rho^2)}{\partial p(j) \partial p(k)} &= 2d(d+1)\delta_{jk} + 2(d+1)p(j), \\
\frac{\partial^2 \mathrm{Tr}\,(\rho^3)}{\partial p(j) \partial p(k)} &= 6d(d+1)\left[(d+1)\sum_i S_{ijk} p(i) + p(j) + p(k)\right] + 6(d+1)\left[\sum_i p(i) - 1\right] \\
&\quad + 6d(d+1)\delta_j k \left[(d-1)p(j) + \sum_i p(i) - 1\right].
\end{aligned} \tag{5.51}$$

Then the elements of $H(M)$ are given by

$$\begin{aligned}
\frac{\partial^2 M(\vec{p})}{\partial p(j) \partial p(k)} = \frac{1}{6}\Bigg[& 6\mathrm{Tr}\,(\rho) \frac{\partial \mathrm{Tr}\,(\rho)}{\partial p(j)} \frac{\partial \mathrm{Tr}\,(\rho)}{\partial p(k)} + 3\mathrm{Tr}\left(\rho^2\right) \frac{\partial^2 \mathrm{Tr}\,(\rho)}{\partial p(j) \partial p(k)} \\
& - 3\frac{\partial \mathrm{Tr}\,(\rho)}{\partial p(j)} \frac{\partial \mathrm{Tr}\,(\rho^2)}{\partial p(k)} - 3\mathrm{Tr}\,(\rho) \frac{\partial^2 \mathrm{Tr}\,(\rho^2)}{\partial p(j) \partial p(k)} \\
& - 3\frac{\partial^2 \mathrm{Tr}\,(\rho)}{\partial p(j) \partial p(k)} \mathrm{Tr}\left(\rho^2\right) - 3\frac{\partial \mathrm{Tr}\,(\rho)}{\partial p(k)} \frac{\partial \mathrm{Tr}\,(\rho^2)}{\partial p(j)} + 2\frac{\partial^2 \mathrm{Tr}\,(\rho^3)}{\partial p(j) \partial p(k)}\Bigg],
\end{aligned}$$

which we would like to evaluate when $\vec{p}$ is a probability distribution.

Using a specific example, we consider the example of pure states generated by the most exceptional SIC in $d = 3$. We performed numerical calculations where we randomly generate a pure state ρ and solve for its corresponding SIC probability distribution $\vec{p}$. Then, we computed the eigenvalues of the Hessian matrix $H(M)$ and we found that for all qutrit pure states considered

$$\text{eigenvalues of } H(M) = \{36, -12, -12, -12, 0, 0, 0, 0, 0\}. \tag{5.52}$$

Because it contains zero eigenvalues, according to the second partial derivative test for multivariable functions, this result is unfortunately inconclusive. However, in the case of qutrits, we do know what are stationary points are because the function $M(\vec{p})$ just corresponds to the determinant of the corresponding density operator, $\text{Det}(\rho)$. Since this is equal to the product of the eigenvalues, it is minimized by any boundary state, which has $M(\vec{p}) = 0$. But only the pure states will have vanishing first partial derivatives so they are the minimum points among quantum states.

This observation also helps us explain the set of eigenvalues we obtained in Eq. (5.52). There are 5 zeroes because the manifold of pure states for qutrits is 5-dimensional, so in all likelihood, the zeroes correspond to directions that move towards other pure states. We have one positive eigenvalue so $M(\vec{p})$ increases parallel to the direction of that eigenvector. We have 3 negative eigenvalues which means $M(\vec{p})$ can become negative in these directions; however, moving in those directions probably takes you out of the subset of the probability simplex associated with quantum states.

Something similar can be used to argue that the pure states are minimum points in higher dimensions. Although $M(\vec{p})$ is no longer the determinant, nonetheless it corresponds to the sum of various combinations of the product of 3 eigenvalues of ρ. (Alternatively, it corresponds to one of the coefficients in Kimura's lemma 5.8.1 in Sec.!5.8, which should be non-negative for a density operator.) This means that it will always be bounded below by zero, which is attained by any boundary state with at most 2 nonzero eigenvalues. But the results in this section tell us that only the pure states corresponds to stationary points so they are the minimum ones for probability vectors corresponding to quantum states.

5.5 Rotations relating different SIC representations

In any finite dimension, a pair of distinct SICs will yield two different sets of probability vectors corresponding to quantum states. In this section, we will show that the probability vectors obtained from any two distinct SICs are related by an orthogonal transformation.

Suppose we have a pair of distinct SICs with elements Π'_i and Π_j, respectively. In the vector space of operators, Π'_i and Π_j correspond to the vertices of 2 identical regular simplices, which means there must exist a rotation between them. Formally, because a SIC forms a basis in the space of operators, we can write

$$\Pi'_i = \sum_k R_{ij} \Pi_j. \tag{5.53}$$

Since every SIC element has unit trace, taking the trace on both sides of Eq. (5.53) gives us

$$\sum_j R_{ij} = 1. \tag{5.54}$$

Multiplying the left-hand side of Eq.(5.53) by Π'_j and the right-hand side by $\sum_l R_{jl}\Pi_l$, we have

$$\begin{aligned} \mathrm{Tr}\left(\Pi'_i \Pi'_j\right) &= \sum_{k,l} R_{ik} R_{jl} \mathrm{Tr}\left(\Pi_k \Pi_l\right), \\ \frac{d\delta_{ij}+1}{d+1} &= \sum_{k,l} R_{ik} R_{jl} \left(\frac{d\delta_{kl}+1}{d+1}\right), \\ \delta_{ij} &= \sum_k R_{ik} R_{jk}. \end{aligned} \tag{5.55}$$

which is just the orthogonality condition.

We can also use Eq. (5.53) to calculate triple products:

$$\mathrm{Tr}\left(\Pi'_i \Pi'_j \Pi'_k\right) = \mathrm{Tr}\left(\sum_x R_{ix}\Pi_x \sum_y R_{jy}\Pi_y \sum_z R_{kz}\Pi'_z\right). \tag{5.56}$$

This tells us that

$$T'_{ijk} = \sum_{x,y,z} R_{ix} R_{jy} R_{kz} T_{x,y,z}. \tag{5.57}$$

Recall the cubic equation for pure states in Eq. (5.5). Suppose we are given two probability distributions $\vec{p}'$ and $\vec{p}$ representing the same quantum state but for two SICs with elements

Π'_i and Π_j, respectively. Using Eq. (5.57) we have

$$\begin{aligned}\frac{d+7}{(d+1)^3} &= \sum_{i,j,k} T'_{ijk} p'(i)p'(j)p'(k), \\ &= \sum_{i,j,k} \left(\sum_{x,y,z} R_{ix}R_{jy}R_{kz}T_{x,y,z} \right) p'(i)p'(j)p'(k), \\ &= \sum_{x,y,z} T_{x,y,z} p(x)p(y)p(z), \end{aligned} \tag{5.58}$$

which implies that

$$p'(i) = \sum_j R_{ij} p(j). \tag{5.59}$$

Thus, the rotation between two probability vectors representing the same quantum state for two distinct SICs is the same rotation that relates the SICs in Eq. (5.53). Because the pure states are the extreme points, by linearity, the same condition holds for their convex hull, that is, all other quantum states.

Eq. (5.55) tells us that the sum of every row of the matrix R_{ij} is one, as it should be if $R\vec{p}$ is to be a valid probability distribution. We also expect $\sum_i R_{ij} = 1$ since R must be unital, that is, it should preserve the uniform distribution, which corresponds to the maximally mixed state for any SIC.

Observe that Eq. (5.53) tells us exactly how to construct the orthogonal transformation between the probability spaces generated by distinct SICs: the expansion coefficients that we get when we express one SIC in terms of the other form the elements of the desired orthogonal matrix. Furthermore, the resulting orthogonal matrix is such that it maps between probability vectors representing the same quantum state. In Sec. 6.2, we will use this to explicitly construct the rotation matrix between the probability vectors of two distinct qutrit SICs.

5.6 Fidelity of quantum states

In quantum information, the fidelity of quantum states gives a measure for how similar two quantum states can be. For a pair of pure states $|\psi\rangle$ and $|\phi\rangle$, the fidelity $F(|\psi\rangle, |\phi\rangle)$ is given by

$$F(|\psi\rangle, |\phi\rangle) = |\langle\phi|\psi\rangle|. \tag{5.60}$$

In general, given density matrices ρ and σ,

$$F(\rho, \sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right) \tag{5.61}$$

where $M = \sqrt{\rho}$ is the unique square root matrix obtained from the spectral decomposition of ρ such that $M^2 = \rho$. This definition can be thought of as the quantum generalization of the fidelity of two probability vectors $\vec{p}$ and $\vec{q}$:

$$F(\vec{p}, \vec{q}) = \sum_i \sqrt{p(i)q(i)}. \tag{5.62}$$

Now consider two density operators ρ and σ expressed in terms of the same SIC $\{\Pi_i\}_{i=1}^{d^2}$:

$$\rho = \sum_{r=1}^{d^2} \left[(d+1)p_r - \frac{1}{d}\right] \Pi_r, \qquad \sigma = \sum_{s=1}^{d^2} \left[(d+1)q_s - \frac{1}{d}\right] \Pi_s. \tag{5.63}$$

The commutator $[\rho, \sigma]$ is given by

$$\begin{aligned}[\rho, \sigma] &= (d+1)^2 \sum_{r,s} p_r q_s \left(\Pi_r \Pi_s - \Pi_s \Pi_r\right) \\ &= (d+1)^2 \sum_{r,s} p_r q_s \sum_t J_{rst} \Pi_t \end{aligned} \tag{5.64}$$

where J_{rst} are the structure constants of the Lie algebra $\mathfrak{gl}_d(\mathbb{C})$ in the basis defined by the SIC that we saw in Sec. 3.7. So ρ and σ commute when

$$\sum_{r,s,t} p_r q_s J_{rst} \Pi_t = 0. \tag{5.65}$$

Because a SIC is a linearly independent set, this implies that for all t,

$$\sum_{r,s} p_r J_{rst} q_s = \vec{p} \cdot J_t \vec{q} = 0, \tag{5.66}$$

where J_t is the matrix with elements $(J_t)_{rs} = J_{rst}$. For probability vectors $\vec{p}$ and $\vec{q}$ that obey Eq. (5.66), the fidelity of the corresponding quantum states is given by the classical formula in Eq. (5.62) but with the probability vectors replaced by the vector of eigenvalues for ρ and σ.

To obtain the fidelity for SIC probabilities in the general case, it is useful to consider first an example involving Bloch vectors. Suppose we have qubit states ρ_1, ρ_2 with Bloch representations

$$\rho_j = \frac{1}{2}\left(I + \vec{r_j} \cdot \vec{\sigma}\right) \qquad (j = 1, 2), \tag{5.67}$$

where $\vec{\sigma}$ is the vector of Pauli operators, with components

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.68}$$

It can be shown that the fidelity in this case is given by [46]

$$F(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}\left(1 + \vec{r}_1 \cdot \vec{r}_2 + \sqrt{1 - r_1^2}\sqrt{1 - r_2^2}\right)^{\frac{1}{2}} \tag{5.69}$$

where $r_i = \|\vec{r}_i\|$. We would like to find a similar expression when quantum states are expressed in terms of SICs.

Given a 2-dimensional SIC $\{\Pi_i\}_{i=1}^4$, the density operator for a qubit can be written as

$$\rho_i = 3\sum_j p_i(j)\Pi_j - I \tag{5.70}$$

where $p_i(j) = \frac{1}{2}\mathrm{Tr}\left(\rho_i \Pi_j\right)$. For a single density operator ρ, we can relate its Bloch vector $\vec{r}$ to a SIC probability vector $\vec{p}$ using

$$r(i) = \mathrm{Tr}\left(\rho\sigma_i\right) = 3\sum_j p(j)\mathrm{Tr}\left(\sigma_i \Pi_j\right) \tag{5.71}$$

where we used $\mathrm{Tr}\left(\sigma_i\right) = 0$ for Pauli matrices.

For concreteness, let us choose the qubit SIC with the following projectors:

$$\Pi_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Pi_2 = \frac{1}{3}\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix},$$
$$\Pi_3 = \frac{1}{3}\begin{pmatrix} 1 & \omega^2\sqrt{2} \\ \omega\sqrt{2} & 2 \end{pmatrix}, \qquad \Pi_4 = \frac{1}{3}\begin{pmatrix} 1 & \omega\sqrt{2} \\ \omega^2\sqrt{2} & 2 \end{pmatrix}. \tag{5.72}$$

Using Eq. (5.72) and Eq.(5.68), we calculate the Bloch vector components to be

$$r(1) = \sqrt{2}[2p(2) - p(3) - p(4)], \tag{5.73}$$
$$r(2) = \sqrt{6}[p(3) - p(4)], \tag{5.74}$$
$$r(3) = 3p(1) - p(2) - p(3) - p(4). \tag{5.75}$$

Thus, we have

$$r^2 = \vec{r} \cdot \vec{r} = 9 \sum_i p(i)^2 - 3 \sum_{i \neq j} p(i)p(j) = 12 \sum_i p(i)^2 - 3. \tag{5.76}$$

If the Bloch vector $\vec{r}_i$ corresponds to the SIC vector $\vec{p}_i$ with probabilities $p_i(j)$, we can also calculate $\vec{r}_1 \cdot \vec{r}_2$:

$$\vec{r}_1 \cdot \vec{r}_2 = 9 \sum_i p_1(i)p_2(i) - 3 \sum_{i \neq j} p_1(i)p_2(j) = 12 \sum_i p_1(i)p_2(i) - 3. \tag{5.77}$$

Therefore, the fidelity for SIC probability vectors is obtained by substituting Eq. (5.77) and Eq. (5.76) for both $\vec{r}_1$ and $\vec{r}_2$ into Eq. (5.69). This gives us

$$F(\vec{p}_1, \vec{p}_2) = \frac{1}{\sqrt{2}} \left(12\vec{p}_1 \cdot \vec{p}_2 - 2 + 4\sqrt{1 - 3p_1^2}\sqrt{1 - 3p_2^2} \right)^{\frac{1}{2}} \tag{5.78}$$

where $p_i^2 = \vec{p}_i \cdot \vec{p}_i$.

If we examine the possible d-dependence of the fidelity function for qubits, a good guess of the general form would be

$$F(\vec{p}_1, \vec{p}_2) = \frac{1}{\sqrt{2}} \left(2d(d+1)\vec{p}_1 \cdot \vec{p}_2 - 2 + 4\sqrt{1 - \frac{d(d+1)}{2}p_1^2}\sqrt{1 - \frac{d(d+1)}{2}p_2^2} \right)^{\frac{1}{2}}. \tag{5.79}$$

In particular, Eq. (5.79) gives the correct values for 2 special cases: (i) when $\rho_1 = \rho_2$ and (ii) when ρ_1 is orthogonal to ρ_2 for any two pure states ρ_1, ρ_2. It also turns out that Mendonca et al. [98] have proposed an alternative fidelity function F_N,

$$F_N(\rho_1, \rho_2) = \mathrm{Tr}\,(\rho_1\rho_2) - \sqrt{1 - \mathrm{Tr}\,(\rho_1^2)}\sqrt{1 - \mathrm{Tr}\,(\rho_2^2)}, \tag{5.80}$$

which when converted into SIC terms yields the same expression as the square of Eq.(5.79).

5.7 Purification of mixed states

In quantum mechanics, any mixed state can be thought of as the reduced state of some pure state on a larger Hilbert space. Formally, consider a density operator ρ on $\mathcal{H}_A$. Then there exists a Hilbert space $\mathcal{H}_B$ and a pure state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ such that

$$\mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \rho. \tag{5.81}$$

We say that $|\psi\rangle$ is a purification of ρ. In general, a square matrix A is positive semidefinite if and only if it can be purified in such a way. In this section, we examine what the SIC representation tells us about purifications of mixed states.

To start, consider a mixed state σ on a Hilbert space $\mathcal{H}^d$. Expressed in terms of a SIC denoted by $\{\Lambda\}_{i=1}^{d^2}$, we have

$$\sigma = (d+1)\sum_{i=1}^{d^2} q_i \Lambda_i - I_d. \tag{5.82}$$

Note that here we change the notation to Λ_i since we want to reserve Π_i for the SIC acting on the composite space $\mathcal{H}_A \otimes \mathcal{H}_B$, where the purification is located.

Suppose that the structure coefficients β_{ijk} for the SIC Λ_i are given by

$$\Lambda_i \Lambda_j = \sum_k \beta_{ijk} \Lambda_k. \tag{5.83}$$

The trace conditions for a mixed state tells us that

$$\frac{1}{d} \leq \mathrm{Tr}\left(\sigma^2\right) < 1, \qquad \frac{1}{d^2} \leq \mathrm{Tr}\left(\sigma^3\right) < 1, \tag{5.84}$$

which imply that the probability distribution $\vec{q}$ must satisfy the conditions

$$\frac{1}{d^2} \leq \sum_{i=1}^{d^2} q_i^2 < \frac{2}{d(d+1)}, \qquad \frac{1}{d^3} \leq \sum_{i,j,k} \beta_{ijk} q_i q_j q_k < \frac{4}{d(d+1)^2} \tag{5.85}$$

since

$$\begin{aligned} \mathrm{Tr}\left(\sigma^2\right) &= d(d+1)\sum_i q_i^2 - 1, \\ \mathrm{Tr}\left(\sigma^3\right) &= d(d+1)^2 \sum_{i,j,k} \beta_{ijk} q_i q_j q_k - 2d(d+1)\sum_i q_i^2 + 1. \end{aligned} \tag{5.86}$$

There exists a purification for any mixed state of dimension d by considering it as the reduction of a pure state $|\psi\rangle$ of dimension d^2. More specifically, for any mixed state with rank $r \leq d$, you can find a purification in a Hilbert space of dimension at least r^2. This follows from the fact that if we consider σ as the state of system A that is part of a composite system AB with state ρ. Taking the diagonal form of σ,

$$\sigma = \sum_{i=1}^{r} p_i |e_i\rangle\langle e_i| \tag{5.87}$$

where r is the rank of σ and $\{e_i\}_{i=1}^r \subset \mathcal{H}_A$ is some orthonormal set, then given any orthonormal set $\{f_i\}_{i=1}^r \in \mathcal{H}_B$, then $\rho = |\psi\rangle\langle\psi|$ with

$$|\psi\rangle = \sum_{i=1}^{r} \sqrt{p_i}\, |e_i\rangle \otimes |f_i\rangle \tag{5.88}$$

gives a purification for σ since

$$\mathrm{Tr}_B(\sigma) = \sum_{i,j} \sqrt{p_i p_j} \mathrm{Tr}_B \left(|e_i\rangle\langle e_j| \otimes |f_i\rangle\langle f_j| \right) = \sum_{i,j} \sqrt{p_i p_j} \delta_{ij} |e_i\rangle\langle e_j| = \sigma. \tag{5.89}$$

Treating the partial trace as a quantum operation on ρ, we can construct an operator-sum representation for it with Kraus operators E_i such that

$$\sum_i E_i \rho E_i^\dagger = \sigma. \tag{5.90}$$

Let $\rho \in \mathcal{L}(\mathcal{H}_{d^2})$ be a purification of σ. Writing

$$\rho = (d^2+1) \sum_{i=1}^{d^4} p_i \Pi_i - I_{d^2}, \tag{5.91}$$

for some particular choice of a SIC $\{\Pi_i\}_{i=1}^{d^4}$ with dimension d^2, with structure coefficients α_{ijk} given by

$$\Pi_i \Pi_j = \sum_k \alpha_{ijk} \Pi_k. \tag{5.92}$$

Thinking of ρ as an operator on $\mathcal{H}_d \otimes \mathcal{H}_d$, ρ is a purification of σ if we there is a probability vector $\vec{p}$ such that

$$\begin{aligned} \sum_{i=1}^{d^4} p_i^2 &= \frac{2}{d^2(d^2+1)}, \\ \sum_{i,j,k} \alpha_{ijk} p_i p_j p_k &= \frac{4}{d^2(d^2+1)}. \end{aligned} \tag{5.93}$$

It is then straightforward to work out how $\vec{q}$ is related to $\vec{p}$ from $\mathrm{Tr}_B(\rho) = \sigma$:

$$(d^2+1)\sum_{i,j} p_i E_j \Pi_i E_j^\dagger - \sum_j E_j E_j^\dagger = (d+1)\sum_i q_i \Lambda_i - I. \tag{5.94}$$

But we know that the partial trace gives us

$$\mathrm{Tr}_B(I_{d^2}) = dI_d \tag{5.95}$$

so this implies that

$$\sum_j E_j E_j^\dagger = dI_d. \tag{5.96}$$

Thus, if we plug this into Eq. (5.94), multiply throughout by Λ_k and take the trace, we obtain

$$(d^2+1)\sum_{i,j} p_i \mathrm{Tr}\left(E_j \Pi_i E_j^\dagger \Lambda_k\right) = (d+1)\sum_i q_i \mathrm{Tr}(\Lambda_i \Lambda_k) + (d-1)\mathrm{Tr}(\Lambda_k). \tag{5.97}$$

Usually we want $\vec{p}$ in terms of $\vec{q}$ but if we solve for $\vec{q}$ instead, we get the somewhat simpler expression

$$q_k = \frac{(d^2+1)}{d}\sum_{i,j} p_i \mathrm{Tr}\left(E_j \Pi_i E_j^\dagger \Lambda_k\right) - 1. \tag{5.98}$$

One way to verify the above result is to check, for instance, that $\sum_{k=1}^{d^2} q_k = 1$:

$$\begin{aligned}
\sum_k q_k =& \frac{(d^2+1)}{d}\sum_{i,j} p_i \mathrm{Tr}\left[E_j \Pi_i E_j^\dagger \left(\sum_k \Lambda_k\right)\right] - \sum_k (1), \\
=& \frac{(d^2+1)}{d}\sum_{i,j} p_i \mathrm{Tr}\left(E_j \Pi_i E_j^\dagger \sum_k dI_d\right) - d^2, \\
=& (d^2+1)\sum_i p_i \mathrm{Tr}(\Pi_i I_{d^2}) - d^2, \\
=& (d^2+1)\sum_i p_i - d^2 = 1
\end{aligned} \tag{5.99}$$

Thus, to look for the purification $\vec{p}$ of some mixed state $\vec{q}$, we solve Eq. (5.98) for $\vec{p}$ subject to the constraints of Eq. (5.93). Note that because there is a unitary freedom in the auxiliary system, there are many pure states $\vec{p}$ associated with each $\vec{q}$.

For the purposes of illustration, let us consider the example of a qubit mixed state σ. In the SIC representation, the density operator for a qubit $\sigma \in \mathcal{B}(\mathcal{H}^2)$ is given by

$$\sigma = 3\sum_{i=1}^{4} q_i \Lambda_i - I_2, \tag{5.100}$$

where a conventional choice for the SIC operators would be

$$\Lambda_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Lambda_2 = \frac{1}{3}\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}, \tag{5.101}$$
$$\Lambda_3 = \frac{1}{3}\begin{pmatrix} 1 & \omega\sqrt{2} \\ \omega^2\sqrt{2} & 2 \end{pmatrix}, \qquad \Lambda_4 = \frac{1}{3}\begin{pmatrix} 1 & \omega^2\sqrt{2} \\ \omega\sqrt{2} & 2 \end{pmatrix}$$

where $\omega = e^{2i\pi/3}$. We can find a purification for σ by considering it as the reduced state of a 4-dimensional pure state $\rho = |\psi\rangle\langle\psi|$. The density operator for ρ can be written in terms of ququart SICs as follows:

$$\rho = 5\sum_{i=1}^{16} p_i \Pi_i - I_4, \tag{5.102}$$

where for specificity we choose the ququart Weyl-Heisenberg SIC with fiducial vector

$$|\psi\rangle = \frac{1}{2\sqrt{3+\phi}} \begin{pmatrix} 1 + e^{-\frac{i\pi}{4}} \\ e^{-\frac{i\pi}{4}} + i\phi^{-3/2} \\ 1 - e^{-\frac{i\pi}{4}} \\ e^{-\frac{i\pi}{4}} - i\phi^{-3/2} \end{pmatrix} \tag{5.103}$$

where ϕ is the golden ratio $\phi = (\sqrt{5}-1)/2$. We denote the structure coefficients of this SIC by α_{ijk}. Treating ρ as a state for $\mathcal{H}_2 \otimes \mathcal{H}_2$, it is a purification of σ if we can find some $\vec{p}$ so that

$$\sum_{i=1}^{16} p_i^2 = \frac{1}{10} \qquad \text{and} \qquad \sum_{i,j,k} \alpha_{ijk} p_i p_j p_k = \frac{1}{25}, \tag{5.104}$$

and of course, $\mathrm{Tr}_B(\rho) = \sigma$.

The partial trace is a straightforward calculation that yields how $\vec{q}$ should be related to the desired purification $\vec{p}$. However, to get the simplest relation possible, we present the following method.

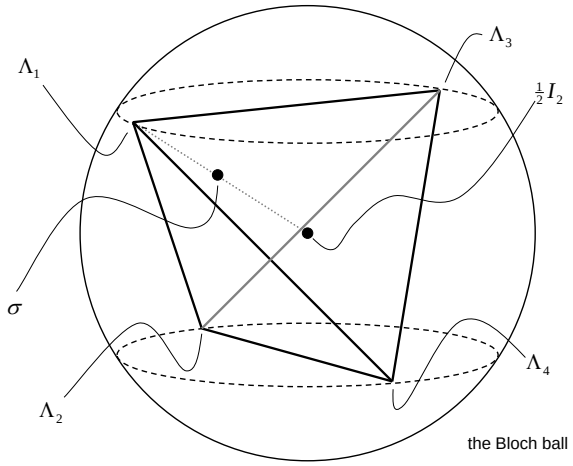


Figure 5.1: Purifying qubit mixed states in the SIC representation. In solving for the purification of σ, it is best to use a qubit SIC with an element 'parallel' to σ.

In a purification problem, we start with a qubit mixed state σ. For any such state, there exists a unitary operator U such that

$$U\sigma U^\dagger = \tilde{\sigma}, \tag{5.105}$$

where $\tilde{\sigma}$ is a diagonal matrix with eigenvalue entries μ and $1-\mu$. For conceptual purposes, one can think of $\tilde{\sigma}$ as representing the same state but using the SIC with elements

$$\tilde{\Lambda}_i = U^\dagger \Lambda_i U. \tag{5.106}$$

Alternatively, one can think of $\tilde{\sigma}$ as the same state as σ but expressed in its eigenbasis, which is what we actually use during the computation. Thus, we have

$$\tilde{\sigma} = 3\sum_i q_i \tilde{\Lambda}_i - I_2. \tag{5.107}$$

If we think of the Bloch sphere, $\{\tilde{\Lambda}_i\}$ corresponds to choosing the SIC with one of the projectors aligned to the Bloch vector for σ, as depicted in Fig. 5.1.

Why is this a better choice? Because when we diagonalize σ, we get

$$q_1 = \frac{\mu}{2}, \qquad q_2 = q_3 = q_4 = \frac{1}{3}\left(1 - \frac{\mu}{2}\right) \tag{5.108}$$

where μ and $1-\mu$ are the eigenvalues of σ. To be consistent with $\{\tilde{\Lambda}_i\}$, we should define the purification $\tilde{\rho}$ in terms of the ququart SIC $\tilde{\Pi}_i = V^\dagger \Pi_i V$,

$$\tilde{\rho} = 5 \sum_{i=1}^{16} p_i \tilde{\Pi}_i - I_4. \tag{5.109}$$

where $V = U \otimes I_2$.

One systematic way to solve this would be to let

$$\begin{aligned} v_1 &= p_1 + p_3 - p_6 - p_8, \\ v_2 &= p_5 + p_7 - p_2 - p_4, \\ v_3 &= p_9 + p_{11} - p_{13} - p_{15}, \\ v_4 &= p_{14} + p_{16} - p_{10} - p_{12}. \end{aligned} \tag{5.110}$$

Using the partial trace equations for this particular choice of SICs, the conditions for obtaining the purification $\vec{p}$ are

$$\begin{aligned} v_1 &= -v_2, \\ v_4 &= -v_3, \\ v_1 - v_4 &= \frac{(-5+3\sqrt{5})(4\mu-2)}{10(\sqrt{10}-2\sqrt{2}) - 5\left(-1+\sqrt{5}\right)^{3/2}} \end{aligned} \tag{5.111}$$

subject to constraints in Eq. (5.104). Because the structure coefficients α_{ijk} are quite non-trivial, this appears to be the best simplification we can achieve.

5.8 Entanglement in 2-qubit systems

Entanglement is considered to be one of the defining characteristics of quantum phenomena. Parts of an entangled system exhibit strong correlations that cannot be achieved by any classical system and these correlations can be observed even when the parts are separated at arbitrarily large distances. A 1935 paper by Albert Einstein, Boris Podolsky, and Nathan Rosen is often cited as the originator of the idea of quantum entanglement, where they used it in an argument that was supposed to demonstrate the inadequacy of quantum mechanics, known widely as the EPR paradox. Here we examine the entanglement of 2-qubit systems in the SIC representation.

Consider two quantum systems A and B with states on Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. The composite system AB has states on the tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, which just means that if $\{|a_i\rangle\}$ is an orthonormal basis for $\mathcal{H}_A$ and $\{|b_j\rangle\}$ is an orthonormal basis for $\mathcal{H}_B$, then the vectors $|a_i, b_j\rangle \equiv |a_i\rangle \otimes |b_j\rangle$ constitute an orthonormal basis for $\mathcal{H}_{AB}$.

We say that $|\psi\rangle$ is a separable state of $\mathcal{H}_{AB}$ if there exists $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$ such that

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \,. \tag{5.112}$$

Otherwise, the state is entangled. More generally, a density operator ρ_{AB} on $\mathcal{H}_{AB}$ is deemed separable if it can be expressed as

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i \tag{5.113}$$

where ρ_A^i and ρ_B^i are states on subsystems A and B, respectively. In other words, ρ_{AB} is separable if it is probability distribution over product states for the composite system. A quantum state is said to be entangled if it is not separable in this more general sense.

In quantum information theory, an entanglement witness is an geometric object that can be evaluated for quantum states and whose values distinguish an entangled state from separable ones. Deciding whether an arbitrary quantum state is separable or not is a highly nontrivial problem especially when one extends to multipartite case. However, we will just focus on a specific result for bipartite systems due to Asher Peres, and the Horodecki trio of Michał, Paweł, and Ryzard.

The Peres-Horodecki criterion [88, 105] gives a necessary condition for the separability of a joint density operator $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ that depends only on the positivity of its partial transpose. In the case where systems A and B are both qubits or one is a qubit while the other is a qutrit, then it is also a sufficient condition. For those 2 situations, it provides a definitive test for entanglement in mixed states.

The theorem is stated as follows: For any state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$,

$$\rho = \sum_{i,j,k,l} \omega_{ijkl} \, |i\rangle \langle j| \otimes |k\rangle \langle l| \tag{5.114}$$

the partial transpose with respect to system B is defined by

$$\rho^{T_B} = \sum_{i,j,k,l} \omega_{ijkl} \, |i\rangle \langle j| \otimes (|k\rangle \langle l|)^T = \sum_{i,j,k,l} \omega_{ijkl} \, |i\rangle \langle j| \otimes |l\rangle \langle k|. \tag{5.115}$$

The partial transpose with respect to system A is defined similarly. If ρ is separable then ρ^{T_B} is positive definite (same for ρ^{T_A}). Thus, if ρ^{T_B} has negative eigenvalues, ρ is necessarily entangled.

It is instructive to consider two examples involving 2-qubit states. The first example involves density operators of the form

$$\rho = q|\psi_1\rangle\langle\psi_1| + (1-q)|\psi_2\rangle\langle\psi_2| \tag{5.116}$$

where

$$\begin{aligned} |\psi_1\rangle &= a\,|e_1\rangle \otimes |e_1\rangle + b\,|e_2\rangle \otimes |e_2\rangle\,, \\ |\psi_2\rangle &= a\,|e_1\rangle \otimes |e_2\rangle + b\,|e_2\rangle \otimes |e_1\rangle\,, \end{aligned} \tag{5.117}$$

for $a, b > 0$ and some orthonormal basis $\{|e_1\rangle\,, |e_2\rangle\} \subset \mathbb{C}^2$. Such states have been considered in the study of Bell inequalities.

We can write out ρ and its partial transpose in matrix form using the basis $\{|e_i\rangle \otimes |e_j\rangle\}$.

$$\rho = \begin{pmatrix} qa^2 & 0 & 0 & qab \\ 0 & (1-q)a^2 & (1-q)ab & 0 \\ 0 & (1-q)ab & (1-q)b^2 & 0 \\ qab & 0 & 0 & qb^2 \end{pmatrix}, \tag{5.118}$$

$$\rho^{T_B} = \begin{pmatrix} qa^2 & 0 & 0 & (1-q)ab \\ 0 & (1-q)a^2 & qab & 0 \\ 0 & qab & (1-q)b^2 & 0 \\ (1-q)ab & 0 & 0 & qb^2 \end{pmatrix}. \tag{5.119}$$

To check for the positivity of ρ^{T_B}, we only need to consider two principal minors or determinants, since for Hermitian operators, positive semi-definiteness is guaranteed if all the principal minors are non-negative:

$$D_1 = \begin{vmatrix} qa^2 & (1-q)ab \\ (1-q)ab & qb^2 \end{vmatrix}, \qquad D_2 = \begin{vmatrix} (1-q)a^2 & qab \\ qab & (1-q)b^2 \end{vmatrix} = -D_1. \tag{5.120}$$

We can easily see that

$$D_1 = a^2b^2[q^2 - (1-q)^2] = a^2b^2(2q-1). \tag{5.121}$$

Thus, for $ab \neq 0, q \neq \frac{1}{2}$, either D_1 or D_2 is negative. When $q = \frac{1}{2}$, $\rho^{T_B} = \rho$.

What would the criterion look like when states are represented using SICs? In terms of a ququart SIC $\{\Pi_i\}_{i=1}^{16}$,

$$\rho = \sum_i \left(5p(i) - \frac{1}{4}\right) \Pi_i. \tag{5.122}$$

Substituting the ρ in Eq. (5.118), we find that

$$\sum_i p(i)^2 = \frac{1}{10}\left(1 - q + q^2\right) \tag{5.123}$$

in which case $q = \frac{1}{2}$ corresponds to a right-hand side of $\frac{3}{40}$. Therefore, for this particular family of states, the SIC probability distributions $p(i)$ such that

$$\frac{3}{40} < \sum_i p(i)^2 \leq \frac{1}{10} \tag{5.124}$$

are associated with entangled 2-qubit states.

The other example we consider are Werner states

$$\rho = q|\Psi^-\rangle\langle\Psi^-| + \left(\frac{1-q}{4}\right) I_4 \tag{5.125}$$

where $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state. These states can be thought of as mixtures of a Bell state with the completely mixed state.

In the computational basis, ρ and its partial transpose are given by

$$\begin{aligned}
\rho &= \frac{1}{4}\begin{pmatrix} 1-q & 0 & 0 & 0 \\ 0 & 1+q & -2q & 0 \\ 0 & -2q & 1+q & 0 \\ 0 & 0 & 0 & 1-q \end{pmatrix}, \\
\rho^{T_B} &= \frac{1}{4}\begin{pmatrix} 1-q & 0 & 0 & -2q \\ 0 & 1+q & 0 & 0 \\ 0 & 0 & 1+q & 0 \\ -2q & 0 & 0 & 1-q \end{pmatrix}.
\end{aligned} \tag{5.126}$$

It is straightforward to calculate the eigenvalues of ρ^{T_B}, which are

$$\frac{1}{4}(1-3q), \frac{1+q}{4}, \frac{1+q}{4}, \frac{1+q}{4}. \tag{5.127}$$

The first one is the smallest one for any $0 \leq q \leq 1$ and it is non-negative only when $q \leq \frac{1}{3}$. Thus, ρ of Eq. (5.126) is entangled for $q > \frac{1}{3}$. Expressing the Werner state in Eq. (5.125) in the SIC representation, we get

$$\sum_i p(i)^2 = \frac{1}{80}\left(5+3q^2\right). \tag{5.128}$$

Hence, the Werner state ρ is entangled with respect to the product basis for two qubits when

$$\frac{1}{15} < \sum_i p(i)^2 \leq \frac{1}{10}. \tag{5.129}$$

In general, we can consider the linear operator

$$H = (d+1)\sum_i p_i \Pi_i - I, \tag{5.130}$$

with $d = 4$ and the set $\{\Pi_i\}_{i=1}^{16}$ is a ququart SIC. When $\vec{p}$ is a probability vector, H is necessarily a unit trace Hermitian operator. To be a density operator, H must also be positive semidefinite, which can be enforced using a lemma first stated by Gen Kimura [78]:

Lemma 5.8.1. Consider the monic polynomial of degree $d \geq 1$

$$f(x) = \sum_{j=0}^{d} (-1)^j a_j x^{d-j} = \prod_{i=1}^{d} (x - x_i) \tag{5.131}$$

where $a_0 = 1$ and which has only real roots $x_i \in \mathbb{R}, i = 1, 2, ..., d$ when $f(x) = 0$. Then for all the roots x_i to be positive semidefinite, all coefficients a_i must be positive semidefinite, and vice-versa:

$$\forall i, \quad x_i \geq 0 \iff a_i \geq 0. \tag{5.132}$$

Applying the lemma to the characteristic polynomial of H, we get the desired positivity constraints using Newton-Girard identities that relate coefficients of a monic polynomial

with the sums of powers of its roots [72]:

$$a_1 = \mathrm{Tr}\left(H\right) = 1, \tag{5.133}$$

$$a_2 = \frac{1}{2}\left(1 - \mathrm{Tr}\left(H^2\right)\right) \geq 0, \tag{5.134}$$

$$a_3 = \frac{1}{6}\left(1 - 3\mathrm{Tr}\left(H^2\right) + 2\mathrm{Tr}\left(H^3\right)\right) \geq 0, \tag{5.135}$$

$$a_4 = \frac{1}{24}\left(1 - 6\mathrm{Tr}\left(H^4\right) + 8\mathrm{Tr}\left(H^3\right) - 6\mathrm{Tr}\left(H^2\right) + 3\mathrm{Tr}\left(H^2\right)^2\right) \geq 0. \tag{5.136}$$

where the coefficients a_j refer to the characteristic polynomial of H, defined to be consistent with Eq. (5.131).

In terms of SICs, it is easy to verify that

$$\begin{aligned}
\mathrm{Tr}\left(H^2\right) &= d(d+1)\sum_i p_i^2 - 1, \\
\mathrm{Tr}\left(H^3\right) &= d(d+1)^2 \sum_{i,j,k} S_{ijk} p_i p_j p_k - 2d(d+1)\sum_i p_i^2 + 1, \\
\mathrm{Tr}\left(H^4\right) &= d(d+1)^3 \sum_{i,j,k,l,m} S_{ijm} S_{klm} p_i p_j p_k p_l - 4d(d+1)^2 \sum_{i,j,k} S_{ijk} p_i p_j p_k \\
&\quad + d^2(d+1)\left(\sum_i p_i^2\right)^2 + 4d(d+1)\sum_i p_i^2 - 1,
\end{aligned} \tag{5.137}$$

where S_{ijk} are the structure coefficients for the ququart SIC, which leads to the following positivity conditions on H:

$$\begin{aligned}
a_1 &= 1, \\
a_2 &= 1 - \frac{d(d+1)}{2}\sum_i p_i^2 \geq 0, \\
a_3 &= \frac{1}{6}\left[6 - 7d(d+1)\sum_i p_i^2 + 2d(d+1)^2 \sum_{i,j,k} S_{ijk} p_i p_j p_k\right] \geq 0. \\
a_4 &= \frac{1}{24}\Bigg(24 - 6d(d+1)^3 \sum_{i,j,k,l,m} S_{ijm} S_{klm} p_i p_j p_k p_l + 32d(d+1)^2 \sum_{i,j,k} S_{ijk} p_i p_j p_k \\
&\quad + (3(d+1) - 6)d^2(d+1)\left(\sum_i p_i^2\right)^2 - 52d(d+1)\sum_i p_i^2\Bigg) \geq 0.
\end{aligned} \tag{5.138}$$

Since this result works for any hermitian operator, it can also be applied to the partial transpose of H, in which case the Peres-Horodecki criterion tells us that H corresponds to a separable state if and only if the coefficients a_i for H^{T_B} are all non-negative.

5.9 Quantum operations as affine maps on SIC probabilities

An important topic in the study of quantum information processing involves the study of quantum operations, particularly maps of density operators that describe open-system evolution. Studying such maps is useful for various applications such as developing methods in quantum control or for identifying new witnesses for quantum entanglement. [91], Mark Byrd et al., provide a description of general maps of density operators in terms of affine maps of the corresponding Bloch vector representation of quantum states. In this section, we provide a similar description using the SIC representation.

Let the Kraus decomposition for a general (completely positive) map between density operators ρ, ρ' of the same dimension be given by

$$\Phi(\rho) = \sum_{i=1}^{r} A_i \rho A_i^\dagger = \rho' \tag{5.139}$$

where A_i are the Kraus operators. Note that if the map is trace preserving then

$$\sum_{i=1}^{r} A_i^\dagger A_i = I. \tag{5.140}$$

What we want is to get the corresponding description in terms of a linear or affine map on SIC probability vectors. In this case, we write density operators in terms of the SIC Π_i,

$$\rho = \sum_{i=1}^{d^2} \left[(d+1)p_i - \frac{1}{d} \right] \Pi_i. \tag{5.141}$$

Because SICs form a basis for the space of operators, we can do the same for the Kraus operators

$$A_i = \sum_{j=1}^{d^2} a_{ij} \Pi_j. \tag{5.142}$$

Two special classes of maps to consider are unital and trace-preserving maps. If Φ is unital then $\Phi(I) = I$. Translating in terms of SICs,

$$\sum_{i=1}^{r} A_i A_i^\dagger = \sum_{i=1}^{r} \sum_{j,k=1}^{d^2} a_{ij} a_{ik}^* \Pi_j \Pi_k$$
$$= \sum_{i=1}^{r} \sum_{j,k,l=1}^{d^2} a_{ij} a_{ik}^* S_{jkl} \Pi_l = I \tag{5.143}$$

where we used the structure coefficients,

$$\Pi_j \Pi_k = \sum_{l=1}^{d^2} S_{jkl} \Pi_l. \tag{5.144}$$

Because $\sum_{l=1}^{d^2} \Pi_l = dI$ for SICs, we obtain that for any l

$$\sum_{i=1}^{r} \sum_{j,k=1}^{d^2} a_{ij} a_{ik}^* S_{jkl} = \frac{1}{d}. \tag{5.145}$$

Observe that the trace-preserving condition in Eq. (5.140) reads

$$\sum_{i=1}^{r} \sum_{j,k=1}^{d^2} a_{ij}^* a_{ik} S_{jkl} = \frac{1}{d}. \tag{5.146}$$

Moreover, from $\mathrm{Tr}\left(\Phi(\rho)\right) = \mathrm{Tr}\left(\rho\right) = 1$, we get

$$\sum_{i=1}^{r} \sum_{j,k,l=1}^{d^2} a_{ij} q_k a_{il}^* T_{jkl} = 1 \tag{5.147}$$

where we used the triple products

$$T_{jkl} = \mathrm{Tr}\left(\Pi_j \Pi_k \Pi_l\right) \tag{5.148}$$

and

$$q_k = (d+1) p_k - \frac{1}{d}. \tag{5.149}$$

In general,

$$\Phi(\rho) = \sum_{i=1}^{r} \sum_{j,k,l} a_{ij} q_k a_{il}^* \Pi_j \Pi_k \Pi_l \tag{5.150}$$

Let $\Pi_j \Pi_k = \sum_m S_{jkm} \Pi_m$ and $\Pi_m \Pi_l = \sum_n S_{mln} \Pi_n$. Then

$$\Phi(\rho) = \sum_{i=1}^{r} \sum_{j,k,l=1}^{d^2} \sum_{m,n=1}^{d^2} a_{ij} q_k a_{il}^* S_{jkm} S_{mln} \Pi_n = \sum_{n=1}^{d^2} \left[\sum_{i=1}^{r} \sum_{j,k,l,m=1}^{d^2} a_{ij} q_k a_{il}^* S_{jkm} S_{mln} \right] \Pi_n. \tag{5.151}$$

Thus, Eq. (5.151) says that $\rho \mapsto \Phi(\rho)$ corresponds to the linear map

$$q_n \mapsto q_n' = \sum_{i=1}^{r} \sum_{j,k,l,m=1}^{d^2} a_{ij} q_k a_{il}^* S_{jkm} S_{mln}, \tag{5.152}$$

that is, each component of $\vec{q}$ given by Eq. (5.149) gets mapped to another vector given in terms of the expansion coefficients of the Kraus operators and the structure coefficients.

Writing it out explicitly, we have

$$q_n' = (d+1) \sum_{i=1}^{r} \sum_{j,k,l,m=1}^{d^2} a_{ij} p_k a_{il}^* S_{jkm} S_{mln} - \frac{1}{d} \sum_{i=1}^{r} \sum_{j,k,l,m=1}^{d^2} a_{ij} a_{il}^* S_{jkm} S_{mln}. \tag{5.153}$$

If the map in Eq. (5.152) is unital then we can write the second term above as

$$\frac{1}{d} \sum_{i=1}^{r} \sum_{j,l,m=1}^{d^2} a_{ij} a_{il}^* S_{mln} d\delta_{jm} = \sum_{i=1}^{r} \sum_{j,l=1}^{d^2} a_{ij} a_{il}^* S_{jln} = \frac{1}{d}. \tag{5.154}$$

where we used $\sum_k S_{jkm} = d\delta_{jm}$ and Eq. (5.145). This says that if the map is unital then the map applies to the probability vector $\vec{p}$ directly:

$$p_n \mapsto p_n' = \sum_{i=1}^{r} \sum_{j,k,l,m=1}^{d^2} a_{ij} p_k a_{il}^* S_{jkm} S_{mln}. \tag{5.155}$$

Of course, Eq. (5.145) will not be true for non-unital maps so generally we would get an affine map $\vec{p} \mapsto \vec{p}' = M\vec{p} + \vec{t}$ where

$$M_{jk} = \sum_{i=1}^{r} \sum_{l,m,n=1}^{d^2} a_{im} a_{in}^* S_{mkl} S_{lnj}, \qquad t_j = \frac{1}{d(d+1)} \left(1 - \sum_{k=1}^{d^2} M_{jk} \right). \tag{5.156}$$

The following are some examples of simple maps for density operators:

1. Trace-preserving inversion map:

$$\rho \mapsto \frac{1}{d-1}\left(I-\rho\right). \tag{5.157}$$

For pure states, this maps $|\psi\rangle$ to some state orthogonal to $|\psi\rangle$. We have

$$\begin{aligned}\rho &\mapsto \frac{1}{d-1}\left[I-(d+1)\sum_i p_i\Pi_i+I\right]\\ &=\sum_i\left[\frac{2}{d(d-1)}-\frac{d+1}{d-1}p_i\right]\Pi_i\\ &\equiv\sum_i\left[(d+1)p_i'-\frac{1}{d}\right]\Pi_i,\end{aligned} \tag{5.158}$$

which implies that

$$p_i'=\frac{1}{d-1}\left(\frac{1}{d}-p_i\right). \tag{5.159}$$

For instance, if we consider the SIC basis state

$$\vec{p}=\left(\frac{1}{d},\frac{1}{d(d+1)},\ldots,\frac{1}{d(d+1)}\right)^T, \tag{5.160}$$

that is, the probability vector for $\rho=\Pi_1$, then

$$\vec{p}'=\left(0,\frac{1}{d^2-1},\ldots,\frac{1}{d^2-1)}\right)^T. \tag{5.161}$$

This means that

$$\vec{p}\cdot\vec{p}'=\frac{1}{d(d+1)} \tag{5.162}$$

which is indeed the value of the scalar product for orthogonal pure states.

2. Completely positive inversion map (approximate universal NOT gate):

$$\rho \mapsto \frac{1}{d^2-1}\left(dI-\rho\right). \tag{5.163}$$

We have

$$\begin{aligned}\rho &\mapsto \frac{1}{d^2-1}\left[dI + I - (d+1)\sum_i p_i \Pi_i\right] \\ &= \sum_i \left[\frac{1}{d(d-1)} - \frac{p_i}{d-1}\right]\Pi_i \\ &\equiv \sum_i \left[(d+1)p_i' - \frac{1}{d}\right]\Pi_i, \end{aligned} \tag{5.164}$$

which implies that

$$p_i' = \frac{1}{d^2-1}\left(1 - p_i\right). \tag{5.165}$$

For instance, if we consider the SIC basis state in Eq. (5.160) then

$$\vec{p}' = \left(\frac{1}{d(d+1)}, \frac{d^2+d-1}{d(d-1)(d+1)^2}, \ldots, \frac{d^2+d-1}{d(d-1)(d+1)^2}\right)^T. \tag{5.166}$$

3. Trace-preserving projection onto a random (average) pure state:

$$\rho \mapsto \frac{1}{d+1}\left(I + \rho\right). \tag{5.167}$$

We have

$$\begin{aligned}\rho &\mapsto \frac{1}{d+1}\left[I + (d+1)\sum_i p_i \Pi_i - I\right] \\ &= \sum_i p_i \Pi_i \\ &\equiv \sum_i \left[(d+1)p_i' - \frac{1}{d}\right]\Pi_i, \end{aligned} \tag{5.168}$$

which implies that

$$p_i' = \frac{1}{d+1}\left(p_i + \frac{1}{d}\right). \tag{5.169}$$

For instance, if we consider the SIC basis state in Eq. (5.160) then

$$\vec{p}' = \left(\frac{2}{d(d+1)}, \frac{d+2}{d(d+1)^2}, \ldots, \frac{d+2}{d(d+1)^2}\right)^T. \tag{5.170}$$

4. Qudit depolarizing channel:

$$\rho \mapsto (1-\epsilon)\rho + \frac{\epsilon}{d} I. \tag{5.171}$$

We have

$$\begin{aligned}
\rho &\mapsto (1-\epsilon)\left[(d+1)\sum_i p_i \Pi_i - I\right] + \frac{\epsilon}{d} I \\
&= (1-\epsilon)\left[(d+1)\sum_i p_i \Pi_i\right] - \left(1-\epsilon-\frac{\epsilon}{d}\right) I \\
&= \left[(d+1)p_i(1-\epsilon) - \frac{1}{d}\left(1-\frac{(d+1)\epsilon}{d}\right)\right]\Pi_i \\
&\equiv \sum_i \left[(d+1)p_i' - \frac{1}{d}\right]\Pi_i,
\end{aligned} \tag{5.172}$$

which implies that

$$p_i' = (1-\epsilon)p_i + \frac{\epsilon}{d^2}. \tag{5.173}$$

For instance, if we consider the SIC basis state in Eq. (5.160) then

$$\vec{p}' = \left(\frac{d(1-\epsilon)+\epsilon}{d^2}, \frac{d+\epsilon}{d^2(d+1)}, \ldots, \frac{d+\epsilon}{d^2(d+1)}\right)^T. \tag{5.174}$$

5. Qudit flip channel:

$$\rho \mapsto \sum_{i=0}^{d-1} \epsilon_i X^i \rho X^{i\dagger} \tag{5.175}$$

where ϵ_i represents the probability of $|j\rangle \mapsto |j \oplus i\rangle$ for the orthonormal basis $\{|j\rangle\}_{j=0}^{d-1}$ that is shifted by X,

$$X|j\rangle = |j \oplus 1\rangle \tag{5.176}$$

with $i \oplus j = i + j \mod d$ and later, $i \ominus j = i - j \mod d$. In this example, it is convenient to restrict ourselves to Weyl-Heisenberg SICs,

$$|\psi_{jk}\rangle = X^j Z^k |\psi\rangle, \qquad \Pi_{jk} = |\psi_{jk}\rangle\langle\psi_{jk}| \tag{5.177}$$

so the relation between density operators and probability vectors can be written as

$$\rho = \sum_{j,k=0}^{d-1} (d+1) p_{jk} \Pi_{jk} - I. \tag{5.178}$$

Thus,

$$\begin{aligned}\rho &\mapsto \sum_{i=0}^{d-1}\sum_{j,k=0}^{d-1}\epsilon_i\left[(d+1)p_{jk}\right]X^i\Pi_{jk}X^{i\dagger} - \sum_{i=0}^{d-1}\epsilon_i X^i X^{i\dagger} \\ &= (d+1)\sum_{i=0}^{d-1}\sum_{j,k=0}^{d-1}\epsilon_i p_{jk}\Pi_{(j\oplus i)k} - I \\ &= (d+1)\sum_{i=0}^{d-1}\sum_{j,k=0}^{d-1}\epsilon_i p_{(j\ominus i)k}\Pi_{jk} - I, \end{aligned} \tag{5.179}$$

which implies that

$$p'_{jk} = \sum_{i=0}^{d-1}\epsilon_i p_{(j\ominus i)k}. \tag{5.180}$$

For instance, if $\epsilon_i = \frac{1}{d}$ for all i and we consider the SIC basis state

$$p_{jk} = \frac{d\delta_{j0}\delta_{k0}+1}{d(d+1)} \tag{5.181}$$

then

$$p'_{jk} = \begin{cases} \dfrac{2}{d(d+1)} & \text{if } k=0, \\ \dfrac{1}{d(d+1)} & \text{if } k\neq 0. \end{cases} \tag{5.182}$$

6. Qutrit phase damping channel:

$$\rho \mapsto (1-\epsilon)\rho + \epsilon Z\rho Z^\dagger \tag{5.183}$$

where

$$Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix} \tag{5.184}$$

and $\omega = e^{i\frac{2\pi}{3}}$. Here we consider Weyl-Heisenberg SICs generated from the vectors

$$|\psi_{jk}\rangle = X^i Z^j |\psi\rangle \tag{5.185}$$

where

$$X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{5.186}$$

and $|\psi\rangle$ is some fiducial vector. Writing the density operator as

$$\rho = \sum_{j,k=0}^{2} (d+1) p_{jk} \Pi_{jk} - I \tag{5.187}$$

and

$$\vec{p} = \big(p_{00}, p_{01}, p_{02}, p_{10}, p_{11}, p_{12}, p_{20}, p_{21}, p_{22}\big)^T, \tag{5.188}$$

it is straightforward to show that $\vec{p} \mapsto M\vec{p}$ where

$$M = \begin{pmatrix} B & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{pmatrix}, \qquad B = \begin{pmatrix} 1-\epsilon & 0 & \epsilon \\ \epsilon & 1-\epsilon & 0 \\ 0 & \epsilon & 1-\epsilon \end{pmatrix}. \tag{5.189}$$

7. Qutrit amplitude damping channel, with Kraus operators

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{1-\epsilon} & 0 \\ 0 & 0 & \sqrt{1-\epsilon} \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & \sqrt{\epsilon} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & \sqrt{\epsilon} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{5.190}$$

where $\epsilon = 1 - e^{-\Gamma t}$ represents the decoherence parameter. We can compute the affine map directly from these matrices A_i if we express them as

$$A_i = \sum_{j,k=1}^{3} A_{ijk} E_{jk} \tag{5.191}$$

where E_{jk} are just the standard basis matrices

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \ldots, \quad E_{33} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.192}$$

We can then compute a_{im} in Eq. (5.156) using

$$a_{il} = \sum_{j,k} A_{ijk} \mathrm{Tr}\left(E_{jk} D_l\right). \tag{5.193}$$

where D_l is the dual operator to the SIC projector Π_l

$$D_l = \frac{1}{3}\left(4\Pi_l - I\right). \tag{5.194}$$

We can choose any qutrit SIC for computing the structure coefficients, which might be tedious but is straightforward. We find that the affine map $\vec{p} \mapsto M\vec{p} + \vec{t}$ for the amplitude damping noise is given by

$$M = \frac{1}{3}\begin{pmatrix} B_1 & -B_2 & -B_2 \\ B_2 & B_3 & B_2 \\ B_2 & B_2 & B_3 \end{pmatrix} \tag{5.195}$$

where

$$B_1 = \begin{pmatrix} 3-4\epsilon & -\epsilon & -\epsilon \\ -\epsilon & 3-4\epsilon & -\epsilon \\ -\epsilon & -\epsilon & 3-4\epsilon \end{pmatrix}, \quad B_2 = \begin{pmatrix} \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon \end{pmatrix}, \quad B_3 = \begin{pmatrix} 1+2\epsilon' & 1-\epsilon' & 1-\epsilon' \\ 1-\epsilon' & 1+2\epsilon' & 1-\epsilon' \\ 1-\epsilon' & 1-\epsilon' & 1+2\epsilon' \end{pmatrix}, \tag{5.196}$$

with $\epsilon' = \sqrt{1-\epsilon}$, and

$$\vec{t} = \frac{1}{6}\begin{pmatrix} 2\vec{v} \\ -\vec{v} \\ -\vec{v} \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} \epsilon \\ \epsilon \\ \epsilon \end{pmatrix}. \tag{5.197}$$

Chapter 6

Geometric features of qutrit state space

A qutrit is any 3-level quantum system whose pure states are described by unit vectors $|\psi\rangle \in \mathcal{H}_3$. More generally, a qutrit is described by a density operator

$$\rho = \frac{1}{3}I + \tau \tag{6.1}$$

where τ is a traceless Hermitian matrix that makes ρ positive semidefinite.

Analogous to the Bloch ball representation for qubits, states for qutrits are often described using 8 linearly independent generators of $SU(3)$, also known as the Gell-Mann matrices λ_i. We can express any qutrit density operator ρ in terms of λ_i and I,

$$\rho = \frac{1}{3}\left(I + \sqrt{3}\sum_i r_i\lambda_i\right), \tag{6.2}$$

where a pure state corresponds to some unit vector $\vec{r} \in \mathbb{R}^8$ on a 7-dimensional sphere.

The convex set of all qutrit states is such that it contains an in-sphere of radius r and is inscribed by an out-sphere of radius R, where the radii of the 2 spheres are related by

$$\frac{r}{R} = \frac{1}{2}. \tag{6.3}$$

In this chapter, we identify several interesting geometric properties of qutrit state space in terms of SIC probability vectors.

6.1 Qutrit pure states in terms of SICs

Harkening back to some of our results in Chapter 3 and Chapter 5, we will now describe the pure states of a qutrit in terms of SICs. It is useful to recall the general case

$$\sum_i p(i)^2 = \frac{1}{d(d+1)} \tag{6.4}$$

$$\sum_{i,j,k} \tilde{S}_{ijk} p(i)p(j)p(k) = \frac{4}{d(d+1)^2} \tag{6.5}$$

where $\tilde{S}_{ijk}$ are the real parts of the structure coefficients for the Lie algebra defined by the basis of SIC operators. Thinking about the probability vectors $\vec{p}$ as points in an affine space, the former condition defines the sphere containing pure states while the latter condition picks out the points on that sphere that actually correspond to quantum states. As mentioned before, the set of points that describe d-dimensional pure states is topologically identical to the points in complex projective space $\mathbb{C}\mathrm{P}^{d+1}$.

Specializing to the case $d = 3$, a SIC probability vector $\vec{p}$ corresponds to a qutrit pure state if and only if

$$\sum_i p(i)^2 = \frac{1}{6},$$

$$\sum_{i,j,k} \tilde{S}_{ijk} p(i)p(j)p(k) = \frac{1}{12}. \tag{6.6}$$

A particularly elegant formula for pure states is obtained by considering the most exceptional SIC, which is the Weyl-Heisenberg SIC with fiducial vector

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{6.7}$$

The probability distributions $\vec{p}$ for the most exceptional SIC are given by

$$\sum_i p(i)^2 = \frac{1}{6}, \tag{6.8}$$

$$\frac{1}{3} \sum_i p(i)^3 = \sum_{(ijk) \in Q} p(i)p(j)p(k) \tag{6.9}$$

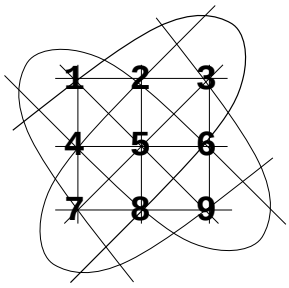


Figure 6.1: The index triples (ijk) for the pure states of the most exceptional SIC, depicted as the 12 lines of a finite affine plane over GF(3), where each point represents an index.

where the set Q of index triples (ijk) is

$$Q = \{(123), (456), (789), (147), (159), (168), (249), (258), (267), (348), (357), (369)\} \quad (6.10)$$

Diagrammatically, the set Q corresponds to the affine lines shown in Fig. 6.1, which coincides with a finite affine plane of order 3, also known as the unique two-$(9, 3, 1)$-design, where each index triple (ijk) corresponds to 3 points that belong to the same line, and Q corresponds to the set of 12 lines [64].

6.2 Rotations between qutrit state spaces

In Sec. 5.5 of Chapter 5, we saw that the probabilities representing a quantum state for two different SICs are connected by a doubly stochastic rotation matrix. Here we are interested in looking at the specific case of $d = 3$ and we will determine the specific rotation matrix between any two qutrit SICs. In particular, we can use it to construct the rotation R_t that takes the probabilities of the most exceptional SIC to the probabilities of some other qutrit SIC in Eq. (3.45).

Let Π_i be the most exceptional SIC and let $\Pi_i^{(r\pm)}(t)$ be the SIC generated by the fiducial vector $\left|\psi_t^{(r\pm)}\right\rangle$. Let

$$Q_i = \frac{1}{3}\left(4\Pi_i - I\right) \quad (6.11)$$

be the dual basis to Π_i (so $\mathrm{Tr}(Q_i\Pi_j) = \delta_{ij}$). Then the orthogonal matrix which takes Π_i onto $\Pi_i^{(r\pm)}(t)$ is

$$R_{ij}^{(r\pm)}(t) = \mathrm{Tr}\left[\Pi_i^{(r\pm)}(t)Q_j\right] \quad (6.12)$$

It turns out that the matrices $R^{(r\pm)}(t)$ have a very simple form. Define the permutations

$$p^{(0+)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{pmatrix}, \qquad p^{(0-)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 3 & 2 & 4 & 6 & 5 & 7 & 9 & 8 \end{pmatrix},$$
$$p^{(1+)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 5 & 9 & 2 & 6 & 7 & 3 & 4 & 8 \end{pmatrix}, \qquad p^{(1-)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 9 & 5 & 2 & 7 & 6 & 3 & 8 & 4 \end{pmatrix},$$
$$p^{(2+)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 6 & 8 & 2 & 4 & 9 & 3 & 5 & 7 \end{pmatrix}, \qquad p^{(2-)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 8 & 6 & 2 & 9 & 4 & 3 & 7 & 5 \end{pmatrix},$$
$$p^{(3+)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 4 & 7 & 2 & 5 & 8 & 3 & 6 & 9 \end{pmatrix}, \qquad p^{(3-)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 7 & 4 & 2 & 8 & 5 & 3 & 9 & 6 \end{pmatrix}. \tag{6.13}$$

Let $P^{(r\pm)}$ be the permutation matrix corresponding to $p^{(r\pm)}$, with matrix elements

$$P^{(r\pm)}_{ij} = \delta_{j,p^{(r\pm)}(i)}. \tag{6.14}$$

Also define

$$a(t) = \frac{1}{3}(1 + 2\cos 2t), \tag{6.15}$$

$$A(t) = \begin{pmatrix} a(t) & a\left(t - \frac{\pi}{3}\right) & a\left(t + \frac{\pi}{3}\right) \\ a\left(t + \frac{\pi}{3}\right) & a(t) & a\left(t - \frac{\pi}{3}\right) \\ a\left(t - \frac{\pi}{3}\right) & a\left(t + \frac{\pi}{3}\right) & a(t) \end{pmatrix}, \tag{6.16}$$

$$R(t) = \begin{pmatrix} A(t) & 0 & 0 \\ 0 & A(t) & 0 \\ 0 & 0 & A(t) \end{pmatrix}. \tag{6.17}$$

It is then straightforward, though somewhat tedious, to verify that

$$R^{(r\pm)}(t) = \left[P^{(r\pm)}\right]^{-1} R(t) P^{(r\pm)}. \tag{6.18}$$

Since $R(t) = I \otimes A(t)$, it follows that

$$\mathrm{Det}\left[R^{(r\pm)}(t)\right] = \mathrm{Det}[R(t)] = \left\{\mathrm{Det}\left[A(t)\right]\right\}^3. \tag{6.19}$$

Because $A(t)$ is a circulant matrix, its eigenvalues are given by

$$\begin{aligned} \lambda_m &= a(t) + \omega^m a\left(t - \frac{\pi}{3}\right) + \omega^{-m} a\left(t + \frac{\pi}{3}\right) \\ &= e^{2itm} \end{aligned} \tag{6.20}$$

for $m = -1, 0, 1$. This implies that $\mathrm{Det}\,[A(t)] = 1$. Thus, $\mathrm{Det}\left[R^{(r\pm)}(t)\right] = 1$ and $R^{(r\pm)}(t)$ is, in fact, a rotation matrix.

It is easily seen that

$$R(t_1)R(t_2) = R(t_1 + t_2), \qquad R(0) = I. \tag{6.21}$$

So the matrices $R(t)$ form a 1-parameter subgroup of the orthogonal group.

We have just established how the qutrit state spaces of different qutrit SICs are related by 3-dimensional rotations. It suggests a particular symmetry of the probability state space in $d = 3$ that we identify in the next section.

6.3 A permutation symmetry for qutrit pure states

Having established the rotational equivalence of the probability spaces for different qutrit SICs, we can now ask the question, what are the symmetry properties possessed by the convex body of qutrits? For comparison, we know that the unitary symmetry for qubits corresponds to the spherical symmetry for the convex body of qubits, which is most apparent in the Bloch representation. We would like to gain a similar insight into the more complicated geometry of qutrits.

Because of the results of the previous section, we know we can focus our attention on the probabilities for most exceptional SIC, whose cubic condition for pure states given by Eq. (6.6) already implies a certain kind of permutation symmetry, which we attempt to make more explicit. To do so, it is better to consider the pure-state condition $\rho = \rho^2$, which in SIC terms is just

$$p(k) = \frac{(d+1)}{3} \sum_{i,j} S_{ijk} p(i) p(j) + \frac{2}{3d(d+1)} \tag{6.22}$$

for $k = 1, 2, \ldots, d^2$. For the most exceptional SIC, Eq. (6.22) becomes

$$12p(k)^2 - 6p(k) - 12 \sum_{(ijk)\in Q} p(i)p(j) + 1 = 0, \tag{6.23}$$

for $k = 1, 2, ..., 9$. Observe that we can rewrite Eq. (6.23) in matrix form using

$$A(\vec{p}) = \begin{pmatrix} f_1 & g_3 & g_2 & g_7 & g_9 & g_8 & g_4 & g_6 & g_5 \\ g_3 & f_2 & g_1 & g_9 & g_8 & g_7 & g_6 & g_5 & g_4 \\ g_2 & g_1 & f_3 & g_8 & g_7 & g_9 & g_5 & g_4 & g_6 \\ g_7 & g_9 & g_8 & f_4 & g_6 & g_5 & g_1 & g_3 & g_2 \\ g_9 & g_8 & g_7 & g_6 & f_5 & g_4 & g_3 & g_2 & g_1 \\ g_8 & g_7 & g_9 & g_5 & g_4 & f_6 & g_2 & g_1 & g_3 \\ g_4 & g_6 & g_5 & g_1 & g_3 & g_2 & f_7 & g_9 & g_8 \\ g_6 & g_5 & g_4 & g_3 & g_2 & g_1 & g_9 & f_8 & g_7 \\ g_5 & g_4 & g_6 & g_2 & g_1 & g_3 & g_8 & g_7 & f_9 \end{pmatrix}, \tag{6.24}$$

which defines a mapping from probability distributions to a matrix whose diagonal entries are given by

$$f_i = \frac{5}{6} - 2p(i) \tag{6.25}$$

and off-diagonal entries are given by

$$g_j = p(j) - \frac{1}{6}. \tag{6.26}$$

We see that $\vec{p}$ corresponding to pure states must satisfy the pseudolinear equation

$$A(\vec{p})\vec{p} = 0. \tag{6.27}$$

It is worthwhile to point out some interesting properties of $A = A(\vec{p})$:

1. A is a symmetric matrix, i.e., $A^T = A$.
2. A is a Latin square in the $p(i)$, that is, each probability component appears just once in each row or column.
3. A is composed of 2 types of 3×3 blocks, one of which is

$$A_1(i, j, k) = \begin{pmatrix} \frac{5}{6} - 2p(i) & p(k) - \frac{1}{6} & p(j) - \frac{1}{6} \\ p(k) - \frac{1}{6} & \frac{5}{6} - 2p(j) & p(i) - \frac{1}{6} \\ p(j) - \frac{1}{6} & p(i) - \frac{1}{6} & \frac{5}{6} - 2p(k) \end{pmatrix} \tag{6.28}$$

for the diagonal blocks, while the other one is

$$A_2(i,j,k) = \begin{pmatrix} p(i) - \frac{1}{6} & p(k) - \frac{1}{6} & p(j) - \frac{1}{6} \\ p(k) - \frac{1}{6} & p(j) - \frac{1}{6} & p(i) - \frac{1}{6} \\ p(j) - \frac{1}{6} & p(i) - \frac{1}{6} & p(k) - \frac{1}{6} \end{pmatrix} \tag{6.29}$$

for the off-diagonal blocks. In particular, we can write A as

$$A = \begin{pmatrix} A_1(1,2,3) & A_2(7,8,9) & A_2(4,5,6) \\ A_2(7,8,9) & A_1(4,5,6) & A_2(1,2,3) \\ A_2(4,5,6) & A_2(1,2,3) & A_1(7,8,9) \end{pmatrix}. \tag{6.30}$$

However, we can replace the index triples with any striation in Q, that is, any set of index triples (ijk) for the most exceptional SIC that correspond to 3 affine lines with the same slope. This is because such a change does not alter the condition for pure states specified by $A(\vec{p})\vec{p} = 0$.

4. Observe that the diagonal blocks mixes probabilities in the same index triple (ijk), i.e., $A_1(i,j,k)$ always meets $(p(i), p(j), p(k))$. On the other hand, the off-diagonal block mixes the other two triples, e.g. $A_2(4,5,6)$ multiplies with $(p(7), p(8), p(9)$.

Properties (3) and (4) of A imply the following about the solutions $\vec{p}$ to Eq. (6.27):

(a) If we denote by $(123, 456, 789)$ the components of a probability distribution $\vec{p}$ corresponding to pure states, that is,

$$\vec{p} = \begin{pmatrix} p(1) \\ p(2) \\ \vdots \\ p(9) \end{pmatrix} \tag{6.31}$$

but thought of as grouped into triples, then $(123, 789, 456)$ and all other transpositions of such triples of components of $\vec{p}$ also correspond to valid pure states.

(b) Also, applying a cyclic permutation to each triple of probability components, such as taking $(123, 456, 789)$ into $(231, 564, 897)$ or $(312, 645, 978)$ also yields valid pure states.

(c) For any $\vec{p}$ corresponding to a pure state, a permutation of its components that corresponds to some combination of the above lead to some other probability distribution that is also a pure state.

For specificity, we will sometimes call these two conditions the *affine-line permutation symmetry* of qutrit pure states, in reference to the finite affine plane associated with the indices of the most exceptional SIC.

Eq. (6.27) gives us another way to characterize SIC probability vectors corresponding to pure states in terms of the mapping $A_{\vec{p}}$. That is, a probability distribution $\vec{p}$ is associated with a qutrit pure state if and only if

(i) The vector $\vec{p}$ lies on a sphere given by $\vec{p} \cdot \vec{p} = \frac{1}{6}$, and

(ii) The probability vector $\vec{p}$ is in the null space of the matrix $A(\vec{p})$.

Furthermore, if $\vec{p}$ satisfy these 2 conditions, the affine-line permutations of its components also correspond to other pure states.

6.4 The boundary of qutrit state space

One concrete way to understand the geometry of qutrit state space is to figure out what the convex body looks like. In this regard, we want to consider not just the pure states but all boundary points of the set. Some valuable insight into the shape of the boundary is gained by looking at the distance of the boundary states from the center of the space, the maximally mixed state $\rho = \frac{1}{d}I$ as a function of direction. Specifically, we can write the SIC probabilities in the form

$$p(i) = \frac{1}{d^2} + rn(i) \tag{6.32}$$

where $\vec{n}$ is a direction vector with

$$\sum_i n(i) = 0, \qquad \sum_i n(i)^2 = 1. \tag{6.33}$$

Let $r(\vec{n})$ be the value of the r corresponding to the quantum state on the boundary, which is given by Eq. (6.32). Here we calculate this function for the most exceptional SIC. Put differently, we are looking for the polar equation describing its boundary states.

The boundary is determined using the following lemma:

Lemma 6.4.1. Let ρ be an arbitrary Hermitian operator on a 3-dimensional Hilbert space. Then

(i) ρ is a density operator if and only if

$$\mathrm{Tr}\,(\rho) = 1, \quad \mathrm{Tr}\left(\rho^2\right) \leq 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) \leq 1; \tag{6.34}$$

(ii) ρ is a density operator for a boundary state if and only if

$$\mathrm{Tr}\,(\rho) = 1, \quad \mathrm{Tr}\left(\rho^2\right) \leq 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) = 1; \tag{6.35}$$

(iii) ρ is a density operator for a pure quantum state if and only if

$$\mathrm{Tr}\,(\rho) = 1, \quad \mathrm{Tr}\left(\rho^2\right) = 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) = 1. \tag{6.36}$$

Proof. Suppose ρ is a density operator. It immediately follows that $\mathrm{Tr}\,(\rho) = 1$ and $\mathrm{Tr}\,(\rho^2) \leq 1$. To prove the remaining inequality, let $\alpha, \beta, 1-\alpha-\beta$ be the eigenvalues of ρ. Then

$$3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) - 1 = -6\alpha\beta(1-\alpha-\beta) \leq 0. \tag{6.37}$$

For ρ to be a boundary state, that means at least one of its eigenvalues must vanish, in which case

$$3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) - 1 = 0. \tag{6.38}$$

In addition, if ρ is a pure state then $\mathrm{Tr}\,(\rho^2) = 1$. This shows the above conditions are necessary.

To prove sufficiency, let $\rho = \rho^\dagger$ such that

$$\mathrm{Tr}\,(\rho) = 1, \quad \mathrm{Tr}\left(\rho^2\right) \leq 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) < 1.$$

The first equality means that we can take the eigenvalues of ρ to be $\alpha, \beta, 1-\alpha-\beta$. From Eq. (6.37), we get

$$\alpha\beta(1-\alpha-\beta) > 0 \tag{6.39}$$

Thus, either (i) all eigenvalues are nonnegative or (ii) two of them are negative. We can show that (ii) is impossible. Assume the contrary to hold. Without loss of generality, $\alpha, \beta < 0$ implies that $1-\alpha-\beta > 1$, which then suggests that $\mathrm{Tr}\,(\rho^2) > 1$, contradicting the hypothesis on ρ. Thus, ρ is positive semidefinite and therefore a density operator.

Next assume that

$$\mathrm{Tr}\,(\rho) = 1, \quad \mathrm{Tr}\left(\rho^2\right) \leq 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) = 1.$$

Then,

$$\alpha\beta(1-\alpha-\beta) = 0, \tag{6.40}$$

which means that at least one eigenvalue vanishes. Without loss of generality, take $\alpha = 0$. Then

$$\mathrm{Tr}\left(\rho^2\right) = 2\left(\beta - \frac{1}{2}\right)^2 + \frac{1}{2} \leq 1 \tag{6.41}$$

implying that $0 \leq \beta \leq 1$. Thus, ρ is again a positive semidefinite operator. It is also on the boundary since one of its eigenvalues is zero.

Finally assume

$$\mathrm{Tr}\left(\rho\right) = 1, \quad \mathrm{Tr}\left(\rho^2\right) = 1, \quad 3\mathrm{Tr}\left(\rho^2\right) - 2\mathrm{Tr}\left(\rho^3\right) = 1.$$

Using the same argument above, ρ is a density operator with eigenvalues $0, \alpha, 1-\alpha$. Since $\mathrm{Tr}\left(\rho^2\right) = 1$, it must be that $\alpha = 0$ or $\alpha = 1$ and therefore ρ is also a rank-1 projection operator. □

We can use the lemma for the quantum states associated with the most exceptional SIC. To this end, recall its structure coefficients in Eq. (3.60). We use these to calculate $\mathrm{Tr}\left(\rho^2\right)$ and $\mathrm{Tr}\left(\rho^3\right)$ for ρ given by Eq. (5.1). We find that for the most exceptional SIC

$$\begin{aligned} \mathrm{Tr}\left(\rho^2\right) &= 12\sum_i p(i)^2 - 1, \\ \mathrm{Tr}\left(\rho^3\right) &= 1 + 24\sum_i p(i)^3 - 72\sum_{(ijk)\in Q} p(i)p(j)p(k). \end{aligned} \tag{6.42}$$

Substituting Eq. (6.32) into the probabilities above, we obtain

$$\begin{aligned} \sum_i p(i)^2 &= \frac{1}{9} + r^2, \\ \sum_i p(i)^3 &= \frac{1}{81} + \frac{r^2}{3} + r^3\sum_i n(i)^3, \\ \sum_{(ijk)\in Q} p(i)p(j)p(k) &= \frac{8}{81} - \frac{r^2}{3} + 6r^3\sum_{(ijk)\in Q} n(i)n(j)n(k). \end{aligned} \tag{6.43}$$

Consequently, we can restate the conditions in Lemma 6.4.1 as follows:

(i) ρ is a density operator if and only if

$$r^2 \leq \frac{1}{18}, \qquad 4r^3F(\vec{n}) - r^2 + \frac{1}{54} \geq 0; \tag{6.44}$$

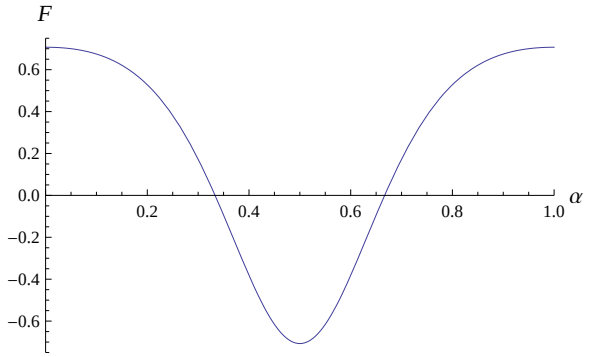


Figure 6.2: A plot of F as a function of one of the eigenvalues α for boundary states.

(ii) ρ is a density operator on the boundary of the state space if and only if

$$r^2 \leq \frac{1}{18}, \qquad 4r^3 F(\vec{n}) - r^2 + \frac{1}{54} = 0; \tag{6.45}$$

(iii) ρ is a density operator for a pure state if and only if

$$r^2 = \frac{1}{18}, \qquad F(\vec{n}) = \frac{1}{\sqrt{2}}; \tag{6.46}$$

where

$$F(\vec{n}) = \sum_i n(i)^3 - 3 \sum_{(ijk)\in Q} n(i)n(j)n(k). \tag{6.47}$$

Thus, the value of $r(\vec{n})$ giving the distance of a boundary state from the completely mixed state along the direction of $\vec{n}$ is the smallest positive root of

$$4r^3 F(\vec{n}) - r^2 + \frac{1}{54} = 0. \tag{6.48}$$

We can find the bounds on $F = F(\vec{n})$ if we express it in terms of the eigenvalues of the density operator ρ. If we denote the eigenvalues of ρ by $\alpha, \beta, 1-\alpha-\beta$, we can perform some algebra using Eqs. (6.42) and (6.43) to get

$$F(\alpha,\beta) = \frac{[2-3(\alpha+\beta)]\,(3\alpha-1)(3\beta-1)}{-2\sqrt{2}\,[1-3(\alpha+\beta+\alpha\beta)+3(\alpha+\beta)^2]^{3/2}} \tag{6.49}$$

which is just F in terms of two of the eigenvalues of ρ.

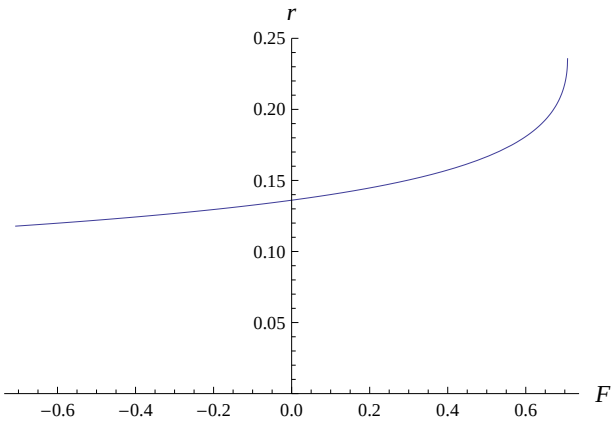


Figure 6.3: The radial distance r of boundary states from the maximally mixed state, as a function of $F = F(\vec{n})$.

Differentiating F with respect to α and β, we can find the critical points of F and solve for its maximum and minimum values. From this, it is straightforward to show that

$$-\frac{1}{\sqrt{2}} \leq F \leq \frac{1}{\sqrt{2}} \tag{6.50}$$

where the upper (respectively, lower) bound is achieved when two of the eigenvalues are identical and $< \frac{1}{3}$ (respectively, $> \frac{1}{3}$). If all the eigenvalues are equal to $\frac{1}{3}$ this corresponds to the maximally mixed state, for which F is undefined. For boundary states, at least one of the eigenvalues must be zero. So the only case we need to consider for F is when $\beta = 1 - \alpha$. Fig. 6.2 shows F as a function of α, provided that one of the eigenvalues vanishes.

In terms of F, the desired root in Eq. (6.48) is then actually given by

$$r = \begin{cases} \dfrac{1}{12F}\left(1 + \dfrac{g}{\omega^s} + \dfrac{\omega^s}{g}\right) & \text{if } F \neq 0 \\ \dfrac{1}{3\sqrt{6}} & \text{if } F = 0, \end{cases} \tag{6.51}$$

where $s = \mathrm{sgn}(F)$ and g is the cube root with the smallest positive part in

$$g^3 = 1 - 4F^2 + 2F\sqrt{4F^2 - 2}. \tag{6.52}$$

When $F = \pm\frac{1}{\sqrt{2}}$, $g = e^{i\frac{\pi}{3}}$ and we get the bounds for r:

$$\frac{1}{6\sqrt{2}} \leq r \leq \frac{1}{3\sqrt{2}}. \tag{6.53}$$

A plot of r as a function of F is shown in Fig. 6.3.

6.5 Plane sections of qutrit state space

To be able to visualize the shape of qutrit state space, we need to go beyond just an algebraic description of its boundary. Since we are unable to plot it as a subset of $\mathbb{R}^9$, the best we can do is look at lower-dimensional cross-sections of the space. A number of authors [56, 68, 78, 79, 97] have previously explored 2-dimensional and 3-dimensional cross-sections of the generalized Bloch representation of a qutrit. Here we review some of their results and compare it with what one would get for the SIC representation.

For Hilbert spaces of dimension two, it is convenient to represent the set of quantum states on a Bloch ball. There are two main reasons why Bloch vectors provide an elegant representation for quantum states:

(i) Every quantum state is represented by a real-valued vector whose components refer to expectation values of some minimal informationally-complete measurement.

(ii) The dynamics of a quantum system is manifested by intuitive operations on the Bloch ball—for example, unitary evolution corresponds to a rotations of the ball, while a depolarizing quantum channel corresponds to shrinking the ball radius.

We can extend the notion of a Bloch sphere in higher dimensions by using the infinitesimal generators of the special unitary group $SU(d)$ as the generalization of Pauli matrices, which is a set of $d^2 - 1$ operators λ_i that together with the identity operator form an orthogonal basis for $d \times d$ complex matrices. The generalized Bloch vector space for $d > 2$ is a proper subset of a ball in $\mathbb{R}^{d^2-1}$.

Any density operator ρ can then be written as

$$\rho = \frac{1}{d}\left(I + \frac{d}{2}\vec{r}\cdot\vec{\lambda}\right), \tag{6.54}$$

where $\vec{\lambda}$ is a vector of the generators of $SU(d)$ and $\vec{r}$ is the generalized Bloch vector. In higher dimensions, not all of the points on the surface of the sphere correspond to pure states. The simplest way to explain the discrepancy is to say that vectors $\vec{r}$ that do not correspond to positive semidefinite ρ are excluded since they do not describe quantum states. In essence, the richness in the geometric structure of quantum state space is a consequence of positivity.

To establish the required constraints on ρ, we employ once again Lemma 5.8.1 in Sec. 5.8, which when applied to the characteristic polynomial of Eq. (6.54) leads to the following theorem:

Theorem 6.5.1. Let $a_i(\vec{r})$ be the coefficients of the characteristic polynomial $\det(xI - \rho)$ of ρ and define the Bloch vector space B,

$$B\left(\mathbb{R}^{d^2-1}\right) = \{\vec{r} \in \mathbb{R}^{d^2-1} \mid a_i(\vec{r}) \geq 0\}. \tag{6.55}$$

Then the map

$$\vec{r} \in B\left(\mathbb{R}^{d^2-1}\right) \mapsto \rho = \frac{1}{d}\left(I + \frac{d}{2}\vec{r}\cdot\vec{\lambda}\right) \in \mathcal{D}, \tag{6.56}$$

is a bijection from $B\left(\mathbb{R}^{d^2-1}\right)$ to the density operator space $\mathcal{D}$.

The positivity conditions $a_i(\vec{r}) \geq 0$ can then be expressed in terms of trace powers of ρ using the Newton-Girard identities. Specializing to $d = 3$, the theorem says that ρ is positive semidefinite when

$$\begin{aligned} a_1 &= \mathrm{Tr}\,(\rho) = 1, \\ a_2 &= \frac{1 - \mathrm{Tr}\,(\rho^2)}{2} \geq 0, \\ a_3 &= \frac{1 - 3\mathrm{Tr}\,(\rho^2) + 2\mathrm{Tr}\,(\rho^3)}{6} \geq 0. \end{aligned} \tag{6.57}$$

Kimura [78] provides a nice visualization of qutrit states by plotting the 2-dimensional sections of the generalized Bloch vector space. In this section, we compare those plane sections with the ones obtained for SIC probability vectors.

Recall the density operator expressed in terms of SIC probabilities,

$$\rho = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right]\Pi_i. \tag{6.58}$$

We simply plug in this ρ into the positivity conditions of Eq. (6.57). By definition, $a_1 = \mathrm{Tr}\,(\rho) = 1$ for a density operator. For the constraint $a_2(\vec{p}) \geq 0$ we have

$$a_2(\vec{p}) = 1 - \frac{d(d+1)}{2}\sum_i p(i)^2 \geq 0. \tag{6.59}$$

On the other hand, for $a_3(\vec{p}) \geq 0$ we have

$$a_3(\vec{p}) = \frac{1}{6}\left[6 - 7d(d+1)\sum_i p(i)^2 + 2d(d+1)^2 \sum_{i,j,k} S_{ijk}p(i)p(j)p(k)\right] \geq 0. \tag{6.60}$$

Note that the pure states correspond to the case $a_2 = 0$ and $a_3 = 0$. We can then compare the 2-dimensional sections for the SIC probability vectors with the cross-sections

$$S_2(i,j) = \{\vec{r} \in B(\mathbb{R}^8) | \vec{r} = (0, ..., r_i, 0, ..., r_j, ..., 0)\} \tag{6.61}$$

obtained by Kimura for Bloch vectors [78]. It is helpful to know the mapping from $\vec{r}$ to $\vec{p}$ obtained from equating Eq. (6.54) and Eq. (6.58),

$$\frac{1}{3}I + \frac{1}{2}\vec{r}\cdot\vec{\lambda} = 4\vec{p}\cdot\vec{\Pi} - I, \tag{6.62}$$

where $\vec{\Pi}$ is the vector of SIC projectors of the most exceptional SIC. One way to work this out explicitly is to expand each SIC element in terms of the basis defined by the Gell-Mann matrices and the identity operator. The cross-sections $S_2(i,j)$ expressed in terms of SIC probabilities are listed in Table 6.1 and shown in Fig. 6.4.

Samples of each type of 2-dimensional cross-section are displayed in Fig. 6.4. It is important to note that these plane sections are identical to Kimura's graphs [78] if plotted with respect to the two parameters defining each subspace. This is consistent with Rosado's observation that the set of Bloch vectors and the set of SIC probability vectors corresponding to quantum states are identical, up to an appropriately chosen scaling factor [114].

Such 2-dimensional qutrit sections have been considered before and here we briefly revisit some of the observations made by previous authors regarding the unitary equivalence of cross-sections.

There are a total of ${}^8\mathrm{C}_2 = 28$ plane sections to consider, where a general 2-dimensional cross-section corresponds to density operators of the form

$$\rho = \frac{1}{3}\left(I + x\lambda_1 + y\lambda_2\right) \tag{6.63}$$

where λ_1, λ_2 are traceless hermitian operators and $x, y \in \mathbb{R}$. Two 2-dimensional cross-sections are said to be unitarily equivalent if for some unitary operator U in $SU(3)$,

$$(U\lambda_1 u^\dagger, U\lambda_2 U^\dagger) = (\tilde{\lambda}_1, \tilde{\lambda}_2) \tag{6.64}$$

Type I $\left(r_8 \leq \frac{1}{\sqrt{3}}, r_8 \geq \pm\sqrt{3}r_i - \frac{2}{\sqrt{3}}; i = 1,2,3\right)$	
$S_2(1,8)$	$\frac{1}{36}\Big(4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8,$ $4-\sqrt{3}r_8, 4+9r_1+2\sqrt{3}r_8, 4+9r_1+2\sqrt{3}r_8, 4-9r_1+2\sqrt{3}r_8\Big)$
$S_2(2,8)$	$\frac{1}{36}\Big(4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8,$ $4-3\sqrt{3}r_2+2\sqrt{3}r_8, 4+3\sqrt{3}r_2+2\sqrt{3}r_8, 4+2\sqrt{3}r_8\Big)$
$S_2(3,8)$	$\frac{1}{36}(4-3r_3-\sqrt{3}r_8, 4-3r_3-\sqrt{3}r_8, 4-3r_3-\sqrt{3}r_8, 4+3r_3-\sqrt{3}r_8,$ $4+3r_3-\sqrt{3}r_8, 4-3r_3-\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8)$
Type II $\left(r_j \leq \frac{2}{3}, r_j \geq \frac{3}{2}r_i^2 - \frac{2}{3}; i = 3, j = 4,5,6,7\right)$	
$S_2(3,4)$	$\frac{1}{36}\left(4-3r_3, 4-3r_3, 4-3r_3, 4+3r_3+3r_4, 4+3r_3+3r_4, 4+3r_3-6r_4, 4, 4, 4\right)$
$S_2(3,5)$	$\frac{1}{36}\left(4-3r_3, 4-3r_3, 4-3r_3, 4+3r_3+3\sqrt{3}r_5, 4+3r_3-3\sqrt{3}r_5, 4+3r_3, 4, 4, 4\right)$
$S_2(3,6)$	$\frac{1}{36}\left(4+3r_3+3r_6, 4+3r_3+3r_6, 4+3r_3-6r_6, 4-3r_3, 4-3r_3, 4-3r_3, 4, 4, 4\right)$
$S_2(3,7)$	$\frac{1}{36}\left(4+3r_3-3\sqrt{3}r_7, 4+3r_3+3\sqrt{3}r_7, 4+3r_3, 4-3r_3, 4-3r_3, 4-3r_3, 4, 4, 4\right)$
Type III $\left(r_i^2 + \frac{16}{9}\left(r_8 + \frac{\sqrt{3}}{6}\right)^2 \leq 1; i = 4,5,6,7\right)$	
$S_2(4,8)$	$\frac{1}{36}\Big(4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4+3r_4-\sqrt{3}r_8, 4+3r_4-3\sqrt{3}r_8,$ $4-6r_4-\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8\Big)$
$S_2(5,8)$	$\frac{1}{36}\Big(4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4+3\sqrt{3}r_5-\sqrt{3}r_8, 4+3\sqrt{3}r_5-\sqrt{3}r_8,$ $4-\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8\Big)$
$S_2(6,8)$	$\frac{1}{36}\Big(4+3r_6-\sqrt{3}r_8, 4+3r_6-\sqrt{3}r_8, 4-6r_6-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8,$ $4-\sqrt{3}r_8, 4+\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8\Big)$
$S_2(7,8)$	$\frac{1}{36}\Big(4-3\sqrt{3}r_7-\sqrt{3}r_8, 4+3\sqrt{3}r_7-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8, 4-\sqrt{3}r_8,$ $4-\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8, 4+2\sqrt{3}r_8\Big)$

Table 6.1: Plane cross-sections of the SIC probability space matched to Kimura's classification. All remaining cross-sections are Type IV $\left(r_i^2 + r_j^2 = \frac{4}{9}\right)$.

Type I: Triangle	**Type II:** Parabola	**Type III:** Ellipse	**Type IV:** Circle
18, 28, 38	34, 35, 36, 37	48, 58, 68, 78	12, 13, 23, 45, 67
			14, 15, 16, 17, 24, 25, 26, 27, 46, 47, 56, 57

Table 6.2: The 28 plane sections arranged according to type, with ij specifying the subspace spanned by (λ_i, λ_j).

where (λ_1, λ_2) and $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ are pairs of traceless hermitian operators that determine some 2-dimensional subspace according to Eq. (6.63).

The conventional choice for the λ_i are the Gell-Mann matrices and they give us ${}^8\mathrm{C}_2 = 28$ possible pairs of operators that span a plane section, which partition into the 4 distinct types described by Kimura.

For concreteness, let us choose the λ_i to be the operators

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}. \tag{6.65}$$

The plane sections then fall under the distinct types as listed in Table 6.2.

The triangle, parabola, and ellipse cross-sections are unitarily equivalent to each other, where the equivalences can be seen when we interchange the roles played by the Gell-Mann operators. The circle cross-sections in Table 6.2 is divided into two subsets, one with 5 elements generated from a pair of anti-commuting λ-matrices, which is not in the same equivalence class as the remaining 12 plane sections.

Gniewomir Sarbicki and Ingemar Bengtsson complete the details by classifying the set of all possible 2-dimensional cross-sections of a qutrit [115]. They show that a general

plane section is bounded by a plane cubic curve, which expressed in a suitable basis, is given by

$$\frac{1}{9} - x^2 - y^2 + 3\left(2abc\cos\phi - k(1-k^2-3a^2)\right)y^3 + gkx^2y + 3(b^2-c^2)xy^2 = 0 \qquad (6.66)$$

where $b \geq c$, $a^2+b^2+c^2+3k^2 = 1$ and ϕ can be any angle. In Eq. (6.66), we are considering density operators of the form

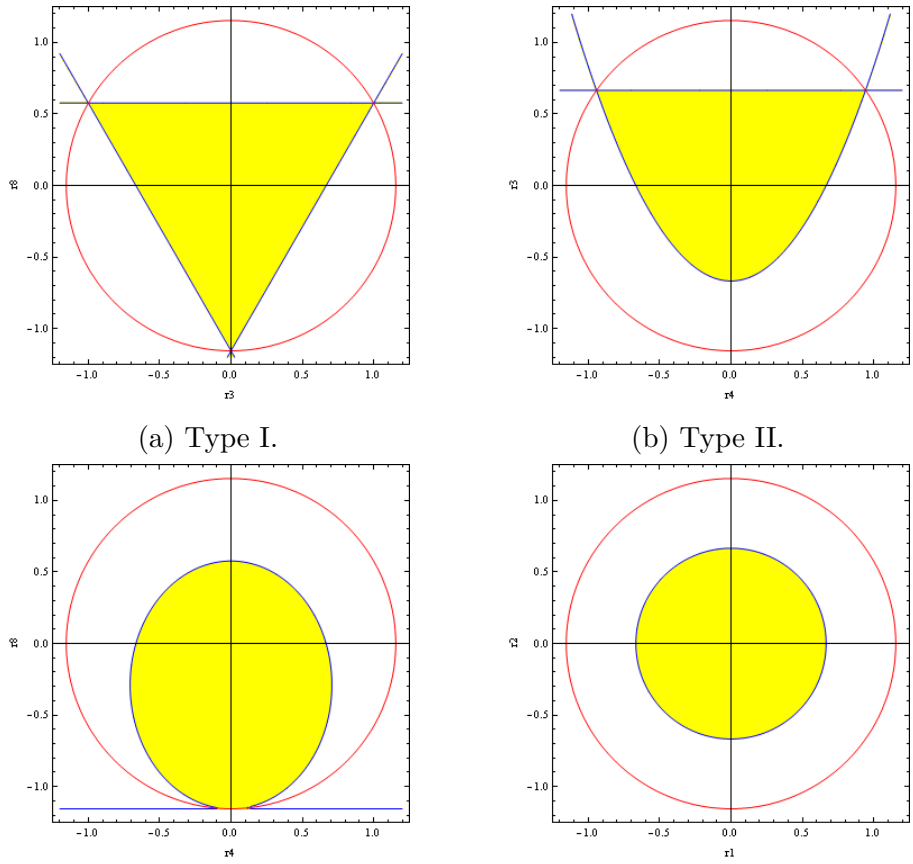
$$\rho = \frac{1}{3}I + xA + yB \qquad (6.67)$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} k & ae^{i\phi} & be^{i\phi} \\ ae^{-i\phi} & k & ce^{i\phi} \\ be^{-i\phi} & ce^{-i\phi} & -2k \end{pmatrix} \qquad \text{for } abc \neq 0 \qquad (6.68)$$

and

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} k & a & b \\ a & k & c \\ b & c & -2k \end{pmatrix} \qquad \text{for } abc = 0. \qquad (6.69)$$



(a) Type I.

(b) Type II.

(c) Type III.

(d) Type IV.

Figure 6.4: 2-dimensional cross-sections of the SIC probability space corresponding to Types I, II, III, and IV, in Kimura's classification. The yellow shaded regions represent valid quantum states. The axes correspond to 2 parameters that define the 2-dimensional subspace for the SIC probabilities given in Table 6.1.

Chapter 7

Quantum state space as a maximal consistent set

The Born rule specifies the probabilities for the outcome of any quantum measurement. In the SIC representation where states and measurements are expressed in terms of a SIC $\{\Pi\}_{i=1}^{d^2}$,

$$q(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i) \tag{7.1}$$

expresses the probability of result j for some measurement $\{E_j\}_{j=1}^{N}$ described by the conditional probabilities

$$r(j|i) = \mathrm{Tr}\left(E_j \Pi_i\right). \tag{7.2}$$

At first glance, Eq. (7.1) looks like an affine transformation

$$p(i) \mapsto (d+1)p(i) - \frac{1}{d} \tag{7.3}$$

on the probabilities. This hints at the intriguing possibility that quantum theory can be understood as some type of extension of classical probability theory that includes certain events that require a different rule for relating marginal to conditional probabilities.

If we think of the Born rule as such an extension, it implies that quantum mechanics corresponds to a restriction on the probability simplex $\Delta^{d^2-1} \subset \mathbb{R}^{d^2}$, since not all possible $p(i)$ or $r(j|i)$ will yield a valid probability distribution for $q(j)$. In fact, some of the constraints implied by the Born rule can be expressed in terms of a set of inequalities, which we will use to define the notion of maximal consistency, and which will serve as the jumping-off point for our efforts to reconstruct quantum state space.

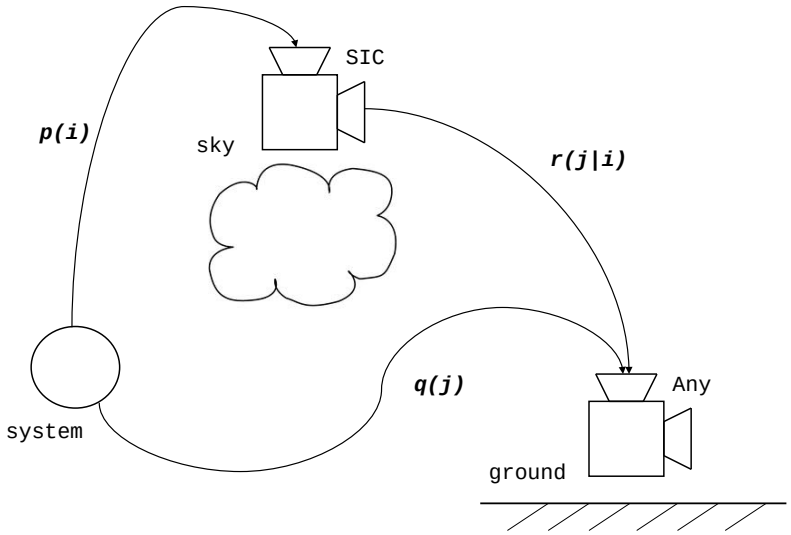


Figure 7.1: A conceptual diagram for the Born rule in terms of a SIC. It shows how the Born rule relates a counterfactual SIC measurement in the sky with some arbitrary measurement on the ground.

7.1 Quantum states from the Born rule

In order to appreciate the conceptual significance of Eq. (7.1), it is helpful for us to consider the situation depicted in Fig. 7.1. Here we are concerned with performing some arbitrary measurement on a quantum system represented in the diagram by a measurement on the ground, where $q(j)$ is the probability associated with each outcome j. We imagine that before performing the ground measurement, there is some fixed SIC measurement in the sky that occurs independently of the ground. The probabilities $p(i)$ describe the SIC outcomes in the sky while the conditional probabilities $r(j|i)$ correspond to the outcomes of the ground measurement assuming outcome i is obtained from the SIC. We can assign probabilities for the ground measurement depending on whether the SIC in the sky is realized or not.

When the sky measurement is actually performed, the outcomes j of the ground measurement are given by

$$q(j) = \sum_i p(j) r(j|i) \tag{7.4}$$

in accordance with the usual laws of probability. This is how we would assign probabilities for a classical system. The quantum case corresponds to the scenario where the sky measurement does not occur. In this case, the system is fed directly to the ground measurement and we use the Born rule in the form of Eq. (7.1), which expresses formally the idea that unperformed measurements have no outcomes.

Moreover, the similarity between Eq. (7.1) and Eq. (7.4) suggests that we may think of the Born rule as a certain kind of coherence condition on actual measurements done in the laboratory with a counterfactual measurement, which we fix to be a SIC measurement, so that quantum probabilities are not independent of the probabilities involved in Eq. (7.4) but are intimately connected to them.

Thus, if we would like to make the assumption that the Born rule constitutes a normative addition to probability theory that leads to a formal description of quantum states in terms of SIC probabilities, where we demand that

(i) the Born rule must yield $q(j)$ that is a probability, and

(ii) using the Bayes' rule on conditional probabilities lead to valid states.

To describe the actual conditions on the SIC probabilities in terms of inequalities, we obviously need (i) if we want to stay within the simplex while for (ii) we know that it should at least hold in the case when we know both the ground and sky measurements are performed but we are completely ignorant of the outcomes in the sky, that is, when

$$\vec{p} = \frac{1}{d^2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \tag{7.5}$$

then Bayes rule tells us that

$$r(i|j) = \frac{r(j|i)p(i)}{q(j)} = \frac{r(j|i)\left(\frac{1}{d^2}\right)}{\sum_k \left(\frac{1}{d^2}\right) r(j|k)} = \frac{r(j|i)}{\sum_k r(j|k)} \equiv s(i) \tag{7.6}$$

where it means that if we ignore the dependence on j, then the probability distribution $\vec{s}$ must correspond to a quantum state itself.

If we substitute Eq. (7.6) into Eq. (7.1) we have

$$q(j) = \left[(d+1)\sum_i p(i)s(i) - \frac{1}{d}\right] \sum_k r(j|k) \tag{7.7}$$

Since $q(j) \geq 0$ for $\vec{q}$ to be a proper probability distribution, we get

$$\vec{p} \cdot \vec{s} \geq \frac{1}{d(d+1)} \tag{7.8}$$

where $\cdot$ is the usual scalar product for points in $\mathbb{R}^{d^2}$. We can get an upper bound from other considerations. For instance, one may take the special case of the sky measurement being the same as the ground measurement, in which case,

$$p(j) = (d+1)\sum_i p(i)\bar{r}(j|i) - \frac{1}{d}\sum_i \bar{r}(j|i), \tag{7.9}$$

where $\bar{r}(j|i)$ denotes the same measurement for the ground and sky. Furthermore, we demand that whenever we assign a uniform distribution for the sky measurement, we should assign a uniform distribution for the ground, a property called *in-step unpredictability* [47]. So with $p(i) = p(j) = \frac{1}{d^2}$, we get

$$\sum_i \bar{r}(j|i) = 1. \tag{7.10}$$

We can use this special case in Eq. (7.7) and the constraint $q(j) \leq 1$ gives us

$$\vec{p} \cdot \vec{s} \leq \frac{2}{d(d+1)}. \tag{7.11}$$

The same upper bound can be obtained by demanding that one can have at most d points in any set S where the dot product of any pair of probability vectors in S is equal to the lower bound $\frac{1}{d(d+1)}$ (in Hilbert space, the set S will just be an orthonormal basis).

Thus, we have

$$\frac{1}{d(d+1)} \leq \vec{p} \cdot \vec{s} \leq \frac{2}{d(d+1)} \tag{7.12}$$

Because we do not specify any particular $\vec{p}$ or $\vec{s}$, this condition must hold for any pair of SIC probabilities that represent quantum states. And because it establishes a necessary binary relation between quantum states, we refer to Eq. (7.12) as the *consistency condition* for SIC probabilities.

Since Eq. (7.12) is a constraint on SIC probabilities obtained from the Born rule, we can start with the postulate that any probability distribution corresponding to a pure state must satisfy this constraint. The remainder of the task, then, is to see what additional structure is needed in order to reconstruct the space of SIC probabilities isomorphic to quantum state space. Thus, we imagine the space of d-dimensional quantum states to be

a consistent subset of the $d^2 - 1$ probability simplex, with a global geometric structure reflecting the projective unitary symmetry of the associated density operators.

So far, the constraints on the SIC probabilities from Eq. (7.12) just come from insisting that $\vec{q}$ always be a probability vector. Without any other assumptions, the only other thing we can demand is *maximality*. Suppose S is the set of probability vectors representing quantum states. S is said to maximal if it is the largest subset of the probability simplex that is consistent. That is, if we consider any probability vector $\vec{v} \notin S$, then for some $\vec{p} \in S$ it must be the case that

$$\vec{p} \cdot \vec{v} < \frac{1}{d(d+1)} \quad \text{or} \quad \vec{p} \cdot \vec{v} > \frac{2}{d(d+1)}. \tag{7.13}$$

Since S is both consistent and maximal, we refer to it as a *maximal consistent set*. But before we describe our approach to reconstructing quantum state space as a maximal consistent set, it is useful to consider some of the properties of any maximal consistent set. We do this in the next section by introducing a more general definition of maximal consistent sets.

7.2 General maximal consistent sets

A more general notion of a maximal consistent set begins with identifying a normed real vector space V where the points are embedded in, also known as the ambient space. In the case of SIC probabilities, this would be the probability simplex Δ^{d^2-1} over d^2 outcomes. In general, we consider any finite-dimensional real inner product space, which means V is isomorphic to some subset of the real Euclidean n-space equipped with its canonical inner product, the dot product

$$\langle \vec{p}, \vec{q} \rangle = \vec{p} \cdot \vec{q} = \sum_{i=1}^{n} p_i q_i \tag{7.14}$$

for any $\vec{p}, \vec{q} \in \mathbb{R}^n$.

Consider some finite-dimensional real vector space V with inner product $\langle p, q \rangle$. For some constant $k \in \mathbb{R}$, a subset $S \subset V$ is said to be consistent if

$$\langle p, q \rangle \geq k \tag{7.15}$$

for all $p, q \in S$. As a matter of terminology, when $\langle p, q \rangle = k$, p and q are said to be maximally distant.

Observe that in this general definition of consistency, we omit the upper bound that is part of the definition in Eq. (7.12). Later we will see that the upper bound can be made to follow from a the boundedness of quantum state space [36].

A consistent set $S \subset V$ is said to be maximal if it is not contained in any larger consistent set in V, that is, for any $r \in V$ but $r \notin S$, there is at least one $q \in S$ such that

$$\langle q, r \rangle < k. \tag{7.16}$$

In other words, no further points in V can be added to S without breaking its consistency. An alternate definition of maximality would be that for any consistent set $S' \subset V$ containing S must be identical to S, i.e.,

$$S' \supseteq S \implies S' = S. \tag{7.17}$$

Next we consider some interesting examples of maximal consistent sets.

Example 7.2.1. Let $V = \mathbb{R}^n$ with the usual dot product and $k = -1$. The n-dimensional unit ball centered at the origin

$$\mathcal{B}^n = \{\vec{v} \in V \mid \|\vec{v}\| \leq 1\} \tag{7.18}$$

is a maximal consistent set.

Proof. To show that $\mathcal{B}^n$ is consistent, observe that since $\|\vec{p}\| \leq 1$ for any point from the ball, Cauchy-Schwarz inequality tells us that

$$\langle \vec{p}, \vec{q} \rangle \geq -\|\vec{p}\|\|\vec{q}\| \geq -1 \tag{7.19}$$

for any $\vec{p}, \vec{q} \in \mathcal{B}^n$.

To show that $\mathcal{B}^n$ is maximal, consider any $\vec{r} \in V$ but not in $\mathcal{B}^n$. This must mean that $\|\vec{r}\| > 1$. Now consider the vector

$$\vec{q} = -\frac{\vec{r}}{\|\vec{r}\|}. \tag{7.20}$$

Since

$$\|\vec{q}\| = \frac{\vec{q} \cdot \vec{q}}{\|\vec{q}\|^2} = 1 \tag{7.21}$$

so $\vec{q} \in \mathcal{B}^n$. But

$$\langle \vec{q}, \vec{r} \rangle = -\frac{\vec{r} \cdot \vec{r}}{\|\vec{r}\|} = -\|\vec{r}\| < -1. \tag{7.22}$$

Because $\vec{q} \in \mathcal{B}^n$ for any $\vec{r} \in V$ not in the ball, $\mathcal{B}^n$ is not a proper subset of any consistent set in V. Thus, $\mathcal{B}^n$ is maximal. □

Note that for antipodal points on the surface of the ball $\mathcal{B}^n$,

$$\vec{p} = -\vec{q}, \quad \|\vec{p}\| = 1 \implies \langle \vec{p}, \vec{q} \rangle = -1, \tag{7.23}$$

so $\vec{p}$ and $\vec{q}$ give the lower bound. It is because of this that we call points that satisfy the lower bound *maximally distant*.

The next two examples show that physically interesting state spaces can be defined as maximal consistent sets.

Example 7.2.2. Let $V \subset \mathbb{R}^n$ be the set of vectors $\vec{v}$ whose components in the standard basis sum to zero,

$$\sum_{i=1}^{n} v_i = 0. \tag{7.24}$$

With the consistency condition

$$\langle \vec{v}, \vec{w} \rangle = \vec{v} \cdot \vec{w} = -\frac{1}{n} \tag{7.25}$$

the set

$$S = \left\{ v \in V \;\middle|\; v_i + \frac{1}{n} \geq 0 \right\} \tag{7.26}$$

forms a maximal consistent set. Observe that the vectors $\vec{p}$

$$p_i = v_i + \frac{1}{n} \tag{7.27}$$

is the set of n-dimensional probability distributions so S is effectively the space of classical probabilities (it is equal to the regular simplex up to a translation.)

Proof. To show that S is consistent, take $\vec{v}, \vec{w} \in S$. Since

$$v_i + \frac{1}{n} \geq 0, \qquad w_i + \frac{1}{n} \geq 0, \tag{7.28}$$

for all i, we have

$$\sum_{i=1}^{n} \left(v_i + \frac{1}{n} \right) \left(w_i + \frac{1}{n} \right) \geq 0 \tag{7.29}$$

Because the components of $\vec{v}$ and $\vec{w}$ sum to zero,

$$\sum_i \left(v_i + \frac{1}{n}\right)\left(w_i + \frac{1}{n}\right) = \sum_i v_i w_i + \frac{1}{n}\left(\sum_i v_i + \sum_i w_i\right) + \sum_i \frac{1}{n^2} \tag{7.30}$$

$$= \langle \vec{v}, \vec{w} \rangle + \frac{1}{n}. \tag{7.31}$$

Using inequality 7.29, this implies that

$$\langle \vec{v}, \vec{w} \rangle \geq -\frac{1}{n}. \tag{7.32}$$

Thus S is consistent.

To show that S is maximal, choose any $\vec{u}$ in V but not in S. For at least one component of $\vec{u}$, it must be the case that

$$u_i + \frac{1}{n} < 0. \tag{7.33}$$

Without loss of generality, suppose this holds for u_1. Let $\vec{v} \in S$ be the vector with components

$$v_i = \delta_{1i} - \frac{1}{n}. \tag{7.34}$$

This leads to

$$\langle \vec{u}, \vec{v} \rangle = \left(1 - \frac{1}{n}\right) u_1 - \frac{1}{n} \sum_{i \neq 1} u_i. \tag{7.35}$$

Because $\vec{u} \in V$, its components add to zero so

$$u_1 = -\sum_{i \neq 1} u_i \tag{7.36}$$

and therefore,

$$\langle \vec{u}, \vec{v} \rangle = \left(1 - \frac{1}{n}\right) u_1 - \frac{1}{n} u_1 = u_1 < -\frac{1}{n}. \tag{7.37}$$

Thus, no set containing both $\vec{u}$ and $\vec{v}$ can be consistent. Since $\vec{v} \in S$ and $\vec{u}$ can be any element in V not in S this means that S is not contained in any larger consistent set in V. Therefore, S is maximal. □

Example 7.2.3. Let V be the space of traceless Hermitian operators A on $\mathbb{C}^d$,

$$A = A^\dagger, \qquad \mathrm{Tr}\,(A) = 0 \tag{7.38}$$

with inner product

$$\langle A, B\rangle = \mathrm{Tr}\,(AB)\,. \tag{7.39}$$

With consistency given by

$$\langle A, B\rangle \geq -\frac{1}{d} \tag{7.40}$$

then the set

$$S = \left\{ A \in V \;\middle|\; A + \frac{1}{d} I \geq 0 \right\} \tag{7.41}$$

forms a maximal consistent set. Observe that if we take

$$\rho = A + \frac{1}{d} I \tag{7.42}$$

for all A in S, we get the space of density operators, so S is effectively the set of quantum states.

Proof. First, we show S is consistent. For any $A, B \in S$

$$\mathrm{Tr}\left[\left(A + \frac{1}{d} I\right)\left(A + \frac{1}{d} I\right)\right] \geq 0 \tag{7.43}$$

since for any pair of positive semidefinite operators P, Q

$$\mathrm{Tr}\,(PQ) \geq 0. \tag{7.44}$$

Because A and B are traceless, we have

$$\mathrm{Tr}\left[\left(A + \frac{1}{d} I\right)\left(A + \frac{1}{d} I\right)\right] = \mathrm{Tr}\,(AB) + \frac{1}{d}\left(\mathrm{Tr}\,(A) + \mathrm{Tr}\,(B)\right) + \frac{1}{d^2}\mathrm{Tr}\,(I) \tag{7.45}$$

$$= \mathrm{Tr}\,(AB) + \frac{1}{d} \tag{7.46}$$

which in conjunction with inequality 7.43 implies that

$$\langle A, B\rangle = \mathrm{Tr}\,(AB) \geq -\frac{1}{d}. \tag{7.47}$$

Thus, S is consistent.

To show S is maximal, take any traceless Hermitian operator C not in S. Let $\{\hat{e}_j\}_{j=1}^n$ be an orthonormal basis in which C is diagonal, that is,

$$C\hat{e}_j = c_j \hat{e}_j \tag{7.48}$$

where c_j is an eigenvalue of C.

Because $C + \frac{1}{d}I$ is not positive semidefinite, at least one of the eigenvalues of C must be less than $-\frac{1}{d}$. Without loss of generality, suppose that $c_1 < -\frac{1}{d}$. Let $B \in S$ be the operator given by the eigenvalue equations

$$B\hat{e}_j = \left(\delta_{1j} - \frac{1}{d}\right)\hat{e}_j \tag{7.49}$$

for $j = 1, 2, ..., d$. Observe that

$$\langle B, C\rangle = \mathrm{Tr}\,(BC) = \left(1 - \frac{1}{d}\right)c_1 - \frac{1}{d}\sum_{j\neq 1} c_j \tag{7.50}$$

Since $\mathrm{Tr}\,(C) = 0$,

$$c_1 = -\sum_{j\neq 1} c_j \tag{7.51}$$

and thus

$$\langle B, C\rangle = \left(1 - \frac{1}{d}\right)c_1 + \frac{1}{d}c_1 = c_1 < -\frac{1}{d}. \tag{7.52}$$

Because $B \in S$ and C can be any element in V not in S, S does not belong to any larger consistent set in V. Thus, S is maximal. □

Example 7.2.4. Let $V = \mathbb{R}^n$ and choose some orthonormal basis $E = \{\hat{e_j}\}_{j=1}^n$ for V. Take the usual scalar product with consistency given by

$$\langle \vec{p}, \vec{q}\rangle = 0. \tag{7.53}$$

The set

$$S = \left\{ \vec{p} = \sum_{i=1}^{n} p_i \hat{e_i} \;\middle|\; p_i \geq 0 \right\} \tag{7.54}$$

forming the positive n-orthant of V in the basis E is a maximal consistent set [36].

Proof. To show S is consistent, let $\vec{p}, \vec{q} \in V$ with nonnegative components in E. Then

$$\langle \vec{p}, \vec{q}\rangle = \vec{p}\cdot\vec{q} = \sum_{i,j} p_i q_j \hat{e_i}\cdot\hat{e_j} = \sum_i p_i q_i \geq 0 \tag{7.55}$$

since $p_i, q_i \geq 0$ for all i. Thus, S is consistent.

To show S is maximal, suppose that $\vec{r}$ is a vector with at least one of its coordinates is negative. In that case,

$$\langle \vec{q}, \vec{r} \rangle < 0 \tag{7.56}$$

for at least one $\vec{q} \in E$, that is, if we write $\vec{r}$ in coordinate form $\vec{r} = (r_1, \ldots, r_n)$ then it will be inconsistent with one of the vectors

$$(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots (0, 0, \ldots, 1). \tag{7.57}$$

Since these vectors are elements of the basis E and belong in S, and $\vec{r}$ can be any vector in V that is not in S, then there is no larger consistent set in V that contains S. Thus, S is maximal. □

Example 7.2.5. Choose $k > c > 0$. Consider two concentric spheres of radii $r < R$ where

$$r = \frac{c}{\sqrt{k}}, \qquad R = \sqrt{k}. \tag{7.58}$$

Let $\hat{u}$ be some unit vector in $V = \mathbb{R}^n$ and

$$\begin{aligned} S &= \left\{ \vec{p} \in V \;\middle|\; \|\vec{p}\| = R, \langle \vec{p}, \hat{u} \rangle \geq \sqrt{\frac{k-c}{2}} \right\} \\ T &= \left\{ -\frac{c}{k}\vec{p} \;\middle|\; \vec{p} \in S \right\}. \end{aligned} \tag{7.59}$$

Then the convex hull of $S \cup T$ is a maximal consistent set.

Proof. First we show that $S \cup T$ is consistent. The consistency of the convex hull follows from the property that the convex hull of any consistent set is consistent, which we prove in Proposition 7.3.2.

Take any two points $\vec{p}, \vec{q} \in S$. We can express them in the form

$$\vec{p} = a\hat{u} + \sqrt{k - a^2}\hat{v} \qquad \vec{q} = b\hat{u} + \sqrt{k - b^2}\hat{w} \tag{7.60}$$

for unit vectors $\hat{v}$ and $\hat{w}$ orthogonal to $\hat{u}$, that is, we can decompose $\vec{p}$ and $\vec{q}$ into components parallel and perpendicular to direction $\hat{u}$. Since

$$\langle \vec{p}, \hat{u} \rangle \geq \sqrt{\frac{k-c}{2}}, \qquad \langle \vec{q}, \hat{u} \rangle \geq \sqrt{\frac{k-c}{2}} \tag{7.61}$$

then

$$a \geq \sqrt{\frac{k-c}{2}}, \quad b \geq \sqrt{\frac{k-c}{2}}, \quad \implies \quad ab \geq \frac{k-c}{2}. \tag{7.62}$$

We also have

$$k - a^2 \leq \frac{k+c}{2}, \quad k - b^2 \leq \frac{k+c}{2}, \tag{7.63}$$

which imply that

$$\sqrt{(k-a^2)\,(k-b^2)} \leq \frac{k+c}{2}. \tag{7.64}$$

Thus,

$$ab - \sqrt{(k-a^2)\,(k-b^2)} \geq \frac{k-c}{2} - \left(\frac{k+c}{2}\right) = -c. \tag{7.65}$$

Using the Cauchy-Schwarz inequality,

$$\begin{aligned} \langle \vec{p}, \vec{q} \rangle &= ab\,\langle \hat{u}, \hat{u} \rangle + \sqrt{(k-a^2)\,(k-b^2)}\,\langle \hat{v}, \hat{w} \rangle \\ &\geq ab - \sqrt{(k-a^2)\,(k-b^2)} \end{aligned} \tag{7.66}$$

which combined with inequality (7.65) gives

$$\langle \vec{p}, \vec{q} \rangle \geq -c. \tag{7.67}$$

Moreover, since $\|\vec{p}\| = \|\vec{q}\| = \sqrt{k}$, $\langle \vec{p}, \vec{q} \rangle \leq k$. Thus,

$$-c \leq \langle \vec{p}, \vec{q} \rangle \leq k \tag{7.68}$$

for any $\vec{p}, \vec{q}$ in S. Because any point in T is just $-\frac{c}{k}\vec{p}$ for some point $\vec{p} \in S$, then

$$-c \leq -\frac{c^3}{k^2} \leq \left\langle \left(-\frac{c}{k}\vec{p}\right), \left(-\frac{c}{k}\vec{q}\right) \right\rangle \leq \frac{c^2}{k^2}k = \frac{c^2}{k} \leq k \tag{7.69}$$

and

$$-c \leq \left\langle \left(-\frac{c}{k}\vec{p}\right), \vec{q} \right\rangle \leq \frac{c^2}{k} \leq k \tag{7.70}$$

since $\frac{c}{k} \leq 1$. Hence,

$$-c \leq \langle \vec{p}, \vec{q} \rangle \leq k \tag{7.71}$$

for any $\vec{p}, \vec{q} \in S \cup T$, and therefore $S \cup T$ is consistent. It follows from Proposition 7.3.2 that $\mathrm{Conv}(S \cup T)$ is consistent.

To show that $\mathrm{Conv}(S \cup T)$ is maximal, consider some point $\vec{x} \in V$. There is at least one 2-dimensional subspace containing both $\hat{u}$ and $\vec{x}$. The intersection of this plane with the convex hull of $S \cup T$ is shown in Fig. 7.2.

One of the following situations must hold:

(i) If $\vec{x}$ is within the blue region 1, there exists $\vec{y} \in S$ such that

$$\langle \vec{x}, \vec{y} \rangle < -c \tag{7.72}$$

so no set with both $\vec{x}$ and $\text{Conv}(S \cup T)$ can be consistent.

(ii) If $\vec{x}$ in within the green region 2, there exists $\vec{y} \in T$ such that

$$\langle \vec{x}, \vec{y} \rangle < -c \tag{7.73}$$

so no set with both $\vec{x}$ and $\text{Conv}(S \cup T)$ can be consistent.

(iii) If $\vec{x}$ is within the yellow region 3 on the left, it lies to the left of a line through the points $\vec{s}_a$ and $\vec{t}_b$. Since

$$\langle \vec{s}_a, \vec{s}_b \rangle = -c, \qquad \langle \vec{s}_b, \vec{t}_b \rangle = -c \tag{7.74}$$

it suggests that

$$\langle \vec{x}, \vec{s}_b \rangle < -c. \tag{7.75}$$

Any set containing both $\vec{x}$ and $\text{Conv}(S \cup T)$ can not be consistent.

(iv) If $\vec{x}$ is within the yellow region 4 on the right, it lies to the right of a line through the points $\vec{s}_b$ and $\vec{t}_a$. Since

$$\langle \vec{s}_a, \vec{s}_b \rangle = -c, \qquad \langle \vec{s}_a, \vec{t}_a \rangle = -c \tag{7.76}$$

it suggests that

$$\langle \vec{x}, \vec{s}_a \rangle < -c. \tag{7.77}$$

Any set containing both $\vec{x}$ and $\text{Conv}(S \cup T)$ can not be consistent.

Because any point outside of $\text{Conv}(S \cup T)$ must lie within one of the 4 regions described above, it implies that $\text{Conv}(S \cup T)$ can not belong to any larger consistent set in V. Thus, it is maximal.

□

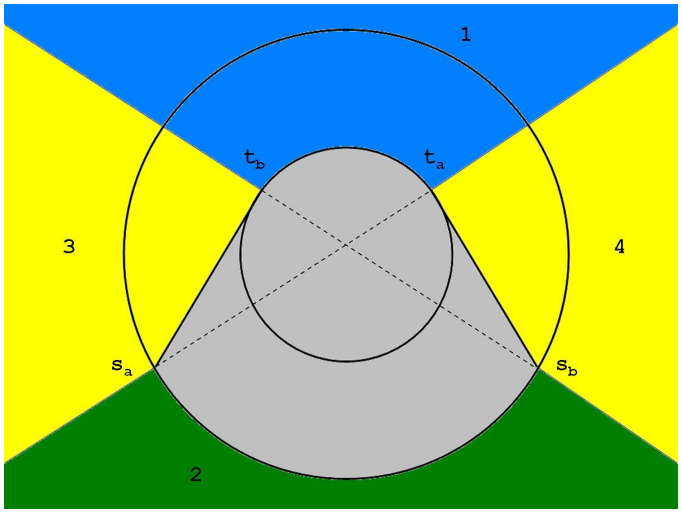


Figure 7.2: Gumdrop with an insphere of radius $\frac{c}{\sqrt{k}}$ and an outsphere of $\sqrt{k}$. The light gray area is a cross section of $\mathrm{Conv}(S \cup T)$.

7.3 Properties of maximal consistent sets

In this section, we prove several propositions concerning basic properties of general maximal consistent sets regarding their topology, convex structure, and boundary features [36].

Proposition 7.3.1. Every maximal consistent set is closed.

Proof. Let S be a maximal consistent set. Let $\vec{p}$ be a point in the closure of S. From the natural topology of the normed vector space V there exists a sequence $\vec{p}_1, \vec{p}_2, \ldots \subset S$ that converges to $\vec{p}$. Take any $\vec{q} \in S$. Consistency of S implies that

$$\langle \vec{q}, \vec{p}_j \rangle \geq k \tag{7.78}$$

The function $f_{\vec{q}} : V \to \mathbb{R}$ $(\vec{r} \mapsto f_{\vec{q}}(\vec{r}) = \langle \vec{q}, \vec{r} \rangle)$ is continuous so the sequence $\langle \vec{q}, \vec{p}_j \rangle$ converges to $\langle \vec{q}, \vec{p} \rangle$. Hence, $\langle \vec{q}, \vec{p} \rangle \geq k$.

Since $\langle \vec{q}, \vec{p} \rangle \geq k$ for any $\vec{q} \in S$, $S \cup \{\vec{p}\}$ is consistent. Because S is maximal, it must be that $\vec{p} \in S$. Since $\vec{p}$ can be any point in the closure of S, this means the closure of S is contained in S, that is, S is closed. □

Perhaps the most useful property of maximal consistent sets is the following:

Proposition 7.3.2. Every maximal consistent set is convex.

Proof. First we show that if R is consistent, then the convex hull $\mathrm{Conv}(R)$ of R is also consistent. Suppose $\vec{p}, \vec{q} \in \mathrm{Conv}(R)$. Then there exists $\vec{x}_1, \ldots, \vec{x}_M \in R$ and $\vec{y}_1, \ldots, \vec{y}_M \in R$ such that

$$\vec{p} = \sum_i \alpha_i \vec{x}_i, \qquad \vec{q} = \sum_j \beta_j \vec{y}_j \tag{7.79}$$

where $\sum_j \alpha_i = 1 = \sum_j \beta_j$ with $\alpha_i, \beta_j \geq 0$, that is, we can write $\vec{p}$ and $\vec{q}$ as a convex combination of elements in R. Because $\vec{x}_i$ and $\vec{y}_j$ are elements of R, they are consistent, i.e.,

$$\langle \vec{x}_i, \vec{y}_j \rangle \geq k. \tag{7.80}$$

Since $\alpha_i, \beta_j \geq 0$, it follows that $\alpha_i \beta_j \langle \vec{x}_i, \vec{y}_j \rangle \geq \alpha_i \beta_j k$. Thus,

$$\langle \vec{p}, \vec{q} \rangle = \sum_{i,j} \alpha_i \beta_j \langle \vec{x}_i, \vec{y}_j \rangle \geq \sum_{i,j} \alpha_i \beta_j k = k \tag{7.81}$$

so $\vec{p}$ and $\vec{q}$ are also consistent. The convexity of maximal consistent sets follows from the fact that given S is maximal, for any consistent subset $R \subseteq S$, then $\mathrm{Conv}(R) \subseteq S$. In particular, if $R = S$ then $\mathrm{Conv}(S) = S$. □

An example of a maximal consistent set that is unbounded would be the positive quadrant Q^{++} in $\mathbb{R}^2$ since

$$\forall \vec{p}, \vec{q} \in Q^{++}, \quad \langle \vec{p}, \vec{q} \rangle \geq 0. \tag{7.82}$$

Typically, we are interested only in bounded consistent sets since physically interesting sets such as classical and quantum state spaces are examples of bounded maximal consistent sets. It turns out that a simple condition is sufficient to guarantee that a consistent set is bounded:

Proposition 7.3.3. Let R be a consistent set. If R contains the ball $B = \{\vec{p} \in V | \ \|\vec{p}\| \leq C, C > 0\}$, then R is contained in the ball $B' = \left\{\vec{p} \in V | \|\vec{p}\| \leq \frac{k}{C}\right\}$.

Proof. First, observe that if R contains a ball, R must contain a pair of points whose inner product is negative. Therefore, k is negative.

Suppose $\vec{q} \in V$ with $\langle \vec{q}, \vec{q} \rangle > \left(\frac{k}{C}\right)^2$. If we write $\|\vec{q}\| = \sqrt{\langle q, q \rangle}$ then

$$\vec{q}' = -\frac{C}{\|\vec{q}\|}\vec{q}, \tag{7.83}$$

which is an element of R since

$$\langle \vec{q}', \vec{q}' \rangle = \frac{C^2}{\langle \vec{q}, \vec{q} \rangle} \langle \vec{q}, \vec{q} \rangle = C^2 \tag{7.84}$$

meaning it lies on the ball B contained in R. Because

$$\langle \vec{q}, \vec{q}' \rangle = -\frac{C}{\|\vec{q}\|} \langle \vec{q}, \vec{q} \rangle < -\frac{Ck^2}{\|\vec{q}\| C^2} = -\frac{k^2}{C\|\vec{q}\|} < k, \tag{7.85}$$

any set that contains both $\vec{q}$ and $\vec{q}'$ is inconsistent. Since $\vec{q}' \in R$, $\vec{q} \notin R$ for any $\vec{q} \in V$. Thus, R can not include any point $\vec{q}$ with $\|\vec{q}\| > \frac{k}{C}$. □

Proposition 7.3.4. Let S be a maximal consistent set. If S is contained in the ball $\{\vec{p} \in V |\ \|\vec{p}\| = C\}$ for $C > 0$ then the ball $\left\{\vec{p} \in V |\ \|\vec{p}\| = \frac{k}{C}\right\}$ is contained in S.

Proof. Let p be a point with $\langle \vec{p}, \vec{p} \rangle = \left(\frac{k}{C}\right)^2$. Take any $\vec{q} \in S$. By assumption $\langle \vec{q}, \vec{q} \rangle < C^2$. From the Cauchy-Schwarz inequality, we are guaranteed that

$$|\langle \vec{p}, \vec{q} \rangle|^2 \leq \langle \vec{p}, \vec{p} \rangle \langle \vec{q}, \vec{q} \rangle = \frac{k^2}{C^2}(C^2) = k^2 \tag{7.86}$$

Because $k < 0$, $\langle \vec{p}, \vec{q} \rangle \geq k$ for any $\vec{q} \in S$. Thus, $S \cup \{\vec{p}\}$ is a consistent set. However, S is maximal so $\vec{p} \in S$. The only condition we imposed on $\vec{p}$ is $\langle \vec{p}, \vec{p} \rangle = \frac{k^2}{C^2}$ so it follows that $\{\vec{p} \in V |\ \langle \vec{p}, \vec{p} \rangle = \frac{k^2}{C^2}\}$ is included in S. □

Proposition 7.3.5. Every maximal consistent set that includes any open neighborhood of the origin is bounded and compact.

Proof. Proposition 7.3.3 tells us that a consistent set which includes a ball around the origin is necessarily bounded. Proposition 7.3.4 suggests that a bounded maximal consistent set includes some finite ball around the origin. That takes care of the bounded part. Compactness follows from the fact that a maximal consistent set is closed and bounded in Euclidean space. □

The following proposition states that there is a minimum size to any maximal consistent set:

Proposition 7.3.6. No maximal consistent set can fit inside the ball

$$\left\{\vec{p} \in V |\ \|\vec{p}\| = C < \sqrt{|k|}\right\}. \tag{7.87}$$

Proof. Assume that S is a maximal consistent set contained in $\{\vec{p} \in V |\ \|\vec{p}\| < \sqrt{|k|}\}$. By Proposition 7.3.1, S is closed so

$$B = \max_{\vec{p} \in S} \langle \vec{p}, \vec{p} \rangle \tag{7.88}$$

is well-defined and S is contained in $\{\vec{p} \in V |\ \langle \vec{p}, \vec{p} \rangle = B\}$. It follows from our assumption that $B < |k|$. By Proposition 7.3.4, $\{\vec{p} \in V |\ \langle \vec{p}, \vec{p} \rangle = \frac{k^2}{C^2}\} \subset S$. But $C^2 < |k|$ implies that $\frac{k^2}{C^2} > |k| > B$, which is a contradiction. Thus, S can not be contained within the ball of radius C. □

The next few results concern boundary properties of maximal consistent sets.

Lemma 7.3.1. Let S be bounded maximal consistent set and $\vec{p} \in V$. If $\langle \vec{p}, \vec{q} \rangle > k$ for all $\vec{q} \in S$ then there exists $\epsilon > 0$ such that $\langle \vec{p}, \vec{q} \rangle \geq k + \epsilon$.

Proof. The function $f_{\vec{p}} : S \to \mathbb{R}$ $(\vec{q} \mapsto \langle \vec{p}, \vec{q} \rangle)$ is a continuous function with a compact domain. Therefore, $f_{\vec{p}}$ has a minimum value. If $f_{\vec{p}}(\vec{q}) > k$ for all $\vec{q} \in S$ then $\min f_{\vec{p}}(\vec{q}) \geq k + \epsilon$ for all $\vec{q} \in S$ then we can always choose

$$\epsilon = \frac{\min f_{\vec{p}} - k}{2}. \tag{7.89}$$

□

Lemma 7.3.2. Let S be a bounded maximal consistent set and let $\vec{p} \in S$. If there exists $\epsilon > 0$ such that $\langle \vec{p}, \vec{q} \rangle \geq k + \epsilon$ for all $\vec{q} \in S$ then $\vec{p}$ is not in the boundary of S.

Proof. Suppose there is some $\epsilon > 0$ such that $\langle \vec{p}, \vec{q} \rangle \geq k + \epsilon$ for all $\vec{q} \in S$. Then

$$\langle \vec{p} + \vec{v}, \vec{q} \rangle \geq k + \epsilon + \langle \vec{v}, \vec{q} \rangle \tag{7.90}$$

for all $\vec{q} \in S$ and $\vec{v} \in V$. Because S is bounded there exists $M > 0$ such that $\langle \vec{q}, \vec{q} \rangle \geq M^2$. By the Cauchy-Schwarz inequality

$$|\langle \vec{v}, \vec{q} \rangle|^2 \leq \langle \vec{v}, \vec{v} \rangle \langle \vec{q}, \vec{q} \rangle = \epsilon^2 \tag{7.91}$$

for $\vec{v} \in V$ such that $\langle \vec{v}, \vec{v} \rangle = \left(\frac{\epsilon}{M}\right)^2$. Thus $\langle \vec{p} + \vec{v}, \vec{q} \rangle \geq k$ for all $\vec{v}$ with $\langle \vec{v}, \vec{v} \rangle = \left(\frac{\epsilon}{M}\right)^2$. Hence S contains an open ball of radius $\frac{\epsilon}{M}$ centered on $\vec{p}$. Therefore, $\vec{p}$ is not on the boundary of S. □

Proposition 7.3.7. Let S be a maximal consistent set and $\vec{p} \in S$. The point $\vec{p}$ is on the boundary of S if and only if there exists $\vec{q} \in S$ such that $\langle \vec{p}, \vec{q} \rangle = k$.

Proof. Suppose $\vec{p}$ is on the boundary of S. By Lemma 7.3.2, there is no $\epsilon > 0$ such that $\langle \vec{p}, \vec{q} \rangle \geq k + \epsilon$. By Lemma 7.3.1, $\langle \vec{p}, \vec{q} \rangle > k$ does not hold for all $\vec{q} \in S$. However, since S is consistent, it must be that $\langle \vec{p}, \vec{q} \rangle \leq k$. Therefore, there exists $\vec{q} \in S$ such that $\langle \vec{p}, \vec{q} \rangle = k$.

Now suppose there is $\vec{q} \in S$ such that $\langle \vec{p}, \vec{q} \rangle = k$. Any open neighborhood of $\vec{p}$ contains $(1+\delta)\vec{p}$ for some $\delta > 0$. Thus $\langle (1+\delta)\vec{p}, \vec{q} \rangle = (1+\delta)k < k$ since $k < 0$. So $(1+\delta)\vec{p} \notin S$. Hence, every open neighborhood of $\vec{p}$ contains a point outside of S. Since every open neighborhood of $\vec{p}$ also contains a point in $\vec{p}$, namely $\vec{p}$ itself, then $\vec{p}$ is on the boundary of S. □

Corollary 7.3.3. Let S be a maximal consistent set and $\vec{p}, \vec{q} \in S$. If $\langle \vec{p}, \vec{q} \rangle = k$ then $\vec{p}, \vec{q}$ are on the boundary of S.

Proposition 7.3.8. Let S be a maximal consistent set and $\vec{p} \in S$. If $\|\vec{p}\| = \max_{\vec{q} \in S} \|\vec{q}\|$ then both $\vec{p}$ and $\frac{k}{\|\vec{p}\|^2}\vec{p}$ are on the boundary of S.

Proof. By Proposition 7.3.4, S contains the ball $\left\{ \vec{q} \in V |\ \|\vec{q}\| < \frac{k}{\|p\|} \right\}$ Therefore, $\frac{k}{\|\vec{p}\|^2}\vec{p} \in S$. Since $\left\langle \frac{k}{\|\vec{p}\|^2}\vec{p}, \vec{p} \right\rangle = k$, Corollary 1 implies that $\vec{p}$ and $\frac{k}{\|\vec{p}\|^2}\vec{p}$ are on the boundary of S. □

The next propositions indicate that a suitably placed origin for a consistent set determines the value for $\|\vec{p}\|$ for $\vec{p} \in S$:

Proposition 7.3.9. Let R be a consistent set with $\langle \vec{p}, \vec{q} \rangle \geq k, k < 0$. If for some $\vec{p}, \vec{q} \in R$ the origin $\mathcal{O} = \alpha\vec{p} + (1-\alpha)\vec{q}$ for $\alpha \in [0, \frac{1}{2})$ then $\vec{p}, \vec{q} \in \left\{ \vec{r} \in V |\ \|\vec{r}\| \leq \sqrt{\frac{|k|(1-\alpha)}{\alpha}} \right\}$.

Proof. For the origin, $\|\mathcal{O}\|^2 = 0$ so

$$\begin{aligned} 0 &= \langle \alpha\vec{p} + (1-\alpha)\vec{q}, \alpha\vec{p} + (1-\alpha)\vec{q} \rangle \\ &= \alpha^2 \|\vec{p}\|^2 + 2\alpha(1-\alpha) \langle \vec{p}, \vec{q} \rangle + (1-\alpha)^2 \|\vec{q}\|^2 \end{aligned} \tag{7.92}$$

which means that

$$\alpha^2 \|\vec{p}\|^2 + (1-\alpha)^2 \|\vec{q}\|^2 \leq -2\alpha(1-\alpha) \langle \vec{p}, \vec{q} \rangle . \tag{7.93}$$

But $\langle \vec{p}, \vec{q} \rangle \geq k \implies -\langle \vec{p}, \vec{q} \rangle \leq 2|k|$ so

$$\alpha^2 \|\vec{p}\|^2 + (1-\alpha)^2 \|\vec{q}\|^2 \leq 2|k|\alpha(1-\alpha). \tag{7.94}$$

Let $0 \leq \alpha < \frac{1}{2}$. Since $\alpha\vec{p} = -(1-\alpha)\vec{q}$, we can write

$$\|\vec{p}\| = \left|\frac{1-\alpha}{\alpha}\right| \|\vec{q}\| \tag{7.95}$$

and so we get

$$\|\vec{q}\|^2 \leq \frac{|k|\alpha}{(1-\alpha)} \qquad \text{and} \qquad \|\vec{p}\|^2 \leq \frac{|k|(1-\alpha)}{\alpha}. \tag{7.96}$$

□

Proposition 7.3.10. Let S be a consistent set with $\langle \vec{p}, \vec{q} \rangle \geq k$ for all $\vec{p}, \vec{q} \in S$.If

$$\{\vec{p}_i\} +_{i=1}^{N+1} \subset S \tag{7.97}$$

with $\langle \vec{p}_i, \vec{p}_j \rangle = C$ when $i \neq j$, for $k \leq C < 0$ and $\sum_{i=1}^{N+1} \vec{p}_i = 0$, then $\|\vec{p}_i\| = \sqrt{N|C|}$.

Proof. For any $j \neq k$, $\langle \vec{p}_j, \vec{p}_k \rangle = C$. Note that

$$-\vec{p}_j = \sum_{i=1, i\neq j}^{N+1} \vec{p}_i. \tag{7.98}$$

Thus,

$$\begin{aligned} \langle \vec{p}_j, \vec{p}_k \rangle &= -\sum_{i=1, i\neq j}^{N+1} \langle \vec{p}_i, \vec{p}_k \rangle \\ &= -\langle \vec{p}_k, \vec{p}_k \rangle - \sum_{i=1, i\neq j, i\neq k}^{N+1} \langle \vec{p}_i, \vec{p}_k \rangle \end{aligned} \tag{7.99}$$

implies that

$$\|\vec{p}_k\|^2 = -C - (N+1-2)C = N|C|. \tag{7.100}$$

□

7.4 Maximization of consistent sets: examples

We want to consider the following problem. Suppose we are given some consistent set $R \subset V$ with consistency given by $\langle p, q \rangle > k$, for any $p, q \in R$. We want to add all points

$v_i \in V$ such that $\langle r, q\rangle \geq k$ for all points in $S \cup \{v_1, v_2, \ldots\}$, that is, we add all points in V such that S is a maximal consistent set. We denote this process by $S = \text{Max}(R)$. Is it then possible to characterize the types of maximal consistent sets you can obtain from consistent sets with a finite number of points? For instance, are there cases where the consistent set leads to a unique maximization? Here we want describe a few examples showing that this is indeed the case, which opens the possibility that quantum state space itself is a maximization of a suitably chosen initial consistent set. If we can find the simplest consistent set then it means that the essential geometric aspects of the set of quantum states is embodied by the symmetries possessed by the initial consistent set. But before we get to that, it is useful to have a better understanding of what it means to maximize a consistent set, which is, in fact, a highly nontrivial task.

The first example we consider is a consistent set with two vectors whose inner product defines the consistency condition. It demonstrates that, for identifying the maximization, it is often crucial to wisely choose the origin and the orientation of coordinates.

Example 7.4.1. Let $V = \mathbb{R}^2$ and $R = \{\vec{a}, \vec{b}\}$ is a consistent set with

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{and} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{7.101}$$

and consistency is given by $k = \left\langle \vec{a}, \vec{b} \right\rangle$. The midpoint of $\vec{a}$ and $\vec{b}$ is given by

$$\vec{m} = \begin{pmatrix} \frac{a_1+b_1}{2} \\ \frac{a_2+b_2}{2} \end{pmatrix}. \tag{7.102}$$

Performing the translation $\vec{x} \mapsto \vec{x}' = \vec{x} - \vec{m}$ so that

$$\vec{m} \mapsto \vec{m}' = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \vec{a} \mapsto \vec{a}' = \begin{pmatrix} \frac{b_1-a_1}{2} \\ \frac{b_2-a_2}{2} \end{pmatrix}, \quad \vec{b} \mapsto \vec{b}' = \begin{pmatrix} \frac{a_1-b_1}{2} \\ \frac{a_2-b_2}{2} \end{pmatrix}, \tag{7.103}$$

we get the translated consistent set R' with consistency given by $k' = \left\langle \vec{a}', \vec{b}' \right\rangle$. Furthermore, we can rotate coordinates using

$$O = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \text{where} \quad \tan\theta = \frac{a_2'}{a_1'} = \frac{b_2 - a_2}{b_1 - a_1}. \tag{7.104}$$

Note that the minus sign for the rotation matrix is on the second row since we are performing a passive transformation. After the rotation, we obtain the consistent set R'' with elements

$$\begin{aligned} O\vec{m}' &= \vec{m}', \\ \vec{a}'' = O\vec{a}' &= \begin{pmatrix} s \\ 0 \end{pmatrix}, \\ \vec{b}'' = O\vec{b}' &= \begin{pmatrix} -s \\ 0 \end{pmatrix}, \end{aligned} \tag{7.105}$$

where

$$s = \sqrt{\langle \vec{a}', \vec{a}' \rangle} = \sqrt{\left\langle \vec{b}', \vec{b}' \right\rangle} = \frac{1}{4} \left[(x_1 - x_2)^2 + (y_1 - y_2)^2 \right]. \tag{7.106}$$

Hence, the consistency condition for $\mathrm{Max}\,(R'')$ is given by

$$-s^2 \leq \langle \vec{p}'', \vec{q}'' \rangle \leq s^2, \qquad \forall \vec{p}'', \vec{q}'' \in \mathrm{Max}\,(R'')\,. \tag{7.107}$$

The upper bound just follows from the fact that $\vec{a}''$ and $\vec{b}''$ are equidistant from the origin and maximally distant from each other.

Suppose

$$\vec{p}'' = \begin{pmatrix} s\cos\alpha \\ s\sin\alpha \end{pmatrix}, \qquad \text{and} \qquad \vec{q}'' = \begin{pmatrix} s\cos\beta \\ s\sin\beta \end{pmatrix}. \tag{7.108}$$

Then we have

$$\begin{aligned} \langle \vec{p}'', \vec{q}'' \rangle &= s^2 \cos\alpha\cos\beta + s^2 \sin\alpha\sin\beta \\ &= s^2 \cos(\alpha - \beta). \end{aligned} \tag{7.109}$$

Since $-1 \leq \cos(\alpha - \beta) \leq 1 \quad \forall \alpha, \beta \in \mathbb{R}$

$$-s^2 \leq \langle \vec{p}'', \vec{q}'' \rangle \leq s^2 \tag{7.110}$$

for all such $\vec{p}'', \vec{q}''$. This means that the set of points

$$T = \left\{ (\vec{x}, \vec{y}) |\ \vec{x}^2 + \vec{y}^2 = s^2 \right\} \tag{7.111}$$

is consistent with R''. Moreover, any such point $\vec{p}$ has $\langle \vec{p}, \vec{p} \rangle = s^2$, so it must lie on the boundary of $\mathrm{Max}\,(R'')$.

Since T is consistent, $\mathrm{Conv}(T)$ is consistent. Thus,

$$\mathrm{Conv}(T) \subseteq \mathrm{Max}\left(R''\right). \tag{7.112}$$

To show that they are in fact equal, consider any point $\vec{r}$ lying outside $\mathrm{Conv}(T)$. If you find the point $\vec{t}$ on T that is closest to $\vec{r}$, it is easy to show that the point opposite to that on the circle will be inconsistent with $\vec{r}$, i.e. $\left\langle \vec{r}, -\vec{t} \right\rangle < -s^2$.

Of course, to get $\mathrm{Max}(R)$, just undo the rotation and translation on $\mathrm{Max}\left(R''\right)$. The generalization for a pair of points from $\mathbb{R}^d$ is that the maximal consistent set is a ball where the two points lie on the surface.

Next we consider the vertices of a regular simplex. If the consistency condition is chosen appropriately, the simplex itself can be considered as a maximal consistent set.

Example 7.4.2. Suppose we have for pairwise maximally distant probability vectors $\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4$ from the 3-simplex. Without loss of generality, these probability vectors can be specified with coordinates

$$\vec{p}_1 = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \qquad \vec{p}_2 = \begin{pmatrix} b \\ c \\ d \\ a \end{pmatrix}, \qquad \vec{p}_3 = \begin{pmatrix} c \\ d \\ a \\ b \end{pmatrix}, \qquad \vec{p}_4 = \begin{pmatrix} d \\ a \\ b \\ c \end{pmatrix}, \tag{7.113}$$

and since these must be probabilities,

$$a + b + c + d = 1 \tag{7.114}$$

for $a, b, c, d \geq 0$, and with the added constraint

$$(ab + bc + cd + da) = 2(ac + bd) \tag{7.115}$$

to ensure that pairwise inner products are equal. Using these coordinates, consistency is given by

$$\langle \vec{p}, \vec{q} \rangle \geq ab + bc + cd + da. \tag{7.116}$$

Performing the same manipulations as before, we translate the origin to

$$\vec{m} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}, \tag{7.117}$$

which would be equivalent to the translation $\vec{p}_i \mapsto \vec{q}_i = \vec{p}_i - \vec{m}$ for $i = 1, 2, 3, 4$. If we denote the squared norm of $\vec{q}_j$ by $N = \langle \vec{q}_j, \vec{q}_j \rangle$ for any j, it is easy to verify that

$$\langle \vec{q}_i, \vec{q}_j \rangle = -\frac{N}{3} \tag{7.118}$$

for any distinct pair $i, j = 1, 2, 3, 4$. Now consider the matrix

$$Q = \begin{pmatrix} \langle q_1, q_1 \rangle & \langle q_1, q_2 \rangle & \langle q_1, q_3 \rangle & \langle q_1, q_4 \rangle \\ \langle q_2, q_1 \rangle & \langle q_2, q_2 \rangle & \langle q_2, q_3 \rangle & \langle q_2, q_4 \rangle \\ \langle q_3, q_1 \rangle & \langle q_3, q_2 \rangle & \langle q_3, q_3 \rangle & \langle q_3, q_4 \rangle \\ \langle q_4, q_1 \rangle & \langle q_4, q_2 \rangle & \langle q_4, q_3 \rangle & \langle q_4, q_4 \rangle \end{pmatrix} = \frac{N}{3} \begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}. \tag{7.119}$$

The matrix Q has exactly 1 zero eigenvalue corresponding to eigenvector $(1, 1, 1, 1)^T$. Note that the linear dependency condition

$$\sum_j a_j \vec{q}_j = 0 \tag{7.120}$$

holds if and only if the vector

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} \tag{7.121}$$

is a zero eigenvector of Q. Therefore, the space spanned by $\{\vec{q}_1, \vec{q}_2, \vec{q}_3, \vec{q}_4\}$ is exactly 3-dimensional. Coupled with the fact that the angles between any pair of vectors $\vec{q}_i$ are the same, these vectors point to vertices of a regular tetrahedron. To see this, it may be convenient to rotate the coordinates of $\vec{q}_j$ such that

$$\vec{q}_1 = \sqrt{N} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{q}_2 = \frac{\sqrt{N}}{3} \begin{pmatrix} \sqrt{8} \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{q}_3 = \frac{\sqrt{N}}{3} \begin{pmatrix} -\sqrt{2} \\ \sqrt{6} \\ -1 \\ 0 \end{pmatrix}, \quad \vec{q}_4 = \frac{\sqrt{N}}{3} \begin{pmatrix} -\sqrt{2} \\ -\sqrt{6} \\ -1 \\ 0 \end{pmatrix}. \tag{7.122}$$

From here, it is straightforward to show that on the sphere S given by

$$x_1^2 + x_2^2 + x_3^2 = N, \qquad x_4 = 0. \tag{7.123}$$

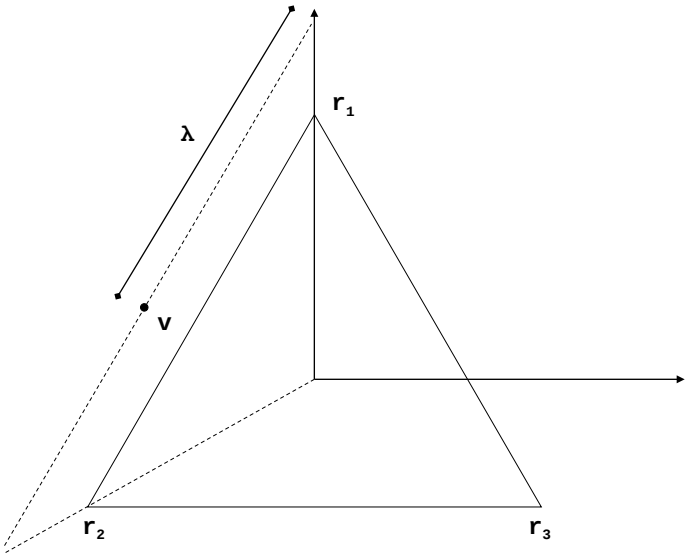


Figure 7.3: Regular triangle on a plane as a maximal consistent set.

the only points consistent with the set consisting of the origin and the vectors $\vec{q}_j$ are the $\vec{q}_j$ themselves. This can be done by choosing one particular point, say $\vec{q}_1$ and showing that a point within the sphere S is consistent with q_1 if it lies on the same half-space defined by the plane passing through $\vec{q}_2, \vec{q}_3$, and q_4. The exact same argument works for finding points consistent with $\vec{q}_2, \vec{q}_3$, and $\vec{q}_4$ and so as we add all such consistent points, we are left with a set composed of points on and inside the tetrahedron inscribed in the sphere S. This is a maximal consistent set.

We can show that the above situation generalizes to all regular simplices, or more generally, the N-dimensional convex polytope formed from the convex hull of $N+1$ equidistant points is a bounded maximal consistent set in $\mathbb{R}^N$. To illustrate the proof technique, first we discuss the example of a regular triangle in $\mathbb{R}^2$.

Example 7.4.3. Let $V = \mathbb{R}^2$ and consider the consistent set $R \subset V$ with elements given by

$$\vec{r}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \vec{r}_2 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix} \qquad \vec{r}_3 = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{7.124}$$

Let consistency be given by $\langle \vec{r}_i, \vec{r}_j \rangle = k = -\frac{1}{2}$ when $i \neq j$. We will show that the maximization of R is just the convex hull of R.

Proof. Consistency of $\text{Conv}(R)$ follows naturally from the consistency of R. To show maximality, suppose we consider some $\vec{v} \in V$ where $\vec{v}$ is not inside the convex hull of R. For any such $\vec{v}$, we can write

$$v = \lambda\alpha\vec{p}_i + (1-\lambda)\beta\vec{p}_j \tag{7.125}$$

where $i, j = 1, 2, 3$, $i \neq j$, $\alpha, \beta \geq 1$, and $0 \leq \lambda \leq 1$. For example, for the particular $\vec{v}$ depicted in Fig. 7.3, it can be written as the convex combination of rescaled vectors $\vec{r}_1$ and $\vec{r}_2$.

Thus, upon maximizing R, we require that

$$\langle \vec{v}, \vec{p}_k \rangle = \lambda\alpha \langle \vec{p}_i, \vec{p}_k \rangle + (1-\lambda)\beta \langle \vec{p}_j, \vec{p}_k \rangle \geq k \tag{7.126}$$

for all $k = 1, 2, 3$. However, when $i \neq j \neq k$, we get

$$\langle \vec{v}, \vec{p}_k \rangle = [\lambda\alpha + (1-\lambda)\beta]\, k \leq k \tag{7.127}$$

since $k < 0$ and $\lambda\alpha + (1-\lambda)\beta \geq 1$, with equality if and only if $\alpha = \beta = 1$. However, if that holds, then it would mean that $\vec{v} \in \text{Conv}(R)$, which is a contradiction. Therefore, if $\vec{v} \in V$ is not in the convex hull of R, it will be inconsistent with one of the elements of R. □

It is not hard to see how procedure generalizes to higher dimensions:

Theorem 7.4.1. Let $R \subset V$ be the consistent set

$$R = \left\{ \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_{N+1} \mid \sum_{i=1}^{N+1} r_i = 0 \right\}, \tag{7.128}$$

where consistency is given by$\langle \vec{r}_i, \vec{r}_j \rangle = k$ for all $i \neq j$ and $V = \text{Span}(R)$. Then $\text{Conv}(R)$ is the maximal consistent set $\text{Max}(R)$ obtained from maximizing R.

Proof. Consistency follows trivially from the consistency of R. Maximality can be shown by contradiction as follows: suppose $\vec{v} \in V$ and $\vec{v} \notin \text{Conv}(R)$. Because R spans V, there exists sets of N numbers $\{\lambda_i\}_{i=1}^N$ and $\{\alpha_i\}_{i=1}^N$ such that

$$\vec{v} = \sum_{i \neq j} \lambda_i \left(\alpha_i \vec{r}_i\right) \tag{7.129}$$

where

$$0 \leq \lambda_i \leq 1, \qquad \sum_{i \neq j} \lambda_i = 1 \tag{7.130}$$

and $\alpha_i \geq 1$. This merely tells us that can write $\vec{v}$ as a convex combination of N stretched vectors in R. Maximizing R by adding all such consistent $\vec{v}$, we have

$$\langle \vec{v}, \vec{r}_k \rangle = \sum_{i \neq j} \lambda_i \langle \alpha_i \vec{r}_i, \vec{r}_k \rangle \geq k. \tag{7.131}$$

for all $k = 1, 2, ..., N+1$. However, we note that in the case $k = j$, we obtain

$$\langle \vec{v}, \vec{r}_j \rangle = \sum_{i \neq j} \lambda_i \langle \alpha_i \vec{r}_i, \vec{r}_j \rangle = \left[\sum_{i \neq j} \lambda_i \alpha_i \right] k \leq k \tag{7.132}$$

with equality if and only if $\alpha_i = 1$ for all i. But that happens only when $\vec{v} \in \text{Conv}(R)$, which contradicts our initial assumption on $\vec{v}$. Therefore, $\vec{v} \in V$ is included in $\text{Max}(R)$ if and only if it is in the convex hull of R. □

The important lesson to take away from the examples of this section is that for consistent sets with the appropriate consistency condition, there corresponds a unique maximal consistent set. This suggests the possibility of constructing quantum state space from the maximization of a suitable consistent set.

7.5 Qutrit state space as a maximal consistent set

In this section, we sketch some of the details for constructing quantum state space in $d = 3$ as the maximization of some initial consistent set. As a matter of convenience, we adopt the bracket notation, where $|p\rangle$ and $|q\rangle$ represent two vectors in $\mathbb{R}^{d^2}$ and the bracket $\langle p|q\rangle$ is just the dot product between the two real vectors. For qutrits, consistency between SIC probability vectors given by

$$\frac{1}{12} \leq \langle p|q\rangle \leq \frac{1}{6}. \tag{7.133}$$

Let $\{|j\rangle\}_{j=1}^{d^2}$ form the standard basis in $\mathbb{R}^{d^2}$. To start, we consider the consistent set $\mathcal{S}_0 \subset \Delta^{d^2-1}$ of probability vectors,

$$\mathcal{S}_0 = \{|E_k\rangle, |Z_{ijk}\rangle \mid i, j, k = 1, 2, \ldots, 9 \text{ for } i \neq j \neq k\}, \tag{7.134}$$

where the $|E_k\rangle$ are called the basis distributions

$$|E_k\rangle = \sum_{j=1}^{9} \frac{3\delta_{jk} + 1}{12} |j\rangle = \frac{1}{12}(1, ..., \overbrace{4}^{k^{th}}, ...1), \tag{7.135}$$

which are just the probability vectors generated from the SIC elements by the SIC representation, and the $|Z_{rst}\rangle$ are uniform maximal-zero distributions

$$|Z_{rst}\rangle = \sum_{j=1}^{9} \frac{1}{9-\beta}(1-\delta_{rj}-\delta_{sj}-\delta_{tj})\,|j\rangle\,, \qquad \text{for } r \neq s \neq t \tag{7.136}$$

where β is the maximum number of zero-valued components allowed, which happens to $\beta = 3$ in this case. Writing $|Z_{rst}\rangle$ out explicitly gives

$$|Z_{rst}\rangle = \frac{1}{6}(1, .., \overbrace{0}^{r^{th}}, \ldots, \overbrace{0}^{s^{th}}, \ldots, \overbrace{0}^{t^{th}}, \ldots, 1) \tag{7.137}$$

What we want is a recipe for building up a maximal consistent set from $\mathcal{S}_0$.

First, we add all the probability vectors $|Q^r_{st}\rangle$ such that is maximally distant with both $|E_r\rangle$ and $|Z_{rst}\rangle$:

$$\langle E_r|Q^r_{st}\rangle = \langle Z_{rst}|Q^r_{st}\rangle = \frac{1}{12}. \tag{7.138}$$

This is possible because if we consider the corresponding vectors in Hilbert space, this just says that for every pair of orthogonal pure states, there is another pure state such that the three unit vectors form an orthonormal basis.

For example, if we take $|E_1\rangle$ and $|Z_{123}\rangle$ then it turns out that there is only one choice for a third point $|Q^1_{23}\rangle$, which is

$$\left|Q^1_{23}\right\rangle = \frac{1}{12}(0, 3, 3, 1, 1, 1, 1, 1, 1) \tag{7.139}$$

where our notation is meant to suggest that it is maximally distant to the probability vectors $|E_1\rangle$ and $|Z_{123}\rangle$. We say that $|E_1\rangle$, $|Z_{123}\rangle$ and $|Q^1_{23}\rangle$ form a *maximally distant triplet*.In geometric terms, $|Q^1_{23}\rangle$ is the third point along with $|Z_{123}\rangle$ and $|E_1\rangle$ that forms an equilateral triangle inscribed in the out-sphere of radius $\sqrt{\frac{2}{d(d+1)}} = \frac{1}{\sqrt{6}}$ and there can only be one such point.

We want to include all such probability vectors for all possible maximally distant pairs $|E_r\rangle$ and $|Z_{rst}\rangle$. We find that the vectors $|Q^r_{st}\rangle$ come with a generic form: its r^{th} entry is zero, the s^{th} and t^{th} entries are both $\frac{1}{4}$ and all remaining entries are $\frac{1}{12}$.

However, in order to maintain consistency in the new set $\mathcal{S}_1 = \mathcal{S}_0 \cup \{|Q^r_{st}\rangle\}$, the $|Z_{rst}\rangle$ that are inconsistent with the $|Q\rangle^r_{st}$ have to be excluded. For instance, in the example with $|E_1\rangle\,, |Z_{123}\rangle\,,$ and $|Q^1_{23}\rangle$, we observe that

$$\langle Q^1_{23}|Z_{23k}\rangle = \frac{5}{72} < \frac{1}{12}, \qquad k = 4, 5, ..., 9 \tag{7.140}$$

so the price for including $|Q^1_{23}\rangle$ to S while retaining consistency is to remove 6 uniform maximal-zero distributions from the set.

If we keep on adding as many $|Q^r_{st}\rangle$ as we can, how many $|Z_{rst}\rangle$ remain in the new set $\mathcal{S}_1$? There are a number of different ways to get at the answer but by a simple counting argument, we can see that we will be left with 12 pairs before all $|Z_{rst}\rangle$ are exhausted. As we saw in the example above, adding $|Q^r_{st}\rangle$ requires removing six points $|Z_{kst}\rangle, k \neq r, s, t$, it follows that since there are

$$ {}^9\mathrm{C}_3 = \binom{9}{3} = 84 \tag{7.141}$$

uniform maximal-zero distributions $|Z_{rst}\rangle$ and each step of adding $|Q^r_{st}\rangle$ involves 7 such vectors, only one of which is kept in the new set, then we can add up to $84 \div 7 = 12$ $|Q_{rst}\rangle$ before we end up considering all $|Z_{rst}\rangle$. The actual 12 that are left the order in which we consider pairs $|E_r\rangle$ and $|Z_{rst}\rangle$ but whichever set we end up with will be related to another possible set by a permutation of indices.

It is always possible to perform the procedure such that the 12 maximal-zero distributions that remain are

$$|Z_{123}\rangle, |Z_{456}\rangle, |Z_{789}\rangle, |Z_{147}\rangle, |Z_{258}\rangle, |Z_{369}\rangle, |Z_{159}\rangle, |Z_{267}\rangle, |Z_{348}\rangle, |Z_{168}\rangle, |Z_{249}\rangle, |Z_{357}\rangle, \tag{7.142}$$

that is, $|Z_{rst}\rangle$ with index triples (rst) that correspond to the affine lines in

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \tag{7.143}$$

which is just the sort of finite affine-plane structure we have encountered before.

So each pair $|E_r\rangle, |Z_{rst}\rangle$ allows us to add $|Q^r_{st}\rangle$. Certainly, we could have instead considered $|E_s\rangle, |Z_{rst}\rangle$ which would tell us to add $|Q^s_{tr}\rangle$. For example,

$$\begin{aligned} \langle E_1|Q^2_{31}\rangle &= \left[\frac{1}{12}(4,1,1,1,1,1,1,1,1,1)\right] \cdot \left[\frac{1}{12}(3,0,3,1,1,1,1,1,1,1)\right] \\ &= \frac{1}{144}(12+3+6) = \frac{21}{144} = \frac{7}{48} > \frac{1}{12}. \end{aligned} \tag{7.144}$$

We may also consider pairs $|E\rangle_q$ and $|Z\rangle_{rst}$ that do not share any indices. As an example,

$$\begin{aligned} \langle E_5|Q^2_{31}\rangle &= \left[\frac{1}{12}(1,1,1,1,3,1,1,1,1)\right] \cdot \left[\frac{1}{12}(3,0,3,1,1,1,1,1,1,1)\right] \\ &= \frac{1}{144}(3+3+3+5) = \frac{14}{144} = \frac{7}{72} > \frac{1}{12}. \end{aligned} \tag{7.145}$$

It is not difficult to see that in general

$$\langle E_r|Q^r_{st}\rangle = \frac{1}{12}, \quad \langle E_s|Q^r_{st}\rangle = \frac{7}{48}, \quad \langle E_q|Q^r_{st}\rangle = \frac{7}{72} \qquad \text{for distinct } q,r,s,t = 1,2,...,9. \tag{7.146}$$

and $\frac{1}{12} < \frac{7}{72} < \frac{7}{48} < \frac{1}{6}$ so we can add all such $|Q^r_{st}\rangle$ and the new set remains consistent. If we do so, the consistent set we end up with is

$$\mathcal{S}' = \{|E_r\rangle, |Z_{rst}\rangle, |Q^r_{st}\rangle \mid r,s,t = 1,2,...,9\} \tag{7.147}$$

where if we count elements we have 9 basis distributions, 12 uniform maximal-zero distributions, and 36 $|Q^r_{st}\rangle$.

To see all the maximal-distance relations, consider $|E_1\rangle$. From the 12 maximal-zero vectors in Eq. (7.142), 4 of them are maximally distant to it: $|Z_{123}\rangle, |Z_{147}\rangle, |Z_{159}\rangle$, and $|Z_{168}\rangle$. To each one there is a unique $|Q^1_{st}\rangle$ that is maximally distant to both $|E_1\rangle$ and $|Z_{1st}\rangle$. Since each basis distribution generates 4 $|Q^r_{st}\rangle$, there must be 36 in total.

Because $\langle E_q|Q^r_{st}\rangle > \frac{1}{12}$ for $q \neq r$, the set S' is actually exhaustive with regard to forming maximally distant triplets, that is, for any pair of vectors in $\mathcal{S}'$ that are maximally distant from each other, the third vector forming a maximally distant triplet with the pair is already in $\mathcal{S}'$.

Before going further, we ask, how do the elements of $\mathcal{S}'$ map back to unit vectors in Hilbert space? This is definitely not unique since the basis distributions may correspond to any SIC in $d = 3$. However, if we express everything in terms of the most exceptional SIC given by the fiducial

$$|\psi_0\rangle = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \tag{7.148}$$

then we observe that

(i) the basis distributions $|E_r\rangle$ are the vectors corresponding to SIC projectors,

(ii) the uniform maximal-zero distributions $|Z_{rst}\rangle$ are the set of vectors associated with a complete set of mutually unbiased bases (MUBs) in $d = 3$. In fact, if we consider

the mapping $|Z_{rst}\rangle \mapsto \vec{\mu} \in \mathcal{H}_3$, we have

$$|Z_{123}\rangle \mapsto \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |Z_{147}\rangle \mapsto \begin{pmatrix}1\\1\\1\end{pmatrix}, \quad |Z_{159}\rangle \mapsto \begin{pmatrix}1\\\omega^2\\\omega^2\end{pmatrix}, \quad |Z_{168}\rangle \mapsto \begin{pmatrix}1\\\omega\\\omega\end{pmatrix},$$
$$|Z_{456}\rangle \mapsto \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad |Z_{258}\rangle \mapsto \begin{pmatrix}1\\\omega\\\omega^2\end{pmatrix}, \quad |Z_{267}\rangle \mapsto \begin{pmatrix}\omega^2\\1\\\omega^2\end{pmatrix}, \quad |Z_{249}\rangle \mapsto \begin{pmatrix}\omega\\1\\\omega\end{pmatrix},$$
$$|Z_{789}\rangle \mapsto \begin{pmatrix}0\\0\\1\end{pmatrix}, \quad |Z_{369}\rangle \mapsto \begin{pmatrix}1\\\omega^2\\\omega\end{pmatrix}, \quad |Z_{348}\rangle \mapsto \begin{pmatrix}\omega^2\\\omega^2\\1\end{pmatrix}, \quad |Z_{357}\rangle \mapsto \begin{pmatrix}\omega\\\omega\\1\end{pmatrix}, \quad (7.149)$$

where $\omega = e^{\frac{2\pi i}{3}}$ and the Hilbert space vectors are listed in unnormalized fashion, and

(iii) the 36 added vectors $|Q^r_{st}\rangle$ correspond to the 4 distinct SICs corresponding to $t = \frac{\pi}{6}$ in the parametrization in Sec 3.3. In particular, if we denote the set $\{|Q^r_{st}\rangle\}$ of 9 vectors as mapping onto a SIC by

$$|Q^r_{st}\rangle \text{ for } (r,s,t) = \text{ 3 triples of indices + permutations } \mapsto \left|\psi^{(i)}_{\frac{\pi}{6}}\right\rangle \quad (7.150)$$

then the 36 vectors that form maximally distant triplets with the 9 basis distributions

and 12 uniform maximal-zero distributions are

$$|Q^r_{st}\rangle \text{ for } (r,s,t) = (1,2,3),(4,5,6),(7,8,9) \quad \mapsto \left|\psi^{(1)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix},$$

$$|Q^r_{st}\rangle \text{ for } (r,s,t) = (1,4,7),(2,5,8),(3,6,9) \quad \mapsto \left|\psi^{(2)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix},$$

$$|Q^r_{st}\rangle \text{ for } (r,s,t) = (1,5,9),(2,6,8),(3,4,7) \quad \mapsto \left|\psi^{(3)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}\omega\\1\\-2\end{pmatrix},$$

$$|Q^r_{st}\rangle \text{ for } (r,s,t) = (1,6,8),(2,4,9),(3,5,7) \quad \mapsto \left|\psi^{(4)}_{\frac{\pi}{6}}\right\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}\omega^2\\1\\-2\end{pmatrix}. \qquad (7.151)$$

Thus, the steps we have taken so far amounts to building up qutrit states in Hilbert space from the most exceptional SIC and 12 MUB vectors, and then adding 36 vectors that themselves form 4 other distinct SICs. The structure corresponding to the set $\mathcal{S}_1$ is actually equivalent to a geometric construct in complex projective space known as the Hesse configuration, which is a configuration of 9 points and 12 lines obtained from the set of inflection points of an elliptic curve. Note that in Hilbert space, the 9 points represent unit vectors and the 12 lines represent 2-dimensional subspaces that contain 3 of these unit vectors at a time.

The precise correspondence between the two can be described as follows:

(i) Each vector in the most exceptional SIC is orthogonal to 4 MUB vectors. In the Hesse configuration of an affine plane over a field of three elements, every point in the set of 9 points in an affine plane belong to 4 lines.

(ii) Each MUB vectors corresponds to the vector normal to the plane that contains 3 vectors of the most exceptional SIC. Thus, in terms of the affine plane, each MUB vector represents one of the 12 lines.

(iii) Because every MUB vector $|w\rangle$ is normal to some plane in the Hesse configuration, we can think about which pairs of orthogonal vectors $|u\rangle, |v\rangle$ span such a plane. If we

insist that $|u\rangle$ is a vector from the most exceptional SIC then the set $\{|u\rangle, |v\rangle, |w\rangle\}$ in Hilbert space corresponds exactly to one of the maximally distant triplets in $\mathcal{S}_1$.

(iv) Furthermore, the set of 36 vectors $\{|v\rangle\}$ determined from such a construction can actually be grouped into the 4 SICs associated with fiducials $\left|\psi_{\frac{\pi}{6}}^{(i)}\right\rangle, i = 1, 2, 3, 4$.

Now we are in a position to ask, do we have enough points such that maximizing $\mathcal{S}_1$ yields qutrit state space? It would be rather surprising if the answer were yes, and indeed, it turns out the answer is negative. In fact, it is not hard to construct an example of a probability vector that is consistent with $\mathcal{S}_1$ but does not correspond to a quantum state.

One possibility is to consider probability vectors of the form

$$|R\rangle = (a, a, b, b, b, b, b, b, a) \tag{7.152}$$

such that $a > b$ and $\langle R|Z_{123}\rangle = \frac{1}{12}$. Because $|R\rangle$ must be a probability distribution, we have constraints

$$\begin{aligned} 3a + 6b &= 1, \\ \frac{1}{6}a + \frac{5}{6}b &= \frac{1}{12}, \end{aligned} \tag{7.153}$$

which have a unique solution

$$|R\rangle = \frac{1}{18}(4, 4, 1, 1, 1, 1, 1, 1, 4). \tag{7.154}$$

It is easy to check that for the inner products that are closest to breaking consistency

$$\langle R|E_1\rangle = \frac{5}{36}, \qquad\qquad \langle R|Q_{12}^3\rangle = \frac{11}{72} \tag{7.155}$$

and so $|R\rangle$ is consistent with $\mathcal{S}_1$ but mapping it back to a vector in Hilbert space, we get

$$|R\rangle \mapsto \rho = \frac{1}{3}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \tag{7.156}$$

with eigenvalues given by

$$x^3 - 3x^2 + \frac{1}{9} = 0, \tag{7.157}$$

which has a negative root according to Descartes' rule of signs, i.e., the two sign changes means there are two or zero positive roots. Thus, it is possible to maximize $\mathcal{S}_1$ into a set that is not quantum state space.

Because the construction of maximally distant triplets is exhausted by $\mathcal{S}_1$, how can we further add points to make a larger consistent set? We can no longer add maximally distant triples but we can still take pairs of points that are not maximally distant and find a point that is maximally distant to both.

More specifically, we add a third vector $\left|{}^{p}_{q}P^{r}_{st}\right\rangle$ to the pair $|E_q\rangle, |Q^r_{st}\rangle$ where

$$\langle E_q|Q^r_{st}\rangle > \frac{1}{12} \tag{7.158}$$

such that

$$\langle E_q|{}^{p}_{q}P^{r}_{st}\rangle = \langle Q^r_{st}|{}^{p}_{q}P^{r}_{st}\rangle = \frac{1}{12}, \qquad \langle {}^{p}_{q}P^{r}_{st}|{}^{p}_{q}P^{r}_{st}\rangle = \frac{1}{6}. \tag{7.159}$$

To see what the generic vector looks like, it suffices to consider two examples. First, suppose we take the pair $|E_1\rangle, |Q^4_{56}\rangle$. Then

$$\left|{}^{7}_{1}P^{4}_{56}\right\rangle = \frac{1}{18}(0,3,3,4,1,1,4,1,1). \tag{7.160}$$

On the other hand, if we take $|E_2\rangle, |Q^5_{64}\rangle$ then

$$\left|{}^{8}_{2}P^{5}_{64}\right\rangle = \frac{1}{18}(3,0,3,1,4,1,1,4,1). \tag{7.161}$$

The pattern that holds in general is the following: given the pair $|E_q\rangle, |Q^r_{st}\rangle$, let

$$\{(q,q',q''),(r,s,t),(p,p',p'')\} \tag{7.162}$$

be a set of index triples corresponding to one of the 12 lines in

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \tag{7.163}$$

that is, the set of three affine lines of same slope, so

$$\left|{}^{p}_{q}P^{r}_{st}\right\rangle = 0\,|q\rangle + \frac{1}{6}\left(|q'\rangle + |q''\rangle\right) + \frac{2}{9}(|r\rangle + |p\rangle) + \frac{1}{18}(|s\rangle + |t\rangle + |p'\rangle + |p''\rangle). \tag{7.164}$$

It is not as complicated as it looks because relevant 5 indices correspond exactly to 2 intersecting lines in the affine plane, with the intersection point given by the raised index r in $|Q^r_{st}\rangle$ and the intersecting lines (p,q,r) and (r,s,t) are uniquely specified by q and s,t.

It is not difficult to check that we can add all such vectors since for the most important inner products we get

$$\langle Z_{pqr}|{}^p_qP^r_{st}\rangle = \frac{5}{54}, \qquad \langle Q^p_{p'p''}|{}^p_qP^r_{st}\rangle = \frac{1}{12}. \tag{7.165}$$

Note that because the second inner product corresponds to a new pair of maximally distant points, it is possible to add more points by constructing maximally distant triplets with such pairs if desired. Also, at least one of $\left|{}^p_qP^r_{st}\right\rangle$ will be inconsistent with probability vectors of the form $|R\rangle$ in Eq. (7.154) so those non-quantum points are precluded from any maximization of the consistent set $\mathcal{S}_1 \cup \{\left|{}^p_qP^r_{st}\right\rangle\}$.

So how many vectors $\left|{}^p_qP^r_{st}\right\rangle$ can we add to set $\mathcal{S}_1$? Each $\left|{}^p_qP^r_{st}\right\rangle$ is determined by 5 indices, (r,s,t) corresponding to an affine line. From the index triple (r,s,t), we pick 1 index, r, as the intersection point and we select 1 of the 3 other lines that pass through r, specifically the line that passes through q. Thus, counting all possible combinations, $12 \times 3 \times 3 = 108$.

Hence, after adding all possible $\left|{}^p_qP^r_{st}\right\rangle$ we have

$$\mathcal{S}_2 = \left\{|E_q\rangle\,, |Z_{rst}\rangle\,, |Q^r_{st}\rangle\,, \left|{}^p_qP^r_{st}\right\rangle \;\middle|\; \text{all distinct } q,r,s,t = 1,2,...,9\right\}. \tag{7.166}$$

so the cardinality of the new consistent set is $|\mathcal{S}_2| = 108 + 57 = 165$.

So how far is $\text{Max}(\mathcal{S}_2)$ from quantum state space? Unfortunately, right now, all we can say is that quantum state space is one of its possible maximizations. It turns out that the discussion in Sec. 7.6 will show that a lot more structure is needed to pick out the quantum case uniquely.

As it stands, the two most promising possibilities for getting quantum state space from $\mathcal{S}_2$ is the following:

(i) The way we add more points to enlarged the set must involve some additional symmetry requirement. It might be possible to indirectly impose the symmetry by a identifying special points in the boundary of quantum state space that are not extreme points as having some geometrical significance. The problem with adding such points is that in the way we have been enlarging the consistent set, it seems very unnatural to add points that do not correspond to pure states.

(ii) The other possibility is to continue adding a maximally distant point to all pairs in $\mathcal{S}_2$ that is consistent with all points already in the set. The added points will then introduce new pairings to the set and we can proceed with the process. We do not know if the process continues indefinitely but if it does, the conjecture is that $\mathcal{S}_n$ converges to a set that densely covers the set of all pure states.

7.6 Duality and symmetry of maximal consistent sets

To make further progress in our attempts to build up quantum state space from a finite consistent set, we need more mathematical tools to see how far maximal consistency can take us. To this end, we introduce the concept of a dual set for consistent sets and examine its implications for the symmetry of the maximization of consistent sets.

Consider any finite-dimensional normed vector space V and $C > 0$. For some subset $S \subset V$, we can define the dual set

$$\tilde{S} = \text{Dual}(S) = \{\vec{v} \in V \mid \langle \vec{v}, \vec{w} \rangle \geq -C \quad \forall \vec{v} \in S\}. \tag{7.167}$$

The value C determines the scale of dual vectors and is not crucial in most cases.

Denoting the dual set of S by $\tilde{S} = \text{Dual}(S)$, some of the basic properties of dual sets are

(i) The dual set is closed and convex.

(ii) The dual of the dual set $\text{Dual}(\tilde{S})$ is the convex hull of S if S contains the origin.

(iii) The dual of the union of 2 sets S and T is the intersection of the dual set for each, i.e.,

$$\text{Dual}(S \cup T) = \tilde{S} \cap \tilde{T}. \tag{7.168}$$

(iv) The dual of the intersection of 2 closed convex sets S and T that include the origin is the convex hull of the union of their dual sets, i.e.,

$$\text{Dual}(S \cap T) = \text{Conv}(\tilde{S} \cup \tilde{T}). \tag{7.169}$$

We are interested in the dual of consistent sets. Given a consistent set S whose centroid or geometric center is at the origin, and with consistency given by

$$\langle \vec{p}, \vec{q} \rangle = k \quad \forall\ \vec{p}, \vec{q} \in S, \tag{7.170}$$

then for any non-zero vector $\vec{p} \in S$ we can define a dual vector $\tilde{\vec{p}} = \text{Dual}(\vec{p})$,

$$\tilde{\vec{p}} = \frac{k}{\|\vec{p}\|^2}\vec{p} \tag{7.171}$$

so that

$$\left\langle \vec{p}, \tilde{\vec{p}} \right\rangle = \frac{k \left\langle \vec{p}, \vec{p} \right\rangle}{\|\vec{p}\|^2}. \tag{7.172}$$

Geometrically, $\tilde{\vec{p}}$ is the vector oppositely directed to $\vec{p}$ along a line through the origin such that the points corresponding to $\vec{p}$ and $\tilde{\vec{p}}$ are maximally distant.

For consistent sets on the probability simplex Δ^{d^2-1} relevant to the quantum case,

$$k = -\frac{1}{d^2(d+1)} \tag{7.173}$$

and for each such set S, there are 3 spheres of significance:

$$\begin{aligned} &\text{an in-sphere with radius} && r = \sqrt{\frac{1}{d^2(d^2-1)}} \\ &\text{an out-sphere with radius} && R = \sqrt{\frac{d-1}{d^2(d+1)}} \\ &\text{a middle sphere with radius} && |k| = \sqrt{\frac{1}{d^2(d+1)}} \end{aligned} \tag{7.174}$$

The out-sphere is the smallest sphere that contains $\text{Conv}(S)$. The in-sphere is the largest sphere that is contained in any maximization of S. The in-sphere and out-sphere are dual to each other. The middle sphere is another sphere whose radius is the geometric mean of the radii of the in-sphere and the out-sphere, and it contains points whose antipodal points are maximally distant to each other, that is, it is a self-dual sphere.

Several properties of maximal consistent sets follow immediately from the notion of dual sets, that is, if S is a maximal consistent set then

(i) S contains d^2 basis vectors. This follows from the duality between the simplex formed by the basis distributions and the probability simplex Δ^{d^2-1} itself;

(ii) for every extreme point in S, there exists at least one point maximally distant to it; and

(iii) S is self-dual and it is given by the intersection of all half-spaces obtained from its extreme points through the lower bound for consistency.

Huangjun Zhu has proven a number of lemmas that follow from duality considerations [136] but we mention a few of them that provide some valuable insights for the quantum reconstruction effort.

Lemma 7.6.1. Suppose S is a maximal consistent set with infinitely many points on the out-sphere. Then any polytope inscribed by S can not contain the in-sphere.

Proof. Suppose we do have a polytope $P \subset S$ and the in-sphere of S is contained in P. Then $\text{Dual}(P) = \tilde{P} \supset S$ and is contained in the out-sphere. Consequently, S must have a finite number of points in the out-sphere. In particular, it must be less than the number of vertices of the dual polytope $\tilde{P}$ to P, which contradicts the assumption that S has infinitely many points on the out-sphere. □

Lemma 7.6.2. Let R be a consistent set and let C be the intersection of all maximal consistent sets that include R. If R contains C then there exists a maximization of R that has the same points on the out-sphere as R.

Proof. The crucial element in the proof is the observation that a vector $\vec{p}$ on the in-sphere of R lies on the boundary of the dual of R if and only if the dual vector $\tilde{\vec{p}}$ belongs to $R$and is found on the out-sphere or R.

Define R_o to be the set of points in R that lie on the out-sphere. Let

$$T = R \cup (\tilde{R} \cup S_{\max}) \tag{7.175}$$

where $S_{\max}$ is the maximum inscribed sphere corresponding to (any maximization of) R. T is consistent by construction and the intersection of its boundary ∇T with the in-sphere S_{in}consists of the dual vectors for the elements in R, i.e.,

$$\nabla T \cup S_{\text{in}} = \left\{ \tilde{\vec{p}} \mid \forall\ \vec{p} \in R \right\} \tag{7.176}$$

so $\vec{p} \in T$ when $\tilde{\vec{p}} \in R$. Thus, any maximization of T will fulfill the lemma. □

Without some additional structure, duality seems to strongly imply that the only way to uniquely obtain a set isomorphic to quantum state space from a consistent set is to

actually include all the pure states. Thus, we supplement the general properties of maximal consistent sets and their dual sets with geometric constraints coming from symmetry considerations.

The symmetry group $\mathrm{Sym}(S)$ of a subset $S \subset V$ of the normed vector space V is the group of isometries that leave the set invariant. If S is bounded and its centroid is at the origin, then $\mathrm{Sym}(S)$ is a subgroup of the orthogonal group $O(d^2-1)$. Because the particular symmetry group of the simplex Δ^{d^2-1} can be identified with the permutation group of d^2 objects, it seems natural to consider the symmetries implied by subgroups of the symmetric group $\mathrm{Sym}(d^2)$.

Theorem 7.6.3. Given any subgroup G of the permutation group, there exists infinitely many maximal consistent sets whose symmetry groups are G.

The theorem tells us that symmetry requirements alone typically do not lead to unique maximization of consistent sets.

In the quantum case, there is an important connection between permutation symmetry and unitary invariance:

Theorem 7.6.4. A unitary invariant maximal consistent set in the simplex Δ^{d^2-1} exists if and only if there exists a SIC in dimension d.

This result manifests the significant role played by SICs in relating maximal consistent sets and quantum state space through a connection between SICs and the unitary group, which might be independently relevant in the context of the SIC existence problem.

Theorem 7.6.5. When $d \geq 3$, the symmetry group of maximal consistent sets isomorphic to quantum state space are strictly smaller than the permutation group.

A corollary to this result says that the group of all projective unitary or anti-unitary operations that leave the SIC invariant is strictly smaller than the permutation group.

Chapter 8

Summary and outlook

Any proper description of a physical theory distinguishes between states, which refer to how systems are prepared, and observables, which determine the measurements that can be performed on systems. Quantum mechanics, however, has a property called self-duality, which means states and observables are the same. This self-duality is reflected clearly in the Born rule since $p(j) = \mathrm{Tr}\,(\rho E_j)$ tell us that the state ρ and the observable E_j are both positive semidefinite operators, up to normalization in the states, which is needed only so that $p(j)$ sum to 1. Thus, understanding the structure of quantum state space should go a long way in revealing fundamental aspects of quantum mechanics as a statistical theory of nature.

It has been our contention that the Born rule is a central idea in achieving this goal. In line with the interpretative program of QBism [43, 45, 47], we consider that the Born rule as a normative law that extends the classical probability theory to the domain of quantum systems. By treating quantum states as objects encapsulating an agent's beliefs about the future behavior of quantum systems, we can characterize quantum theory as the set of rules that dictate how to translate those beliefs into probability assignments that remain consistent with the potentialities presented by counterfactual measurements.

If quantum mechanics is indeed about how to structure one's degrees of belief, then the theory is best understood in the language of Bayesian probabilities. Moreover, its characteristics as a statistical theory should stem directly from the convex geometry of the set of probabilities describing its states, and, since it is self-dual, its observables. This thesis represents such an attempt to explain the geometric structure of quantum states in the framework of SIC probabilities, in particular, the hope for an intuitive grasp on why Hilbert space describes quantum theory so naturally.

8.1 Recap of main results

8.1.1 Practical implementation of SICs

SICs are important in practical quantum state tomography because they provide an unbiased estimate of an unknown state using an optimal number of measurements. With Medendorp, et al., we have developed a method that allows arbitrary POVMs to be approximated in linear optics to arbitrary precision [95].

To implement the SIC-POVM, we place a sequence of d^2 weak projections, using partial polarizing beam splitters, into a storage loop with an optical switch designed to trap the photons in a cavity. If the projections are sufficiently weak, the state describing the photons is virtually unaltered as they propagate through the loop. The SIC measurement is achieved by an X gate with a beam displacer crystal, and 3 Z gates with a liquid crystal wave plate that applies a $\frac{2\pi}{3}$ phase shift between linear polarizations, and a complete cycle of measurement consists of 3 rounds around the loop.

Although the storage loop experiment is very efficient in terms of optical elements, one disadvantage is that it is difficult to keep the losses low enough so that it does not distort the photon signal. For instance, without performing post-measurement adjustments to the raw data, the errors in the Medendorp experiment can be as large as 20%. This is why we propose an alternate method for realizing SICs with an integrated optics [124].

The design proposed is for a photonic multiport device, an optical circuit for performing an arbitrary discrete unitary operator as a sequence of qubit operations. We employ Naimark's theorem in turning a SIC-POVM in dimension d to a projective measurement in dimension d^2. The idea behind the multiport experiment then is to realize the SIC-POVM as a rotation into the basis corresponding to the Naimark projective measurement and then detecting the state of the photons in the standard basis. The resulting statistics then have the same outcome probabilities as that of the corresponding SIC-POVM. For the particular qubit and qutrit schemes presented in Sec. 4.4, we describe circuits that use about half as many elements as would be needed in a general decomposition of unitary operators into beam splitters and phase shifters [112].

8.1.2 Characterizing states in the SIC representation

The SIC representation of quantum states establishes a one-to-one correspondence between the probabilities for the outcomes of a SIC measurement with density operators represent-

ing quantum states via the mapping

$$\rho = \sum_i \left[(d+1)p(i) - \frac{1}{d} \right] \Pi_i. \tag{8.1}$$

In particular, the trace conditions on ρ for pure states, $\mathrm{Tr}\,(\rho) = \mathrm{Tr}\,(\rho^2) = \mathrm{Tr}\,(\rho^3) = 1$, translates into quadratic and cubic constraints on the SIC probability vectors,

$$\sum_i p(i)^2 = \frac{2}{d(d+1)}, \qquad \sum_{i,j,k} S_{ijk} p(i)p(j)p(k) = \frac{4}{d(d+1)^2} \tag{8.2}$$

where S_{ijk} are the expansion coefficients of operator multiplication in the basis of SIC projectors, that is,

$$\Pi_i \Pi_j = \sum_k S_{ijk} \Pi_k. \tag{8.3}$$

If we substitute the values 1 and $\frac{1}{d+1}$ for the cases when some of the indices of S_{ijk} are identical, then the cubic equation for pure states becomes $F(\vec{p}) = 0$ where

$$F(\vec{p}) = \frac{d-1}{d+1} \sum_i p(i)^3 + \sum_{i \neq j \neq k} S_{ijk} p(i)p(j)p(k). \tag{8.4}$$

Using the derivatives $\frac{\partial F}{\partial p(k)}$ of the function F with respect to the components of $\vec{p}$, we can define a mapping G,

$$G : p(k) \mapsto \frac{d+1}{3} \frac{\partial F(\vec{p})}{\partial p(k)} + \frac{2}{d(d+1)} \tag{8.5}$$

such that for probability vectors satisfying

$$G(\vec{p}) = \vec{p}, \qquad \vec{p} \cdot \vec{p} = \frac{2}{d(d+1)} \tag{8.6}$$

implies that $\vec{p}$ is a pure state. So pure states are fixed points of G that lie on a certain sphere. Moreover, if we introduce the mapping

$$M(\vec{p}) = \frac{1}{6} \left[1 - 3\mathrm{Tr}\left(\rho^2\right) + 3\mathrm{Tr}\left(\rho^3\right) \right] \tag{8.7}$$

where $\rho = \rho(\vec{p})$ is the density operator written in terms of SICs, then

$$M(\vec{p}) = \vec{\nabla} M(\vec{p}) = 0, \qquad \vec{p} \cdot \vec{p} = \frac{2}{d(d+1)} \tag{8.8}$$

corresponds to a pure quantum state. In this case, pure states are stationary points of the function M.

Note that in any arbitrary finite dimension, two distinct SICs Π'_i and Π_j are related by an orthogonal matrix

$$\Pi'_i = \sum_j R_{ij} \Pi_j \tag{8.9}$$

and the same orthogonal matrix R_{ij} can be used to relate their corresponding probability vectors.

The SIC representation is useful not just for translating pure states in terms of probability vectors; we can also express other properties of quantum states in terms of SIC probabilities. This includes the fidelity between pairs of states, the purification of mixed states, 2-qubit entanglement, and quantum operations in terms of affine maps on SIC probabilities, which are discussed from Sec. 5.6 to Sec. 5.9.

8.1.3 A most exceptional SIC for qutrits

Consider the extended Clifford group that consists of Weyl-Heisenberg displacement operators

$$D_{\mathbf{p}} = \tau^{p_1 p_2} X^{p_1} Z^{p_2}, \qquad \tau = -e^{i\frac{\pi}{d}} \tag{8.10}$$

where X and Z are the generalized shift and phase operators, and symplectic unitaries and anti-unitaries U_F that permute those displacement operators, i.e.,

$$U_F D_{\mathbf{p}} U_F^\dagger = D_{F\mathbf{p}}. \tag{8.11}$$

If we take the orbits of the extended Clifford group elements in $d = 3$, we can identify 8 SIC-families in the generic case, which are given by the fiducial vectors

$$\left|\psi_t^{(0\pm)}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ e^{\mp it} \\ -e^{\pm it} \end{pmatrix}, \qquad \left|\psi_t^{(k\pm)}\right\rangle = \sqrt{\frac{2}{3}} \begin{pmatrix} \omega^k \sin t \\ \sin\left(t \pm \frac{2\pi}{3}\right) \\ \sin\left(t \mp \frac{2\pi}{3}\right) \end{pmatrix}, \tag{8.12}$$

where $k = 1, 2, 3$ and $0 \leq t \leq \frac{\pi}{6}$.

The most special qutrit SIC corresponds to the single SIC obtained when $t = 0$, which we call the most exceptional SIC. It is generated by the fiducial vector

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{8.13}$$

Some properties that make the most exceptional SIC stand out are

(i) It has the fewest number of possible values for the real structure coefficients $\tilde{S}_{ijk}$ among qutrit SICs

$$\tilde{S}_{ijk} = \begin{cases} 1 & \text{if all indices are the same,} \\ \frac{1}{4} & \text{if 2 distinct indices,} \\ -\frac{1}{4} & \text{if 3 distinct, colinear indices,} \\ 0 & \text{if 3 distinct, non-colinear indices,} \end{cases} \tag{8.14}$$

where colinear means that the index triple (ijk) belong to the same line in Fig. 6.1.

(ii) It has the maximal number of linear dependent sets of $d = 3$ vectors among qutrit SICs, where the 12 normal vectors that span the subspace orthogonal to each linear dependent set form a complete of mutually unbiased bases in dimension 3 (see Table 3.3.)

(iii) The imaginary parts of its triple products form a simple Lie algebraic structure, which is equivalent to determining a SIC from its structure coefficients S_{ijk}.

(iv) Its SIC representation has the simplest algebraic conditions for probability vectors $\vec{p}$ corresponding to pure states:

$$\begin{aligned} \sum_i p(i)^2 &= \frac{1}{6}, \\ \frac{1}{3}\sum_i p(i)^3 &= \sum_{(ijk)\in Q} p(i)p(j)p(k) \end{aligned} \tag{8.15}$$

where Q is the set of index triples given by the lines marked in Fig. 6.1.

(v) In Section 6.2, we show that the probability spaces for qutrit SICs are related to that of the most exceptional SIC by a matrix that is fully described in terms of a circulant 3-dimensional rotation given by Eq. (6.16).

8.1.4 Maximal consistent sets and quantum states

When we write the Born rule in terms of SICs, we get the so-called quantum law of total probability

$$q(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i) \tag{8.16}$$

where $p(i)$ is the probability for SIC outcome i, $r(j|i)$ is the conditional probability for outcome j of an arbitrary measurement given the state Π_i, and $q(j)$ is the probability for outcome j of the arbitrary measurement. If we consider Eq. (8.16) as an addition to classical probability theory and demand that the Bayes rule applied to conditional probabilities yields probability distributions that represent valid states, then we establish a consistency condition

$$\frac{1}{d(d+1)} \leq \vec{p} \cdot \vec{q} \leq \frac{2}{d(d+1)} \tag{8.17}$$

for any $\vec{p}, \vec{q}$ that correspond to quantum states. Moreover, the set of quantum states is the largest subset of the probability simplex Δ^{d^2-1} for which Eq. (8.17) holds, that is, it is a maximal consistent set.

In Sec. 7.5, we presented a way for reconstructing quantum state space from probabilities by identifying an appropriate finite consistent set and imposing maximality to this set, subject to some additional symmetry constraint that is not yet known.

8.2 List of unresolved questions

The following is a partial list of open problems that can be considered as a natural follow-up to the results presented in this thesis:

1. Hoan Dang, et al. have shown that the linear dependencies in the Zauner subspaces defined by Eq. (3.38) where $F = F_Z$ of Eq. (3.41) can be extended to all finite dimensions if we consider all eigenvalues. Furthermore, the linear dependencies can be found by considering the eigenspace of some other unitary matrix [30]. Thus, it would be interesting to know, to what extent is the linear dependency structure sufficient to determine a SIC?

2. In Sec. 3.7, we considered the Lie algebra associated with a SIC treated as the basis for the space of Hermitian operators. Are there any further features of the Lie algebraic structure of SICs that would help determine their existence in all finite dimensions?

3. In our discussion of the properties of different qutrit SICs, we found it convenient to use an index generating matrix that can be thought of as a set of 9 points in a finite affine plane. The lines of this finite plane are particularly important for the most exceptional SIC. So what is the full symmetry associated with this finite affine plane? Moreover, how much of this symmetry is responsible for the geometric structure of the state space described by the most exceptional SIC?

4. In order to convince an experimenter to actually perform the multiport scheme with linear optics, it is necessary to analyze its feasibility in terms of how much small errors in the circuit propagate as a function of the circuit's size (number of optical elements) and depth (number of layers of parallel elements)?

5. The optical multiport scheme for SICs presents a specific implementation for qubit and qutrit SICs. However, there is no clear pattern for obtaining a decomposition into qubit operations that can be performed with the smallest number of beam splitters and phase shifters for a general SIC in higher dimensions. Is there a more systematic approach to achieve such a decomposition for any SIC-POVM, or is there at least an achievable lower bound to the number of elements needed for a sufficiently symmetric or sparse Naimark unitary matrix?

6. Is there any more geometric or algebraic relevance for the function $F(\vec{p})$ in Eq. (5.19) that defines not just the pure state but also the boundary of qutrit states? When the pure states are characterized as stationary points of the map $M(\vec{p})$ in Eq. (5.35), do they represent minimum, maximum, or saddle points?

7. What is the group associated with the rotational equivalence of qutrit SICs? Is the corresponding symmetry somehow related to the unitary symmetry of quantum states?

8. What is the conceptual significance of the specific convex geometry of quantum state space as demonstrated by the SIC probabilities and how is it related to the complex structure of quantum theory exhibited by its Hilbert space representation?

9. Zhu has shown that without any further demands, a finite set of consistent extreme points can belong to any one of an infinite class of maximal consistent sets [136]. If we are to somehow reconstruct quantum state space from the assumption of maximal consistency, this suggests that we will need to impose additional structure. Then, what would be a reasonable condition to include so that we can uniquely distinguish quantum state space as the maximal consistent set from the myriad of possibilities that correspond to any finite consistent subset of it?

10. Zhu also pointed out the equivalence of a unitarily invariant maximal consistent set with quantum state space. What are the conditions on SIC probabilities, thought of as a convex body in a real Euclidean space, that is equivalent to the unitary symmetry of the set of quantum states?

8.3 Closing remarks

If one postulates the Born rule, the complex structure of quantum mechanics then follows from the specific way it restricts the set of allowed probabilities. More precisely, the restriction requires that counterfactual events maintain a form of coherence with actual ones, expressed quantitatively by the quantum law of total probability, which is just the Born rule written terms of SICs.

To gain an intuitive understanding of what the restriction is, we think of it in geometric terms and picture the set of allowed probabilities as a convex set embedded in a real Euclidean space, where its shape embodies the structure of quantum mechanics as a statistical theory.

An example should help clarify what we mean by this. Consider the state space of classical mechanics, which for discrete systems is just a $d-1$ dimensional regular simplex for a system with d perfectly distinguishable states. The regularity of the simplex tells us that the pure states represented by the vertices are of equal footing, none preferred over the others. The symmetry of the vertices correspond to a set of transformations that take you from one pure state to another. Taking the Euclidean distance of the embedding space, we can also think of distances between states, where points closer to each other represent states that are in a sense similar, the extreme case being the vertices, which correspond to the perfectly distinguishable pure states. The distance of the points from the vertices shows how mixed a particular state is, where the center of the simplex is the maximally mixed state. More specifically, every point can be expressed uniquely as a mixture of the pure states. We may also consider the polygons formed by sections of the boundary called faces. Closed faces correspond to certain physical properties and extreme points correspond to properties with the finest graining. The faces can be given a partial ordering structure leading to the usual Boolean algebra of classical properties [101]. Knowing the convex geometry of the state space reveals much of the qualitative and sometimes even quantitative characteristics of a statistical theory. Quantum theory should be no exception.

What we know so far is that quantum state space can be represented geometrically by a convex body in d^2-1 dimensions for a system with sets of d perfectly distinguishable states. Those d states form a regular simplex within the convex body but it can also be rotated in various directions corresponding to a certain subgroup of the special orthogonal group. The collection of pure states form a lower-dimensional subspace of a sphere whose radius is $d-1$ times bigger than the largest sphere inscribed by the quantum set.

The fact that there exists a continuous set of transformations between pure states makes the set of quantum states very different from the simplices of classical systems.

In particular, a general mixed state can be decomposed into a mixture of pure states in infinitely many ways and all of these decompositions are physically equivalent. The quantum set also accommodates a $d^2 - 1$-dimensional simplex, at least for dimensions where SICs are known to exist. Our analysis assumes that this structure exists in all finite dimensions and explores what consequences follow from allowing this symmetry.

It turns out that although we know how to define density operators, we really know little of their geometric structure as a convex set of operators. We do not really have an intuitive way of visualizing what we mean by positivity or unitary invariance in higher dimensions in a way that is comparable to how well we understand the geometry of the Bloch ball for qubits. We believe the structure of quantum states in higher dimensions is worth examining in detail.

From recent investigations and from the results of this thesis, we at least know a whole lot more about qutrits. We know how to characterize the 2-dimensional cross-sections in terms of cubic curves that describe its boundary. We know how to classify some of the 3-dimensional cross-sections that include the center of the convex set for qutrits. We know that if we start with a regular simplex with d^2 points whose coordinates are chosen to coincide with the most exceptional SIC, we can construct sets of d pairwise maximally distant points that correspond to a Hesse configuration in Hilbert space.

We have also shown that the probability vectors for pure states possess a type of permutation symmetry that respects the symmetries of a finite affine plane with 9 points and 12 lines. The same symmetry plays a role for all boundary states in that the same function that defines pure states can be used to find the radial distance of the boundary from the uniform distribution for any direction within the probability simplex.

Granted there is still much to know about the geometry of quantum states, even for qutrits. Ultimately, we expect that quantum mechanics is formally the way it is because the shape of its state space has a certain geometric character that is both intuitive and simple, that its structure can be described in clear, unambiguous terms that will also feel like it is obvious and that in many ways, it had to be that way.

APPENDICES

Appendix A

From sets to probabilities

Set theory is the language underlying much of contemporary abstract mathematics and probability theory makes extensive use of its operations. This appendix serves both to provide a self-contained introduction to probability theory and an attempt to persuade the reader that probabilities emerge naturally from the notions of sets [2].

Measures on sets

A set S is any list or collection of objects, which are called the elements of S. We write $s \in S$ to mean that the element s belongs to set S. Note that we distinguish between element s and the singleton set $\{s\}$. A subset R of S is a set that contains some elements in S, written as $R \subset S$. The cardinality $n(S)$ of S is the number of elements in S. A set with a finite number of elements is called a finite set. Many important sets in mathematics are infinite, such as the natural numbers $\mathbb{N}$, the integers $\mathbb{Z}$, or the real numbers $\mathbb{R}$. Here we will be concerned mainly with finite sets.

There are three basic set operations:

(i) The complement $\overline{R}$ of set R is the set of elements that does not belong to the set.

(ii) The union $A \cup B$ of two sets A and B is the set that combines elements in A and B.

(iii) The intersection $A \cap B$ of two sets A and B is the set of elements shared by A and B.

It is also useful to define $A - B = A \cap \overline{B}$, that is $A - B$ is the set of all elements in A that are not in B.

The empty set $\emptyset$ is the unique set with no elements. It is considered a subset of any set. For instance, $A \cap R = \emptyset$ above since they do not share any elements.

To introduce measures we need the concept of a power set. The power set $\mathcal{P}(S)$ of set S is the set of all subsets of S. For any finite set S with $n(S) = N$, the cardinality of the power set is given by $n(\mathcal{P}(S)) = 2^N$.

A power set with the usual set operations is actually an example of a Boolean algebra. More generally, $\mathcal{B}(S) \subseteq \mathcal{P}(S)$ forms a Boolean algebra if

(i) for any $A \in \mathcal{B}(S)$, its complement $\overline{A}$ should be an element $\mathcal{B}(S)$, and

(ii) for every pair $A, B \in \mathcal{B}(S)$ then $A \cup B \in \mathcal{B}(S)$.

For any Boolean set with $A, B \in \mathcal{B}(S)$, $A \cap B \in \mathcal{B}(S)$ since $A \cap B = \overline{\overline{A} \cup \overline{B}}$. The empty set $\emptyset$ and S are also in $\mathcal{B}(S)$ since if $A \in \mathcal{B}(S)$, $A \cup \overline{A} = S \in \mathcal{B}(S)$ by (ii) and also $\overline{S} = \emptyset \in \mathcal{B}(S)$ by (i).

If we associate logical propositions with subsets of a Boolean algebra, we obtain the structure of Boolean logic, whose laws can be expressed in terms of set operations, which we list in Table A.1, where $A, B, C \in \mathcal{B}(S)$. To each subset A of S in the Boolean algebra $\mathcal{B}(S)$, we can a weight function called a measure which tells us the size of A relative to $\mathcal{B}(S)$. More precisely, a measure $m : \mathcal{B}(S) \to \mathbb{R}$ is a mapping that assigns to every subset $A \in \mathcal{B}(S)$ a number $m(A) \geq 0$ such that if $A, B \in \mathcal{B}(S)$ and $A \cap B = \emptyset$, then

$$m(A \cup B) = m(A) + m(B). \tag{A.1}$$

Technically speaking, we need a σ-algebra to also define measures on infinite sets and introduce the concept of probabilities. A σ-algebra $\Sigma(S)$ over set S is a nonempty set of subsets of S that is closed under complements and countable unions, i.e.,

(i) if $A \in \Sigma(S)$ then $\overline{A} \in \Sigma(S)$ and

(ii) if $A_1, A_2, \ldots, A_k \in \Sigma(S)$ then $A_1 \cup A_2 \cup \ldots \cup A_k \in \Sigma(S)$ for any $k \in \mathbb{N}$.

Note that k can be infinite. For finite sets, the distinction between Boolean and σ-algebras is not crucial.

The following theorem describes important properties of a measure:

Idempotence	$A \cup A = A$ $A \cap A = A$
Commutativity	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
Distributivity	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
de Morgan's laws	$\overline{A \cap B} = \overline{A} \cup \overline{B}$ $\overline{A \cup B} = \overline{A} \cap \overline{B}$
Complements	$A \cup \overline{A} = S$ $A \cap \overline{A} = \emptyset$
Identity on S	$A \cap S = A$ $A \cup S = S$
Empty set rules	$A \cup \emptyset = A$ $A \cap \emptyset = \emptyset$

Table A.1: Boolean laws on sets

Theorem A.0.1. Let m be a measure on the Boolean algebra $\mathcal{B}(S)$ of subsets of S. For any $A, B \in \mathcal{B}(S)$,

(i) if $B \subseteq A$ then $m(A - B) = m(A) - m(B)$;

(ii) $m(B) \leq m(A)$

(iii) $m(\emptyset) = 0$;

(iv) $m(A \cup B) = m(A) + m(B) - m(A \cap B)$.

Proof. To prove (i), we use De Morgan's laws and the rest follow as a consequence of (i):

(i) $(A-B)\cup B = (A\cap\overline{B})\cup B = A\cap(\overline{B}\cup B) = A$. $(A-B)\cap B(A\cap\overline{B})\cap B = A\cap(\overline{B}\cup B) = \emptyset$. Thus, since $A - B$ and B are disjoint, $m(A) = m(A - B) + m(B)$.

(ii) $m(A - B) \geq 0$ so it follows immediately from (i).

(iii) $m(\emptyset) = m(S - S) = m(S) - m(S) = 0$, where we used (i).

(iv) $A \cup B = (A \cup B) \cup \emptyset = (A \cup B) \cup [(A \cap B) - (A \cap B)] = [A - (A \cap B)] \cup [B - (A \cap B)] U (A \cap B)$. Therefore, $m(A \cup B) = m(A) - m(A \cap B) + m(B) - m(A \cap B) + m(A \cap B) = m(A) + m(B) - m(A \cap B)$.

□

The notion of probability

When a set S represents the possible outcomes of some experiment, we say that S is the sample space for the experiment. An event is a subset of outcomes in the sample space S. More specifically, it is the subset of outcomes for which a certain proposition holds true.

Kolmogorov's axioms states that for every event E in an experiment, we can assign a number $p(E)$ called the probability of event E such that

(i) $p(E)$ is non-negative,

(ii) an event corresponding to the entire sample space has $p(S) = 1$, and

(iii) for disjoint events A and B, that is, $A \cap B = \emptyset$, then

$$p(A \cup B) = p(A) + p(B). \tag{A.2}$$

But these properties are nothing more than a measure p on the Boolean algebra $\mathcal{B}(S)$ of subsets of S with unit total mass. The triple $\{S, \mathcal{B}(S), p\}$ denotes the probability space with sample space S, space of events $\mathcal{B}(S)$ and probability measure p.

Let $A, B \in \mathcal{B}(S)$. The probability measure p on $\mathcal{B}(S)$ has the following properties:

1. $p(\emptyset) = 0, p(S) = 1$.
2. $0 \leq p(A) \leq p(S)$.
3. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.
4. $p(\overline{A}) = 1 - p(A)$.

5. $p(A) \leq p(B)$ whenever $A \subseteq B$.

6. Let $\mathcal{E} = \{E_1, \ldots, E_N\}$ be a partition of S into N disjoint subsets, i.e., $E_i \cap E_j = \emptyset$ for any $i \neq j$, then

$$\sum_{j=1}^{N} p(E_j) = 1. \tag{A.3}$$

Properties (i) to (iii) are conditions on p that satisfy Kolmogorov's axioms.

The main role of p is to assign a weight to each event in $\mathcal{B}(S)$ based on its plausibility, in which case probability-one means an event is certain while probability-zero means an event is never going to occur.

When probabilities are determined with the assumption that certain events occurred, they are called conditional probabilities. The conditional probability $P(A|B)$ is the probability of event A given that event B is true. Conditional probabilities are especially important in Bayesian inference, where all probabilities are conditional on some prior knowledge, which are then updated when new information is acquired.

Intuitively, we want conditional probabilities to have the following properties:

(i) Since it is still a measure but only on a smaller sample space, we expect $p(B|B = 1$ and $p(\emptyset|B) = 0$.

(ii) For any $A \in \mathcal{B}(S)$, $p(A - B|B) = p(A \cap \overline{B}) = 0$.

(iii) p as a measure also means

$$\begin{aligned} p(A|B) &= p[(A \cap B) \cup (A \cap \overline{B})|B] \\ &= p(A \cap B|B) + p(A - B|B) \\ &= p(A \cap B|B). \end{aligned} \tag{A.4}$$

Thus, if we know the values for $p(A|B)$ for all subsets $A \subseteq B$ then we know everything about the probabilities conditioned on B.

(iv) If E_1 and E_2 represent disjoint events then

$$p(E_1|B) + p(E_2|B) = p(E_1 \cup E_2|B). \tag{A.5}$$

These properties are satisfied if we define the conditional probability of A given B to be

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \tag{A.6}$$

for $p(B) > 0$.

Conditional probabilities allow us to decompose probabilities of intersections into products of probabilities:

$$p(A_1 \cap A_2 \cap \ldots \cap A_N) = p(A_1)p(A_2|A_1)p(A_3|A_1 \cap A_2) \ldots p(A_N|A_1 \cap A_2 \cap \ldots \cap A_{N-1}). \tag{A.7}$$

For unions of events, the most important case is when we consider a partition of S, that is, $\{E_i : i = 1, 2, \ldots, N\}$ such that $E_i \cap E_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{N} E_i = S$. Using disjoint additivity,

$$p(A) = p\left[\bigcup_{i=1}^{N}(E_i \cap A)\right] = \sum_{i=1}^{N} p(E_i \cap A), \tag{A.8}$$

which leads us to the law of total probability:

Theorem A.0.2 (Law of total probability)**.** Let $\{S, \mathcal{B}(S), p\}$ be a probability space and $\{E_i : i = 1, 2, \ldots, N\}$ be a partition of S into disjoint subsets. The probability of event A is given by

$$p(A) = \sum_{i=1}^{N} p(E_i)p(A|E_i). \tag{A.9}$$

In many practical situations, the probability of an event is conditioned on some hidden states of the experiment. If we have some beliefs about the likelihood of each hidden state, the law of total probability tells us how to compute the event's probability.

Finally, we note that two events A and B are said to be statistically independent if

$$p(A \cap B) = p(A)p(B), \tag{A.10}$$

which implies that $p(A|B) = p(B)$ and $p(B|A) = p(A)$. That is, the plausibility of A is not affected by B, and vice-versa.

Bayesian updating

In daily life, we are interested in the values of all sorts of quantities. However, many of them are difficult to determine directly and must be inferred from various forms of

evidence. With probability models, the situation is reversed: there is a model for assigning probabilities to events provided a certain hypothesis is true. When there is available data serving as evidence, the aim becomes assessing the probability that the initial hypothesis is correct. Unfortunately, there is no unique, objective way to accomplish this. However, if we are willing to assign prior probabilities based on any subjective judgements, then probability theory tells us how to revise our initial beliefs in the face of new information.

Suppose we have some initial hypothesis represented by set H and with probability $p(H)$. From the inclusion-exclusion principle, $p(\overline{H}) = 1 - p(H)$. Now suppose we observe some evidence which we call event E. Also, there is a known probability model which provides us with values for the probabilities $p(E|H)$ and $p(E|\overline{H})$. To update the probability of H from its prior value $p(H)$ to some posterior value $p(H|E)$, we could use

$$p(H|E) = \frac{p(H \cap E)}{p(E)} \tag{A.11}$$

but generally we do not know any of the probabilities on the right-hand side. Instead we compute

$$p(H \cap E)p(H)p(E|H) \tag{A.12}$$

and use the law of total probability for the partition $\{H, \overline{H}\}$ to obtain

$$p(H|E) = \frac{p(H)p(E|H)}{p(H)p(E|H) + p(\overline{H})p(E|\overline{H})}. \tag{A.13}$$

If there is more than a single hypothesis to consider $(H_j, j = 1, \ldots, N)$, we would use

$$p(H_j|E) = \frac{p(H_j)p(E|H_j)}{\sum_{k=1}^{N} p(H_k)p(E|H_k)}. \tag{A.14}$$

Eq. (A.14) is called Bayes' theorem and it describes the rational way for modifying initial probabilities to account for new evidence.

Interpreting probabilities

What exactly do we mean when we say that the probability of event E is $p(E)$? One of the earliest known attempts to rigorously define probabilities is due to Pierre-Simon Laplace. It is called the classical theory of probability and was developed from an analysis of various games of chance. Its main idea is the principle of symmetry for all possible

outcome, provided those outcomes are reasonably deemed equally likely. Formally, if an experiment can yield one of N mutually exclusive, equally possible outcomes, and N_A of those outcomes are associated with event A then we define the probability of event A as

$$p(A) = \frac{N_A}{N}. \tag{A.15}$$

It is clear that the definition is faulty: first, it seems to work only in the cases with finite outcomes but more importantly, it is difficult to see how one avoids running into a circular argument when trying to identify outcomes that are equally likely. Nevertheless, there are certain instances when it seems intuitively justified, like when we consider ensembles in statistical mechanics, where microstates that correspond to the same macrostate, which are indistinguishable with regard to their thermodynamic properties, are deemed equiprobable.

Arguably the most widely accepted interpretation of probabilities is in terms of relative frequencies. The idea behind frequentism is that the probability of an event is reflected by its relative frequency over the course of many repetitions of the same experiment. If N_A denotes the number of occurrences of event A then if

$$\lim_{N\to\infty} \frac{N_A}{N} = p_A \tag{A.16}$$

we say that the probability of A is p_A.

The frequentist point of view has its own problems, the main one being the impossibility of performing identical experiments infinitely many times to confirm that the relative frequency indeed tends to the probability. It may be argued that the limit can be thought of as the limit for an imagined infinite ensemble of identically prepared, independently distributed systems in thought experiments. Still, you run into the problem of justifying why a certain limit is guaranteed when there is no way to actual observe the infinite relative frequency. For instance, say we observe that a coin flipped 20 times yields 10 heads and 10 tails. We would be inclined to believe that the coin is fair, with equal probability for heads or tails. But if is not inconceivable that if you actually flip it a million times, you would get 600,000 heads, which would hardly make the coin a fair one. Without actually doing the experiments, it is impossible to verify either way.

There are some attempts to think of relative frequency in terms of a finite number of trials but the relative frequency becomes ambiguous. Even with a marginal increase in the number of trials, relative frequencies can change significantly and thus, there is no discernible way to determine a unique probability.

One can choose to move away from relative frequencies to the camp of propensity theorists, who contend that the observed relative frequencies are mere indicators of some

physical property of the system to generate the particular outcome associated with it. The idea is to think of probability as an objective property akin to mass or electric charge. However, the problem with the propensity view is without expounding on the features of that physical property, then propensity becomes just a different name for probability.

Finally we have the camp of subjectivists, whose basic tenets is that all probabilities are conditional on an agent's subjective degrees of beliefs resulting from assessing the uncertainty surrounding a particular experiment. Whereas the probability associated with propensity would be called *objective chance*, subjectivists would refer to it as a *credence function*.Subjectivists are sometimes called Bayesians, since the methods for updating probabilities are often performed under the framework of subjective prior probabilities. The idea behind Bayesian probabilities is to assign probabilities that may be subjective but are nonetheless consistent. One way to motivate this is through the Dutch book argument for evaluating gambling odds, which we describe in the next section.

The rational gambler

A simple way to explain subjective probabilities is to argue that probability assignments are correct only in so much as they are developed from rational decision making. In this way, probabilities are really the generalization of logical thinking in the presence of uncertainty. Just as logic is concerned with valid reasoning about propositions but does not say anything about the truth values of those propositions, probability theory is concerned with proper methods for dealing with uncertainty without making a commitment to a fixed quantification of that uncertainty. For concreteness, it is conventional to use the example of a betting scenario.

Imagine there is a bookmaker who offers a bet for some event E to occur. A bettor willing to accept this bet has to pay a stake of pX dollars up front. The bookmaker pays X if E happens and nothing if it does not. In terms of gambling odds, the bettor is making a bet at odds of $(1-p)/p$ is to 1. Recall that for odds of x to y, it means the bettor is willing to bet a fraction $x/(x+y)$ of his money that he will win the bet. Thus, the probability of the bettor winning is given by

$$p = \frac{x}{x+y}. \tag{A.17}$$

Now since the bookmaker chooses the value of X, he might try to choose it so that the bettor never wins. This is called an inconsistent betting scenario, and rational gambler would never participate in that.

The important quantity to consider is the bettor's average gain G,

$$G = X - pX = (1 - p)X, \tag{A.18}$$

which is the net amount the bettor earns if he wins.

Suppose that event E has two outcomes called 'yes' and 'no', with associated winnings of Y and N, respectively. Let $p(E = \text{yes}) = q$ and $p(E = \text{no}) = 1 - q$ with corresponding stakes qY and $(1 - q)N$. Normally the bookmaker doesn't want to pay out anything if E does not happen so he can choose $N = 0$. If $E = \text{yes}$ then $G_{\text{yes}} = (1 - q)Y$; if $E = \text{no}$ then $G_{\text{no}} = -qY$ since the bettor has to pay for the stake in any situation.

The bookmaker can make both G_{yes} and G_{no} negative except when $1 - q$ and q have the same sign, which implies that

$$\frac{1 - q}{q} > 0. \tag{A.19}$$

If $q < 0$ then $1 - q < 0$ or $q > 1$, which is false by assumption. So $q > 0$, from which it follows that $1 - q > 0$ or $q < 1$, that is,

$$0 < q < 1. \tag{A.20}$$

The argument can be extended to include both certain $q = 1$ and impossible $q = 0$ events if the gambler wants to avoid the situation where he believes E might occur but his best gain is zero. The analysis can be extended to demonstrate other properties of probabilities such as disjoint additivity, when one considers betting on mutually exclusive events, and Bayes' theorem, when we consider conditional bets.

We emphasize that only the gambler assigns probabilities. The only concern of the bookmaker is to persuade the gambler to place bet in situations that are not so favorable to the bettor. The Dutch book argument is not an analysis of the competition between gambler and bookmaker; rather it is entirely about consistent betting behavior for the gambler alone. The only assumption is that a rational gambler will always avoid a scenario where he can never win.

Appendix B

Convex sets in real Euclidean spaces

The state space of a typical algebra has a natural convex structure: one can use probabilities μ_j to obtain mixtures

$$\sum_j \mu_j \omega_j \tag{B.1}$$

of states ω_j. If we assume finite dimensional $\mathcal{A}$ then the $S(\mathcal{A})$ is generally a convex subset of a finite-dimensional real space.

A subset K of $\mathbb{R}^N$ is convex if the line segments joining points of K lie in K.

For any arbitrary subset $X \in \mathbb{R}^N$, we can form its convex hull

$$\mathrm{Conv}(X) = \{\sum_j \mu_j x_j : x_j \in X, \mu_j \geq 0, \sum_j \mu_j = 1\}. \tag{B.2}$$

Consider the convex boundary $\partial_C K$ of a closed, compact subset of $\mathbb{R}^N$, it is the set of extreme points of K that admit only trivial decompositions in K. More explicitly, k is extreme if

$$k = \lambda k_1 + (1-\lambda)k_2, \quad 0 \leq \lambda \leq 1, k_1, k_2 \in K \implies k_1 = k_2 = k. \tag{B.3}$$

Two basic results in convex theory are the Krein-Milman theorem and Carathéodory's theorem. The Krein-Milman theorem [84] states that a compact convex set $K \in \mathbb{R}^N$ is the closed convex hull of its extreme points. Carathéodory's theorem states that every point in a convex subset of $\mathbb{R}^N$ admits a mixture of at most $N+1$ points.

A distinguished class of compact convex sets of $\mathbb{R}^N$ are the simplices Δ^N, which are characterized by the unique decomposition of any point in Δ^N into a convex combination of extreme points.

An important class of convex sets is the set of bistochastic matrices. These are square matrices B_{jk} of dimension N such that for all j, k

$$B_{jk} \geq 0, \qquad \sum_j B_{ij} = \sum_k B_{kl} = 1. \tag{B.4}$$

Bistochastic matrices define a class of Markov processes with transition probabilities that leave the uniform distribution invariant. Birkhoff's theorem says that the extreme points of the set of bistochastic matrices are the set of permutation matrices.

Probabilities can be stochastically ordered, that is, given probability distributions $\vec{p}$ and $\vec{q}$, we say that p majorizes q if for all $j = 1, 2, \ldots, N$

$$\max\{p_1 + p_2 + ... + p_j\} \geq \max\{q_1 + q_2 + \ldots + q_j\}. \tag{B.5}$$

Hardy-Littlewood-Polya theorem defines this order relation as some measure of mixedness: $\vec{p} \succ \vec{q}$ if and only iff $\vec{q}$ is obtained from $\vec{p}$ by a convex combination of permutations, i.e.,

$$\vec{q} = B\vec{p} \tag{B.6}$$

where B is some bistochastic matrix.

Majorized ordering is important in comparing certain quantities such as Shannon entropy $H(\vec{p})$

$$H(\vec{p}) = -\sum_{j=1}^{N} p_j \ln p_j \tag{B.7}$$

which is a concave function

$$H(\alpha\vec{p} + (1-\alpha)\vec{q}) \geq \alpha H(\vec{p}) + (1-\alpha)H(\vec{q}). \tag{B.8}$$

By virtue of Hardy-Littlewood-Polya theorem and since the entropy is invariant under permutations,

$$H(\vec{q}) \geq H(\vec{p}) \tag{B.9}$$

whenever $\vec{q} = B\vec{p}$ for bistochastic B. Thus, more mixed probability distributions have higher entropy.

Appendix C

Basic facts about linear operators

Here is a quick reminder of properties of linear operators that are relevant in the context of quantum mechanics.

1. We can write linear operators on a Hilbert space in the form

$$A = \sum_{i,j}^{d} a_{ij} |a_i\rangle\langle a_j| \tag{C.1}$$

by including enough terms in the decomposition, which is possible since

$$\mathcal{B}(\mathcal{H}) = \mathcal{H} \otimes \mathcal{H}^*. \tag{C.2}$$

2. A linear operator $A \in \mathcal{B}(\mathcal{H})$ is diagonalized by some unitary matrix U,

$$UU^\dagger = U^\dagger U = I, \tag{C.3}$$

if and only if it is normal, that is,

$$[A, A^\dagger] = AA^\dagger - A^\dagger A = 0. \tag{C.4}$$

3. Any normal operator can be expressed in the form

$$A = \sum_{i=1}^{r} a_i |a_i\rangle\langle a_i| \tag{C.5}$$

where the set $\{|a_i\rangle\}$ consists of an orthogonal set of eigenvectors corresponding to the r nonzero eigenvalues a_i of A. Eq. (C.5) is known as the spectral decomposition of A.

4. An operator P is said to be positive semidefinite if

$$\langle\psi|P\,|\psi\rangle \geq 0 \tag{C.6}$$

for all $|\psi\rangle \in \mathcal{H}$. For the eigenvectors of A, this means that the eigenvalues of A are all non-negative. Alternatively, one can write P as

$$P = AA^{\dagger} \tag{C.7}$$

for some matrix $A \in \mathcal{B}(\mathcal{H})$. If we choose A to be a lower triangular matrix with strictly positive diagonal entries, then Eq. (C.7) is the Cholesky decomposition of P. Any convex combination of positive operators is again a positive operator, which implies that the set $\mathcal{P}$ of all positive operators form a convex cone.

5. Any square complex matrix A can be decomposed into the form

$$A = UP \tag{C.8}$$

where U is a unitary matrix and P is a positive operator. This is the matrix analog of polar decomposition for complex numbers,

$$z = re^{i\theta} \tag{C.9}$$

where $r \geq 0$ and $0 \leq \theta \leq 2\pi$ for any $z \in \mathbb{C}$.

6. One way to impose positivity conditions on a parameterized matrix is by considering the spectrum of an arbitrary operator A derived from its characteristic polynomial

$$\det\left(aI - A\right) = a^N - c_1 a^{N-1} + c_2 a^{N-2} + \ldots + (-1)^N c_N = 0. \tag{C.10}$$

Recall that the coefficients c_i of this polynomial can be expressed in terms of symmetric functions of the roots, which are actually the eigenvalues a of A in this case, through formulas called Newton-Girard identities [72]:

$$\mathrm{Tr}\left(A^k\right) + c_{N-1}\mathrm{Tr}\left(A^{k-1}\right) + \ldots + c_{N-k+1}\mathrm{Tr}\left(A\right) = -kc_{N-k}, \qquad (1 \leq k \leq N), \tag{C.11}$$

$$\mathrm{Tr}\left(A^k\right) + c_{N-1}\mathrm{Tr}\left(A^{k-1}\right) + \ldots + c_0\mathrm{Tr}\left(A^{k-N}\right) = 0, \qquad (k > N). \tag{C.12}$$

Appendix D

Measuring distances between probability distributions

There are two important notions of distances connected with probabilities: the more traditional Euclidean distance when probability distributions are considered as vectors in a real linear space, and the perhaps lesser known notion of statistical distance, which distinguishes probability distributions that are almost identical in light of uncertainty.

Hilbert-Schmidt distance between probability distributions

Consider an experiment with N possible outcomes or some random variable X corresponding to it, with range $\{x_i : i = 1, 2, \ldots, N\}$. Let the probability for each outcome be given by

$$p(X = x_i) = p_i \tag{D.1}$$

In many situations, the outcomes themselves are not very significant. What matters more is the probability distribution $p(X)$ treated as a collection of N numbers $p(i)$ such that $p(i) \geq 0, \sum_i p(i) = 1$. We consider these numbers to be the components of vector $\vec{p}$ that belongs to some N-dimensional real vector space. Technically the probabilities reside in an $(N-1)$-dimensional space because of the linear constraint associated with normalization. Nevertheless, we keep to the components $p(i)$ and consider the set of all probabilities, known as the probability simplex, as embedded on $\mathbb{R}^N$. One reason to do this is that

it allows to adopt the usual geometric notions in real Euclidean spaces such as norms or distances.

Given the Cartesian coordinates $p = (p_1, p_2, \ldots, P_N)$ and $q = (q_1, q_2, \ldots, q_N)$ of two probability distributions, the inner product is the usual dot product for a real Euclidean N-space

$$\langle \vec{p}, \vec{q} \rangle = \vec{p} \cdot \vec{q} = \sum_{i=1}^{N} p_i q_i. \tag{D.2}$$

This natural comes with a definition for the distance between $\vec{p}$ and $\vec{q}$,

$$D(\vec{p}, \vec{q}) = D(\vec{q}, \vec{p}) = \sqrt{\sum_{i=1}^{N} (q_i - p_i)^2}. \tag{D.3}$$

and the length of a probability distribution

$$||\vec{p}|| = \sqrt{\sum_{i=1}^{N} p_i^2}. \tag{D.4}$$

Mathematically, a probability distribution is simply a measure on a sample space constrained to have unit total mass. In its vector space representation, the probability space corresponds to a set of points that forms a convex set. For instance, we know that the entire set of N-outcome probabilities forms an $(N-1)$-dimensional simplex Δ^{N-1}, which when embedded in a Euclidean space is simply a regular hypertriangle lying on the hyperplane given by $p_1 + p_2 + \ldots + p_N = 1$, where the standard basis vectors correspond to its vertices.

Statistical geometry and the Fubini-Study metric

When we represent density operators using real vectors, it is natural to assume that the these vectors are embedded in a d^2-dimensional real Euclidean space, so that we can talk about how two states are different from each other in terms of Euclidean or Hilbert-Schmidt distances. There is, however, an alternate way of comparing a pair of states, one that quantifies how easy or hard they are to distinguish from each other from measurement results. Here we review the situation discussed by Wootters in Ref. [130].

Consider a finite ensemble of identically prepared quantum systems analyzed by a fixed measuring device. The relative frequencies we obtain from measurements will typically be different from the actual probabilities if the preparation is fully known. Because of this statistical uncertainty, it is not always possible to distinguish between two marginally different preparations in a limited number of trials. This leads to the notion of statistical distance between quantum states, which involves counting the number of distinguishable states between any pair of states, where distinguishability is determined entirely by the size of statistical fluctuations.

In general, for an experiment with d possible outcomes, the relative frequencies are distributed according to a multinomial distribution, which is approximately Gaussian in the limit of very large N. Two probability distributions $\vec{p}$ and $\vec{q}$ are said to be distinguishable in N trials, if and only if

$$\frac{\sqrt{N}}{2}\left[\sum_{i=1}^{d}\frac{(\Delta p_i)^2}{p_i}\right]^{1/2} > 1 \tag{D.5}$$

where $\Delta p_i = p_i - q_i$. This condition guarantees that $\vec{p}$ and $\vec{q}$ will not have overlapping regions of uncertainty.

Geometrically, the statistical distance between $\vec{p}$ and $\vec{q}$ can be computed as the length of a smooth curve connected $\vec{p}$ and $\vec{q}$; in fact, it is the one with minimum length

$$L = \min_{\vec{p}(t)} \frac{1}{2}\int_0^1 dt \sqrt{\sum_{i=1}^{d}\frac{1}{p_i(t)}\left[\frac{dp_i(t)}{dt}\right]^2} \tag{D.6}$$

where $\vec{p}(t)$ is a parameterized curve such that $\vec{p}(0) = \vec{p}$ and $\vec{p}(1) = \vec{q}$.

Let $x_i = \sqrt{p_i}$ and $y_i = \sqrt{q_i}$. Then L becomes

$$L = \int_0^1 dt \sqrt{\sum_{i=1}^{d}\left(\frac{dx_i}{dt}\right)^2}, \tag{D.7}$$

which we recognize as the Euclidean length in the space of $\vec{x}$ and $\vec{y}$. The requirement that the curve $\vec{p}(t)$ must lie within the space of probabilities is imposed by the constraint

$$\sum_{i=1}^{d} p_i(t) = \sum_{i=1}^{d} x_i^2(t) = 1, \tag{D.8}$$

which means that the curve $\vec{x}(t)$ must lie on the unit hypersphere in the Euclidean space of $\vec{x}$-type vectors.

The statistical distance, also called the Bhattacharya distance, between probability distributions $\vec{p}$ and $\vec{q}$ is therefore the shortest distance between two points $\vec{x}$ and $\vec{y}$ along a unit sphere provided that $x_i = \sqrt{p_i}$ and $y_i = \sqrt{q_i}$. Formally, we would write

$$D_{\text{st}}(\vec{p}, \vec{q}) = \arccos\left(\sum_{i=1}^{d} \sqrt{p_i}\sqrt{q_i}\right). \tag{D.9}$$

Now consider two quantum states $|\psi\rangle$ and $|\phi\rangle$ in a d-dimensional Hilbert space. For some orthonormal basis $\{|1\rangle, \ldots, |d\rangle\}$, we can write

$$|\psi\rangle = \sum_{i=1}^{d} \alpha_i |i\rangle, \qquad |\phi\rangle = \sum_{j=1}^{d} \beta_j |j\rangle. \tag{D.10}$$

The statistical distance between $|\psi\rangle$ and $|\phi\rangle$ can then be computed from the probabilities $|\alpha_i|^2$ and $|\beta_i|^2$,

$$D_{\text{st}}(|\psi\rangle, |\phi\rangle) = \arccos\sum_{i=1}^{d} |\alpha_i||\beta_i| = \arccos\sum_{i=1}^{2} |\langle i|\psi\rangle||\langle\phi|i\rangle|. \tag{D.11}$$

The distance $D_{\text{st}}(|\psi\rangle, |\phi\rangle)$ reaches its maximum value when the measurement projects onto one of the basis vectors $|i\rangle$ that is also an eigenstate of the observable being measured, e.g. if we were trying to measure observable A, which has spectral decomposition

$$A = \sum_{i=1}^{d} a_i |i\rangle\langle i|, \tag{D.12}$$

then the absolute statistical distance D^*_{st} is

$$D^*_{\text{st}}(|\psi\rangle, |\phi\rangle) = \arccos|\langle\phi|\psi\rangle|. \tag{D.13}$$

This is actually the exact same expression for the angle between two rays $|\psi\rangle$ and $|\phi\rangle$ in Hilbert space. More specifically, the angle represents the geodesic length of a line in the projective Hilbert space $\mathbb{C}\text{P}^{d-1}$, more commonly known as the Fubini-Study metric. Thus, we have just established that the Fubini-Study metric measures the experimental distinguishability of pure quantum states in terms of the statistical distance between the probability distributions they generate from a measurement with a finite number of trials.

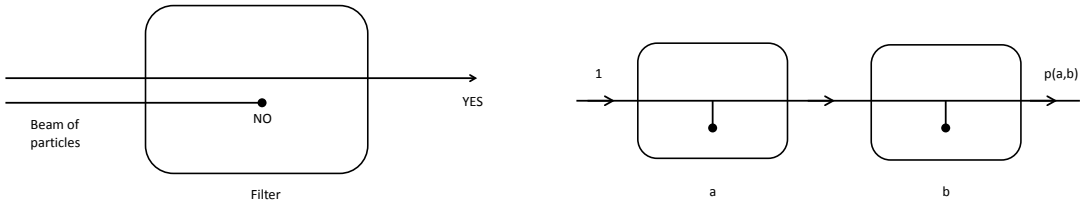
Appendix E

Theory of filters and transition probabilities

John von Neumann's treatise on quantum mechanics established the most widely accepted framework for the theory, where each quantum system is associated with a Hilbert space $\mathcal{H}$, unit vectors $|\psi\rangle \in \mathcal{H}$ specifies a possible state, and self-adjoint operators $A \in \mathcal{B}(\mathcal{H})$ represent observables. Does this particular representation impose limits on the allowed structure for quantum state space?

Birkhoff and von Neumann attempted to answer this question by examining the structure of pure states in terms of the set of yes-no measurements, obtaining a lattice of Hilbert space projection operators with features akin to a logical system. To understand the projection lattice operationally, we model the yes-no measurements as filters that select certain beams of particles, with 'yes' for particles that pass through a filter and 'no' for those absorbed, as shown in Fig. E.1a.

Consider 2 filters $a, b \in Q$ placed in sequence like in Fig. E.1b. In general, the intensity of the beam that passes through a will be diminished further by filter b and the relative decrease is constant in certain cases. For instance, the intensity of light is reduced by $\cos^2\theta$ as it pass through to linear polarizers oriented at an angle θ relative to each other. In analogy with light, we can introduce an absorption coefficient for every pair of filters corresponding to the transition probability of going from state a to b, and vice-versa. We say that an absorption coefficient $p(a, b)$ exists between filters $a, b \in Q$ if the relative decrease in intensity for a beam exiting a and passing into b is $p(a, b)$ and is the same when the a and bs are interchanged. Having a definite absorption coefficient for every pair of filters characterizes a geometric property for filters. For example, orthogonal filters have



(a) Yes-no measurement depicted as a filter. (b) Absorption coefficient for a and b.

Figure E.1: Transition probability from filters

the geometric property that the absorption coefficient vanishes.

A filter that makes the most fine-grained selection possible is called a minimal filter and such each minimal filter is associated with some pure quantum state. In the framework of Birkhoff and von Neumann, we represent filters by projection operators via the mapping $a \mapsto P(a)$ from minimal filter a to a projector $P(a)$, which projects onto a 1-dimensional subspace $\Psi(a) \subset \mathcal{H}$. Let $a \mapsto \psi(a)$ be the mapping from filter a to a unit vector $\psi(a) \in \mathcal{H}$ that lies within the subspace $\Psi(a)$. The relation between absorption coefficients of filters to the geometry of vectors in Hilbert space is then given by

$$p(a,b) = |\langle \psi(a), \psi(b) \rangle|^2 . \tag{E.1}$$

This is just the Born rule for calculating transition probabilities between pure states. It shows that we can reconstruct the geometry of minimal filters from the relative angles of their corresponding unit vectors in Hilbert space. However, the mapping $a \mapsto P(a)$ does not necessarily imply Eq. (E.1), which means that $a \mapsto P(a)$ does not fully explain the role played by Hilbert spaces in quantum theory.

We can extend absorption coefficients to all possible filters, which leads to the notion of a geometric system. A set $S \subset Q$ of filters is called a geometric system if for every pair of filters $a, b \in S$, $P(a,b)$ exists and has a definite value. If $S \subset S'$ where S, S' are both geometric systems, then S' is said to be an extension of S. When S' is the largest or maximal extension of S, then S' forms the complete state space. Filters that belong to a maximal geometric system are called states.

For any 2 states, we define the transition probability to be given by the absorption coefficient between those filters. The transition probabilities provide us with an abstract representation for states called a *probability space*.

Definition E.0.1. The space of transition probabilities or probability space (S, p) consists of a nonempty set S with function p defined on $S \times S$ such that

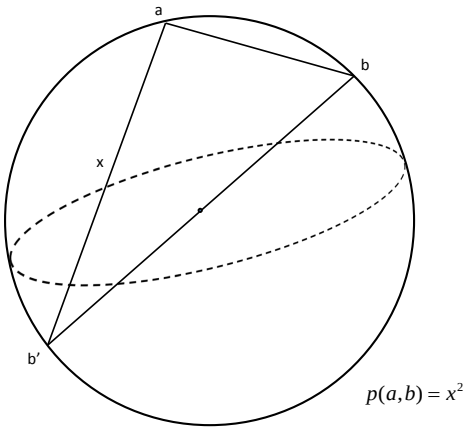


Figure E.2: Representing transition probabilities with $S(2, n)$.

(i) $0 \leq p(a, b) \leq 1$,

(ii) $p(a, b) = p(b, a)$,

(iii) $p(a, b) = 1$ implies $a = b, p(a, b) = 0$ if and only if $a \perp b$, and

(iv) for $R = \{r_i\}$ where $r_i \perp r_j$ whenever $i \neq j$, $\sum_i p(a, r_i) = 1$ implies that R is a basis.

A subset $R \subset S$ is an orthogonal system if every distinct pair of elements $a, b \in R$, $p(a, b) = 0$. R is called a basis if it is not contained in any larger orthogonal system in S. The existence of at least one basis follows from Zorn's lemma.

Let $S(\mathcal{H})$ be the set of 1-dimensional subspaces in $\mathcal{H}$. A probability space is said to be emebeddable in some Hilbert space, or Hilbertian, if for rays $\Psi_1, \Psi_2 \subset \mathcal{H}$,

$$p(\Psi_1, \Psi_2) = |\langle \psi_1, \psi_2 \rangle|^2 \tag{E.2}$$

where $\psi_1 \in \Psi_1, \psi_2 \in \Psi_2$ are unit vectors.

To find out which probability spaces are Hilbertian, we consider some examples of 2-dimensional probability spaces. First, we define an element called the complement:

Definition E.0.2. Let (S, p) be a 2-dimensional probability space. For every $a \in S$ there exists at least one a' such that $a \perp a'$. The complement a' is unique since if there exists another $\tilde{a}' \perp a$, then

$$p(\tilde{a}', a) + p(\tilde{a}', a') = 1 = p(\tilde{a}', a') \tag{E.3}$$

since $\{a, a'\}$ is a basis and $p(\tilde{a}', a) = 0$.

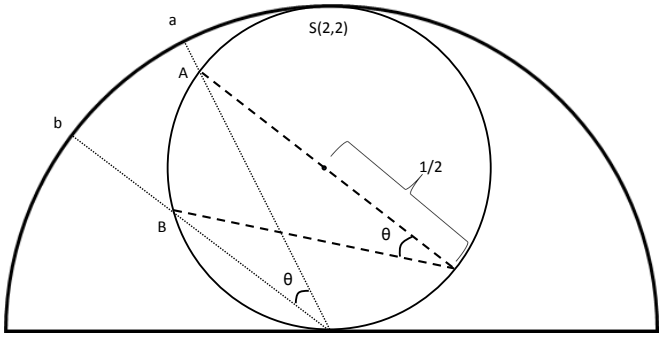


Figure E.3: Isomorphism between $S(2,2)$ and $\mathcal{H}_2(\mathbb{R})$ via stereographic projection.

Now consider a sphere of radius $1/2$ in $\mathbb{R}^n$. Let S be the set of points on the $(n-1)$–dimensional surface of the sphere. For any two points $a, b \in S$ define $p(a,b)$ as the squared distance between a and the antipode of b, i.e., the point diametrically opposite to b. As illustrated in Fig. E.2, this sphere represents a 2-dimensional probability space since

(i) $p(b,a) = p(a,b)$ since the distance between b and the antipode of a is the same as the distance between a and the antipode of b;

(ii) there exists a unique complement a' to each point $a \in S$; in particular, a' is the antipode of a; and

(iii) for each pair $\{a, a'\}$ and for every $b \in S$, $p(a,b) + p(a',b) = 1$ follows from the Pythagorean theorem.

Denote this probability space by $S(2,n)$. The subspaces of $S(2,n)$ are the 1-element subsets and 2-dimensional subsets that are invariant when the roles of a and a' are exchanged. Spaces $S(2,n)$ form a family of concentric spheres of increasing dimension, where $S(2,n)$ may be embedded in $S(2,m)$ whenever $n < m$.

Let us consider some examples of Hilbertian spaces $S(2,n)$. Consider the 2-dimensional real Hilbert space $\mathcal{H}_2(\mathbb{R})$. The unit circle in $\mathcal{H}_2(\mathbb{R})$ has two oppositely directed unit vectors corresponding to the same ray. Thus, the rays may be represented by a semicircle. Because the transition probability $p(\psi_1, \psi_2)$ between two rays ψ_1, ψ_2 is defined by the square of the cosine of the angle between ψ_1 and ψ_2, the isomorphism between $S(2,2)$ and $\mathcal{H}_2(\mathbb{R})$ is established via stereographic projection, as shown in Fig. E.3.

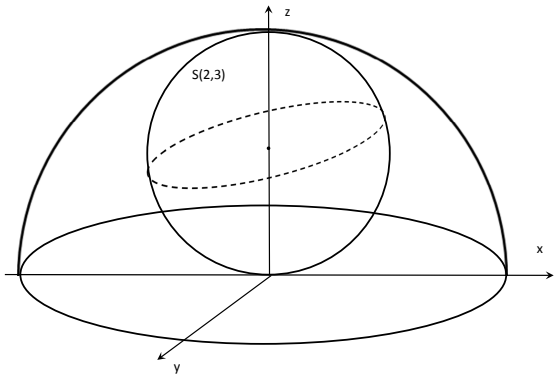


Figure E.4: Isomorphism between $S(2,3)$ and $\mathcal{H}_2(\mathbb{C})$ via stereographic projection.

Let $\mathcal{H}_2(\mathbb{C})$ be a 2-dimensional complex Hilbert space. Rays in $\mathcal{H}_2(\mathbb{C})$ can be represented using vectors

$$\psi = \begin{pmatrix} x \\ y + iz \end{pmatrix}, \qquad x, y, z \in \mathbb{R}, \quad x^2 + y^2 + z^2 = 1. \tag{E.4}$$

These points are in one-to-one correspondence with points $\vec{r} = (x, y, z)$ on the surface of a unit hemisphere in a 3-dimensional Euclidean space. By defining the transition probability for points $\vec{r}_1 = (x_1, y_1, z_1)$ and $\vec{r}_2 = (x_2, y_2, z_2)$ by

$$p(\vec{r}_1, \vec{r}_2) = (x_1 x_2 + y_1 y_2 + z_1 z_2)^2 + (y_1 z_2 - y_2 z_1)^2 \tag{E.5}$$

and by identifying all the points on the circle $x = 0$, we obtain a model for the for complex Hilbert space that is isomorphic to $S(2,3)$, where the isomorphism is again established by stereographic projection, as shown in Fig. E.4.

Appendix F

Projective geometry of quantum phase space

Most theories in physics are explicitly formulated in geometric terms. Examples include classical mechanics based on symplectic geometry, general relativity based on Riemannian manifolds, and Yang-Mills theory based on fiber bundles. On the other hand, the traditional language of quantum mechanics is linear algebra and functional analysis. In 1979, Tom Kibble showed that quantum mechanics can be cast in Hamiltonian form, similar to classical mechanics, but different in that the underlying phase space is complex projective Hilbert space. This phase space geometry offers new insights into various aspects of quantum mechanics such as linear superposition, uncertainty relations, geometric phases, and entanglement.

The space of pure quantum states is complex projective Hilbert space $\mathbb{C}\mathrm{P}^{d-1}$, where points correspond to rays in the Hilbert space $\mathcal{H}_d$. In finite dimensions, $\mathcal{H}_d \cong \mathbb{C}^d$, where normalized vectors define a sphere

$$S^{2d-1} = \{|\psi\rangle \in \mathbb{C}^d : \langle\psi|\psi\rangle = 1\}. \tag{F.1}$$

Two points $|\psi\rangle$ and $|\phi\rangle$ represent the same state if

$$|\phi\rangle = e^{i\alpha}\,|\psi\rangle\,, \qquad \alpha \in \mathbb{R} \tag{F.2}$$

so $\mathbb{C}\mathrm{P}^{d-1} = S^{2d-1}/U(1)$. For example, $\mathbb{C}\mathrm{P}^1 = S^3/U(1) \cong S^2$, which is the Bloch ball.

In general, $\mathbb{C}\mathrm{P}^{d-1}$ is a complex space equipped with a symplectic form and a Fubini-Study metric inherited from Hilbert space. If we decompose the Hilbert space scalar

product into its real and imaginary parts, $\langle\psi|\phi\rangle = R(\psi,\phi) + iJ(\psi,\phi)$ where $R(\psi,\phi) = \mathrm{Re}\left[\langle\psi|\phi\rangle\right]$ and $J(\psi,\phi) = \mathrm{Im}\left[\langle\psi|\phi\rangle\right]$, then

$$R(\psi,\phi) = R(\phi,\psi), \quad J(\psi,\phi) = -J(\phi,\psi), \quad J(\psi,\phi) = R(\psi,i\phi). \tag{F.3}$$

This means that both Hilbert space and its projective counterpart are Kähler spaces.

Expressing $|\psi\rangle = \sum_\alpha \psi_\alpha |e_\alpha\rangle$ for some orthonormal basis $\{|e_\alpha\rangle\}_{i=1}^d \subset \mathbb{C}^d$, the Fubini-Study metric is given by

$$g_{\alpha\beta} = \frac{\langle\psi|\psi\rangle\delta_{\alpha\beta} - \psi_\alpha\bar{\psi}_\beta - \psi_\beta\bar{\psi}_\alpha}{\langle\psi|\psi\rangle^2}, \tag{F.4}$$

that is, infinitesimal distances ds are defined by

$$ds^2 = \sum_{\alpha,\beta} g_{\alpha\beta} d\psi_\alpha d\bar{\psi}_\beta. \tag{F.5}$$

For $d = 2$ this is just the standard round metric on the surface of the sphere S^2.

The map from S^{2d-1} to $\mathbb{C}\mathrm{P}^{d-1}$ defines the so-called *Hopf fibration* or Hopf bundle. The simplest Hopf fibration is for $\mathbb{C}\mathrm{P}^1 \cong S^2$, which gives rise to the projection $S^3 \to S^2$. For example, consider a unit vector $|\psi\rangle$ for a qubit,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{F.6}$$

Define the quantity

$$n_j = \langle\psi|\sigma_j|\psi\rangle \tag{F.7}$$

where $\sigma_j, j = 1,2,3$ are the Pauli matrices. Then, it follows that $n_1^2 + n_2^2 + n_3^2 = 1$ so $|\psi\rangle \mapsto \vec{n}$ establishes a mapping from S^3 to S^2. For qubit Hilbert space, the Hopf fibration of S^3 consists of the fiber S^1, which represents global phases for $|\psi\rangle$ and the base space S^2, which is the Bloch ball. To visualize the Hopf fibration, one usually performs a stereographic projection from S^3 to $\mathbb{R}^3$.

Heinz Hopf found that the maps $f : S^3 \to S^2$ are characterized by a number called Hopf invariant, $\mathrm{Hopf}(f) \in \mathbb{Z}$. For example, the map $|\psi\rangle \to \langle\psi|\vec{\sigma}|\psi\rangle$ has Hopf invariant of 1. Any two maps f and g such that $\mathrm{Hopf}(f) = \mathrm{Hopf}(g)$ can be continuously deformed into one another. In this case, f and g are said to be homotopically equivalent.

Another way to think about Hopf fibrations is to consider rotations of a 2-sphere embedded in a 3-dimensional space. Its rotation group $SO(3)$ has a double cover that is diffeomorphic to the 3-sphere, which can be identified with the special unitary group $SU(2)$.

Any point on the 3-sphere is equivalent to a quaternion $q = q_0 + q_1 i + q_2 j + q_3 k$, which in turn is equivalent to a particular rotation R_q in three dimensions,

$$R_q = \begin{pmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 + q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & 1 - 2(q_2^2 + q_1^2) \end{pmatrix}. \tag{F.8}$$

The set of all possible quaternions produces the set of all possible rotations, which moves the tip of one unit vector of such a coordinate frame (say, the z-axis vector) to all points on a unit 2-sphere. However, fixing the z-axis vector does not specify the rotation fully: a further rotation is possible about the z-axis. Thus, the 3-sphere is mapped onto the 2-sphere, plus a single rotation. We say that the 3-sphere is the principle circle bundle over the 2-sphere, which is a Hopf fibration. (Circle bundles are actually the natural setting for Maxwell's theory of electromagnetism, in which the quantization of electric charge is exhibited by cohomology groups that correspond to gauge transformations.)

Only 3 spheres can be equipped with group structures:

(i) $S^1 \to$ complex numbers, $\mathbb{C}$,

(ii) $S^3 \to$ quaternions, $\cong SU(2)$,

(iii) $S^7 \to$ octonions.

The Hopf fibration $f : S^7 \to S^4$ gives rise to coset spaces S^7/S^3 which are important in quantum entanglement.

Suppose we have 2 quantum systems A and B with Hilbert spaces $\mathcal{H}^A$ and $\mathcal{H}^B$. The composite system AB has Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^B$, where $\otimes$ denotes the tensor product, meaning that if $\{|a_i\rangle\}$ is an orthonormal basis for $\mathcal{H}^A$ and $\{|b_j\rangle\}$ is an orthonormal basis for $\mathcal{H}^B$, then $\{|a_i, b_j\rangle \equiv |a_i\rangle \otimes |b_j\rangle\}$ forms an orthonormal basis for $\mathcal{H}^{AB}$.

We say that $|\psi\rangle$ is a separable state of $\mathcal{H}^{AB}$ if there exists $\left|\psi^A\right\rangle \in \mathcal{H}^A$ and $\left|\psi^B\right\rangle \in \mathcal{H}^B$ such that

$$|\psi\rangle = \left|\psi^A\right\rangle \otimes \left|\psi^B\right\rangle. \tag{F.9}$$

Otherwise, the state is entangled.

For any $|\psi\rangle \in \mathcal{H}^{AB}$, there exists an orthonormal basis $\{|e_i\rangle\} \subset \mathcal{H}^A$ and $\{|f_j\rangle\} \subset \mathcal{H}^B$ such that

$$|\psi\rangle = \sum_{k=1}^{r} \alpha_k |e_k\rangle \otimes |f_k\rangle \tag{F.10}$$

where $\alpha_k > 0, \sum_k^r \alpha_k^2 = 1$ and $r \leq \min\{n, m\}$. Eq. (F.10) is called the Schmidt decomposition of $|\psi\rangle$. A state $|\psi\rangle$ is said to be separable if and only if there is just one term in its Schmidt decomposition.

The most important example of entangled states are the 2-qubit Bell states

$$\left|\Psi^{\pm}\right\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle \pm |10\rangle\right), \qquad \left|\Phi^{\pm}\right\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right), \tag{F.11}$$

where $\{|0\rangle, |1\rangle\}$ is the standard basis for a qubit. The four Bell states form an orthonormal basis in $\mathbb{C}^4$.

Any normalized 2-qubit state can be expressed as

$$|\psi\rangle = \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle, \qquad |\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta^2| = 1. \tag{F.12}$$

This represents a unit sphere S^7 since $\alpha, \beta, \delta, \gamma \in \mathbb{C}$. A 2-qubit state $|\phi\rangle$ is separable if it is a tensor product, i.e., if $\left|\phi^A\right\rangle = a\,|0\rangle + b\,|1\rangle$ and $\left|\phi^B\right\rangle = c\,|0\rangle + d\,|1\rangle$ then

$$|\phi\rangle = \left|\phi^A\right\rangle \otimes \left|\phi^B\right\rangle = ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle. \tag{F.13}$$

Thus, Eq. (F.12) is separable if and only if $\alpha\delta = \beta\gamma$.

If we introduce the anti-linear map $E : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$,

$$E\,|\psi\rangle = (\sigma_2 \otimes \sigma_2)\,|\psi^*\rangle \tag{F.14}$$

where $|\psi^*\rangle$ is the complex conjugate of $|\psi\rangle$ and if we define

$$X_0 = \langle \sigma_3 \otimes I \rangle, \qquad X_3 = \mathrm{Re}\left[\langle E \rangle\right] \tag{F.15}$$

$$X_1 = \langle \sigma_1 \otimes I \rangle, \qquad X_4 = \mathrm{Im}\left[\langle E \rangle\right] \tag{F.16}$$

$$X_2 = \langle \sigma_2 \otimes I \rangle, \tag{F.17}$$

then $\sum_i X_i^2 = 1$ and so the mapping $|\psi\rangle \to (X_0, X_1, X_2, X_3)$ corresponds to the Hopf fibration $S^7 \to S^4$. We note that

$$C(\psi) = |\langle \psi | E\,|\psi\rangle\,| \tag{F.18}$$

is, in fact, just Wootters' concurrence function, a widely known entanglement measure for 2-qubit systems.

Note that $X_3 + iX_4 = 2\,(\alpha\delta - \beta\gamma)$ and $X_3 = X_4 = 0$ yield a separable state. Therefore the Hopf map $S^7 \to S^4$ is sensitive to entanglement and can be used to detect it. In particular, a 2-qubit state is separable if and only if its image under the Hopf map belongs to the 2-dimensional sphere defined by the intersection of S^4 with the 3-dimensional hyperplane defined by $X_3 = 0 = X_4$.

References

[1] J. Anandan. A geometric approach to quantum mechanics. *Found. Phys.*, 21:1265–1284, 1991.

[2] D. Applebaum. *Probability and information*. Cambridge University Press, Cambridge, 2nd edition, 2008.

[3] D. M. Appleby. SIC-POVMs and the Extended Clifford Group. *J. Math. Phys.*, 46:052107, 2005.

[4] D. M. Appleby. Symmetric Informationally Complete Measurements of Arbitrary Rank. *Opt. Spectr.*, 103:416, 2007.

[5] D. M. Appleby. Properties of the extended Clifford group with applications to SIC-POVMs and MUBs. *arXiv:0909.5233*, 2009.

[6] D. M. Appleby, H. Dang, and C. A. Fuchs. Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum-Uncertainty States. *arXiv:0707.2071*, 2007.

[7] D. M. Appleby, Steven T. Flammia, and Christopher A. Fuchs. The Lie Algebraic Significance of Symmetric Informationally Complete Measurements. *J. Math. Phys.*, 52:022202, 2011.

[8] D. M. Appleby, Å. Ericsson, and C. A. Fuchs. Properties of QBist State Spaces. *Found. Phys.*, 41:564–579, 2011.

[9] J. S. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447–452, 1966.

[10] J.S. Bell. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. Cambridge University Press, Cambridge, 2004.

[11] E. G. Beltrametti and S. Bugajski. Effect algebras and statistical physical theories. *J. Math. Phys.*, 38:3020–31, 1997.

[12] I. Bengtsson and H. Granström. The Frame Potential, on Average. *Open Sys. & Information Dyn.*, 16:145, 2009.

[13] I. Bengtsson, S. Weis, and K. Życzkowski. Geometry of the set of mixed states: an apophatic approach. In P. Kielanowski, S. T. Ali, A. Odzijewicz, M. Schlichenmaier, and T. Voronov, editors, *Geometric Methods in Physics XXX Workshop 2011*, Trends in Mathematics, pages 175–197, Basel, 2013. Springer.

[14] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States*. Cambridge University Press, Cambridge, 2006.

[15] Ingemar Bengtsson. From SICs and MUBs to Eddington. *J. Phys. Conf. Ser.*, 254:012007, 2010.

[16] Ingemar Bengtsson, Stephan Weis, and Karol Życzkowski. Geometry of the set of mixed quantum states: An apophatic approach. *arXiv:1112.2347*, 2011.

[17] G. Birkhoff and J. von Neumann. The logic of quantum mechanics. *Ann. of Math.*, 37:823–843, 1936.

[18] Niels Bohr. The Quantum Postulate and the Recent Developments of Atomic Theory. *Nature*, 121:580–590, 1928.

[19] D. Bonneau, M. Lobino, P. Jiang, C. M. Natarajan, Michael G. Tanner, R. H. Hadfield, S. N. Dorenbos, V. Zwiller, M. G. Thompson, and J. L. O'Brien. Fast path and polarisation manipulation of telecom wavelength single photons in lithium niobate waveguide devices. *Phys. Rev. Lett.*, 108:053601, 2012.

[20] Max Born. Zur Quantenmechanik der Stossvorgänge. *Zeitschrift für Physik*, 37:863–867, 1926.

[21] D. C. Brody and L. P. Hughston. Geometric quantum mechanics. *J. Geom. Phys.*, 38:19–53, 2001.

[22] Jeffrey Bub. *Interpreting the Quantum World*. Cambridge University PRess, Cambridge, 1999.

[23] P. Busch. Informationally complete sets of physical quantities. *Int. J. Theor. Phys.*, 30:1217, 1991.

[24] C. M. Caves. Symmetric Informationally Complete POVMs. internal report, 1999.

[25] C. M. Caves, C. A. Fuchs, and R. Schack. Conditions for compatibility of quantum-state assignments. *Phys. Rev. A*, 65:022305, 2002.

[26] C. M. Caves, C. A. Fuchs, and R. Schack. Unknown quantum states: the quantum de Finetti representation. *J. Math. Phys.*, 43:4537, 2002.

[27] D. Chruscinski. Geometric Aspects of Quantum Mechanics and Quantum Entanglement. *J. Phys. Conf. Ser.*, 30:9–16, 2006.

[28] R. Cleve, D. Gottesman, and H.-K. Lo. How to share a quantum secret. *Phys. Rev. Lett.*, 83:648–651, 1999.

[29] S. Colin, J. Corbett, T. Durt, and D. Gross. About SIC-POVMs and Discrete Wigner Distributions. *J. Opt. B*, 7:S778, 2005.

[30] Hoan Bui Dang, Kate Blanchfield, Ingemar Bengtsson, and D. M. Appleby. Linear Dependencies in Weyl-Heisenberg Orbits. *arXiv:1211.0215*, 2012.

[31] E. B. Davies and J. T. Lewis. An Operational Approach to Quantum Probability. *Commun. math. Phys.*, 7:239–260, 1970.

[32] K. M. Davis, K. Miura, N. Sugimoto, and K. Hirao. Writing waveguides in glass with a femtosecond laser. *Opt. Lett.*, 21(21):1729–1731, 1996.

[33] J. Du, M. Sun, X. Peng, and T. Durt. Realization of entanglement-assisted qubit-covariant symmetric-informationally-comple positive-operator-valued measurements. *Phys. Rev. A*, 74:042341, 2006.

[34] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling. Wigner tomography of two qubit states and quantum cryptography. *Phys. Rev. A*, 78:042338, 2008.

[35] Berthold-Georg Englert. Fringe Visibility and Which-Way Information: An Inequality. *Phys. Rev. Lett.*, 77:2152–2157, 1996.

[36] Aaron Fenyes. Non-commutative probability and maximal consistent sets. Master's thesis, University of Waterloo, Waterloo, ON, Canada, 2010.

[37] Christopher Ferrie and Joseph Emerson. Framed Hilbert space: Hanging the quasi-probability pictures of quantum theory. *New J. Phys.*, 11:063040, 2009.

[38] Ronald A. Fisher. On the dominance ratio. *Proc. Roy. Soc. Edinb.*, 42:321–341, 1922.

[39] M. Fitzi, N. Gisin, and U. Maurer. A quantum solution to the byzantine agreement problem. *Phys. Rev. Lett.*, 87:217901, 2001.

[40] S. Flammia. On SIC-POVMs in Prime Dimensions. *J. Phys. A*, 39:13483, 2006.

[41] S. T. Flammia. unpublished, 2004.

[42] C. A. Fuchs. On the Quantumness of a Hilbert Space. *Quant. Inf. Comput.*, 4:467, 2004.

[43] C. A. Fuchs. QBism, the Perimeter of Quantum Bayesianism. *arXiv:1003.5209*, 2010.

[44] C. A. Fuchs and Asher Peres. Quantum theory needs no interpretation. *Physics Today*, 53:70, 2000.

[45] C. A. Fuchs and R. Schack. A quantum-bayesian route to quantum-state space. *Found. Phys.*, 41:34556, 2011.

[46] Christopher A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, University of New Mexico, Albuquerque, 1996.

[47] Christopher A. Fuchs and Ruediger Schack. Quantum-Bayesian Coherence: The No-Nonsense Version. *arXiv:1301.3274*, 2013.

[48] Sascha Gaertner, Mohamed Bourennane, Christian Kurtsiefer, Adan Cabello, and Harald Weinfurter. Experimental demonstration of a quantum protocol for Byzantine agreement and liar detection. *Phys. Rev. Lett.*, 100:070504, 2008.

[49] E. F. Galvao, M. B. Plenio, and S. Virmani. Tripartite entanglement and quantum relative entropy. *J. Phys. A*, 33:8809, 2000.

[50] I. M. Gelfand and M. A. Naimark. On the imbedding of normed rings into the ring of operators on a hilbert space. *Math. Sbornik*, 12(2):197–217, 1943.

[51] A. Gleason. Measures on the Closed Subspaces of a Hilbert Space. *Indiana Univ. Math. J.*, 6(4):885–893, 1957.

[52] C. Godsil and A. Roy. Equiangular Lines, Mutually Unbiased Bases, and Spin Models. *Euro. J. Combin.*, 30:246–62, 2009.

[53] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. Available online as arXiv:quant-ph/9705052 (1997).

[54] D. Gottesman. A Theory of Fault-Tolerant Quantum Computation. *Phys. Rev. A*, 57:127, 1998.

[55] D. Gottesman. *The Heisenberg Representation of Quantum Computers*, pages 32–43. Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics. International Press, Cambridge, MA, 1999.

[56] Sandeep K. Goyal, B. Neethi Simon, Rajeev Singh, and Sudhavathani Simon. Geometry of the generalized Bloch sphere for qutrit. *arXiv:1111.4427*, 2012.

[57] M. Grassl. Computing Equiangular Lines in Complex Space. *Lecture Notes in Computer Science*, 5393:89–104, 2008.

[58] A. D. Greentree, S. G. Schirmer, F. Green, L. Hollenberg, A. R. Hamilton, and R. G. Clark. Maximizing the Hilbert space for a finite number of distinguishable quantum states. *Phys. Rev. Lett.*, 92:097901, 2004.

[59] S. Groeblacher, T. Jennewein, A. Vaziri, G. Weihs, and A. Zeilinger. Experimental quantum cryptography with qutrits. *New J. Phys.*, 8:75, 2006.

[60] S. Gudder, S. Pulmannova, S. Bugajski, and E. Beltrametti. Convex and Linear Effect Algebras. *Rep. Math. Phys.*, 44:359–379, 1999.

[61] N. Harrigan and R.W. Spekkens. Einstein, incompleteness, and the epistemic view of quantum states. *Found. Phys.*, 40(2):125–157, 2010.

[62] Werner Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43:172–198, 1927.

[63] D. Hilbert, J. von Neumann, and L. Nordheim. Über die Grundlangen der Quantenmechanik. *Mathematische Annalen*, 98:1–30, 1927.

[64] J. W. P. Hirschfeld. *Projective Geometries over Finite Fields*. Oxford University Press, Oxford, 1998.

[65] Lane P. Hughston. d=3 SIC-POVMs and Elliptic Curves. Seminar at the Perimeter Institute. Available at http://pirsa.org/07100040/, 2007.

[66] W. Hunziker. A note on symmetry operations in quantum mechanics. *Helve. Phys. Acta*, 45:233, 1972.

[67] I. D. Ivanović. Geometrical description of quantal state determination. *J. Phys. A*, 14:3241, 1981.

[68] L. Jakobczyk and M. Siennicki. Geometry of Bloch vectors in two-qubit system. *Phys. Lett. A*, 286:383, 2001.

[69] J. M. Jauch. *Foundations of Quantum Mechanics*. Addison-Wesley, Reading, 1968.

[70] N. S. Jones and N. Linden. Parts of quantum states. *Phys. Rev. A*, 71:012324, 2005.

[71] Amir Kalev, Jiangwei Shang, and Berthold-Georg Englert. Symmetric minimal quantum tomography by successive measurements. *Phys. Rev. A*, 2012.

[72] Dan Kalman. A Matrix Proof of Newton's Identities. *Math. Mag.*, 73(3):313–314, October 2000.

[73] K. Kato, M. Ishii, , and Y. Inoue. Packaging of large-scale planar lightwave circuits. *IEEE Trans. Compon. Packag. Manuf. Technol. B: Adv. Packag.*, 21(2):121–129, 1998.

[74] M. Kawachi. Silica waveguides on silicon and their application to integrated-optic components. *Opt. Quantum Electron.*, 22(5):391416, 1990.

[75] M. Khatirinejad. On Weyl-Heisenberg Orbits of Equiangular Lines. *J. Russ. Laser. Res.*, 28:333, 2008.

[76] T. W. B. Kibble. Geometrization of Quantum Mechanics. *Commun. Math. Phys.*, 65:189–201, 1979.

[77] I. H. Kim. Quantumness, Generalized Spherical 2-Design and Symmetric Informationally Complete POVM. *Quant. Inf. Comp*, 7:730–737, 2007.

[78] Gen Kimura. The Bloch Vector for N-Level Systems. *Phys. Lett. A*, 314:339, 2003.

[79] Gen Kimura and A. Kossakowski. The Bloch Vector for N-Level Systems: the Spherical-Coordinate Point of View. *Open Sys. & Info. Dyn.*, 12:207–229, 2005.

[80] A. Klappenecker, M. Rotteler, I. Shparlinski, and A. Winterhof. On approximately symmetric informationally complete positive operator-valued measures and related systems of quantum states. *J. Math. Phys.*, 46:082104, 2005.

[81] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46 – 52, 2001.

[82] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *J. Math. Mech.*, 17:59–87, 1967.

[83] Karl Kraus. *States, Effects, and Operations*, volume 190 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1983.

[84] M. Krein and D. Milman. On extreme points of regular convex sets. *Studia Mathematica*, 9:133–138, 1940.

[85] Y. L. Lim and A. Beige. Multiphoton entanglement through a bell multiport beam splitter. *Phys. Rev. A*, 71:062311, 2005.

[86] N. Linden, S. Pospecu, and P. Skrzypczyk. How small can thermal machines be? The smallest possible refrigerator. *Phys. Rev. Lett.*, 105:130401, 2010.

[87] A. Ling, K. P. Soh, A. Lamas-Linares, and C. Kurtsiefer. Experimental polarization state tomography using optimal polarimeters. *Phys. Rev. A*, 74:022309, 2006.

[88] R. Horodecki M. Horodecki, P. Horodecki. Separability of Mixed States: Necessary and Sufficient Conditions. *Phys. Lett. A*, 223:1–8, 1996.

[89] G. Mackey. *Mathematical Foundations of Quantum Mechanics*. W. A. Benjamin, New York, 1963.

[90] G. Mackey. Quantum Mechanics and Hilbert Space. *American Math. Monthly*, 64:45–57, 1967.

[91] Yong-Cheng Ou Mark S. Byrd, C. Allen Bishop. General open-system quantum evolution in terms of affine maps of the polarization vector. *Phys. Rev. A*, 83:012301, 2011.

[92] Graham D. Marshall, Alberto Politi, Jonathan C. F. Matthews, Peter Dekker, Martin Ams, Michael J. Withford, and Jeremy L. OBrien. Laser written waveguide photonic quantum circuits. *Opt. Exp.*, 17(15):12546–12554, 2009.

[93] A. Martin, A. Issautier, H. Herrmann, W. Sohler, D. B. Ostrowsky, O. Alibart, and S. Tanzilli. A quantum relay chip based on telecommunication integrated optics technology. *New J. Phys.*, 12:103005, 2010.

[94] J. C. F. Matthews, A. Politi, A. Stefanov, and J. L. OBrien. Manipulation of multi-photon entanglement in waveguide quantum circuits. *Nat. Photon.*, 3:346–350, 2009.

[95] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg. Experimental characterization of qutrits using SIC-POVMs. *Phys. Rev. A*, 83:051801R, 2011.

[96] A. Melikidze, V. V. Dobrovitski, H. A. De Raedt, M. I. Katsnelson, and B. N. Harmon. Parity effects in spin decoherence. *Phys. Rev. B*, 70:014435, 2004.

[97] I. Mendaš. The classification of three-parameter density matrices for a qutrit. *J. Phys. A: Math. Gen.*, 39:11313–11324, 2006.

[98] Paulo E. M. F. Mendonca, Reginaldo d. J. Napolitano, Marcelo A. Marchiolli, Christopher J. Foster, and Yeong-Cherng Liang. Alternative fidelity measure for quantum states. *Phys. Rev A*, 78:052330, 2008.

[99] B. Mielnik. Geometry of quantum states. *Commun. Math. Phys.*, 9:55–80, 1968.

[100] B. Mielnik. Theory of filters. *Commun. Math. Phys.*, 15:1–45, 1969.

[101] B. Mielnik. Convex Geometry: A Travel to the Limits of Our Knowledge. In P. Kielanowski, S. T. Ali, A. Odzijewicz, M. Schlichenmaier, and T. Voronov, editors, *Geometric Methods in Physics XXX Workshop 2011*, Trends in Mathematics, pages 253–264, Basel, 2013. Springer.

[102] S. Nolte, M. Will, J. Burghoff, and A. Tuennermann. Femtosecond waveguide writing: a new avenue to three-dimensional integrated optics. *Appl. Phys. A*, 71(1):109–111, 2003.

[103] M. G. A. Paris and J. Řeháček, editors. *Quantum State Estimation*, volume 649 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 2004.

[104] R. Peierls. *Surprises in Theoretical Physics*. Princeton University Press, Princeton, 1979.

[105] A. Peres. Separability Criterion for Density Matrices. *Phys. Rev. Lett.*, 77:1413–5, 1996.

[106] Asher Peres. Neumark's theorem and quantum inseparability. *Found. Phys.*, 12:1441 – 1453, 1990.

[107] A. Politi, M. J. Cryan, J. G. Rarity, S. Yu, and J. L. OBrien. Silica-on-silicon waveguide quantum circuits. *Science*, 320:646–649, 2008.

[108] I. R. Porteous. *Clifford algebras and the classical groups*. Cambridge University Press, Cambridge, 1995.

[109] R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger. High-speed linear optics quantum computing using active feed-forward. *Nature*, 445:65–69, 2007.

[110] E. Prugovečki. Information-theoretical aspects of quantum measurement. *Int. J. Theor. Phys.*, 16:321, 1977.

[111] M.F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nature Physics*, page 2309, 2012.

[112] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental Realization of Any Discrete Unitary Operator. *Phys. Rev. Lett.*, 73:58 – 61, 1994.

[113] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves. Symmetric informationally complete quantum measurements. *J. Math. Phys.*, 45:2171, 2004.

[114] J. I. Rosado. Representation of quantum states as points in a probability simplex associated to a SIC-POVM. *Found. Phys.*, 41:1200–1213, 2011.

[115] Gniewomir Sarbicki and Ingemar Bengtsson. Dissecting the qutrit. *arXiv:1208.2118*, 2012.

[116] C. Schaeff, R. Polster, R. Lapkiewicz, R. Fickler, S. Ramelow, and A. Zeilinger. Scalable fiber integrated source for higher dimensional path-entangled photonic qunits. *Opt. Exp.*, 20(15):16145–16153, 2012.

[117] C. Schaeff, R. Polster, R. Lapkiewicz, R. Fickler, S. Ramelow, and A. Zeilinger. Scalable fiber integrated source for higherdimensional path-entangled photonic qunits. *Opt. Exp.*, 20(10), 2012.

[118] A. J. Scott and M. Grassl. SIC-POVMs: A new computer study. *J. Math. Phys.*, 51:042203, 2010.

[119] Andrew J. Scott. Tight Informationally Complete Quantum Measurements. *J. Phys. A*, 39:13507, 2006.

[120] M. O. Scully, B.-G. Englert, and H. Walther. Quantum optical tests of complementarity. *Nature*, 375:368, 1991.

[121] I. E. Segal. Irreducible representations of operator algebras. *Bull. Am. Math. Soc.*, 53:73–88, 1947.

[122] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75:032110, 2007.

[123] E. C. G. Stükelberg. Quantum theory in real Hilbert space. *Helv. Phys. Acta*, 33:621–28, 1960.

[124] G. N. M. Tabia. Experimental scheme for qubit and qutrit symmetric informationally complete positive operator-valued measurements using multiport devices. *Phys. Rev. A*, 86:062107, 2012.

[125] G. N. M. Tabia and D. Marcus Appleby. Exploring the geometry of qutrit state space using symmetric informationally complete probabilities. *Phys. Rev. A*, 88:012131, 2013.

[126] A. Uhlmann and B. Crell. *Geometry of state spaces*, volume 768 of *Lecture Notes in Physics*, chapter Entanglement and Decoherence, pages 1–60. Springer, Berlin, 2009.

[127] John von Neumann. Mathematische Begründung der Quantenmechanik. *Nachr. Ges. Wiss. Göttingen*, 98:1–57, 1927.

[128] S. Waldron and L. Bos. Some remarks on Heisenberg frames and sets of equiangular lines. *New Zealand J. Math.*, 36:113–137, 2007.

[129] J. A. Wheeler. How Come the Quantum? *Annals of the New York Academy of Sciences*, 480:304–316, 1986.

[130] W. K. Wootters. Statistical distance and Hilbert space. *Phys. Rev. D*, 23:357, 1981.

[131] W. K. Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Ann. Phys.*, 176:1–21, 1987.

[132] W. K. Wootters. Quantum Measurements and Finite Geometry. *Found. Phys.*, 36:112–126, 2006.

[133] William K. Wootters. *The acquisition of information from quantum measurements*. PhD thesis, University of Texas at Austin, Texas, USA, May 1980.

[134] G. Zauner. *Quantendesigns. Grundzüge einer nicht- kommutativen Designtheorie*. PhD thesis, Univ. Wien, 1999. translated to English in Int. J. Quant. Inf. **9**, 445 (2011).

[135] Huangjun Zhu. SIC-POVMs and Clifford Group in prime dimensions. *J. Phys. A: Math. Theor.*, 43:305305, 2010.

[136] Huangjun Zhu. Maximal consistent sets and quantum state space. February 2013.

[137] M. Zukowski, A. Zeilinger, and M. A. Horne. Realizable higher-dimensional two-particle entanglements via multiport beam splitters. *Phys. Rev. A*, 55:2564–2579, 1997.

[138] Karol Życzkowski and Hans-Jürgen Sommers. Hilbert-Schmidt volume of the set of mixed quantum states. *J. Phys. A: Math. Gen.*, 36:10115–10130, 2003.