Rethinking Quantum States: An Introduction to QBism

A Major Qualifying Project Submitted to the Faculty of

Worcester Polytechnic Institute

in partial fufillment of the requirements for the Degree in Bachelor of Science in

Physics

By

Robert Kramer

Date

June 13, 2022

Advisor

Professor P.K. Aravind

Abstract

Quantum Bayesianism (QBism) is an alternative approach to quantum mechanics which aims to understand systems through their probability distributions rather than their wavefunctions. Measuring systems using symmetric informationally complete positive operator value measures (SIC-POVMs or SICs) allows this approach to be implemented in practice. This MQP will explore the ways in which QBism differs from traditional quantum mechanics, the structure of SICs and some of their applications.

Contents

List of Figures

List of Tables

1 Introduction

Quantum Bayesianisn, or QBism, is a new approach to quantum mechanics that replaces the use of a wavefunction to describe a quantum system by a list of probabilities. These probabilities, together with a set of rules, allow all the usual properties of the system to be calculated. In addition to the new calculational approach it takes to quantum mechanics, QBism also suggests an alternative interpretation of many of the features of quantum mechanics to the one advanced in the usual Copenhagen interpretation.

This MQP focused on a study of the basic features of QBism as laid out in the review article by Fuchs and Schack [1]. QBism itself grew out of an earlier probabilistic approach to quantum mechanics pioneered by Feynman [2]. Feynman advanced the very unconventional view that if one was willing to attach negative probabilities to certain events that quantum mechanics said could never occur, one was then in a position to calculate the correct values of all actually observed quantities without the use of a wavefunction or even the Born rule. QBism can be viewed as an elaboration of Feynman's viewpoint, but with a different definition of the fundamental probabilities that eliminates the negative values that were present in Feynman's scheme.

The main technical tool used in the development of QBism is the notion of a Symmetric Informationally Complete Positive Operator Measure (SIC-POVM or SIC). The SICs describe a special type of measurement that can be made on a quantum system to determine the defining probabilities on which the entire description of a quantum system depends. A good discussion of SICs can be found in the papers [1], [3], and [4].

This report is organized as follows. Section 2 provides an account of a quantum two-state systems, or qubits, based on the traditional formulation quantum mechanics. Section 3 provides an alternative treatment of qubits based on Feynman's viewpoint, and also indicates how the approach can be generalized to deal with systems of qubits. Section 4 presents the QBist account of two-state systems and also indicates how it can be generalized to d-state systems. Since SICs play a pivotal role in the QBist approach, the notion of a SIC is introduced and its basic properties are discussed. Although the QBist approach leads to the description of a quantum system by a probability vector, not all probability vectors correspond to possible states of quantum systems. A discussion is therefore given of which probability vectors qualify and how they can be picked out. Section 5 discusses how SICs can actually be constructed in dimensions 2 through 4 by the use of a group theoretical technique based on the Heisenberg-Weyl group. This technique actually works in higher dimensions as well, although that is not pursued here. Finally Section 6 ends with a brief discussion of some applications of SICs in a few select fields.

2 Two-State Quantum Systems

We begin by discussing the simplest type of quantum system, the qubit. A qubit is any two state quantum system whose most general state can be specified in terms of two orthogonal vectors in a 2-d Hilbert space. This description is applicable to any 2-state system, including spin states of spin-$\frac{1}{2}$ particles and the polarization of photons.

2.1 Standard Framework and the Bloch Sphere

Given orthogonal eigenstates $|0\rangle$ and $|1\rangle$, the wavefunction of an arbitrary qubit state is given as a superposition of these two states as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \qquad |\alpha|^2 + |\beta|^2 = 1 \tag{2.1}$$

where α and β are scalars. They are both allowed to be complex, but because the wavefunction is defined only up to an overall phase, it is possible to multiply $|\psi\rangle$ by a phase factor to make α real. Doing this is the typical convention because it removes redundancies and simplifies the comparison of wavefunctions.

We now introduce the Pauli spin matricies σ_x, σ_y, and σ_z

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.2}$$

The expectation value for a spin measurement along the x, y, or z axis is given by the expectation value of σ_x, σ_y, or σ_z in the state $|\psi\rangle$

$$\langle S_x\rangle = \langle\psi|\sigma_x|\psi\rangle\,, \langle S_y\rangle = \langle\psi|\sigma_y|\psi\rangle\,, \langle S_x\rangle = \langle\psi|\sigma_y|\psi\rangle \tag{2.3}$$

For instance, consider the state

$$|\psi\rangle = \sqrt{\frac{1}{2\sqrt{2}(\sqrt{2}-1)}}\,|0\rangle - i\sqrt{\frac{\sqrt{2}-1}{2\sqrt{2}}}\,|1\rangle \tag{2.4}$$

Using (2.4) in (2.3) and doing some algebra shows that the expectation values of the spin operator in these states is

$$\langle\psi|\sigma_x|\psi\rangle = 0 \quad \langle\psi|\sigma_y|\psi\rangle = -\frac{1}{\sqrt{2}} \quad \langle\psi|\sigma_z|\psi\rangle = \frac{1}{\sqrt{2}} \tag{2.5}$$

The spin components turn out to be very simple, but there is no way to anticipate this from (2.4). An alternative description of the state that makes it easy to anticipate what the expectation values of the spin components can be obtained as follows. The condition that α is real and the wavefunction must be normalized means that the wavefunction can be completely described by just 2 scalar parameters, which we can take to be the polar and azimuthal angles of a point on the unit sphere

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\,; \quad 0 \le \theta \le \pi \quad 0 \le \phi \le 2\pi \tag{2.6}$$

which can be represented graphically as

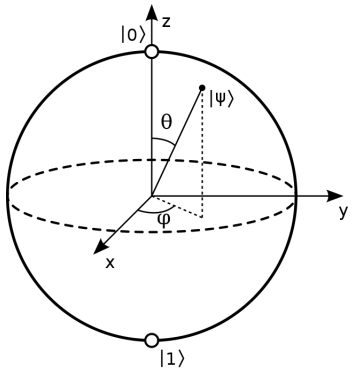

Figure 1: The Bloch sphere, a geometric representation of all single-qubit systems.

This way of depicting the state is known as the Bloch sphere representation. One can determine θ and ϕ from the parameters α and β in (2.1) via the equations

$$\theta = 2\cos^{-1}(\alpha); \quad \tan\phi = \frac{\mathrm{Im}\{\beta\}}{\mathrm{Re}\{\beta\}} \tag{2.7}$$

The vector from the origin of the Bloch sphere to the point $|\psi\rangle$ is called the pseudospin vector

$$\vec{\mathbf{s}} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z} \tag{2.8}$$

This pseudospin vector offers an intuitive explanation for expectation values of spin. The expectation value of $|\psi\rangle$ along some arbitrary direction $\hat{\mathbf{n}}$ is simply the scalar product $\vec{\mathbf{s}} \cdot \hat{\mathbf{n}}$.

This means that the expectation values of spin along x, y, and z are just the x, y, and z coordinates of $\vec{\mathbf{s}}$.

For instance, consider the state from (2.4). Using (2.7), we obtain $\theta = \frac{\pi}{4}$ and $\phi = \frac{3\pi}{2}$. Using (2.8), we obtain the spin components as

$$\vec{\mathbf{s}}_x = 0, \quad \vec{\mathbf{s}}_y = -\frac{1}{\sqrt{2}}, \quad \vec{\mathbf{s}}_z = \frac{1}{\sqrt{2}} \tag{2.9}$$

which agrees with the expectation values from (2.5)

It is clear that the pseudospin vectors that lie on the x,y, or z axis are the eigenvectors of σ_x, σ_y, or σ_z respectively. Thus, the direction of $|\psi\rangle$ as given by θ and ϕ provides information about the expectation values of the state's spin when measured along the principal axes.

A more general description of a state can be given by what is called a density matrix ρ defined as the outer product of the statevector with itself

$$\rho = |\psi\rangle\langle\psi| \,. \tag{2.10}$$

For a general qubit state, the density matrix can be written as

$$\rho = \frac{1}{2}[I + s_x\sigma_x + s_y\sigma_y + s_z\sigma_z] \tag{2.11}$$

$$= \frac{1}{2}[I + \vec{\mathbf{s}} \cdot \vec{\sigma}] \tag{2.12}$$

where the Pauli vector $\vec{\sigma}$ is defined as:

$$\vec{\sigma} = \sigma_x\hat{\mathbf{x}} + \sigma_y\hat{\mathbf{y}} + \sigma_z\hat{\mathbf{z}}. \tag{2.13}$$

If a system is in a state described by the density matrix ρ, the probability of observing an eigenstate $|\psi_j\rangle$ as a result of a measurement is given by

$$p(j) = \mathrm{tr}\left(|\psi_j\rangle\langle\psi_j|\,\rho\right) = \mathrm{tr}(\Pi_j\rho) \tag{2.14}$$

where we have introduced the projection operator $\Pi_j = |\psi_j\rangle\langle\psi_j|$.

As a check, by using (2.12) in (2.14), we can calculate the probabilities of finding the spin up along the x, y, and z axes and, from them, recover the expectation values of the spin componentes s_x, s_y, and s_z along x, y, and z.

3 Feynman's Approach

In 1987, Richard Feynman wrote a paper in which he suggested a slight change in the axioms of classical probability theory that would allow it to be used to describe quantum states and also make calculations of all their properties [2]. The change consists of assigning negative probabilities to certain events, but only those that can never be observed (such as the state of an object having both a definite position and a definite momentum). At first, this seems very unnatural, but provided that one restricts oneself to physically observable events, the probabilities are all between 0 and 1 as required. It is only the probabilities of intermediate (and unobservable) eventes that occur in the middle of a calculation that are allowed to be negative.

Feynman considers a qubit, but instead of starting from a wavefunction, he pretends that it is possible to measure the spin of a quantum particle along both the x and z directions simultaneously. This is of course, impossible, but for the moment Feynman ignores this difficulty. He introduces the following four basic "probabilities"

f_{++}: The probability of finding the spin up along both axes

f_{--}: The probability of finding the spin down along both axes

f_{+-}: The probability of finding the spin up along z and down along x

f_{-+}: The probability of finding the spin down along z and up along x

These can be defined as the expectation values of the 4 operators

$$F_{++} = \frac{1}{4}[I + \sigma_x + \sigma_y + \sigma_z] \quad F_{--} = \frac{1}{4}[I - \sigma_x + \sigma_y - \sigma_z] \tag{3.1}$$

$$F_{+-} = \frac{1}{4}[I - \sigma_x - \sigma_y + \sigma_z] \quad F_{-+} = \frac{1}{4}[I + \sigma_x - \sigma_y - \sigma_z] \tag{3.2}$$

Notice that these can have negative expectation values, so that the probabilities they generate can be negative. Feynman's point is that one can use these probabilities to get answers to all the questions one can ask about qubits if one ignores the negative values that occur at various points during the calculation and looks only at the probabilities of the final, physically meaningful results. It is also worth noting that though Feynman drops the non-negative restriction on the individual probabilities, the four probabilities should still sum to 1

$$f_{++} + f_{--} + f_{+-} + f_{-+} = 1 \tag{3.3}$$

Naturally, the probability of finding the particle spin up along the z-axis is given as the sum

of the probabilities of the two events in which the particle is spin up along z, f_{++} and f_{+-}

$$p(z\uparrow) = f_{++} + f_{+-} = \langle\psi|F_{++}|\psi\rangle + \langle\psi|F_{+-}|\psi\rangle \tag{3.4}$$
$$= \frac{1}{4}[\langle\psi|I + \sigma_x + \sigma_y + \sigma_z|\psi\rangle + \langle\psi|I - \sigma_x - \sigma_y + \sigma_z|\psi\rangle] \tag{3.5}$$
$$= \frac{1}{2}[\langle\psi|I|\psi\rangle + \langle\psi|\sigma_z|\psi\rangle] \tag{3.6}$$
$$= \langle\psi|\frac{1}{2}[I + \sigma_z]|\psi\rangle \tag{3.7}$$

Thus we obtain an observable $F_z = \frac{1}{2}[I + \sigma_z]$ which will have an expectation value equal to the probability for finding the system spin up along z. The same can be done for x and y

$$p(x\uparrow) = \langle\psi|F_{++} + F_{-+}|\psi\rangle \quad p(y\uparrow) = \langle\psi|F_{++} + F_{--}|\psi\rangle \tag{3.8}$$

Using this and the definitions of s_x, s_y, and s_z, we can relate Feynamn's probabilities to the pseudospin vector

$$s_x = p(x\uparrow) - p(x\downarrow) = 2p(x\uparrow) - 1 = 2\langle\psi|F_{++} + F_{-+}|\psi\rangle - 1 = 2(f_{++} + f_{-+}) - 1 \tag{3.9}$$
$$s_y = p(y\uparrow) - p(y\downarrow) = 2p(y\uparrow) - 1 = 2\langle\psi|F_{++} + F_{--}|\psi\rangle - 1 = 2(f_{++} + f_{--}) - 1 \tag{3.10}$$
$$s_z = p(z\uparrow) - p(z\downarrow) = 2p(z\uparrow) - 1 = 2\langle\psi|F_{++} + F_{+-}|\psi\rangle - 1 = 2(f_{++} + f_{+-}) - 1 \tag{3.11}$$

Now we can set up a system of equations to find the pseudospin vector from the Feynman probabilities or vice versa. For instance, consider the example qubit from (2.4) and the spin vector components from (2.9)

$$|\psi\rangle = \sqrt{\frac{1}{2\sqrt{2}(\sqrt{2}-1)}}\,|0\rangle - i\sqrt{\frac{\sqrt{2}-1}{2\sqrt{2}}}\,|1\rangle\,; \quad \vec{\mathbf{s}}_x = 0, \quad \vec{\mathbf{s}}_y = -\frac{1}{\sqrt{2}}, \quad \vec{\mathbf{s}}_z = \frac{1}{\sqrt{2}}$$

Using (3.9), (3.10), and (3.11)

$$f_{++} = \frac{1}{4}, \quad f_{+-} = \frac{1+\sqrt{2}}{4}, \quad f_{-+} = \frac{1}{4}, \quad f_{--} = \frac{1-\sqrt{2}}{4} \tag{3.12}$$

Notice that the value of f_{--} is negative! This is okay becuase f_{--} is the probability for the qubit to be measured spin down along the x and z axes simultaneously, which is not physically possible.

It is worth noting that one can combine (3.9), (3.10), and (3.11) with the condition $|\vec{\mathbf{s}}| \leq 1$ to show that

$$f_{++}^2 + f_{-+}^2 + f_{+-}^2 + f_{--}^2 \leq \frac{1}{2} \tag{3.13}$$

Additionally, (3.9), (3.10), and (3.11) can be combined with (2.11) and the normalization condition (3.3) to cast the density matrix in terms of the Feynman probabilities

$$\rho = \frac{1}{2}[I + (f_{++} + f_{-+} - f_{+-} - f_{--})\sigma_x + (f_{++} + f_{--} - f_{+-} - f_{-+})\sigma_y + (f_{++} + f_{+-} - f_{-+} - f_{--})\sigma_z] \tag{3.14}$$

In a similar manner, it is possible to find the probability of measuring the spin up along an arbitrary direction $\vec{\mathbf{V}}$ when the system is in the state described by the pseudospin vector ($\vec{\mathbf{s}}$) or the Feynman probabilities (f_{++}, etc.)

$$\begin{aligned} p(\uparrow) =& \frac{1}{2}[1 + \vec{\mathbf{s}} \cdot \vec{\mathbf{V}}] = \frac{1}{2}[1 + s_x V_x + s_y V_y + s_z V_z] && (3.15) \\ =& \frac{1}{2}[1 + (f_{++} + f_{-+} - f_{+-} - f_{--})V_x + (f_{++} + f_{--} - f_{+-} - f_{-+})V_y ... \\ & + (f_{++} + f_{+-} - f_{-+} - f_{--})V_z] && (3.16) \\ =& \frac{1}{2}[(1 + V_x + V_y + V_z)f_{++} + (1 - V_x - V_y + V_z)f_{+-} ... \\ & + (1 + V_x - V_y - V_z)f_{-+} + (1 - V_x + V_y - V_z)f_{--}] && (3.17) \end{aligned}$$

This matches the result given by Feynman. To make the connection to classical probability theory more apparent, he introduces the shortened form of (3.17)

$$p(\vec{\mathbf{V}} \uparrow) = \sum_a f_a p_a(\vec{\mathbf{V}} \uparrow) \quad a = (++, +-, -+, --) \tag{3.18}$$

$$\begin{aligned} p_{++}(\vec{\mathbf{V}} \uparrow) = \frac{1}{2}(1 + V_x + V_y + V_z), \quad p_{+-}(\vec{\mathbf{V}} \uparrow) = \frac{1}{2}(1 - V_x - V_y + V_z), \\ p_{-+}(\vec{\mathbf{V}} \uparrow) = \frac{1}{2}(1 + V_x - V_y - V_z), \quad p_{--}(\vec{\mathbf{V}} \uparrow) = \frac{1}{2}(1 - V_x + V_y - V_z) \end{aligned} \tag{3.19}$$

In this form, it is clear that one should interpret $p_a(\vec{\mathbf{V}} \uparrow)$ as the probability to measure the system's spin as up along $\vec{\mathbf{V}}$ while in the state a. These $p_a(\vec{\mathbf{V}} \uparrow)$ are weighted by the probability, f_a, of the particle being found in state a and then summed over all values for a. This approach can easily be generalized to a system of two qubits. For an arbitrary state of two qubits, the joint probability that a measurement of the first qubit along $\vec{\mathbf{V}}$ and the second along $\vec{\mathbf{U}}$ will show them both to be up is given by

$$p(\vec{\mathbf{V}} \uparrow_a \vec{\mathbf{U}} \uparrow_b) = \sum_{a,b} P_{ab} p_a(\vec{\mathbf{V}} \uparrow) p_b(\vec{\mathbf{U}} \uparrow) \tag{3.20}$$

Here P_{ab} is the probability of finding the first qubit in state a and the second in state b.

As an example, consider a 2-qubit Bell state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{3.21}$$

Using $P_{ab} = \langle\psi|F_a \otimes F_b|\psi\rangle$ gives

P_{ab}	$++$	$+-$	$-+$	$--$
$++$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$+-$	$\frac{1}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$-+$	$\frac{1}{8}$	$\frac{1}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$
$--$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$-\frac{1}{8}$

Table 1: The "probability" of finding the state ψ in the configuration f_{ab}. The columns represent the state of the first qubit, f_a and the rows represent the state of the second, f_b.

Note that the values of P_{ab} are negative for $a = b$, which would be problematic if they were meant to be interpreted as true probabilities. Proceeding using (3.19) and Table 1 in (3.20) gives

$$p(\vec{\mathbf{V}} \uparrow_a \vec{\mathbf{U}} \uparrow_b) = \frac{1}{4}(1 - V_x U_x - V_y U_y - V_z U_z) = \frac{1}{4}(1 - \vec{\mathbf{V}} \cdot \vec{\mathbf{U}}) \tag{3.22}$$

As expected, the final probability we obtain is a real probabilty between 0 and 1. For $\vec{\mathbf{V}} = \vec{\mathbf{U}}$, the probability of finding both particles spin as up along the shared spin vector is given by (3.22) as 0. This is consistent with the understood properties of the entangled Bell state, since a spin measurement on one qubit fixes the second in the opposite direction. That is to say, if one qubit is measured as spin up along $\vec{\mathbf{V}}$, the other qubit will, by the nature of quantum entanglment, be found as spin-down along $\vec{\mathbf{V}}$. Thus Feynman's method remains valid for multi-qubit systems.

The significance of what Feynman has done is in managing to describe the probabilities and expectation values of quantum systems without defining a wavefunction or using the Born rule. Instead, he imagines the system as having some probability f_a to be in each of the 4 unobservable states ($a =++,+-,-+,--$), and then shows how other, physically meaningful probabilities can be calculated from (3.18). The negative probabilities are not an issue so long as the 4 probabilities together obey (3.3) and (3.13).

Also important to mention is the ease with which this approach can be generalized to a large number of qubits. Extrapolating from (3.18) and (3.20), the probability for n mutually interacting qubits to be found spin-up with respect to n vectors $\vec{\mathbf{V}}_\mathbf{i}$, $(i = 0, 1, ...n)$ can be written as

$$p(\uparrow_1, \uparrow_2, ... \uparrow_n) = \sum_{[x_1,...,x_n]=1}^{4} \left[\prod_{i=1}^{n} f_{x_i} \times \prod_{i=1}^{n} p_{x_i}(\vec{\mathbf{V}}_\mathbf{i}) \right] \tag{3.23}$$

Although calculations with such an expression could prove unweildy, they nevertheless allow us to obtain results using only the ideas of conventional probability theory (expanded now to allow the occurence of negative probability).

4 Quantum Bayesianism and SICs

Expanding on the ideas presented by Feynman and others, Q-Bism was developed as a way to shift the description of a quantum system based on a wavefunction to a description based on a set of probabilities. In the Q-Bist view, a quantum state can be described by a probability vector - a list of probabilities that describe the outcomes of a particular set of measurements carried out on the system. From a suitable probability vector, it is possible to find all the observables of the system that one can get using the wavefunction. This approach expands upon the ideas of Feynman because it is applicable not just to qubits but qudits (or d-state systems). It also departs from Feynman's scheme in other ways that will become clear as we proceed.

This chapter contains a brief account of the Q-Bist description of quantum states, as laid out in the article by Fuchs and Shack [1]. The treatment is self-contained and detailed enough to be accessible to someone with no prior knowledge of this topic.

4.1 Introduction of SICs

The density matrix of a d-state quantum system is a dxd Hermitian matrix, and so characterized by d^2 independent parameters. The fact that the density matrix must have unit trace reduces the number of independent parameters to $d^2 - 1$. From this it follows that any probability vector describing the system must have d^2 components and have the form

$$||p\rangle\rangle = (p(1), p(2), ...p(d^2))^T \tag{4.1}$$

where the component $p(i)$ is the probability of the system being found in the state labeled by the index i. The fact that the system must be found in one of the states i (where $i = 1, ..., d$) implies that the sum of the components of this vector must be unity, and this again yields a total of $d^2 - 1$ parameters describing the system.

The probability of measuring the system to be in the ith state is given by

$$p(i) = \mathrm{tr}\, \rho E_i \tag{4.2}$$

where E_i is an operator, known as a POVM, that represents one of the SIC measurements that can be carried out on the system.

It follows from (4.2) that the mapping $\rho \to ||p\rangle\rangle$ is injective, meaning that it is one to one and has an inverse. However, the mapping is not surjective, so not all probability vectors represent valid physical states.

To define a probability vector via the relation (4.2), one can choose the operators E_i to be the elements of a SIC-POVM or SIC (Symmetric, Informationally Complete Positive Operator Value Measure) in dimension d. A SIC in d-dimensions is a set of operators $\{E_i\}, i = 1, ..., d^2$ defined as

$$E_i = \frac{1}{d} |\psi_i\rangle\langle\psi_i| = \frac{1}{d}\Pi_i, \qquad \mathrm{Tr}\,\Pi_i\Pi_j = \left| \langle\psi_i|\psi_j\rangle \right|^2 = \frac{d\delta_{ij} + 1}{d+1}, \qquad \sum_{i=1}^{d^2} E_i = 1 \tag{4.3}$$

For the moment, we set aside the problem of how one finds d^2 vectors $|\psi_i\rangle$ that form the members of a SIC. It is widely believed that SICs exist in all dimensions, although this still remains to be proven.

We will proceed simply by assuming that we can construct a SIC in any dimension and see how we can use it to achieve two goals: constructing an injective (or one-to-one) mapping from the density matrix into a probability vector and then determining what constraints the probability vector must satisfy if it is to represent a valid density matrix.

The probabiliy vector can be constructed from the density operator ρ by using the SIC operators E_i introduced earlier in (4.3). The i-th component of the probability operator is obtained as

$$p(i) = \mathrm{tr}\,\rho E_i = \frac{1}{d}\,\mathrm{Tr}\,\rho\Pi_i \tag{4.4}$$

This gives a mapping from an arbitrary density matrix ρ to the probability vector $||p\rangle\rangle$. One can arrive at the reverse mapping by expressing the density matrix as an expansion in the components of the SIC

$$\rho = \sum_{i=1}^{d^2} c_i E_i = \frac{1}{d}\sum_{i=1}^{d^2} c_i \Pi_i \tag{4.5}$$

Putting this into (4.4) gives

$$p(i) = \frac{1}{d^2}\sum_j c_i \operatorname{Tr}(\Pi_j \Pi_i) = \frac{1}{d^2}\left[c_i + \frac{1}{d+1}\sum_{j\neq i} c_j\right] = \frac{1}{d^2}\left[c_i + \frac{1}{d+1}(N - c_i)\right]$$
$$= \frac{1}{d(d+1)}(c_i + 1). \tag{4.6}$$

The second step uses the second part of (4.3), the third step uses $N = c_i + \sum_{j\neq i} c_j$, and the final step uses the fact that $N = d$, which is obtained by taking the trace of (4.5)

$$\operatorname{Tr}\rho = \frac{1}{d}\sum_{i=1}^{d^2} c_i \operatorname{Tr}(\Pi_i) = \frac{N}{d}. \tag{4.7}$$

Since $\operatorname{Tr}\rho = \operatorname{Tr}(\Pi_i) = 1$, $N = d$. Rearranging (4.6) gives

$$c_i = d(d+1)p(i) - 1 \tag{4.8}$$

which can be inserted into (4.5) to obtain

$$\rho = \sum_{i=1}^{d^2}\left((d+1)p(i) - \frac{1}{d}\right)\Pi_i \tag{4.9}$$

With this we have sucessfully found a mapping $||p\rangle\rangle \to \rho$. Now all that remains is to determine the constraints on $||p\rangle\rangle$ that determine whether a given probability vector actually represents a physical state.

To begin, we introduce structure coefficients α_{ijk}

$$\Pi_i \Pi_j = \sum_k \alpha_{ijk}\Pi_k \tag{4.10}$$

Taking the trace of both sides and using (4.3) gives

$$\sum_k \alpha_{ijk} = \frac{d\delta_{ij} + 1}{d+1} \tag{4.11}$$

Multiplying (4.10) on the right by Π_k and taking the trace gives

$$\begin{aligned}\mathrm{Tr}(\Pi_i\Pi_j\Pi_k) &= \sum_m \alpha_{ijm}\,\mathrm{Tr}(\Pi_m\Pi_k) = \sum_m \alpha_{ijm}\frac{d\delta_{km}+1}{d+1} = \alpha_{ijk} + \frac{1}{d+1}\sum_{m\neq k}\alpha_{ijm} \\ &= \alpha_{ijk}(1-\frac{1}{d+1}) + \frac{1}{d+1}\sum_m \alpha_{ijm} = \alpha_{ijk}(1-\frac{1}{d+1}) + \frac{1}{d+1}\frac{d\delta_{ij}+1}{d+1} \\ &= \frac{1}{d+1}\left[d\alpha_{ijk} + \frac{d\delta_{ij}+1}{d+1}\right] \end{aligned} \tag{4.12}$$

where in the second step I use the second part of (4.3), and in the last step of the second line I use (4.11).

Now (4.12) can be rearranged into

$$\alpha_{ijk} = \frac{1}{d}\Big((d+1)\,\mathrm{Tr}(\Pi_i\Pi_j\Pi_k) - \frac{d\delta_{ij}+1}{d+1}\Big) \tag{4.13}$$

Summing this over i gives

$$\sum_i \alpha_{ijk} = \frac{1}{d}\left((d+1)\,\mathrm{Tr}\left(\sum_i \Pi_i(\Pi_j\Pi_k)\right) - \frac{d+d^2}{d+1}\right) = (d+1)\,\mathrm{Tr}(\Pi_j\Pi_k) - 1 = (d\delta_{jk}+1) - 1 = d\delta_{jk} \tag{4.14}$$

Proceeding in the same manner with the summation over j

$$\sum_j \alpha_{ijk} = d\delta_{ik} \tag{4.15}$$

We now investigate the conditions a probability vector must satisfy if it is to represent a a pure state, for which the density matrix satsfies the idempotency condtion $\rho = \rho^2$. Writing $\rho = \sum c_i\Pi_i$, with c_i given by (4.8), and using this in $\rho = \sum c_i\Pi_i$ gives

$$\begin{aligned}\sum_k c_k\Pi_k &= \sum_{ij} c_ic_j\Pi_i\Pi_j = \sum_{ijk} c_ic_j\alpha_{ijk}\Pi_k \\ \Rightarrow \quad c_k &= \sum_{ij} c_ic_j\alpha_{ijk} \end{aligned} \tag{4.16}$$

Now replacing c_i, c_j and c_k using (4.8) and doing the summation

$$\begin{aligned}(d+1)p(k) - \frac{1}{d} &= \sum_{ij}\left[\big((d+1)p(k) - \frac{1}{d}\big)\big((d+1)p(k) - \frac{1}{d}\big)\alpha_{ijk}\right] \\ &= (d+1)^2\Big[\sum_{ij} p(i)p(j)\alpha_{ijk}\Big] - 2(d+1)p(k) + \frac{1}{d} \\ \Rightarrow p(k) &= \frac{1}{3}(d+1)\sum_{ij}\alpha_{ijk}p(i)p(j) + \frac{2}{3d(d+1)} \end{aligned} \tag{4.17}$$

This is the condition that the components of the probability vector must satisfy if it is to represent a pure state of a d-state system. However, this single condition can be broken up into two much simpler conditions if one uses the fact that a pure state in dimension 3 or larger obeys the conditions $\mathrm{Tr}(\rho^2) = 1$ and $\mathrm{Tr}(\rho^3) = 1$.

Using the fact that $\mathrm{Tr}\,\rho^2 = 1$ and (4.5) leads to the result

$$\mathrm{Tr}\,\rho^2 = \sum_{ijk} c_i c_j \alpha_{ijk}\,\mathrm{Tr}(\Pi_k) = \sum_{ij} c_i c_j \Big[\frac{d\delta_i j + 1}{d+1}\Big] = \sum_i c_i^2 + \frac{1}{d+1}\sum_{i,j\neq i} c_i c_j \tag{4.18}$$

Now $\mathrm{Tr}\,\rho = 1$ implies that

$$\sum_i c_i\,\mathrm{Tr}(\Pi_i) = 1 \Rightarrow \sum_i c_i = 1 \tag{4.19}$$

so

$$\sum_{i,j\neq i} c_i c_j = \sum_i c_i(1-c_i) = \sum_i c_i - \sum_i c_i^2 = 1 - \sum_i c_i^2. \tag{4.20}$$

Using (4.20) in (4.18) gives

$$\begin{aligned}\mathrm{Tr}\,\rho^2 &= \frac{1}{d+1} + \Big(1 - \frac{1}{d+1}\Big)\sum_i c_i^2 = \frac{1}{d+1} + \frac{d}{d+1}\sum_i\Big[(d+1)^2 p(i)^2 - \frac{2(d+1)}{d}p(i) + \frac{1}{d^2}\Big] \\ &= \frac{1}{d+1} + d(d+1)\sum_i p(i)^2 - 2 + \frac{d}{d+1} = d(d+1)\sum_i p(i)^2 - 1 \end{aligned} \tag{4.21}$$

Setting the final expression equal to 1 leads to

$$\sum_i p(i)^2 = \frac{2}{d(d+1)} \tag{4.22}$$

In a similar manner as (4.18 - 4.22), we can use $\mathrm{Tr}\,\rho^3 = 1$ to get

$$\rho^3 = \rho^2\rho = \sum_{ijk} c_i c_j \alpha_{ijk} \Pi_k \sum_l c_l \Pi_l = \sum_{ijkl} c_i c_j c_l \alpha_{ijk} \Pi_k \Pi_l = \sum_{ijklm} c_i c_j c_l \alpha_{ijk} \alpha_{klm} \Pi_m = 1 \tag{4.23}$$

So

$$1 = \sum_{ijkl} c_i c_j c_l \alpha_{ijk} \Big[\frac{d\delta_{kl}+1}{d+1}\Big] \quad \text{or} \quad \sum_{ijk} c_i c_j c_k \alpha_{ijk} + \frac{1}{d+1} \sum_{ijk,l\neq k} c_i c_j c_l \alpha_{ijk} = 1 \tag{4.24}$$

The second term on the left of (4.24) can be simplified as follows

$$\begin{aligned}
\sum_{ijk,l\neq k} c_i c_j c_l \alpha_{ijk} &= \sum_{ijk} c_i c_j (1-c_k)\alpha_{ijk} = \sum_{ij} c_i c_j \sum_k \alpha_{ijk} - \sum_{ijk} c_i c_j c_k \alpha_i jk \\
&= \sum_{ij} c_i c_j \frac{d\delta_i j + 1}{d+1} - \sum_{ijk} c_i c_j c_k \alpha_{ijk} = \sum_i c_i^2 + \frac{1}{d+1}\sum_i c_i(1-c_i) - \sum_{ijk} c_i c_j c_k \alpha_{ijk} \\
&= \frac{d}{d+1}\sum_i c_i + \frac{d}{d+1} - \sum_{ijk} c_i c_j c_k \alpha_{ijk}
\end{aligned} \tag{4.25}$$

Putting (4.25) into (4.24) gives

$$\frac{d}{d+1}\sum_{ijk} c_i c_j c_k \alpha_{ijk} + \frac{d}{(d+1)^2}\sum_i c_i^2 + \frac{1}{(d+1)^2} = 1 \tag{4.26}$$

Now

$$\sum_i c_i^2 = \sum_i \Big[(d+1)^2 p(i)^2 - \frac{2(d+1)}{d} p(i) + \frac{1}{d^2}\Big] = (d+1)^2 \frac{2}{d(d+1)} - \frac{2(d+1)}{d} + 1 = 1 \quad (4.27)$$

where (4.22) is used to simplify the first term at the second step. Putting (4.27) in (4.26) and simplifying gives

$$\sum_{ijk} c_i c_j c_k \alpha_{ijk} = 1 \quad (4.28)$$

Using $c_i = (d+1)p(i) - \frac{1}{d}$ allows to rewrite (4.28) as

$$\begin{aligned} &(d+1)^3 \sum_{ijk} \alpha_{ijk} p(i)p(j)p(k) - \frac{(d+1)^2}{d} \sum_{ijk} \alpha_{ijk} p(i)p(j) - \frac{2(d+1)^2}{d} \sum_{ijk} \alpha_{ijk} p(j)p(k) \\ &... + \frac{d+1}{d^2} \sum_{ijk} \alpha_{ijk} [p(i) + p(j) + p(k)] - \frac{1}{d^3} \sum_{ijk} \alpha_{ijk} = 1 \end{aligned} \quad (4.29)$$

The various sums on the left can be evaluated as follows

$$\begin{aligned} \sum_{ijk} \alpha_{ijk} p(i)p(j) &= \sum_{ij} p(i)p(j) \frac{d\delta_i j + 1}{d+1} = \sum_i \Big[p(i)^2 + \frac{1}{d+1} p(i)[1-p(i)]\Big] \\ &= \frac{d}{d+1} \sum_i p(i)^2 + \frac{1}{d+1} = \frac{2}{(d+1)^2} + \frac{1}{d+1} + \frac{d+3}{(d+1)^2} \end{aligned} \quad (4.30)$$

$$\sum_{ijk} \alpha_{ijk} p(j)p(k) = \sum_{jk} p(j)p(k) d\delta_{jk} = d \sum_j p(j)^2 = \frac{2}{d+1} \quad (4.31)$$

$$\begin{aligned} \sum_{ijk} \alpha_{ijk} p(i) &= \sum_{ijk} \alpha_{ijk} p(j) = \sum_{ijk} \alpha_{ijk} p(k) \\ &= d \sum_{ik} \delta_{ik} p(i) = d \sum_k p(k) = d \end{aligned} \quad (4.32)$$

Using (4.30), (4.31), and (4.32) in (4.29) allows it to be simplified to

$$(d+1)^3 \sum_{ijk} \alpha_{ijk} p(i)p(j)p(k) - 3 - \frac{4}{d} = 1 \tag{4.33}$$

from which we finally obtain

$$\sum_{ijk} \alpha_{ijk} p(i)p(j)p(k) = \frac{4}{d(d+1)^2} \tag{4.34}$$

.

Thus, we have acheived our final goal of obtaining the constraints on the probability vector if it is to represent a pure state. The constraints are the two equations (4.22) and (4.34). The discussion could conclude here, but there remains the complexity of α_{ijk} in (4.17) and (4.34). There are advantages to working with real numbers, and real forms of these equations can be obtained by replacing α_{ijk} with a real, symmetric tensor c_{ijk} with elements defined by

$$c_{ijk} = \mathrm{Re}\,\mathrm{Tr}(\Pi_i \Pi_j \Pi_k) \tag{4.35}$$

.

It can be seen that the elments are symmetric by writing c_{ijk} as

$$c_{ijk} = \mathrm{Re}\,\mathrm{Tr}(\Pi_i \Pi_j \Pi_k) = \langle\psi_i|\psi_j\rangle \langle\psi_j|\psi_k\rangle \langle\psi_k|\psi_i\rangle + \langle\psi_j|\psi_i\rangle \langle\psi_k|\psi_j\rangle \langle\psi_i|\psi_k\rangle \tag{4.36}$$

and noting that an exchange of any two indicies sends each term into the other.

Since $p(k)$ is purely real, only the real part of α_{ijk} is relevant in (4.17), so it can be written as

$$p(k) = \frac{1}{3}(d+1) \sum_{ij} \mathrm{Re}(\alpha_{ijk}) p(i)p(j) + \frac{2}{3d(d+1)} \tag{4.37}$$

From (4.13), we find that

$$\sum_{ij} \mathrm{Re}(\alpha_{ijk}) p(i)p(j) = \frac{d+1}{d} \sum_{ij} c_{ijk} p(i)p(j) - \frac{1}{d(d+1)} \sum_{ij} (d\delta_{ij} + 1) p(i)p(j) \tag{4.38}$$

The last term can be solved explicitly

$$\sum_{ij}(d\delta_{ij}+1)p(i)p(j) = d\sum_{i}p(i)^2+1 = \frac{2}{d+1}+1 = \frac{d+3}{d+1} \tag{4.39}$$

so (4.38) becomes

$$\sum_{ij}\mathrm{Re}(\alpha_{ijk})p(i)p(j) = \frac{d+1}{d}\sum_{ij}c_{ijk}p(i)p(j) - \frac{d+3}{d(d+1)^2} \tag{4.40}$$

Putting (4.40) in (4.37) gives

$$p(k) = \frac{(d+1)^2}{3d}\sum_{ij}c_{ijk}p(i)p(j) - \frac{d+3}{3d(d+1)} + \frac{2}{3d(d+1)} = \frac{(d+1)^2}{3d}\sum_{ij}c_{ijk}p(i)p(j) - \frac{1}{3d} \tag{4.41}$$

as a purely real form of (4.17).

Taking the real part of (4.34) and using (4.13) gives

$$\sum_{ijk}\mathrm{Re}(\alpha_{ijk})p(i)p(j)p(k) = \frac{d+1}{d}\sum_{ijk}c_{ijk}p(i)p(j)p(k) - \frac{1}{d(d+1)}\sum_{ijk}[d\delta_{ij}+1]p(i)p(j)p(k) \tag{4.42}$$

But

$$\sum_{ijk}[d\delta_{ij}+1]p(i)p(j)p(k) = d\sum_{i}p(i)^2+1 = \frac{2}{d+1}+1 = \frac{d+3}{d+1} \tag{4.43}$$

So (4.42) becomes

$$\sum_{ijk}\mathrm{Re}(\alpha_{ijk})p(i)p(j)p(k) = \frac{d+1}{d}\sum_{ijk}c_{ijk}p(i)p(j)p(k) - \frac{d+3}{d(d+1)^2} \tag{4.44}$$

Putting the right side of 4.44 equal to $\frac{4}{d(d+1)^2}$ and simplifying leads to

$$\sum_{ijk} c_{ijk} p(i) p(j) p(k) = \frac{d+7}{(d+1)^3} \tag{4.45}$$

With this, we have accomplished all our goals. To reiterate, we have shown how a SIC can be used to map the density matrix of a d-state system into a d^2-component probability vector (see (4.4)) and, conversely, how the probability vector can be mapped back into the density matrix to which it corresponds (see (4.9)). Further, we have shown what conditions the probability vector must satisfy if it is to represent the density matrix of a pure state of d-state system; these are given by (4.22) and either (4.34) or (4.45). Some illustrations of these results will be discussed for two-state systems in the next chapter.

5 SICs in 2 and 3 Dimensions

With an understanding of how SICs can be used to represent quantum systems in any dimension, we see how the formalism works out for the lowest values of d. We first discuss some features of the SIC description in $d = 2$ and then do the same in $d = 3$.

5.1 d=2

The SIC basis for a two-state system is defined by the four operators [1]

$$\begin{aligned}
\Pi_1 &= |\psi_1\rangle\langle\psi_1| = \frac{1}{2}\Big(I + \frac{1}{\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z)\Big) \\
\Pi_2 &= |\psi_2\rangle\langle\psi_2| = \frac{1}{2}\Big(I + \frac{1}{\sqrt{3}}(\sigma_x - \sigma_y - \sigma_z)\Big) \\
\Pi_3 &= |\psi_3\rangle\langle\psi_3| = \frac{1}{2}\Big(I + \frac{1}{\sqrt{3}}(-\sigma_x - \sigma_y + \sigma_z)\Big) \\
\Pi_4 &= |\psi_4\rangle\langle\psi_4| = \frac{1}{2}\Big(I + \frac{1}{\sqrt{3}}(-\sigma_x + \sigma_y - \sigma_z)\Big)
\end{aligned} \tag{5.1}$$

Although it remains to be explained where this set of operators comes from, it can be verified that they constitute a SIC by checking that they satisfy (4.3) with $d = 2$, i.e.

$$\mathrm{Tr}(\Pi_i\Pi_j) = \begin{cases} \frac{1}{3}, & \text{if } i \neq j \\ 1, & \text{if } i = j. \end{cases} \tag{5.2}$$

The wavevectors $|\psi_1\rangle$-$|\psi_4\rangle$ which generate (5.1) and their respective pseudospin vectors $\vec{\mathbf{s_1}}$-$\vec{\mathbf{s_4}}$ are

$$\begin{aligned}
|\psi_1\rangle &= \frac{1}{\sqrt{2\sqrt{3}}}\left[\sqrt{\sqrt{3}+1}\,|0\rangle + e^{\frac{i\pi}{4}}\sqrt{\sqrt{3}-1}\,|1\rangle\right], \quad \vec{\mathbf{s_1}} = \frac{1}{\sqrt{3}}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \\
|\psi_2\rangle &= \frac{1}{\sqrt{2\sqrt{3}}}\left[\sqrt{\sqrt{3}-1}\,|0\rangle + e^{\frac{7i\pi}{4}}\sqrt{\sqrt{3}+1}\,|1\rangle\right], \quad \vec{\mathbf{s_2}} = \frac{1}{\sqrt{3}}(\hat{\mathbf{x}} - \hat{\mathbf{y}} - \hat{\mathbf{z}}) \\
|\psi_3\rangle &= \frac{1}{\sqrt{2\sqrt{3}}}\left[\sqrt{\sqrt{3}+1}\,|0\rangle + e^{\frac{5i\pi}{4}}\sqrt{\sqrt{3}-1}\,|1\rangle\right], \quad \vec{\mathbf{s_3}} = \frac{1}{\sqrt{3}}(-\hat{\mathbf{x}} - \hat{\mathbf{y}} + \hat{\mathbf{z}}) \\
|\psi_4\rangle &= \frac{1}{\sqrt{2\sqrt{3}}}\left[\sqrt{\sqrt{3}-1}\,|0\rangle + e^{\frac{3i\pi}{4}}\sqrt{\sqrt{3}+1}\,|1\rangle\right], \quad \vec{\mathbf{s_4}} = \frac{1}{\sqrt{3}}(-\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})
\end{aligned} \tag{5.3}$$

In the Bloch sphere picture, these vectors form a regular tetrahedron

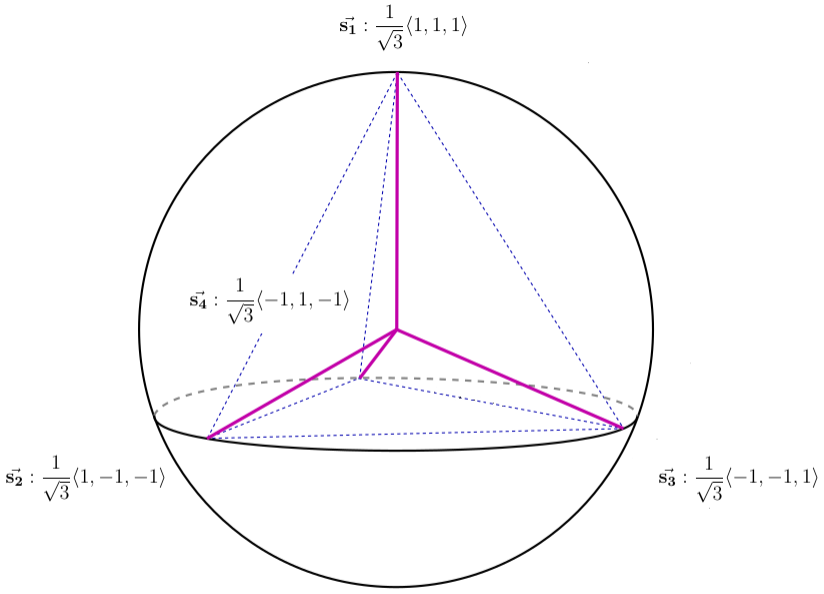

Figure 2: The Bloch sphere with the pseudospin vectors that generate the SICs in 2 dimensions. Notice that the sphere has been rotated to make the tetrahedron more obvious.

The density matrix with the pseudospin vector $\vec{\mathbf{s}} = (s_x, s_y, s_z)$ is represented by the probability vector with the components

$$\begin{aligned} p(1) &= \frac{1}{4}\Big[1 + \frac{1}{\sqrt{3}}\big(s_x + s_y + s_z\big)\Big] \quad p(2) = \frac{1}{4}\Big[1 + \frac{1}{\sqrt{3}}\big(s_x - s_y - s_z\big)\Big] \\ p(3) &= \frac{1}{4}\Big[1 + \frac{1}{\sqrt{3}}\big(-s_x - s_y + s_z\big)\Big] \quad p(4) = \frac{1}{4}\Big[1 + \frac{1}{\sqrt{3}}\big(-s_x + s_y - s_z\big)\Big] \end{aligned} \tag{5.4}$$

The four probabilites sum to 1, so this represents a valid probability distribution. It is also possible to verify (4.22)

$$\sum_i p(i)^2 = \frac{2}{d(d+1)} = \frac{1}{3} \tag{5.5}$$

We now verify (4.45), which requires calculating the c_{ijk} tensor. Putting (5.1) into (4.35)

allows us to calculate c_{ijk}, which is written here as 4 4x4 C_k matricies ($(C_k)_{\text{ij}} = c_{\text{ijk}}$)

$$C_1 = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{bmatrix} C_2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix} C_3 = \begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 1 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{bmatrix} C_4 = \begin{bmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{bmatrix} \tag{5.6}$$

By inspection, we observe properties of the c_{ijk} tensor specific to the qubit

1. If $i = j = k$, then $c_{\text{ijk}} = 1$
2. If $i \neq j$, $j \neq k$, $i \neq k$, then $c_{\text{ijk}} = 0$
3. Otherwise, $c_{\text{ijk}} = \frac{1}{3}$

These rules can be used in (4.45)

$$\sum_{ijk} c_{ijk} p(i) p(j) p(k) = \frac{d+7}{(d+1)^3} \tag{5.7}$$

to give the simpler expression

$$\sum_i p(i)^3 + \sum_i \sum_{j \neq i} p(i)^2 p(j) = \frac{d+7}{(d+1)^3} \tag{5.8}$$

The first term comes from rule 1, and the second comes from rule 3, noting that the set of indicies that fit rule 3 come in triplets, so the factor of $\frac{1}{3}$ cancels. The two terms on the left of (5.8) can be combined into a single term to give

$$\sum_i p(i)^2 \cdot \sum_i p(i) = \frac{d+7}{(d+1)^3} \tag{5.9}$$

or

$$\frac{2}{d(d+1)} = \frac{d+7}{(d+1)^3} \tag{5.10}$$

Which is always satisfied for $d = 2$.

5.2 d=3

The SIC basis for a three-state system (qutrit) is defined by the nine operators [1]

$$\begin{aligned}
|\psi_0\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\end{pmatrix}, \quad |\psi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}, \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \\
|\psi_3\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix}0\\\omega\\-\bar{\omega}\end{pmatrix}, \quad |\psi_4\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\\bar{\omega}\end{pmatrix}, \quad |\psi_5\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-\omega\\0\end{pmatrix}, \\
|\psi_6\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix}0\\\bar{\omega}\\-\omega\end{pmatrix}, \quad |\psi_7\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\\omega\end{pmatrix}, \quad |\psi_8\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-\bar{\omega}\\0\end{pmatrix}
\end{aligned} \tag{5.11}$$

where ω is the cube root of unity $e^{\frac{2\pi i}{3}}$ and $\bar{\omega}$ is its complex conjugate.

As before with the $d = 2$ case, although it remains to be explained where this set of operators comes from, it can be verified that they constitute a SIC by checking that they satisfy (4.3) with $d = 3$, i.e.

$$\mathrm{Tr}(\Pi_i \Pi_j) = \begin{cases} \frac{1}{4}, & \text{if } i \neq j \\ 1, & \text{if } i = j. \end{cases} \tag{5.12}$$

Unlike the qubit, the phase space of a qutrit cannot be expressed as the set of points on a Bloch sphere. However, it is possible to represent a qutrit as a symmetric outer product of 2 qubit states. At this point, it is convenient to use stereographic projection to map the points on the unit sphere to the points on the complex plane. Imagine the Bloch sphere to be bisected by the x-y plane, with the x- and y- axes being taken to be real and imaginary and imaginary axes for complex numbers in the plane. The line drawn from the south pole of the sphere to the tip of the pseudospin vector $\vec{\mathbf{a}}$ will intersect this complex plane at the unique point α

$$\vec{\mathbf{a}}(\theta, \phi) \to \alpha = \cos\frac{\theta}{2}\cos\phi + i\sin\frac{\theta}{2}\sin\phi. \tag{5.13}$$

It was shown in the early decades of Quantum Mechanics by Majorana [5] [6] that any pure state of a qutrit can be expressed in terms of two pseudospin vectors. If these vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ go into the complex numbers α and β under stereographic projection, then the qubit state can be expressed in terms of them as

$$|\alpha, \beta\rangle = \sqrt{\frac{1}{8}\frac{1}{(1+|\alpha|^2)(1+|\beta|^2)+|\alpha+\beta|^2}}\Big[\,|0\rangle + \frac{\alpha+\beta}{\sqrt{2}}\,|1\rangle + \alpha\beta\,|2\rangle\,\Big] \tag{5.14}$$

Using (4.4), it is possible to compute the 9 entries of the probabilty vector p_i. Each term begins with an identical constant F

$$F = \frac{1}{24}\frac{1}{(1+|\alpha|^2)(1+|\beta|^2)+|\alpha+\beta|^2}; \tag{5.15}$$

$$\begin{aligned}
p_1 &= F[|\alpha|^2+|\beta|^2+2|\alpha|^2|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(|\alpha|^2\overline{\beta}+\overline{\alpha}|\beta|^2+|\alpha|^2\beta)+\alpha|\beta|^2] \\
p_2 &= F[2+2|\alpha|^2|\beta|^2-2(\alpha\beta+\overline{\alpha}\overline{\beta})] \\
p_3 &= F[2+|\alpha|^2+|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(\alpha+\overline{\alpha}+\beta+\overline{\beta})] \\
p_4 &= F[|\alpha|^2+|\beta|^2+2|\alpha|^2|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(\omega|\alpha|^2\overline{\beta}+\omega\overline{\alpha}|\beta|^2+\overline{\omega}|\alpha|^2\beta)+\overline{\omega}\alpha|\beta|^2] \\
p_5 &= F[2+2|\alpha|^2|\beta|^2-2(\omega\alpha\beta+\overline{\omega\alpha}\overline{\beta})] \\
p_6 &= F[2+|\alpha|^2+|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(\omega\alpha+\overline{\omega\alpha}+\omega\beta+\overline{\omega}\overline{\beta})] \\
p_7 &= F[|\alpha|^2+|\beta|^2+2|\alpha|^2|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(\overline{\omega}|\alpha|^2\overline{\beta}+\overline{\omega\alpha}|\beta|^2+\omega|\alpha|^2\beta)+\omega\alpha|\beta|^2] \\
p_8 &= F[2+2|\alpha|^2|\beta|^2-2(\overline{\omega}\alpha\beta+\omega\overline{\alpha}\overline{\beta})] \\
p_9 &= F[2+|\alpha|^2+|\beta|^2+\alpha\overline{\beta}+\overline{\alpha}\beta-\sqrt{2}(\overline{\omega}\alpha+\omega\overline{\alpha}+\overline{\omega}\beta+\omega\overline{\beta})]
\end{aligned} \tag{5.16}$$

Doing a bit of algebra and using the identity $1+\omega+\overline{\omega}=0$, it can be shown that the p_i values sum to 1, so this clearly works as a probability distribution. It is also possible to verify (both numerically and algebraically) (4.22)

$$\sum_i p(i)^2 = \frac{2}{d(d+1)} = \frac{1}{6} \tag{5.17}$$

6 Group Theory of SICs

Despite a general belief in those who use and study SICs that they exist in every dimension, a proof of this has yet to be demonstrated. Despite this, work has been done to find ways to generate SICs. All known SICs share the property of being group covariant. This means that the basis kets $|\psi_i\rangle$ can be generated from the orbit of a fiducial vector under a particular group. In this section we discuss this group, following the presentation in Fuchs et al [3] and show how SICs can be generated for $d = 2, 3, 4$.

6.1 Weyl-Heisenberg Group

Consider a d-dimensional Hilbert space with an orthonormal basis given by $\{|0\rangle, |1\rangle, ... |d-1\rangle\}$. Define a shift operator X and a phase operator Z such that

$$X|j\rangle = |j+1\rangle \qquad Z|j\rangle = \omega^j |j\rangle \tag{6.1}$$

where $\omega = e^{2\pi i/d}$.

These operators satisfy the commutation relation

$$X^l Z^\alpha = \omega^{-l\alpha} Z^\alpha X^l \tag{6.2}$$

From this we can construct the displacement operators

$$D_{l\alpha} = (-e^{i\pi/d})^{l\alpha} X^l Z^\alpha \tag{6.3}$$

The product of two displacement operators is another displacement operator up to a phase factor

$$D_{l\alpha} D_{m\beta} = (-e^{i\pi/d})^{\alpha m - \beta l} D_{l+m,\alpha+\beta} \tag{6.4}$$

By allowing the multiplication of the group elements by a phase factor, X and Z become the generators of a group known as the Weyl-Heisenberg group. Assuming one is able to determine a suitable fiducial vector $|\psi_0\rangle$, the displacement operators can be applied to $|\psi_0\rangle$ to generate SICs. The only challenge that remains to would-be SIC finders is producing a valid fiducial vector in the given dimension d.

6.2 Generation of SICs

6.2.1 d=2

For $d = 2$, the operators X and Z are given as

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{6.5}$$

which are simply the Pauli matricies σ_x and σ_z.

Using the fiducial vector given as

$$|\psi_0\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3+\sqrt{3}} \\ e^{i\pi/4}\sqrt{3-\sqrt{3}} \end{pmatrix} \tag{6.6}$$

we can apply all the distinct elements of the Weyl-Heisenberg group to get the remaining members of the SIC as

$$|\psi_1\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3+\sqrt{3}} \\ -e^{i\pi/4}\sqrt{3-\sqrt{3}} \end{pmatrix}, \quad |\psi_2\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3-\sqrt{3}} \\ e^{-i\pi/4}\sqrt{3+\sqrt{3}} \end{pmatrix}, \tag{6.7}$$
$$|\psi_3\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3-\sqrt{3}} \\ -e^{-i\pi/4}\sqrt{3+\sqrt{3}} \end{pmatrix}$$

Under (4.3), this meets the criteria for a 2-d SIC. It is not the same as the basis given in (5.1), however there exists a unitary transformation that rotates the Bloch sphere onto itself such that the two basis are equivalent.

6.2.2 d=3

For $d = 3$, the operators X and Z are given as

$$X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{bmatrix} \tag{6.8}$$

where $\omega = e^{2\pi i/3}$

Using the fiducial vector given as

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \tag{6.9}$$

we can apply all the distinct elements of the Weyl-Heisenberg group to get the remaining members of the SIC as

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}, \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix},$$
$$|\psi_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\\omega\\-\bar{\omega}\end{pmatrix}, \quad |\psi_4\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\\bar{\omega}\end{pmatrix}, \quad |\psi_5\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-\omega\\0\end{pmatrix}, \tag{6.10}$$
$$|\psi_6\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\\bar{\omega}\\-\omega\end{pmatrix}, \quad |\psi_7\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\\omega\end{pmatrix}, \quad |\psi_8\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-\bar{\omega}\\0\end{pmatrix}$$

This matches (5.11) therefore meeting the criteria for a 3-d SIC.

6.2.3 d=4

For $d = 4$, the operators X and Z are given as

$$X = \begin{bmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}, \quad Z = \begin{bmatrix}1&0&0&0\\0&i&0&0\\0&0&-1&0\\0&0&0&-i\end{bmatrix} \tag{6.11}$$

a particular fiducial vector is given as

$$|\psi_0\rangle = \sqrt{\frac{5-\sqrt{5}}{40}}\begin{pmatrix}2\cos\frac{\pi}{8}\\i(e^{\frac{-i\pi}{8}} + e^{\frac{i\pi}{8}}(2+\sqrt{5})^{\frac{1}{2}})\\2i\sin\frac{\pi}{8}\\i(e^{\frac{-i\pi}{8}} - e^{\frac{i\pi}{8}}(2+\sqrt{5})^{\frac{1}{2}})\end{pmatrix} \tag{6.12}$$

The set of 16 $|\psi_i\rangle$ vectors generated by the action of the Weyl-Heisenberg group on this fiducial vector can be verified to satisfy the conditions for a SIC. We do not list the SIC vectors explicitly here as we will not use them.

7 Applications

Quantum State Tomography

Quantum state tomography describes the set of approaches one may use to recover the state of a quantum system. This is a core problem in quantum mechanics, dating back to 1933 when Pauli asked if it were possible to determine a particle's wavefunction from the probability densities of its position and momentum. Since measuring a quantum system destroys it, an ensemble of identical states must be prepared and measured in order to recover any useful information about the system. For the problem of quantum computation, minimizing the size of this ensemble is of paramount importance. The fewer measurements required to asses the configuration of a quantum computer bit, the less redundant bits are required to make quantum computer chips. By their definition, SICs are a perfect option as a minimum-measurement basis for any d-dimensional quantum system. Extensive work has been done demonstrating the supreme efficiency of SICs as a basis [7], as well as expanding on their utility in quantum state tomography. One recent example is given by Czerwinski where it is possible to use existing understanding of the time evolution of a qubit or qutrit to reconstruct a quantum state even with limited measurement potential [8].

Quantum Key Distribution

Beyond just quantum computing, SICs also have potential applications in quantum cryptography. Quantum communication offers an additional level of security over classical methods. Because quantum systems are affected by measurement, careful choice of a communication protocol makes it possible to detect eavesdropping. One example is quantum key distribution (QKD), where both parties cooperate to create a shared secret key, the true secrecy of which can be verified via quantum mechanics and standard techniques from classical cryptography.

Tavakoli et al. have used SICs to develop a quantum key distribution method which outperforms conventional methods [9]. In order to achieve this, they first define a structure called a SIC-compound. This is a set of d^3 quantum states, each denoted $|\psi_{jk}\rangle$ for $j \in [d^2]$ and $k \in [d]$, where each set of d^2 states $\{|\psi_{jk}\rangle\}_k$ form a SIC and each set of d states $\{|\psi_{jk}\rangle\}_j$ form an orthonormal basis. For the $d = 2$ case, a SIC-compound can be constructed from the 4 states in any 2d SIC and the 4 states orthogonal to the first 4. In the Bloch sphere picture, this is represented as the tetrahdron from Figure 2 and 4 more vectors pointing in the opposite direction. It is shown that no such SIC-compund can be constructed for $d = \{3, 5, 6, 7, 8\}$, however such a construction is possible for $d = 4$. It is further demonstrated that a QKD protocol using the 16 orthonormal bases formed by this 4d SIC-compound outperformes other competing schemes.

Signal Analysis

Although we have developed SICs here as an aid in our approach to quantum mechanics, the mathematical properties of SICs have enabled them to be utilized by those working in areas outside of Quantum Information Theory. Particularly, the symmetric properties of the Heisenberg-Weyl group present a convenient basis to explore concepts in signal

analysis. In one instance, the Heisenberg-Weyl group has been used in the development of adaptive radar, specifically spreading sequeneces and error-correcting codes [10]. Through the lens of group theory, Howard et al. make the observation that a host of error correcting codes are associated with abelian decomposition of the Heisienberg-Weyl group. As such, this group makes it possible to greatly simplify and unify many ideas in the area of radar and communications. In another case, the SIC framework has been applied to signal reconstruction, specifically voice recognition. The problem of signal reconstruction involves reconstructing a vector in a Hilbert space when only the outcomes of non-complete measurements of that vector are known. Interestingly, this is just the problem of quantum state tomography, however, as demonstrated in [11], there are applications of this problem to speech analysis and recogntion.

8 Conclusion

This MQP has given a basic introduction to SICs and the role they play in the QBist interpretation of quantum mechanics, following the treatment given in the review article by Fuchs and Schack [1]. Nevertheless, there remains plenty more to be discussed with respect to SICs, both with respect to their mathematical properties and their broad applications in science. For future MQPs, there are plenty of areas that could be explored in addition to what we have presented here. For instance, the unsolved problem of finding SICs in all dimensions (should they even exist in the first place) has been an area of intense study that seems worth exploring in greater detail. The applications of SICs and the Weyl-Heisenberg group to quantum cryptography protocols also seems worth studying in view of its important practical applications.

References

[1] Fuchs C., & Schack R. (2013). "Quantum-Bayesian coherence." *Reviews of Modern Physics*, 85(4), 1693-1715.

[2] Feynman R. (1987). "Negative probabilities." *Quantum Implications: Essays in Honor of David Bohm*, 235-248.

[3] Fuchs C., Hoang, M. & Stacey B. (2017). "The SIC Question: History and State of Play" *Axioms*, 6(4), 21-.

[4] Appleby D. (2005). "SIC-POVMs and the Extended Clifford Group." *Journal of Mathematical Physics*, 46, 052107.

[5] Majorana E. (1932) "Atomi Orientati in Campo Magnetico Variabile" *Nuovo Cimento*, 9, 43-50.

[6] Aravind P. (2017) "MUBs and SIC-POVMs of a spin-1 system from the Majorana approach"

[7] Scott A. (2006). "Tight informationally complete quantum measurements." *Journal of Physics A: Mathematical and General*, 39 13507.

[8] Czerwinski A. (2021). "Quantum state tomography with informationally complete POVMs generated in the time domain." *Quantum Information Processing*, 20 105.

[9] Tavakoli A., Bengtsson I., Gisin N., & Renes J. (2020). "Compounds of symmetric informationally complete measurements and their application in quantum key distribution." *Physical Review Research* 2, 043122.

[10] Howard S., Calderbank A., & Moran, W. (2006) "The Finite Heisenberg-Weyl Groups in Radar and Communications" *EURASIP Journal on Applied Signal Processing*, 85685, 1-12.

[11] Balan R., Bodmann B., Casazza P., & Edidin D. (2009) "Painless reconstruction from magnitudes of frame coefficients" *Journal of Fourier Analysis and Applications*, 15 488-501.