ADVANCES IN ATOMIC, MOLECULAR AND OPTICAL PHYSICS, VOL. 52

PHOTONIC STATE TOMOGRAPHY

J.B. ALTEPETER, E.R. JEFFREY and P.G. KWIAT

Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract

Quantum state tomography is the process by which an identical ensemble of unknown quantum states is completely characterized. A sequence of identical measurements within a series of different bases allow the reconstruction of a complete quantum wavefunction. This article reviews state representation and notation, lays out the theory of ideal tomography, and details the full experimental realization (measurement, electronics, error correction, numerical analysis, measurement choice, and estimation of uncertainties) of a tomographic system applied to polarized photonic qubits.

ISSN 1049-250X
DOI 10.1016/S1049-250X(05)52003-2

Unlike their classical counterparts, quantum states are notoriously difficult to measure. In one sense, the spin of an electron can be in only one of two states, up or down. A simple experiment can discover which state the electron occupies, and further measurements on the same electron will always confirm this answer. However, the simplicity of this picture belies the complex, complete nature of an electron which always appears in one of exactly two states—states which *change* depending on how it is measured.

Quantum state tomography is the process by which any quantum system, including the spin of an electron, can be characterized completely using an ensemble of many identical particles. Measurements of multiple types reconstruct a quantum state from different eigenbases, just as classical tomography can image a three-dimensional object by scanning it from different physical directions. Additional measurements in any single basis bring that dimension into sharper relief.

This article is structured into two major partitions:[1] the theory of tomography (Sections 1 and 2) and the experimental tomography of photonic systems (Sections 3–6). The theoretical sections provide a foundation for quantum state tomography, and should be applicable to any system, including photons (White et al., 1999; Sanaka et al., 2001; Mair et al., 2001; Nambu et al., 2002; Giorgi et al., 2003; Yamamoto et al., 2003; Sergienko et al., 2003; Pittman et al., 2003; O'Brien et al., 2003; Marcikic et al., 2003), spin-$\frac{1}{2}$ particles (as, e.g., are used in NMR quantum computing (Cory et al., 1997; Jones et al., 1997; Weinstein et al., 2001; Laflamme et al., 2002)), and (effectively) 2-level atoms (Monroe, 2002; Schmidt-Kaler et al., 2003). Section 1 provides an introduction to state representation and the notation of this article. Section 2 describes the theory of tomographic reconstruction assuming error-free, exact measurements. The second part of the article contains not only information specific to the experimental measurement of photon polarization (e.g., how to deal with imperfect waveplates), but extensive information on how to deal with real, error-prone systems; information useful to anyone implementing a real tomography system. Section 3 concerns the collection of experimental data (projectors, electronics, systematic error correction) and Section 4 deals with its analysis (numerical techniques for reconstructing states). Sections 5 and 6 describe how to choose which measurements to make and how to estimate the uncertainty in a tomography, respectively.

[1] This manuscript is based on a shorter article (Altepeter et al., 2004) which appeared in the special volume *Quantum State Estimation*; here we have rewritten that article in order to be specific to polarization-based photonic tomography and extended the results to include qudits, imperfect waveplates, a new type of maximum likelihood techniques, and information on the choice of measurements. Because the conceptual background is identical, some of the text and figures have been borrowed from that earlier work.

In order to facilitate the use of these techniques by groups and individuals working in any field, a website is available which provides both further details about these techniques and working, documented code for implementing them.[2]

1. State Representation

Before states can be analyzed, it is necessary to understand their representation. In particular, the reconstruction of an unknown state is often simplified by a specific state parametrization.

1.1. Representation of Single-Qubit States

Rather than begin with a general treatment of tomography for an arbitrary number of qubits, throughout this chapter the single-qubit case is investigated initially. This provides the opportunity to strengthen an intuitive grasp of the fundamentals of state representation and tomography before moving on to the more complex (and more useful) general case. In pursuance of this goal, we use graphical representations available only at the single-qubit level.

1.1.1. Pure States, Mixed States, and Diagonal Representations

In general, any single qubit in a pure state can be represented by

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1}$$

where α and β are complex and $|\alpha|^2 + |\beta|^2 = 1$ (Nielsen and Chuang, 2000). If the normalization is written implicitly and the global phase is ignored, this can be rewritten as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|1\rangle. \tag{2}$$

Example 1 (*Pure states*). Throughout this chapter, examples will be provided using qubits encoded into the electric field polarization of photons. For a single photon, this system has two levels, e.g., horizontal ($|H\rangle \equiv |0\rangle$) and vertical ($|V\rangle \equiv |1\rangle$), with all possible pure polarization states constructed from coherent superpositions of these two states. For example, diagonal, antidiagonal, right-circular and left-circular light are respectively represented by

[2] http://www.physics.uiuc.edu/research/QuantumPhotonics/Tomography/

$$
\begin{aligned}
|D\rangle &\equiv \frac{|H\rangle + |V\rangle}{\sqrt{2}}, \\
|A\rangle &\equiv \frac{|H\rangle - |V\rangle}{\sqrt{2}}, \\
|R\rangle &\equiv \frac{|H\rangle + \mathrm{i}|V\rangle}{\sqrt{2}}, \quad \text{and} \\
|L\rangle &\equiv \frac{|H\rangle - \mathrm{i}|V\rangle}{\sqrt{2}}.
\end{aligned}
\tag{3}
$$

This representation enables the tomography of an ensemble of *identical* pure states, but is insufficient to describe either an ensemble containing a variety of *different* pure states or an ensemble whose members are not pure (perhaps because they are entangled to unobserved degrees of freedom). In this case the overall state is *mixed*.

In general, these mixed states may be described by a probabilistically weighted incoherent sum of pure states, i.e., they behave as if any particle in the ensemble has a specific probability of being in a given pure state, and this state is distinguishably labelled in some way. If it were not distinguishable, the total state's constituent pure states would add coherently (with a definite relative phase), yielding a single pure state.

A mixed state can be represented by a density matrix $\hat{\rho}$, where

$$
\hat{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i| = \begin{matrix} \\ |0\rangle \\ |1\rangle \end{matrix} \begin{matrix} \begin{matrix} \langle 0| & \langle 1| \end{matrix} \\ \begin{pmatrix} A & Ce^{\mathrm{i}\phi} \\ Ce^{-\mathrm{i}\phi} & B \end{pmatrix} \end{matrix}. \tag{4}
$$

P_i is the probabilistic weighting ($\sum_i P_i = 1$), A, B and C are all real and non-negative, $A + B = 1$, and $C \leq \sqrt{AB}$ (Nielsen and Chuang, 2000).

While any ensemble of pure states can be represented in this way, it is also true that *any* ensemble of single-qubit states can be represented by an ensemble of only *two* orthogonal pure states. (Two pure states $|\psi_i\rangle$ and $|\psi_j\rangle$ are orthogonal if $|\langle\psi_i|\psi_j\rangle| = 0$.) For example, if the matrix from Eq. (4) were diagonal, then it would clearly be a probabilistic combination of two orthogonal states, as

$$
\begin{matrix} \\ |0\rangle \\ |1\rangle \end{matrix} \begin{matrix} \begin{matrix} \langle 0| & \langle 1| \end{matrix} \\ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \end{matrix} \equiv A|0\rangle\langle 0| + B|1\rangle\langle 1|. \tag{5}
$$

However, *any* physical density matrix can be diagonalized, such that

$$\hat{\rho} = \begin{matrix} \\ |\psi\rangle \\ |\psi^{\perp}\rangle \end{matrix} \begin{matrix} \langle\psi| \; \langle\psi^{\perp}| \\ \end{matrix} \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} = E_1 |\psi\rangle\langle\psi| + E_2 |\psi^{\perp}\rangle\langle\psi^{\perp}|, \tag{6}$$

where $\{E_1, E_2\}$ are the eigenvalues of $\hat{\rho}$, and $\{|\psi\rangle, |\psi^{\perp}\rangle\}$ are the eigenvectors (recall that these eigenvectors are always mutually orthogonal, denoted here by the $\perp$ symbol). Thus the representation of any quantum state, no matter how it is constructed, is identical to that of an ensemble of two orthogonal pure states.[3]

EXAMPLE 2 (*A mixed state*). Now consider measuring a source of photons which emits a one-photon wave packet each second, but alternates—perhaps randomly—between horizontal, vertical, and diagonal polarizations. Their emission time labels these states (in principle) as distinguishable, and so if we ignore that timing information when they are measured, we must represent their state as a density matrix $\hat{\rho}$:

$$\begin{aligned} \hat{\rho} &= \frac{1}{3}\big(|H\rangle\langle H| + |V\rangle\langle V| + |D\rangle\langle D|\big) \\ &= \frac{1}{3}\left(\begin{matrix} \\ |H\rangle \\ |V\rangle \end{matrix} \begin{matrix} \langle H| \; \langle V| \\ \end{matrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{matrix} |H\rangle \\ |V\rangle \end{matrix} \begin{matrix} \langle H| \; \langle V| \\ \end{matrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{matrix} |H\rangle \\ |V\rangle \end{matrix} \begin{matrix} \langle H| \; \langle V| \\ \end{matrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \right) \\ &= \frac{1}{6}\left(\begin{matrix} \\ |H\rangle \\ |V\rangle \end{matrix} \begin{matrix} \langle H| \; \langle V| \\ \end{matrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \right). \end{aligned} \tag{7}$$

[3] It is an interesting question whether all physical states described by a mixed state—e.g., Eq. (6)—are indeed completely equivalent. For example, Lehner, Leonhardt, and Paul discussed the notion that two types of unpolarized light could be considered, depending on whether the incoherence between polarization components arose purely due to an averaging over rapidly varying phases, or from an entanglement with another quantum system altogether (Lehner et al., 1996). This line of thought can even be pushed further, by asking whether all mixed states necessarily arise only from tracing over some unobserved degrees of freedom with which the quantum system has become entangled, or if indeed such entanglement may 'collapse' when the systems involved approach macroscopic size (Kwiat and Englert, 2004). If the latter were true, then there would exist mixed states that could *not* be seen as pure in some larger Hilbert space. In any event, these subtleties of interpretation do not in any way affect experimental results, at least insofar as state tomography is concerned.

When diagonalized,

$$\hat{\rho} = \frac{1}{3}\begin{pmatrix} & \langle D| & \langle A| \\ |D\rangle & \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \\ |A\rangle & \end{pmatrix} = \frac{2}{3}|D\rangle\langle D| + \frac{1}{3}|A\rangle\langle A|, \tag{8}$$

which, as predicted in Eq. (6), is a sum of only *two* orthogonal states.

Henceforth, the 'bra' and 'ket' labels will be suppressed from written density matrices where the basis is $\{|0\rangle, |1\rangle\}$ or $\{|H\rangle, |V\rangle\}$.

1.1.2. The Stokes Parameters and the Poincaré Sphere

Any single-qubit density matrix $\hat{\rho}$ can be represented uniquely by three parameters $\{S_1, S_2, S_3\}$:

$$\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i. \tag{9}$$

The $\hat{\sigma}_i$ matrices are

$$\begin{aligned} \hat{\sigma}_0 &\equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \hat{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \hat{\sigma}_2 &\equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \hat{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{10}$$

and the S_i values are given by

$$S_i \equiv \mathrm{Tr}\{\hat{\sigma}_i \hat{\rho}\}. \tag{11}$$

For all pure states, $\sum_{i=1}^{3} S_i^2 = 1$; for mixed states, $\sum_{i=1}^{3} S_i^2 < 1$; for the completely mixed state, $\sum_{i=1}^{3} S_i^2 = 0$. Due to normalization, S_0 will always equal one.

Physically, each of these parameters corresponds directly to the outcome of a specific pair of projective measurements:

$$\begin{aligned} S_0 &= P_{|0\rangle} + P_{|1\rangle}, \\ S_1 &= P_{1/\sqrt{2}(|0\rangle+|1\rangle)} - P_{1/\sqrt{2}(|0\rangle-|1\rangle)}, \\ S_2 &= P_{1/\sqrt{2}(|0\rangle+i|1\rangle)} - P_{1/\sqrt{2}(|0\rangle-i|1\rangle)}, \\ S_3 &= P_{|0\rangle} - P_{|1\rangle}, \end{aligned} \tag{12}$$

where $P_{|\psi\rangle}$ is the probability to measure the state $|\psi\rangle$. As we shall see below, these relationships between probabilities and S parameters are extremely useful

in understanding more general operators. Because $P_{|\psi\rangle} + P_{|\psi^{\perp}\rangle} = 1$, these can be simplified in the single-qubit case, and

$$P_{|\psi\rangle} - P_{|\psi^{\perp}\rangle} = 2P_{|\psi\rangle} - 1. \tag{13}$$

The probability of projecting a given state $\hat{\rho}$ into the state $|\psi\rangle$ (the probability of measuring $|\psi\rangle$) is given by (Gasiorowitz, 1996):

$$P_{|\psi\rangle} = \langle\psi|\hat{\rho}|\psi\rangle = \mathrm{Tr}\{|\psi\rangle\langle\psi|\hat{\rho}\}. \tag{14}$$

In Eqs. (12) above, the S_i are defined with respect to three states, $|\phi\rangle_i$:

$$\begin{aligned} |\phi\rangle_1 &= \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \\ |\phi\rangle_2 &= \frac{1}{\sqrt{2}}\big(|0\rangle + \mathrm{i}|1\rangle\big) \\ |\phi\rangle_3 &= |0\rangle, \end{aligned} \tag{15}$$

and their orthogonal complements, $|\phi^{\perp}\rangle$. Parameters similar to these and serving the same function can be defined with respect to any three arbitrary states, $|\psi_i\rangle$, as long as the matrices $|\psi_i\rangle\langle\psi_i|$ along with the identity are linearly independent. Operators analogous to the $\hat{\sigma}$ operators can be defined relative to these states:

$$\hat{\tau}_i \equiv |\psi_i\rangle\langle\psi_i| - \big|\psi_i^{\perp}\big\rangle\big\langle\psi_i^{\perp}\big|. \tag{16}$$

We can further define an 'S-like' parameter T, given by:

$$T_i \equiv \mathrm{Tr}\{\hat{\tau}_i\hat{\rho}\}. \tag{17}$$

Continuing the previous convention and to complete the set, we define $\hat{\tau}_0 \equiv \hat{\sigma}_0$, which then requires that $T_0 = 1$. Note that the S_i parameters are simply a special case of the T_i, for the case when $\hat{\tau}_i = \hat{\sigma}_i$.

Unlike the specific case of the S parameters which describe *mutually unbiased*[4] (MUB) measurement bases, for biased measurements

$$\hat{\rho} \neq \frac{1}{2}\sum_{i=0}^{3} T_i\hat{\tau}_i. \tag{18}$$

In order to reconstruct the density matrix, the T parameters must first be transformed into the S parameters using Eq. (21).

[4] Two measurement bases, $\{\langle\psi_i|\}$ and $\{\langle\psi_j|\}$, are mutually unbiased if $\forall_{i,j}|\langle\psi_i|\psi_j\rangle|^2 = 1/d$, where d is the dimension of the system (for a system of n qubits, $d = 2^n$). A set of measurement bases are mutually unbiased if each basis in the set is mutually unbiased with respect to every other basis in the set. In single-qubit Poincaré space, the axes indicating mutually unbiased measurement bases are at right angles (Lawrence et al., 2002).

EXAMPLE 3 (*The Stokes parameters*). For photon polarization, the S_i are the famous Stokes parameters (though normalized), and correspond to measurements in the D/A, R/L, and H/V bases (Stokes, 1852). In terms of the $\hat{\tau}$ matrices just introduced, we would define a set of basis states $|\psi_1\rangle \equiv |D\rangle$, $|\psi_2\rangle \equiv |R\rangle$, and $|\psi_3\rangle \equiv |H\rangle$. For these analysis bases, $\hat{\tau}_1 = \hat{\sigma}_1$, $\hat{\tau}_2 = \hat{\sigma}_2$, and $\hat{\tau}_3 = \hat{\sigma}_3$ (and therefore $T_i = S_i$ for this specific choice of analysis bases).

As the simplest example, consider the input state $|H\rangle$. Applying Eq. (11), we find that

$$\begin{aligned} S_0 &= \mathrm{Tr}\{\hat{\sigma}_0\hat{\rho}_H\} = 1, \\ S_1 &= \mathrm{Tr}\{\hat{\sigma}_1\hat{\rho}_H\} = 0, \\ S_2 &= \mathrm{Tr}\{\hat{\sigma}_2\hat{\rho}_H\} = 0, \\ S_3 &= \mathrm{Tr}\{\hat{\sigma}_3\hat{\rho}_H\} = 1, \end{aligned} \tag{19}$$

which from Eq. (9) implies that

$$\hat{\rho}_H = \frac{1}{2}(\hat{\sigma}_0 + \hat{\sigma}_3) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{20}$$

When the Stokes parameters (S_i) are used as coordinates in 3-space, all physically possible states fall within a sphere of radius one (the Poincaré sphere for polarization, the Bloch sphere for electron spin or other two-level systems; see Born and Wolf (1987)). The pure states are found on the surface, states of linear polarization on the equator, circular states at the poles, mixed states within, and the totally mixed state—corresponding to completely unpolarized photons—at the center of the sphere. This provides a very convenient way to visualize one-qubit states (see Fig. 1). The θ and ϕ values from Eq. (2) allow any pure state to be easily mapped onto the sphere surface. These values are the polar coordinates of the pure state they represent on the Poincaré sphere.[5] In addition to mapping states, the sphere can be used to represent any unitary operation as a rotation about an arbitrary axis. For example, waveplates implement rotations about an axis that passes through the equator.

Any state $|\psi_0\rangle$ and its orthogonal partner, $|\psi_0^{\perp}\rangle$, are found on opposite points of the Poincaré sphere. The line connecting these two points forms an axis of the sphere, useful for visualizing the outcome of a measurement in the $|\psi_0\rangle/|\psi_0^{\perp}\rangle$ basis. The projection of any state $\hat{\rho}$ (through a line perpendicular to the $|\psi_0\rangle/|\psi_0^{\perp}\rangle$

[5] These polar coordinates are by convention rotated by 90°, so that $\theta = 0$ is on the equator corresponding to the state $|H\rangle$ and $\theta = 90°$, $\phi = 90°$ is at the North Pole corresponding to the state $|R\rangle$. This 90° rotation is particular to the Poincaré representation of photon polarization (Peters et al., 2003); representations of two-level systems on the Bloch sphere do not introduce it.

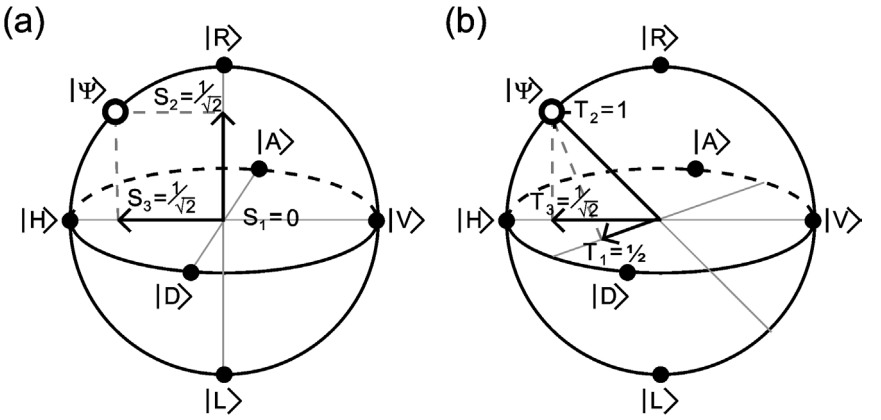

FIG. 1. The Poincaré (or Bloch) sphere. Any single-qubit quantum state $\hat{\rho}$ can be represented by three parameters $T_i = \mathrm{Tr}\{\hat{\tau}_i \hat{\rho}\}$, as long as the operators $\hat{\tau}_i$ in addition to the identity are linearly independent. Physically, the T_i parameters directly correspond to the outcome of a specific projective measurement: $T_i = 2P_i - 1$, where P_i is the probability of success for the measurement. The T_i may be used as coordinates in 3-space. Then all 1-qubit quantum states fall on or within a sphere of radius one. The surface of the sphere corresponds to pure states, the interior to mixed states, and the origin to the totally mixed state. Shown is a particular pure state $|\psi\rangle$, which is completely specified by its projection onto a set of nonparallel basis vectors. (a) When $\hat{\tau}_i = \hat{\sigma}_i$ (the Pauli matrices), the basis vectors are orthogonal, and in this particular case the T_i are equal to the S_i, the well known Stokes parameters, corresponding to measurements of diagonal (S_1), right-circular (S_2), and horizontal (S_3) polarizations. (b) A nonorthogonal coordinate system in Poincaré space. It is possible to represent a state using its projection onto nonorthogonal axes in Poincaré space. This is of particular use when attempting to reconstruct a quantum state from mutually biased measurements. Shown here are the axes corresponding to measurements of 22.5° linear (T_1), elliptical light rotated 22.5° from H towards R(T_2), and horizontal (T_3). Taken from Altepeter et al. (2004).

axis), will lie a distance along this axis corresponding to the relevant Stokes-like parameter ($T = \langle\psi_0|\hat{\rho}|\psi_0\rangle - \langle\psi_0^{\perp}|\hat{\rho}|\psi_0^{\perp}\rangle$).

Thus, just as any point in three-dimensional space can be specified by its projection onto three linearly independent axes, any quantum state can be specified by the three parameters $T_i = \mathrm{Tr}\{\hat{\tau}_i \hat{\rho}\}$, where $\hat{\tau}_{i=1,2,3}$ are linearly independent matrices equal to $|\psi_i\rangle\langle\psi_i| - |\psi_i^{\perp}\rangle\langle\psi_i^{\perp}|$. The $\hat{\tau}_i$ correspond to general Stokes-like parameters for any three linearly independent axes on the Poincaré sphere. However, they can differ from the canonical Stokes axes and need not even be orthogonal. See Fig. 1(b) for an example of state representation using nonorthogonal axes.

In order to use these mutually biased Stokes-like parameters, it is necessary to be able to transform a state from the mutually biased representation to the Stokes representation and vice-versa. In general, for any two representations $S_i =$

$\mathrm{Tr}\{\hat{\sigma}_i\hat{\rho}\}$ and $T_i = \mathrm{Tr}\{\hat{\tau}_i\hat{\rho}\}$ it is possible to transform between them by using

$$\begin{pmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_3\} \end{pmatrix}\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}. \quad (21)$$

This relation allows S parameters to be transformed into any set of T parameters. In order to transform from T to S, we can invert the 4 by 4 matrix in Eq. (21) and multiply both sides by this new matrix. This inversion is possible because we have chosen the $\hat{\tau}_i$ operators to be linearly independent; otherwise the T_i parameters would not specify a single point in Hilbert space.

1.2. Representation of Multiple Qubits

With the extension of these ideas to cover multiple qubits, it becomes possible to investigate nonclassical features, including the quintessentially quantum mechanical phenomenon of entanglement.

1.2.1. Pure States, Mixed States, and Diagonal Representations

As the name implies, multiple-qubit states are constructed out of individual qubits. As such, the Hilbert space of a many qubit system is spanned by state vectors which are the tensor product of single-qubit state vectors. A general n-qubit system can be written as

$$|\psi\rangle = \sum_{i_1,i_2,\ldots,i_n=0,1} \alpha_{i_1,i_2,\ldots,i_n}|i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_n\rangle. \quad (22)$$

Here the α_i are complex, $\sum_i |\alpha_i|^2 = 1$, and $\otimes$ denotes a tensor product, used to join component Hilbert spaces. For example, a general two-qubit pure state can be written as

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle, \quad (23)$$

where $|00\rangle$ is shorthand for $|0\rangle_1 \otimes |0\rangle_2$.

As before, we represent a general mixed state through an incoherent sum of pure states:

$$\hat{\rho} = \sum_i P_i|\psi_i\rangle\langle\psi_i|. \quad (24)$$

And, as before, this 2^n-by-2^n density matrix representing the n-qubit state may always be diagonalized, allowing any state to be written as

$$\hat{\rho} = \sum_{i=1}^{2^n} P_i|\phi_i\rangle\langle\phi_i|. \quad (25)$$

(24) differs from (25) in that the ϕ_i are necessarily orthogonal ($\langle\phi_i|\phi_j\rangle = \delta_{ij}$), and there are at most 2^n of them; in (24) there could be an arbitrary number of $|\psi_i\rangle$.

EXAMPLE 4 (*A general two-qubit polarization state*). Any two-qubit polarization state can be written as

$$\hat{\rho} = \begin{array}{c} \\ |HH\rangle \\ |HV\rangle \\ |VH\rangle \\ |VV\rangle \end{array} \begin{array}{c} \begin{array}{cccc} \langle HH| & \langle HV| & \langle VH| & \langle VV| \end{array} \\ \begin{pmatrix} A_1 & B_1 e^{i\phi_1} & B_2 e^{i\phi_2} & B_3 e^{i\phi_3} \\ B_1 e^{-i\phi_1} & A_2 & B_4 e^{i\phi_4} & B_5 e^{i\phi_5} \\ B_2 e^{-i\phi_2} & B_4 e^{-i\phi_4} & A_3 & B_6 e^{i\phi_6} \\ B_3 e^{-i\phi_3} & B_5 e^{-i\phi_5} & B_6 e^{-i\phi_6} & A_4 \end{pmatrix} \end{array}, \tag{26}$$

where $\hat{\rho}$ is positive and Hermitian with unit trace. Henceforth, the 'bra' and 'ket' labels will be omitted from density matrices presented in this standard basis.

EXAMPLE 5 (*The Bell states*). Perhaps the most famous examples of pure two-qubit states are the Bell states (Bell, 1964):

$$\begin{aligned} |\phi^{\pm}\rangle &= \frac{1}{\sqrt{2}}\big(|HH\rangle \pm |VV\rangle\big), \\ |\psi^{\pm}\rangle &= \frac{1}{\sqrt{2}}\big(|HV\rangle \pm |VH\rangle\big). \end{aligned} \tag{27}$$

Mixed states of note include the Werner states (Werner, 1989),

$$\hat{\rho}_W = P|\gamma\rangle\langle\gamma| + (1-P)\frac{1}{4}I, \tag{28}$$

where $|\gamma\rangle$ is a maximally entangled state and $\frac{1}{4}I$ is the totally mixed state, and the maximally entangled mixed states (MEMS), which possess the maximum amount of entanglement for a given amount of mixture (Munro et al., 2001).

Measures of entanglement and mixture may be derived from the density matrix; for reference, we now describe several such measures used to characterize a quantum state.

Fidelity. Fidelity is a measure of state overlap:

$$F(\rho_1, \rho_2) = \left(\mathrm{Tr}\left\{\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right\}\right)^2, \tag{29}$$

which—for ρ_1 and ρ_2 pure—simplifies to $\mathrm{Tr}\{\rho_1\rho_2\} = |\langle\psi_1|\psi_2\rangle|^2$ (Jozsa, 1994).[6]

[6] Note that some groups use an alternative definition of fidelity, equal to the square root of the formula presented here.

Tangle. The concurrence and tangle are measures of the nonclassical properties of a quantum state (Wooters, 1998; Coffman et al., 2000). For two qubits,[7] concurrence is defined as follows: consider the non-Hermitian matrix $\widehat{R} = \hat{\rho}\widehat{\Sigma}\hat{\rho}^{\mathrm{T}}\widehat{\Sigma}$ where the superscript T denotes transpose and the 'spin flip matrix' $\widehat{\Sigma}$ is defined by

$$\hat{\Sigma} \equiv \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{30}$$

If the eigenvalues of $\widehat{R}$, arranged in decreasing order, are given by $r_1 \geq r_2 \geq r_3 \geq r_4$, then the concurrence is defined by

$$C = \mathrm{Max}\{0, \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3} - \sqrt{r_4}\}. \tag{31}$$

The tangle is calculated directly from the concurrence:

$$T \equiv C^2. \tag{32}$$

The tangle (and the concurrence) range from 0 for product states (or, more generally, any incoherent mixture of product states) to a maximum value of 1 for Bell states.

Entropy and the linear entropy. The Von Neuman entropy quantifies the degree of mixture in a quantum state, and is given by

$$S \equiv -\mathrm{Tr}\{\hat{\rho}\ln[\hat{\rho}]\} = -\sum_i p_i \ln\{p_i\}, \tag{33}$$

where the p_i are the eigenvalues of ρ. The linear entropy (White et al., 1999) is a more analytically convenient description of state mixture. The linear entropy for a two-qubit system is defined by

$$S_L = \frac{4}{3}\left(1 - \mathrm{Tr}\{\hat{\rho}^2\}\right) = \frac{4}{3}\left(1 - \sum_{a=1}^{4} p_a^2\right), \tag{34}$$

where p_a are the eigenvalues of ρ. Note that for pure states, $\hat{\rho}^2 = \hat{\rho}$, and $\mathrm{Tr}[\hat{\rho}]$ is always 1, so that S_L ranges from 0 for pure states to 1 for the completely mixed state.

[7] The analysis in this subsection applies to the two-qubit case only. Measures of entanglement for mixed n-qubit systems are a subject of on-going research: see, for example, (Terhal, 2001) for a recent survey. In some restricted cases it may be possible to measure entanglement directly, without quantum state tomography; this possibility was investigated in Sancho and Huelga (2000). Also, one can detect the presence of nonzero entanglement, without quantifying it, using so-called "entanglement witnesses" (Lewenstein et al., 2000). Elsewhere we describe the trade-offs associated with these other entanglement characterization schemes (Altepeter et al., 2005).

1.2.2. Multiple Qubit Stokes Parameters

Extending the single-qubit density matrix representation of Eq. (9), any n-qubit state $\hat{\rho}$ may be represented as

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1,i_2,\ldots,i_n=0}^{3} S_{i_1,i_2,\ldots,i_n} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \cdots \otimes \hat{\sigma}_{i_n}. \tag{35}$$

Normalization requires that $S_{0,0,\ldots,0} = 1$, leaving $4^n - 1$ real parameters (the multiple-qubit analog of the single-qubit Stokes parameters) to identify any point in Hilbert space, just as three parameters determined the exact position of a one-qubit state in the Bloch/Poincaré sphere. Already for two qubits, the state space is much larger, requiring 15 independent real parameters to describe it. For this reason, there is no convenient graphical picture of this space, as there was in the single-qubit case (see, however, the interesting approaches made by Zyczkowski (2000, 2001)).

For multiple qubits the link between the multiple-qubit Stokes parameters (James et al., 2001; Abouraddy et al., 2002) and measurement probabilities still exists. The formalism of $\hat{\tau}$ operators also still holds for larger qubit systems, so that

$$T = \mathrm{Tr}\{\hat{\tau}\hat{\rho}\}. \tag{36}$$

For 'local' measurements (a local measurement is the tensor product of a number of single-qubit measurements: the first projecting qubit one along $\hat{\tau}_{i_1}$, the second qubit two along $\hat{\tau}_{i_2}$, etc.), $\hat{\tau} = \hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \cdots \otimes \hat{\tau}_{i_n}$. Combining Eqs. (35) and (36),

$$\begin{aligned} T_{i_1,i_2,\ldots,i_n} &= \mathrm{Tr}\{(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \cdots \otimes \hat{\tau}_{i_n})\hat{\rho}\} \\ &= \frac{1}{2^n} \sum_{j_1,j_2,\ldots,j_n=0}^{3} \mathrm{Tr}\{\hat{\tau}_{i_1}\hat{\sigma}_{j_1}\}\mathrm{Tr}\{\hat{\tau}_{i_2}\hat{\sigma}_{j_2}\} \cdots \mathrm{Tr}\{\hat{\tau}_{i_n}\hat{\sigma}_{j_n}\} S_{j_1,j_2,\ldots,j_n}. \end{aligned} \tag{37}$$

Recall that for single qubits,

$$\begin{aligned} &T_{i=1,2,3} = P_{|\psi_i\rangle} - P_{|\psi_i^{\perp}\rangle}, \\ &T_0 = P_{|\psi\rangle} + P_{|\psi^{\perp}\rangle} = 1, \quad \forall\psi. \end{aligned} \tag{38}$$

Therefore, for an n-qubit system,

$$\begin{aligned} T_{i_1,i_2,\ldots,i_n} = &\left(P_{|\psi_{i_1}\rangle} \pm P_{|\psi_{i_1}^{\perp}\rangle}\right) \otimes \left(P_{|\psi_{i_2}\rangle} \pm P_{|\psi_{i_2}^{\perp}\rangle}\right) \otimes \cdots \\ &\otimes \left(P_{|\psi_{i_n}\rangle} \pm P_{|\psi_{i_n}^{\perp}\rangle}\right), \end{aligned} \tag{39}$$

where the plus sign is used for a zero index and the minus sign is used for a nonzero index. For a two-qubit system where $i_1 \neq 0$ and $i_2 \neq 0$, T_{i_1,i_2} simplifies

dramatically, giving

$$\begin{aligned} T_{i_1,i_2} &= (P_{|\psi_{i_1}\rangle} - P_{|\psi_{i_1}^{\perp}\rangle}) \otimes (P_{|\psi_{i_2}\rangle} - P_{|\psi_{i_2}^{\perp}\rangle}) \\ &= P_{|\psi_{i_1}\rangle|\psi_{i_2}\rangle} - P_{|\psi_{i_1}\rangle|\psi_{i_2}^{\perp}\rangle} - P_{|\psi_{i_1}^{\perp}\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}^{\perp}\rangle|\psi_{i_2}^{\perp}\rangle}. \end{aligned} \tag{40}$$

This relation will be crucial for rebuilding a two-qubit state from local measurements.

As before, we are not restricted to multiple-qubit Stokes parameters based only on mutually unbiased operators. Extending Eq. (21) to multiple qubits, and again assuming two representations $S_{i_1,i_2,\ldots,i_n} = \mathrm{Tr}\{(\hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \cdots \otimes \hat{\sigma}_{i_n})\hat{\rho}\}$, and $T_{i_1,i_2,\ldots,i_n} = \mathrm{Tr}\{(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \cdots \otimes \hat{\tau}_{i_n})\hat{\rho}\}$,

$$\begin{aligned} T_{i_1,i_2,\ldots,i_n} = \frac{1}{2^n} \sum_{j_1,j_2,\ldots,j_n=0}^{3} \mathrm{Tr}\big\{&\big(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \cdots \otimes \hat{\tau}_{i_n}\big) \\ &\times \big(\hat{\sigma}_{j_1} \otimes \hat{\sigma}_{j_2} \otimes \cdots \otimes \hat{\sigma}_{j_n}\big)\big\} S_{j_1,j_2,\ldots,j_n}. \end{aligned} \tag{41}$$

In general, a given $\hat{\tau}$ operator is not uniquely mapped to a single pair of analysis states. For example, consider measurements of $|H\rangle$ and $|V\rangle$ corresponding to $\hat{\tau}_H = |H\rangle\langle H| - |V\rangle\langle V| = \hat{\sigma}_3$ and $\hat{\tau}_V = |V\rangle\langle V| - |H\rangle\langle H| = -\hat{\sigma}_3$. Therefore, $\hat{\tau}_{H,H} \equiv \hat{\sigma}_3 \otimes \hat{\sigma}_3 = -\hat{\sigma}_3 \otimes -\hat{\sigma}_3 \equiv \hat{\tau}_{V,V}$. This artifact of the mathematics does not in practice affect the results of a tomography.

EXAMPLE 6 (*A separable two-qubit polarization state*). Consider the state $|HH\rangle$. Following the example in Eqs. (20),

$$\begin{aligned} \hat{\rho}_{HH} &= |HH\rangle\langle HH| \\ &= \frac{1}{2}\big(\hat{\sigma}_0 + \hat{\sigma}_3\big) \otimes \frac{1}{2}\big(\hat{\sigma}_0 + \hat{\sigma}_3\big) \\ &= \frac{1}{4}\big(\hat{\sigma}_0 \otimes \hat{\sigma}_0 + \hat{\sigma}_0 \otimes \hat{\sigma}_3 + \hat{\sigma}_3 \otimes \hat{\sigma}_0 + \hat{\sigma}_3 \otimes \hat{\sigma}_3\big). \end{aligned} \tag{42}$$

This implies that for this state there are exactly four nonzero two-qubit Stokes parameters: $S_{0,0}$, $S_{0,3}$, $S_{3,0}$, and $S_{3,3}$—all of which are equal to one. (As earlier, for the special case when $\hat{\tau}_{i,j} = \hat{\sigma}_{i,j}$, we relabel the $T_{i,j}$ as $S_{i,j}$, the two-qubit Stokes parameters (James et al., 2001; Abouraddy et al., 2002).) The separable nature of this state makes it easy to calculate the two-qubit Stokes decomposition.

EXAMPLE 7 (*The singlet state*). If instead we investigate the entangled state $|\psi^-\rangle \equiv (|HV\rangle - |VH\rangle)/\sqrt{2}$, it will be necessary to calculate each two-qubit Stokes parameter from the $\hat{\sigma}$ matrices. As an example, consider $\hat{\sigma}_{3,3} \equiv \hat{\sigma}_3 \otimes \hat{\sigma}_3$,

for which

$$S_{3,3} = \mathrm{Tr}\{\hat{\sigma}_{3,3}|\psi^-\rangle\langle\psi^-|\} = -1. \tag{43}$$

We could instead calculate $S_{3,3}$ directly from probability outcomes of measurements on $|\psi^-\rangle$:

$$\begin{aligned} S_{3,3} &= (P_H - P_V) \otimes (P_H - P_V) \\ &= P_{HH} - P_{HV} - P_{VH} + P_{VV} \\ &= 0 - \frac{1}{2} - \frac{1}{2} + 0 = -1. \end{aligned} \tag{44}$$

Continuing on, we measure $S_{0,3}$:

$$\begin{aligned} S_{0,3} &= (P_H + P_V) \otimes (P_H - P_V) \\ &= P_{HH} - P_{HV} + P_{VH} - P_{VV} \\ &= 0 - \frac{1}{2} + \frac{1}{2} - 0 = 0. \end{aligned} \tag{45}$$

Here the signs of the probabilities changed due to the zero index in $S_{0,3}$. These results would have been the same even if the analysis bases of the first qubit had been shifted to any other orthogonal basis, i.e., $S_{0,3} = (P_\psi + P_{\psi^\perp}) \otimes (P_H - P_V)$.

If the method above is continued for all the Stokes parameters, one concludes that

$$\begin{aligned} \hat{\rho}_{\psi^-} &= \frac{1}{2}\big(|HV\rangle - |VH\rangle\big)\big(\langle HV| - \langle VH|\big) \\ &= \frac{1}{4}\big(\hat{\sigma}_0 \otimes \hat{\sigma}_0 - \hat{\sigma}_1 \otimes \hat{\sigma}_1 - \hat{\sigma}_2 \otimes \hat{\sigma}_2 - \hat{\sigma}_3 \otimes \hat{\sigma}_3\big). \end{aligned} \tag{46}$$

1.3. Representation of Nonqubit Systems

Although most interest within the field of quantum information and computation has focused on two-level systems (qubits) due to their simplicity, availability, and similarity to classical bits, nature contains a multitude of many-level systems, both discrete and continuous. A discussion of continuous systems is beyond the scope of this work—see Leonhardt (1997), but we will briefly address here the representation and tomography of discrete, d-level systems ("qudits"). For a more detailed description of qudit tomography, see Thew et al. (2002).

1.3.1. Pure, Mixed, and Diagonal Representations

Directly extending Eqs. (1) and (2), a d-level qudit can be represented as

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_{d-1}|d-1\rangle, \tag{47}$$

where $\sum_i |\alpha_i|^2 = 1$. Mixed qudit states can likewise be represented by generalizing Eqs. (4) and (6):

$$\rho = \sum_k P_k |\phi_k\rangle\langle\phi_k| \tag{48}$$

$$= \sum_{i=0}^{d-1} P_i |\psi_i\rangle\langle\psi_i|. \tag{49}$$

Here $\{|\phi_k\rangle\}$ is completely unrestricted while $|\langle\psi_i|\psi_j\rangle| = \delta_{ij}$. In other words, while any mixed state is an incoherent superposition of an undetermined number of pure states, any mixed state can be represented by an incoherent superposition of only n orthogonal states (the diagonal representation).

EXAMPLE 8 (*Orbital angular momentum modes*). The Laguerre–Gaussian modes of an optical field propagating in the z direction possess z components of orbital angular momentum that serve as the quantum numbers of a multiple-level photonic system that has recently been studied for quantum information (Mair et al., 2001; Langford et al., 2004; Arnaut and Barbosa, 2000). Consider a qudit system with an infinite number of levels representing the quantization of orbital angular momentum. A superposition of the three lowest angular momentum levels would look like

$$|\psi\rangle = |+1\rangle + |0\rangle + |-1\rangle, \tag{50}$$

where $|+1\rangle$ ($|-1\rangle$) corresponds to a mode where each photon has $+\hbar$ ($-\hbar$) z-component of orbital angular momentum, and $|0\rangle$ corresponds to a zero angular momentum mode, e.g., a mode having a Gaussian transverse profile. Using specially designed holograms, these states can be measured and interconverted (Allen et al., 2004).

1.3.2. Qudit Stokes Parameters

In order to completely generalize the qubit mathematics laid out previously to the qudit case, it is necessary to find Stokes-like parameters which satisfy the following conditions:

$$\hat{\rho} = \sum_{i=0}^{n} S_i \hat{\sigma}_i, \tag{51}$$

$$S_i \equiv \mathrm{Tr}\{\hat{\sigma}_i \hat{\rho}\}. \tag{52}$$

In addition, in order to easily generalize the tomographic techniques of the next section, it will be necessary to find S_i as a function of measurable probabilities:

$$S_i = \mathcal{F}(\{P_{|\psi\rangle}\}). \tag{53}$$

Obviously, it would be ideal to find a simple form similar to the qubit $\hat{\sigma}$ matrices. Conveniently, the general qudit sigma matrices and corresponding S_i parameters can be divided into three groups ($\{\hat{\sigma}_i^X, \hat{\sigma}_i^Y, \hat{\sigma}_i^Z\}$ and $\{S_i^X, S_i^Y, S_i^Z\}$), according to their similarity to $\hat{\sigma}_x = \hat{\sigma}_1$, $\hat{\sigma}_y = \hat{\sigma}_2$, and $\hat{\sigma}_z = \hat{\sigma}_3$, respectively (Thew et al., 2002). Using these divisions, we can expand Eq. (51):

$$\hat{\rho} = S_0\hat{\sigma}_0 + \sum_{\substack{j,k\in\{0,1,\ldots,n-1\}\\ j\neq k}} \left(S_{j,k}^X\hat{\sigma}_{j,k}^X + S_{j,k}^Y\hat{\sigma}_{j,k}^Y\right) + \sum_{r=1}^{n-1} S_r^Z\hat{\sigma}_r^Z. \tag{54}$$

Investigating the simplest group first, it is unsurprising that

$$\hat{\sigma}_0 = I, \qquad S_0 = 1, \tag{55}$$

continuing the previous qubit convention. The X and Y related variables are defined almost identically to their predecessors:

$$\hat{\sigma}_{j,k}^X = |j\rangle\langle k| + |k\rangle\langle j|, \tag{56}$$

$$S_{j,k}^X = P_{1/\sqrt{2}(|j\rangle+|k\rangle)} - P_{1/\sqrt{2}(|j\rangle-|k\rangle)}, \tag{57}$$

$$\hat{\sigma}_{j,k}^Y = -\mathrm{i}\left(|j\rangle\langle k| - |k\rangle\langle j|\right), \tag{58}$$

$$S_{j,k}^Y = P_{1/\sqrt{2}(|j\rangle+\mathrm{i}|k\rangle)} - P_{1/\sqrt{2}(|j\rangle-\mathrm{i}|k\rangle)}. \tag{59}$$

The definitions for $\hat{\sigma}_i^Z$ and S_i^Z are slightly more complicated:

$$\hat{\sigma}_r^Z = \sqrt{\frac{2}{r(r+1)}}\left[\left(\sum_{j=0}^{r-1}|j\rangle\langle j|\right) - r|r\rangle\langle r|\right], \tag{60}$$

$$S_r^Z = \sqrt{\frac{2}{r(r+1)}}\left[\left(\sum_{j=0}^{r-1}P_{|j\rangle}\right) - rP_{|r\rangle}\right]. \tag{61}$$

These definitions complete the set of n^2 sigma matrices, and have a slightly more complex form in order to satisfy $\mathrm{Tr}[\hat{\sigma}_i] = 0$ and $\mathrm{Tr}[\hat{\sigma}_i\hat{\sigma}_j] = 2\delta_{ij}$ (these definitions apply to all $\hat{\sigma}_i$ except $\hat{\sigma}_0$).

EXAMPLE 9 (*The qutrit*). For a 3-level system ($|0\rangle$, $|1\rangle$, and $|2\rangle$), the $\hat{\sigma}$ matrices can be defined as:

$$\hat{\sigma}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \hat{\sigma}_1^Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\hat{\sigma}_2^Z = \sqrt{\frac{1}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \qquad \hat{\sigma}_{1,2}^X = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\hat{\sigma}_{1,3}^X = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \hat{\sigma}_{2,3}^X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\hat{\sigma}_{1,2}^Y = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \hat{\sigma}_{1,3}^Y = \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix},$$

$$\hat{\sigma}_{2,3}^Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}.$$

Expanding the S_i parameters in terms of probabilities, we find that:

$$\begin{aligned}
S_0 &= 1, \\
S_{1,2}^X &= P_{1/\sqrt{2}(|0\rangle+|1\rangle)} - P_{1/\sqrt{2}(|0\rangle-|1\rangle)}, \\
S_{1,3}^X &= P_{1/\sqrt{2}(|0\rangle+|2\rangle)} - P_{1/\sqrt{2}(|0\rangle-|2\rangle)}, \\
S_{2,3}^X &= P_{1/\sqrt{2}(|1\rangle+|2\rangle)} - P_{1/\sqrt{2}(|1\rangle-|2\rangle)}, \\
S_{1,2}^Y &= P_{1/\sqrt{2}(|0\rangle+i|1\rangle)} - P_{1/\sqrt{2}(|0\rangle-i|1\rangle)}, \\
S_{1,3}^Y &= P_{1/\sqrt{2}(|0\rangle+i|2\rangle)} - P_{1/\sqrt{2}(|0\rangle-i|2\rangle)}, \\
S_{2,3}^Y &= P_{1/\sqrt{2}(|1\rangle+i|2\rangle)} - P_{1/\sqrt{2}(|1\rangle-i|2\rangle)}, \\
S_1^Z &= P_{|0\rangle} - P_{|1\rangle}, \\
S_2^Z &= \frac{1}{\sqrt{3}}\big(P_{|0\rangle} + P_{|1\rangle} - 2P_{|2\rangle}\big).
\end{aligned}$$

2. Tomography of Ideal Systems

The goal of tomography is to reconstruct the density matrix of an ensemble of particles through a series of measurements. In practice, this can never be performed exactly, as an infinite number of particles would be required to eliminate statistical error. If exact measurements were taken on an infinite number of identically

prepared systems, each measurement would yield an exact probability of success, which could then be used to reconstruct a density matrix. Though unrealistic, it is highly illustrative to examine this exact tomography before considering the more general treatment. Hence, this section will treat all measurements as yielding exact probabilities, and ignore all sources of error in those measurements.

2.1. Single-Qubit Tomography

Although reconstructive tomography of any size system follows the same general procedure, beginning with tomography of a single qubit allows the visualization of each step using the Poincaré sphere, in addition to providing a simpler mathematical introduction.

2.1.1. Visualization of Single-Qubit Tomography

Exact single-qubit tomography requires a sequence of three linearly independent measurements. Each measurement exactly specifies one degree of freedom for the measured state, reducing the free parameters of the unknown state's possible Hilbert space by one.

As an example, consider measuring R, D, and H on the partially mixed state

$$\hat{\rho} = \begin{pmatrix} \frac{5}{8} & \frac{-\mathrm{i}}{2\sqrt{2}} \\ \frac{\mathrm{i}}{2\sqrt{2}} & \frac{3}{8} \end{pmatrix}. \tag{62}$$

Rewriting the state using Eq. (9) as

$$\hat{\rho} = \frac{1}{2}\left(\hat{\sigma}_0 + \frac{1}{\sqrt{2}}\hat{\sigma}_2 + \frac{1}{4}\hat{\sigma}_3\right) \tag{63}$$

allows us to read off the normalized Stokes parameters corresponding to these measurements:

$$S_1 = 0, \quad S_2 = \frac{1}{\sqrt{2}}, \quad \text{and} \quad S_3 = \frac{1}{4}. \tag{64}$$

As always, $S_0 = 1$ due to normalization. Measuring R (which determines S_2) first, and looking to the Poincaré sphere, we determine that the unknown state must lie in the $z = 1/\sqrt{2}$ plane (as $S_2 = 1/\sqrt{2}$). A measurement in the D basis (with the result $P_D = P_A = \frac{1}{2}$) further constrains the state to the $y = 0$ plane, resulting in a total confinement to a line parallel to and directly above the x axis. The final measurement of H pinpoints the state. This process is illustrated in Fig. 2(a). Obviously the order of the measurements is irrelevant: it is the intersection point of three orthogonal planes that defines the location of the state.

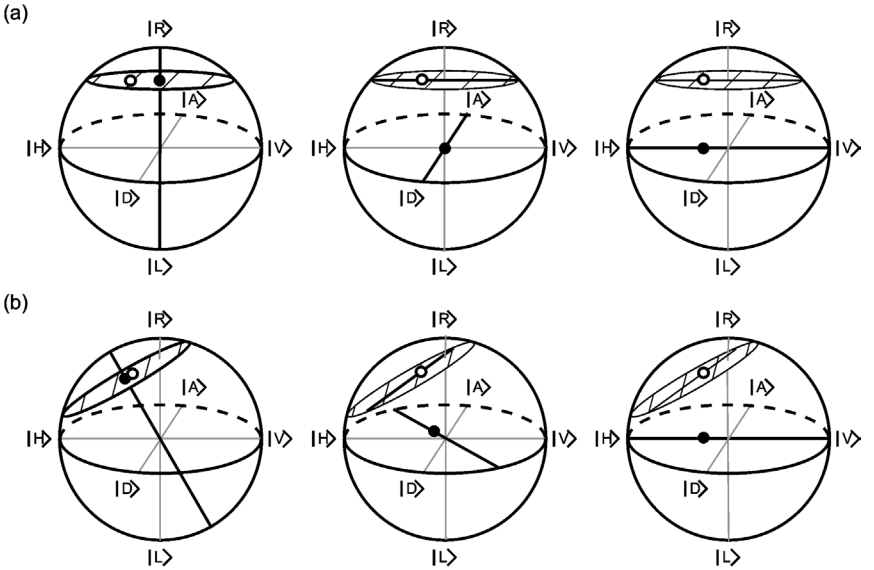

FIG. 2. A sequence of three linearly independent measurements isolates a single quantum state in Hilbert space (shown here as an open circle in the Poincaré sphere representation). The first measurement isolates the unknown state to a plane perpendicular to the measurement basis. Further measurements isolate the state to the intersections of nonparallel planes, which for the second and third measurements correspond to a line and finally a point. The black dots shown correspond to the projection of the unknown state onto the measurement axes, which determines the position of the aforementioned planes. (a) A sequence of measurements along the right-circular, diagonal, and horizontal axes. (b) A sequence of measurements on the same state taken using nonorthogonal projections: elliptical light rotated 30° from H towards R, 22.5° linear, and horizontal. Taken from Altepeter et al. (2004).

If instead measurements are made along nonorthogonal axes, a very similar picture develops, as indicated in Fig. 2(b). The first measurement always isolates the unknown state to a plane, the second to a line, and the third to a point.

Of course, in practice, the experimenter has no knowledge of the unknown state before a tomography. The set of the measured probabilities, transformed into the Stokes parameters as above, allows a state to be directly reconstructed.

2.1.2. A Mathematical Look at Single-Qubit Tomography

Using the tools developed in the first section of this chapter, single-qubit tomography is relatively straightforward. Recall Eq. (9), $\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i$. Considering that S_1, S_2, and S_3 completely determine the state, we need only measure them to complete the tomography. From Eq. (13), $S_{j>0} = 2P_{|\psi\rangle} - 1$, therefore three measurements in the $|0\rangle$, $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $\frac{1}{\sqrt{2}}(|0\rangle + \mathrm{i}|1\rangle)$ bases will completely

specify the unknown state. If instead measurements are made in another basis, even a nonorthogonal one, they can be easily related back to the S_i parameters, and therefore the density matrix, by means of Eq. (21).

While this procedure is straightforward, there is one subtlety which will become important in the multiple-qubit case. Projective measurements generally refer to the measurement of a single basis state and return a single value between zero and one. This corresponds, for example, to an electron beam passing through a Stern–Gerlach apparatus with a detector placed at one output. While a single detector and knowledge of the input particle intensity will—in the one-qubit case—completely determine a single Stokes parameter, one could collect data from both outputs of the Stern–Gerlach device. This would measure the probability of projecting not only onto the state $|\psi\rangle$, but also onto $|\psi^{\perp}\rangle$, without needing to know the input intensity. All physical measurements on single qubits, regardless of implementation, can in principle be measured this way (though in practice measurements of some qubit systems may typically detect a population in only *one* of the states, as in Kielpinski et al. (2001)). We will see below that although one detector functions as well as two in the single-qubit case, this situation will not persist into higher dimensions.

2.2. Multiple-Qubit Tomography

The same methods used to reconstruct an unknown single-qubit state can be applied to multiple-qubit systems. Just as each single-qubit Stokes vector can be expressed in terms of measurable probabilities—Eq. (12), each multiple-qubit Stokes vector can be measured in terms of the probabilities of projecting the multiple-qubit state into a sequence of separable bases—Eq. (39).

Using the most naive method, an n-qubit system, represented by 4^n Stokes parameters, would require $4^n \times 2^n$ probabilities to be reconstructed (2^n probabilities for each of 4^n Stokes parameters).

Of course, because an n-qubit density matrix contains $4^n - 1$ free parameters, the $4^n \times 2^n$ measured probabilities must be linearly dependent. As expected, by using the extra information that measurements of complete orthogonal bases must sum to one (e.g., $P_{HH} + P_{HV} + P_{VH} + P_{VV} = 1$, $P_{HH} + P_{HV} = P_{HD} + P_{HA}$), we find that only $4^n - 1$ probability measurements are necessary to reconstruct a density matrix.

While we can easily construct a minimum measurement set for an n-qubit system by measuring every combination of $\{H, V, D, R\}$ at each qubit, i.e.,

$$\{M\} = \{H, V, D, R\}_1 \otimes \{H, V, D, R\}_2 \otimes \cdots \{H, V, D, R\}_n, \tag{65}$$

this is almost never optimal (see Section 5). See Section 2.4 for a formal method for testing whether a specific set of measurements is sufficient for tomography.

EXAMPLE 10 (*An ideal 2-qubit tomography of photon pairs*). Consider measuring a state in nine complete four-element bases, for a total of 36 measurement results. These results are compiled below, with each row representing a single basis, and therefore a single two-qubit Stokes parameter.

$$
\begin{aligned}
S_{1,1} &= \underset{\frac{1}{3}}{+P_{DD}} \quad \underset{\frac{1}{6}}{-P_{DA}} \quad \underset{\frac{1}{6}}{-P_{AD}} \quad \underset{\frac{1}{3}}{+P_{AA}} = \frac{1}{3},\\
S_{1,2} &= \underset{\frac{1}{4}}{+P_{DR}} \quad \underset{\frac{1}{4}}{-P_{DL}} \quad \underset{\frac{1}{4}}{-P_{AR}} \quad \underset{\frac{1}{4}}{+P_{AL}} = 0,\\
S_{1,3} &= \underset{\frac{1}{4}}{+P_{DH}} \quad \underset{\frac{1}{4}}{-P_{DV}} \quad \underset{\frac{1}{4}}{-P_{AH}} \quad \underset{\frac{1}{4}}{+P_{AV}} = 0,\\
S_{2,1} &= \underset{\frac{1}{4}}{+P_{RD}} \quad \underset{\frac{1}{4}}{-P_{RA}} \quad \underset{\frac{1}{4}}{-P_{LD}} \quad \underset{\frac{1}{4}}{+P_{LA}} = 0,\\
S_{2,2} &= \underset{\frac{1}{6}}{+P_{RR}} \quad \underset{\frac{1}{3}}{-P_{RL}} \quad \underset{\frac{1}{3}}{-P_{LR}} \quad \underset{\frac{1}{6}}{+P_{LL}} = -\frac{1}{3},\\
S_{2,3} &= \underset{\frac{1}{4}}{+P_{RH}} \quad \underset{\frac{1}{4}}{-P_{RV}} \quad \underset{\frac{1}{4}}{-P_{LH}} \quad \underset{\frac{1}{4}}{+P_{LV}} = 0,\\
S_{3,1} &= \underset{\frac{1}{4}}{+P_{HD}} \quad \underset{\frac{1}{4}}{-P_{HA}} \quad \underset{\frac{1}{4}}{-P_{VD}} \quad \underset{\frac{1}{4}}{+P_{VA}} = 0,\\
S_{3,2} &= \underset{\frac{1}{4}}{+P_{HR}} \quad \underset{\frac{1}{4}}{-P_{HL}} \quad \underset{\frac{1}{4}}{-P_{VR}} \quad \underset{\frac{1}{4}}{+P_{VL}} = 0,\\
S_{3,3} &= \underset{\frac{1}{3}}{+P_{HH}} \quad \underset{\frac{1}{6}}{-P_{HV}} \quad \underset{\frac{1}{6}}{-P_{VH}} \quad \underset{\frac{1}{3}}{+P_{VV}} = \frac{1}{3}.
\end{aligned}
\tag{66}
$$

Measurements are taken in each of these nine bases, determining the above nine two-qubit Stokes parameters. The six remaining required parameters, listed below, are dependent upon the same measurements.

$$
\begin{aligned}
S_{0,1} &= \underset{\frac{1}{3}}{+P_{DD}} \quad \underset{\frac{1}{6}}{-P_{DA}} \quad \underset{\frac{1}{6}}{+P_{AD}} \quad \underset{\frac{1}{3}}{-P_{AA}} = 0,\\
S_{0,2} &= \underset{\frac{1}{6}}{+P_{RR}} \quad \underset{\frac{1}{3}}{-P_{LR}} \quad \underset{\frac{1}{3}}{+P_{RL}} \quad \underset{\frac{1}{6}}{-P_{LL}} = 0,\\
S_{0,3} &= \underset{\frac{1}{3}}{+P_{HH}} \quad \underset{\frac{1}{6}}{-P_{HV}} \quad \underset{\frac{1}{6}}{+P_{VH}} \quad \underset{\frac{1}{3}}{-P_{VV}} = 0,\\
S_{1,0} &= \underset{\frac{1}{3}}{+P_{DD}} \quad \underset{\frac{1}{6}}{+P_{DA}} \quad \underset{\frac{1}{6}}{-P_{AD}} \quad \underset{\frac{1}{3}}{-P_{AA}} = 0,\\
S_{2,0} &= \underset{\frac{1}{6}}{+P_{RR}} \quad \underset{\frac{1}{3}}{+P_{LR}} \quad \underset{\frac{1}{3}}{-P_{RL}} \quad \underset{\frac{1}{6}}{-P_{LL}} = 0,\\
S_{3,0} &= \underset{\frac{1}{3}}{+P_{HH}} \quad \underset{\frac{1}{6}}{+P_{HV}} \quad \underset{\frac{1}{6}}{-P_{VH}} \quad \underset{\frac{1}{3}}{-P_{VV}} = 0.
\end{aligned}
\tag{67}
$$

These terms will not in general be zero. Recall—cf. Eq. (42)—that for $|HH\rangle$, $S_{0,3} = S_{3,0} = 1$. Of course, $S_{0,0} = 1$. Taken together, these two-qubit Stokes parameters determine the density matrix:

$$
\hat{\rho} = \frac{1}{4}\left(\hat{\sigma}_0 \otimes \hat{\sigma}_0 + \frac{1}{3}\hat{\sigma}_1 \otimes \hat{\sigma}_1 - \frac{1}{3}\hat{\sigma}_2 \otimes \hat{\sigma}_2 + \frac{1}{3}\hat{\sigma}_3 \otimes \hat{\sigma}_3\right)
$$

$$= \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{68}$$

This is the final density matrix, a Werner state, as defined in Eq. (28).

2.3. Tomography of Nonqubit Systems

By making use of the qudit extensions to the Stokes parameter formalism—Eqs. (57)–(61), we can reconstruct any qudit system in exactly the same manner as qubit systems. For a single particle d-level system, a single Stokes parameter is dependent on $d-1$ independent probabilities, and $d+1$ Stokes parameters are necessary to reconstruct the density matrix.

Multiple-qudit systems can be reconstructed by using separable projectors (Thew et al., 2002) upon which the multiple qudit Stokes parameters are dependent (these dependencies were laid out in Section 1.3.2). Likewise, the following section on general tomography, while specific to qubits, can be easily adapted to qudit systems.

2.4. General Qubit Tomography

As discussed earlier, qubit tomography will require $4^n - 1$ probabilities in order to define a complete set of T_i parameters. In practice, this will mean that 4^n measurements are necessary in order to normalize counts to probabilities. By making projective measurements on each qubit and only taking into account those results where a definite result is obtained (e.g., the photon was transmitted by the polarizer), it is possible to reconstruct a state using the results of 4^n measurements.

Our first task is to represent the density matrix in a useful form. To this end, define a set of $2^n \times 2^n$ matrices which have the following properties:

$$\begin{aligned} &\operatorname{Tr}\{\widehat{\Gamma}_\nu \cdot \widehat{\Gamma}_\mu\} = \delta_{\nu,\mu}, \\ &\hat{A} = \sum_\nu \widehat{\Gamma}_\nu \operatorname{Tr}\{\widehat{\Gamma}_\nu \cdot \hat{A}\} \quad \forall \hat{A}, \end{aligned} \tag{69}$$

where $\hat{A}$ is an arbitrary $2^n \times 2^n$ matrix. A convenient set of $\widehat{\Gamma}$ matrices to use are tensor-products of the $\hat{\sigma}$ matrices used throughout this paper:

$$\widehat{\Gamma}_\nu = \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \cdots \otimes \hat{\sigma}_{i_n}, \tag{70}$$

where ν is simply a short-hand index by which to label the Γ matrices (there are 4^n of them) which is more concise than $i_1, i_2, \ldots, i_n$. Transforming Eq. (35) into

this notation, we find that

$$\hat{\rho} = \frac{1}{2^n} \sum_{\nu=1}^{4^n} \widehat{\Gamma}_\nu S_\nu. \tag{71}$$

Next, it is necessary to consider exactly which measurements to use. In particular, we now wish to determine the necessary and sufficient conditions on the 4^n measurements to allow reconstruction of any state.[8] Let $|\psi_\mu\rangle$ ($\mu = 1$ to 4^n) be the measurement bases, and define the probability of the μth measurement as $P_\mu \equiv \langle\psi_\mu|\hat{\rho}|\psi_\mu\rangle$.

Combining this with Eq. (71),

$$P_\mu = \langle\psi_\mu|\frac{1}{2^n} \sum_{\nu=1}^{4^n} \widehat{\Gamma}_\nu S_\nu|\psi_\mu\rangle = \frac{1}{2^n} \sum_{\nu=1}^{4^n} B_{\mu,\nu} S_\nu, \tag{72}$$

where the $4^n \times 4^n$ matrix $B_{\mu,\nu}$ is given by

$$B_{\mu,\nu} = \langle\psi_\mu|\widehat{\Gamma}_\nu|\psi_\mu\rangle. \tag{73}$$

Immediately we find a necessary and sufficient condition for the completeness of the set of tomographic states $\{|\psi_\mu\rangle\}$: if the matrix $B_{\mu,\nu}$ is nonsingular, then Eq. (72) can be inverted to give

$$S_\nu = 2^n \sum_{\mu=1}^{4^n} \left(B^{-1}\right)_{\mu,\nu} P_\mu. \tag{74}$$

While this provides an exact solution if exact probabilities are known, it leads to a number of difficulties in real systems. First, it is possible for statistical errors to cause a set of measurements to lead to an illegal density matrix. Second, if more than the minimum number of measurements are taken and they contain any error, they will overdefine the problem, eliminating the possibility of a single analytically calculated answer. To solve these problems it is necessary to analyze the data in a fundamentally different way, in which statistically varying probabilities are assumed from the beginning and optimization algorithms find the state most likely to have resulted in the measured data (Section 4.2).

[8] If exact probabilities are known, only $4^n - 1$ measurements are necessary. However, often only numbers of counts (successful measurements) are known, with no information about the number of counts which would have been measured by detectors in orthogonal bases. In this case an extra measurement is necessary to normalize the inferred probabilities.

3. Collecting Tomographic Measurements

Before discussing the analysis of real experimental data, it is necessary to understand how that experimental data is collected. This chapter outlines the experimental implementation of tomography on polarization entangled qubits generated from spontaneous parametric downconversion (Kwiat et al., 1999), though the techniques for projection and particularly systematic error correction will be applicable to many systems. We filter these photon pairs using both spatial filters (irises used to isolate a specific k-vector, necessary because our states are angle-dependent) and frequency filters (interference filters, typically 5–10 nm wide, FWHM).

After this initial filtering, measurement collection involves two central issues: projection (into an ideally arbitrary range of states) and systematic error correction (to compensate for any number of experimental problems ranging from imperfect optics to accidental coincidences).

3.1. Projection

Any tomography, in fact any measurement on a quantum system, depends on state projection; for the purposes of this chapter, these projections will be *separable*. While tomography could be simplified by using arbitrary projectors (e.g., joint measurements on two qubits), this is experimentally difficult. We therefore focus on the ability to create arbitrary single-qubit projectors which will then be easily chained together to create any separable projector.

3.1.1. Arbitrary Single-Qubit Projection

An arbitrary polarization measurement and its orthogonal complement can be realized using, in order, a quarter-wave plate, a half-waveplate, and a polarizing beam splitter. Waveplates implement unitary operations, and in the Poincaré sphere picture, act as rotations about an axis lying within the linear polarization plane (the equator) (Born and Wolf, 1987). Specifically, a waveplate whose optic axis is oriented at angle θ with respect to the horizontal induces a rotation on the Poincaré sphere about an axis 2θ from horizontal, in the linear plane. The magnitude of this rotation is equal to the waveplate's retardance (90° for quarter-wave plates and 180° for half-wave plates). For the remainder of this chapter we adopt the convention that polarizing beam splitters transmit horizontally polarized light and reflect vertically polarized light—though for some types the roles are reversed.

This analysis, while framed in terms of waveplates acting on photon polarization, is directly applicable to other systems, e.g., spin-$\frac{1}{2}$ particles (Cory et al.,

1997; Jones et al., 1997; Weinstein et al., 2001; Laflamme et al., 2002) or two-level atoms (Monroe, 2002; Schmidt-Kaler et al., 2003). In these systems, measurements in arbitrary bases are obtained using suitably phased π- and $\frac{\pi}{2}$-pulses (externally applied electromagnetic fields) to rotate the state to be measured into the desired analysis basis.

To derive the settings for these waveplates as a function of the projection state desired, we use the Poincaré sphere (see Fig. 3). For any state on the surface of the sphere, a 90° rotation about a linear axis directly below it will rotate that state into a linear polarization (see Fig. 3(b)). Assume the desired projection state is

$$|\psi_P\rangle = \cos\left(\frac{\theta}{2}\right)|H\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|V\rangle. \tag{75}$$

Simple coordinate transforms from spherical to Cartesian coordinates reveal that a quarter-waveplate at $\theta_{QWP} = \frac{1}{2}\cos^{-1}\{\sin(\theta)\tan(\phi)\}$ will rotate the projection state (75) into a linear state

$$|\psi'_P\rangle = \cos\left(\frac{\theta'}{2}\right)|H\rangle + \sin\left(\frac{\theta'}{2}\right)|V\rangle. \tag{76}$$

A half-waveplate at $\frac{1}{4}\theta'$ (with respect to horizontal orientation) will then rotate this state to $|H\rangle$.[9] Finally, the PBS will transmit the projected state and reflect its orthogonal complement.

Mathematically, this process of rotation and projection can be described using unitary transformations. The unitary transformations for half- and quarter-waveplates in the H/V basis are

$$\begin{aligned} U_{\mathrm{HWP}}(\theta) &= \begin{bmatrix} \cos^2(\theta) - \sin^2(\theta) & 2\cos(\theta)\sin(\theta) \\ 2\cos(\theta)\sin(\theta) & \sin^2(\theta) - \cos^2(\theta) \end{bmatrix}, \\ U_{\mathrm{QWP}}(\theta) &= \begin{bmatrix} \cos^2(\theta) + \mathrm{i}\sin^2(\theta) & (1-\mathrm{i})\cos(\theta)\sin(\theta) \\ (1-\mathrm{i})\cos(\theta)\sin(\theta) & \sin^2(\theta) + \mathrm{i}\cos^2(\theta) \end{bmatrix}, \end{aligned} \tag{77}$$

with θ denoting the rotation angle of the waveplate with respect to horizontal. Assume that during the course of a tomography, the νth measurement setting requires that the QWP be set to $\theta_{\mathrm{QWP},\nu}$ and the HWP to $\theta_{\mathrm{HWP},\nu}$. Therefore, the total unitary[10] for the νth measurement setting will be

$$U_\nu = U_{\mathrm{HWP}}(\theta_{\mathrm{HWP},\nu})U_{\mathrm{QWP}}(\theta_{\mathrm{QWP},\nu}). \tag{78}$$

[9] $\theta' = \cos^{-1}/\cos^{-1}\{\sin(\theta)\tan(\phi)\} - \cos^{-1}/\cos^{-1}\{\cot(\theta)\cot(\phi)\}$. In practice, care must be taken that consistent conventions are used (e.g., right- vs. left-circular polarization), and it may be easier to calculate this angle directly from waveplate operators and the initial state.

[10] Note the order of the unitary matrices for the HWP and QWP. Incoming light encounters the QWP first, and therefore U_{QWP} is last when defining U_ν.

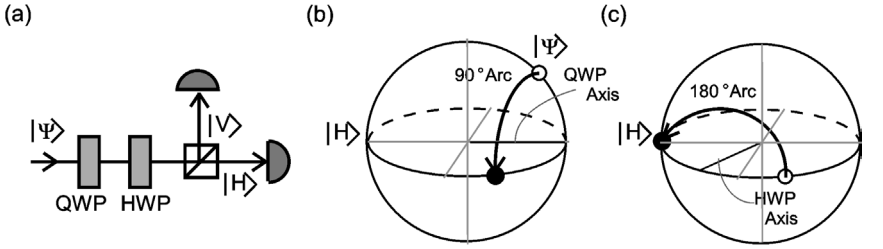

FIG. 3. A quarter-waveplate (QWP), half-waveplate (HWP), and polarizing beam splitter (PBS) are used to make an arbitrary polarization measurement. Both a diagram of the experimental apparatus (a) and the step-by-step evolution of the state on the Poincaré sphere are shown. (b) The quarter-waveplate rotates the projection state (the state we are projecting into, *not* the incoming unknown state) into the linear polarization plane (the equator). (c) The half-waveplate rotates this linear state to horizontal. The PBS transmits the projection state (now $|H\rangle$) and reflects its orthogonal complement (now $|V\rangle$), which can then both be measured.

For multiple qubits, we can directly combine these unitaries such that

$$U_v = {}^1U_v \otimes {}^2U_v \otimes \cdots \otimes {}^nU_v, \tag{79}$$

where ${}^qU_v$ denotes the qth qubit's unitary transformation due to waveplates. The total projection operator for this system is therefore $\langle 0|U_v$, where $|0\rangle$ is the first computational basis state (the state which passes through the beamsplitters—most likely $|H\rangle$ for each qubit). The measurement state (the state which will pass through the measurement apparatus and be measured every time) is therefore $U_v^\dagger|0\rangle$.

Of course, these calculations assume that we are using waveplates with retardances equal to exactly π or $\frac{\pi}{2}$ (or Rabi pulses producing perfect phase differences). Imperfect yet well characterized waveplates will lead to measurements in slightly different, yet known, bases. This can still yield an accurate tomography, but first these results must be transformed from a biased basis into the canonical Stokes parameters using Eq. (21). As discussed below (see Section 3), the maximum likelihood technique provides a different but equally effective way to accommodate for imperfect measurements.

3.1.2. Compensating for Imperfect Waveplates

While the previous section shows that it is possible using a quarter- and half-waveplate to project into an arbitrary single qubit state, perfect quarter- and half-waveplates are experimentally impossible to obtain. More likely, the experimenter will have access to waveplates with known retardances slightly different than the ideal values of π (HWP) and $\frac{\pi}{2}$ (QWP). Even in this case, it is often possible to obtain arbitrary single-qubit projections. (Note that this is the second solution to the problem of imperfect waveplates. Imperfect waveplates could be used at

virtually any angles during a tomography—such as the same angles at which perfect waveplates would be used to measure in the canonical basis—resulting in a set of biased bases. The tomography mathematics have already been shown to function for either mutually biased or unbiased bases, as long as the set of bases is complete. In contrast, this section describes how—*even using imperfect waveplates*—one can still measure in the canonical, mutually unbiased bases.)

Analytically finding the angles where this is possible proves to be inconvenient and, for some waveplates, impossible. Rather than solve a system of equations based on the unitary waveplate matrices, we will examine the effect of these waveplates graphically using the Poincaré sphere. For the remainder of this discussion, we will assume that the experimenter has access to two waveplates, WP_1 and WP_2, which will respectively take the place of the QWP and HWP normally present in the experimental setup. We constrain the retardances of these waveplates to be $0 \leq \phi_1 \leq \phi_2 \leq \pi$.

In order to project into an arbitrary state $|\psi\rangle$, WP_1 and WP_2 must together rotate the state $|\psi\rangle$ into the state $|H\rangle$ (assuming a horizontal polarizer is used after the waveplates—any linear polarizer is equivalent). Taking a piecewise approach, first consider which states are possible after acting on the input state $|\psi\rangle$ with WP_1. Figure 4(a) shows several example cases on the Poincaré sphere, each resulting in a curved band of possible states that can be reached by varying the orientation of WP_1. Next consider which states could be rotated by WP_2 into the target state $|H\rangle$. Figure 4(b) shows several examples of these states, which also take the form of a curved band, traversed by varying the orientation of WP_2. In order for state $|\psi\rangle$ to be rotatable into state $|H\rangle$, these two bands of potential states (shown in Figs. 4(a) and (b)) must overlap.

Briefly examining the geometry of this system, it appears that for most states this will be possible as long as the waveplate phases do not differ too much from the ideal HWP and QWP. Further consideration reveals that it is sufficient to be able to project into the states on the H-R-V-L great circle. There are two conditions under which this will *not* occur. First, if WP_2 is too close to a HWP, with WP_1 far from a QWP, the states at the poles (close to $|R\rangle$ and $|L\rangle$) will be unreachable from $|H\rangle$ (see Fig. 4(c)). Quantifying this condition, we require that

$$2\left|\frac{\pi}{2} - \phi_1\right| \leq \pi - \phi_2. \tag{80}$$

Put another way, the error in the QWP must be less than half the error in the HWP. Second, the combined retardances from both waveplates can be insufficient to reach $|V\rangle$ (see Fig. 4(c)):

$$\phi_1 + \phi_2 \geq \pi. \tag{81}$$

Given these two conditions, numerical simulations confirm that arbitrary single-qubit projectors can be constructed with two waveplates.

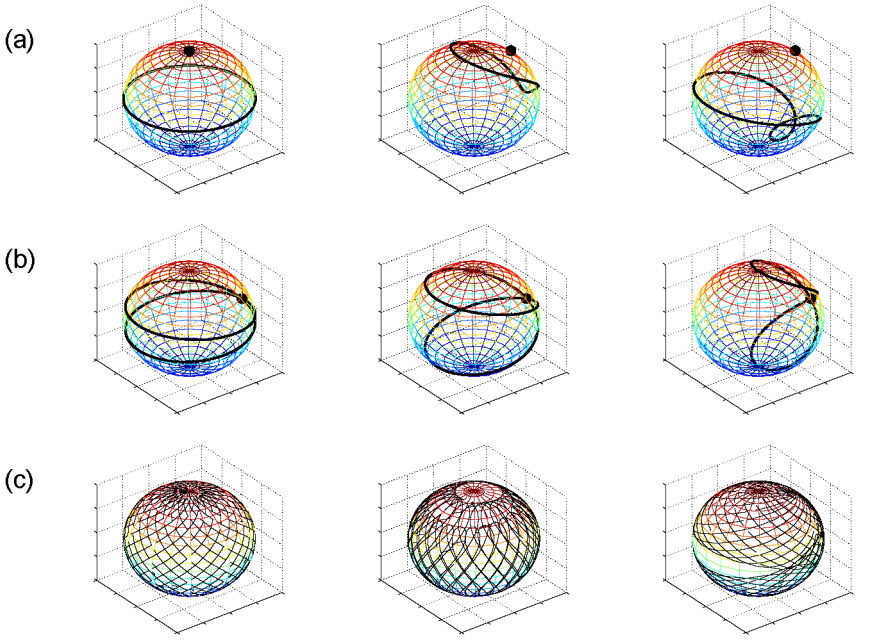

FIG. 4. Possible projectors simulated by waveplates and a stationary polarizer, graphically shown on Poincaré spheres. (a) WP_1, depending on its orientation, can rotate an incoming state into a variety of possible output states. Shown here on three Poincaré spheres are an initial incoming state (represented by a solid dot) and the set of all output states that WP_1 can rotate it into (represented by a dark band on the surface of the sphere). From left to right, the spheres depict $|R\rangle$ transformed by a $(\pi/2)$-waveplate, $|\gamma\rangle = \cos(\pi/8)|H\rangle + \mathrm{i}\sin(\pi/8)|V\rangle$ transformed by a $(\pi/3)$-waveplate, and $|\gamma\rangle$ transformed by a $(2\pi/3)$-waveplate. (b) WP_2, depending on its orientation, can rotate a variety of states into the target state $|H\rangle$. Shown here from left to right are the states able to be rotated into $|H\rangle$ by a $(11\pi/12)$-waveplate, a $(3\pi/4)$-waveplate, and a $(\pi/2)$-waveplate. (c) The possible projectors able to be produced by two waveplates and a horizontal polarizer. A series of arcs blanketing the Poincaré sphere show the areas of the sphere representing achievable projectors for each waveplate combination. From left to right, the spheres show the states (in this case, all of them) accessible from an ideal QWP and HWP, the states accessible using $(\pi/3)$- and $(11\pi/12)$-waveplates (groups of states near the poles are inaccessible), and the states accessible using $(\pi/3)$- and $(3\pi/5)$-waveplates (states on the equator are inaccessible). Note that the spheres shown in (c) are *not* simply combinations of the spheres above it, but include retardance values chosen to illustrate the possible failure modes of imperfect waveplates.

To clarify, as discussed in the previous section, one does not *require* arbitrary single-qubit projectors, since an accurate tomography can be obtained with *any* set of linearly independent projectors as long as they are known. In fact, one advantage to this approach is that the exact same tomography measurement system can be used on photons with different wavelengths (on which the waveplates' birefringent phase retardances depend), simply by entering in the analysis pro-

gram what the actual phase retardances are at the new wavelengths (Peters et al., 2005).

Wedged waveplates It is an experimental reality that all commercially available waveplates have some degree of *wedge* (i.e., the surfaces of the waveplate are not parallel). This leads to a number of insidious difficulties which the experimenter must confront, grouped into two categories: (1) The thickness of the waveplate will change along its surface, providing a corresponding change in the phase retardance of the waveplate. This means that during a tomography when the waveplate is routinely rotated to different orientations, its total phase after rotation will change according to a much more complex—and often very difficult to calculate—formula. (In fact, if a large collection aperture is used, then different parts of the beam will experience different phase shifts.) (2) The direction (k-vector) of a beam will be deflected after passing through a wedged waveplate. This deflection will again depend on waveplate orientation, therefore changing throughout a tomography. This can have the effect of changing detector efficiencies (if, as in our case, a lens is used to focus to a portion of a very small detector area, different pieces of which have different efficiencies). This deflection will also affect any interferometric effects that depend on the beam direction being stable under waveplate rotation. Some of these problems can be mitigated (e.g., by taking care to pass through the exact center of the waveplate), but in general the best solution is to select waveplates with faces very close to parallel.

3.1.3. Multiple-Qubit Projections and Measurement Ordering

For multiple-qubit systems, separable projectors can be implemented by using in parallel the single-qubit projectors described above. This, by construction, allows the implementation of arbitrary separable projectors.

In practice, depending on the details of a specific tomography (see Section 5 for a discussion of how to choose which and how many measurements to use), multiple-qubit tomographies can require a large number of measurements. If the time to switch from one measurement to another varies depending on which measurements are switched between (as is the case with waveplates switching to different values for each projector), minimizing the time spent switching is a problem equivalent to the travelling salesman problem (Cormen et al., 2001). A great deal of time can be saved by implementing a simple, partial solution to this canonical problem (e.g., a genetic algorithm which is not guaranteed to find the optimal solution but likely to find a comparably good solution).

3.2. n VS. $2n$ DETECTORS

Until now, this chapter has discussed the use of an array of n detectors to measure a single separable projector at a time. While this is conceptually simple, there is

an extension to this technique which can dramatically improve the efficiency and accuracy of a tomography: using an array of $2n$ detectors, project every incoming n-qubit state into one of 2^n basis states. This is the generalization of simultaneously measuring both outputs in the single-qubit case (the two detectors used for single-qubit measurement are shown in Fig. 3(a)), or all four basis states (HH, HV, VH, and VV) in the two-qubit case; in the general case $2n$ detectors will measure in n-fold coincidence with 2^n possible outcomes.

It should be emphasized that these additional detectors are not some 'trick', effectively masking a number of sequential settings of n detectors. If only n detectors are used, then over the course of a tomography most members comprising the input ensemble will never be measured. For example, consider measuring the projection of an unknown state into the $|00\rangle$ basis using two detectors. While this will give some number of counts, unmeasured coincidences will be routed into the $|01\rangle$, $|10\rangle$, and $|11\rangle$ modes. The information of how many coincidences are routed to which mode will be lost, unless another two detectors are in place in the '1' modes to measure it.

Returning to the notation of Section 3.1, recall that the state which passes through every beamsplitter is $U_v^\dagger|0\rangle$, but when $2n$ detectors are employed, the states $U_v^\dagger|r\rangle$ can all be measured, where r ranges from 0 to $2^n - 1$ and $|r\rangle$ denotes the rth element of the canonical basis (the canonical basis is chosen/enforced by the beamsplitters themselves).

EXAMPLE 11 (*The $|r\rangle$ notation for two qubits*). For two qubits, each incident on separate beamsplitters which transmit $|H\rangle$ and reflect $|V\rangle$, we can define the following values of $|r\rangle$, the canonical basis:

$$|0\rangle \equiv |HH\rangle, \qquad |1\rangle \equiv |HV\rangle, \qquad |2\rangle \equiv |VH\rangle, \qquad |3\rangle \equiv |VV\rangle. \tag{82}$$

The usefulness of this notation will become apparent during the discussion of the Maximum Likelihood algorithm in Section 4.2.

The primary advantage to using $2n$ detectors is that every setting of the analysis system (every group of the projector and its orthogonal complements) generates exactly enough information to determine a single multiple-qubit Stokes vector. Expanding out the probabilities that a multiple-qubit Stokes vector (which for now we limit to those with only nonzero indices) is based on,

$$\begin{aligned} S_{i_1,i_2,\ldots,i_n} &= \left(P_{\psi_1} - P_{\psi_1^\perp}\right) \otimes \left(P_{\psi_2} - P_{\psi_2^\perp}\right) \otimes \cdots \otimes \left(P_{\psi_n} - P_{\psi_n^\perp}\right) \\ &= P_{\psi_1,\psi_2,\ldots,\psi_n} - P_{\psi_1,\psi_2,\ldots,\psi_n^\perp} - \cdots \pm P_{\psi_1^\perp,\psi_2^\perp,\ldots,\psi_n^\perp}, \end{aligned} \tag{83}$$

where the sign of each term on the last line is determined by the parity of the number of orthogonal ($\perp$) terms.

These probabilities are precisely those measured by a single setting of the entire analysis system followed by a $2n$ detector array. Returning to our primary

decomposition of the density matrix from Eq. (35),

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1,i_2,\ldots,i_n=0}^{3} S_{i_1,i_2,\ldots,i_n} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \cdots \otimes \hat{\sigma}_{i_n},$$

we once again need only determine all of the multiple-qubit Stokes parameters to exactly characterize the density matrix. At first glance this might seem to imply that we need to use $4^n - 1$ settings of the analysis system, in order to find all of the multiple-qubit Stokes parameters save $S_{0,0,\ldots,0}$, which is always one.

While this is certainly sufficient to solve for $\hat{\rho}$, many of these measurements are redundant. In order to choose the smallest possible number of settings, note that the probabilities that constitute some multiple-qubit Stokes parameters overlap exactly with the probabilities for other multiple-qubit Stokes parameters. Specifically, any multiple-qubit Stokes parameter with at least one 0 subscript is derived from a set of probabilities that at least one other multiple-qubit Stokes vector (with no 0 subscripts) is also derived from. As an example, consider that

$$S_{0,3} = P_{|00\rangle} - P_{|01\rangle} + P_{|10\rangle} - P_{|11\rangle}, \tag{84}$$

while

$$S_{3,3} = P_{|00\rangle} - P_{|01\rangle} - P_{|10\rangle} + P_{|11\rangle}. \tag{85}$$

These four probabilities, measured simultaneously, will provide enough information to determine both values. This dependent relationship between multiple-qubit Stokes vectors is true, in general, as can be seen by returning to Eq. (83). Each subscript with nonzero value for S contributes a term to the tensor product on the right that looks like $(P_{\psi_i} - P_{\psi_i^\perp})$. Had there been subscripts with value zero, however, they each would have contributed a $(P_{\psi_i} + P_{\psi_i^\perp})$ term; as an aside, terms with zero subscripts are always dependent on terms will all positive subscripts. This reduces the minimum number of analysis settings to 3^n, a huge improvement in multiple qubit systems (e.g., 9 vs. 15 settings for 2-qubit tomography, 81 vs. 255 for 4-qubit tomography, etc.). Note that, as discussed earlier, this benefit is only possible if one employs $2n$ detectors, leading to a total of 6^n measurements (2^n measurements for each of 3^n analysis settings).[11]

Because Eq. (41) can be used to transform any set of nonorthogonal multiple-qubit Stokes parameters into the canonical form, orthogonal measurement sets

[11] These measurements, even though they result from the minimum number of analysis settings for $2n$ detectors, are overcomplete. A density matrix has only $4^n - 1$ free parameters, which implies that only $4^n - 1$ measurements are necessary to specify it (see n-detector tomography). Because the overcomplete set of 6^n measurements is not linearly independent, it can be reduced to a $4^n - 1$ element subset and still completely specify an unknown state.

need not be used. One advantage of the option to use nonorthogonal measurement sets is that an orthogonal set may not be experimentally achievable, for instance, due to waveplate imperfections, as discussed in Section 3.1.2.

3.3. Electronics and Detectors

Single photon detectors and their supporting electronics are crucial to any photonic tomography. Figure 5 shows a simple diagram of the electronics used to count in coincidence from a pair of Si-avalanche photodiodes. An electrical pulse from a single-photon generated avalanche in the silicon photodiode sends a signal to a discriminator, which, after receiving a pulse of the appropriate amplitude and width, produces in a fan-out configuration several TTL (transistor–transistor logic) signals which are fed into the coincidence circuitry. In order to avoid pulse reelections, a fan-out configuration is used in preference to repeatedly splitting one signal.

The signals from these discriminators represent physical counts, with the number of discriminator signals sent to a detector equal to the singles counts for that detector. A copy of this signal, after travelling through a variable length delay line, is input into an AND gate with a similar pulse (with a static delay) from a complimentary detector. The pulses sent from the discriminators are variable width, typically about 2 ns, producing a 4-ns window in which the AND gate can produce a signal. (The coincidence window is chosen to be as small as conveniently possible, in order to reduce the number of "accidental" coincidences, discussed below.) This signal is also sent to the counters and is recorded as a coincidence between its two parent detectors.

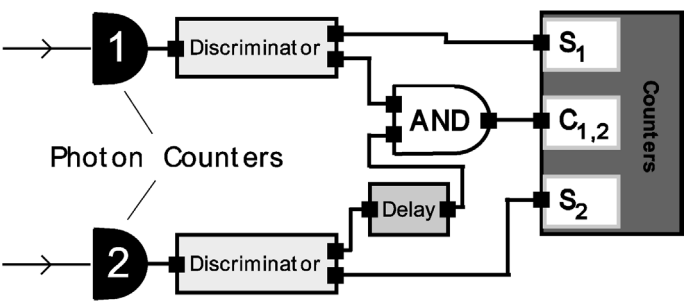

FIG. 5. A simple diagram of the electronics necessary to operate a coincidence-based photon counting circuit. While this diagram depicts a two-detector counting circuit, it is easily extendible to multiple detectors; by adding additional detectors each fed into a discriminator and a fan-out, we gain the signals necessary for one singles counter per detector and an AND gate for each pair of detectors capable of recording a coincidence.

As with any system of this sort, the experimenter must be wary of reflected pulses generating false counts, delay lines being properly matched for correct AND gate operation, and system saturation for high count rates.

3.4. Collecting Data and Systematic Error Correction

The projection optics and electronics described above will result in a list of coincidence counts each tied to a single projective measurement. Incorporating the projectors defined earlier in this section, we can now make a first estimate on the number of counts we expect to receive for a given measurement of the state $\hat{\rho}$:

$$\begin{aligned} \bar{n}_{\nu,r} &= I_0 \mathrm{Tr}\{\widetilde{M}_{\nu,r}\hat{\rho}\}, \\ \widetilde{M}_{\nu,r} &= U_\nu^\dagger |r\rangle\langle r| U_\nu. \end{aligned} \tag{86}$$

Our eventual strategy (see Section 4.2) will be to vary $\hat{\rho}$ until our expectations optimally match our actual measured counts. Here $\bar{n}_{\nu,r}$ is the expectation value of the number of counts recorded for the νth measurement setting on the rth pair of detectors (this is the pair of detectors which projects into the canonical basis state $|r\rangle$). The density matrix to be measured is denoted by $\hat{\rho}$ and I_0 is a constant scaling factor which takes into account the duration of a measurement and the rate of state production. Note that regardless of whether n or $2n$ detectors are used, each distinct measurement setting will be indexed by ν. For n detectors, there will be a single value of r for each value of ν, as each measurement setting projects into a single state. For $2n$ detectors, there will be 2^n values of r, one for each pair of detectors capable of registering coincidences.

Throughout this section we will modify Eq. (86) to give a more complete estimate of the expected count rates, taking into account real errors and statistical deviations. In particular, without adjustment, the expected coincidence counts will likely be inaccurate without adjustment due to experimental factors including accidental coincidences, imperfect optics, mismatched detector efficiencies, and drifts in state intensity. Below we will discuss each of these in turn.

3.4.1. Accidental Coincidences

In general, the spontaneous generation of photon pairs from downconversion processes can result in several pairs of photons being generated at the same time. These multiple-pair generation events can lead to two uncorrelated photons being detected as a coincidence, which will tend to raise all measured counts and lead to state tomographies resulting in states closer to the maximally mixed state.[12]

[12] There is also a similar but generally smaller contribution from one real photon and a detector noise count, and a smaller contribution still from two detector noise counts.

We can model these accidental coincidences for the two-qubit case by considering the probability that *any* given singles count will be detected during the coincidence window of a conjugate photon. This model implies that the accidental coincidences for the νth measurement setting on the rth detector pair ($n_{\nu,r}^{\text{accid}}$) will be dependent on the singles totals in each channel (${}^1S_{\nu,r}$ and ${}^2S_{\nu,r}$), the total coincidence window (Δt_r, approximately equal to twice the pulse width produced by the discriminators[13]), and the total measurement time (T_ν). When the singles channels are far from saturation (${}^{1,2}S_{\nu,r}\Delta t_{\text{dead}} \ll T_\nu$, where Δt_{dead} is the dead time of the detectors, i.e., the time it takes after a detector registers a singles count before it can register another), the percentage of time that a channel is triggered (able to produce a coincidence) is approximated by ${}^{1,2}S_{\nu,r}\Delta t_r/T_\nu$. The probability that the other channel will produce a coincidence within this time (again in the unsaturated regime) is proportional to the singles counts on that channel. This allows us to approximate the total accidentals as

$$n_{\nu,r}^{\text{accid}} \simeq \frac{{}^1S_{\nu,r}\,{}^2S_{\nu,r}\Delta t_r}{T_\nu}, \tag{87}$$

implying that

$$\bar{n}_{\nu,r} = I_0 \text{Tr}\left\{\widetilde{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}}. \tag{88}$$

Because the accidental rate will be necessary for analyzing the data, these expected accidental counts will need to be calculated from the singles rates for each measurement and recorded along with the actual measured coincidence counts.[14]

3.4.2. Beamsplitter Crosstalk

In most experimental implementations, particularly those involving $2n$ detectors, the polarizer used for single-qubit projection will be a beamsplitter, either based on dielectric stacks, or crystal birefringence. In practice, all beamsplitters function with some levels of crosstalk and absorption, i.e., some probability of reflecting or absorbing the polarization which should be transmitted and vice versa. By measuring these crosstalk probabilities and adjusting the measured counts accordingly, it is possible to recreate the approximate measurement values that would have resulted from a crosstalk-free system.

We can characterize a beamsplitter using four numbers $\mathcal{C}_{r'\to r}$ which represent the probability that state r' will be measured as state r.

[13] If the pulses are not square, or the AND logic has speed limitations, this approximation may become inaccurate.

[14] It is advisable to initially experimentally determine Δt_r by directly measuring the accidental coincidence rate (by introducing an extra large time delay into the variable time delay before the AND gate, shown in Figure 5), and using Eq. (87) to solve for Δt_r. This should be done for every pair of detectors, and ideally at several count rates, in case there are nonlinear effects in the detectors.

EXAMPLE 12 (*A faulty beamsplitter*). Assume that we have measured a beamsplitter which transmits 90% and absorbs 10% of incident horizontal light (state 0), while reflecting 80% and transmitting 10% of vertical light (state 1). We would therefore use

$$\begin{aligned} \mathcal{C}_{0\to 0} &= 0.9, \\ \mathcal{C}_{0\to 1} &= 0, \\ \mathcal{C}_{1\to 0} &= 0.1, \\ \mathcal{C}_{1\to 1} &= 0.8, \end{aligned} \tag{89}$$

to characterize the behavior of this beamsplitter.

EXAMPLE 13 (*Two-qubit crosstalk*). Consider two faulty beamsplitters identical to the one presented in Example 12, with crosstalk coefficients $C^{A}_{r'\to r}$ and $C^{B}_{r'\to r}$. Assuming that we label the two qubit canonical basis $|r\rangle$ as $|0\rangle \equiv |HH\rangle$, $|1\rangle \equiv |HV\rangle$, $|2\rangle \equiv |VH\rangle$, and $|3\rangle \equiv |VV\rangle$, we can derive the general two-qubit crosstalk coefficients $C_{r'\to r}$ by multiplying the single-qubit crosstalk coefficients, according to the rule:

$$C_{(2r'_A+r'_B)\to(2r_A+r_B)} \equiv C^{A}_{r'_A\to r_A} C^{B}_{r'_B\to r_B}. \tag{90}$$

Thus, the total crosstalk matrix will be

$$C_{r'\to r} \equiv \begin{array}{c} \\ \to 0 \\ \to 1 \\ \to 2 \\ \to 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0' & 1' & 2' & 3' \end{array} \\ \begin{pmatrix} 0.81 & 0.09 & 0.09 & 0.01 \\ 0 & 0.72 & 0 & 0.08 \\ 0 & 0 & 0.72 & 0.08 \\ 0 & 0 & 0 & 0.64 \end{pmatrix} \end{array}. \tag{91}$$

If we use this notation to modify Eq. (88) for predicted counts, we find that

$$\begin{aligned} \bar{n}_{v,r} &= I_0 \mathrm{Tr}\{\widehat{M}_{v,r}\hat{\rho}\} + n^{\mathrm{accid}}_{v,r}, \\ \widehat{M}_{v,r} &\equiv \sum_{r'} (C_{r'\to r})\widetilde{M}_{v,r'}. \end{aligned} \tag{92}$$

3.4.3. Detector-Pair Efficiency Calibration

Because single-photon detectors will in general have different efficiencies, it may be necessary to measure the relative efficiencies of any detector pairs used in the course of a tomography. For the n-detector case, this is unnecessary, as all recorded counts will be taken with the same detectors and scaled equally. For the $2n$-detector configuration, this can be a noticeable problem, with each of the n^2

measurement bases using a different combination of detectors, with a different total coincidence efficiency. By measuring the relative efficiencies of each combination, it is possible to correct the measured counts by dividing them by the appropriate relative efficiency.

Note that it is not necessary to know the *absolute* efficiency of each detector combination, but only the *relative* efficiencies. Knowing only the relative efficiencies leaves a single scaling factor that is applied to all counts, but as the error on a set of counts is dependent on the *measured* counts, rather than the total number of incident states, this ambiguity does not affect the tomography results.

The tomography process itself may be used to conveniently determine the full set of relevant relative efficiencies. By performing enough measurements to perform an n-detector tomography while using $2n$ detectors, it is possible to perform a tomography for *each* detector combination, using only the results of that detector combination's measurements. Each of these sets will be sufficient to perform a tomography, and the tomography algorithm (see Section 4) will necessarily determine the total state intensity. The ratios between these state intensities (one for each detector combination) will provide the relative efficiencies of each detector combination. In the two-qubit case, this means using four detectors and 36 measurement settings, for a total of 144 measurements to calibrate the relative efficiencies.

In order to continue to update our equation for $\bar{n}_{v,r}$, we define an efficiency $\widehat{E}_r$ which describes the relative efficiency of the rth detector combination. This allows us to correct our previous equation to

$$\bar{n}_{v,r} = I_0 E_0 E_r \mathrm{Tr}\left\{\widehat{M}_{v,r}\hat{\rho}\right\} + n_{v,r}^{\mathrm{accid}}, \tag{93}$$

where E_0 is a constant scalar, which combined with the easier to measure relative efficiency E_r, gives the absolute efficiency of each detector pair.

3.4.4. Intensity Drift

In polarization experiments based on downconversion sources, a major cause of error can be drift in the intensity (or direction) of the pump, which causes a drift in the rate of downconversion and therefore state production. If this intensity drift is recorded, then the prediction of the expected number of counts can be adjusted to account for this additional information. Alternatively, if $2n$ detectors are used, the sum of the counts from each of the detectors will automatically give the normalized intensity for each measurement setting, since the sum of the counts in orthonormal bases must add up to the total counts (assuming no state-dependent losses, e.g., in the polarizing beamsplitters). However, when summing the counts from a complete basis like this, the measurements *must be taken at the same time*, and the summed counts must take other sources of error, like detector inefficiency, accidental counts, and beamsplitter crosstalk into account.

By whatever method it is measured, assume that the relative size of the ensemble subject to the νth measurement setting is given by I_ν. Then

$$\bar{n}_{\nu,r} = I_0 E_0 I_\nu E_r \mathrm{Tr}\{\widehat{M}_{\nu,r}\hat{\rho}\} + n_{\nu,r}^{\mathrm{accid}}. \tag{94}$$

Now it becomes clear that I_0 is the factor (not necessarily the total number of pairs produced) which, combined with the relative efficiency I_ν, gives the total number of incident states for the νth measurement setting.

4. Analyzing Experimental Data

As discussed earlier, any real experiment will contain statistical and systematic errors which preclude the use of the ideal tomography described in Section 2. Instead, it is necessary to use an algorithm (the Maximum Likelihood technique) which assumes some uncertainty or error in measurement results, and returns a state which is the most likely to have produced the measured results.

In order to describe real tomography, we will first discuss the types of errors which are present in an experiment, the Maximum Likelihood algorithm, and some details of the optimization of the entire process using numerical search techniques. We first list the information

U_ν—measurement settings,
$n_{\nu,r}$—counts recorded,
$n_{\nu,r}^{\mathrm{accid}}$—accidental counts,
$\mathcal{C}_{r'\to r}$—crosstalk coefficients,
E_r—relative efficiencies,
I_ν—relative intensities (not used with $2n$ detectors)

that should have been gathered during the experimental phase of the experiment, followed by the formulae

$$\begin{aligned}
\bar{n}_{\nu,r} &\equiv I_0 E_0 I_\nu E_r \mathrm{Tr}\{\widehat{M}_{\nu,r}\hat{\rho}\} + n_{\nu,r}^{\mathrm{accid}},\\
\widehat{M}_{\nu,r} &\equiv \sum_{r'} (C_{r'\to r})\widetilde{M}_{\nu,r'},\\
\widetilde{M}_{\nu,r} &\equiv U_\nu^\dagger |r\rangle\langle r| U_\nu
\end{aligned}$$

used to determine the expected number of measured counts for the νth measurement setting on the rth detector combination.

Given this information, we are able to numerically estimate which state was most likely to return the measured results. Note that the relative intensities I_ν are optional, but can be included for an n-detector tomography. For a $2n$ detector tomography, the I_ν parameters are varied as part of the optimization algorithm, and do not need to be provided as part of the experimental data.

4.1. Types of Errors and State Estimation

Errors in the measurement of a density matrix fall into three main categories: errors in the measurement basis, errors from counting statistics, and errors from experimental stability. The first problem can be addressed by increasing the accuracy of the measurement apparatus (e.g., obtaining higher tolerance waveplates, better controlling the Rabi pulses, etc.) while the second problem is reduced by performing each measurement on a larger ensemble (counting for a longer time). The final difficulty is drift which occurs over the course of the tomography.[15] This drift occurs either in the state produced or the efficiency of the detection system, and can constrain the data-collection time. Figure 6(a) shows what a basis error looks like on the Poincaré sphere and how that error affects the ability to isolate a state in Poincaré space. This picture indicates that a basis error is more pronounced when measuring a pure state, but actually has no effect when measuring a totally mixed state (because all bases give the same answer).

Figure 6(b) shows the same analysis of errors in counting statistics. Any real measurement can be carried out only on a limited size ensemble. Though the details of the statistics will be dealt with later, the detection events are accurately

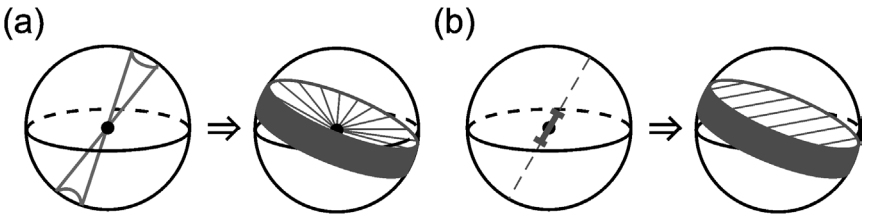

Fig. 6. Graphical representation of errors in a single-qubit tomography. (a) Basis errors. Errors in the setting of measurement apparatus can result in an accurate measurement being taken in an unintended basis. Shown graphically is the effect that an uncertainty in the measurement basis can have on the reconstruction of a state. Instead of a single axis on the Poincaré sphere, the possible measurement axes form uncertainty cones touching at the center, since all possible measurement axes pass through the origin. This uncertainty in axis is then translated into an uncertainty in the state (shown on the right). Instead of isolating the state to a plane, all possible measurement axes trace out a volume with large uncertainty near the surface of the sphere and low uncertainty near the center. (b) Counting errors. Even if the measurement basis is exactly known, only a limited number of qubits can be measured to gain an estimate of a state's projection onto this axis (taken directly from the probability of a successful measurement). This uncertainty results in an unknown state being isolated to a one-dimensional Gaussian (approximately) in three-dimensional space, rather than to a plane.

[15] These are the main sources of error that are likely to be present to some degree in *any* qubit implementation. In addition, each implementation may have its own unique errors, such as the wedged waveplates described earlier or accidental background counts from noisy detectors. Here we neglect such system-specific difficulties.

described by a Poissonian distribution, which for large numbers of counts is well approximated by a Gaussian distribution. This will cause the resultant knowledge about the unknown state to change from a plane (in the exact case) to a thick disk (uniformly thick for pure and mixed states), a one-dimensional Gaussian distribution plotted in three-dimensional space.

After all sources of error are taken into account, a single measurement results in a distribution over all possible states describing the experimenter's knowledge of the unknown state. This distribution represents the likelihood that a particular state would give the measured results, relative to another state. When independent measurements are combined, these distributions are multiplied, and ideally the knowledge of the unknown state is restricted to a small ball in Poincaré space, similar to a three-dimensional Gaussian (as a large uncertainty in any one direction will lead to a large uncertainty in the state). State isolation occurs regardless of which measurements are taken, as long as they are linearly independent, and is shown graphically in Fig. 7 for a set of orthogonal measurements.

In contrast to the ideal case in the previous section, for which the accuracy of a reconstructed state did not depend on whether mutually unbiased measurements were made, with real measurements the advantage of mutually unbiased measurement bases becomes clear. In contrast to the measurements shown in Fig. 7, mutually biased measurements result in a nonsymmetric error ball, increasing the error in state estimation in one direction in Hilbert space.

Even after tomography returns a distribution of likelihood over Poincaré space, one final problem remains. It is very possible, especially with low counts or with

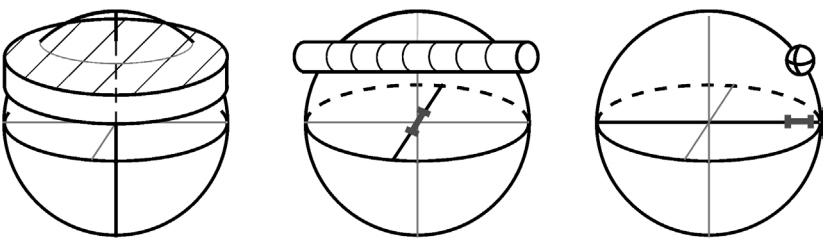

FIG. 7. Isolation of a quantum state through inexact measurements. Although a series of real measurements (those with uncertainties) will never be able to exactly isolate an unknown quantum state, they can isolate it to a region of Hilbert space that is far more likely than any other region to contain the unknown state. Consider a series of three measurements, each containing counting errors, along orthogonal axes. From left to right, the area of Hilbert space containing the unknown state is truncated from a one-dimensional Gaussian probability distribution (the disk in the left figure) to a two-dimensional Gaussian (the cylinder in the middle figure) and finally to a three-dimensional Gaussian (the ball in the right figure). This results in an 'error ball' which approximates the position of the unknown state. The global maximum, however, can often be outside allowed Hilbert space (outside the Poincaré sphere), which is one reason a maximum likelihood technique must be used to search over only allowed quantum states.

the measurement of very pure states, that state estimation will return an "illegal" state. For example, in Fig. 7, the measurements seem to place the error ball just on the edge of the sphere and slightly outside it. As all legal states have a radius of less than or equal to one in Poincaré space, it is necessary to find a way to return the most likely *legitimate* state reconstructed from a set of measurements.

4.2. The Maximum Likelihood Technique

The problem of reconstructing illegal density matrices is resolved by selecting the legitimate state most likely to have returned the measured counts (James et al., 2001; Hradil and Rehacek, 2001). In practice, analytically calculating this maximally likely state is prohibitively difficult, and a numerical search is necessary. Three elements are required: a manifestly legal parametrization of a density matrix, a likelihood function which can be maximized, and a technique for numerically finding this maximum over a search of the density matrix's parameters.

The Stokes parameters are an unacceptable parametrization for this search, as there are clearly combinations of these parameters which result in an illegal state (e.g., $S_1 = S_2 = S_3 = 1$). In this context, a legitimate state refers to a nonnegative definite Hermitian density matrix of trace one. The property of nonnegative definiteness for any matrix $\widehat{\mathcal{G}}$ is written mathematically as

$$\langle\psi|\widehat{\mathcal{G}}|\psi\rangle \geq 0 \quad \forall|\psi\rangle. \tag{95}$$

Any matrix that can be written in the form $\widehat{\mathcal{G}} = \widehat{T}^\dagger\widehat{T}$ must be nonnegative definite. To see that this is the case, substitute into Eq. (95):

$$\langle\psi|\widehat{T}^\dagger\hat{T}|\psi\rangle = \langle\psi'|\psi'\rangle \geq 0, \tag{96}$$

where we have defined $|\psi'\rangle = \widehat{T}|\psi\rangle$. Furthermore $(\widehat{T}^\dagger\widehat{T})^\dagger = \widehat{T}^\dagger(\widehat{T}^\dagger)^\dagger = \widehat{T}^\dagger\widehat{T}$, i.e., $\widehat{\mathcal{G}} = \widehat{T}^\dagger\widehat{T}$ must be Hermitian. To ensure normalization, one can simply divide by the trace. Thus the matrix $\hat{g}$ given by the formula

$$\hat{g} = \frac{\widehat{T}^\dagger\widehat{T}}{\mathrm{Tr}\{\widehat{T}^\dagger\widehat{T}\}} \tag{97}$$

has all three of the mathematical properties required for density matrices.

For the one-qubit system, we have a 2×2 density matrix with 3 independent real parameters (although we will search over 4 in order to fit the intensity of the data). Since it will be useful to be able to invert relation (97), it is convenient to choose a tri-diagonal form for $\widehat{T}$:

$$\widehat{T}(\vec{t}\,) = \begin{pmatrix} t_1 & 0 \\ t_3 + \mathrm{i}t_4 & t_2 \end{pmatrix} \tag{98}$$

where $\vec{t}$ is a vector containing each t_i. The multiple-qubit form of the same equation is given by:

$$\widehat{T}(\vec{t}) = \begin{pmatrix} t_1 & 0 & \cdots & 0 \\ t_{2^n+1} + \mathrm{i}t_{2^n+2} & t_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ t_{4^n-1} + \mathrm{i}t_{4^n} & t_{4^n-3} + \mathrm{i}t_{4^n-2} & t_{4^n-5} + \mathrm{i}t_{4^n-4} & t_{2^n} \end{pmatrix}. \tag{99}$$

The manifestly 'physical' density matrix $\hat{\rho}_p$ is then given by the formula

$$\hat{\rho}_p(\vec{t}) = \frac{\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})}{\mathrm{Tr}\{\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})\}}. \tag{100}$$

This satisfies the first criterion for a successful maximum likelihood search, by providing an explicitly physical parametrization for $\hat{\rho}$. The second criterion, a likelihood function, will in general depend on the specific measurement apparatus used and the physical implementation of the qubit (as these will determine the statistical distributions of counts, and therefore their relative weightings). If we assume Gaussian counting statistics, then we can easily provide a suitable likelihood function.

Let $n_{\nu,r}$ be the result for the νth measurement setting on the rth detector combination. Let $\bar{n}_{\nu,r}$ be the counts that would be expected from the state $\hat{\rho}$, given all information about the system:

$$\bar{n}_{\nu,r} \equiv I_0 E_0 I_\nu E_r \mathrm{Tr}\{\widehat{M}_{\nu,r}\hat{\rho}\} + n_{\nu,r}^{\mathrm{accid}} \tag{101}$$

$$\widehat{M}_{\nu,r} \equiv \sum_{r'} (C_{r'\to r})\widetilde{M}_{\nu,r'} \tag{102}$$

$$\widetilde{M}_{\nu,r} \equiv U_\nu^\dagger |r\rangle\langle r| U_\nu. \tag{103}$$

Given that we wish to search over the parameters of $\vec{t}$, rather than $\hat{\rho}$, we will rewrite this equation as

$$\bar{n}_{\nu,r} = I_\nu E_r \mathrm{Tr}\{\widehat{M}_{\nu,r}\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})\} + n_{\nu,r}^{\mathrm{accid}}. \tag{104}$$

Notice that the unknown scalars I_0 and E_0 have been absorbed into the unnormalized $\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})$, allowing our numerical search to discover what their combined effect is without ever knowing their individual values.

Given these definitions, the probability of obtaining the νth measurement on the rth set of detectors, $n_{\nu,r}$, from the search parameters $\vec{t}$ is proportional to

$$\exp\left[-\frac{(\bar{n}_{\nu,r} - n_{\nu,r})^2}{2\hat{\sigma}_{\nu,r}^2}\right], \tag{105}$$

where $\hat{\sigma}_{\nu,r}$ is the standard deviation of the νth measurement (given approximately by $\sqrt{\bar{n}_{\nu,r}}$). Therefore, the total probability of $\hat{\rho}$ yielding the counts $\{n_{\nu,r}\}$ is given

by:

$$P(n_{v,r}) = \frac{1}{Norm} \prod_{v,r} \exp\left[-\frac{(\bar{n}_{v,r} - n_{v,r})^2}{2\bar{n}_{v,r}} \right], \tag{106}$$

where *Norm* is the normalization constant. In order to find the ideal $\vec{t}$, and therefore the ideal $\hat{\rho}$, we need to maximize the probability function above. This is equivalent to maximizing the log of the same function, or equivalently, minimizing its negation, giving us our final likelihood function (notice that the normalization constant is ignored for this function, as it will not affect the minimum):

$$\mathcal{L}(\vec{t}) = \sum_{v,r} \frac{(\bar{n}_{v,r} - n_{v,r})^2}{2\bar{n}_{v,r}}. \tag{107}$$

The final piece in the maximum likelihood technique is an optimization routine, of which there are many available. The authors' implementation will be discussed in the next subsection.[16] After a minimum is found, $\hat{\rho}$ can be reconstructed from the values of $\vec{t}$.

EXAMPLE 14 (*A single-qubit tomography*). Photon pairs generated via spontaneous parametric downconversion from a nonlinear crystal can be used to generate single-photon states. Measuring a photon in one arm collapses the state of its partner to a single-qubit Fock state (Hong and Mandel, 1986). An ensemble of these photons can be characterized using the maximum likelihood technique. The following data was taken from an experiment in "Remote State Preparation" (Peters et al., 2005):

$$H = 6237, \qquad D = 5793,$$
$$V = 8333, \qquad R = 6202.$$

For this first example we will assume that no intensity normalization or crosstalk compensation needs to occur (see Example 15 for a more thorough example). After minimizing the likelihood function, we obtain the following $\widehat{T}$ matrix

$$\widehat{T} = \begin{pmatrix} 73.4 & 0 \\ -29.0 - 1.2\mathrm{i} & 77.1 \end{pmatrix}, \tag{108}$$

from which we can derive the density matrix,

$$\hat{\rho} = \frac{\widehat{T}^\dagger \widehat{T}}{\mathrm{Tr}\{\widehat{T}^\dagger \widehat{T}\}} = \begin{pmatrix} 0.5121 & 0.1837 + 0.0075\mathrm{i} \\ 0.1837 - 0.0075\mathrm{i} & 0.4879 \end{pmatrix}. \tag{109}$$

[16] For freely available code and further examples, see: http://www.physics.uiuc.edu/research/QuantumPhotonics/Tomography/

Note that the maximum likelihood technique easily adapts to measurements in mutually biased bases (e.g., due to imperfect yet well characterized waveplates) and overcomplete measurements (taking more measurements than is necessary). In the first case the set of $|\psi\rangle$ is mutually biased (i.e., not in the canonical bases), though still governed by the mathematics of tomography we have laid out; in the second case the sum in Eq. (107) is extended beyond the minimum number of measurement settings.

4.3. Optimization Algorithms and Derivatives of the Fitness Function

In order to complete a tomography, the likelihood function $\mathcal{L}(\vec{t})$ must be minimized. A number of optimization programs exist which can search over a large number of parameters (e.g., $\vec{t}$) in order to minimize a complex function. The authors use the *Matlab 7.0* function lsqnonlin, which is optimized to minimize a sum of squares. This type of optimized algorithm is more efficient than a generic search, such as the *Matlab* function fminunc. In order for this minimization to work most effectively, it takes as parameters $f(\vec{t})$ and $\partial f(\vec{t})/\partial t_i$, where $\mathcal{L}(\vec{t})$ is of the form

$$\mathcal{L}(\vec{t}) = \sum_x \bigl[f_x(\vec{t})\bigr]^2. \tag{110}$$

For the problem of tomography, we can write

$$f_{v,r} \equiv \frac{\bar{n}_{v,r} - n_{v,r}}{\sqrt{2\bar{n}_{v,r}}} \tag{111}$$

$$= \frac{I_v E_r \mathrm{Tr}\{\widehat{M}_{v,r}\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})\} + n_{v,r}^{\mathrm{accid}} - n_{v,r}}{\sqrt{2(I_v E_r \mathrm{Tr}\{\widehat{M}_{v,r}\widehat{T}^\dagger(\vec{t})\widehat{T}(\vec{t})\} + n_{v,r}^{\mathrm{accid}})}}. \tag{112}$$

With some effort, we can analytically derive the partial derivatives of these terms, allowing the optimization algorithm to not only run much faster, but to converge quickly regardless of the initial search condition:

$$\begin{aligned}\frac{\partial f_{v,r}}{\partial t_i} &= \frac{\bigl[\frac{\partial}{\partial t_i}(\bar{n}_{v,r} - n_{v,r})\bigr]\sqrt{\bar{n}_{v,r}} - (\bar{n}_{v,r} - n_{v,r})\bigl(\frac{\partial}{\partial t_i}\sqrt{\bar{n}_{v,r}}\bigr)}{\sqrt{2}\bar{n}_{v,r}} \\ &= \frac{\bigl(\frac{\partial \bar{n}_{v,r}}{\partial t_i}\bigr)(\bar{n}_{v,r})^{1/2} - (\bar{n}_{v,r} - n_{v,r})\bigl[\frac{1}{2}(\bar{n}_{v,r})^{-1/2}\frac{\partial \bar{n}_{v,r}}{\partial t_i}\bigr]}{\sqrt{2}\bar{n}_{v,r}} \\ &= \frac{1}{2\sqrt{2\bar{n}_{v,r}}}\frac{\partial \bar{n}_{v,r}}{\partial t_i}\left(1 + \frac{n_{v,r}}{\bar{n}_{v,r}}\right).\end{aligned} \tag{113}$$

Note that it is impossible for this function to go to zero unless $\partial \bar{n}_{v,r}/\partial t_i$ goes to zero, important when considering whether or not the maximum likelihood function will have several local minima. Because $\widehat{T}$ is a linear function, we can easily write down

$$\begin{aligned}\frac{\partial \bar{n}_{v,r}}{\partial t_i} &= I_v E_r \frac{\partial}{\partial t_i} \mathrm{Tr}\big\{\widehat{M}_{v,r}\widehat{T}^{\dagger}(\vec{t})\widehat{T}(\vec{t})\big\} \\ &= I_v E_r \mathrm{Tr}\big\{\widehat{M}_{v,r}\big[\widehat{T}^{\dagger}(\vec{t})\widehat{T}(\vec{\delta}_{ij}) + \widehat{T}^{\dagger}(\vec{\delta}_{ij})\widehat{T}(\vec{t})\big]\big\}, \end{aligned} \tag{114}$$

where $\vec{\delta}_{ij}$ is a j-element vector whose ith element is equal to one. All other elements of $\vec{\delta}_{ij}$ are equal to zero. (Here j is the length of $\vec{t}$.)

Even using these derivatives (and especially if they are not used), it is important to choose an initial condition for the search which is as close as possible to the correct answer. This amounts to making the best analytic guess possible using the ideal tomographic techniques presented in Section 2. It is possible that those ideal techniques will result in an illegal density matrix, i.e., some of its eigenvalues will be negative. If this happens (indeed, this happening is the reason we *need* the Maximum Likelihood technique), we simply set those negative eigenvalues to zero, renormalize the positive eigenvalues, and use this truncated state as the starting condition for the search.

5. Choice of Measurements

After describing how to take measurements and how to analyze them, it is necessary to discuss how to choose which measurements to take. This general problem includes several choices, from whether to measure projectors independently or simultaneously as part of a complete basis (n versus $2n$ detectors) to how many measurements to take (the minimum set or an overcomplete set full of redundancy).

5.1. How Many Measurements?

The choice of how many measurements to perform depends on the details of the experimental setup and the goals of the experiment. For most applications, the speed of a tomography is paramount. At first it might appear obvious that the optimal number of measurements to perform would also be the *minimum* number of measurements to perform, seemingly implying the minimum time necessary to finish a tomography. This intuitive assumption is in fact false; it is often true that taking *more* measurements can in fact *reduce* the total time to achieve a specified level of accuracy for a complete tomography.

If changing between measurements requires little time or effort, as is the case when using fast, automated waveplates, then the primary consideration will be the total number of state copies necessary to run an accurate tomography. If, however, changing bases is time-intensive or error-prone, such as when moving waveplates by hand, it may be desirable to minimize the number of measurement settings.

As discussed in Section 2, minimizing the number of measurements will result in 4^n distinct measurement settings for n detectors, or 3^n settings for $2n$ detectors. In order to minimize state copies necessary, even for n detectors, every measurement should be accompanied by its complementary orthogonal measurements, in effect simulating a $2n$ detector measurement with only n detectors. For example, when measuring a two-qubit system, one should use 9 measurement settings (with 4 detectors) or 36 measurement settings (with only 2 detectors).

The reason that making more measurements can be so beneficial is simple: in order to transform the measured counts into probabilities, complete bases are necessary. If no complete bases are measured, then there is no scaling information about the rate at which state copies are being measured. Taking this information redundantly within every measured basis very precisely determines the state intensity, *allowing every other measurement to become more accurate*. Alternatively, this can function as a detector of systematic errors, allowing the experimenter to notice when the sum of the measured counts in one basis differs from the sum in another basis, e.g., due to a drift in the production rate of the source.

5.2. How Many Counts per Measurement?

In order to quantify the conclusions of the previous section, Monte Carlo simulations were used to estimate the number of state copies per measurement that would be necessary to achieve an average of 99% fidelity (between the tomography result and the "unknown" state) for a variety of two-qubit states, using both 2- and 4-detector measurements. Keep in mind that the 4-detector results, each of which require nine measurements, could be achieved by using 36 measurements with two detectors, which in many cases is still superior to the 16-measurement case. (For example, the Bell state requires 150 counts per measurement setting for 16 measurements on two detectors—a total of 150×16 counts, while at the same time 36 measurements on two detectors requires a total of 50×36 counts—a factor of two improvement.) Table I shows the state copies *per measurement* necessary to achieve a 99% fidelity for five families of states.

- The Werner state

$$\hat{\rho}_1(\lambda) = (1-\lambda)\hat{\rho}\frac{1}{\sqrt{2}}\big(|HH\rangle + |VV\rangle\big) + \lambda\frac{I}{4},$$

Table I
Counts per measurement setting necessary to achieve, on average, a 99% fidelity state. To calculate the total counts necessary for a 2- (4-)detector tomography, multiply the listed number by 16 (9).

	λ	**0**	**0.01**	**0.1**	**0.25**	**0.5**	**1.0**
$\hat{\rho}_1$	2 det.	150	900	1300	4900	2400	1600
	4 det.	50	150	1200	600	350	350
$\hat{\rho}_2$	2 det.	150	200	250	300	350	350
	4 det.	50	100	150	150	150	150
$\hat{\rho}_3$	2 det.	75	200	350	350	300	350
	4 det.	50	100	150	150	150	150
$\hat{\rho}_4$	2 det.	150	1100	1600	1500	1000	700
	4 det.	50	100	300	500	250	200
$\hat{\rho}_5$	2 det.	75	300	6800	3500	2100	1600
	4 det.	50	100	1000	450	350	350

- The Bell state $\longrightarrow$ The half-mixed state

$$\hat{\rho}_2(\lambda) = \frac{1}{2}\big(|HH\rangle\langle HH| + |VV\rangle\langle VV|\big) + \frac{1-\lambda}{2}\big(|HH\rangle\langle VV| + |VV\rangle\langle HH|\big),$$

- HH $\longrightarrow$ The half-mixed state

$$\hat{\rho}_3(\lambda) = \left(1 - \frac{\lambda}{2}\right)|HH\rangle\langle HH| + \frac{\lambda}{2}|VV\rangle\langle VV|,$$

- Maximally Entangled Mixed States (MEMS)

$$\hat{\rho}_4\left(\lambda < \frac{1}{3}\right) = \left(\frac{1-\lambda}{2}\right)\big(|HH\rangle + |VV\rangle\big)\big(\langle HH| + \langle VV|\big) + \lambda|HV\rangle\langle HV|$$

$$\hat{\rho}_4\left(\lambda \geq \frac{1}{3}\right) = \left(\frac{1-\lambda}{2}\right)\big(|HH\rangle\langle VV| + |VV\rangle\langle HH|\big) + \frac{1}{3}\big(|HH\rangle\langle HH| + |HV\rangle\langle HV| + |VV\rangle\langle VV|\big),$$

- HH $\longrightarrow$ The totally mixed state

$$\hat{\rho}_5(\lambda) = (1-\lambda)|HH\rangle\langle HH| + \lambda\frac{I}{4}.$$

In order to understand these numbers, consider just the Werner state's behavior as it transitions from a maximally entangled state ($\lambda = 0$) to a totally mixed state ($\lambda = 1$). Figure 8 shows the counts necessary to achieve either a 99% (dotted line) or 99.9% (solid line) fidelity for the full range of λ.

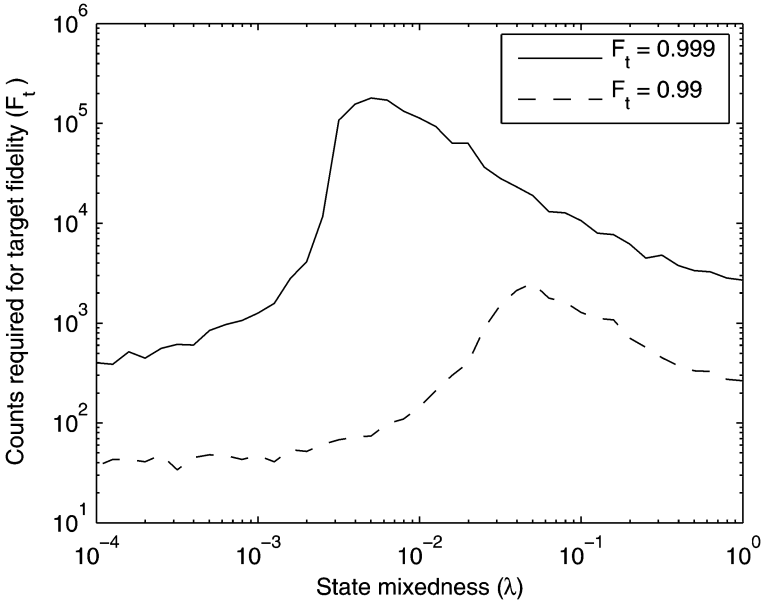

FIG. 8. This plot shows the counts per measurement for a 4-detector system using the minimum nine measurement settings to achieve either a 99% (dotted line) or a 99.9% (solid line) target fidelity (F_t) with the ideal state. These results used Monte Carlo simulations on the Werner state $(1-\lambda)\hat{\rho}_{1/\sqrt{2}|HH+VV\rangle} + \lambda I/4$, for a range of values of λ.

For very low λ (very pure states), almost no state copies are necessary. When measuring these states, several measurements return a value close to zero counts. This value of zero probability guarantees that the state will be near the border of allowable states (e.g., near the surface of the Poincaré sphere for one-qubit states). This one measurement, because it isolates the allowable states to a very small area, can very quickly lead to a high fidelity. For example, if after measuring 100 copies in the H–V basis, one has recorded 100 H counts and 0 V counts, then there is a *very* high probability that the state $|H\rangle$ is the target state, because every other axis in Hilbert space has also been isolated by this measurement (in this specific example, the results of measurements in the D–A and R–L basis—50 counts in each—can be inferred from this one H–V measurement).

As the state becomes more mixed, it moves away from the border in Hilbert space and now many counts are necessary to isolate the position of the state in those other directions. Continuing the example above, if our first measurement was instead 90 H and 10 V, we cannot infer the results of a D–A measurement. If, for instance, it results in 50 D and 50 A, these counts have a great deal more uncertainty than a 90/10 split, due to the nature of the fitness function. This accounts

for the sharp transition in Fig. 8, a transition which moves closer to the pure states for a higher fidelity cutoff, as expected.

The last behavior shown in the graph, the gradual decrease in necessary counts for increasing mixedness, occurs because fidelity becomes less sensitive for states of greater mixedness (Peters et al., 2004).

6. Error Analysis

Error analysis of reconstructed density matrices is in practice a nontrivial process. The traditional method of error analysis involves analytically solving for the error in each measurement due to each source of error, then propagating these errors through a calculation of any derived quantity. In the photon case, for example, errors in counting statistics and waveplate settings were analyzed in some detail in reference (James et al., 2001), giving errors in both density matrices and commonly derived quantities, such as the tangle and the linear entropy. In practice, however, these errors appear to be too large: We have experimentally repeated some of our measurements many times, and observed a spread in the value of derived quantities which is approximately an order of magnitude smaller that the spread predicted from an analytic calculation of the uncertainty. Obviously, the correctness of the analytic calculation is questionable. Thus it is worthwhile to discuss alternate methods of error analysis.

One promising numerical method is the 'Monte Carlo' technique, whereby additional numerically simulated data is used to provide a statistical distribution over any derived quantity. Once an error distribution is understood over a single measurement (e.g., Gaussian for waveplate setting errors or Poissonian over count statistics), a set of 'simulated' results can be generated. These results are simulated using the known error distributions in such a way as to produce a full set of numerically generated data which could feasibly have come from the same system. These data are numerically generated (at the measured counts level), and each set is used to calculate a density matrix via the maximum likelihood technique. This set of density matrices is then used to calculate the standard error on any quantity implicit in or derived from the density matrix.

As an example, consider the application of the Monte Carlo technique to the downconversion results from Example 15. Two polarization-encoded qubits are generated within ensembles that obey Poissonian statistics, and these ensembles are used to generate a density matrix using the maximum likelihood technique. In order to find the error on a quantity derived from this density matrix (e.g., the tangle), 36 new measurement results are numerically generated, each drawn randomly from a Poissonian distribution with mean equal to the original number of counts. These 36 numerically generated results are then fed into the maximum likelihood technique, in order to generate a new density matrix, from which, e.g., the tangle may be calculated. This process is repeated many times, generating

many density matrices and a distribution of tangle values, from which the error in the initial tangle may be determined. In practice, additional sets of simulated data must be generated until the error on the quantity of interest converges to a single value. For the data in Examples 14 and 15, a total of 100 simulations were used.

7. A Complete Example of Tomography

In order to demonstrate how all of the concepts presented in this chapter are actually applied, we have included an example which from start to finish uses laboratory parameters and data, taken from a two-qubit entangled photon source. Throughout this example we will use our usual convention for the canonical basis: $|0\rangle \equiv |HH\rangle$, $|1\rangle \equiv |HV\rangle$, $|2\rangle \equiv |VH\rangle$, and $|3\rangle \equiv |VV\rangle$.

EXAMPLE 15 (*A complete two-qubit tomography*). Before collecting tomography data, there are several measurement parameters that must be measured. After experimentally determining that each of our beamsplitters has negligible absorption, a 0.8% chance to reflect $|H\rangle$, and a 0.5% chance to transmit $|V\rangle$, we can determine that

$$C_{r'\to r} \equiv \begin{array}{c} \\ \to 0 \\ \to 1 \\ \to 2 \\ \to 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0' & 1' & 2' & 3' \end{array} \\ \begin{pmatrix} 0.9842 & 00049 & 0.0049 & 0.0000 \\ 0.0079 & 0.9871 & 0.0000 & 0.0050 \\ 0.0079 & 0.0000 & 0.9871 & 0.0050 \\ 0.0001 & 0.0079 & 0.0079 & 0.9901 \end{pmatrix} \end{array}. \tag{115}$$

Rather than measuring intensity fluctuations by picking off a part of the pump laser, we will choose during this tomography to fit the intensity parameters I_ν as part of the maximum likelihood technique (we use four detectors, which will allow us to fit a relative intensity to each measurement setting by using the measured counts from each of four orthogonal projectors).

Because this particular tomography will use a total of nine measurement settings (the minimum number required), there will not be enough information to fit for the detector-pair efficiencies. A previous tomography (using 36 measurement settings and not shown here) was used to solve for the E_r, using a two-detector tomography applied to the 36 measurement results from *each* of the four pairs of detectors:

$$E_1 = 0.9998, \qquad E_3 = 0.9195,$$
$$E_2 = 1.0146, \qquad E_4 = 0.9265.$$

To simplify the example, we will make all measurements in the canonical bases (this could be accomplished using either ideal waveplates or, in some cases, imperfect waveplates—see Section 3.1.2).

With these parameters recorded, we can now take the data. The following counts were recorded for a slightly mixed Bell state (close to $\frac{1}{\sqrt{2}}(|HH\rangle+\mathrm{i}|VV\rangle)$):

$$
\begin{array}{llll}
n_{1,r}\colon HH = 3708 & HV = 77 & VH = 51 & VV = 3642 \\
n_{2,r}\colon HD = 1791 & HA = 1987 & VD = 2096 & VA = 3642 \\
n_{3,r}\colon HR = 2048 & HL = 1854 & VR = 1926 & VL = 1892 \\
n_{4,r}\colon DH = 1766 & DV = 1914 & AH = 2153 & AV = 1741 \\
n_{5,r}\colon DD = 1713 & DA = 1945 & AD = 2208 & AA = 1647 \\
n_{6,r}\colon DR = 3729 & DL = 91 & AR = 102 & AL = 3662 \\
n_{7,r}\colon RH = 2017 & RV = 1709 & LH = 1917 & LV = 1955 \\
n_{8,r}\colon RD = 3686 & RA = 102 & LD = 109 & LA = 3651 \\
n_{9,r}\colon RR = 2404 & RL = 1474 & LR = 1712 & LL = 2209,
\end{array}
$$

with the corresponding accidental counts (calculated using the measured singles rates and the previously determined coincidence window Δt_r (c.f., Eq. (87)).

$$
\begin{array}{llll}
n_{1,1}^{\mathrm{accid}} = 5.4 & n_{1,2}^{\mathrm{accid}} = 5.6 & n_{1,3}^{\mathrm{accid}} = 5.9 & n_{1,4}^{\mathrm{accid}} = 6.0 \\
n_{2,1}^{\mathrm{accid}} = 5.2 & n_{2,2}^{\mathrm{accid}} = 5.5 & n_{2,3}^{\mathrm{accid}} = 5.6 & n_{2,4}^{\mathrm{accid}} = 6.0 \\
n_{3,1}^{\mathrm{accid}} = 5.3 & n_{3,2}^{\mathrm{accid}} = 5.5 & n_{3,3}^{\mathrm{accid}} = 5.6 & n_{3,4}^{\mathrm{accid}} = 5.9 \\
n_{4,1}^{\mathrm{accid}} = 5.2 & n_{4,2}^{\mathrm{accid}} = 5.3 & n_{4,3}^{\mathrm{accid}} = 6.0 & n_{4,4}^{\mathrm{accid}} = 6.1 \\
n_{5,1}^{\mathrm{accid}} = 5.2 & n_{5,2}^{\mathrm{accid}} = 5.4 & n_{5,3}^{\mathrm{accid}} = 5.9 & n_{5,4}^{\mathrm{accid}} = 6.2 \\
n_{6,1}^{\mathrm{accid}} = 5.2 & n_{6,2}^{\mathrm{accid}} = 5.3 & n_{6,3}^{\mathrm{accid}} = 5.9 & n_{6,4}^{\mathrm{accid}} = 6.1 \\
n_{7,1}^{\mathrm{accid}} = 5.3 & n_{7,2}^{\mathrm{accid}} = 5.4 & n_{7,3}^{\mathrm{accid}} = 5.9 & n_{7,4}^{\mathrm{accid}} = 6.1 \\
n_{8,1}^{\mathrm{accid}} = 5.4 & n_{8,2}^{\mathrm{accid}} = 5.9 & n_{8,3}^{\mathrm{accid}} = 6.1 & n_{8,4}^{\mathrm{accid}} = 6.6 \\
n_{9,1}^{\mathrm{accid}} = 5.3 & n_{9,2}^{\mathrm{accid}} = 5.4 & n_{9,3}^{\mathrm{accid}} = 6.0 & n_{9,4}^{\mathrm{accid}} = 6.2.
\end{array}
$$

After minimizing the likelihood function, we obtain the following $\widehat{T}$ matrix

$$
\widehat{T} = \begin{pmatrix}
0 & 0 & 0 & 0 \\
2.401 + 3.167\mathrm{i} & 2.372 & 0 & 0 \\
-6.381 - 2.649\mathrm{i} & 3.919 - 0.897\mathrm{i} & 2.674 & 0 \\
-8.975 + 58.630\mathrm{i} & 1.356 - 2.106\mathrm{i} & 1.685 - 1.514\mathrm{i} & 60.08
\end{pmatrix}, \tag{116}
$$

from which we can derive the density matrix,

$$
\begin{aligned}
\hat{\rho} &= \frac{\widehat{T}^{\dagger}\widehat{T}}{\mathrm{Tr}\{\widehat{T}^{\dagger}\widehat{T}\}} \\
&= \begin{pmatrix}
0.50 & -0.02 - 0.01\mathrm{i} & -0.02 - 0.01\mathrm{i} & -0.07 - 0.49\mathrm{i} \\
-0.02 + 0.01\mathrm{i} & 0.00 & 0.00 + 0.00\mathrm{i} & 0.01 + 0.02\mathrm{i} \\
-0.02 + 0.01\mathrm{i} & 0.00 - 0.01\mathrm{i} & 0.00 & 0.01 + 0.01\mathrm{i} \\
-0.07 + 0.49\mathrm{i} & 0.01 - 0.02\mathrm{i} & 0.01 - 0.01\mathrm{i} & 0.50
\end{pmatrix}.
\end{aligned} \tag{117}
$$

Our search algorithm returned this density matrix because it minimized not only the main search parameters $\vec{t}$, but the intensities I_ν:

$$\begin{array}{lll} I_1 = 7647 & I_2 = 7745 & I_3 = 7879 \\ I_4 = 7725 & I_5 = 7669 & I_6 = 7754 \\ I_7 = 7751 & I_8 = 7716 & I_9 = 7967, \end{array}$$

allowing us to calculate the expected counts $\bar{n}_{\nu,r}$ (for the final density matrix):

$$\begin{array}{llll} \bar{n}_{1,1} = 3792 & \bar{n}_{1,2} = 81 & \bar{n}_{1,3} = 67 & \bar{n}_{1,4} = 3544 \\ \bar{n}_{2,1} = 1794 & \bar{n}_{2,2} = 1956 & \bar{n}_{2,3} = 2106 & \bar{n}_{2,4} = 1735 \\ \bar{n}_{3,1} = 2046 & \bar{n}_{3,2} = 1787 & \bar{n}_{3,3} = 1933 & \bar{n}_{3,4} = 1956 \\ \bar{n}_{4,1} = 1815 & \bar{n}_{4,2} = 1895 & \bar{n}_{4,3} = 2108 & \bar{n}_{4,4} = 1758 \\ \bar{n}_{5,1} = 1618 & \bar{n}_{5,2} = 2050 & \bar{n}_{5,3} = 2247 & \bar{n}_{5,4} = 1604 \\ \bar{n}_{6,1} = 3792 & \bar{n}_{6,2} = 97 & \bar{n}_{6,3} = 103 & \bar{n}_{6,4} = 3594 \\ \bar{n}_{7,1} = 2032 & \bar{n}_{7,2} = 1699 & \bar{n}_{7,3} = 1901 & \bar{n}_{7,4} = 1966 \\ \bar{n}_{8,1} = 3758 & \bar{n}_{8,2} = 103 & \bar{n}_{8,3} = 105 & \bar{n}_{8,4} = 3583 \\ \bar{n}_{9,1} = 2271 & \bar{n}_{9,2} = 1580 & \bar{n}_{9,3} = 1751 & \bar{n}_{9,4} = 2206. \end{array}$$

Using the error analysis techniques presented in Section 6, we can estimate this state's fidelity with the Bell state $\frac{1}{\sqrt{2}}(|HH\rangle + \mathrm{i}|VV\rangle)$ to be $98.4 \pm 0.2\%$.

8. Outlook

This chapter represents our best efforts to date to experimentally optimize quantum tomography of discrete systems, but fails to address several key areas that future research will need to include: (1) As discussed earlier, tomographic error analysis is still in the nascent stages of development, and to date the only acceptable error estimates have come from Monte Carlo simulations. Analytic solutions to the problem of error estimation could greatly speed up this computationally intensive task. (2) The study of *adaptive* tomography has motivated a number of the improvements presented in this chapter, notably the use of more measurement settings to achieve a more efficient and accurate tomography in less time. A general theory of how to adapt measurement settings and measurement times based on the data that has already been collected—which can be experimentally applied in real time—has yet to be fully realized (see, however, D'Ariano et al. (2004)). (3) The number of measurements necessary to perform tomography grows exponentially with the number of qubits; it will eventually be necessary to partially characterize states using fewer measurements. This will be particularly important for error—as the analysis of large systems takes more and more time, it will become less and less feasible to use Monte Carlo simulations to estimate the error. (4) Each distinct qubit implementation provides a unique challenges, which will need to be explored by the experimental groups specializing in those systems; hopefully,

the study of each system's differences will illuminate new areas for tomographic improvement.

9. Acknowledgements

We would like to thank Daniel James and Andrew White for their assistance in the development of the theory of tomography and for helpful discussions throughout the preparation of this manuscript. We would like to thank Rob Thew for helpful conversations concerning qudit tomography. This work was supported by the National Science Foundation (Grant No. EIA-0121568) and the MURI Center for Photonic Quantum Information Systems (ARO/ARDA program DAAD19-03-1-0199).

10. References

Abouraddy, A.F., Sergienko, A.V., Saleh, B.E.A., Teich, M.C. (2002). Quantum entanglement and the two-photon Stokes parameters. *Opt. Commun.* **201**, 93.

Allen, L., Barnett, S.M., Padgett, M.J. (2004). "Optical Angular Momentum". Institute of Physics Publishing, London.

Altepeter, J.B., James, D.F.V., Kwiat, P.G. (2004). "Quantum State Estimation". *Lecture Notes in Phys*. Springer, Berlin.

Altepeter, J.B., Jeffrey, E.R., Kwiat, P.G., Tanzilli, S., Gisin, N., Acin, A. (2005). Experimental methods for detecting entanglement. Submitted to *Phys. Rev. Lett*.

Arnaut, H.H., Barbosa, G.A. (2000). Orbital and intrinsic angular momentum of single photons and entangled pairs of photons generated by parametric down-conversion. *Phys. Rev. Lett.* **85**, 286–289.

Bell, J.S. (1964). On the Einstein–Podolsky–Rosen paradox. *Physics* **1**, 195–200.

Born, M., Wolf, E. (1987). "Principles of Optics". Pergamon Press, Oxford, UK.

Coffman, V., Kundu, J., Wootters, W.K. (2000). Distributed entanglement. *Phys. Rev. A* **61**, 052306.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C. (2001). "Introduction to Algorithms", 2nd Edn. MIT Press, London, UK.

Cory, D.G., Fahmy, A.F., Havel, T.F. (1997). Ensemble quantum computing by NMR spectroscopy. *Proc. Nat. Acad. Sci. USA* **94**, 1634.

D'Ariano, G.M., Paris, M.G.A., Sacchi, M.F. (2004). "Advances in Imaging and Electron Physics", vol. 128. Academic Press, New York.

Gasiorowitz, S. (1996). "Quantum Physics". Wiley, New York.

Giorgi, G., Nepi, G.D., Mataloni, P., Martini, F.D. (2003). A high brightness parametric source of entangled photon states. *Laser Phys.* **13**, 350.

Hong, C.K., Mandel, L. (1986). Experimental realization of a localized one-photon state. *Phys. Rev. Lett.* **56**, 58.

Hradil, Z., Rehacek, J. (2001). Efficiency of maximum-likelihood reconstruction of quantum states. *Fortschr. Phys.* **49**, 1083.

James, D.F.V., Kwiat, P.G., Munro, W.J., White, A.G. (2001). Measurement of qubits. *Phys. Rev. A* **64**, 052312.

Jones, J.A., Mosca, M., Hansen, R.H. (1997). Implementation of a quantum search algorithm on a quantum computer. *Nature* **392**, 344.

Jozsa, R. (1994). Fidelity for mixed quantum states. *J. Modern Opt.* **41**, 12.

Kielpinski, D., Meyer, V., Rowe, M.A., Sackett, C.A., Itano, W.M., Monroe, C., Wineland, D.J. (2001). A decoherence-free quantum memory using trapped ions. *Science* **291**, 1013.

Kwiat, P.G., Englert, B.-G. (2004). "Science and Ultimate Reality: Quantum Theory, Cosmology and Complexity". Cambridge Univ. Press, Cambridge, UK.

Kwiat, P.G., Waks, E., White, A.G., Appelbaum, I., Eberhard, P.H. (1999). Ultrabright source of polarization-entangled photons. *Phys. Rev. A* **60**, R773.

Laflamme, R., Knill, E., Cory, D., Fortunato, E., Havel, T., Miquel, C., Martinez, R., Negrevergne, C., Ortiz, G., Pravia, M., Sharf, Y., Sinha, S., Somma, R., Viola, L. (2002). Introduction to NMR quantum information processing. quant-ph/0207172.

Langford, N.K., Dalton, R.B., Harvey, M.D., O'Brien, J.L., Pryde, G.J., Gilchrist, A., Bartlett, S.D., White, A.G. (2004). Measuring entangled qutrits and their use for quantum bit commitment. *Phys. Rev. Lett.* **93**, 053601.

Lawrence, J., Brukner, C., Zeilinger, A. (2002). Mutually unbiased binary observable sets on *n* qubits. *Phys. Rev. A* **65**, 032320.

Lehner, J., Leonhardt, U., Paul, H. (1996). Unpolarized light: Classical and quantum states. *Phys. Rev. A* **53**, 2727.

Leonhardt, U. (Ed.) (1997). "Measuring the Quantum State of Light". Cambridge Univ. Press, Cambridge, UK.

Lewenstein, M., Kraus, B., Cirac, J.I., Horodecki, P. (2000). Optimization of entanglement witnesses. *Phys. Rev. A* **62**, 052310.

Mair, A., Vaziri, A., Weihs, G., Zeilinger, A. (2001). Entanglement of the orbital angular momentum states of photons. *Nature* **412**, 313.

Marcikic, I., de Riedmatten, H., Tittel, W., Zbinden, H., Gisin, N. (2003). Long-distance teleportation of qubits at telecommunication wavelengths. *Nature* **421**, 509.

Monroe, C. (2002). Quantum information processing with atoms and photons. *Nature* **416**, 238.

Munro, W.J., James, D.F.V., White, A.G., Kwiat, P.G. (2001). Maximizing the entanglement of two mixed qubits. *Phys. Rev. A* **64**, 030302.

Nambu, Y., Usami, K., Tsuda, Y., Matsumoto, K., Nakamura, K. (2002). Generation of polarization-entangled photon pairs in a cascade of two type-*i* crystals pumped by femtosecond pulses. *Phys. Rev. A* **66**, 033816.

Nielsen, M.A., Chuang, I.L. (2000). "Quantum Computation and Quantum Information". Cambridge Univ. Press, Cambridge, UK.

O'Brien, J.L., Pryde, G.J., White, A.G., Ralph, T.C., Branning, D. (2003). Demonstration of an all-optical quantum controlled-not gate. *Nature* **426**, 264.

Peters, N., Altepeter, J.B., Jeffrey, E.R., Branning, D., Kwiat, P.G. (2003). Precise creation, characterization and manipulation of single optical qubits. *Quantum Information and Computation* **3**, 503.

Peters, N., Wei, T.-C., Kwiat, P.G. (2004). Mixed-state sensitivity of several quantum-information benchmarks. *Phys. Rev. A* **70**, 052309.

Peters, N.A., Barreiro, J.T., Goggin, M.E., Wei, T.-C., Kwiat, P.G. (2005). Remote state preparation: Arbitrary remote control of photon polarization. To appear in *Phys. Rev. Lett*.

Pittman, T.B., Fitch, M.J., Jacobs, B.C., Franson, J.D. (2003). Experimental controlled-not logic gate for single photons in the coincidence basis. *Phys. Rev. A* **68**, 032316.

Sanaka, K., Kawahara, K., Kuga, T. (2001). New high-efficiency source of photon pairs for engineering quantum entanglement. *Phys. Rev. Lett.* **86**, 5620.

Sancho, J.M.G., Huelga, S.F. (2000). Measuring the entanglement of bipartite pure states. *Phys. Rev. A* **61**, 042303.

Schmidt-Kaler, F., Häffner, H., Riebe, M., Gulde, S., Lancaster, G.P.T., Deuschle, T., Becher, C., Roos, C.F., Eschner, J., Blatt, R. (2003). Realization of the ciraczoller controlled-not quantum gate. *Nature* **422**, 408.

Sergienko, A.V., Giuseppe, G.D., Atatüre, M., Saleh, B.E.A., Teich, M.C. (2003). Entangled-photon state engineering. In: Shapiro, J.H., Hirota, O. (Eds.), "Proceedings of the Sixth International Conference on Quantum Communication, Measurement and Computing (QCMC)", Rinton, Princeton, p. 147.

Stokes, G.C. (1852). On the composition and resolution of streams of polarized light from different sources. *Trans. Cambridge Phil. Soc.* **9**, 399.

Terhal, B.M. (2001). Detecting quantum entanglement. quant-ph/0101032.

Thew, R.T., Nemoto, K., White, A.G., Munro, W.J. (2002). Qudit quantum-state tomography. *Phys. Rev. A* **66**, 012303.

Weinstein, Y.S., Pravia, M.A., Fortunato, E.M., Lloyd, S., Cory, D.G. (2001). Implementation of the quantum Fourier transform. *Phys. Rev. Lett.* **86**, 1889.

Werner, R. (1989). Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40**, 4277.

White, A.G., James, D.F.V., Eberhard, P.H., Kwiat, P.G. (1999). Nonmaximally entangled states: Production, characterization, and utilization. *Phys. Rev. Lett.* **83**, 3103.

Wooters, W.K. (1998). Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245.

Yamamoto, T., Koashi, M., Özdemir, S.K., Imoto, N. (2003). Experimental extraction of an entangled photon pair from two identically decohered pairs. *Nature* **421**, 343.

Zyczkowski, K. (2000). Geometry of entangled states. quant-ph/0006068.

Zyczkowski, K. (2001). cp^n, or, entanglement illustrated. quant-ph/0108064.