Measurement of qubits

Daniel F. V. James,[1,*] Paul G. Kwiat,[2,3] William J. Munro,[4,5] and Andrew G. White[2,4]
[1]*Theoretical Division T-4, Los Alamos National Laboratory, Los Alamos, New Mexico 87545*
[2]*Physics Division P-23, Los Alamos National Laboratory, Los Alamos, New Mexico 87545*
[3]*Department of Physics, University of Illinois, Urbana-Champaign, Illinois 61801*
[4]*Department of Physics, University of Queensland, Brisbane, Queensland 4072, Australia*
[5]*Hewlett-Packard Laboratories, Filton Road, Stoke Gifford, Bristol BS34 8QZ, United Kingdom*
(Received 20 March 2001; published 16 October 2001)

We describe in detail the theory underpinning the measurement of density matrices of a pair of quantum two-level systems (''qubits''). Our particular emphasis is on qubits realized by the two polarization degrees of freedom of a pair of entangled photons generated in a down-conversion experiment; however, the discussion applies in general, regardless of the actual physical realization. Two techniques are discussed, namely, a tomographic reconstruction (in which the density matrix is linearly related to a set of measured quantities) and a maximum likelihood technique which requires numerical optimization (but has the advantage of producing density matrices that are always non-negative definite). In addition, a detailed error analysis is presented, allowing errors in quantities derived from the density matrix, such as the entropy or entanglement of formation, to be estimated. Examples based on down-conversion experiments are used to illustrate our results.

DOI: 10.1103/PhysRevA.64.052312 PACS number(s): 03.67.−a, 42.50.−p

I. INTRODUCTION

The ability to create, manipulate, and characterize quantum states is becoming an increasingly important area of physical research, with implications for areas of technology such as quantum computing, quantum cryptography, and communications. With a series of measurements on a large enough number of identically prepared copies of a quantum system, one can infer, to a reasonable approximation, the quantum state of the system. Arguably, the first such experimental technique for determining the state of a quantum system was devised by George Stokes in 1852 [1]. His famous four parameters allow an experimenter to determine uniquely the polarization state of a light beam. With the insight provided by nearly 150 years of progress in optical physics, we can consider coherent light beams to be an ensemble of two-level quantum mechanical systems, the two levels being the two polarization degrees of freedom of the photons; the Stokes parameters allow one to determine the density matrix describing this ensemble. More recently, experimental techniques for the measurement of the more subtle quantum properties of light have been the subject of intensive investigation (see Ref. [2] for a comprehensive and erudite exposition of this subject). In various experimental circumstances it has been found reasonably straightforward to devise a simple linear tomographic technique in which the density matrix (or Wigner function) of a quantum state is found from a linear transformation of experimental data. However, there is one important drawback to this method, in that the recovered state might not correspond to a physical state because of experimental noise. For example, density matrices for any quantum state must be Hermitian, positive semidefinite matrices with unit trace. The tomographically measured matrices often fail to be positive semidefinite, especially when measuring low-entropy states. To avoid this problem the ''maximum likelihood'' tomographic approach to the estimation of quantum states has been developed [3–7]. In this approach the density matrix that is ''mostly likely'' to have produced a measured data set is determined by numerical optimization.

In the past decade several groups have successfully employed tomographic techniques for the measurement of quantum mechanical systems. In 1990 Ashburn *et al.* reported the measurement of the density matrix for the nine sublevels of the $n=3$ level of hydrogen atoms formed following collision between H^+ ions and He atoms, in conditions of high symmetry which simplified the tomographic problem [8]. Since then, in 1993 Smithey *et al.* made a homodyne measurement of the Wigner function of a single mode of light [9]. Other explorations of the quantum states of single mode light fields have been made by Breitenbach *et al.* [10] and Wu *et al.* [11]. Other quantum systems whose density matrices have been investigated experimentally include the vibrations of molecules [12], the motion of ions and atoms [13,14], and the internal angular momentum quantum state of the $F=4$ ground state of a cesium atom [15]. The quantum states of multiple spin-$\frac{1}{2}$ nuclei have been measured in the high-temperature regime using NMR techniques [16], albeit in systems of such high entropy that the creation of entangled states is necessarily precluded [17]. The measurement of the quantum state of entangled qubit pairs, realized using the polarization degrees of freedom of a pair of photons created in a parametric down-conversion experiment, was reported by us recently [18].

In this paper we will examine in detail techniques for quantum state measurement as it applies to multiple correlated two-level quantum mechanical systems (or ''qubits'' in the terminology of quantum information). Our particular em-

*Corresponding author. Mailing address: Mail stop B-283, Los Alamos National Laboratory, Los Alamos NM 87545. FAX: (505) 667-1931. Email address: dfvj@lanl.gov

phasis is qubits realized via the two polarization degrees of freedom of photons, data from which we use to illustrate our results. However, these techniques are readily applicable to other technologies proposed for creating entangled states of pairs of two-level systems. Because of the central importance of qubit systems to the emergent discipline of quantum computation, a thorough explanation of the techniques needed to characterize the qubit states will be of relevance to workers in the various diverse experimental fields currently under consideration for quantum computation technology [19]. This paper is organized as follows. In Sec. II we explore the analogy with the Stokes parameters, and how they lead naturally to a scheme for measurement of an arbitrary number of two-level systems. In Sec. III, we discuss the measurement of a pair of qubits in more detail, presenting the validity condition for an arbitrary measurement scheme and introducing the set of 16 measurements employed in our experiments. Sec. IV deals with our method for maximum likelihood reconstruction and in Sec. V we demonstrate how to calculate the errors in such measurements, and how these errors propagate to quantities calculated from the density matrix.

II. THE STOKES PARAMETERS AND QUANTUM STATE TOMOGRAPHY

As mentioned above, there is a direct analogy between the measurement of the polarization state of a light beam and the measurement of the density matrix of an ensemble of two-level quantum mechanical systems. Here we explore this analogy in more detail.

A. Single qubit tomography

The Stokes parameters are defined from a set of four intensity measurements [20]: (i) with a filter that transmits 50% of the incident radiation, regardless of its polarization; (ii) with a polarizer that transmits only horizontally polarized light; (iii) with a polarizer that transmits only light polarized at 45° to the horizontal; and (iv) with a polarizer that transmits only right-circularly polarized light. The number of photons counted by a detector, which is proportional to the classical intensity, in these four experiments is as follows:

$$
\begin{aligned}
n_0 &= \frac{\mathcal{N}}{2}(\langle H|\hat{\rho}|H\rangle + \langle V|\hat{\rho}|V\rangle) = \frac{\mathcal{N}}{2}(\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle),\\
n_1 &= \mathcal{N}(\langle H|\hat{\rho}|H\rangle)\\
&= \frac{\mathcal{N}}{2}(\langle R|\hat{\rho}|R\rangle + \langle R|\hat{\rho}|L\rangle + \langle L|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle),\\
n_2 &= \mathcal{N}(\langle \overline{D}|\hat{\rho}|\overline{D}\rangle)\\
&= \frac{\mathcal{N}}{2}(\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle - i\langle L|\hat{\rho}|R\rangle + i\langle R|\hat{\rho}|L\rangle),\\
n_3 &= \mathcal{N}(\langle R|\hat{\rho}|R\rangle).
\end{aligned}
\tag{2.1}
$$

Here $|H\rangle$, $|V\rangle$, $|\overline{D}\rangle = (|H\rangle - |V\rangle)/\sqrt{2} = \exp(i\pi/4)(|R\rangle + i|L\rangle)/\sqrt{2}$, and $|R\rangle = (|H\rangle - i|V\rangle)/\sqrt{2}$ are the kets representing qubits polarized in the linear horizontal, linear vertical, linear diagonal (45°), and right-circular senses respectively, $\hat{\rho}$ is the (2×2) density matrix for the polarization degrees of the light (or for a two-level quantum system), and $\mathcal{N}$ is a constant dependent on the detector efficiency and light intensity. The *Stokes parameters*, which fully characterize the polarization state of the light, are then defined by

$$
\begin{aligned}
\mathcal{S}_0 &\equiv 2n_0 = \mathcal{N}(\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle),\\
\mathcal{S}_1 &\equiv 2(n_1 - n_0) = \mathcal{N}(\langle R|\hat{\rho}|L\rangle + \langle L|\hat{\rho}|R\rangle),\\
\mathcal{S}_2 &\equiv 2(n_2 - n_0) = \mathcal{N}i(\langle R|\hat{\rho}|L\rangle - \langle L|\hat{\rho}|R\rangle),\\
\mathcal{S}_3 &\equiv 2(n_3 - n_0) = \mathcal{N}(\langle R|\hat{\rho}|R\rangle - \langle L|\hat{\rho}|L\rangle).
\end{aligned}
\tag{2.2}
$$

We can now relate the Stokes parameters to the density matrix $\hat{\rho}$ by the formula

$$
\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3}\frac{\mathcal{S}_i}{\mathcal{S}_0}\hat{\sigma}_i, \tag{2.3}
$$

where $\hat{\sigma}_0 = |R\rangle\langle R| + |L\rangle\langle L|$ is the single qubit identity operator and $\hat{\sigma}_1 = |R\rangle\langle L| + |L\rangle\langle R|$, $\hat{\sigma}_2 = i|L\rangle\langle R| - |R\rangle\langle L|$, and $\hat{\sigma}_3 = |R\rangle\langle R| - |L\rangle\langle L|$ are the Pauli spin operators. Thus the measurement of the Stokes parameters can be considered equivalent to a tomographic measurement of the density matrix of an ensemble of single qubits.

B. Multiple beam Stokes parameters: Multiple qubit tomography

The generalization of the Stokes scheme to measure the state of multiple photon beams (or multiple qubits) is reasonably straightforward. One should, however, be aware that important differences exist between the one-photon and the multiple photon cases. Single photons, at least in the current context, can be described in a purely *classical* manner, and the density matrix can be related to the purely classical concept of the coherency matrix [21]. For multiple photons one has the possibility of nonclassical correlations occurring, with quintessentially quantum mechanical phenomena such as entanglement being present. We will return to the concept of entanglement and how it may be measured later in this paper.

An *n*-qubit state is characterized by a density matrix which may be written as follows:

$$
\hat{\rho} = \frac{1}{2^n}\sum_{i_1,i_2,\ldots,i_n=0}^{3} r_{i_1,i_2,\ldots,i_n}\hat{\sigma}_{i_1}\otimes\hat{\sigma}_{i_2}\otimes\cdots\otimes\hat{\sigma}_{i_n}, \tag{2.4}
$$

where the 4^n parameters $r_{i_1,i_2,\ldots,i_n}$ are real numbers. The normalization property of the density matrices requires that $r_{0,0,\ldots,0} = 1$, and so the density matrix is specified by $4^n - 1$ real parameters. The symbol $\otimes$ represents the tensor

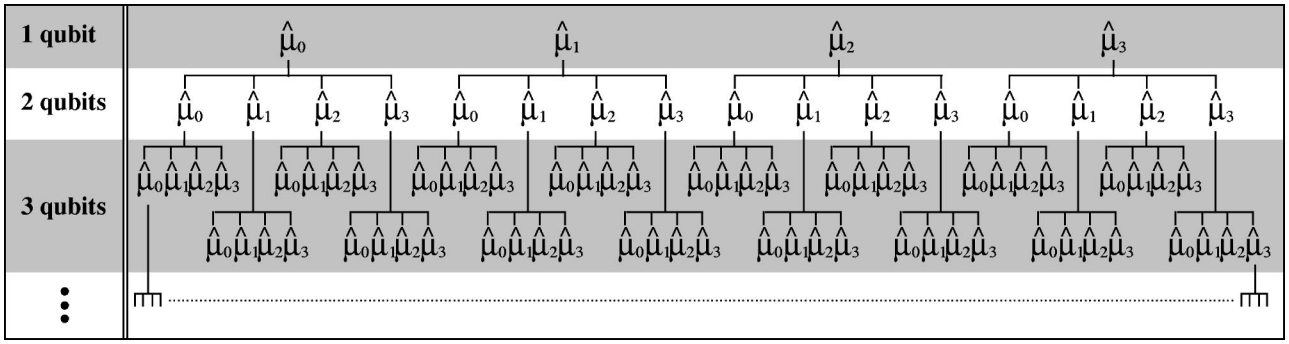

FIG. 1. Tree diagram representing number and type of measurements necessary for tomography. For a single qubit, the measurements $\{\hat{\mu}_0,\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3\}$ suffice to reconstruct the state, e.g., measurements of the horizontal, vertical, diagonal, and right-circular polarization components, (H,V,D,R). For two qubits, 16 double-coincidence measurements are necessary ($\{\hat{\mu}_0\hat{\mu}_0,\hat{\mu}_0\hat{\mu}_1,\ldots,\hat{\mu}_3\hat{\mu}_3\}$), increasing to 64 three-coincidence measurements for three qubits ($\{\hat{\mu}_0\hat{\mu}_0\hat{\mu}_0,\hat{\mu}_0\hat{\mu}_0\hat{\mu}_1,\ldots,\hat{\mu}_3\hat{\mu}_3\hat{\mu}_3\}$), and so on, as shown.

product between operators acting on the Hilbert spaces associated with the separate qubits.

As Stokes showed, the state of a single qubit can be determined by taking a set of four projection measurements which are represented by the four operators $\hat{\mu}_0=|H\rangle\langle H|+|V\rangle\langle V|$, $\hat{\mu}_1=|H\rangle\langle H|$, $\hat{\mu}_2=|\overline{D}\rangle\langle\overline{D}|$, $\hat{\mu}_3=|R\rangle\langle R|$. Similarly, the state of two qubits can be determined by the set of 16 measurements represented by the operators $\hat{\mu}_i\otimes\hat{\mu}_j$ $(i,j=0,1,2,3)$. More generally the state of an n-qubit system can be determined by 4^n measurements given by the operators $\hat{\mu}_{i_1}\otimes\hat{\mu}_{i_2}\otimes\cdots\otimes\hat{\mu}_{i_n}$ $(i_k=0,1,2,3$ and $k=1,2,\ldots,n)$. This ''tree'' structure for multiqubit measurement is illustrated in Fig. 1.

The proof of this conjecture is reasonably straightforward. The outcome of a measurement is given by the formula

$$n=\mathcal{N}\mathrm{Tr}\{\hat{\rho}\hat{\mu}\}, \tag{2.5}$$

where $\hat{\rho}$ is the density matrix, $\hat{\mu}$ is the measurement operator, and $\mathcal{N}$ is a constant of proportionality which can be determined from the data. Thus in our n-qubit case the outcomes of the various measurement are

$$n_{i_1,i_2,\ldots,i_n}=\mathcal{N}\,\mathrm{Tr}\,\{\hat{\rho}(\hat{\mu}_{i_1}\otimes\hat{\mu}_{i_2}\otimes\cdots\otimes\hat{\mu}_{i_n})\}. \tag{2.6}$$

Substituting from Eq. (2.4) we obtain

$$n_{i_1,i_2,\ldots,i_n}=\frac{\mathcal{N}}{2^n}\sum_{j_1,j_2,\ldots,j_n=0}^{3}\mathrm{Tr}\{\hat{\mu}_{i_1}\hat{\sigma}_{j_1}\}\times\mathrm{Tr}\{\hat{\mu}_{i_2}\hat{\sigma}_{j_2}\}\cdots\mathrm{Tr}\{\hat{\mu}_{i_n}\hat{\sigma}_{j_n}\}r_{i_1,i_2,\ldots,i_n}. \tag{2.7}$$

As can be easily verified, the single qubit measurement operators $\hat{\mu}_i$ are linear combinations of the Pauli operators $\hat{\sigma}_j$, i.e., $\hat{\mu}_i=\Sigma_{j=0}^{3}\Upsilon_{ij}\hat{\sigma}_j$, where Υ_{ij} are the elements of the matrix

$$\Upsilon=\begin{pmatrix}1&0&0&0\\1/2&1/2&0&0\\1/2&0&1/2&0\\1/2&0&0&1/2\end{pmatrix}. \tag{2.8}$$

Further, we have the relation $\mathrm{Tr}\{\hat{\sigma}_i\hat{\sigma}_j\}=2\delta_{ij}$ (where δ_{ij} is the Kronecker delta). Hence Eq. (2.7) becomes

$$n_{i_1,i_2,\ldots,i_n}=\mathcal{N}\sum_{j_1,j_2,\ldots,j_n=0}^{3}\Upsilon_{i_1j_1}\Upsilon_{i_2j_2}\cdots\Upsilon_{i_nj_n}r_{i_1,i_2,\ldots,i_n}. \tag{2.9}$$

Introducing the left inverse of the matrix Υ, defined so that $\Sigma_{k=0}^{3}(\Upsilon^{-1})_{ik}\Upsilon_{kj}=\delta_{ij}$ and whose elements are

$$\Upsilon^{-1}=\begin{pmatrix}1&0&0&0\\-1&2&0&0\\-1&0&2&0\\-1&0&0&2\end{pmatrix}, \tag{2.10}$$

we can find a formula for the parameters $r_{i_1,i_2,\ldots,i_n}$ in terms of the measured quantities $n_{i_1,i_2,\ldots,i_n}$, viz.,

$$\begin{aligned}\mathcal{N}r_{i_1,i_2,\ldots,i_n}&=\sum_{j_1,j_2,\ldots,j_n=0}^{3}(\Upsilon^{-1})_{i_1j_1}(\Upsilon^{-1})_{i_2j_2}\cdots(\Upsilon^{-1})_{i_nj_n}\times n_{i_1,i_2,\ldots,i_n}\\&\equiv\mathcal{S}_{i_1,i_2,\ldots,i_n}.\end{aligned} \tag{2.11}$$

In Eq. (2.11) we have introduced the n-photon Stokes parameter $\mathcal{S}_{i_1,i_2,\ldots,i_n}$, defined in an analogous manner to the single photon Stokes parameters give in Eq. (2.2).

Since, as already noted, $r_{0,0,\ldots,0}=1$, we can make the identification $\mathcal{S}_{0,0,\ldots,0}=\mathcal{N}$, and so the density matrix for the n-qubit system can be written in terms of the Stokes parameters as follows:

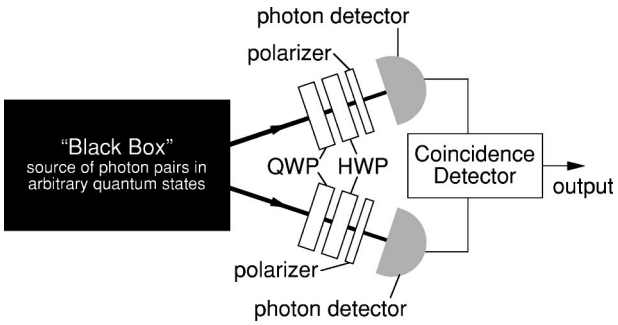

FIG. 2. Schematic illustration of the experimental arrangement. QWP stands for quarter-wave plate, HWP for half-wave plate; the angles of both pairs of wave plates can be set independently giving the experimenter four degrees of freedom with which to set the projection state. In the experiment, the polarizers were realized using polarizing prisms, arranged to transmit vertically polarized light.

$$\hat{\rho}=\frac{1}{2^n}\sum_{i_1,i_2,\ldots,i_n=0}^{3}\frac{\mathcal{S}_{i_1,i_2,\ldots,i_n}}{\mathcal{S}_{0,0,\ldots,0}}\,\hat{\sigma}_{i_1}\otimes\hat{\sigma}_{i_2}\otimes\cdots\otimes\hat{\sigma}_{i_n}. \quad (2.12)$$

This is a recipe for measurement of the density matrices which, assuming perfect experimental conditions and the complete absence of noise, will always work. It is important to realize that the set of four Stokes measurements $\{\hat{\mu}_0,\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3\}$ is not unique: there may be circumstances in which it is more convenient to use some other set, which is equivalent. A more typical set, at least in optical experiments, is $\hat{\mu}_0'=|H\rangle\langle H|$, $\hat{\mu}_1'=|V\rangle\langle V|$, $\hat{\mu}_2'=|D\rangle\langle D|$, $\hat{\mu}_3'=|R\rangle\langle R|$.

In the following section we will explore more general schemes for the measurement of two qubits, starting with a discussion, in some detail, of how the measurements are actually performed.

III. GENERALIZED TOMOGRAPHIC RECONSTRUCTION OF THE POLARIZATION STATE OF TWO PHOTONS

A. Experimental setup

The experimental arrangement used in our experiments is shown schematically in Fig. 2. An optical system consisting of lasers, polarization elements, and nonlinear optical crystals (collectively characterized for the purposes of this paper as a ''black box,'') is used to generate pairs of qubits in an almost arbitrary quantum state of their polarization degrees of freedom. A full description of this optical system and how such quantum states can be prepared can be found in Refs. [22–24].[1] The output of the black box consists of a pair of beams of light, whose quanta can be measured by means of photodetectors. To project the light beams onto a polarization state of the experimenter's choosing, three optical elements are placed in the beam in front of each detector: a polarizer (which transmits only vertically polarized light), a quarter-wave plate, and a half-wave plate. The angles of the fast axes of both of the wave plates can be set arbitrarily, allowing the $|V\rangle$ projection state fixed by the polarizer to be rotated into any polarization state that the experimenter may wish.

Using the Jones calculus notation, with the convention

$$\begin{pmatrix}0\\1\end{pmatrix}=|V\rangle,\quad \begin{pmatrix}1\\0\end{pmatrix}=|H\rangle, \quad (3.1)$$

where $|V\rangle$ ($|H\rangle$) is the ket for a vertically (horizontally) polarized beam, the effects of quarter- and half-wave plates whose fast axes are at angles q and h with respect to the vertical axis, respectively, are given by the 2×2 matrices

$$\hat{U}_{QWP}(q)=\frac{1}{\sqrt{2}}\begin{pmatrix}i-\cos(2q) & \sin(2q)\\ \sin(2q) & i+\cos(2q)\end{pmatrix},$$

$$\hat{U}_{HWP}(q)=\begin{pmatrix}\cos(2h) & -\sin(2h)\\ -\sin(2h) & -\cos(2h)\end{pmatrix}. \quad (3.2)$$

Thus the projection state for the measurement in one of the beams is given by

$$|\psi_{proj}^{(1)}(h,q)\rangle=\hat{U}_{QWP}(q)\cdot\hat{U}_{HWP}(h)\cdot\begin{pmatrix}0\\1\end{pmatrix}$$
$$=a(h,q)|H\rangle+b(h,q)|V\rangle, \quad (3.3)$$

where, neglecting an overall phase, the functions $a(h,q)$ and $b(h,q)$ are given by

$$a(h,q)=\frac{1}{\sqrt{2}}\{\sin(2h)-i\sin[2(h-q)]\},$$

$$b(h,q)=-\frac{1}{\sqrt{2}}\{\cos(2h)+i\cos[2(h-q)]\}. \quad (3.4)$$

The projection state for the two beams is given by

$$|\psi_{proj}^{(2)}(h_1,q_1,h_2,q_2)\rangle=|\psi_{proj}^{(1)}(h_1,q_1)\rangle\otimes|\psi_{proj}^{(1)}(h_2,q_2)\rangle$$
$$=a(h_1,q_1)a(h_2,q_2)|HH\rangle$$
$$+a(h_1,q_1)b(h_2,q_2)|HV\rangle$$
$$+b(h_1,q_1)a(h_2,q_2)|VH\rangle$$
$$+b(h_1,q_1)b(h_2,q_2)|VV\rangle. \quad (3.5)$$

We shall denote the projection state corresponding to one particular set of wave plate angles $\{h_{1,\nu},q_{1,\nu},h_{2,\nu},q_{2,\nu}\}$ by

[1]It is important to realize that the entangled photon pairs are produced in a *nondeterministic* manner: one cannot specify with certainty when a photon pair will be emitted; indeed there is a small probability of generating four or six or a higher number of photons. Thus we can only *postselectively* generate entangled photon pairs: i.e., one only knows that the state was created after if has been measured.

the ket $|\psi_\nu\rangle$;[2] thus the projection measurement is represented by the operator $\hat{\mu}_\nu = |\psi_\nu\rangle\langle\psi_\nu|$. Consequently, the average number of coincidence counts that will be observed in a given experimental run is

$$n_\nu = \mathcal{N}\langle\psi_\nu|\hat{\rho}|\psi_\nu\rangle \tag{3.6}$$

where $\hat{\rho}$ is the density matrix describing the ensemble of qubits, and $\mathcal{N}$ is a constant dependent on the photon flux and detector efficiencies. In what follows, it will be convenient to consider the quantities s_ν defined by

$$s_\nu = \langle\psi_\nu|\hat{\rho}|\psi_\nu\rangle. \tag{3.7}$$

B. Tomographically complete set of measurements

In Sec. II we have given one possible set of projection measurements $\{|\psi_\nu\rangle\langle\psi_\nu|\}$ which uniquely determine the density matrix $\hat{\rho}$. However, one can conceive of situations in which these will not be the most convenient set of measurements to make. Here we address the problem of finding other sets of suitable measurements. The smallest number of states required for such measurements can be found by a simple argument: there are 15 real unknown parameters that determine a 4×4 density matrix, plus there is the single unknown real parameter $\mathcal{N}$, making a total of 16.

In order to proceed it is helpful to convert the 4×4 matrix $\hat{\rho}$ into a 16-dimensional column vector. To do this we use a set of 16 linearly independent 4×4 matrices $\{\hat{\Gamma}_\nu\}$ which have the following mathematical properties:

$$\mathrm{Tr}\{\hat{\Gamma}_\nu\cdot\hat{\Gamma}_\mu\} = \delta_{\nu,\mu}$$

$$\hat{A} = \sum_{\nu=1}^{16} \hat{\Gamma}_\nu \mathrm{Tr}\{\hat{\Gamma}_\nu\cdot\hat{A}\}\ \forall\hat{A}, \tag{3.8}$$

where $\hat{A}$ is an arbitrary 4×4 matrix. Finding a set of $\hat{\Gamma}_\nu$ matrices is in fact reasonably straightforward: for example, the set of (appropriately normalized) generators of the Lie algebra $SU(2)\otimes SU(2)$ fulfill the required criteria (for reference, we list this set in Appendix A). These matrices are of course simply a relabeling of the two-qubit Pauli matrices $\hat{\sigma}_i\otimes\hat{\sigma}_j$ $(i,j=0,1,2,3)$ discussed above. Using these matrices the density matrix can be written as

$$\hat{\rho} = \sum_{\nu=1}^{16} \hat{\Gamma}_\nu r_\nu, \tag{3.9}$$

where r_ν is the νth element of a 16-element column vector, given by the formula

$$r_\nu = \mathrm{Tr}\{\hat{\Gamma}_\nu\cdot\hat{\rho}\}. \tag{3.10}$$

[2]Here the first subscript on the wave plate angle refers to one of the two photon beams; the second subscript distinguishes which of the 16 different experimental states is under consideration.

TABLE I. The tomographic analysis states used in our experiments. The number of coincidence counts measured in projection measurements provides a set of 16 data that allow the density matrix of the state of the two modes to be estimated. We have used the notation $|D\rangle \equiv (|H\rangle + |V\rangle)/\sqrt{2}$, $|L\rangle \equiv (|H\rangle + i|V\rangle)/\sqrt{2}$, and $|R\rangle \equiv (|H\rangle - i|V\rangle)/\sqrt{2}$. Note that, when the measurements are taken in the order given by the table, only one wave plate angle has to be changed between measurements.

ν	Mode 1	Mode 2	h_1	q_1	h_2	q_2
1	$\|H\rangle$	$\|H\rangle$	45°	0	45°	0
2	$\|H\rangle$	$\|V\rangle$	45°	0	0	0
3	$\|V\rangle$	$\|V\rangle$	0	0	0	0
4	$\|V\rangle$	$\|H\rangle$	0	0	45°	0
5	$\|R\rangle$	$\|H\rangle$	22.5°	0	45°	0
6	$\|R\rangle$	$\|V\rangle$	22.5°	0	0	0
7	$\|D\rangle$	$\|V\rangle$	22.5°	45°	0	0
8	$\|D\rangle$	$\|H\rangle$	22.5°	45°	45°	0
9	$\|D\rangle$	$\|R\rangle$	22.5°	45°	22.5°	0
10	$\|D\rangle$	$\|D\rangle$	22.5°	45°	22.5°	45°
11	$\|R\rangle$	$\|D\rangle$	22.5°	0	22.5°	45°
12	$\|H\rangle$	$\|D\rangle$	45°	0	22.5°	45°
13	$\|V\rangle$	$\|D\rangle$	0	0	22.5°	45°
14	$\|V\rangle$	$\|L\rangle$	0	0	22.5°	90°
15	$\|H\rangle$	$\|L\rangle$	45°	0	22.5°	90°
16	$\|R\rangle$	$\|L\rangle$	22.5°	0	22.5°	90°

Substituting from Eq. (3.9) into Eq. (3.6), we obtain the following linear relationship between the measured coincidence counts n_ν and the elements of the vector r_μ:

$$n_\nu = \mathcal{N}\sum_{\mu=1}^{16} B_{\nu,\mu} r_\mu \tag{3.11}$$

where the 16×16 matrix $B_{\nu,\mu}$ is given by

$$B_{\nu,\mu} = \langle\psi_\nu|\hat{\Gamma}_\mu|\psi_\nu\rangle. \tag{3.12}$$

Immediately we find a necessary and sufficient condition for the completeness of the set of tomographic states $\{|\psi_\nu\rangle\}$: if the matrix $B_{\nu,\mu}$ is nonsingular, then Eq. (3.11) can be inverted to give

$$r_\nu = (\mathcal{N})^{-1}\sum_{\mu=1}^{16} (B^{-1})_{\nu,\mu} n_\mu. \tag{3.13}$$

The set of 16 tomographic states that we employed is given in Table I. They can be shown to satisfy the condition that $B_{\nu,\mu}$ is nonsingular. By no means are these states unique in this regard: these were the states chosen principally for experimental convenience.

These states can be realized by setting specific values of the half- and quarter-wave plate angles. The appropriate values of these angles (measured from the vertical) are given in Table I. Note that overall phase factors do not affect the results of projection measurements.

Substituting Eq. (3.13) into Eq. (3.9), we find that

$$\hat{\rho} = (\mathcal{N})^{-1} \sum_{\nu=1}^{16} \hat{M}_\nu n_\nu = \sum_{\nu=1}^{16} \hat{M}_\nu s_\nu , \tag{3.14}$$

where the sixteen 4×4 matrices $\hat{M}_\nu$ are defined by

$$\hat{M}_\nu = \sum_{\nu=1}^{16} (B^{-1})_{\nu,\mu} \hat{\Gamma}_\mu . \tag{3.15}$$

The introduction of the $\hat{M}_\nu$ matrices allows a compact form of linear tomographic reconstruction, Eq. (3.14), which will be most useful in the error analysis that follows. These $\hat{M}_\nu$ matrices, valid for our set of tomographic states, are listed in Appendix B, together with some of their important properties. We can use one of these properties, Eq. (B6), to obtain the value of the unknown quantity $\mathcal{N}$. That relationship implies

$$\sum_\nu \mathrm{Tr}\{\hat{M}_\nu\} |\psi_\nu\rangle\langle\psi_\nu| \hat{\rho} = \hat{\rho}. \tag{3.16}$$

Taking the trace of this formula, and multiplying by $\mathcal{N}$ we obtain

$$\sum_\nu \mathrm{Tr}\{\hat{M}_\nu\} n_\nu = \mathcal{N}. \tag{3.17}$$

For our set of tomographic states, it can be shown that

$$\mathrm{Tr}\{\hat{M}_\nu\} = \begin{cases} 1 & \text{if } \nu = 1,2,3,4 \\ 0 & \text{if } \nu = 5,\ldots,16; \end{cases} \tag{3.18}$$

hence the value of the unknown parameter $\mathcal{N}$ in our experiments is given by

$$\begin{aligned} \mathcal{N} &= \sum_{\nu=1}^{4} n_\nu \\ &= \mathcal{N}(\langle HH|\hat{\rho}|HH\rangle + \langle HV|\hat{\rho}|HV\rangle \\ &\quad + \langle VH|\hat{\rho}|VH\rangle + \langle VV|\hat{\rho}|VV\rangle). \end{aligned} \tag{3.19}$$

Thus we obtain the final formula for the tomographic reconstruction of the density matrices of our states:

$$\hat{\rho} = \left(\sum_{\nu=1}^{16} \hat{M}_\nu n_\nu \right) \Big/ \left(\sum_{\nu=1}^{4} n_\nu \right). \tag{3.20}$$

As an example, the following set of 16 counts were taken for the purpose of tomographically determining the density matrix for an ensemble of qubits all prepared in a specific quantum state: $n_1 = 34\,749$, $n_2 = 324$, $n_3 = 35\,805$, $n_4 = 444$, $n_5 = 16\,324$, $n_6 = 17\,521$, $n_7 = 13\,441$, $n_8 = 16\,901$, $n_9 = 17\,932$, $n_{10} = 32\,028$, $n_{11} = 15\,132$, $n_{12} = 17\,238$, $n_{13} = 13\,171$, $n_{14} = 17\,170$, $n_{15} = 16\,722$, $n_{16} = 33\,586$. Applying Eq. (3.20) we find

$$\hat{\rho} = \begin{pmatrix} 0.4872 & -0.0042 + i0.0114 & -0.0098 - i0.0178 & 0.5192 + i0.0380 \\ -0.0042 - i0.0114 & 0.0045 & 0.0271 - i0.0146 & -0.0648 - i0.0076 \\ -0.0098 + i0.0178 & 0.0271 + i0.0146 & 0.0062 & -0.0695 + i0.0134 \\ 0.5192 - i0.0380 & -0.0648 + i0.0076 & -0.0695 - i0.0134 & 0.5020 \end{pmatrix}. \tag{3.21}$$

This matrix is shown graphically in Fig. 3(left).

Note that, by construction, the density matrix is normalized, i.e., $\mathrm{Tr}\{\hat{\rho}\} = 1$, and Hermitian, i.e., $\hat{\rho}^\dagger = \hat{\rho}$. However, when one calculates the eigenvalues of this measured density matrix, one finds the values 1.021 55, 0.068 123 8, $-0.065\,274$, and $-0.024\,396$; and also $\mathrm{Tr}\{\hat{\rho}^2\} = 1.053$. Density matrices for all physical states must have the property of positive semidefiniteness, which (in conjunction with the normalization and Hermiticity properties) implies that all of the eigenvalues must lie in the interval $[0,1]$, their sum being 1; this in turn implies that $0 \leqslant \mathrm{Tr}\{\hat{\rho}^2\} \leqslant 1$. Clearly, the density matrix reconstructed above by linear tomography violates this condition. From our experience of tomographic measurements of various mixed and entangled states prepared experimentally, this seems to happen roughly 75% of the time for low-entropy, highly entangled states; it seems to have a higher probability of producing the correct result for states of higher entropy, but the cautious experimenter should check every time. The obvious culprit for this problem is experimental inaccuracies and statistical fluctuations of coincidence counts, which mean that the actual numbers of counts recorded in a real experiment differ from those that can be calculated by Eq. (3.6). Thus the linear reconstruction is of limited value for states of low entropy (which are of most experimental interest because of their application to quantum information technology); however, as we shall see, the linear approach does provide a useful starting point for the numerical optimization approach to density matrix estimation which we will discuss in the next section.

IV. MAXIMUM LIKELIHOOD ESTIMATION

As mentioned in Sec. III, the tomographic measurement of density matrices can produce results that violate important basic properties such as positivity. To avoid this problem, the maximum likelihood estimation of density matrices may be employed. Here we describe a simple realization of this technique.

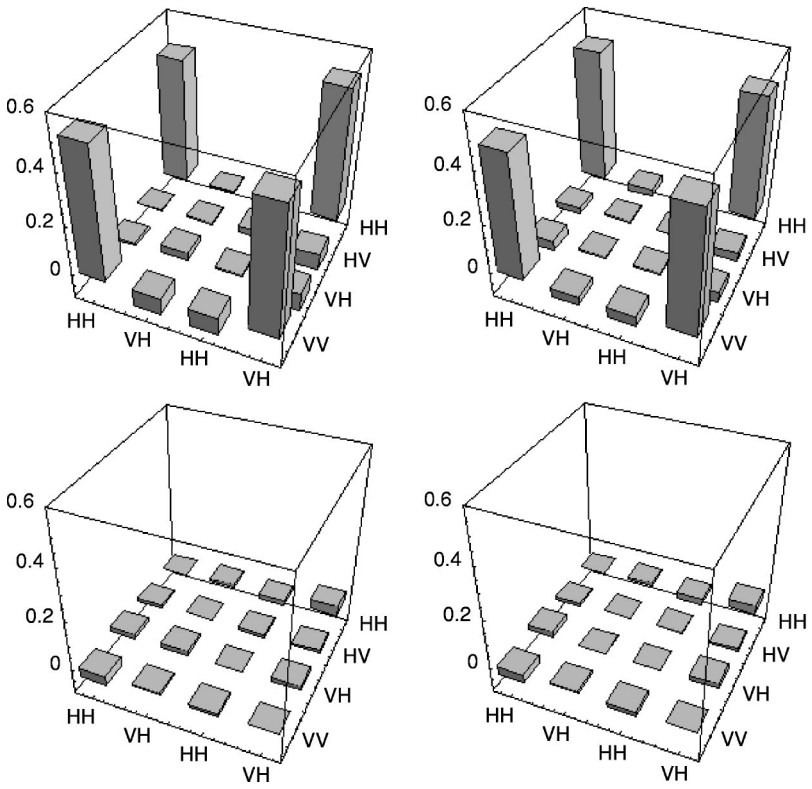

FIG. 3. Graphical representation of the density matrix of a state as estimated by linear tomography (left) and by maximum likelihood tomography (right) from the experimental data given in the text. The upper plot is the real part of $\hat{\rho}$, the lower plot the imaginary part.

A. Basic approach

Our approach to the maximum likelihood estimation of the density matrix is as follows.

(i) Generate a formula for an explicitly ''physical'' density matrix, i.e., a matrix that has the three important properties of normalization, Hermiticity, and positivity. This matrix will be a function of 16 real variables (denoted $\{t_1,t_2,\ldots,t_{16}\}$). We will denote the matrix as $\hat{\rho}_p(t_1,t_2,\ldots,t_{16})$.

(ii) Introduce a ''likelihood function'' which quantifies how good the density matrix $\hat{\rho}_p(t_1,t_2,\ldots,t_{16})$ is in relation to the experimental data. This likelihood function is a function of the 16 real parameters t_ν and of the 16 experimental data n_ν. We will denote this function as $\mathcal{L}(t_1,t_2,\ldots,t_{16};n_1,n_2,\ldots,n_{16})$.

(iii) Using standard numerical optimization techniques, find the optimum set of variables $\{t_1^{(opt)},t_2^{(opt)},\ldots,t_{16}^{(opt)}\}$ for which the function $\mathcal{L}(t_1,t_2,\ldots,t_{16};n_1,n_2,\ldots,n_{16})$ has its maximum value. The best estimate for the density matrix is then $\hat{\rho}(t_1^{(opt)},t_2^{(opt)},\ldots,t_{16}^{(opt)})$.

The details of how these three steps can be carried out are described in the next three subsections.

B. Physical density matrices

The property of non-negative definiteness for any matrix $\hat{\mathcal{G}}$ is written mathematically as

$$\langle\psi|\hat{\mathcal{G}}|\psi\rangle\geqslant 0 \quad \forall|\psi\rangle. \tag{4.1}$$

Any matrix that can be written in the form $\hat{\mathcal{G}}=\hat{T}^\dagger\hat{T}$ must be non-negative definite. To see that this is the case, substitute into Eq. (4.1)

$$\langle\psi|\hat{T}^\dagger\hat{T}|\psi\rangle=\langle\psi'|\psi'\rangle\geqslant 0, \tag{4.2}$$

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

$$\hat{g}=\hat{T}^\dagger\hat{T}/\mathrm{Tr}\{\hat{T}^\dagger\hat{T}\} \tag{4.3}$$

has all three of the mathematical properties that we require for density matrices.

For the two-qubit system, we have a 4×4 density matrix with 15 independent real parameters. Since it will be useful to be able to invert relation (4.3), it is convenient to choose a tridiagonal form for $\hat{T}$:

$$\hat{T}(t)=\begin{pmatrix} t_1 & 0 & 0 & 0 \\ t_5+it_6 & t_2 & 0 & 0 \\ t_{11}+it_{12} & t_7+it_8 & t_3 & 0 \\ t_{15}+it_{16} & t_{13}+it_{14} & t_9+it_{10} & t_4 \end{pmatrix}. \tag{4.4}$$

Thus the explicitly ''physical'' density matrix $\hat{\rho}_p$ is given by the formula

$$\hat{\rho}_p(t)=\hat{T}^\dagger(t)\hat{T}(t)/\mathrm{Tr}\{\hat{T}^\dagger(t)\hat{T}(t)\}. \tag{4.5}$$

For future reference, the inverse relationship, by which the elements of $\hat{T}$ can be expressed in terms of the elements of $\hat{\rho}$, is as follows:

$$\hat{T}=\begin{pmatrix} \sqrt{\dfrac{\Delta}{\mathcal{M}_{11}^{(1)}}} & 0 & 0 & 0 \\ \dfrac{\mathcal{M}_{12}^{(1)}}{\sqrt{\mathcal{M}_{11}^{(1)}\mathcal{M}_{11,22}^{(2)}}} & \sqrt{\dfrac{\mathcal{M}_{11}^{(1)}}{\mathcal{M}_{11,22}^{(2)}}} & 0 & 0 \\ \dfrac{\mathcal{M}_{12,23}^{(2)}}{\sqrt{\rho_{44}}\sqrt{\mathcal{M}_{11,22}^{(2)}}} & \dfrac{\mathcal{M}_{11,23}^{(2)}}{\sqrt{\rho_{44}}\sqrt{\mathcal{M}_{11,22}^{(2)}}} & \sqrt{\dfrac{\mathcal{M}_{11,22}^{(2)}}{\rho_{44}}} & 0 \\ \dfrac{\rho_{41}}{\sqrt{\rho_{44}}} & \dfrac{\rho_{42}}{\sqrt{\rho_{44}}} & \dfrac{\rho_{43}}{\sqrt{\rho_{44}}} & \sqrt{\rho_{44}} \end{pmatrix}. \tag{4.6}$$

Here we have used the notation $\Delta=\mathrm{Det}(\hat{\rho})$; $\mathcal{M}_{ij}^{(1)}$ is the first minor of $\hat{\rho}$, i.e., the determinant of the 3×3 matrix formed by deleting the ith row and jth column of $\hat{\rho}$; $\mathcal{M}_{ij,kl}^{(2)}$ is the second minor of $\hat{\rho}$, i.e., the determinant of the 2×2 matrix formed by deleting the ith and kth rows and jth and lth columns of $\hat{\rho}$ ($i\neq k$ and $j\neq l$).

C. The likelihood function

The measurement data consist of a set of 16 coincidence counts n_ν ($\nu=1,2,\ldots,16$) whose expected value is $\overline{n}_\nu=\mathcal{N}\langle\psi_\nu|\hat{\rho}|\psi_\nu\rangle$. Let us assume that the noise on these coincidence measurements has a Gaussian probability distribution. Thus the probability of obtaining a set of 16 counts $\{n_1,n_2,\ldots n_{16}\}$ is

$$P(n_1,n_2,\ldots,n_{16})=\frac{1}{N_{\mathrm{norm}}}\prod_{\nu=1}^{16}\exp\left[-\frac{(n_\nu-\overline{n}_\nu)^2}{2\sigma_\nu^2}\right], \tag{4.7}$$

where σ_ν is the standard deviation for the νth coincidence measurement (given approximately by $\sqrt{\overline{n}_\nu}$) and N_{norm} is the normalization constant. For our physical density matrix $\hat{\rho}_p$ the number of counts expected for the νth measurement is

$$\overline{n}_\nu(t_1,t_2,\ldots,t_{16})=\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1,t_2,\ldots,t_{16})|\psi_\nu\rangle. \tag{4.8}$$

Thus the likelihood that the matrix $\hat{\rho}_p(t_1,t_2,\ldots,t_{16})$ could produce the measured data $\{n_1,n_2,\ldots,n_{16}\}$ is

$$P(n_1,n_2,\ldots,n_{16})=\frac{1}{N_{\mathrm{norm}}}\prod_{\nu=1}^{16}\exp\left[-\frac{[\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1,t_2,\ldots,t_{16})|\psi_\nu\rangle-n_\nu]^2}{2\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1,t_2,\ldots,t_{16})|\psi_\nu\rangle}\right], \tag{4.9}$$

where $\mathcal{N}=\Sigma_{\nu=1}^{4}n_\nu$.

Rather than find the maximum value of $P(t_1,t_2,\ldots,t_{16})$ it simplifies things somewhat to find the maximum of its logarithm (which is mathematically equivalent).[3] Thus the optimization problem reduces to finding the *minimum* of the following function:

$$\mathcal{L}(t_1,t_2,\ldots,t_{16})=\sum_{\nu=1}^{16}\frac{[\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1,t_2,\ldots,t_{16})|\psi_\nu\rangle-n_\nu]^2}{2\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1,t_2,\ldots,t_{16})|\psi_\nu\rangle}. \tag{4.10}$$

This is the ''likelihood'' function that we employed in our numerical optimization routine.

D. Numerical optimization

We used the MATHEMATICA 4.0 routine FINDMINIMUM which executes a multidimensional Powell direction set algorithm (see Ref. [25] for a description of this algorithm). To execute this routine, one requires an initial estimate for the values of $t_1,t_2,\ldots,t_{16}$. For this, we used the tomographic estimate of the density matrix in the inverse relation (4.6), allowing us to determine a set of values for $t_1,t_2,\ldots t_{16}$. Since the tomographic density matrix may not be nonnegative definite, the values of the t_ν's deduced in this manner are not necessarily real. Thus for our initial guess we used the real parts of the t_ν's deduced from the tomographic density matrix.

For the example given in Sec. II, the maximum likelihood estimate is

[3]Note that here we neglect the dependence of the normalization constant on $t_1,t_2,\ldots,t_{16}$, which only weakly affects the solution for the most likely state.

$$\hat{\rho}=\begin{pmatrix} 0.5069 & -0.0239+i0.0106 & -0.0412-i0.0221 & 0.4833+i0.0329 \\ -0.0239-i0.0106 & 0.0048 & 0.0023+i0.0019 & -0.0296-i0.0077 \\ -0.0412+i0.0221 & 0.0023-i0.0019 & 0.0045 & -0.0425+i0.0192 \\ 0.4833-i0.0329 & -0.0296+i0.0077 & -0.0425-i0.0192 & 0.4839 \end{pmatrix}. \tag{4.11}$$

This matrix is illustrated in Fig. 3 (right). In this case, the matrix has eigenvalues 0.986 022, 0.013 977 7, 0, and 0; and $\mathrm{Tr}\{\hat{\rho}^2\}=0.972\,435$, indicating that, while the linear reconstruction gave a nonphysical density matrix, the maximum likelihood reconstruction gives a legitimate density matrix.

V. ERROR ANALYSIS

In this section we present an analysis of the errors inherent in the tomographic scheme described in Sec. III. Two sources of error are found to be important: the shot noise error in the measured coincidence counts n_ν and the uncertainty in the settings of the angles of the wave plates used to make the tomographic projection states. We will analyze these two sources separately.

In addition to determining the density matrix of a pair of qubits, one is often also interested in quantities derived from the density matrix, such as the entropy or the entanglement of formation. For completeness, we will also derive the errors in some of these quantities.

A. Errors due to count statistics

From Eq. (3.20) we see that the density matrix is specified by a set of 16 parameters s_ν defined by

$$s_\nu=n_\nu/\mathcal{N}, \tag{5.1}$$

where n_ν are the measured coincidence counts and $\mathcal{N}=\Sigma_{\nu=1}^{4}n_\nu$. We can determine the errors in s_ν using the following formula [26]:

$$\overline{\delta s_\nu \delta s_\mu}=\sum_{\lambda,\kappa=1}^{16}\left(\frac{\partial s_\nu}{\partial n_\lambda}\right)\left(\frac{\partial s_\mu}{\partial n_\kappa}\right)\overline{\delta n_\lambda \delta n_\kappa}, \tag{5.2}$$

where the overbar denotes the ensemble average of the random uncertainties δs_ν and δn_λ. The measured coincidence counts n_λ are statistically independent Poissonian random variables, which implies the following relation:

$$\overline{\delta n_\lambda \delta n_\kappa}=n_\lambda \delta_{\lambda,\kappa}, \tag{5.3}$$

where $\delta_{\lambda,\kappa}$ is the Kronecker delta.

Taking the derivative of Eq. (5.1), we find that

$$\frac{\partial s_\mu}{\partial n_\nu}=\frac{1}{\mathcal{N}}\delta_{\mu\nu}-\frac{n_\mu}{\mathcal{N}^2}D_\nu, \tag{5.4}$$

where

$$D_\nu=\sum_{\lambda=1}^{4}\delta_{\lambda,\nu}=\begin{cases}1 & \text{if } 1\leqslant\nu\leqslant 4\\ 0 & \text{if } 5\leqslant\nu\leqslant 16.\end{cases} \tag{5.5}$$

Substituting from Eq. (5.4) into Eq. (5.2) and using Eq. (5.3), we obtain the result

$$\overline{\delta s_\nu \delta s_\mu}=\frac{n_\mu}{\mathcal{N}^2}\delta_{\nu,\mu}+\frac{n_\nu n_\mu}{\mathcal{N}^3}(1-D_\mu-D_\nu). \tag{5.6}$$

In most experimental circumstances $\mathcal{N}\gg 1$, and so the second term in Eq. (5.6) is negligibly small in comparison to the first. We shall therefore ignore it, and use the approximate expression in the subsequent discussion:

$$\overline{\delta s_\nu \delta s_\mu}\approx\frac{n_\mu}{\mathcal{N}^2}\delta_{\nu,\mu}\equiv\frac{s_\mu}{\mathcal{N}}\delta_{\nu,\mu}. \tag{5.7}$$

B. Errors due to angular settings uncertainties

Using the formula (3.7) for the parameters s_ν we can find the dependence of the measured density matrix on errors in the tomographic states. The derivative of s_ν with respect to some generic wave plate setting angle θ is

$$\frac{\partial s_\nu}{\partial\theta}=\left\{\frac{\partial}{\partial\theta}\langle\psi_\nu|\right\}\hat{\rho}|\psi_\nu\rangle+\langle\psi_\nu|\hat{\rho}\left\{\frac{\partial}{\partial\theta}|\psi_\nu\rangle\right\}, \tag{5.8}$$

where $|\psi_\nu\rangle$ is the ket of the νth projection state [see Eq. (3.5)]. Substituting from Eq. (3.14) we find

$$\frac{\partial s_\nu}{\partial\theta}=\sum_{\mu=1}^{16}s_\mu\left[\left\{\frac{\partial}{\partial\theta}\langle\psi_\nu|\right\}\hat{M}_\mu|\psi_\nu\rangle+\langle\psi_\nu|\hat{M}_\mu\left\{\frac{\partial}{\partial\theta}|\psi_\nu\rangle\right\}\right]. \tag{5.9}$$

For convenience, we shall label the four wave plate angles $\{h_{1,\nu},q_{1,\nu},h_{2,\nu},q_{2,\nu}\}$, which specify the νth state by $\{\theta_{\nu,1},\theta_{\nu,2},\theta_{\nu,3},\theta_{\nu,4}\}$, respectively. Clearly the μth state does not depend on any of the νth set of angles. Thus we obtain the following expression for the derivatives of s_ν with respect to wave plate settings:

$$\frac{\partial s_\nu}{\partial\theta_{\lambda,i}}=\delta_{\nu,\lambda}\sum_{\mu=1}^{16}s_\mu f_{\nu,\mu}^{(i)}, \tag{5.10}$$

where

$$f_{\nu,\mu}^{(i)}=\left\{\frac{\partial}{\partial\theta_{\nu,i}}\langle\psi_\nu|\right\}\hat{M}_\mu|\psi_\nu\rangle+\langle\psi_\nu|\hat{M}_\mu\left\{\frac{\partial}{\partial\theta_{\nu,i}}|\psi_\nu\rangle\right\}. \tag{5.11}$$

The 1024 quantities $f^{(i)}_{\nu,\mu}$ can be determined by taking the derivatives of the functional forms of the tomographic states given by Eqs. (3.4) and (3.5), and evaluating those derivatives at the appropriate values of the arguments (see Table I).

The errors in the angles are assumed to be uncorrelated, as would be the case if each wave plate were adjusted for each of the 16 measurements. In reality, for qubit experiments, only one or two of the four wave plates are adjusted between one measurement and the next. However, the assumption of uncorrelated angular errors greatly simplifies the calculation (which is, after all, only an *estimate* of the errors), and seems to produce reasonable figures for our error bars.[4] Thus, with the assumption

$$\overline{\delta\theta_{\nu,i}\delta\theta_{\mu,j}}=\delta_{\nu,\mu}\delta_{i,j}(\Delta\theta)^2 \tag{5.12}$$

(where $\Delta\theta$ is the rms uncertainty in the setting of the wave plate, with an estimated value of 0.25° for our apparatus), we obtain the following expression for the errors in s_ν due to angular settings:

$$\overline{\delta s_\nu \delta s_\mu}=\delta_{\nu,\mu}\sum_{i=1}^{4}\sum_{\epsilon,\lambda=1}^{16} f^{(i)}_{\nu,\epsilon} f^{(i)}_{\nu,\lambda} s_\epsilon s_\lambda . \tag{5.13}$$

Combining Eqs. (5.13) and (5.7) we obtain the following formula for the total error in the quantities s_ν:

$$\overline{\delta s_\nu \delta s_\mu}=\delta_{\nu,\mu}\Lambda_\nu \tag{5.14}$$

where

$$\Lambda_\nu=\left[\frac{s_\nu}{\mathcal{N}}+\sum_{i=1}^{4}\sum_{\epsilon,\lambda=1}^{16} f^{(i)}_{\nu,\epsilon} f^{(i)}_{\nu,\lambda} s_\epsilon s_\lambda\right]. \tag{5.15}$$

These 16 quantities can be calculated using the parameters s_ν and the constants $f^{(i)}_{\nu,\epsilon}$. Note that the same result can be obtained by assuming *a priori* that the errors in the s_ν are all uncorrelated, with $\Lambda_\nu=\delta s_\nu^2$; the more rigorous treatment given here is necessary, however, to demonstrate this fact. For a typical number of counts, say $\mathcal{N}=10\,000''''$ it is found that the contribution of errors from the two causes is roughly comparable; for larger numbers of counts, the angular settings will become the dominant source of error.

Based on these results, the errors in the values of the various elements of the density matrix estimated by the linear tomographic technique described in Sec. III are as follows:

$$(\Delta\rho_{i,j})^2=\sum_{\nu,\mu=1}^{16}\frac{\partial\rho_{i,j}}{\partial s_\nu}\frac{\partial\rho_{i,j}}{\partial s_\mu}\overline{\delta s_\nu\delta s_\mu}=\sum_{\nu=1}^{16}(M_{\nu(i,j)})^2\Lambda_\nu \tag{5.16}$$

[4]In other experimental circumstances, such as the measurement of the joint state of two spin-1/2 particles, the tomography would be realized by performing unitary operations on the spins prior to measurement. In this case, an assumption analogous to ours would be wholly justified.

where $M_{\nu(i,j)}$ is the i,j element of the matrix $\hat{M}_\nu$.

A convenient way in which to estimate errors for a maximum likelihood tomographic technique (rather than a linear tomographic technique) is to employ the above formulas, with the slight modification that the parameter s_ν should be recalculated from Eq. (3.7) using the estimated density matrix $\hat{\rho}_{est}$. This does not take into account errors inherent in the maximum likelihood technique itself.

C. Errors in quantities derived from the density matrix

When calculating the propagation of errors, it is actually more convenient to use the errors in the s_ν parameters [given by Eq. (5.15)], rather than the errors in the elements of density matrix itself (which have non-negligible correlations).

1. Von Neumann entropy

The von Neumann entropy is an important measure of the purity of a quantum state $\hat{\rho}$. It is defined by [27]

$$\mathcal{S}=-\mathrm{Tr}\{\hat{\rho}\log_2(\hat{\rho})\}=-\sum_{a=1}^{4} p_a\log_2 p_a , \tag{5.17}$$

where p_a is an eigenvalue of $\hat{\rho}$, i.e.,

$$\hat{\rho}|\phi_a\rangle=p_a|\phi_a\rangle, \tag{5.18}$$

$|\phi_a\rangle$ being the ath eigenstate ($a=1,\ldots,4$). The error in this quantity is given by

$$(\Delta\mathcal{S})^2=\sum_{\nu=1}^{16}\left(\frac{\partial\mathcal{S}}{\partial s_\nu}\right)^2\Lambda_\nu . \tag{5.19}$$

Applying the chain rule, we find

$$\left(\frac{\partial\mathcal{S}}{\partial s_\nu}\right)=\sum_{a=1}^{4}\left(\frac{\partial p_a}{\partial s_\nu}\right)\left(\frac{\partial\mathcal{S}}{\partial p_a}\right). \tag{5.20}$$

The partial differential of an eigenvalue can be easily found by perturbation theory. As is well known (e.g., [28]) the change in the eigenvalue λ_a of a matrix $\hat{W}$ due to a perturbation in the matrix $\delta\hat{W}$ is

$$\delta\lambda_a=\langle\phi_a|\delta\hat{W}|\phi_a\rangle, \tag{5.21}$$

where $|\phi_a\rangle$ is the eigenvector of $\hat{W}$ corresponding to the eigenvalue λ_a. Thus the derivative of λ_a with respect to some variable x is given by

$$\frac{\partial\lambda_a}{\partial x}=\left\langle\phi_a\left|\frac{\partial\hat{W}}{\partial x}\right|\phi_a\right\rangle. \tag{5.22}$$

Since $\hat{\rho}=\Sigma_{\nu=1}^{16}\hat{M}_\nu s_\nu$, we find that

$$\frac{\partial p_a}{\partial s_\nu}=\langle\phi_a|\hat{M}_\nu|\phi_a\rangle \tag{5.23}$$

and so, taking the derivative of Eq. (5.17), Eq. (5.20) becomes

$$\left(\frac{\partial \mathcal{S}}{\partial s_\nu}\right) = -\sum_{a=1}^{4} \langle \phi_a | \hat{M}_\nu | \phi_a \rangle \frac{[1+\ln p_a]}{\ln 2}. \tag{5.24}$$

Hence

$$(\Delta \mathcal{S})^2 = \sum_{\nu=1}^{16} \left(\sum_{a=1}^{4} \langle \phi_a | \hat{M}_\nu | \phi_a \rangle \frac{[1+\ln p_a]}{\ln 2} \right)^2 \Lambda_\nu . \tag{5.25}$$

For the experimental example given above, $\mathcal{S}=0.106 \pm 0.049$.

2. Linear entropy

The ''linear entropy'' is used to quantify the degree of mixture of a quantum state in an analytically convenient form, although unlike the von Neumann entropy it has no direct information theoretic implications. In a normalized form (defined so that its value lies between zero and 1), the linear entropy for a two-qubit system is defined by

$$\mathcal{P} = \frac{4}{3}(1 - \mathrm{Tr}\{\hat{\rho}^2\}) = \frac{4}{3}\left(1 - \sum_{a=1}^{4} p_a^2\right). \tag{5.26}$$

To calculate the error in this quantity, we need the following partial derivative:

$$\begin{aligned} \frac{\partial \mathcal{P}}{\partial s_\nu} &= -\frac{8}{3}\sum_{a=1}^{4} p_a \frac{\partial p_a}{\partial s_\nu} \\ &= -\frac{8}{3}\sum_{a=1}^{4} p_a \langle \phi_a | \hat{M}_\nu | \phi_a \rangle \\ &= -\frac{8}{3}\mathrm{Tr}\{\hat{\rho}\hat{M}_\nu\} \\ &= -\frac{8}{3}\sum_{\mu=1}^{16} \mathrm{Tr}\{\hat{M}_\mu \hat{M}_\nu\} s_\mu . \end{aligned} \tag{5.27}$$

Hence the error in the linear entropy is

$$(\Delta \mathcal{P})^2 = \sum_{\nu=1}^{16} \left(\frac{\partial \mathcal{P}}{\partial s_\nu}\right)^2 \Lambda_\nu = \sum_{\nu}^{16} \left(\frac{8}{3}\sum_{\mu=1}^{16} \mathrm{Tr}\{\hat{M}_\mu \hat{M}_\nu\} s_\mu \right)^2 \Lambda_\nu . \tag{5.28}$$

For the example given in Secs. III and IV, $\mathcal{P}=0.037 \pm 0.026$.

3. Concurrence, entanglement of Formation, and tangle

The concurrence, entanglement of formation, and tangle are measures of the quantum coherence properties of a mixed quantum state [29]. For two qubits,[5] concurrence is defined as follows: consider the non-Hermitian matrix $\hat{R} = \hat{\rho}\hat{\Sigma}\hat{\rho}^{\mathrm{T}}\hat{\Sigma}$ where the superscript T denotes the transpose and the ''spin flip matrix'' $\hat{\Sigma}$ is defined by

$$\hat{\Sigma} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{5.29}$$

Note that the definition of $\hat{\Sigma}$ depends on the basis chosen; we have assumed here the ''computational basis'' $\{|HH\rangle, |HV\rangle, |VH\rangle, |VV\rangle\}$. In what follows, it will be convenient to write $\hat{R}$ in the following form:

$$\hat{R} = \frac{1}{2}\sum_{\mu,\nu=1}^{16} \hat{q}_{\mu,\nu} s_\mu s_\nu , \tag{5.30}$$

where $\hat{q}_{\mu,\nu} = \hat{M}_\mu \hat{\Sigma} \hat{M}_\nu^{\mathrm{T}} \hat{\Sigma} + \hat{M}_\nu \hat{\Sigma} \hat{M}_\mu^{\mathrm{T}} \hat{\Sigma}$. The left and right eigenstates and eigenvalues of the matrix $\hat{R}$ we shall denote by $\langle \xi_a |$, $|\zeta_a\rangle$, and r_a, respectively, i.e.,

$$\begin{aligned} \langle \xi_a | \hat{R} &= r_a \langle \xi_a |, \\ \hat{R} | \zeta_a \rangle &= r_a | \zeta_a \rangle. \end{aligned} \tag{5.31}$$

We shall assume that these eigenstates are normalized in the usual manner for biorthogonal expansions, i.e., $\langle \xi_a | \zeta_b \rangle = \delta_{a,b}$. Further we shall assume that the eigenvalues are numbered in decreasing order, so that $r_1 \geqslant r_2 \geqslant r_3 \geqslant r_4$. The concurrence is then defined by the formula

$$\begin{aligned} C &= \max\{0, \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3} - \sqrt{r_4}\} \\ &= \max\left\{0, \sum_{a=1}^{4} \mathrm{sgn}\left(\frac{3}{2} - a\right)\sqrt{r_a}\right\}, \end{aligned} \tag{5.32}$$

where $\mathrm{sgn}(x)=1$ if $x>0$ and $\mathrm{sgn}(x)=-1$ if $x<0$. The tangle is given by $T=C^2$ and the entanglement of formation by

$$E = h\left(\frac{1+\sqrt{1-C^2}}{2}\right), \tag{5.33}$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$. Because $h(x)$ is a monotonically increasing function, these three quantities are to some extent equivalent measures of the entanglement of a mixed state.

To calculate the errors in these rather complicated functions, we must employ the perturbation theory for non-Hermitian matrices (see Appendix C for more details). We need to evaluate the following partial derivative:

[5]The analysis in this subsection applies to the two-qubit case only. Measures of entanglement for mixed *n*-qubit systems are a subject of ongoing research: see, for example, [30] for a recent survey. It may be possible to measure entanglement directly, without quantum state tomography; this possibility was investigated in [31].

$$\frac{\partial C}{\partial s_\nu} = \sum_{a=1}^{4} \mathrm{sgn}\left(\frac{3}{2}-a\right)\frac{1}{2\sqrt{r_a}}\frac{\partial r_a}{\partial s_\nu}$$
$$= \sum_{a=1}^{4} \mathrm{sgn}\left(\frac{3}{2}-a\right)\frac{1}{2\sqrt{r_a}}\left\langle \xi_a \left| \frac{\partial \hat{R}}{\partial s_\nu} \right| \zeta_a \right\rangle$$
$$= \sum_{a=1}^{4}\sum_{\mu=1}^{16} \mathrm{sgn}\left(\frac{3}{2}-a\right)\frac{1}{2\sqrt{r_a}}\langle \xi_a|\hat{q}_{\mu,\nu}s_\mu|\zeta_a\rangle, \tag{5.34}$$

where the function $\mathrm{sgn}(x)$ is the sign of the quantity x: it takes the value 1 if $x>0$ and -1 if $x<0$. Thus $\mathrm{sgn}(3/2-a)$ is equal to $+1$ if $a=1$ and -1 if $a=2,3$, or 4. Hence the error in the concurrence is

$$(\Delta C)^2 = \sum_{\nu=1}^{16}\left(\frac{\partial C}{\partial s_\nu}\right)^2 \Lambda_\nu$$
$$= \sum_{\nu=1}^{16}\left[\sum_{a=1}^{4}\sum_{\mu=1}^{16} \times \mathrm{sgn}\left(\frac{3}{2}-a\right)\frac{1}{2\sqrt{r_a}}\langle \xi_a|\hat{q}_{\mu,\nu}s_\mu|\zeta_a\rangle\right]^2 \Lambda_\nu. \tag{5.35}$$

For our example the concurrence is 0.963 ± 0.018.

Once we know the error in the concurrence, the errors in the tangle and the entanglement of formation can be found straightforwardly:

$$\Delta T = 2C\Delta C, \tag{5.36}$$

$$\Delta E = \frac{C}{\sqrt{1-C^2}}h'\left(\frac{1+\sqrt{1-C^2}}{2}\right)\Delta C, \tag{5.37}$$

where $h'(x)$ is the derivative of $h(x)$. For our example the tangle is 0.928 ± 0.034 and the entanglement of formation is 0.947 ± 0.025.

VI. CONCLUSIONS

In conclusion, we have presented a technique for reconstructing density matrices of qubit systems, including a full error analysis. We have extended the latter through to calculation of quantities of interest in quantum information, such as the entropy and concurrence. Without loss of generality, we have used the example of polarization qubits of entangled photons, but we stress that these techniques can be adapted to any physical realization of qubits.

ACKNOWLEDGMENTS

The authors would like to thank Joe Altepeter, Mauro d'Ariano, Zdenek Hradil, Susana Huelga, Kurt Jacobs, Poul Jessen, James D. Malley, Michael Nielsen, Mike Raymer, Sze Tan, and Jaroslav Řeháček for useful discussions and correspondence. This work was supported in part by the U.S. National Security Agency, and Advanced Research and Development Activity (ARDA), by the Los Alamos National Laboratory LDRD program, and by the Australian Research Council.

APPENDIX A: THE $\hat{\Gamma}$ MATRICES

One possible set of $\hat{\Gamma}$ matrices are generators of $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$, normalized so that the conditions given in Eq. (3.8) are fulfilled. These matrices are

$$\hat{\Gamma}_1=\frac{1}{2}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad \hat{\Gamma}_2=\frac{1}{2}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix},$$

$$\hat{\Gamma}_3=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix},\quad \hat{\Gamma}_4=\frac{1}{2}\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix},$$

$$\hat{\Gamma}_5=\frac{1}{2}\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix},\quad \hat{\Gamma}_6=\frac{1}{2}\begin{pmatrix}0&0&0&-i\\0&0&i&0\\0&-i&0&0\\i&0&0&0\end{pmatrix},$$

$$\hat{\Gamma}_7=\frac{1}{2}\begin{pmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{pmatrix},\quad \hat{\Gamma}_8=\frac{1}{2}\begin{pmatrix}0&0&-i&0\\0&0&0&-i\\i&0&0&0\\0&i&0&0\end{pmatrix},$$

$$\hat{\Gamma}_9=\frac{1}{2}\begin{pmatrix}0&0&0&-i\\0&0&-i&0\\0&i&0&0\\i&0&0&0\end{pmatrix},\quad \hat{\Gamma}_{10}=\frac{1}{2}\begin{pmatrix}0&0&0&-1\\0&0&1&0\\0&1&0&0\\-1&0&0&0\end{pmatrix},$$

$$\hat{\Gamma}_{11}=\frac{1}{2}\begin{pmatrix}0&0&-i&0\\0&0&0&i\\i&0&0&0\\0&-i&0&0\end{pmatrix},\quad \hat{\Gamma}_{12}=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},$$

$$\hat{\Gamma}_{13}=\frac{1}{2}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{pmatrix},\quad \hat{\Gamma}_{14}=\frac{1}{2}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&i\\0&0&-i&0\end{pmatrix},$$

$$\hat{\Gamma}_{15}=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix},\quad \hat{\Gamma}_{16}=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}. \tag{A1}$$

As noted in the text, this is only one possible choice for these matrices, and the final results are independent of the choice.

APPENDIX B: THE $\hat{M}$ MATRICES AND SOME OF THEIR PROPERTIES

The $\hat{M}$ matrices, defined by Eq. (3.15), are as follows:

$$\hat{M}_1=\frac{1}{2}\begin{pmatrix}2&-(1-i)&-(1+i)&1\\-(1+i)&0&i&0\\-(1-i)&-i&0&0\\1&0&0&0\end{pmatrix},$$

$$\hat{M}_2=\frac{1}{2}\begin{pmatrix}0&-(1-i)&0&1\\-(1+i)&2&i&-(1+i)\\0&-i&0&0\\1&-(1+i)&0&0\end{pmatrix},$$

$$\hat{M}_3=\frac{1}{2}\begin{pmatrix}0&0&0&1\\0&0&i&-(1+i)\\0&-i&0&-(1-i)\\1&-(1-i)&-(1+i)&2\end{pmatrix},$$

$$\hat{M}_4=\frac{1}{2}\begin{pmatrix}0&0&-(1+i)&1\\0&0&i&0\\-(1-i)&-i&2&-(1-i)\\1&0&-(1+i)&0\end{pmatrix},$$

$$\hat{M}_5=\frac{1}{2}\begin{pmatrix}0&0&2i&-(1+i)\\0&0&(1-i)&0\\-2i&(1+i)&0&0\\-(1-i)&0&0&0\end{pmatrix},$$

$$\hat{M}_6=\frac{1}{2}\begin{pmatrix}0&0&0&-(1+i)\\0&0&(1-i)&2i\\0&(1+i)&0&0\\-(1-i)&-2i&0&0\end{pmatrix},$$

$$\hat{M}_7=\frac{1}{2}\begin{pmatrix}0&0&0&-(1+i)\\0&0&-(1-i)&2\\0&-(1+i)&0&0\\-(1-i)&2&0&0\end{pmatrix},$$

$$\hat{M}_8=\frac{1}{2}\begin{pmatrix}0&0&2&-(1+i)\\0&0&-(1-i)&0\\2&-(1+i)&0&0\\-(1-i)&0&0&0\end{pmatrix}, \tag{B1}$$

$$\hat{M}_9=\begin{pmatrix}0&0&0&i\\0&0&-i&0\\0&i&0&0\\-i&0&0&0\end{pmatrix},\quad \hat{M}_{10}=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix},$$

$$\hat{M}_{11}=\begin{pmatrix}0&0&0&i\\0&0&i&0\\0&-i&0&0\\-i&0&0&0\end{pmatrix},$$

$$\hat{M}_{12}=\frac{1}{2}\begin{pmatrix}0&2&0&-(1+i)\\2&0&-(1+i)&0\\0&-(1-i)&0&0\\-(1-i)&0&0&0\end{pmatrix},$$

$$\hat{M}_{13}=\frac{1}{2}\begin{pmatrix}0&0&0&-(1+i)\\0&0&-(1+i)&0\\0&-(1-i)&0&2\\-(1-i)&0&2&0\end{pmatrix},$$

$$\hat{M}_{14}=\frac{1}{2}\begin{pmatrix}0&0&0&-(1-i)\\0&0&-(1-i)&0\\0&-(1+i)&0&-2i\\-(1+i)&0&2i&0\end{pmatrix},$$

$$\hat{M}_{15}=\frac{1}{2}\begin{pmatrix}0&-2i&0&-(1-i)\\2i&0&(1-i)&0\\0&(1+i)&0&0\\-(1+i)&0&0&0\end{pmatrix},$$

$$\hat{M}_{16}=\begin{pmatrix}0&0&0&1\\0&0&-1&0\\0&-1&0&0\\1&0&0&0\end{pmatrix}.$$

The form of these matrices is independent of the chosen set of matrices $\{\hat{\Gamma}_\nu\}$ used to convert the density matrix into a column vector. However, the $\hat{M}_\nu$ matrices *do* depend on the set of tomographic states $|\psi_\nu\rangle$.

There are some useful properties of these matrices which we will now derive. From Eq. (3.15), we have

$$\langle \psi_\mu|\hat{M}_\nu|\psi_\mu\rangle = \sum_\lambda \langle \psi_\mu|\hat{\Gamma}_\lambda|\psi_\mu\rangle (B^{-1})_{\lambda,\nu}. \tag{B2}$$

From Eq. (3.12) we have $\langle \psi_\mu|\hat{\Gamma}_\lambda|\psi_\mu\rangle = B_{\mu,\lambda}$; thus we obtain the result

$$\langle \psi_\mu|\hat{M}_\nu|\psi_\mu\rangle = \delta_{\mu,\nu}. \tag{B3}$$

If we denote the basis set for the four-dimensional Hilbert space by $\{|i\rangle\ (i=1,2,3,4)\}$, then Eq. (3.14) can be written as follows:

$$\langle i|\hat{\rho}|j\rangle = \sum_{k,l}\sum_\nu \langle i|\hat{M}_\nu|j\rangle\langle\psi_\nu|k\rangle\langle l|\psi_\nu\rangle\langle k|\hat{\rho}|l\rangle. \tag{B4}$$

Since Eq. (B4) is valid for arbitrary states $\hat{\rho}$, we obtain the following relationship:

$$\sum_\nu \langle i|\hat{M}_\nu|j\rangle\langle\psi_\nu|k\rangle\langle l|\psi_\nu\rangle = \delta_{ik}\delta_{jl}. \tag{B5}$$

Contracting Eq. (B5) over the indices (i,j) we obtain

$$\sum_\nu \mathrm{Tr}\{\hat{M}_\nu\}|\psi_\nu\rangle\langle\psi_\nu| = \hat{I}, \tag{B6}$$

where $\hat{I}$ is the identity operator for our four-dimensional Hilbert space.

A second relationship can be obtained by contracting Eq. (B5), viz.,

$$\sum_\nu \langle i|\hat{M}_\nu|j\rangle = \delta_{ij}, \tag{B7}$$

or, in operator notation,

$$\sum_\nu \hat{M}_\nu = \hat{I}. \tag{B8}$$

APPENDIX C: PERTURBATION THEORY FOR NON-HERMITIAN MATRICES

Whereas perturbation theory for Hermitian matrices is covered in most quantum mechanics textbooks, the case of non-Hermitian matrices is less familiar, and so we will present it here. The problem is as follows. Given the eigenspectrum of a matrix $\hat{R}_0$ [32], i.e.,

$$\langle \xi_a|\hat{R}_0 = r_a\langle\xi_a|, \tag{C1}$$

$$\hat{R}_0|\zeta_a\rangle = r_a|\zeta_a\rangle, \tag{C2}$$

where

$$\langle \xi_a|\zeta_b\rangle = \delta_{a,b}, \tag{C3}$$

we wish to find expressions for the eigenvalues r'_a and eigenstates $\langle\xi'_a|$ and $|\zeta'_a\rangle$ of the perturbed matrix $\hat{R}' = \hat{R}_0 + \delta\hat{R}$.

We start with the standard assumption of perturbation theory, i.e., that the perturbed quantities r'_a, $\langle\xi'_a|$, and $|\zeta'_a\rangle$ can be expressed as power series of some parameter λ:

$$r'_a = r_a^{(0)} + \lambda r_a^{(1)} + \lambda^2 r_a^{(2)} + \cdots, \tag{C4}$$

$$|\zeta'_a\rangle = |\zeta_a^{(0)}\rangle + \lambda|\zeta_a^{(1)}\rangle + \lambda^2|\zeta_a^{(2)}\rangle + \cdots, \tag{C5}$$

$$\langle\xi'_a| = \langle\xi_a^{(0)}| + \lambda\langle\xi_a^{(1)}| + \lambda^2\langle\xi_a^{(2)}| + \cdots, \tag{C6}$$

Writing $\hat{R}' = \hat{R}_0 + \lambda\,\delta\hat{R}$, and comparing terms of equal powers of λ in the eigenequations, one obtains the following formulas:

$$\hat{R}_0|\zeta_a^{(0)}\rangle = r_a^{(0)}|\zeta_a^{(0)}\rangle, \tag{C7}$$

$$\langle\xi_a^{(0)}|\hat{R}_0 = r_a^{(0)}\langle\xi_a^{(0)}|, \tag{C8}$$

$$(\hat{R}_0 - r_a^{(0)}\hat{I})|\zeta_a^{(1)}\rangle = -(\delta\hat{R} - r_a^{(1)})|\zeta_a^{(0)}\rangle, \tag{C9}$$

$$\langle\xi_a^{(1)}|(\hat{R}_0 - r_a^{(0)}\hat{I}) = -\langle\xi_a^{(0)}|(\delta\hat{R} - r_a^{(1)}). \tag{C10}$$

Equations (C7) and (C8) imply that, as might be expected,

$$|\zeta_a^{(0)}\rangle = |\zeta_a\rangle, \tag{C11}$$

$$\langle\xi_a^{(0)}| = \langle\xi_a|, \tag{C12}$$

$$r_a^{(0)} = r_a. \tag{C13}$$

Taking the inner product of Eq. (C9) with $\langle\xi_a|$, and using the biorthogonal property Eq. (C3), we obtain

$$r_a^{(1)} = \langle\xi_a|\delta\hat{R}|\zeta_a\rangle. \tag{C14}$$

This implies that

$$\delta r_a \equiv r'_a - r_a \approx \langle\xi_a|\delta\hat{R}|\zeta_a\rangle. \tag{C15}$$

Thus, dividing both sides by some differential increment δx and taking the limit $\delta x \to 0$, we obtain

$$\frac{\partial r_a}{\partial x} = \left\langle \xi_a \middle| \frac{\partial\hat{R}}{\partial x} \middle| \zeta_a \right\rangle. \tag{C16}$$

Using the completeness property of the eigenstates, $\Sigma_b|\zeta_b\rangle\langle\xi_b| = \hat{I}$, and the identity $\hat{R}_0 = \Sigma_b r_b|\zeta_b\rangle\langle\xi_b|$, we obtain the following formula

$$(\hat{R}_0 - r_a\hat{I})^{-1} = \sum_{\substack{b\\ b\neq a}} \frac{1}{r_b - r_a}|\zeta_b\rangle\langle\xi_b|. \tag{C17}$$

Applying this to Eq. (C9) we obtain

$$|\delta\zeta_a^{(1)}\rangle \equiv |\zeta'_a\rangle - |\zeta_a\rangle \approx -\sum_{\substack{b\\ b\neq a}} \left(\frac{\langle\xi_b|\delta\hat{R}|\zeta_a\rangle}{r_b - r_a}\right)|\zeta_b\rangle. \tag{C18}$$

Similarly, Eqs. (C10) and (C17) imply

$$\langle\delta\xi_a| \equiv \langle\delta\xi'_a| - \langle\delta\xi_a| \approx -\sum_{\substack{b\\ b\neq a}} \left(\frac{\langle\xi_a|\delta\hat{R}|\zeta_b\rangle}{r_b - r_a}\right)\langle\xi_b|. \tag{C19}$$

[1] G.C. Stokes, Trans. Cambridge Philos. Soc. **9**, 399 (1852).
[2] U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, Cambridge, 1997).
[3] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).
[4] S.M. Tan, J. Mod. Opt. **44**, 2233 (1997).
[5] K. Banaszek, G.M. D'Ariano, M.G.A. Paris, and M.F. Sacchi, Phys. Rev. A **61**, 010304 (1999).
[6] Z. Hradil, J. Summhammer, G. Badurek, and H. Rauch, Phys. Rev. A **62**, 014101 (2000).
[7] J. Řeháček, Z. Hradil, and M. Ježek, Phys. Rev. A **63**, 040303 (2001).
[8] J.R. Ashburn, R.A. Cline, P.J.M. Vanderburgt, W.B. Westerveld, and J.S. Risley, Phys. Rev. A **41**, 2407 (1990).
[9] D.T. Smithey, M. Beck, M.G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993); see also the discussion of polarization effects in M.G. Raymer, D.F. McAlister, and A. Funk, in *Quantum Communication, Computing, and Measurement '98*, edited by P. Kumar (Plenum, New York, 2000), pp. 147-162.
[10] G. Breitenbach, S. Schiller, and J. Mlynek, Nature (London) **387**, 471 (1997).
[11] J.W. Wu, P.K. Lam, M.B. Gray, and H.-A. Bachor, Opt. Express **3**, 154 (1998).
[12] T.J. Dunn, I.A. Walmsley, and S. Mukamel, Phys. Rev. Lett. **74**, 884 (1995).
[13] D. Leibfried, D.M. Meekhof, B.E. King, C. Monroe, W.M. Itano, and D.J. Wineland, Phys. Rev. Lett. **77**, 4281 (1996); D. Leibfried, T. Pfau, and C. Monroe, Phys. Today **51**(4), 22 (1998).
[14] C. Kurtsiefer T. Pfau, and J. Mlynek, Nature (London) **386**, 150-153 (1997).
[15] G. Klose, G. Smith, and P.S. Jessen, Phys. Rev. Lett. **86**, 4721 (2001).
[16] I.L. Chuang, N. Gershenfeld, and M. Kubinec, Phys. Rev. Lett. **80**, 3408 (1998).
[17] S.L. Braunstein, C.M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack, Phys. Rev. Lett. **83**, 1054 (1999).
[18] A.G. White, D.F.V. James, P.H. Eberhard, and P.G. Kwiat, Phys. Rev. Lett. **83**, 3103 (1999).
[19] A recent overview of many quantum computation technologies is given by S. Braunstein and Ho.-K. Lo, *Scalable Quantum Computers: Paving the Way to Realization* (Wiley, New York, 2001); see also Fortschr. Phys. **48**, (9–11) (2000).
[20] E. Hecht and A. Zajac, *Optics* (Addision-Wesley, Reading, MA, 1974), Sec. 8.12.
[21] E. Wolf, Nuovo Cimento **13**, 1165 (1959); L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995), Chap. 6.
[22] P.G. Kwiat, E. Waks, A.G. White, I. Appelbaum, and P.H. Eberhard, Phys. Rev. A **60**, R773 (1999).
[23] A.G. White, D.F.V. James, W.J. Munro and P.G. Kwiat, e-print quant-ph/0108088.
[24] A. Berglund, Undergraduate thesis, Dartmouth College, 2000; A. Berglund, e-print quant-ph/0010001.
[25] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992), Sec. 10.5.
[26] A.C. Melissinos, *Experiments in Modern Physics* (Academic Press, New York, 1966), Sec, 10.4, pp. 467–479.
[27] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000), Chap. 11.
[28] L.I. Schiff, *Quantum Mechanics*, 3rd ed. (McGraw-Hill, New York, 1968), Eq. (31.8), p. 246.
[29] W.K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998); V. Coffman, J. Kundu, and W.K. Wootters, Phys. Rev. A **61**, 052306 (2000).
[30] B.M. Terhal, e-print quant-ph/0101032.
[31] J.M.G. Sancho and S.F. Huelga, Phys. Rev. A **61**, 042303 (2000).
[32] The properties of the eigenvectors and eigenvalues of non-Hermitian matrices are discussed by P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I, p. 884 *et seq*.