



## On the Measurement of Qubits

Daniel F.V. James<sup>1</sup>, Paul G. Kwiat<sup>2</sup>, William J. Munro, Andrew G. White<sup>3</sup>

Trusted E-Services Laboratory

HP Laboratories Bristol

HPL-2001-160

July 2<sup>nd</sup>, 2001\*

quantum  
tomography,  
state  
reconstruction,  
maximum  
likelihood  
methods,  
tomographic  
reconstruction

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 which 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.

\* Internal Accession Date Only

Approved for External Publication

<sup>1</sup> Theory Division, T-4, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

<sup>2</sup> Dept. of Physics, University of Illinois, Urbana-Champaign, Illinois, USA.

<sup>3</sup> Special Research Centre for Quantum Computer Technology, University of Queensland, Brisbane, Australia

© Copyright Hewlett-Packard Company 2001

# On the Measurement of Qubits

D. F. V. James<sup>1</sup>, P. G. Kwiat<sup>2</sup>, W. J. Munro<sup>3,4</sup> and A. G. White<sup>4</sup>

<sup>1</sup> *Theory Division, T-4, Los Alamos National Laboratory, Los Alamos, New Mexico, USA,*

<sup>2</sup> *Dept. of Physics, University of Illinois, Urbana-Champaign, Illinois, USA.,*

<sup>3</sup> *Hewlett-Packard Laboratories, Filton Road, Stoke Gifford, Bristol BS34 8QZ, UK, and*

<sup>4</sup> *Special Research Centre for Quantum Computer Technology,  
University of Queensland, Brisbane, AUSTRALIA*

(Dated: June 20, 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 which 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.

PACS numbers:

## 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 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 semi-definite matrices with unit trace. The tomographically measured matrices often fail to be positive semi-definite, 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 Risley et al. at North Carolina State University 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. (University of Oregon), 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 Mlynek et al. (University of Konstanz, Germany) [10] and Bachor et al. (Australian National University) [11]. Other quantum systems whose density matrices have been investigated experimentally include the vibrations of molecules [12], the motion 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-half 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 techniques in detail 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 emphasis 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 Section 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 Section 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. Section IV deals with our method for maximum likelihood reconstruction and in Section V we demonstrate how to calculate the errors in such measurements, and how these errors propagate to quantities calculated from the density matrix.

## 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.

### 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  $45^\circ$  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 are as follows:

$$\begin{aligned}
 n_0 &= \frac{\mathcal{N}}{2} (\langle H|\hat{\rho}|H\rangle + \langle V|\hat{\rho}|V\rangle) = \frac{\mathcal{N}}{2} (\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle) \\
 n_1 &= \mathcal{N} (\langle H|\hat{\rho}|H\rangle) \\
 &= \frac{\mathcal{N}}{2} (\langle R|\hat{\rho}|R\rangle + \langle R|\hat{\rho}|L\rangle + \langle L|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle) \\
 n_2 &= \mathcal{N} (\langle \bar{D}|\hat{\rho}|\bar{D}\rangle) \\
 &= \frac{\mathcal{N}}{2} (\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle - i\langle L|\hat{\rho}|R\rangle + i\langle R|\hat{\rho}|L\rangle) \\
 n_3 &= \mathcal{N} (\langle R|\hat{\rho}|R\rangle)
 \end{aligned} \tag{0.1}$$

Here  $|H\rangle$ ,  $|V\rangle$ ,  $|\bar{D}\rangle = (|H\rangle - |V\rangle)/\sqrt{2} = \exp(i\pi/4) (|R\rangle + i|L\rangle)/\sqrt{2}$  and  $|R\rangle = (|H\rangle - i|V\rangle)/\sqrt{2}$  are the kets representing qubits polarized in the linear horizontal, linear vertical, linear diagonal ( $45^\circ$ ) and right-circular senses respectively,  $\hat{\rho}$  is the  $(2 \times 2)$  density matrix for the polarization degrees of the light (or, for a two-level quantum system) and  $\mathcal{N}$  is 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

$$\begin{aligned}
 \mathcal{S}_0 &\equiv 2n_0 = \mathcal{N} (\langle R|\hat{\rho}|R\rangle + \langle L|\hat{\rho}|L\rangle) \\
 \mathcal{S}_1 &\equiv 2(n_1 - n_0) = \mathcal{N} (\langle R|\hat{\rho}|L\rangle + \langle L|\hat{\rho}|R\rangle) \\
 \mathcal{S}_2 &\equiv 2(n_2 - n_0) = \mathcal{N}i (\langle R|\hat{\rho}|L\rangle - \langle L|\hat{\rho}|R\rangle) \\
 \mathcal{S}_3 &\equiv 2(n_3 - n_0) = \mathcal{N} (\langle R|\hat{\rho}|R\rangle - \langle L|\hat{\rho}|L\rangle) .
 \end{aligned} \tag{0.2}$$

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

$$\hat{\rho} = \frac{1}{2} \sum_{i=0}^3 \frac{\mathcal{S}_i}{\mathcal{S}_0} \hat{\sigma}_i, \quad (0.3)$$

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.

### 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 importance 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 photon one has the possibility of non-classical 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:

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1, i_2, \dots, i_n=0}^3 r_{i_1, i_2, \dots, i_n} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \dots \otimes \hat{\sigma}_{i_n}, \quad (0.4)$$

where the  $4^n$  parameters  $r_{i_1, i_2, \dots, i_n}$  are real numbers. The normalization property of the density matrices requires that  $r_{0,0,\dots,0} = 1$ , and so the density matrix is specified by  $4^n - 1$  real parameters. The symbol  $\otimes$  represents the tensor 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  $\hat{\mu}_0 = |H\rangle\langle H| + |V\rangle\langle V|$ ,  $\hat{\mu}_1 = |H\rangle\langle H|$ ,  $\hat{\mu}_2 = |\bar{D}\rangle\langle \bar{D}|$ ,  $\hat{\mu}_3 = |R\rangle\langle R|$ . 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  $\hat{\mu}_{i_1} \otimes \hat{\mu}_{i_2} \otimes \dots \otimes \hat{\mu}_{i_n}$  ( $i_k = 0, 1, 2, 3$ ) and ( $k = 1, 2, \dots, n$ ). This ‘tree’ structure for multi-qubit measurement is illustrated in fig.1.

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

$$n = \mathcal{N} \text{Tr} \{ \hat{\rho} \hat{\mu} \}, \quad (0.5)$$

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

$$n_{i_1, i_2, \dots, i_n} = \mathcal{N} \text{Tr} \{ \hat{\rho} (\hat{\mu}_{i_1} \otimes \hat{\mu}_{i_2} \otimes \dots \otimes \hat{\mu}_{i_n}) \}. \quad (0.6)$$

Substituting from eq.(0.4) we obtain

$$n_{i_1, i_2, \dots, i_n} = \frac{\mathcal{N}}{2^n} \sum_{j_1, j_2, \dots, j_n=0}^3 \text{Tr} \{ \hat{\mu}_{i_1} \hat{\sigma}_{j_1} \} \text{Tr} \{ \hat{\mu}_{i_2} \hat{\sigma}_{j_2} \} \dots \text{Tr} \{ \hat{\mu}_{i_n} \hat{\sigma}_{j_n} \} r_{i_1, i_2, \dots, i_n}. \quad (0.7)$$

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.  $\hat{\mu}_i = \sum_{j=0}^3 \Upsilon_{ij} \hat{\sigma}_j$ , where  $\Upsilon_{ij}$  are the elements of the matrix

$$\Upsilon = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}. \quad (0.8)$$

Further, we have the relation  $\text{Tr} \{ \hat{\sigma}_i \hat{\sigma}_j \} = 2\delta_{ij}$  (where  $\delta_{ij}$  is the Kronecker delta). Hence eq.(0.7) becomes

$$n_{i_1, i_2, \dots, i_n} = \mathcal{N} \sum_{j_1, j_2, \dots, j_n=0}^3 \Upsilon_{i_1 j_1} \Upsilon_{i_2 j_2} \dots \Upsilon_{i_n j_n} r_{i_1, i_2, \dots, i_n}. \quad (0.9)$$

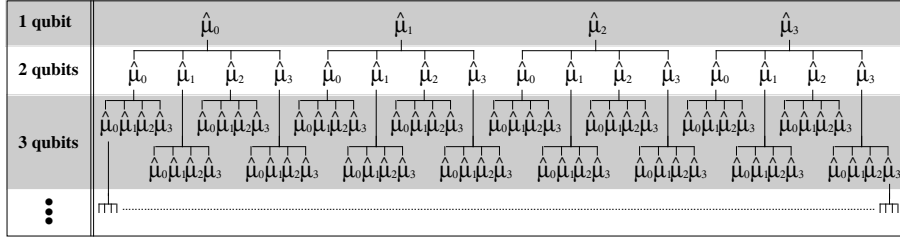


FIG. 1: Tree diagram representing number and type of measurements necessary for tomography. For a single qubit, the measurements  $\{\hat{\mu}_0, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3\}$  suffice to reconstruct the state, e.g., measurements of the horizontal, vertical, diagonal and right-circular polarization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

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

$$\Upsilon^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix}, \quad (0.10)$$

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

$$\begin{aligned} \mathcal{N} r_{i_1, i_2, \dots, i_n} &= \sum_{j_1, j_2, \dots, j_n=0}^3 (\Upsilon^{-1})_{i_1 j_1} (\Upsilon^{-1})_{i_2 j_2} \dots (\Upsilon^{-1})_{i_n j_n} n_{i_1, i_2, \dots, i_n} \\ &\equiv \mathcal{S}_{i_1, i_2, \dots, i_n}. \end{aligned} \quad (0.11)$$

In eq.(0.11) we have introduced the n-photon Stokes parameter  $\mathcal{S}_{i_1, i_2, \dots, i_n}$ , defined in an analogous manner to the single photon Stokes parameters give in eq.(0.2).

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

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1, i_2, \dots, i_n=0}^3 \frac{\mathcal{S}_{i_1, i_2, \dots, i_n}}{\mathcal{S}_{0,0,\dots,0}} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \dots \otimes \hat{\sigma}_{i_n}. \quad (0.12)$$

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  $\{\hat{\mu}_0, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3\}$  are not unique: there may be circumstances in which it is more convenient to use some other set, which are 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.

## GENERALIZED TOMOGRAPHIC RECONSTRUCTION OF THE POLARIZATION STATE OF TWO PHOTONS

### Experimental set-up

The experimental arrangement used in our experiments is shown schematically in Fig.1. An optical system consisting of lasers, polarization elements and non-linear optical crystals (and 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 ref.[22–24] [32] The output of the black box consists of a pair of beams of light, whose quanta can be measured by means of photo-detectors. 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 waveplates 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 following convention,

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = |V\rangle, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |H\rangle, \quad (0.1)$$

where  $|V\rangle$  ( $|H\rangle$ ) is the ket for a vertically (horizontally) polarized beam, the effect 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  $2 \times 2$  matrices

$$\begin{aligned} \hat{U}_{QWP}(q) &= \frac{1}{\sqrt{2}} \begin{pmatrix} i - \cos(2q) & \sin(2q) \\ \sin(2q) & i + \cos(2q) \end{pmatrix}, \\ \hat{U}_{HWP}(q) &= \begin{pmatrix} \cos(2h) & -\sin(2h) \\ -\sin(2h) & -\cos(2h) \end{pmatrix}. \end{aligned} \quad (0.2)$$

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

$$\begin{aligned} |\psi_{proj}^{(1)}(h, q)\rangle &= \hat{U}_{QWP}(q) \cdot \hat{U}_{HWP}(h) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= a(h, q)|H\rangle + b(h, q)|V\rangle, \end{aligned} \quad (0.3)$$

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

$$\begin{aligned} a(h, q) &= \frac{1}{\sqrt{2}} (\sin(2h) - i \sin[2(h - q)]), \\ b(h, q) &= -\frac{1}{\sqrt{2}} (\cos(2h) + i \cos[2(h - q)]). \end{aligned} \quad (0.4)$$

The projection state for the two beams is given by

$$\begin{aligned} |\psi_{proj}^{(2)}(h_1, q_1, h_2, q_2)\rangle &= |\psi_{proj}^{(1)}(h_1, q_1)\rangle \otimes |\psi_{proj}^{(1)}(h_2, q_2)\rangle \\ &= a(h_1, q_1)a(h_2, q_2)|HH\rangle + a(h_1, q_1)b(h_2, q_2)|HV\rangle + \\ &\quad b(h_1, q_1)a(h_2, q_2)|VH\rangle + b(h_1, q_1)b(h_2, q_2)|VV\rangle. \end{aligned} \quad (0.5)$$

We shall denote the projection state corresponding to one particular set of waveplate angles  $\{h_{1,\nu}, q_{1,\nu}, h_{2,\nu}, q_{2,\nu}\}$  [33] by the ket  $|\psi_\nu\rangle$ ; 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

$$n_\nu = \mathcal{N} \langle \psi_\nu | \hat{\rho} | \psi_\nu \rangle \quad (0.6)$$

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  $s_\nu$  defined by

$$s_\nu = \langle \psi_\nu | \hat{\rho} | \psi_\nu \rangle. \quad (0.7)$$

### Tomographically Complete set of Measurements

In Section 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 which determine a  $4 \times 4$  density matrix, plus there is the single unknown real parameter  $\mathcal{N}$ , making a total of 16.

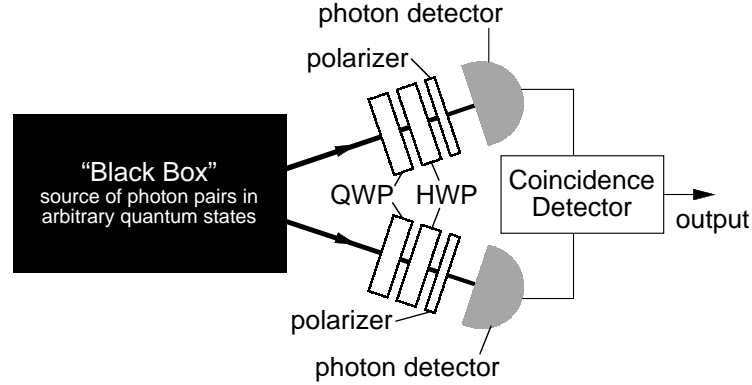


FIG. 2: Schematic illustration of the experimental arrangement. “QWP” stands for quarter waveplate, “HWP” for half waveplate; the angles of both pairs of wave plates can be set independently giving the experimenter four degrees of freedom with which to set the projection state. In the experiment, the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

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

$$\begin{aligned} \text{Tr} \left\{ \hat{\Gamma}_\nu \cdot \hat{\Gamma}_\mu \right\} &= \delta_{\nu,\mu} \\ \hat{A} &= \sum_{\nu=1}^{16} \hat{\Gamma}_\nu \text{Tr} \left\{ \hat{\Gamma}_\nu \cdot \hat{A} \right\} \quad \forall \hat{A}, \end{aligned} \quad (0.8)$$

where  $\hat{A}$  is an arbitrary  $4 \times 4$  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 re-labeling 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

$$\hat{\rho} = \sum_{\nu=1}^{16} \hat{\Gamma}_\nu r_\nu, \quad (0.9)$$

where  $r_\nu$  is the  $\nu$ -th element of a sixteen element column vector, given by the formula

$$r_\nu = \text{Tr} \left\{ \hat{\Gamma}_\nu \cdot \hat{\rho} \right\} \quad (0.10)$$

Substituting from eq.(0.9) into eq.(0.6), we obtain the following linear relationship between the measured coincidence counts  $n_\nu$  and the elements of the vector  $r_\mu$ :

$$n_\nu = \mathcal{N} \sum_{\mu=1}^{16} B_{\nu,\mu} r_\mu \quad (0.11)$$

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

$$B_{\nu,\mu} = \langle \psi_\nu | \hat{\Gamma}_\mu | \psi_\nu \rangle. \quad (0.12)$$

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 non-singular, then eq.(0.11) can be inverted to give

$$r_\nu = (\mathcal{N})^{-1} \sum_{\mu=1}^{16} (B^{-1})_{\nu,\mu} n_\mu. \quad (0.13)$$

The set of sixteen tomographic states which we employed are given in Table 1. They can be shown to satisfy the condition that  $B_{\nu,\mu}$  is non-singular. By no means are these states unique in this regard: these were the states chosen principally for experimental convenience.

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

**Table 1**

TABLE 1: The tomographic analysis states used in our experiments. The number of coincidence counts measured in projections measurements provide 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 measurement are taken in the order given by the table, only one waveplate angle had to be changed between each measurement.

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 1. Note that overall phase factors do not affect the results of projection measurements.

Substituting eq.(0.13) into eq.(0.9), we find that

$$\begin{aligned}\hat{\rho} &= (\mathcal{N})^{-1} \sum_{\nu=1}^{16} \hat{M}_{\nu} n_{\nu} \\ &= \sum_{\nu=1}^{16} \hat{M}_{\nu} s_{\nu},\end{aligned}\tag{0.14}$$

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

$$\hat{M}_{\nu} = \sum_{\mu=1}^{16} (B^{-1})_{\nu,\mu} \hat{\Gamma}_{\mu}.\tag{0.15}$$

The introduction of the  $\hat{M}_{\nu}$  matrices allows a compact form of linear tomographic reconstruction, eq.(0.14), 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.(B.6), to obtain the value of the unknown quantity  $\mathcal{N}$ . That relationship implies

$$\sum_{\nu} \text{Tr} \{ \hat{M}_{\nu} \} |\psi_{\nu}\rangle \langle \psi_{\nu}| \hat{\rho} = \hat{\rho}.\tag{0.16}$$



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

$$\sum_{\nu} \text{Tr} \{ \hat{M}_{\nu} \} n_{\nu} = \mathcal{N}. \quad (0.17)$$

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

$$\sum_{\nu} \text{Tr} \{ \hat{M}_{\nu} \} = \begin{cases} 1 & \text{if } \nu = 1, 2, 3, 4 \\ 0 & \text{if } \nu = 5, \dots, 16, \end{cases} \quad (0.18)$$

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

$$\begin{aligned} \mathcal{N} &= \sum_{\nu=1}^4 n_{\nu} \\ &= \mathcal{N} (\langle HH | \hat{\rho} | HH \rangle + \langle HV | \hat{\rho} | HV \rangle + \langle VH | \hat{\rho} | VH \rangle + \langle VV | \hat{\rho} | VV \rangle). \end{aligned} \quad (0.19)$$

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

$$\hat{\rho} = \left( \sum_{\nu=1}^{16} \hat{M}_{\nu} n_{\nu} \right) / \left( \sum_{\nu=1}^4 n_{\nu} \right). \quad (0.20)$$

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:  $n_1 = 34749, n_2 = 324, n_3 = 35805, n_4 = 444, n_5 = 16324, n_6 = 17521, n_7 = 13441, n_8 = 16901, n_9 = 17932, n_{10} = 32028, n_{11} = 15132, n_{12} = 17238, n_{13} = 13171, n_{14} = 17170, n_{15} = 16722, n_{16} = 33586$  Applying eq.(0.20) we find

$$\hat{\rho} = \begin{pmatrix} 0.4872 & -0.0042 + i0.0114 & -0.0098 - i0.0178 & 0.5192 + i0.0380 \\ -0.0042 - i0.0114 & 0.0045 & 0.0271 - i0.0146 & -0.0648 - i0.0076 \\ -0.0098 + i0.0178 & 0.0271 + i0.0146 & 0.0062 & -0.0695 + i0.0134 \\ 0.5192 - i0.0380 & -0.0648 + i0.0076 & -0.0695 - i0.0134 & 0.5020 \end{pmatrix} \quad (0.21)$$

This matrix is shown graphically in fig.3a.

Note that, by construction, the density matrix is normalized, i.e.  $\text{Tr}\{\hat{\rho}\} = 1$  and Hermitian, i.e.  $\hat{\rho}^{\dagger} = \hat{\rho}$ . However, when one calculates the eigenvalues of this measured density matrix, one finds the values 1.02155, 0.0681238, -0.065274 and -0.024396; and also,  $\text{Tr}\{\hat{\rho}^2\} = 1.053$ . Density matrices for all physical states must have the property of positive semi-definiteness, which (in conjunction with the normalization and Hermiticity properties) imply that all of the eigenvalues must lie in the interval  $[0, 1]$ , their sum being 1; this in turn implies that  $0 \leq \text{Tr}\{\hat{\rho}^2\} \leq 1$ . Clearly the density matrix reconstructed above by linear tomography violates these 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.(0.6). 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.

### MAXIMUM LIKELIHOOD ESTIMATION

As mentioned in Section III, the tomographic measurement of density matrices can produce results which 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.

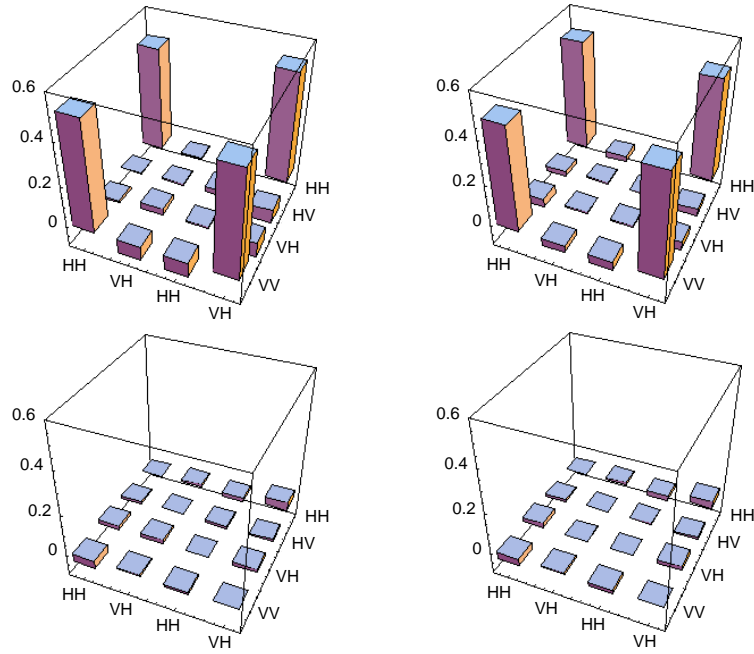


FIG. 3: Graphical representation of the density matrix of a state as estimated by linear tomography (left) and by maximum likelihood tomography (right) from the experimental data given in the text. The upper plot is the real part of  $\hat{\rho}$ , the lower plot the imaginary part.

### 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 which has the three important properties of normalization, Hermiticity and positivity. This matrix will be a function of 16 real variables (denoted  $\{t_1, t_2, \dots, t_{16}\}$ ). We will denote the matrix as  $\hat{\rho}_p(t_1, t_2, \dots, t_{16})$ .
- (ii) Introduce a “likelihood function” which quantifies how good the density matrix  $\hat{\rho}_p(t_1, t_2, \dots, t_{16})$  is in relation to the experimental data. This likelihood function is a function of the 16 real parameters  $t_\nu$  and of the 16 experimental data  $n_\nu$ . We will denote this function as  $\mathcal{L}(t_1, t_2, \dots, t_{16}; n_1, n_2, \dots, n_{16})$ .
- (iii) Using standard numerical optimization techniques, find the optimum set of variables  $\{t_1^{(opt)}, t_2^{(opt)}, \dots, t_{16}^{(opt)}\}$  for which the function  $\mathcal{L}(t_1, t_2, \dots, t_{16}; n_1, n_2, \dots, n_{16})$  has its maximum value. The best estimate for the density matrix is then  $\hat{\rho}(t_1^{(opt)}, t_2^{(opt)}, \dots, t_{16}^{(opt)})$ .

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

### Physical Density Matrices

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

$$\langle \psi | \hat{\mathcal{G}} | \psi \rangle \geq 0 \quad \forall |\psi\rangle. \quad (0.1)$$

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.(0.1):

$$\langle \psi | \hat{T}^\dagger \hat{T} | \psi \rangle = \langle \psi' | \psi' \rangle \geq 0, \quad (0.2)$$

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

$$\hat{g} = \hat{T}^\dagger \hat{T} / \text{Tr}\{\hat{T}^\dagger \hat{T}\} \quad (0.3)$$

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

For the two qubit system, we have a  $4 \times 4$  density matrix with 15 independent real parameters. Since it will be useful to be able to invert relation (0.3), it is convenient to choose a tri-diagonal form for  $\hat{T}$ :

$$\hat{T}(t) = \begin{pmatrix} t_1 & 0 & 0 & 0 \\ t_5 + it_6 & t_2 & 0 & 0 \\ t_{11} + it_{12} & t_7 + it_8 & t_3 & 0 \\ t_{15} + it_{16} & t_{13} + it_{14} & t_9 + it_{10} & t_4 \end{pmatrix}, \quad (0.4)$$

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

$$\hat{\rho}_p(t) = \hat{T}^\dagger(t)\hat{T}(t)/\text{Tr}\{\hat{T}^\dagger(t)\hat{T}(t)\}. \quad (0.5)$$

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:

$$\hat{T} = \begin{pmatrix} \sqrt{\frac{\Delta}{\mathcal{M}_{11}^{(1)}}} & 0 & 0 & 0 \\ \frac{\mathcal{M}_{12}^{(1)}}{\sqrt{\mathcal{M}_{11}^{(1)}\mathcal{M}_{11,22}^{(2)}}} & \sqrt{\frac{\mathcal{M}_{11}^{(1)}}{\mathcal{M}_{11,22}^{(2)}}} & 0 & 0 \\ \frac{\mathcal{M}_{12,23}^{(2)}}{\sqrt{\rho_{44}}\sqrt{\mathcal{M}_{11,22}^{(2)}}} & \frac{\mathcal{M}_{11,23}^{(2)}}{\sqrt{\rho_{44}}\sqrt{\mathcal{M}_{11,22}^{(2)}}} & \sqrt{\frac{\mathcal{M}_{11,22}^{(2)}}{\rho_{44}}} & 0 \\ \frac{\rho_{41}}{\sqrt{\rho_{44}}} & \frac{\rho_{42}}{\sqrt{\rho_{44}}} & \frac{\rho_{43}}{\sqrt{\rho_{44}}} & \sqrt{\rho_{44}} \end{pmatrix}. \quad (0.6)$$

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

### The Likelihood Function

The measurement data consists of a set of 16 coincidence counts  $n_\nu$  ( $\nu = 1, 2, \dots, 16$ ) whose expected value is  $\bar{n}_\nu = \mathcal{N}\langle\psi_\nu|\hat{\rho}|\psi_\nu\rangle$ . 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  $\{n_1, n_2, \dots, n_{16}\}$  is

$$P(n_1, n_2, \dots, n_{16}) = \frac{1}{\text{Norm}} \prod_{\nu=1}^{16} \exp\left[-\frac{(n_\nu - \bar{n}_\nu)^2}{2\sigma_\nu^2}\right], \quad (0.7)$$

where  $\sigma_\nu$  is the standard deviation for  $\nu$ -th coincidence measurement (given approximately by  $\sqrt{\bar{n}_\nu}$ ) and Norm is the normalization constant. For our physical density matrix  $\hat{\rho}_p$  the number of counts expected for the  $\nu$ -th measurement is

$$\bar{n}_\nu(t_1, t_2, \dots, t_{16}) = \mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1, t_2, \dots, t_{16})|\psi_\nu\rangle. \quad (0.8)$$

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

$$P(t_1, t_2, \dots, t_{16}) = \frac{1}{\text{Norm}} \prod_{\nu=1}^{16} \exp\left[-\frac{(\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1, t_2, \dots, t_{16})|\psi_\nu\rangle - n_\nu)^2}{2\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1, t_2, \dots, t_{16})|\psi_\nu\rangle - n_\nu}\right], \quad (0.9)$$

where  $\mathcal{N} = \sum_{\nu=1}^4 N_\nu$ .

Rather than find maximum value of  $P(t_1, t_2, \dots, t_{16})$  it simplifies things somewhat to find the maximum of its logarithm (which is mathematically equivalent) [34]. Thus the optimization problem reduces to finding the *minimum* of the following function:

$$\mathcal{L}(t_1, t_2, \dots, t_{16}) = \sum_{\nu=1}^{16} \frac{(\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1, t_2, \dots, t_{16})|\psi_\nu\rangle - n_\nu)^2}{2\mathcal{N}\langle\psi_\nu|\hat{\rho}_p(t_1, t_2, \dots, t_{16})|\psi_\nu\rangle - n_\nu}. \quad (0.10)$$

This is the ‘‘likelihood’’ function which we employed in our numerical optimization routine.

## 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  $t_1, t_2, \dots, t_{16}$ . For this, we used the tomographic estimate of the density matrix in the inverse relation (0.6), allowing us to determine a set of values for  $t_1, t_2, \dots, t_{16}$ . Since the tomographic density matrix may not be non-negative definite, the values of the  $t_\nu$ 's deduced in this manner are not necessarily real. Thus for our initial guess we used the real parts of the  $t_\nu$ 's deduced from the tomographic density matrix.

For the example given in Section 2, the maximum likelihood estimate is

$$\hat{\rho} = \begin{pmatrix} 0.5069 & -0.0239 + i0.0106 & -0.0412 - i0.0221 & 0.4833 + i0.0329 \\ -0.0239 - i0.0106 & 0.0048 & 0.0023 + i0.0019 & -0.0296 - i0.0077 \\ -0.0412 + i0.0221 & 0.0023 - i0.0019 & 0.0045 & -0.0425 + i0.0192 \\ 0.4833 - i0.0329 & -0.0296 + i0.0077 & -0.0425 - i0.0192 & 0.4839 \end{pmatrix}. \quad (0.11)$$

This matrix is illustrated in Fig.3b. In this case, the matrix has eigenvalues 0.986022, 0.0139777, 0 and 0; and  $\text{Tr}\{\hat{\rho}^2\} = 0.972435$ , indicating, while the linear reconstruction gave a non-physical density matrix, the maximum-likelihood reconstruction gives a legitimate density matrix.

## ERROR ANALYSIS

In this section we present an analysis of the errors inherent in the tomographic scheme described in Section III. Two sources of errors are found to be important: the shot noise error in the measured coincidence counts  $n_\nu$  and the uncertainty in the settings of the angles of the waveplates 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.

### Errors due to Count Statistics

From eq.(0.20) we see that the density matrix is specified by a set of sixteen parameters  $s_\nu$  defined by

$$s_\nu = n_\nu / \mathcal{N}, \quad (0.1)$$

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

$$\overline{\delta s_\nu \delta s_\mu} = \sum_{\lambda, \kappa=1}^{16} \left( \frac{\partial s_\nu}{\partial n_\lambda} \right) \left( \frac{\partial s_\mu}{\partial n_\kappa} \right) \overline{\delta n_\lambda \delta n_\kappa}, \quad (0.2)$$

where the over-bar denotes the ensemble average of the random uncertainties  $\delta s_\nu$  and  $\delta n_\lambda$ . The measured coincidence counts  $n_\lambda$  are statistically independent Poissonian random variables, which implies the following relation:

$$\overline{\delta n_\lambda \delta n_\kappa} = n_\lambda \delta_{\lambda, \kappa}, \quad (0.3)$$

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

Taking the derivative of eq.(0.1), we find that

$$\frac{\partial s_\mu}{\partial n_\nu} = \frac{1}{\mathcal{N}} \delta_{\mu\nu} - \frac{n_\mu}{\mathcal{N}^2} D_\nu, \quad (0.4)$$

where

$$D_\nu = \sum_{\lambda=1}^4 \delta_{\lambda, \nu} = \begin{cases} 1 & \text{if } 1 \leq \nu \leq 4 \\ 0 & \text{if } 5 \leq \nu \leq 16. \end{cases} \quad (0.5)$$

Substituting from eq.(0.4) into eq.(0.2) and using eq.(0.3), we obtain the result

$$\overline{\delta s_\nu \delta s_\mu} = \frac{n_\mu}{\mathcal{N}^2} \delta_{\nu,\mu} + \frac{n_\nu n_\mu}{\mathcal{N}^3} (1 - D_\mu - D_\nu). \quad (0.6)$$

In most experimental circumstances  $\mathcal{N} \gg 1$ , and so the second term in eq.(0.6) is negligibly small in comparison to the first. We shall therefore ignore it, and use the approximate expression in the subsequent discussion;

$$\overline{\delta s_\nu \delta s_\mu} \approx \frac{n_\mu}{\mathcal{N}^2} \delta_{\nu,\mu} \equiv \frac{s_\mu}{\mathcal{N}} \delta_{\nu,\mu}. \quad (0.7)$$

### Errors due to Angular Settings Uncertainties

Using the formula (0.7) for the parameters  $s_\nu$  we can find the dependence of the measured density matrix on errors in the tomographic states. The derivative of  $s_\nu$  with respect to some generic wave-plate settings angle  $\theta$  is

$$\frac{\partial s_\nu}{\partial \theta} = \left\{ \frac{\partial}{\partial \theta} \langle \psi_\nu | \right\} \hat{\rho} | \psi_\nu \rangle + \langle \psi_\nu | \hat{\rho} \left\{ \frac{\partial}{\partial \theta} | \psi_\nu \rangle \right\}, \quad (0.8)$$

where  $|\psi_\nu\rangle$  is the ket of the  $\nu$ -th projection state [see eq.(0.5)]. Substituting from eq.(0.14) we find

$$\frac{\partial s_\nu}{\partial \theta} = \sum_{\mu=1}^{16} s_\mu \left[ \left\{ \frac{\partial}{\partial \theta} \langle \psi_\nu | \right\} \hat{M}_\mu | \psi_\nu \rangle + \langle \psi_\nu | \hat{M}_\mu \left\{ \frac{\partial}{\partial \theta} | \psi_\nu \rangle \right\} \right]. \quad (0.9)$$

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

$$\frac{\partial s_\nu}{\partial \theta_{\lambda,i}} = \delta_{\nu,\lambda} \sum_{\mu=1}^{16} s_\mu f_{\nu,\mu}^{(i)}, \quad (0.10)$$

where

$$f_{\nu,\mu}^{(i)} = \left\{ \frac{\partial}{\partial \theta_{\nu,i}} \langle \psi_\nu | \right\} \hat{M}_\mu | \psi_\nu \rangle + \langle \psi_\nu | \hat{M}_\mu \left\{ \frac{\partial}{\partial \theta_{\nu,i}} | \psi_\nu \rangle \right\}. \quad (0.11)$$

The 1024 quantities  $f_{\nu,\mu}^{(i)}$  can be determined by taking the derivatives of the functional forms of the tomographic states given by eqs.(0.4) and (0.5), and evaluating those derivatives at the appropriate values of the arguments (see Table 1).

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 waveplates 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 [35] Thus with the following assumption

$$\overline{\delta \theta_{\nu,i} \delta \theta_{\mu,j}} = \delta_{\nu,\mu} \delta_{i,j} (\Delta \theta)^2, \quad (0.12)$$

(where  $\Delta \theta$  is the RMS uncertainty in the setting of the waveplate, with an estimated value of  $0.25^\circ$  for our apparatus) we obtain the following expression for the errors in  $s_\nu$  due to angular settings:

$$\overline{\delta s_\nu \delta s_\mu} = \delta_{\nu,\mu} \sum_{i=1}^4 \sum_{\epsilon,\lambda=1}^{16} f_{\nu,\epsilon}^{(i)} f_{\nu,\lambda}^{(i)} s_\epsilon s_\lambda \quad (0.13)$$

Combining eqs.(0.13) and (0.7) we obtain the following formula for the total error in the quantities  $s_\nu$ :

$$\overline{\delta s_\nu \delta s_\mu} = \delta_{\nu,\mu} \Lambda_\nu \quad (0.14)$$

where

$$\Lambda_\nu = \left[ \frac{s_\nu}{\mathcal{N}} + \sum_{i=1}^4 \sum_{\epsilon, \lambda=1}^{16} f_{\nu, \epsilon}^{(i)} f_{\nu, \lambda}^{(i)} s_\epsilon s_\lambda \right]. \quad (0.15)$$

These sixteen quantities can be calculated using the parameters  $s_\nu$  and the constants  $f_{\nu, \epsilon}^{(i)}$ . Note that the same result can be obtained by assuming *a priori* that the errors in the  $s_\nu$  are all uncorrelated, with  $\Lambda_\nu = \delta s_\nu^2$ ; the more rigorous treatment given here is however necessary to demonstrate this fact. For a typical number of counts, say  $\mathcal{N} = \infty$  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 Section 3 are as follows:

$$\begin{aligned} (\Delta \rho_{i,j})^2 &= \sum_{\nu, \mu=1}^{16} \frac{\partial \rho_{i,j}}{\partial s_\nu} \frac{\partial \rho_{i,j}}{\partial s_\mu} \overline{\delta s_\nu \delta s_\mu} \\ &= \sum_{\nu=1}^{16} (M_{\nu(i,j)})^2 \Lambda_\nu \end{aligned} \quad (0.16)$$

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 formulae, with the slight modification that the parameter  $s_\nu$  should be recalculated from eq.(0.7) using the estimated density matrix  $\hat{\rho}_{est}$ . This does not take into account errors inherent in the maximum likelihood technique itself.

### 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  $s_\nu$  parameters, (given by eq.(0.15), rather than the errors in the elements of density matrix itself (which have non-negligible correlations).

#### *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]

$$\begin{aligned} \mathcal{S} &= -\text{Tr} \{ \hat{\rho} \log_2 (\hat{\rho}) \} \\ &= -\sum_{a=1}^4 p_a \log_2 p_a, \end{aligned} \quad (0.17)$$

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

$$\hat{\rho} |\phi_a\rangle = p_a |\phi_a\rangle, \quad (0.18)$$

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

$$(\Delta \mathcal{S})^2 = \sum_{\nu=1}^{16} \left( \frac{\partial \mathcal{S}}{\partial s_\nu} \right)^2 \Lambda_\nu. \quad (0.19)$$

Applying the chain rule, we find

$$\left( \frac{\partial \mathcal{S}}{\partial s_\nu} \right) = \sum_{a=1}^4 \left( \frac{\partial p_a}{\partial s_\nu} \right) \left( \frac{\partial \mathcal{S}}{\partial p_a} \right). \quad (0.20)$$

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

$$\delta\lambda_a = \langle\phi_a|\delta\hat{W}|\phi_a\rangle, \quad (0.21)$$

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

$$\frac{\partial\lambda_a}{\partial x} = \langle\phi_a|\frac{\partial\hat{W}}{\partial x}|\phi_a\rangle. \quad (0.22)$$

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

$$\frac{\partial p_a}{\partial s_\nu} = \langle\phi_a|\hat{M}_\nu|\phi_a\rangle \quad (0.23)$$

and so, taking the derivative of eq.(0.17), eq.(0.20) becomes

$$\left(\frac{\partial\mathcal{S}}{\partial s_\nu}\right) = -\sum_{a=1}^4 \langle\phi_a|\hat{M}_\nu|\phi_a\rangle \frac{[1 + \ln p_a]}{\ln 2}. \quad (0.24)$$

Hence

$$(\Delta\mathcal{S})^2 = \sum_{\nu=1}^{16} \left( \sum_{a=1}^4 \langle\phi_a|\hat{M}_\nu|\phi_a\rangle \frac{[1 + \ln p_a]}{\ln 2} \right)^2 \Lambda_\nu. \quad (0.25)$$

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

#### *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 one), the linear entropy for a two qubit system is defined by:

$$\begin{aligned} \mathcal{P} &= \frac{4}{3} (1 - \text{Tr} \{ \hat{\rho}^2 \}) \\ &= \frac{4}{3} \left( 1 - \sum_{a=1}^4 p_a^2 \right). \end{aligned} \quad (0.26)$$

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

$$\begin{aligned} \frac{\partial\mathcal{P}}{\partial s_\nu} &= -\frac{8}{3} \sum_{a=1}^4 p_a \frac{\partial p_a}{\partial s_\nu} \\ &= -\frac{8}{3} \sum_{a=1}^4 p_a \langle\phi_a|\hat{M}_\nu|\phi_a\rangle \\ &= -\frac{8}{3} \text{Tr} \{ \hat{\rho} \hat{M}_\nu \} \\ &= -\frac{8}{3} \sum_{\mu=1}^{16} \text{Tr} \{ \hat{M}_\mu \hat{M}_\nu \} s_\mu \end{aligned} \quad (0.27)$$

Hence the error in the linear entropy is

$$\begin{aligned} (\Delta\mathcal{P})^2 &= \sum_{\nu=1}^{16} \left( \frac{\partial\mathcal{P}}{\partial s_\nu} \right)^2 \Lambda_\nu, \\ &= \sum_{\nu}^{16} \left( \frac{8}{3} \sum_{\mu=1}^{16} \text{Tr} \{ \hat{M}_\mu \hat{M}_\nu \} s_\mu \right)^2 \Lambda_\nu. \end{aligned} \quad (0.28)$$

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

The concurrence, entanglement of formation and tangle are measures of the quantum-coherence properties of a mixed quantum state [29]. For two qubits [36], concurrence is defined as follows: consider the non-Hermitian matrix  $\hat{R} = \hat{\rho}\hat{\Sigma}\hat{\rho}^T\hat{\Sigma}$  where the superscript T denotes transpose and the ‘‘spin flip matrix’’  $\hat{\Sigma}$  is defined by:

$$\hat{\Sigma} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \quad (0.29)$$

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:

$$\hat{R} = \frac{1}{2} \sum_{\mu, \nu=1}^{16} \hat{q}_{\mu, \nu} s_{\mu} s_{\nu}, \quad (0.30)$$

where  $\hat{q}_{\mu, \nu} = \hat{M}_{\mu} \hat{\Sigma} \hat{M}_{\nu}^T \hat{\Sigma} + \hat{M}_{\nu} \hat{\Sigma} \hat{M}_{\mu}^T \hat{\Sigma}$ . 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.:

$$\begin{aligned} \langle \xi_a | \hat{R} &= r_a \langle \xi_a | \\ \hat{R} |\zeta_a\rangle &= r_a |\zeta_a\rangle. \end{aligned} \quad (0.31)$$

We shall assume that these eigenstates are normalized in the usual manner for bi-orthogonal 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  $r_1 \geq r_2 \geq r_3 \geq r_4$ . The concurrence is then defined by the formula

$$\begin{aligned} C &= \text{Max} \{0, \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3} - \sqrt{r_4}\} \\ &= \text{Max} \left\{ 0, \sum_{a=1}^4 \text{sgn} \left( \frac{3}{2} - a \right) \sqrt{r_a} \right\}, \end{aligned} \quad (0.32)$$

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

$$E = h \left( \frac{1 + \sqrt{1 - C^2}}{2} \right), \quad (0.33)$$

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,

$$\begin{aligned} \frac{\partial C}{\partial s_{\nu}} &= \sum_{a=1}^4 \text{sgn} \left( \frac{3}{2} - a \right) \frac{1}{2\sqrt{r_a}} \frac{\partial r_a}{\partial s_{\nu}} \\ &= \sum_{a=1}^4 \text{sgn} \left( \frac{3}{2} - a \right) \frac{1}{2\sqrt{r_a}} \langle \xi_a | \frac{\partial \hat{R}}{\partial s_{\nu}} | \zeta_a \rangle \\ &= \sum_{a=1}^4 \sum_{\mu=1}^{16} \text{sgn} \left( \frac{3}{2} - a \right) \frac{1}{2\sqrt{r_a}} \langle \xi_a | \hat{q}_{\mu, \nu} s_{\mu} | \zeta_a \rangle, \end{aligned} \quad (0.34)$$

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

$$\begin{aligned} (\Delta C)^2 &= \sum_{\nu=1}^{16} \left( \frac{\partial C}{\partial s_{\nu}} \right)^2 \Lambda_{\nu} \\ &= \sum_{\nu=1}^{16} \left( \sum_{a=1}^4 \sum_{\mu=1}^{16} \text{sgn} \left( \frac{3}{2} - a \right) \frac{1}{2\sqrt{r_a}} \langle \xi_a | \hat{q}_{\mu, \nu} s_{\mu} | \zeta_a \rangle \right)^2 \Lambda_{\nu}. \end{aligned} \quad (0.35)$$



For our example the concurrence is  $0.963 \pm 0.018$ .

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

$$\Delta T = 2C\Delta C \quad (0.36)$$

$$\Delta E = \frac{C}{\sqrt{1-C^2}} h' \left( \frac{1 + \sqrt{1-C^2}}{2} \right) \Delta C, \quad (0.37)$$

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

## 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.

## ACKNOWLEDGEMENTS

The authors would like to thank Joe Altepeter, Mauro d'Ariano, Zdenek Hradil, Kurt Jacobs, Poul Jessen, Michael Neilsen, 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  $SU(2) \otimes SU(2)$ , normalized so that the conditions given in eq.0.8 are fulfilled. These matrices are:

$$\begin{aligned} \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}, & \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}, & \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}, & \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}, & \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}, & \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}, & \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}, & \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}, & \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}. \end{aligned} \quad (\text{A.1})$$

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.(0.15), are as follows:

$$\begin{aligned}
\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}, \\
\hat{M}_9 &= \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, & \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}.
\end{aligned} \tag{B.1}$$

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.(0.15), we have

$$\langle \psi_\mu | \hat{M}_\nu | \psi_\mu \rangle = \sum_\lambda \langle \psi_\mu | \hat{\Gamma}_\lambda | \psi_\mu \rangle (B^{-1})_{\lambda,\nu}. \tag{B.2}$$

From eq.(0.12) we have  $\langle \psi_\mu | \hat{\Gamma}_\lambda | \psi_\mu \rangle = B_{\mu,\lambda}$ , thus we obtain the result

$$\langle \psi_\mu | \hat{M}_\nu | \psi_\mu \rangle = \delta_{\mu,\nu}. \tag{B.3}$$

If we denote the basis set for the four-dimensional Hilbert space by  $\{|i\rangle \ (i = 1, 2, 3, 4)\}$ , then eq.(0.14) can be written as follows:

$$\langle i|\hat{\rho}|j\rangle = \sum_{k,l} \sum_{\nu} \langle i|\hat{M}_{\nu}|j\rangle \langle \psi_{\nu}|k\rangle \langle l|\psi_{\nu}\rangle \langle k|\hat{\rho}|l\rangle. \quad (\text{B.4})$$

Since eq.(B.4) is valid for arbitrary states  $\hat{\rho}$ , we obtain the following relationship:

$$\sum_{\nu} \langle i|\hat{M}_{\nu}|j\rangle \langle \psi_{\nu}|k\rangle \langle l|\psi_{\nu}\rangle = \delta_{ik}\delta_{jl}. \quad (\text{B.5})$$

Contracting eq.(B.5) over the indices  $(i, j)$  we obtain:

$$\sum_{\nu} \text{Tr} \left\{ \hat{M}_{\nu} \right\} |\psi_{\nu}\rangle \langle \psi_{\nu}| = \hat{I}, \quad (\text{B.6})$$

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

A second relationship can be obtained by contracting eq.(B.5), viz:

$$\sum_{\nu} \langle i|\hat{M}_{\nu}|j\rangle = \delta_{ij}, \quad (\text{B.7})$$

or, in operator notation,

$$\sum_{\nu} \hat{M}_{\nu} = \hat{I}. \quad (\text{B.8})$$

### APPENDIX C: PERTURBATION THEORY FOR NON-HERMITIAN MATRICES

Whereas perturbation theory for Hermitian matrices is covered in most quantum mechanics text-books, the case of non-Hermitian matrices is less familiar, and so we will present it here. The problem is, given the eigenspectrum of a matrix  $\hat{R}_0$  [31], i.e.:

$$\langle \xi_a|\hat{R}_0 = r_a \langle \xi_a| \quad (\text{C.1})$$

$$\hat{R}_0|\zeta_a\rangle = r_a|\zeta_a\rangle, \quad (\text{C.2})$$

where

$$\langle \xi_a|\zeta_b\rangle = \delta_{a,b} \quad (\text{C.3})$$

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  $\lambda$ :

$$r'_a = r_a^{(0)} + \lambda r_a^{(1)} + \lambda^2 r_a^{(2)} + \dots \quad (\text{C.4})$$

$$|\zeta'_a\rangle = |\zeta_a^{(0)}\rangle + \lambda |\zeta_a^{(1)}\rangle + \lambda^2 |\zeta_a^{(2)}\rangle + \dots \quad (\text{C.5})$$

$$\langle \xi'_a| = \langle \xi_a^{(0)}| + \lambda \langle \xi_a^{(1)}| + \lambda^2 \langle \xi_a^{(2)}| + \dots \quad (\text{C.6})$$

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

$$\hat{R}_0|\zeta_a^{(0)}\rangle = r_a^{(0)}|\zeta_a^{(0)}\rangle \quad (\text{C.7})$$

$$\langle \xi_a^{(0)}|\hat{R}_0 = r_a^{(0)}\langle \xi_a^{(0)}| \quad (\text{C.8})$$

$$\left(\hat{R}_0 - r_a^{(0)}\hat{I}\right)|\zeta_a^{(1)}\rangle = -\left(\delta\hat{R} - r_a^{(1)}\right)|\zeta_a^{(0)}\rangle \quad (\text{C.9})$$

$$\langle \xi_a^{(1)}|\left(\hat{R}_0 - r_a^{(0)}\hat{I}\right) = -\langle \xi_a^{(0)}|\left(\delta\hat{R} - r_a^{(1)}\right). \quad (\text{C.10})$$

Equations (C.7) and (C.8) imply that, as might be expected,

$$|\zeta_a^{(0)}\rangle = |\zeta_a\rangle \quad (\text{C.11})$$

$$\langle \xi_a^{(0)}| = \langle \xi_a| \quad (\text{C.12})$$

$$r_a^{(0)} = r_a \quad (\text{C.13})$$

Taking the inner product of eq.(C.9) with  $\langle \xi_a|$ , and using the bi-orthogonal property eq.(C.3), we obtain

$$r_a^{(1)} = \langle \xi_a|\delta\hat{R}|\zeta_a\rangle. \quad (\text{C.14})$$

This implies that

$$\begin{aligned} \delta r_a &\equiv r_a' - r_a \\ &\approx \langle \xi_a|\delta\hat{R}|\zeta_a\rangle. \end{aligned} \quad (\text{C.15})$$

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

$$\frac{\partial r_a}{\partial x} = \langle \xi_a|\frac{\partial\hat{R}}{\partial x}|\zeta_a\rangle. \quad (\text{C.16})$$

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

$$\left(\hat{R}_0 - r_a \hat{I}\right)^{-1} = \sum_{\substack{b \\ b \neq a}} \frac{1}{r_b - r_a} |\zeta_b\rangle\langle \xi_b|. \quad (\text{C.17})$$

Applying this to eq.(C.9) we obtain

$$\begin{aligned} |\delta\zeta_a^{(1)}\rangle &\equiv |\zeta_a'\rangle - |\zeta_a\rangle \\ &\approx -\sum_{\substack{b \\ b \neq a}} \left( \frac{\langle \xi_b|\delta\hat{R}|\zeta_a\rangle}{r_b - r_a} \right) |\zeta_b\rangle. \end{aligned} \quad (\text{C.18})$$

Similarly, eqs.(C.10) and (C.17) imply

$$\begin{aligned} \langle \delta\xi_a| &\equiv \langle \xi_a'| - \langle \xi_a| \\ &\approx -\sum_{\substack{b \\ b \neq a}} \left( \frac{\langle \xi_a|\delta\hat{R}|\zeta_b\rangle}{r_b - r_a} \right) \langle \xi_b|. \end{aligned} \quad (\text{C.19})$$

- [1] G. C. Stokes, *Trans. Cambr. Phil. Soc.* **9**, 399 (1852).
- [2] U. Leonhardt, *Measuring the quantum state of light* (Cambridge University Press, 1997).
- [3] Z. Hradil, "Quantum-state estimation," *Phys. Rev. A* **55**, R1561 (1997).
- [4] S. M. Tan, "An inverse problem approach to optical homodyne tomography," *J. Mod. Opt.* **44**, 2233 (1997).
- [5] K. Banaszek, G. M. D'Ariano, M. G. A. Paris and M. F. Sacchi, "Maximum-likelihood estimation of the density matrix," *Phys. Rev. A* **61**, 010304 (1999).
- [6] Z. Hradil, J. Summhammer, G. Badurek and H. Rauch, "Reconstruction of the spin state," *Phys. Rev. A* **62**, 014101 (2000).
- [7] J. Řeháček, Z. Hradil and M. Ježek, "Iterative algorithm for reconstruction of entangled states", quant-ph/0009093 (submitted).
- [8] J.R. Ashburn, R.A. Cline, P.J.M. Vanderburgt, W.B. Westerveld and J.S. Risley, "Experimentally determined density-matrices for H(n=3) formed in H<sup>+</sup>-He collisions from 20 to 100 keV," *Phys. Rev. A* **41**, 2407-2421 (1990).
- [9] D. T. Smithey, M. Beck, M. G. Raymer and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum," *Phys. Rev. Lett.* **70**, 1244 (1993).
- [10] G. Breitenbach, S. Schiller, J. Mlynek, "Measurement of the quantum states of squeezed light," *Nature* **387**, 471-475 (1997).

- [11] J. W. Wu, P. K. Lam, M. B. Gray and H.-A. Bachor, "Optical homodyne tomography of information carrying laser beams," *Optics Express* **3**, 154 (1998).
- [12] T.J. Dunn, I.A. Walmsley and S. Mukamel, "Experimental-determination of the quantum-mechanical state of a molecular vibrational-mode," *Phys. Rev. Lett.* **74** 884-887 (1995).
- [13] Experimental determination of the motional quantum state of a trapped atom D. Leibfried, D.M. Meekhof, B.E. King, C. Monroe, W.M. Itano and D.J. Wineland *Phys. Rev. Lett.* **77**, 4281-4285 (1996); D. Leibfried, T. Pfau and C. Monroe, "Shadows and mirrors: Reconstructing quantum states of atom motion," *Physics Today* **51**(4), 22-28 (April 1998).
- [14] C. Kurtsiefer, T. Pfau, J. Mlynek, "Measurement of the Wigner function of an ensemble of helium atoms," *Nature* **386**, 150-153 (1997).
- [15] G. Klose, G. Smith and P. S. Jessen, "Measuring the Quantum State of a Large Angular Momentum," submitted to *Phys. Rev. Lett.*; Los Alamos e-print archive quant-ph/0101017.
- [16] I.L. Chuang, N. Gershenfeld, M. Kubinec, "Experimental implementation of fast quantum searching," *Phys. Rev. Lett.* **80** 3408-3411 (1998).
- [17] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu and R. Schack, "Separability of very noisy mixed states and implications for NMR Quantum computing," *Phys. Rev. Lett.* **83**, 1054 (1999)
- [18] A. G. White, D. F. V. James, P. H. Eberhard and P. G. Kwiat, "Nonmaximally Entangled States: Production, Characterization, and Utilization," *Phys. Rev. Lett.* **83**, 3103 (1999).
- [19] A recent overview of many quantum computation technologies is given in *Fort. d. Phys.* **48** issue numbers 9-11 (2000) (S.L. Braunstein and H.K. Lo, eds.).
- [20] E. Hecht and A. Zajac *Optics* (Addison-Wesley, Reading MA, 1974) section 8.12.
- [21] E. Wolf, *Il Nuovo Cimento* **13**, 1165 (1959); L. Mandel and E. Wolf *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995), ch.6.
- [22] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum and P. H. Eberhard, "Ultrabright-source of polarization-entangled photons," *Phys. Rev. A* **60**, R773 (1999).
- [23] A. G. White, D. F. V. James, W. J. Munro and P. G. Kwiat, "Exploring Hilbert Space: accurate characterization of quantum information," submitted to *Science* (2001).
- [24] A. Berglund, "Quantum coherence and control in one- and two-photon optical systems," undergraduate thesis, Dartmouth College, June 2000; 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), section 10.5.
- [26] A. C. Melissinos, *Experiments in Modern Physics* (Academic Press, New York, 1966), §10.4, pp.467-479.
- [27] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000), ch.11.
- [28] L. I. Schiff, *Quantum Mechanics* (3rd Edition, McGraw-Hill, New York, 1968), eq.(31.8), p.246.
- [29] W. K. Wootters, "Entanglement of formation of an arbitrary state of two qubits," *Phys. Rev. Lett.* **80**, 2245 (1998); V. Coffman, J. Kundu, W. K. Wootters, "Distributed entanglement," *Phys. Rev. A* **61**, 052306 (2000).
- [30] B. M. Terhal, "Detecting Quantum Entanglement," Quantum Physics preprint quant-ph/0101032.
- [31] The properties of the eigenvectors and eigenvalues of non-Hermitian matrices is discussed in P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), vol.I, p.884 *et sequi*.
- [32] It is important to realize that the entangled photon pairs are produced in a *non-deterministic* 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 higher numbers of photons. Thus we can only *post-selectively* generate entangled photon pairs: i.e. one only knows that the state was created after it has been measured.
- [33] Here the *first* subscript on the waveplate angle refers one of the two photon beams; the second subscript distinguishes which of the sixteen different experimental states is under consideration.
- [34] Note that here we neglect the dependence of the normalization constant on  $t_1, t_2, \dots, t_{16}$ , which only weakly effects solution for the most likely state
- [35] 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 will be wholly justified.
- [36] 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, [30] for a recent survey.