Quantum-enhanced tomography of unitary processes

XIAO-QI ZHOU,1,2,† HUGO CABLE,1,*,† REBECCA WHITTAKER,1,† PETER SHADBOLT,1 JEREMY L. O’BRIEN,1 AND JONATHAN C. F. MATTHEWS1,3

1Centre for Quantum Photonics, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, UK
2e-mail: Xiaoqi.Zhou@bristol.ac.uk
3e-mail: Jonathan.Matthews@bristol.ac.uk
*Corresponding author: Hugo.Cable@bristol.ac.uk

Received 5 February 2015; revised 29 April 2015; accepted 30 April 2015 (Doc. ID 234119); published 25 May 2015

A fundamental task in photonics is to characterize an unknown optical process, defined by properties such as birefringence, spectral response, thickness and flatness. Among many ways to achieve this, single-photon probes can be used in a method called quantum process tomography (QPT). However, the precision of QPT is limited by unavoidable shot noise when implemented using single-photon probes or laser light. In situations where measurement resources are limited, for example, where the process (sample) to be probed is very delicate such that the exposure to light has a detrimental effect on the sample, it becomes essential to overcome this precision limit. Here we devise a scheme for process tomography with a quantum-enhanced precision by drawing upon techniques from quantum metrology. We implement a proof-of-principle experiment to demonstrate this scheme—four-photon quantum states are used to probe an unknown arbitrary unitary process realized with an arbitrary polarization rotation. Our results show a substantial reduction of statistical fluctuations compared to traditional QPT methods—in the ideal case, one four-photon probe state yields the same amount of statistical information as twelve single probe photons. © 2015 Optical Society of America

OCIS codes: (270.5585) Quantum information and processing; (120.3940) Metrology.

http://dx.doi.org/10.1364/OPTICA.2.000510

1. INTRODUCTION

Quantum information protocols promise new capabilities for a range of computational, communication and sensing applications. Successful development and implementation of all these quantum information protocols rely on efficient techniques to characterize quantum devices. The most widely used method for this purpose is quantum process tomography (QPT)—in which a mathematical description of a quantum process is reconstructed by estimating the probabilities of outcomes for a selection of probe states and measurement settings. For example for quantum technology, QPT is used to characterize multi-qubit processors [1] and quantum communication channels [2]; across quantum physics, QPT of some form is often the first experimental investigation of a new physical process—for example, recent research into coherent transport in biological mechanisms [3]. QPT has been demonstrated in a variety of physical systems, including ion traps [4], nuclear magnetic resonance [5], superconducting circuits [6] and nitrogen-vacancy color centers [7]. In the context of photonics, experimental demonstrations have been reported in, e.g., Refs. [1,8–10] for linear-optical systems using few-photon states. A continuous-variable version of QPT has also been demonstrated for general optical quantum processes using coherent-state inputs and homodyne measurements [11].

Despite the success of QPT on small systems, there remains scope for improvement. For example, a well-known problem for QPT is the requirement for an exponentially growing number of measurements for processes on an increasing number of qubits. Many methods have been devised and demonstrated to circumvent this problem, such as efficient state tomography [12] and compressed sensing [13,14]. In photonics, photon loss and interferometric instability present the main challenges, and a method called “super-stable tomography” was recently proposed to address these [15,16]. Here we are concerned with improving measurement precision given a fixed number of particles propagating through the unknown process. It has been found that the precision achieved by tomography methods after a fixed number of measurements is dependent on the unknown state or process. This has naturally led to adaptive schemes [17,18] in which dynamic measurement settings are used. However, regardless of the measurement scheme—adaptive or nonadaptive—the precision is always limited by the unavoidable statistical fluctuation, where the ultimate precision limits are dictated by quantum mechanics.
In the context of single-parameter phase estimation, it is very well explored how quantum resources allow an improvement in the precision limit [19, 20]. Such quantum metrology techniques are particularly important for measuring very light-sensitive biological samples [21] and delicate materials [22], where high power and long exposure time is not permitted. However, when the interaction of a sample with light cannot be represented by a single parameter, for example, an arbitrary polarization rotation, then these techniques cannot be directly applied. While traditional quantum metrology deals with single-parameter phase estimation, QPT can be viewed as multiparameter estimation, and it is natural to consider whether the precision can be improved for this more general problem by using nonclassical resources. Situations for estimating multiple phases in linear optics have been explored for special families of unitaries, with applications in phase imaging [23] and interferometry in waveguides [24]. For the case of general unitary estimation, some theoretical results are known about how precision scaling can be improved using entanglement [25–28]. However, these results were derived in an abstract setting, and no scheme for practical implementation has been proposed so far. It is not even known in principle if the quantum advantage is achievable by using standard methods in quantum photonics (linear optics, photon-number counting, etc.).

There have been several related theoretical investigations that explore how the properties of quantum mechanics, especially quantum entanglement, can improve the precision for abstract unitary estimation [25–28], but none of these offers an explicit mapping onto photon systems. Some theoretical results on estimating a single optical phase parameter together with a nonunitary process parameter can be found in Refs. [29, 30]. The problem of detecting and characterizing isotropic depolarizing noise using two-photon entangled states is tackled recently in Ref. [31].

Here we present a quantum-enhanced process tomography protocol that works for arbitrary unitary optical processes on two modes—this corresponds to estimating three unknown noncommuting phases that cannot, in general, be measured separately. This protocol uses principles from quantum metrology to exploit quantum interferences as a means for minimizing unwanted fluctuation on the quantum-process measurement statistics. We have also implemented a proof-of-principle experiment to demonstrate this protocol and obtained measurement precision beyond that achievable with traditional QPT protocols. We first explain the theory of our protocol, and then present the results of our experimental implementation, which show evidence for a quantum-enhanced precision compared to the conventional QPT approach. Finally, we discuss generalizations and applications of our protocol.

2. RESULTS

A. Protocol for Unitary Estimation

The standard procedure for QPT applies repeated state tomography on a set of input states acted on by the process [1, 32]. A full procedure for process tomography, as illustrated in Fig. 1(a), commonly assumes that the quantum process corresponds mathematically to a completely positive trace-preserving map and physically to quantum evolution, which can include decoherence or dissipation. If the process acts on a l-dimensional system, $l^4 - l^2$ configurations must be tested [32]. Here we consider the unitary case where there are $l^2 - 1$ unknown real parameters, encompassing a broad class of optical devices and processes.

Our protocol is applied to reconstructing unitaries acting on two-mode radiation fields ($l = 2$), $U$, which are parameterized by three real numbers up to a global phase. The key ideas of our protocol can be applied to pairs of modes, whether defined by spatial, polarization, or orbital angular momentum, or temporal degrees of freedom. We will work here with unitary rotations on polarization degrees of freedom labeled $H$ and $V$ (where all other optical degrees of freedom are identical). When $U$ acts on an arbitrary single-photon superposition state, $|\psi\rangle = c_H|1,0\rangle_H + c_V|0,1\rangle_V$, where $c_H$ and $c_V$ are the amplitudes, and $H$ and $V$ denote horizontal and vertical polarization), the corresponding transformation $|\psi\rangle \mapsto U|\psi\rangle$ is given by 

$$U = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}.$$  

Our protocol works for an arbitrary number of photons $N$, for which the Hilbert space is spanned by basis vectors $|N, 0\rangle_H, |N - 1, 1\rangle_H, \ldots, |0, N\rangle_H$. The action of $U$ here corresponds to a transformation using a $N + 1$ by $N + 1$ matrix on this state space (see Supplement 1 and Refs. [33, 34]). So the task to estimate the unknown unitary $U$ becomes to determine the values of $a$, $b$, $c$, and $d$ satisfying the unitarity constraint.

![Fig. 1.](image)

(a) Standard procedure for quantum process tomography. This procedure can be considered in three stages: preparation of a series of probe states, interaction of the unknown process with the probe, and detection of the output in a selection of measurement bases. Based on the measurement outcomes, a mathematical map corresponding to the unknown process can be reconstructed. (b) Experimental setup for probing the unknown unitary processes. An 80 MHz pulsed Ti:sapphire laser centered at 808 nm is upconverted to 404 nm and then focused onto a bismuth borate (BiBO) crystal, phase-matched for Type-I SPDC, creating noncollinear degenerate horizontally polarized photon pairs at 808 nm. After converting one arm to vertical polarization using a half-wave plate (HWP), the two arms are combined by a polarizing beam splitter (PBS). The resulting state is passed through the unknown unitary and then measured in the H/V basis. Approximate photon-number counting (see Appendix A) is implemented on each polarization mode using a 1–8 fan-out array onto avalanche photodiode detectors (APDs). By using different wave plate settings before and after the unknown unitary, the basis for the probe state and measurement can be changed to D/A and R/L. The setup is precalibrated using laser light.
\[ a^2 + b^2 + c^2 + d^2 = 1. \]  

We assume that the unitary is operating on the polarization degrees of freedom and we denote horizontal (diagonal, right circular) and vertical (antidiagonal, left circular) polarization by \( H \) and \( V \) respectively, where \( |D/A\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle) \) and \( |R/L\rangle = \frac{1}{\sqrt{2}}(|H\rangle \pm i|V\rangle) \). By inputting an \( H \) photon to the unitary and measuring the output photon in the \( H/V \) basis \( (D/A,R/L) \) basis, the probability of detecting \( H \) photon at the output is \( p_{HV} = \frac{D}{A}p_{DA} \), where

\[ p_{HV} = a^2 + b^2, \quad p_{DA} = a^2 + d^2, \quad p_{RL} = a^2 + c^2. \]  

By using the unitary constraint of Eq. (1), \( U \) can then be estimated by measuring these three probabilities. There is a discrete set of estimates, which all correspond to the same values for \( p_{HV(DA,RL)} \). (A similar situation exists in interferometric phase estimation, where typically multiple values of the phase are consistent with a particular set of data.) While the sign of \( a \) can always be fixed to positive, the signs of \( b, c, \) and \( d \) need to be resolved. For this we use supplementary standard QPT, using a small and minimal number of probe photons, to provide an initial coarse-grained estimate sufficient to differentiate these alternatives (see Appendix A).

The traditional approach would directly estimate \( p_{HV}, p_{DA} \) and \( p_{RL} \) with single photons by looking at the ratio of detections at the two outputs. The precision of estimating \( p_{HV} \) is \( \Delta p_{HV} = \sqrt{p_{HV}(1 - p_{HV})/N} \), where \( N \) is the number of probing photons, which scales as \( O(N^{-1/2}) \), the standard quantum limit (SQL) for measurement. To go beyond the SQL, our approach uses multiphoton states as probes, determining the three probabilities shown in Eq. (2), indirectly from the multiphoton counting statistics.

### B. Process Reconstruction Using Multiphoton States

The multiphoton input state we use is a \( N \)-photon state split equally between \( H \) and \( V \)-polarization [35], \([N/2, N/2]_{HV}\), where \( N \) is even. After propagating through the unknown unitary, the state is measured in the \( H/V \) basis as in the single-photon case. The probability of detecting \( n_H \) \( H \)-polarized photons and \( n_V \) \( V \)-polarized photons at the output is a function of \( p_{HV} \) alone. The general form of the matrix elements corresponding to the action of \( U \) in the Fock basis is explained in Supplement 1 (Section 1): there is interference due to \( p_{HV} \), but not due to other parameters, which generate only a global phase that depends on the total number of photons (and that has no effect on the measurement statistics). The form of the matrix elements and corresponding probabilities is given by a standard derivation following the Schwinger representation [36], and is given explicitly in Ref. [37] for example. We give the result here expressed using Legendre polynomials:

\[ \mathcal{P}(n_H, n_V, p_{HV}) = \frac{n_V^n}{n_H^n} \left( \frac{\mathcal{L}(n_H - n_V)}{\mathcal{L}(n_H + n_V)} \right)^2 (2p_{HV} - 1)^2, \]  

where \( \mathcal{L}(n_H - n_V)/2 \) denotes the standard associated Legendre polynomial [38] with degree \( n_H + n_V \) and order \( n_H - n_V \) (see Supplement 1). Consequently, \( p_{HV} \) can be estimated from the photon-counting data using a maximum-likelihood technique with a precision of \( \Delta p_{HV} = \sqrt{p_{HV}(1 - p_{HV})/N(N/2 + 1)} \) (as discussed in Ref. [39], for example), which scales with \( O(N^{-1}) \)—a quadratic improvement compared to the SQL. For the \( N = 4 \) case, the probabilities \( \mathcal{P}(n_H, n_V, p_{HV}) \) of each possible outcome \( (n_H, n_V) \) are given in Table 1. The resulting uncertainty of the estimated probability, \( \Delta p_{HV} \), is 70% larger for the standard QPT approach than when using our method. By changing the input state to \([N/2, N/2]_{DA(RL)}\) and the measurement basis to \( D/A(R/L) \), we can obtain \( p_{DA} \) (\( p_{RL} \)) with the same precision as \( p_{HV} \). Our method can be directly extended to the unbalanced input states of the form \([M, N - M] \), with precision improved beyond the SQL, provided \( M \neq 0 \) or \( N \). As a practical consequence, the probe state can be the entire state generated by spontaneous parametric downconversion (SPDC), which can be easily created in the laboratory: the probe state would be a superposition of states \([N/2, N/2] \) having different total photon number generated by a Type-I SPDC source, or a mixture of \([M, N - M] \) states with a heralded Type-II SPDC source [40].

### C. Experimental Setup

To demonstrate our scheme, we use four-photon states generated using standard Type-I SPDC source (see Appendix A), as shown in Fig. 1(b). With a certain probability, the SPDC source will produce two photon pairs (with \( H \) polarization) across the two arms. After rotating one arm to \( V \) polarization using a half-wave plate (HWP), the two arms are then combined on a polarization beam splitter (PBS), thus producing the desired four-photon state \([2, 2]_{HV}\). The state passes through the unknown unitary and is then separated into the \( H \) and \( V \) components using a PBS. The photon number at each output is resolved using a fan-out array that couples to eight avalanche photodiodes (APDs). We postselect the four-photon coincidence events (see Appendix A), and for this case the detection loss of the APDs has little effect on the fidelity of the postselected state and data analysis. We use the measured rates of four-photon outcomes to estimate \( p_{HV} \) by using the maximum-likelihood method based on the theoretical probability distributions shown in Table 1. By changing the input state and the measurement basis to \( D/A(R/L) \), implemented by wave plates before and after the unitary, we estimate \( p_{DA} \) (\( p_{RL} \)) in a similar manner. The experimentally determined \( \tilde{p}_{HV}, \tilde{p}_{DA}, \) and \( \tilde{p}_{RL} \) are then used to construct an estimate of the unknown unitary, \( \tilde{U} \). To quantify the discrepancy between \( \tilde{U} \) and \( U \), we use the process infidelity, defined as \( (1 - \min |\psi(\tilde{U}^\dagger U)\psi|^2) \), where the minimum is taken over all single-photon states \( |\psi\rangle \). For a given choice of probe state and total number of probe photons, the mean and spread of the infidelity depend strongly on the choice of unitary; this is detailed in Supplement 1.

### D. Unitary Reconstructions from Data

To provide evidence that the scheme works for any two-mode unitary, we test it using a number of preselected unitaries randomly sampled from the Haar distribution [41]. For each

<table>
<thead>
<tr>
<th>( n_H, n_V )</th>
<th>( \text{Outcome(s)} )</th>
<th>( \text{Probability} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (0, 4) )</td>
<td>( (0, 4) )</td>
<td>( 6p_{HV}^2(1 - p_{HV})^2 )</td>
</tr>
<tr>
<td>( (1, 3) )</td>
<td>( (3, 1) )</td>
<td>( 6p_{HV}^2(1 - p_{HV})(2p_{HV} - 1)^2 )</td>
</tr>
<tr>
<td>( (2, 2) )</td>
<td>( (2, 2) )</td>
<td>( (6p_{HV}^2 - 6p_{HV} + 1)^2 )</td>
</tr>
</tbody>
</table>

*Analogous expressions hold for the \( D/A \) and \( R/L \) measurements.*
of these unitaries $U$, 200 four-photon probes (800 photons in total) are used to construct the estimate $\tilde{U}$. The process infidelities range from 96.8% to 99.8%, with a mean of 98.8% showing near-ideal performance for the scheme. The relationships between the expected and the experimentally reconstructed unitaries are represented graphically in Fig. 2. This analysis shows qualitatively that our scheme provides high-quality estimates for arbitrary unitaries.

### E. Demonstration of Quantum Advantage for Precision

Now we turn to the central feature of the scheme—the ability to exploit multiphoton quantum interference to improve the estimate precision with a fixed input resource—the total number of photons propagating through the unknown unitary. We randomly choose two unitaries, $U_A = \begin{pmatrix} 0.70 + 0.21i & -0.65 - 0.20i \\ 0.65 - 0.20i & 0.70 - 0.21i \end{pmatrix}$ and $U_B = \begin{pmatrix} 0.29 + 0.34i & 0.33 + 0.83i \\ -0.33 + 0.83i & 0.29 - 0.34i \end{pmatrix}$ (shown in Fig. 2), and look at the variation of $\tilde{U}$ over many repetitions of the experiment. Theoretically, both the mean and standard deviation of the process infidelities of $\tilde{U}$ with respect to $U$ will be improved toward zero using our scheme, as opposed to the traditional approach using single photons. In practice, we use a SPDC source and postselection to simulate exact photon-number states, which inevitably result in systematic errors that prevent a fair comparison between the mean infidelities of four-photon and single-photon probe states [42]. To properly quantify the spread of the estimates $\tilde{U}$ from actual data, we use the standard deviation of the process infidelities of $\tilde{U}$ with respect to a central estimate $U'$, where $U'$ is a single estimate constructed using a large dataset comprised of all the data used to make each $\tilde{U}$. For example, in the case of probing $U_A$ with 1800 photons, 450 four-photon states across all three measurements are used to obtain an estimate $\tilde{U}_A$; this is repeated 240 times. The corresponding central estimate $\tilde{U}_A'$ is obtained by using the same data used to make these 240 estimates.

The analysis of the experimental data for $U_A$ and $U_B$ are shown in Fig. 3. As expected, the mean and the standard deviation of the infidelity decrease as the photon number increases for both schemes. More importantly, for every fixed input resource, the standard deviation of the infidelity is reduced for our scheme compared to the traditional approach. Theoretically, as shown in Fig. 3, using standard QPT one would need approximately 3600 photons overall to obtain the same infidelity and spread as achieved by our protocol using only 1800 probe photons. In the ideal case [Figs. 3(a2) and 3(b2)], where there is no restriction on the number of probe photons, one four-photon probe state yields the same amount of statistical information as twelve single probe photons. The experimental results are closely matched to the predictions of the theoretical simulations as described in Fig. 3; however, our results [Figs. 3(a1) and 3(b1)] show a slightly smaller quantum advantage due to practical limitations on the size of the dataset and imperfections of the four-photon states. For both $U_A$ and $U_B$, our experimental results show that 600 four-photon states (2400 total probing photons) gives a smaller spread of error compared to 3600 single photons—this holds true even for comparing experimental data for the four-photon states with the theoretical simulation using single photons.

Our scheme will obtain a quantum advantage when using $\lfloor N/2 \rfloor$ probes for all $N \geq 2$. In this experiment we demonstrated the advantage for $N = 4$, as opposed to $N = 2$, because the expected scaling of precision for our scheme increases with larger photon number. We note that, while we use single-photon probes to compare the performance of our scheme for unitary estimation, the same results would have been achieved using coherent laser light instead of single photons.

### F. Error Analysis

The deviations of our experimental results from the theoretical predictions originate from four parts of the experiment. (i) **Input states:** Our scheme assumes perfect Fock states that have fixed photon number. To simulate the Fock states experimentally, we used a SPDC source, which inherently contains higher order terms, temporal distinguishability between photon pairs, and spectral distinguishability between the two arms, all of which alter the intended quantum interference. (ii) **Optical components:** There always is imprecision in setting the angles of the manual wave plates, which can be improved by using a motorized bulk-optical system or migrating to an integrated architecture. (iii) **Detection system:** To simulate photon-number resolving detection, we use a 1-to-8 fiber array terminated with eight APDs on each output. The nonuniform efficiency across the 16 APDs and the nonidentical splitting ratio of the fiber arrays result in some bias in the detection system. (iv) **Dark counts:** There is a greater discrepancy between simulation and experiment when using single-photon states over the four-photon states. While the dark-count contribution to the fourfold coincidences is negligible, it does contribute to the results of the single-photon data (compounded by the fact that all 16 APDs were used for all measurements). The effect of dark counts changes depending on the unitary being measured.
which accounts for why there is a larger discrepancy for $U_A$ than for $U_B$. Despite these limitations, we still see a clear quantum advantage of our four-photon data over the traditional method.

### 3. DISCUSSION

Since any unitary linear-optical circuit can be viewed as a nested interferometer [43] that can be decomposed into a series of symmetric beam splitters and phase shifters, it is a promising direction to extend the protocol introduced in this paper to multiparameter estimation of larger unitary operations acting on many modes. One might expect a generalization of our protocol to deal with estimating unknown unitaries on $l$ modes, where there are $l^2 - 1$ parameters, and at least $l^2 - 1$ configurations are required for the input and measurement bases. The results of Ref. [26] go some way to meeting this goal: The author considered the problem of estimating $SU(l)$ quantum operations, when $N$ copies of it are available at the same time. Although not stated in the reference, the results can be applied to the situation of $l$-mode linear-optical unitaries probed by a selection of $N$-photon probe states, and a mathematical isomorphism relates the analysis in Ref. [26] and a first-quantized analysis of an interferometer. However, it remains to find an experimental procedure that is equivalent to the measurements suggested by Ref. [26].

As shown by the theoretical simulations (see Supplement 1) and experimental results of our protocol, the attainable precision is highly dependent on the unknown unitary. As such, we expect to achieve maximum performance by combining our protocol with an adaptive technique [17,18]. This would involve adapting the procedure toward one of the optimal unitaries $U'$, where the precision improvement due to the use of multiphoton probe states and adaptive techniques are separate enhancements that can be combined to optimize the quantum advantage.

In the context of single-parameter quantum metrology, the improvements to measurement precision achieved using nonclassical quantum states have found important applications, such as gravitational-wave detection [44], where the laser power cannot be turned up further due to radiation pressure, and measurement of fragile biological samples, where overexposure to light can...
damage the sample [22]. We expect that our technique will be applied to achieve similar gains for applications such as characterization of optical media (in particular birefringence), quantum logic gates [45], and new types of quantum sensors [19], where the processes in question require multiple parameters for their descriptions. Achieving the same precision with fewer photons results in reduction of power consumption together with a higher throughput of samples being screened, while also enabling probing samples that would otherwise degrade due to light-exposure.

Considering applications outside of photonics, the scheme has a natural geometric interpretation bestowed by the Schwinger representation [37], which provides a one-to-one map from two-mode-N-photon states to the N/2-spin state space. Here the two-mode unitary operations correspond to physical rotations of a spin system, or, equivalently, rotations of the reference frame. Consequently, our experiment shows that a quantum advantage is indeed possible for the task of aligning Cartesian reference frames, as predicted in several theoretical works using other protocols [46]. A spin implementation of our protocol has practical applications for both gyroscopy and magnetometry.

APPENDIX A: METHODS

1. Reconstruction of $U$ from Estimated Quantities $\tilde{p}_{HV}, \tilde{p}_{RL}$, and $\tilde{p}_{DA}$

Linear inversion of Eqs. (1) and (2) allows the unknown unitary to be reconstructed from the experimentally derived estimates $\tilde{p}_{HV}, \tilde{p}_{RL}$, and $\tilde{p}_{DA}$, whenever this leads to nonnegative values for $\tilde{a}^2, \tilde{b}^2, \tilde{c}^2$, or $\tilde{d}^2$. This inversion is given explicitly by

$$
\begin{pmatrix}
\tilde{a}^2 \\
\tilde{b}^2 \\
\tilde{c}^2 \\
\tilde{d}^2
\end{pmatrix} = \frac{1}{2} \begin{pmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1
\end{pmatrix} \begin{pmatrix}
\tilde{p}_{DA} \\
\tilde{p}_{RL} \\
\tilde{p}_{HV}
\end{pmatrix},
$$

(A1)

and it can be applied whenever the following inequalities are all satisfied:

$$
\begin{align*}
\tilde{p}_{DA} + \tilde{p}_{RL} + \tilde{p}_{HV} & \geq 1, \\
\tilde{p}_{DA} + \tilde{p}_{RL} - \tilde{p}_{HV} & \leq 1, \\
\tilde{p}_{DA} - \tilde{p}_{RL} + \tilde{p}_{HV} & \leq 1, \\
-\tilde{p}_{DA} + \tilde{p}_{RL} + \tilde{p}_{HV} & \leq 1.
\end{align*}
$$

Possible values for the probability estimates define the cube with $0 \leq \tilde{p}_{HV}, \tilde{p}_{RD}, \tilde{p}_{DA} \leq 1$, and the inequalities above determine a tetrahedral subregion that we call the physical region—with vertices at $(1,0,0), (0,1,0), (0,0,1)$, and $(1,1,1)$. Outside of the physical region, exactly one of the inequalities fails to hold, and, therefore, we choose the point closest to $(\tilde{p}_{HV}, \tilde{p}_{RD}, \tilde{p}_{DA})$ in the physical region (with respect to the Euclidean metric). Simple expressions for the closest point follow from geometric considerations. As an aside, the maximum-likelihood procedure that is applied to estimate the value of $\tilde{p}_{HV(DA,RL)}$ from data, with four-photon input states, cannot distinguish values $\tilde{p}_{HV(DA,RL)}$ and $1 - \tilde{p}_{HV(DA,RL)}$, which are both consistent with measurement results. This is because the probability distributions in Table 1 and Eq. (3) are symmetric under the mathematical operation $\tilde{p}_{HV} \leftrightarrow 1 - \tilde{p}_{HV}$ (and similarly for $\tilde{p}_{DA(RL)}$). This ambiguity is resolved by the same coarse-grained estimates, obtained with supplementary QPT measurements, which is needed to determine the signs of $b, c$ and $d$.

2. Source

A BiBO nonlinear crystal (2 mm thick), phase-matched for Type-I noncollinear SPDC, is pumped by a frequency-doubled mode-locked Ti:sapphire laser operating with 80 MHz repetition rate. The degenerate SPDC photons centered at 808 nm from the crystal are generated for building the desired quantum states. The bandwidth of the SPDC photons is restricted by using 3 nm FWHM interference filters.

3. Photon-Number Detection

Each of the two outputs is connected to an eight-detector array of APDs through a one-to-eight fiber splitter [Fig. 1(b)]. To process the large number of coincidence patterns that can occur, we developed a highly sophisticated coincidence counting system. This system time-tags incoming photons across 16 channels with ~80 ps timing resolution, and we have developed fast programs that process these time-tags in real time. With this detection system, we can effectively realize an $N$-photon input state by postselecting on $N$-fold coincidences. In our experiment using four-photon input states, there are a total of 1820 possible fourfold coincidence patterns. The five possible outcomes, (4,0), (3,1), (2,2), (1,3), and (0,4), correspond to 70, 448, 784, 448, and 70 different fourfold coincidence patterns, respectively. Our detection system can properly resolve the photon number only when the input photons are all distributed to different channels (detectors). The probability for output state $(n_H, n_V)$ being perfectly resolved is $8! \times 8! ((8 - n_H)(8 - n_V)(8^{n_H+n_V}))^{-1}$, assuming unit efficiency for each APD. As mentioned in the main text, we can ignore the effects of the inefficiency of individual APDs by considering only those cases where photons are detected in fourfold coincidences. To make this approximation valid, the detection efficiency must not be dependent on the specific coincidence pattern. Therefore, we applied a discard rule with keep probabilities 1, 0.625, 0.536, 0.625, and 1 for the five possible outcomes, (4,0), (3,1), (2,2), (1,3), and (0,4), respectively [47].

Engineering and Physical Sciences Research Council (EPSRC); European Research Council (ERC); European Commission FP7 ACTION project: Beyond the Barriers of Optical Integration (BB0I); PHORBITECH; QUANTIP; Army Research Office (ARO) (W911NF-14-1-0133); Air Force Office of Scientific Research (AFOSR); Centre for Nanoscience and Quantum Information (NSQI).

The authors thank Dylan Mahler, Tomek Paterek, Peter Turner, and Sai Vinjanampathy for helpful discussions. J. L. O’Brien acknowledges a Royal Society Wolfson Merit Award and a Royal Academy of Engineering Chair in Emerging Technologies. J.C.F.M. acknowledges support from a Leverhulme Trust Early Career Fellowship.

These authors contributed equally to this work.

See Supplement 1 for supporting content.

REFERENCES
