Superresolution Enhancement in Biphoton Spatial-Mode Demultiplexing

Florence Grenapin,[1] Dilip Paneru,[1] Alessio D'Errico,[1,*] Vincenzo Grillo,[2] Gerd Leuchs,[1,3] and Ebrahim Karimi[1,3]

[1] *Nexus for Quantum Technologies, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

[2] *Istituto Nanoscienze, Consiglio Nazionale delle Ricerche, Via G. Campi 213/A, Modena 41125, Italy*

[3] *Max Planck Institute for the Science of Light, Erlangen 91058, Germany*

Imaging systems measuring intensity in the far field succumb to Rayleigh's curse, a resolution limitation dictated by the finite aperture of the optical system. Many proof-of-principle and some two-dimensional imaging experiments have shown that, by using spatial mode demultiplexing (SPADE), the field information collected is maximal, and, thus, the resolution increases beyond the Rayleigh criterion. Hitherto, the SPADE approaches are based on resolving the lateral splitting of a Gaussian wave function. Here, we consider the case in which the light field originates from a biphoton source, i.e., spontaneous parametric down-conversion, and a horizontal separation is introduced in one of the two photons. We show that a separation induced in the signal photon arm can be superresolved using coincidence measurements after projecting both photons on Hermite-Gauss modes. Remarkably, the Fisher information associated with the measurement is enhanced compared to the ordinary SPADE techniques by $\sqrt{K}$, where K is the Schmidt number of the two-photon state that quantifies the amount of spatial entanglement between the two photons.

I. INTRODUCTION

The resolution of far-field optical imaging systems based on direct intensity measurements is limited by the point spread function (PSF), a diffraction phenomenon dictated by light wavelength and aperture width of the optics involved [1]. The accuracy of this measurement diverges for small separations of two-point sources; an effect also referred to as "Rayleigh's curse." The discovery of this phenomenon has encouraged research efforts in other fields of imaging, such as near-field imaging [2,3] or imaging based on electronic effects [4,5], which goes beyond the optical Rayleigh limit. In 2016, Tsang and his colleagues proposed a simple spatial mode demultiplexing (SPADE) scheme, which is robust to the resolution curse [6]. It shows that our ability to estimate a separation between two incoherent point sources collapses as the separation approaches zero when using conventional intensity measurements, thus assimilating Rayleigh's curse to an intrinsic loss of Fisher information at small separations. The SPADE scheme proposed uses spatial mode sorting, with mode projections tailored to the PSF of the imaging system, which collects maximal Fisher information regardless of the separation magnitude. This has

prompted many further theoretical and experimental findings in recent years, exploring applications and modifications of SPADE for different imaging systems [7–10], coherent point sources [11–13], limitations in the presence of cross-talk or other noise sources [14,15], and even extensions to higher-dimensional objects [16,17]. Although the explicit model estimating all points of an arbitrary two-dimensional (2D) object is vastly complicated to derive in terms of Fisher information, 2D imaging simulations [18] and experiments [19] using postprocessing algorithms (deconvolution, machine learning) have shown an increase in resolution beyond the Rayleigh limit. This solidifies the intuition relating the distinguishability of two-point sources to the overall resolution of a 2D imaging system.

The main idea behind SPADE is to estimate the separation between two-point sources by looking at the alteration of the field's phase, in particular, at the change in phase symmetry, which can be efficiently probed by the Hermite-Gauss (HG) decomposition for the usual Gaussian PSF. In this work, we address the following question: can the spatial correlations emerging from two-photon sources enhance the resolution sensitivity? We show that it is possible to superresolve a separation using mode sorting in coincidence measurements with a PSF that is itself entangled in the form of spontaneous parametric down-conversion (SPDC) light. More specifically, we consider a two-photon

*aderrico@uottawa.ca

wave function that allows a Schmidt decomposition in HG modes, with Schmidt number K. An incoherent displacement, which mimics the effect of a transmitting sample, is applied on one of the two photons, and the displacement is estimated by projecting the resulting state on direct products of HG modes. We note that the somewhat related inverse problem of generating a spatially or temporally shaped pure photon state in one beam of an entangled pair (ghost interference) likewise requires the projection on pure states in the other beam of the entangled pair [20]. Our main result shows that the Fisher information, whose inverse gives a lower bound on the estimation error, scales as $\sqrt{K}/2$, where $K = 1$ corresponds to the separable case of a Gaussian point spread function. Hence, any spatially entangled two-photon source provides an advantage with respect to the ordinary SPADE. Here, we derive this result and present a first proof-of-principle experimental implementation. We conclude by discussing the possible settings in which the two-photon SPADE can give a practical advantage.

II. THEORY

To understand the resolution limits of a given experimental technique, one must evaluate the lower bound on the achievable uncertainty of the estimator δ (which, in our case, is the lateral displacement). This limit is given by the Cramer-Rao bound [6,21]: the estimator's standard error is lower bounded by the inverse of the Fisher information $\Delta\delta \geq 1/\sqrt{n\mathcal{FI}}$ [22,23], where n corresponds to the number of repeated measurements and the Fisher information, $\mathcal{FI}$, can be calculated from the probabilities $\mathcal{P}_j$ of a given measurement outcome j as $\mathcal{FI} = \sum_j (1/\mathcal{P}_j)(\partial\mathcal{P}_j/\partial\delta)^2$. In the following, we consider the problem of resolving an incoherent transverse displacement of a photon that is in a spatially correlated state with another idler photon. We calculate the Fisher information for a biphoton spatial mode demultiplexing scheme and compare the result with the classical SPADE, which emerges from the uncorrelated limit of our scenario.

In a typical imaging setup, the image plane of a point source is described by a two-dimensional PSF $\Psi(\mathbf{r})$ [with $\mathbf{r} = (x, y)$]. In the case of two incoherent point sources separated horizontally by a distance s, the wave function in the image plane can be described by the mixed state $\rho = \frac{1}{2}(|\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|)$, where $\langle x|\Psi^\pm\rangle := \Psi(x \pm s/2, y)$. An intensity measurement, which corresponds to a projective measurement in the position bases, results in outcomes with distribution $I(x, y) = \frac{1}{2}(|\Psi^+|^2 + |\Psi^-|^2)$. The Fisher information for the direct imaging method is known to rapidly fall to zero for separations smaller than the width of the point spread function [6]. However, by mode sorting using modes that form orthogonal bases for space containing the PSF (SPADE), the

total Fisher information remains constant across all values of s [6]. Optimal bases are ones in which most of the Fisher information for small s is captured in the first couple of measurements, making schemes like binary SPADE [10] simple and attractive. We show in Appendix A how a projection on the derivative of the PSF with respect to the coordinate associated with the direction of the displacement is an optimal projection for the small separation regime. More rigorous ways to determine optimal measurement bases have been developed [7].

In a coincidence imaging setup with entangled SPDC photon pairs, there is no pure wave function describing one of the single photons' quantum states. The single particle state is maximally mixed, which can mimic a Gaussian in its intensity pattern when viewed on the camera. The pure state describing the entangled particles is a biphoton state of signal and idler that can be expressed in the Schmidt basis of HG modes [24]:

$$|\Psi\rangle = \sum_{m,n} C_{mn} |m, n\rangle_s \otimes |m, n\rangle_i. \tag{1}$$

Here the C_{mn} coefficients are the Schmidt coefficients of the HG decomposition and $|m, n\rangle$ states are the 2D HG modes of order (m, n) with a beam waist parameter w: $\langle x, y | m, n\rangle := \mathrm{HG}_{m,n}(x, y) = \mathcal{N} \exp(-(x^2 + y^2)/w^2)H_m(\sqrt{2}x/w)H_n(\sqrt{2}y/w)$ with the $H_m(x)$ Hermite polynomials and $\mathcal{N}$ a normalization constant. In this basis, the correlations are diagonal: a signal photon in a particular $|i, j\rangle$ mode is entangled with an idler photon in that exact mode (see Appendix B). In practice, any action occurring on one photon, such as the typical diffraction causing PSFs, is also reflected in the biphoton wave function. The Schmidt coefficients can be given an analytical expression (see Refs. [25,26]): $C_{m,n} = |(\gamma - 1)/(\gamma + 1)|^{m+n}4\gamma/(1 + \gamma)^2$ with $\gamma = (1/w_p)\sqrt{L\lambda_p/2\pi}$ a parameter defined by crystal length L, pump beam waist w_p, and wavelength λ_p. Here, we investigate how two-photon SPADE allows us to superresolve separation introduced in one of the photon paths as if the PSF was of the form of the source (just as in Refs. [9,10])—which is the biphoton wave function described above.

Finding the optimal basis of projection for SPADE is not as straightforward as in the single-photon PSF cases. Although the derivative of the biphoton PSF with respect to the signal photon's coordinates can be calculated directly (Appendix B), this is a nonseparable state for which it is challenging to devise a single-shot projector experimentally. However, we note that interesting progress is being made in that direction [27]). Instead, as in ordinary SPADE, we consider the projection on the basis of direct products of HG modes $\{|m, n\rangle_s \langle m, n| \otimes |m, n\rangle_i \langle m, n|\}$. The expression for a projection onto mode $\mathrm{HG}_{k,l}(x_i, y_i)$ in the idler arm and $\mathrm{HG}_{k',l'}(x_s, y_s)$ in the signal arm can be derived

as

$$_{k,l}^{k',l'}\mathcal{P}_{\text{coinc}} = \frac{1}{2}C_{k',l}^2\delta_{l,l'}\sum_{\pm}|\langle k|k'^{\pm}\rangle|^2. \qquad (2)$$

With $s = 0$, the correlations are diagonal: only projections where $k = k'$ have nonzero outcomes. As s increases, contributions from the first off-diagonal modes appear. Equation (2) can be given an analytical expression using the result $\langle m|n^{\pm}\rangle = \sqrt{m!/n!}\,2^{(m-n)/2}(\pm s/2)^{n-m}e^{-s^2/16}\mathcal{L}_m^{n-m}(s^2/8)$, where $\mathcal{L}_i^{\alpha}$ is the αth generalized Laguerre polynomial of order i, α and $n \geq m$ (see Appendix C for the derivation).

For small values of s, we consider the expansion of $_{k,l}^{k',l'}\mathcal{P}_{\text{coinc}}$ to the order of $(s/2)^2$ by substituting $|m^{\pm}\rangle = |m\rangle \pm s|m\rangle'/4 + O(s^2/16)$ into Eq. (2). The result $\mathcal{P}_{kl}^{k'l'} \approx C_{k',l}^2\{\delta_{k,k'}\{1 - s^2(2k'+1)/8\} + \delta_{k,k'-1}k'/2 + \delta_{k,k'+1}(k'+1)/2\}\delta_{l,l'}$ implies that, for each mode k considered in the signal beam, only the neighboring k' modes give nonzero coincidence counts: $\mathcal{P}_{kl}^{kl} \approx C_{k,l}^2(1 - s^2(2k+1)/8)$, $\mathcal{P}_{kl}^{k+1,l} \approx s^2C_{k+1,l}^2(k+1)/8$, $\mathcal{P}_{kl}^{k-1,l} \approx s^2C_{k-1,l}^2k/8$. Of these terms, only the $k = k', k \pm 1$ modes varying quadratically with the separation will translate to a nonzero Fisher information in the small separation regime, $s \to 0$ (see Appendix D):

$$\mathcal{FI}_{kl}^{kl}(s) \approx \frac{[C_{k,l}^2(2k+1)s/2]^2}{4 - s^2(2k+1)/2} \to 0, \qquad (3a)$$

$$_{k}^{k+1}\mathcal{FI}_{\text{coinc}} = \frac{1}{2}(k+1)C_{k+1,l}^2, \qquad (3b)$$

$$_{k}^{k-1}\mathcal{FI}_{\text{coinc}} = \frac{1}{2}kC_{k-1,l}^2. \qquad (3c)$$

Since the Fisher information in these particular projections is independent of the separation s, the variance of our separation estimates does not suffer from the so-called Rayleigh curse, similar to the SPADE technique for a single photon.

The total Fisher information can then be obtained by summing the individual contributions from all of the detected modes. Interestingly, for small values of s, the total has an upper bound dictated by the strength of the spatial correlation between the photon pair. More precisely,

$$_{2D}^{\text{tot}}\mathcal{FI}_{\text{coinc}} = \gamma + \gamma^{-1} = \frac{\sqrt{K}}{2} \geq \frac{1}{2}, \qquad (4)$$

where $K = \frac{1}{4}[\gamma + 1/\gamma]^2$ is the Schmidt rank often used to quantify the strength of bipartite entanglement (see
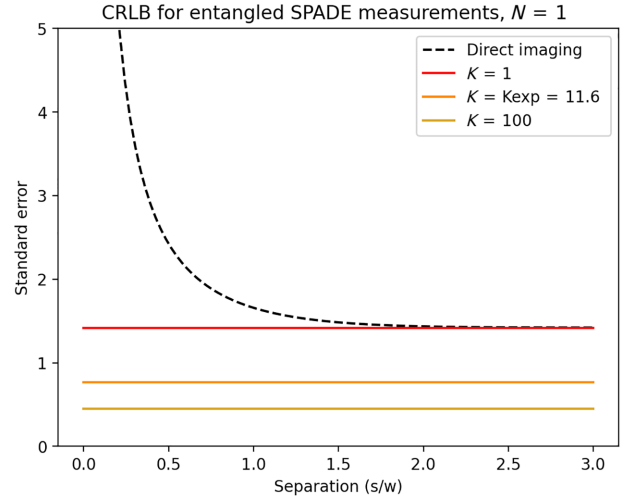

FIG. 1. Cramer Rao lower bounds (CRLBs) versus the separation (units of the Schmidt waist, $w = \sigma_s$). Classical SPADE measurements (lower bounded by $K = 1$) attain the SQL precision. Higher values of entanglement provide an advantage over classical SPADE. Kexp, experimental Schmidt number.

Appendix B for the derivation). This leads to a corresponding lower bound in the achievable precision as

$$\Delta\hat{s}^2 \geq \frac{2}{\sqrt{K}}, \qquad (5)$$

where $\Delta\hat{s}^2$ is the variance in the separation estimates. The result suggests that the precision in the measurement of near-zero separations is limited only by the strength of the spatial correlations, thereby suggesting high-dimensional entanglement as a promising resource for superresolution measurements. Not only does the method circumvent the so-called "Rayleigh curse" like the ordinary SPADE (which is bounded by the special case of $K = 1$; see Fig. 1), but it also offers further enhancement proportionate to the correlations in the system. In what follows, we describe in detail an experimental implementation and results of the method outlined above, along with some simulation results.

III. METHODS

For the purpose of a simpler experimental implementation in the laboratory, we considered a smaller subspace spanned by modes with $l = 0$ (i.e., $HG_{00}, HG_{01}, HG_{02}, \ldots,$ HG_{0n}) for both the signal and idler photons. The total Fisher information associated with this set of joint projective measurements can be written as (see Appendix D)

$$_{1D}^{\text{tot}}\mathcal{FI}_{\text{coinc}} = \frac{1}{2} + \frac{1}{2}\left(\frac{1-\gamma}{1+\gamma}\right)^2. \qquad (6)$$

Experimentally, we reconstructed the correlation matrices in the HG basis for a biphoton state generated via SPDC for different traverse separations. The separation was then inferred from the observed correlation matrices using maximum likelihood estimation.

A. Experimental realization

The down-converted photon pairs are generated by pumping a 0.5-mm-thick type-I barium borate crystal (BBO) with a 200-mW 405-nm pump beam obtained from the second harmonic generation (SHG) of a 1-GHz 100-fs laser beam. The pump beam is focused to a waist of $w_p = 40\ \mu$m at the BBO crystal, which sets a value of $\gamma = 0.15$. The value of γ determines the waist of the Schmidt modes, the Schmidt coefficients in the Hermite Gauss basis, and the Schmidt rank K [26,28]. The generated photon pairs are then split into two paths by a 50:50 beam splitter (BS), and the signal photon travels through an asymmetrical Mach-Zender interferometer (MZI), which introduces a transverse separation between the horizontal and the vertical polarization components of the beam. Two mirrors are placed on translation stages, specific combinations of which allow us to generate a tunable symmetric separation about the center of the undisplaced beam. The MZI also ensures that a path difference of approximately 20 cm exists between the horizontal and the vertical polarization components, which ensures that we have two incoherent point sources. In each of the signal and the idler arms, projections onto the spatial Schmidt modes are realized via the amplitude flattening technique [29,30], whereby a combination of appropriate holograms displayed in spatial

light modulators (SLMs) along with suitable demagnification to the fibre plane approximates the ideal projective measurement for a particular spatial mode. The first order of the diffracted light is selected with a pinhole and coupled into a single-mode fibre (SMF) after being filtered by bandpass filters 810 ± 5 nm to ensure the collection of degenerate pairs. Each SMF is then connected to a single-photon avalanche-diode (SPAD) detector (Excelitas SPCM-AQRH-14-FC), and coincidence counts are registered via a time-correlated single-photon counting system. Further details on the experimental setup with the essential components are illustrated in Fig. 2.

To determine the experimental Schmidt waist, multiple sets of data for differing waists of projection were taken for a zero displacement, and the experimental projection waist was chosen so as to have as close to a diagonal density matrix as possible. For each applied separation, the setup was realigned so that the center of the displaced beam was coupled to the fibre. The coincidence counts in the diagonal modes for zero separation account for at least 82% of the total counts in the 7×7 cross-talk matrix. The deviation from a perfectly diagonal decomposition can be attributed to several factors, including a deviation of the pump beam shape from a Gaussian mode and modal cross-talk in the detection scheme. These factors can be tackled, in principle, by improved control of the pump shape (e.g., through spatial filtering or the use of SLMs) and the use of other detection schemes, respectively. Indeed, the recent advances in spatial mode sorting through multiplane converters allow, in principle, a lossless and low cross-talk detection scheme [31].
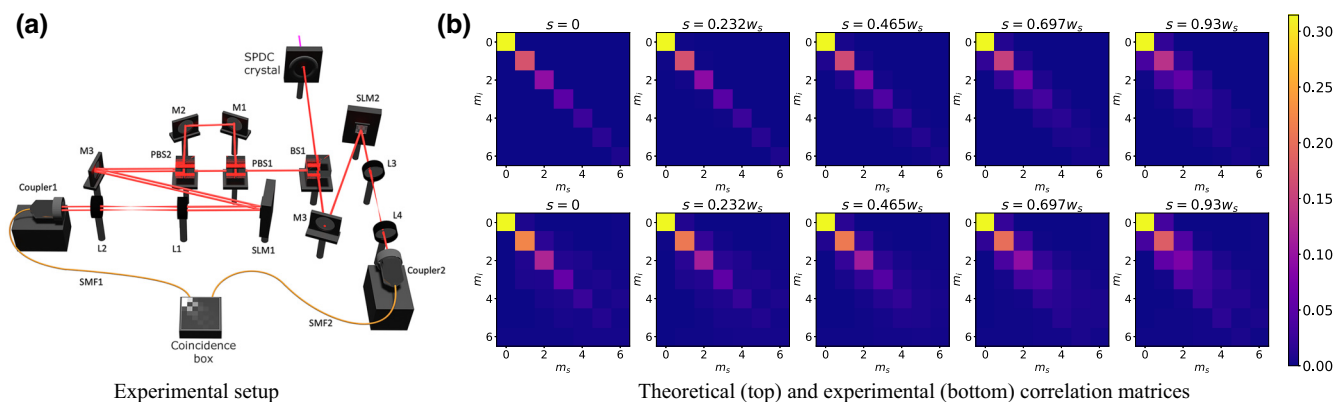

FIG. 2.   (a) A 405-nm laser shines a 1-mm type-I BBO crystal generating degenerate down-converted photon pairs. Idler and signal photons are separated by a 50:50 beam splitter (BS). An unbalanced MZI interferometer implements an incoherent lateral displacement on the signal photon. The two photons are then projected on HG modes by coupling to single-mode fibers (SMFs) after being spatially modified by holograms displayed on spatial light modulators (SLM1 and SLM2). The photons are then detected by a pair of single-photon avalanche-diode (SPAD) detectors and then the coincidence events are detected via a coincidence counting module. PBS, polarizing beam splitter; M1–M3, mirrors; L1–L4, lenses. (b) Theoretical (top) and experimental (bottom) cross-talk matrices showing coincidence counts for the first 7×7 joint projections on signal and idler photons for five different values of separation s between 0 and 0.93w. We can see the off-diagonal coefficients becoming more prominent as the separation increases.
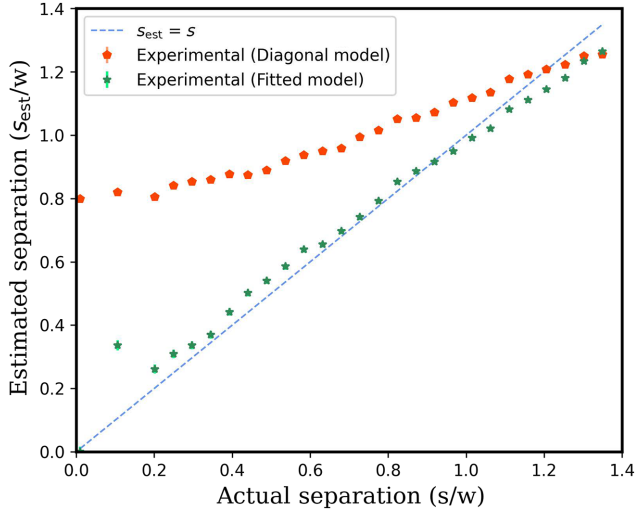
FIG. 3.   Estimates from the data versus the actual separation (units of the Schmidt waist, $w = \sigma_s$). The estimation was first performed assuming an ideal theoretical model, i.e., a perfectly diagonal mode decomposition (red) where probabilities are defined by Eq. (2). We see a significant deviation from the ideal $s_{\text{est}} = s$ estimates, most likely due to modal cross-talk, attenuation, and dark counts. Therefore, in order to account for these effects, a linear fitting of the counts, as functions of the theoretical probabilities, was performed before the estimation, the results of which closely match the ideal curve (green).
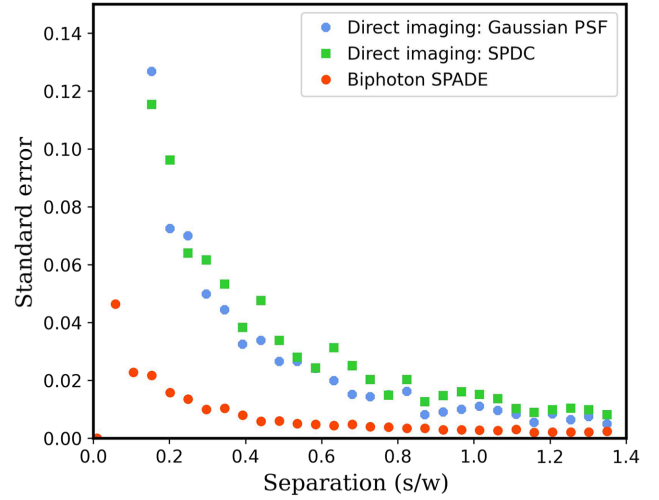

FIG. 4.   Simulation results for biphoton SPADE compared with direct imaging of both Gaussian and SPDC PSFs for $N = 37\,000$ photons (separation scaled in units of the Schmidt waist, $w = \sigma_s$). A 50-pixel array was used for direct imaging simulation, corresponding to intensity measurements at 50 different points. Similarly, for the biphoton SPADE ($K = 11.6$), the simulation was done for 49 mode projections with combinations of Hermite-Gauss modes with indices $n = 0, m = 0, \ldots, 6$ for each signal and idler. A maximum likelihood estimation was performed on generated samples, and the standard error in the separation estimates was plotted.

IV. RESULTS

We applied evenly spaced lateral displacements at increments of $0.0465\sigma_s$, from a separation $s = 0$ to $s = 1.35\sigma_s$, and recorded the outcomes of the joint measurements onto modes HG_{n_i,m_i} in the idler arm and HG_{n_s,m_s} in the signal arm for $n_{i/s} = 0$ and $m_{i/s} = 0, 1, \ldots, 6$. Each set of data is then a 7×7 matrix containing 49 joint projection outcomes, the coincidence counts for each of which were accumulated for up to 60 s. The 7×7 matrix sums were normalized to account for laser intensity fluctuations. For such small separations, to a very good approximation, it can be safely assumed that almost all of the generated photons are in the 7×7 Hilbert space. In total, for each separation, approximately 37 000 photons were collected. The results for a few sets of displacements, along with the corresponding theoretical predicted outcomes [Eq. (2)] are illustrated in the form of correlation matrices [see Fig. 2(b)]. The figure shows that the experimental outcomes closely match the corresponding theoretical predictions. We also see the modes just outside the diagonal starting to increase in significance symmetrically with increasing separation. In order to obtain the estimates of the separation parameter s, first, we perform the least-squares fitting for the normalized counts (or, equivalently, probabilities) to the expression $P_{ij} = \alpha_{ij}\mathcal{P}_{ij} + \beta_{ij}$, where the $\mathcal{P}_{ij}$ are the theoretical probabilities given by Eq. (2), α_{ij} is a scaling factor to account for attenuation for different modes, and β_{ij} is an additive factor to account for the spurious photon counts. For each separation, we then applied a maximum likelihood estimation procedure, maximizing the likelihood function defined as $L(s) := (N!/\prod n_{ij}!)\prod_{ij} P_{ij}(s)^{n_{ij}}$, where the $n_{i,j}$ are the photon counts for the joint projection onto HG_{0i} for the signal and HG_{0j} for the idler, and N is the total photon counts in all of the modes for that particular separation. The results of the estimation are shown in Fig. 3.

In order to compare the precision in the estimates obtained from this method with direct imaging, we also performed simulation runs of the experimental outcomes for the direct imaging of the SPDC point sources whose mode decomposition is characterized by $\gamma = 0.15$. The simulation results are plotted in Fig. 4. We see that the biphoton SPADE outperforms the direct imaging of the SPDC and also the direct imaging of the Gaussian with the same waist as the Schmidt waist, offering an order-of-magnitude advantage especially for near zero separations.

V. DISCUSSION AND CONCLUSION

We have theoretically derived optimal projective measurements for separation estimation using a biphoton wave function generated via the SPDC process. The Fisher information for projection onto these modes remains constant

even for near zero separation, which shows that the measurement scheme is not plagued by the so-called "Rayleigh curse" inherent in direct imaging. Moreover, the precision in the proposed method is limited only by the strength of the spatial correlations inherent in the photon pairs. In particular, we show that the total Fisher information of the separation estimation is proportional to the square root of the Schmidt rank (K) of the system. Therefore, depending upon the strength of the correlations in the system, the method even offers further enhancement that scales with K compared to ordinary SPADE, which corresponds to the trivial case of $K = 1$. As a proof of concept, we also experimentally implemented the scheme using down-converted photon pairs and obtained results that are in good agreement with the theoretical predictions. We also performed simulations comparing the precision in our estimates with the case of direct imaging, where we indeed observed at least an order-of-magnitude improvement, especially for near zero separations, where the errors start to diverge. Our theoretical and experimental results suggest the potential of using high-dimensional entanglement for further enhancing superresolution measurements. Combined with neural networks (see, e.g., Ref. [19]), this technique can provide a new tool for high-resolution imaging in low-light conditions of passive samples with variable transmissivity. As a perspective, we believe that these ideas may also have an appeal in electron microscopy where the resolution problem is very relevant, especially when trying to separate two close pointlike objects such as atoms. Optimal measurements (to which SPADE belongs) in the single electron framework have been already studied [32], but the recent advances in controlling entangled electrons [33,34] could pave the way to increase the resolution below 10 pm, which would be the highest spatial resolution achieved so far. More generally, the Schmidt decomposition of an electron pair could provide a new framework for advanced metrology in quantum transmission electron microscopy.

APPENDIX A: OPTIMAL PROJECTION

Fisher information describes the amount of information a measurement contains about a certain parameter being estimated. The Cramer-Rao lower bound, which is the minimum possible variance on the unbiased estimation of the parameter, is given by the inverse of the Fisher information [22],

$$\sigma_{\text{crlb}} \geq \frac{1}{\sqrt{\mathcal{FI}}} \geq \frac{1}{\sqrt{\text{Tr}[(\partial\rho/\partial\delta)L]}}, \quad (A1)$$

where ρ is the state of the system, δ is the parameter to be estimated, and L is the so-called symmetric logarithmic operator defined as $\rho L + L\rho = 2\partial\rho/\partial\delta$. The higher the Fisher information associated with a particular set of measurements, the better the precision in the estimates.

In superresolution schemes for separation estimation with the so-called spatial mode demultiplexing, a projective measurement is carried out in spatial modes (most commonly in the Hermite-Gauss basis). The Fisher information of a parameter δ, in the basis of such a set of spatial modes $\{|j\rangle\langle j|\}$, is given by

$$\mathcal{FI} = \sum_j \frac{1}{\mathcal{P}_j}\left(\frac{\partial\mathcal{P}_j}{\partial\delta}\right)^2, \quad (A2)$$

where $\mathcal{P}_j$ is the probability associated with the successful projection in the corresponding mode $|j\rangle$.

In the small separation regime, where errors from direct imaging start to diverge, one can find an optimal basis of projective measurements for which the Fisher information remains constant. For a general PSF $|\Psi(x)\rangle$, the "optimal" mode of projection is proportional to its derivative, which can be seen as follows. The displaced PSFs $\langle x_s|\Psi\rangle^{\pm} = \langle x_s|\Psi(x\pm s)\rangle$, where $\delta = 2s$, can be expanded in terms of the derivatives of the original PSF as

$$\langle x_s|\Psi\rangle^{\pm} = \sum_n \frac{(\pm s)^n \langle x_s|\Psi\rangle^{(n)}}{n!}, \quad (A3)$$

where $\langle x_s|\Psi\rangle^{(n)}$ is the nth derivative in x_s of $\langle x_s|\Psi\rangle$. For small separations, we can consider only up to the first-order term in the separation,

$$\langle x_s|\Psi\rangle^{\pm} \approx \langle x_s|\Psi\rangle \pm s\langle x_s|\Psi\rangle^{(1)} + O(s^2). \quad (A4)$$

If the two displaced PSFs are incoherent then the resulting state is characterized by a completely mixed state,

$$\rho = \tfrac{1}{2}[|\Psi\rangle^+\langle\Psi| + |\Psi\rangle^-\langle\Psi|], \quad (A5)$$

where $|\Psi\rangle^{\pm}$ refer to the wave functions displaced by $\pm s$ in the x axis.

If we substitute Eq. (A4) into Eq. (A5), we find the density matrix approximated for small separations,

$$\rho \approx |\Psi\rangle\langle\Psi| + s^2|\Psi\rangle^{(1)}\langle\Psi|^{(1)} + \frac{s^2}{2}(|\Psi\rangle^{(2)}\langle\Psi| + |\Psi\rangle\langle\Psi|^{(2)}). \quad (A6)$$

The probability of successful projection onto the normalized first derivative mode is

$$
\begin{aligned}
\mathcal{P}_{\text{der}} &\approx \frac{1}{\mathcal{N}} \langle \Psi |^{(1)} \rho_{\text{BI}} | \Psi \rangle^{(1)} \\
&\approx \frac{s^2}{\mathcal{N}} | \langle \Psi |^{(1)} | \Psi \rangle^{(1)} |^2 \\
&\approx \frac{s^2}{\mathcal{N}} \\
&= \frac{\delta^2}{4\mathcal{N}},
\end{aligned}
\tag{A7}
$$

where $\mathcal{N}$ is the normalization factor. The Fisher information can be calculated as

$$
\mathcal{FI}_{\text{der}} = \frac{1}{\mathcal{P}_{\text{der}}} \left(\frac{\partial \mathcal{P}_{\text{der}}}{\partial \delta} \right)^2 = \mathcal{N}.
\tag{A8}
$$

We see that in the small separation regime, the Fisher information in the derivative mode is a constant regardless of the separation.

APPENDIX B: SCHMIDT DECOMPOSITION OF SPDC

The biphoton state generated in the SPDC process can be decomposed into a superposition of the tensor product of two linear combinations of Hermite-Gauss modes [Eq. (B1)],

$$
| \Psi \rangle = \sum_{m,n} \sum_{m',n'} C_{mn}^{m'n'} | m, n \rangle_s \otimes | m', n' \rangle_i.
\tag{B1}
$$

Hermite-Gauss $\mathcal{HG}$ modes $| m, n \rangle = | m \rangle \otimes | n \rangle$ form an orthonormal basis of spatial modes for the transverse plane and have the following expressions (2D and 1D, respectively) in adimensional coordinates:

$$
\langle x_o, y_o | m, n \rangle = \frac{1}{\sqrt{2^m 2^n m! \, n! \, \pi}} \mathcal{H}_m(x_o) e^{-x_o^2/2} \otimes \mathcal{H}_n(y_o) e^{-y_o^2/2},
\tag{B2}
$$

$$
\langle x_o | m \rangle = \frac{1}{\sqrt{2^m m! \, \sqrt{\pi}}} \mathcal{H}_m(x_o) e^{-x_o^2/2}.
\tag{B3}
$$

Here $\mathcal{H}_i(x_k)$ is the ith-order Hermite polynomial for variable x_k, and $x_o = \sqrt{2}x/\sigma_s$ and $y_o = \sqrt{2}y/\sigma_s$, as defined in the main text. When expressed in the Schmidt basis of Hermite Gauss $\mathcal{HG}$ modes, the original SPDC state can be

simplified to [24,26,28]

$$
| \Psi \rangle = \sum_{m,n} C_{mn} | m, n \rangle_s \otimes | m, n \rangle_i,
\tag{B4}
$$

where

$$
C_{mn} = \frac{4\gamma}{(1 + \gamma)^2} \left| \frac{1 - \gamma}{1 + \gamma} \right|^{(m+n)}
$$

and γ is a constant determined by source parameters such as the crystal length and pump beam waist. When performing any sort of joint measurement (i.e., coincidence counts or coincidence imaging) on the entangled state, we consider a PSF of the form of the biphoton expression given in Eq. (B4). If a separation is introduced in the signal beam, it becomes a mixed state of two separated beams with the signal photon displaced in the x_s axis. The two-photon density matrix ρ_{BI} resembles Eq. (A5), with $\langle x_s | m^\pm \rangle_s = \sqrt{1/(2^m m! \sqrt{\pi})} \mathcal{H}_m(x_s \pm s) e^{-(x_s \pm s)^2/2}$:

$$
\begin{aligned}
\rho_{\text{BI}} = \sum_{m,n} \sum_{u,v} C_{mn} C_{uv}^* | m, n \rangle_i \, {}_i\langle u, v | \\
\otimes [| m^+, n \rangle_s \, {}_s\langle u^+, v | + | m^-, n \rangle_s \, {}_s\langle u^-, v |].
\end{aligned}
\tag{B5}
$$

The derivative of the PSF [Eq. (B4)] in x_s has the form

$$
| \Psi \rangle^{(1)} = \sum_{m,n} C_{mn} | m \rangle_s^{(1)} | n \rangle_s \otimes | m, n \rangle_i,
$$

where we can find from the recurrence relations of the Hermite polynomials that

$$
| m \rangle_s^{(1)} = \sqrt{\frac{m}{2}} | m - 1 \rangle_s - \sqrt{\frac{m+1}{2}} | m + 1 \rangle_s.
\tag{B6}
$$

So, the derivative becomes a maximally entangled state with the expression

$$
\begin{aligned}
| \Psi \rangle^{(1)} = \sum_{m,n} C_{m,n} \left(\sqrt{\frac{m}{2}} | m - 1 \rangle_s - \sqrt{\frac{m+1}{2}} | m + 1 \rangle_s \right) | n \rangle_s \\
\otimes | m, n \rangle_i.
\end{aligned}
\tag{B7}
$$

Projecting onto such a state is not trivial and is outside the scope of this experiment. We instead chose to use measurements readily available to us experimentally in a setup with entangled photon pairs. A variety of measurements (either categorized as joint or single-photon measurements) are possible that will determine the projection outcomes $\mathcal{P}_i$. Using a joint basis of $\mathcal{HG}$ modes as the mode sorting basis, with elements $| k, l \rangle_s \, {}_s\langle k, l | \otimes | k', l' \rangle_i \, {}_i\langle k', l' |$, the projection outcomes $\mathcal{P}_{kl}^{k'l'}$ can be expressed as

$$
\begin{aligned}
\mathcal{P}_{kl}^{k'l'} &= \text{Tr}[| k, l \rangle_s \, {}_s\langle k, l | \otimes | k', l' \rangle_i \, {}_i\langle k', l' | \rho_{\text{BI}}] \\
&= {}_s\langle k, l | \otimes {}_i\langle k', l' | \rho_{\text{BI}} | k, l \rangle_s \otimes | k', l' \rangle_i \\
&= \tfrac{1}{2} C_{k',l}^2 \{ |\langle k | k'^+ \rangle|^2 + |\langle k | k'^- \rangle|^2 \} \delta_{l,l'}.
\end{aligned}
\tag{B8}
$$

APPENDIX C: ANALYTICAL EXPANSION OF $\langle m | n^\pm \rangle$

The inner product between an $\mathcal{HG}$ function of mode m and a displaced $\mathcal{HG}$ function of mode n can be written as

$$\langle m | n^\pm \rangle = \int_{-\infty}^{+\infty} \mathcal{HG}_m(x) \mathcal{HG}_n(x \pm s) dx.$$

Applying a change of variables $u = x \pm s/2$ and inserting the expressions into Eq. (B3), we obtain

$$\langle m | n^\pm \rangle = \frac{1}{\sqrt{2^m 2^n m! \, n! \, \pi}} \int_{-\infty}^{+\infty} e^{-(1/2)[(u \mp s/2)^2 + (u \pm s/2)^2]}$$
$$\mathcal{H}_m\left(u \mp \frac{s}{2}\right) \mathcal{H}_n\left(u \pm \frac{s}{2}\right) du$$
$$= \frac{e^{-s^2/4}}{\sqrt{2^{m+n} m! n! \pi}} \int_{-\infty}^{+\infty} e^{-x^2}$$
$$\mathcal{H}_m\left(u \mp \frac{s}{2}\right) \mathcal{H}_n\left(u \pm \frac{s}{2}\right) du. \qquad (C1)$$

The integral in Eq. (C1) is given in a table of integrals [35], where, for $n \geq m$,

$$\int_{-\infty}^{+\infty} e^{-x^2} \mathcal{H}_m(x+y) \mathcal{H}_n(x+z) dx$$
$$= 2^n \pi m! \, z^{n-m} \mathcal{L}_m^{n-m}(-2yz) \qquad (C2)$$

with $\mathcal{L}_i^\alpha$ the αth generalized Laguerre α polynomial of order i, α. So, the analytical expression for the inner product for $n \geq m$ is

$$\langle m | n^\pm \rangle = \sqrt{\frac{m!}{n!}} 2^{(m-n)/2} (\pm s)^{n-m} e^{-s^2/4} \mathcal{L}_m^{n-m}\left(\frac{s^2}{2}\right). \quad (C3)$$

APPENDIX D: BIPHOTON SPADE BASED ON PROJECTION ONTO HG MODES

Here, we derive the main result showing the relation between the Fisher information and the Schmidt rank (K) in biphoton SPADE. In the small separation limit [using Eqs. (A4) and (B6)] only three terms in Eq. (B8) are nonzero:

$$\mathcal{P}_{kl}^{k'l'} \approx C_{k',l}^2 \left[\delta_{k,k'}^2 \left\{ 1 - \frac{s^2}{2}(2k'+1) \right\} \right.$$
$$\left. + \delta_{k,k'-1}^2 \frac{k'}{2} + \delta_{k,k'+1}^2 \frac{k'+1}{2} \right] \delta_{l,l'}^2, \qquad (D1)$$

$$\mathcal{P}_{kl}^{kl} \approx C_{k,l}^2 \left(1 - \frac{s^2}{2}(2k+1) \right), \quad \mathcal{P}_{kl}^{k+1,l} \approx s^2 C_{k+1,l}^2 \frac{k+1}{2},$$

$$\mathcal{P}_{kl}^{k-1,l} \approx s^2 C_{k-1,l}^2 \frac{k}{2}. \qquad (D2)$$

For $k = k'$, the $\mathcal{FI}$ about the parameter $\delta = 2s$ tends to 0. For $k = k' \pm 1$, the projection outcome has the same form as for optimal measurements, where $\mathcal{P}_i$ is proportional to δ^2:

$$\mathcal{FI}_{kl}^{kl}(s) \approx \frac{\left[C_{k,l}^2 (2k+1)s \right]^2}{4 - 2s^2(2k+1)} \to 0, \qquad (D3a)$$

$$\mathcal{FI}_{kl}^{k+1,l} \approx \frac{k+1}{2} C_{k+1,l}^2, \qquad (D3b)$$

$$\mathcal{FI}_{kl}^{k-1,l} \approx \frac{k}{2} C_{k-1,l}^2. \qquad (D3c)$$

For example, projecting the signal photon onto $|1,0\rangle_s$ and coupling the idler photon into a single-mode fiber (projection onto $|0,0\rangle_i$) would give a constant Fisher information, for small separations, of $\frac{1}{2} C_{0,0}^2$. The total $\mathcal{FI}$ is obtained by summing the nonvanishing $\mathcal{FI}$ across all HG mode measurements. Writing out the coefficients explicitly,

$$\mathcal{FI}_{kl}^{k+1,l} \approx \frac{k+1}{2} C_{0,0}^2 \left(\frac{1-\gamma}{1+\gamma} \right)^{2(k+1)} \left(\frac{1-\gamma}{1+\gamma} \right)^{2l},$$

$$\mathcal{FI}_{kl}^{k-1,l} \approx \frac{1}{2} k C_{0,0}^2 \left(\frac{1-\gamma}{1+\gamma} \right)^{2(k-1)} \left(\frac{1-\gamma}{1+\gamma} \right)^{2l},$$

and using

$$\sum_k k \left(\frac{1-\gamma}{1+\gamma} \right)^{2k} = \frac{(1-\gamma)^2(1+\gamma)^2}{16\gamma^2},$$

$$\sum_k \left(\frac{1-\gamma}{1+\gamma} \right)^{2k} = \frac{(1+\gamma)^2}{4\gamma},$$

we can calculate the sum over k for the 1D case,

$$\sum_k \mathcal{FI}_{kl}^{k+1,l} = \frac{1}{2} \left(\frac{1-\gamma}{1+\gamma} \right)^2 \approx 0.27, \qquad (D4a)$$

$$\sum_k \mathcal{FI}_{kl}^{k-1,l} = \frac{1}{2}, \qquad (D4b)$$

or the sum over all k, l,

$$\sum_{k,l} \mathcal{FI}_{kl}^{k+1,l} = \frac{(1-\gamma)^2}{8\gamma} \approx 0.60, \qquad (D5a)$$

$$\sum_{k,l} \mathcal{FI}_{kl}^{k-1,l} = \frac{(1+\gamma)^2}{8\gamma} \approx 1.10. \qquad (D5b)$$

The numerical values correspond to the Fisher information value for $\gamma = 0.15$, as in the experiment. Note that,

for strong spatial correlations, the value of γ gets closer to 0 and the total Fisher information increases. In fact, the Schmidt number $K = (\gamma + \gamma^{-1})^2/4$, which is often used as a measure to describe the strength of spatial entanglement in the SPDC pair, can be found in the total Fisher information sum:

$$\sum_{k,l} \mathcal{FI}_{\text{coinc}} = \frac{1}{4}(\gamma + \gamma^{-1}) = \frac{1}{2}\sqrt{K} \geq \frac{1}{2}. \qquad (D6)$$

APPENDIX E: GAUSSIAN CASE ($\gamma = 1$)

Note that $\gamma = 1$ corresponds to a simple separable product state of a two Gaussian wave functions for the signal and idler. In that case, the total Fisher information according to Eq. (D6) reduces to $\frac{1}{2}(N/\sigma_s^2)$, which can be briefly seen as follows. The state after displacement can be represented by a density matrix $\rho = \frac{1}{2}[|0^+0\rangle\langle0^+0| + |0^-0\rangle\langle0^-0|]$. At the small separation regime the optimal mode of the projection is the first-order Hermite-Gauss mode $\mathcal{HG}_{01}$, for which the probability for a successful projection is given by

$$P_1 = \langle10|\rho|10\rangle = \frac{1}{2}[|\langle1|0^+\rangle|^2 + |\langle1|0^-\rangle|^2].$$

Using Eq. (C3), we have $\langle1|0^+\rangle = \langle1^-|0\rangle = -\sqrt{1/2}\,se^{-s^2/4}$ and $\langle1|0^-\rangle = \langle1^+|0\rangle = \sqrt{1/2}\,se^{-s^2/4}$. The probability outcome P_1 is given by

$$P_1 = \frac{1}{2}s^2 e^{-s^2/2}$$

with Fisher information

$$\mathcal{FI}_1 = \frac{1}{P_1}\left(\frac{\partial P_1}{\partial(2s)}\right) = \frac{1}{4}\left(2 - 2s^2 + \frac{s^4}{2}\right)e^{-s^2/2} \approx \frac{1}{2}\left(\frac{N}{\sigma_s^2}\right).$$

The total Fisher information for the simple Gaussian case does indeed correspond to the total Fisher information when $\gamma = 1$. The entanglement scheme thus provides an advantage over classical SPADE for any level of entanglement with $\gamma < 1$.

[1] L. Rayleigh, XXXI. Investigations in optics, with special reference to the spectroscope, The London, Edinburgh, and Dublin, Philos. Mag. J. Sci. **8**, 261 (1879).

[2] E. Ash and G. Nicholls, Super-resolution aperture scanning microscope, Nature **237**, 510 (1972).

[3] E. Betzig and J. K. Trautman, Near-field optics: Microscopy, spectroscopy, and surface modification beyond the diffraction limit, Science **257**, 189 (1992).

[4] M. Von Ardenne, Das elektronen-rastermikroskop, Z. Phys. **109**, 553 (1938).

[5] J. Tersoff and D. R. Hamann, Theory and Application for the Scanning Tunneling Microscope, Phys. Rev. Lett. **50**, 1998 (1983).

[6] M. Tsang, R. Nair, and X.-M. Lu, Quantum Theory of Superresolution for Two Incoherent Optical Point Sources, Phys. Rev. X **6**, 031033 (2016).

[7] R. Kerviche, S. Guha, and A. Ashok, in *2017 IEEE International Symposium on Information Theory (ISIT)* (IEEE, Aachen, Germany, 2017), p. 441.

[8] M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto, and J. Rehacek, Achieving the ultimate optical resolution, Optica **3**, 1144 (2016).

[9] P. Boucher, C. Fabre, G. Labroille, and N. Treps, Spatial optical mode demultiplexing as a practical tool for optimal transverse distance estimation, Optica **7**, 1621 (2020).

[10] W.-K. Tham, H. Ferretti, and A. M. Steinberg, Beating Rayleigh's Curse by Imaging Using Phase Information, Phys. Rev. Lett. **118**, 070801 (2017).

[11] Z. Hradil, J. Řeháček, L. Sánchez-Soto, and B.-G. Englert, Quantum Fisher information with coherence, Optica **6**, 1437 (2019).

[12] K. Liang, S. A. Wadood, and A. Vamivakas, Coherence effects on estimating two-point separation, Optica **8**, 243 (2021).

[13] S. Kurdzialek, Back to sources—The role of losses and coherence in super-resolution imaging revisited, Quantum **6**, 697 (2022).

[14] M. Gessner, C. Fabre, and N. Treps, Superresolution Limits from Measurement Crosstalk, Phys. Rev. Lett. **125**, 100501 (2020).

[15] Y. L. Len, C. Datta, M. Parniak, and K. Banaszek, Resolution limits of spatial mode demultiplexing with noisy detection, Int. J. Quantum Inf. **18**, 1941015 (2020).

[16] J. Yang, S. Pang, Y. Zhou, and A. N. Jordan, Optimal measurements for quantum multiparameter estimation with general states, Phys. Rev. A **100**, 032104 (2019).

[17] S. Zhou and L. Jiang, Modern description of Rayleigh's criterion, Phys. Rev. A **99**, 013808 (2019).

[18] K. K. Bearne, Y. Zhou, B. Braverman, J. Yang, S. Wadood, A. N. Jordan, A. Vamivakas, Z. Shi, and R. W. Boyd, Confocal super-resolution microscopy based on a spatial mode sorter, Opt. Express **29**, 11784 (2021).

[19] A. Pushkina, G. Maltese, J. Costa-Filho, P. Patel, and A. Lvovsky, Superresolution Linear Optical Imaging in the Far Field, Phys. Rev. Lett. **127**, 253602 (2021).

[20] D. Sych, V. Averchenko, and G. Leuchs, Generic method for lossless generation of arbitrarily shaped photons, Phys. Rev. A **96**, 053847 (2017).

[21] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, AVS Quantum Sci. **2**, 024703 (2020).

[22] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).

[23] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[24] F. M. Miatto, H. di Lorenzo Pires, S. M. Barnett, and M. P. van Exter, Spatial Schmidt modes generated in parametric down-conversion, Eur. Phys. J. D **66**, 1 (2012).

[25] M. Fedorov, Y. M. Mikhailova, and P. Volkov, Gaussian modelling and Schmidt modes of SPDC biphoton states, J. Phys. B: At., Mol. Opt. Phys. **42**, 175503 (2009).

[26] F. M. Miatto, T. Brougham, and A. M. Yao, Cartesian and polar Schmidt bases for down-converted photons, Eur. Phys. J. D **66**, 1 (2012).

[27] G. S. Thekkadath, L. Giner, X. Ma, J. Flórez, and J. S. Lundeen, Projecting onto any two-photon polarization state using linear optics, New J. Phys. **20**, 083033 (2018).

[28] S. Straupe, D. Ivanov, A. Kalinkin, I. Bobrov, and S. Kulik, Angular Schmidt modes in spontaneous parametric down-conversion, Phys. Rev. A **83**, 060302 (2011).

[29] E. Bolduc, N. Bent, E. Santamato, E. Karimi, and R. W. Boyd, Exact solution to simultaneous intensity and phase encryption with a single phase-only hologram, Opt. Lett. **38**, 3546 (2013).

[30] F. Bouchard, N. H. Valencia, F. Brandt, R. Fickler, M. Huber, and M. Malik, Measuring azimuthal and radial modes of photons, Opt. Express **26**, 31925 (2018).

[31] N. K. Fontaine, R. Ryf, H. Chen, D. T. Neilson, K. Kim, and J. Carpenter, Laguerre-Gaussian mode sorter, Nat. Commun. **10**, 1 (2019).

[32] F. Troiani, E. Rotunno, S. Frabboni, R. B. Ravelli, P. J. Peters, E. Karimi, and V. Grillo, Efficient molecule discrimination in electron microscopy through an optimized orbital angular momentum sorter, Phys. Rev. A **102**, 043510 (2020).

[33] S. Meier, J. Heimerl, and P. Hommelhoff, Few-electron correlations after ultrafast photoemission from nanometric needle tips, Nat. Phys., 1 (2023).

[34] R. Haindl, A. Feist, T. Domröse, M. Möller, J. H. Gaida, S. V. Yalunin, and C. Ropers, Coulomb-correlated electron number states in a transmission electron microscope beam, Nat. Phys., 1 (2023).

[35] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, San Diego, CA, 2014).