Certification of high-dimensional entanglement and Einstein-Podolsky-Rosen steering with cold atomic quantum memory

Michał Dąbrowski,1,2 Mateusz Mazelanik,1,2,∗Michał Parniak,1,2
Adam Leszczyński,1,2 Michał Lipka,1,2 and Wojciech Wasilewski2

1Institute of Experimental Physics, Faculty of Physics,
University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland
2Centre for Quantum Optical Technologies, Centre of New Technologies,
University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland

The advent of complex, structured and high-dimensional entangled states bring both new possibilities for experimental and theoretical scenarios as well as new challenges for generation and characterization of such states. In particular, spatially-structured photonic states offer applications in quantum imaging, information processing, and quantum key distribution. Here we experimentally generate a spatially entangled high-dimensional state composed of at least 10 Schmidt modes in a quantum memory setup and perform characterization using the entropic EPR-steering inequality, yielding genuine violation of $1.06 \pm 0.15$ bits. The entanglement of formation of at least $0.70 \pm 0.15$ ebits for the measured noisy state is certified using the entropic witness method. We point out and solve the difficulties in estimating the entropy, achieving characterization of the high-dimensional entangled state with highly undersampled data. Finally, the practical supremacy of the entropic EPR-steering witness over a variance-based witness is demonstrated for a wide class of states typical in an experimental scenario, giving prospects for EPR-steering demonstrations and applications in noisy systems or with lossy quantum channels.

\[ \Delta I \geq \log(\pi e) - \left[ h(x^B | x^A) + h(k^B_x | k^A_x) \right]_{\Sigma_h}, \] (1)

where $h(x^B | x^A)$ and $h(k^B_x | k^A_x)$ are conditional differential Shannon entropies for measurement in two conjugate, e.g. position and momentum bases. An estimated positive key rate certifies EPR-steering. Henceforth we use the binary logarithm and thus the entropy is expressed in bits. Another commonly used witness for EPR-steering, apart from the conditional entropies sum $\Sigma_h$ [see Eq. (1)], is the product of variances [15, 21].

\[ \Pi_{\Delta^2} = \langle \Delta^2(x^B - x^A)\rangle \langle \Delta^2(k^B_x + k^A_x)\rangle, \] (2)

which certifies EPR-steering if $\Pi_{\Delta^2} < 1/4$. This witness is well-suited to Gaussian states as any EPR-steerable Gaussian state is properly characterized by the variance-based witness, i.e. it will always indicate EPR-steering when steering is present. In consequence, it has been widely used in demonstrations of entanglement and EPR-steering in continuous-variable systems such as quadratures of light or collective spin [22, 24]. This translates well to QKD where it has been shown that only Gaussian attacks need to be considered to prove security [25]. Although, Eq. (1) is more inclusive than Eq. (2) in a sense that for a broad range of EPR-steerable states Eq. (1) confirms the EPR-steering while Eq. (2) criterion might not [11, 20, 26].

\[ * \text{mateusz.mazelanik@fuw.edu.pl} \]
In this Letter we demonstrate generation and characterization of an HD entangled and EPR-steerable state of a single photon (A) and a collective atomic excitation stored in a cold rubidium ensemble, retrievable after quantum memory storage time as another single photon (B). The HD character relies on a wide space of angular emission modes and the conjugate position-space of the atomic ensemble [27, 29]. The HD EPR-steering is demonstrated for the first time using a single-photon resolving camera, which offers parallel access to all coincidences and is thus necessary to truly utilize the high informational capacity of the generated state [30]. Furthermore, we show that for the emerging domain of HD entanglement in position-momentum [31, 33] (but also time-frequency [34, 36] or optical angle-orbital angular momentum (OAM) [37, 40]) variables of single photons the entropic witness vastly outperforms the variance-based witness and consequently constitutes a promising way of demonstrating HD entanglement in real systems. Finally, we use a recent result of Schneeloch and Howland [41] to estimate the entanglement of formation (EOF) of the generated quantum state. The integration of the quantum memory presented here with a practical QKD scheme provides an instantaneous advantage however here we show that the presence of accidental counts must be effectively doubled the key rate generation. A quantum memory also opens new avenues to fundamental studies of one-particle uncertainty principles [42].

II. CERTIFIABILITY OF EPR-STEERING

The character of the HD entangled state of a photon pair differs significantly from the one of quadrature squeezed states of light [24], traditionally understood as continuous-variable states. Even if the photon pair is prepared in a Gaussian-shaped state [32]

\[ \psi(x^A, x^B) = \frac{\kappa}{\pi \sigma} \exp \left( -\frac{(x^B - x^A)^2}{4\sigma^2} - \kappa^2(x^B + x^A)^2 \right), \]

where \( \sigma \) and \( \kappa^{-1} \) are the s.d. of the two Gaussians, the processes of transferring the state through a quantum channel, detection or the multi-photon pairs will yield an accidental coincidence background. The coincidence distribution is well-described by

\[ P(x^A, x^B) = q|\psi(x^A, x^B)|^2 + (1 - q)p(x^A)p(x^B), \]

where \( p(x) \) is the marginal distribution of photon counts and \( 0 \leq q \leq 1 \) is the fraction of correlated coincidences. This coincidence distribution is non-Gaussian unless the width \( \sigma = \kappa^{-1} \) or \( q=1 \), as the \( p(x^A)p(x^B) \) term is very wide in the \( x^B - x^A \) variable. An analogous scenario will be observed also for the conjugate variable – the wavenumber \( k^A \) and \( k^B \) in our case. As a result, the variance-based witness immediately fails as even for large \( q \) the huge variance of wide distribution will make the

![Figure 1](image-url)

Figure 1. Values of the entropic (a) and variance-based (b) witnesses as a function of Schmidt mode number \( K \) and correlation fraction \( q \). The denoted contour bounds correspond to EPR-steering in the case of \( \log(\pi \sigma) \) for \( \Sigma_h \) and \( 1/4 \) for \( \Pi_{\Delta^2} \), and the demonstration of \( E > 0 \) in the case of the log(2\( \pi \)) bound. The dot with errorbars shows our experimental result for the zero quantum memory storage time. Note that since detection resolution of our experiment is higher than width of correlations we may project the Shannon entropy result onto differential entropies, obtaining the same degree of violation of respective inequalities. In both cases lighter color corresponds to a higher value of the respective witness and consequently less entanglement certified in the system.

III. EXPERIMENTAL SETUP

In order to experimentally demonstrate genuine violation, we use a single-photon resolving I-sCMOS camera [44, 45] (see Appendix A for details) to capture all photons emitted from a cold atomic quantum memory. Note that previous demonstrations of HD entanglement heavily relied on subtraction of accidental background [27, 32, 33] and in consequence only estimated
the number \( K \) of Schmidt modes potentially available while not confirming the presence of certified entanglement or EPR-steering. Other experiments utilize various types of spatial modes but allowed measurement of only a single mode-pair at a time \([11, 26]\). For certification of EPR-steering we use a coarse-grained analogue of the entropic witness \([26, 47]\):

\[
\Delta I \geq \log(\frac{\pi e}{\Delta x \Delta k}) - \left[ H(x^B | x^A) + H(k^B | k^A) \right],
\]

where now \( H(x^B | x^A) \) and \( H(k^B | k^A) \) are conditional discrete Shannon entropies. Note that in our case the coarse-graining resolutions \( \Delta x, \Delta k \) are high enough (\( \Delta x < \sigma, \Delta k < \kappa \)) to well resolve the observed correlations thus the differential criterion (Eq. (1)) is simply a shifted version of Eq. (5) in which the bounds remain at specified levels so any kind of offset value such as \( \Delta I \) remains unchanged.

The EPR-steering generation between Stokes (S) photons stored in the quantum memory and collective atomic excitations retrieved in the form of scattered anti-Stokes (AS) photons relies on the off-resonant Raman scattering in a \( \Lambda \)-type system (see Fig. 2). The experimental setup is based on the cold rubidium-87 ensemble inside a magneto-optical trap (MOT) (see Appendix B for details). The scattered photons are imaged on the sCMOS camera equipped with an image intensifier \([44]\). Both near (position basis) and far field (momentum basis) imaging setups consist of four lenses: near field with magnification \( M = 8 \) and far field with effective focal length \( f_{\text{eff}} = 50 \text{ mm} \). High resolution of the near field imaging with two separate systems for S and AS photons is achieved by overlapping the focal points of the two systems in the longitudinal direction within the Rayleigh range of 38 \( \mu \text{m} \). Write pulse contains \( 10^9 \) photons, resulting in registering on average \( \bar{n} = 0.1 \) Stokes photons per camera frame. We estimate memory readout efficiency of 35\%, taken as the ratio of coincidence rate to S photons rate and accounted for losses \([29]\).

In our experimental setup the coarse-grained elementary cells \([11, 26]\) have the size of \( \Delta x = \delta_0 / M = 3.88 \mu \text{m} \) \( (\Delta k = 2\pi \delta_0 / (\lambda f_{\text{eff}}) = 4.86 \text{ mm}^{-1}) \) for near (far) field, where \( \delta_0 = 31 \mu \text{m} \) is the physical bin size on the detector plane and \( \lambda \) is scattered photon wavelength (equal either 780 nm or 795 nm). The experimental uncertainties of the magnification \( M \) and focal length \( f_{\text{eff}} \) determine the uncertainty of the right-hand-side of Eq. (5), increasing our errorbar on the attainable key rate \( \Delta I \) or entanglement of formation \( E \). On the camera we observe S and AS photons in separate circular regions of 1000\( \delta_0 \) diameter corresponding to an image of the circular region of 0.4 mm diameter for the near field (62 mrad scattering angle for the far field). These regions are selected using iris diaphragms situated in intermediate image and Fourier planes in the part of the imaging setup common to both systems (see Fig. 2(c)).

Figure 2. (a) Representation of spin-wave modes in the atomic ensemble for which the projective measurement is performed. While the far-field mode is close to a transverse plane wave, the near field mode is shaped like a tightly-focused Gaussian beam, having a wide transverse profile at the edges of the ensemble. This highlights the importance of the perfect overlap of detection modes determined by two separate systems in the longitudinal direction. (b) Experimental setup: strong write (W) laser beam produces Stokes (S) photon as a result of Raman scattering in cold \( ^{87}\text{Rb} \) ensemble (MOT). After adjustable storage time anti-Stokes (AS) photon is scattered during readout (R) process. (c) Far and near field imaging setups with iris diaphragms to select observation regions, determining the system instrumental response and the number \( K \) of observed modes. The near and far field setups differ solely by the last lens, the aperture of which is large enough to project all remaining modes onto the camera.

IV. VARIANCE-BASED WITNESS

In Fig. 3(a) and 3(b) we present joint coincidence distributions of scattered Stokes–anti-Stokes pairs for positions \( (x^S, x^{AS}) \) and momenta \( (k_x^A, k_x^{AS}) \) respectively, registered for zero memory storage time. In Fig. 3(a) we post-select from full four-dimensional distribution \( n(x^A, y^A, x^B, y^B) \) the detection events with \( y^S \approx y^S \pm 2\sigma \) (and \( k_x^{AS} = -k_x^S \pm 2\kappa \) from \( (k_x^A, k_x^B) \)) for Fig. 3(b) where \( \sigma(\kappa) \) is the size of the spatial mode in the near (far) field (inferred from Fig. 3(c) and 3(d), thus the \( y \)-dimension can be treated as a herald of coincidence event in \( x \)-direction. Assuming both \( n(x^A, x^B) \) and \( n(k_x^A, k_x^B) \) are described in terms of Gaussian function (see Eq. 3) and its Fourier transform, we estimate the number of post-selected independent spatial modes for one-dimensional situation equal \( K = 10.2 \) (ref. \([43]\)) which corresponds to \( K_{3D} \approx K^2 \approx 104 \) modes for two dimensions. Note that considering only the near or far field (comparing long and short axes of the observed Gaussian-shaped distribution \([29]\) – see Eq. (5)) we obtain slightly higher estimate of the number of modes \( K \approx 14 \) that
V. SHANNON ENTROPY-BASED WITNESS

Estimating entropy of sampled probability distribution from experimental data has been shown to be a hard task, as no unbiased non-parametric estimator exists \[25\]. Here we infer the conditional entropy values from the experimental data (Fig. 3) as a difference of joint and marginal entropy distributions \( H(B|x^A) = H(B|x^A) - H(x^A) \). Figure 3(a) presents the values of the entropic witness \( \Sigma_H \) [Eq. (5)], inferred using four different Shannon entropy estimators \[52\] (see Appendix C for details). A naïve “plugin” estimator, based on relative frequencies directly plugged into the formula for Shannon entropy as probabilities, is strongly biased towards lower entropy values, unless the total number of counts exceeds the total number of available bins by orders of magnitude. In the case of EPR-steering verification this leads to the Eq. (5) being falsely satisfied. In particular, \( \Sigma_H \) estimated using the naïve method drops with the storage time \( \tau \) as an effect of falling number of counts due to the quantum memory retrieval efficiency decay, which in turn results in a higher negative bias. This crucial problem has been addressed by Bayesian estimation through applying prior mixtures highly non-informative about the entropy: a Dirichlet mixture (NSB estimator) \[53\] and the Pitman-Yor mixture (PYM estimator) \[52\]. These estimators are specifically designed to deal with undersampled number of counts. Additionally, we construct a parametric estimator that assumes a Gaussian shape of the distribution on top of a wide accidental background. The coincident counts probability distribution is fitted to experimental data by maximizing the likelihood function and the entropy is calculated directly from the fitted distribution (ML estimator).

The consistency of ML and PYM estimators is a strong indication that the more general PYM estimator can be considered trustworthy in our particular scenario. On the other hand, NSB estimator shows a constant bias for all measurements while still performing better than the “plugin” estimator. The parametric ML and PYM estimators agreeably certify EPR-steering for \( \tau = 0 \) as well as for the \( \tau = 5 \) µs, but fail for the case of \( \tau = 10 \) µs in which the secure information capacity \( \Delta I \) drops to zero as accidental coincidences dominate over the Stokes–anti-Stokes coincidences. The two main limiting factors are the decay of memory retrieval efficiency due to motional dephasing and the increasing noise due to influx of untrapped atoms from the residual gas into the interaction region within the readout beam. Nevertheless, for \( \tau = 0 \) we obtain \( \Delta I = 1.06 \pm 0.15 \) bits per photon. We also certify entanglement of formation \( E \geq 0.70 \pm 0.15 \) ebits. We stress that even for Gaussian states not all EOF is certified with this method \[41\], and for the state generated here more ebits are likely needed to describe correlations. The number of modes \( K \) places an upper limit on EOF equal \( E \leq \log(K) \approx 3.3 \) ebits, however currently there is no method to better estimate the EOF.

Nevertheless, we may still certify the inseparability of
Such witness offers two distinct features: if $\Sigma_{H}^{(sym)} \leq \log(2\pi e/\Delta x \Delta k)$ the state is inseparable, and if $\Sigma_{H}^{(sym)} \leq \log(\pi e/\Delta x \Delta k)$, the state is EPR-steerable in both directions. The latter condition is slightly more exclusive than $\Delta I \geq 0$ [Eq. (5)], but the inseparability can be certified more easily than in the asymmetric (conditional) case. In Fig. 4(b) we plot the experimentally inferred values of the symmetric witness along with the two bounds. We clearly observe that inseparability is unambiguously certified for storage times of up to 5 $\mu$s, with 1.77 $\pm$ 0.15 bits of difference between $\Sigma_{H}^{(sym)}$ and $\log(2\pi e/\Delta x \Delta k)$ for parametric ML estimator (for $\tau = 0$). Note that in the asymmetric case presented in Fig. 4(a) the condition for inseparability overlaps with the condition for EPR-steering.

VI. DISCUSSION

In conclusion, we presented the generation and storage of an atom-photon HD EPR-steerable state without noise subtraction. We showed, using entropic steering inequality, that the atomic part of the generated state can be stored up to 5 $\mu$s, long enough to perform electronic feedback-loop based on the results of Stokes photon measurements [29]. With such system an active multiplexing technique could be employed [54] to obtain a one-dimensional stripe of entangled modes with a high correlated coincidence fraction $q$. We also demonstrate the superiority of the entropic entanglement witness – even if the variance EPR-steering inequality (Eq. 2, 15) is not satisfied, the corresponding entropic version (Eq. 1, 20 47) confirms EPR-steering. We highlighted and solved the difficulties in entropy estimation, achieving characterization of the HD entangled state with highly undersampled data.

We envisage that with higher numbers of modes $K$ the entropic witness [Eq. (1)] may detect EPR-steering even without the detection efficiency loophole. As the value of $q$ required to witness EPR-steering drops with $K$, loophole-free demonstration could be achieved for large $K$ even with low quantum efficiency of the detector [55]. Moreover, in a practical scenario a secure transmission could be established over a very lossy channel. This would be an essential step towards HD free-space QKD, particularly in space applications [56, 57].

ACKNOWLEDGMENTS

We thank I. Białynicki-Birula and R. Demkowicz-Dobrzański for discussions as well as K. Banaszek for generous support. This project has been financed by the National Science Center (Poland) (Grants No. 2015/19/N/ST2/01671, 2016/21/B/ST2/02559, 2017/25/N/ST2/01163 and 2017/25/N/ST2/00713) and Polish MNiSW Diamantowy Grant (Projects No. DI2013 011943 and DI2016 014846). M. D. is supported by the Foundation for Polish Science (FNP).
Appendix A: Detection via I-sCMOS camera

To localize photons with high spatial resolution we use an I-sCMOS camera system consisting of scientific complementary metal-oxide semiconductor (sCMOS) sensor endowed with image intensifier, assembled in our group from commercially available components. The sketch of the I-sCMOS device is depicted in Fig. 5(a).

A third generation image intensifier (Hamamatsu V7090D), powered by an external gating module consists of a photocathode, a micro-channel plate (MCP) and a phosphor screen. The image intensifier power module (Photek, GM300-3), enabling gating down to 5 ns, is gated synchronously with drive laser pulses using the FPGA. High voltage is delivered by a separate module (Photek, FPS30). Photons impinging through the input window on the 18 mm diameter gallium arsenide (GaAs) photocathode are converted to photoelectrons with 20% quantum efficiency. The details of image intensifier module are presented in Fig 5(b).

Photoelectrons are attracted from photocathode to MCP by short pulses of gate voltage \( V_{in} = -200 \) V. Normally the voltage \( V_{in} = +50 \) V and turns off the image intensifier. Electrons reach the adjacent MCP under high voltage \( 0.25 \leq V_{MCP} \leq 1.75 \) kV that accelerate electrons in \( \sim 9 \times 10^6 \) small, separated channels of ca. 6 \( \mu m \) diameter each. The channels are tilted by several degrees from the axis (and twisted in the middle to prevent direct electrons passing) thus the electrons hit the channel walls and produce a stochastically growing charge avalanche. The typical gain, i.e. the mean multiplication factor (tuned by changing the MCP voltage \( V_{MCP} \)) is about \( 10^5 \). Electrons leaving MCP are accelerated by a high voltage \( V_{out} = 6 \) kV and hit the phosphor screen producing bright flashes of green light (66 \( \mu m \) diameter, decay time below 200 ns).

The phosphor flashes emitting light in the wide angle are subsequently imaged onto the fast, low-noise 6.5 \times 6.5 \) mm pixel size sCMOS sensor (Andor Zyla 5.5 Megapixel) by a high numerical aperture relay lens (Stanford Computer Optics, \( f/\# = 1.1 \)). The focal length \( f = 105 \) mm of the relay lens and the magnification \( M = -0.44 \) result in an image size of the phosphor screen corresponding to a circle of 1230 pixels (8mm) in diameter, hence we effectively use about \( 1.2 \times 10^9 \) camera pixels altogether. We operate the camera in the fastest shutter mode enabling the readout of the full frame with 100 fps frame rate which can be significantly increased by reducing the number of readout camera rows. This is essential for collecting large photon statistics in quantum optics experiments where one can operate with low photon numbers per frame.

The individual single photon flashes are registered as 5-pixel FWHM Gaussian spots on the sCMOS sensor. A single pixel signal is digitized to 2048 analog-to-digital units (ADU). The average pixel offset is at the level of 100 ADU and varies less than 5 ADU across the sensor. Moreover, the s.d. of the signal when measuring dark counts is at the level of 2.5 ADU (1.5 electrons). The average intensity of the brightest pixel (highly random brightness is determined by the stochastic avalanche process) is approximately 500 times higher than the camera noise level thus the spot can be easily discriminated from the low-noise background.

A real-time software algorithm (C compiled DLL library) retrieves central positions of the single flashes from a raw image with 1.5 \( \mu m \) subpixel resolution and bins them into camera pixels. The central position of the spot is found by fitting a two-dimensional paraboloid to the logarithm of the pixel intensity, which must exceed the present threshold at least five times, within the three pixels radius. Fast and efficient fitting is performed by evaluating a numerical integral on the neighboring pixel intensities. The real-time algorithm reduces the amount of stored data enormously – instead of large raster camera frames we store only 16-bit unsigned integer numbers describing the transverse positions of the photons.

Appendix B: Cold atoms quantum memory

In the experimental sequence we first trap the atoms using the MOT for 1.4 ms. In the steady state we trap approximately \( N = 10^8 \) atoms at 100 \( \mu \)K. Next, we turn off the trapping magnetic field and detune the cooling laser 35 MHz to the red from the \( 5^2S_{1/2}, F = 2 \rightarrow 5^2P_{3/2}, F = 3 \) transition. The optical molasses is maintained for 300 \( \mu \)s and we reach the final temperature of \( 22 \pm 2 \) \( \mu \)K and optical depth OD = 40. The final atomic ensemble is 6-mm-long and has a transverse width of \( w = 0.6 \pm 0.1 \) mm (which we calculate as the \( 1/e^2 \) diameter of the atomic column density). After the cooling and trapping stage we optically pump the atoms into a single Zeeman \( F = 1, m_F = 1 \) state \([29]\), which takes 40 \( \mu \)s. The quantization axis is set by a weak bias magnetic field of approx. 100 mG.

![Figure 5](image_url)

Figure 5. (a) Scheme of the I-sCMOS camera used to collect single photons. (b) Image intensifier module. Photons impinging on the photocathode produce avalanches of electrons inside micro-channel plate (MCP) which results in phosphor flashes on the screen. See text for details.
The quantum memory is used once per atomic ensemble prepared with this sequence. After a short dark period of 1 µs that allows all residual excitations due to pump fields to decay we generate the bipartite entangled state by illuminating the ensemble with a 100 ns, off-resonant write pulse with wavevector $k_w$ (left-circular polarization, red-detuned by 20 MHz from the $5^{2}S_{1/2}, F = 1 \rightarrow 5^{2}P_{3/2}, F' = 2$ transition). As a result, a Raman-scattered Stokes (S) photon with a wavevector $k_S$ is generated together with a collective atomic excitation (spin wave) with a wavevector $K = k_w - k_s$. After quantum memory storage time, the spin wave is converted on-demand to an anti-Stokes (AS) photon through resonant Raman scattering with a 200 ns read pulse using with a wavevector $k_r$ (readout beam counter-propagating with respect to the write-in, right-circular polarization, resonant with the $5^{2}S_{1/2}, F = 2 \rightarrow 5^{2}P_{1/2}, F' = 2$ transition). Wavevector $k_{AS}$ of the AS photon is determined by the stored atomic excitation: $k_{AS} = K + k_r$.

We additionally use two separate (one for write-in and the other for readout) optically-pumped hot rubidium vapour cells with buffer gas and paraffin coating to obtain spatially-insensitive filtering, essential to observe spatial entanglement. The cells are pumped by strong lasers during the cooling and trapping period of the MOT. Combined average transmission of the imaging and filtering system is 40%. Finally, photons originating from the atomic quantum memory are imaged onto separate regions of the same the I-sCMOS sensor (located in the image or Fourier plane) through a diffraction-limited imaging setup, as depicted in Fig. 2c.

**Appendix C: Shannon entropy estimators**

Here, we describe the method of inferring the value of Shannon entropy estimators presented in Fig. 4 from the experimental data, measured using I-sCMOS camera. In the experiment we register the positions $(x^A, y^A)$, $(x^B, y^B)$ (momenta $(k_x^A, k_y^A)$, $(k_x^B, k_y^B)$) of photons in write-in (A) and readout (B) regions of I-sCMOS camera, depending on imaging system (near or far field).

We are working in the low photon-number regime where rarely more than single count occurs either in A and B region and post-select frames where exactly one photon is register in A and one photon is registered in B. With many frames we arrive at coincidence distribution $n(x^A, y^A, x^B, y^B)$ in the near $(n(k_x^A, k_y^A, k_x^B, k_y^B)$ in the far) field. Assuming the registered coincidences are position-correlated (momenta-anti-correlated), the $y$-dimension can be treated as a herald of coincidence event in $x$-direction. Thus, we post-select the detection events with $y^B = y^A \pm 2\sigma (k_y^B = -k_y^A \pm 2\kappa)$, where $\sigma (\kappa)$ is the size of the spatial mode in the near (far) field, and finally obtain joint, two-dimensional distribution $n(x^A, x^B)$ $(n(k_x^A, k_x^B))$. Based on joint $n(x^A, x^B)$ and marginal $n(x^A)$ coincidence distributions (integrated over all possible $x^B$ coordinate of a second photon), we infer the conditional entropy values as a difference of joint and marginal entropy distributions $H(x^B|x^A) = H(x^B, x^A) - H(x^A)$. Then we make the analogous procedure for the far field. Finally, we calculate the EPR witness [Eq. (5)]: $\Sigma_H = H(x^B|x^A) + H(k_x^B|k_x^A)$.

The details of calculation depends on specific type of estimator, as described below. For the symmetric steering witness $\Sigma_H^{sym} = H(x^B - x^A) + H(k_x^B + k_x^A)$ we use a similar procedure by first binning the coincidences in terms of the $(x^B - x^A)$, $(k_x^B + k_x^A)$ variables for the near and the far field, respectively. The final distributions from which the entropies are inferred are one-dimensional, and thus the issue of undersampling is less severe for this case. In subsequent considerations we keep the two-dimensional notation as a more general case.

A naïve “plugin” estimator is based on relative frequencies directly plugged as probabilities into the formula for Shannon entropy:

$$H_{\text{plug}}(x^A, x^B) = -\sum_{x^A, x^B} p(x^A, x^B) \log p(x^A, x^B), \quad (C1)$$

summed over $x^A, x^B$ (treated as discrete pixel coordinates), with probability $p(x^A, x^B) = n(x^A, x^B)/\sum n(x^A, x^B)$. Figure 6 presents the results of entropy sum $\Sigma_H$ simulation for different number of coincidences $N$. We randomly choose $N$ events from the probability distributions defined by Eq. (4) for the Gaussian-shaped state [Eq. (3)] and its Fourier transform. We assume the same $\sigma$, $\kappa$ and $q$ as in the real experiment and calculate the probability distribution

![Figure 6. Simulation of conditional entropy sum as a function of total number of coincidences. Different curves correspond to the estimators presented in Fig. 4. Dashed vertical line shows the number of coincidences registered in the experiment (approx. 3000 coincidences), while the solid horizontal line depicts the $\Sigma_b$, calculated from probability density distribution defined in Eq. (4). Coincidence maps present the results of near and far field distributions calculated for two particular number of coincidences, from which the entropy estimators were calculated. See text for details.](image-url)
[Eq. (4)] using $100 \times 100$ square grid in both near and far field. Figure 6 shows obtained distributions for two particular number of coincidences as well as mean results for the conditional entropy sum $\Sigma_{\beta}$ obtained using different entropy estimators, averaged over 10 different realizations. Additionally, we plot the conditional differential entropy sum $\Sigma_{h}$ from Eq. (1), corresponding to the asymptotic value when $N \to \infty$.

The “plugin” estimator is strongly biased towards lower entropy values, unless the total number of coincidences exceeds the total number of available bins by orders of magnitude. This crucial problem has been addressed by calculating a Dirichlet prior mixture (NSB estimator) [31] and the Pitman-Yor prior mixture (PYM estimator) [52], following the formulas in the respective articles. These estimators are specifically designed to deal with under-sampled number of coincidences and we can see in Fig. 6 they converge to $\Sigma_{h}$ even for relatively small number of coincidences.

Finally, we construct a parametric ML estimator assuming a Gaussian-shaped distribution on top of a wide, flat accidental background: $\lambda(x^{A}, x^{B}) = \alpha(x^{A}, x^{B})^{2} + \beta$, where $\psi(x^{A}, x^{B})$ is defined in Eq. (3). The analogous formula is obtained for the far field using Fourier transform of $\psi(x^{A}, x^{B})$. The coincidence probability distribution has been fitted to experimental or simulated data by maximizing the likelihood function $\log f(n, \lambda)$ over $\alpha, \beta$, where the averaging is carried over all values of $x$ or $k$, with $f(n, \lambda) = \lambda^{n} e^{-\lambda}/n!$ being Poissonian distribution. The entropy is calculated directly from the fitted distribution $\lambda$ using Eq. (C1).

---

[1] T. Vértesi, S. Pironio, and N. Brunner, Phys. Rev. Lett. 104, 060401 (2010)
[2] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. 74, 145 (2002)
[3] M. Krenn, M. Huber, R. Fickler, R. Lapkiewicz, S. Ramelow, and A. Zeilinger, Proc. Natl. Acad. Sci. USA 111, 6243 (2014)
[4] P. B. Dixon, G. A. Howland, J. Schneeloch, and J. C. Howell, Phys. Rev. Lett. 108, 143603 (2012)
[5] L. Zhang, C. Silberhorn, and I. A. Walmsley, Phys. Rev. Lett. 100, 110504 (2008)
[6] C. Lee et al., Phys. Rev. A 90, 062331 (2014)
[7] T. Zhong et al., New J. Phys. 17, 022002 (2015)
[8] C. Eltschka and J. Siewert, J. Phys. A 47, 424005 (2014)
[9] P. Erker, M. Krenn, and M. Huber, Quantum 1, 22 (2017)
[10] G. A. Howland, J. Schneeloch, D. J. Lum, and J. C. Howell, Phys. Rev. Lett. 112, 253602 (2014)
[11] S. P. Walborn, B. G. Taketani, A. Salles, F. Toscano, and R. L. de Matos Filho, Phys. Rev. Lett. 103, 160505 (2009)
[12] I. Kogias, A. R. Lee, S. Ragi, and G. Adesso, Phys. Rev. Lett. 114, 060403 (2015)
[13] X. Deng, Y. Xiang, C. Tian, G. Adesso, Q. He, Q. Gong, X. Su, C. Xie, and K. Peng, Phys. Rev. Lett. 118, 230501 (2017)
[14] A. Martin, T. Guerreiro, A. Tiranov, S. Designolle, F. Fröwis, N. Brunner, M. Huber, and N. Gisin, Phys. Rev. Lett. 118, 110501 (2017)
[15] M. D. Reid, Phys. Rev. A 40, 913 (1989)
[16] I. Białynicki-Birula and J. Mycielski, Commun. Math. Phys. 44, 129 (1975)
[17] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935)
[18] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. 98, 140402 (2007)
[19] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A 85, 010301 (2012)
[20] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro, Phys. Rev. Lett. 106, 130402 (2011)
[21] S. Mancini, V. Giovannetti, D. Vitali, and P. Tombesi, Phys. Rev. Lett. 88, 120401 (2002)
[22] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng, Phys. Rev. Lett. 68, 3663 (1992)
[23] N. Takei, N. Lee, D. Moriyama, J. S. Neergaard-Nielsen, and A. Furusawa, Phys. Rev. A 74, 060101 (2006)
[24] K. Jensen et al., Nat. Phys. 7, 13 (2011)
[25] F. Grosshans and N. J. Cerf, Phys. Rev. Lett. 92, 047905 (2004)
[26] D. S. Tasca, L. Rudnickii, R. M. Gomes, F. Toscano, and S. P. Walborn, Phys. Rev. Lett. 110, 210502 (2013)
[27] M. Dąbrowski, M. Parniak, and W. Wasilewski, Optica 4, 272 (2017)
[28] R. Chrapkiewicz, M. Dąbrowski, and W. Wasilewski, Phys. Rev. Lett. 118, 063603 (2017)
[29] M. Parniak, M. Dąbrowski, M. Mazelanik, A. Leszczyński, M. Lipka, and W. Wasilewski, Nat. Commun. 8, 2140 (2017)
[30] J. Leach, E. Bolduc, D. J. Gauthier, and R. W. Boyd, Phys. Rev. A 85, 060304 (2012)
[31] J. C. Howell, R. S. Bennink, S. J. Bentley, and R. W. Boyd, Phys. Rev. Lett. 92, 214003 (2004)
[32] M. Edgar, D. Tasca, F. Izdebski, R. Warburton, J. Leach, M. Agnew, G. Buller, R. Boyd, and M. Padgett, Nat. Commun. 3, 984 (2012)
[33] P. A. Moreau, F. Devaux, and E. Lantz, Phys. Rev. Lett. 113, 160401 (2014)
[34] Y.-W. Cho, K.-K. Park, J.-C. Lee, and Y.-H. Kim, Phys. Rev. Lett. 113, 063602 (2014)
[35] M. Kozuma, Phys. Rev. Lett. 103, 160505 (2009)
[36] Y.-W. Cho, K.-K. Park, J.-C. Lee, and Y.-H. Kim, Phys. Rev. Lett. 113, 063602 (2014)
[37] J. Leach et al., Science 329, 662 (2010)
[38] Z.-Q. Zhou et al., Nature (London) 546, 622 (2017)
[39] A. Tiranov, S. Designolle, E. Z. Cruzeiro, J. Lavoie, N. Brunner, M. Afzelius, M. Huber, and N. Gisin, Phys. Rev. A 96, 040303(R) (2017)
[40] J. Leach et al., Science 329, 662 (2010)
[41] R. Inoue, T. Yonehara, Y. Miyamoto, M. Koashi, and M. Koizumi, Phys. Rev. Lett. 103, 110503 (2009)
[42] S. Shi et al., Phys. Rev. A 96, 047905 (2018)
[43] J. Schneeloch and G. A. Howland, Phys. Rev. A 97, 042333 (2018)
[44] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nat. Phys. 6, 659 (2010)
[45] C. K. Law and J. H. Eberly, Phys. Rev. Lett. 92, 127903 (2004)
[44] R. Chrapkiewicz, W. Wasilewski, and K. Banaszek, Opt. Lett. 39, 5090 (2014).
[45] M. Jachura, R. Chrapkiewicz, R. Demkowicz-Dobrzański, W. Wasilewski, and K. Banaszek, Nat. Commun. 7, 11411 (2016).
[46] M. N. O’Sullivan-Hale, I. A. Khan, R. W. Boyd, and J. C. Howell, Phys. Rev. Lett. 94, 220501 (2005).
[47] J. Schneeloch, P. B. Dixon, G. A. Howland, C. J. Broadbent, and J. C. Howell, Phys. Rev. Lett. 110, 130407 (2013).
[48] W. Zhang et al., Nat. Commun. 7, 13514 (2016).
[49] D.-S. Ding, W. Zhang, S. Shi, Z.-Y. Zhou, Y. Li, B.-S. Shi, and G.-C. Guo, Light Sci. Appl. 5, e16157 (2016).
[50] J.-C. Lee, K.-K. Park, T.-M. Zhao, and Y.-H. Kim, Phys. Rev. Lett. 117, 250501 (2016).
[51] L. Paninski, Neural Computation 15, 1191 (2003).
[52] E. Archer, I. M. Park, and J. W. Pillow, Journal of Machine Learning Research 15, 2833 (2014).
[53] I. Nemenman, W. Bialek, and R. de Ruyter van Steveninck, Phys. Rev. E 69, 056111 (2004).
[54] M. Mazelanik, M. Dąbrowski, and W. Wasilewski, Opt. Express 24, 21995 (2016).
[55] G. Vallone, Phys. Rev. A 87, 020101(R) (2013).
[56] G. Vallone, V. D’Ambrosio, A. Sponselli, S. Slussarenko, L. Marrucci, F. Sciarrino, and P. Villoresi, Phys. Rev. Lett. 113, 060503 (2014).
[57] J. Yin et al., Phys. Rev. Lett. 119, 200501 (2017).