

Efficient High-Dimensional Entanglement Imaging with a Compressive-Sensing Double-Pixel Camera

Gregory A. Howland* and John C. Howell

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA
(Received 24 October 2012; revised manuscript received 21 December 2012; published 20 February 2013)

We implement a double-pixel compressive-sensing camera to efficiently characterize, at high resolution, the spatially entangled fields that are produced by spontaneous parametric down-conversion. This technique leverages sparsity in spatial correlations between entangled photons to improve acquisition times over raster scanning by a scaling factor up to $n^2/\log(n)$ for n -dimensional images. We image at resolutions up to 1024 dimensions per detector and demonstrate a channel capacity of 8.4 bits per photon. By comparing the entangled photons' classical mutual information in conjugate bases, we violate an entropic Einstein-Podolsky-Rosen separability criterion for all measured resolutions. More broadly, our result indicates that compressive sensing can be especially effective for higher-order measurements on correlated systems.

DOI: [10.1103/PhysRevX.3.011013](https://doi.org/10.1103/PhysRevX.3.011013)

Subject Areas: Computational Physics, Optics, Quantum Information

I. INTRODUCTION

Spatially entangled biphotons, such as those generated by spontaneous parametric down-conversion (SPDC), exhibit strong Einstein-Podolsky-Rosen (EPR) type correlations [1] in the transverse-position and transverse-momentum degrees of freedom [2]. Because these variables are continuous, the entanglement can be very high dimensional, with a typical Schmidt number greatly exceeding 1000 [3]. The many dimensions provide high information density that can be leveraged to increase channel capacity and security for quantum key distribution [4–6] and dense coding [7,8]. Other applications include ghost imaging [9,10], quantum computing [11], and quantum teleportation [12].

Experimentally characterizing the SPDC state is unfortunately difficult due to weak sources and low-resolution detectors. Spatial entanglement is traditionally imaged by jointly raster scanning photon-counting avalanche photodiodes (APDs) to measure spatial correlations. This technique scales extremely poorly with increasing detector resolution. With a biphoton flux of 4000 coincident detections per second, it would take 55 d to jointly scan a 24×24 pixel region for a signal-to-noise ratio (SNR) of 10. For 32×32 pixels, it would take 310 d [see Eq. (9)].

Other approaches have been tried with mixed success. Intensified CCD cameras can measure the Schmidt number [13] but do not detect single-photon correlations, which renders them ineffective for most quantum applications. Arrays of photon-counting APDs could replace CCDs, but they are currently low resolution, noisy, and resource intensive, especially since each pixel pair must be individually

correlated [14–16]. A recent promising result averages intensity correlations over many images from a single-photon sensitive electron-multiplying CCD, reporting 2500 modes [17]. This technique is limited to a 30-ms exposure time (APDs are sub-ns) and is noisier than using APDs because it does not isolate individual coincident detections.

In Ref. [18], Dixon *et al.* reduce the number of measurements required for a raster scan by only measuring in an area of interest where correlations are expected; they report a channel capacity of 7 bits per photon. While they do not perform a true full-field measurement, they highlight a critical feature of the SPDC field. In *both* position and momentum representations, the distribution of correlations between pairs of detector pixels is very sparse, despite dense (not sparse) single-particle distributions. Applying ideas from the field of compressive sensing, we exploit prior knowledge of this sparsity to beat the “curse of dimensionality” [19] and efficiently characterize the full biphoton field without raster scanning.

In this article, we implement a compressive-sensing photon-counting double-pixel camera that efficiently images single-photon SPDC correlations in the near and far fields at resolutions of up to $32 \times 32 = 1024$ dimensions per detector. At 32×32 resolution, the measurement time is reduced from 310 d to around 8 h for raster scanning. We perform an entropic characterization that shows channel capacities of up to 8.4 bits per photon, equivalent to 337 independent, identically distributed modes. The sums of channel capacities in conjugate bases violate an EPR steering bound [20] by up to 6.6 bits.

II. THEORY

A. Compressive sensing

Compressive sensing (CS) is a technique that employs optimization to measure a sparsely represented N -dimensional signal from $M < N$ incoherent measurements [21–24]. The approach is so named because the

*ghowland@pas.rochester.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

signal is effectively compressed during measurement. Although sparsity is assumed, it is not known prior to measurement which elements contain appreciable amplitude. Compressive sensing must determine which elements are significant, as well as find their values.

To detect a sparsely represented N -dimensional signal vector X , we measure a series of $M < N$ values Y by multiplying X by an $M \times N$ sensing matrix A such that

$$Y = AX + \Gamma, \quad (1)$$

where Γ is a noise vector.

Because $M < N$, this system is undetermined; a given Y does not specify a unique X . The correct X is recovered by minimizing a regularized least-squares objective function

$$\min_X \frac{1}{2} \|Y - AX\|_2^2 + \tau g(X), \quad (2)$$

where, for example, $\|\Omega\|_2^2$ is the ℓ_2 norm of Ω and τ is a scaling constant. The function $g(X)$ is a regularization that promotes sparsity. Common $g(X)$ include X 's ℓ_1 norm, assuming the signal is sparse, and X 's total variation, assuming the signal's gradient is sparse [25]. A must be incoherent with the basis of interest, with the surprising and nonintuitive result that a random binary sensing matrix works well. Given sufficiently large M , the recovered X approaches the exact signal with high probability [26]. For a k -sparse signal, the required M scales as $M \propto k \log(N/k)$.

Incoherent random sampling is particularly beneficial for low-light measurements, as each measurement receives, on average, half the total photon flux $\Phi/2$, as opposed to Φ/n for a raster scan. Compressive sensing is now beginning to be used for quantum applications such as state tomography [27]. Shabani *et al.*, for example, have performed a tomography on a two-qubit photonic gate for polarization-entangled photons [28]. CS has also been used with spatially correlated light for ghost imaging [29,30]. It is important to note that, for ghost imaging, CS is not required to recover the full two-particle probability distribution as in entanglement characterization.

The quintessential example of compressed sensing is the single-pixel camera [31,32]. An object is imaged onto a digital micromirror device (DMD), a 2D binary array of individually addressable mirrors that reflect light either to a single detector or a dump. Rows of the sensing matrix A consist of random binary patterns placed sequentially on the DMD. For an N -dimensional image, by minimizing Eq. (2) one recovers images while using as few as $M = 0.02N$ measurements.

B. Compressive sensing for measuring correlations

The concept of a single-pixel camera naturally adapts to imaging correlations by adding a second detector. Consider placing separate DMDs in the near field or far field of the SPDC signal and idler modes, where ‘‘on’’ pixels are redirected to photon-counting modules. The signal of interest is

$$p_x(u, v) = \int_u d\vec{x}_s \int_v d\vec{x}_i |\psi(\vec{x}_s, \vec{x}_i)|^2, \quad (3)$$

$$p_k(u, v) = \int_u d\vec{k}_s \int_v d\vec{k}_i |\tilde{\psi}(\vec{k}_s, \vec{k}_i)|^2, \quad (4)$$

where $p(u, v)$ represents the probability of a coincident detection between the u th pixel on the signal DMD and the v th pixel on the idler DMD. The functions $\psi(\vec{x}_s, \vec{x}_i)$ and $\tilde{\psi}(\vec{k}_s, \vec{k}_i)$ are approximate position and momentum wave functions for the biphoton

$$\begin{aligned} \psi(\vec{x}_s, \vec{x}_i) &= \mathcal{N} \exp\left(-\frac{(\vec{x}_s - \vec{x}_i)^2}{4\sigma_c^2}\right) \exp\left(-\frac{(\vec{x}_s + \vec{x}_i)^2}{16\sigma_p^2}\right) \tilde{\psi}(\vec{k}_s, \vec{k}_i) \\ &= (4\sigma_p\sigma_c)^2 \mathcal{N} \exp[-\sigma_c^2(\vec{k}_s - \vec{k}_i)^2] \\ &\quad \times \exp[-4\sigma_p^2(\vec{k}_s + \vec{k}_i)^2]. \end{aligned} \quad (5)$$

The subscripts s and i refer to signal and idler photons, respectively; σ_p and σ_c are the pump and correlation widths; \mathcal{N} is a normalizing constant. X of Eq. (2) is simply a one-dimensional reshaping of p_x or p_k .

Like the single-pixel camera, a series of random patterns is placed on the DMDs to form rows of A . For each pair of patterns, correlations between the signal and idler photons form the measurement vector Y . The minimization of Eq. (2) recovers $p(u, v)$.

While a fully random A is preferred, the DMDs only act on their respective signal or idler subspaces, which prevents arbitrary A . Rows of A are therefore outer products of rows of single-particle sensing matrices \mathbf{a} and \mathbf{b} ,

$$A = \begin{pmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 \\ \mathbf{a}_2 \otimes \mathbf{b}_2 \\ \vdots \\ \mathbf{a}_m \otimes \mathbf{b}_m \end{pmatrix}, \quad (6)$$

where rows of \mathbf{a} represent random patterns placed on the signal DMD, and rows of \mathbf{b} represent random patterns placed on the idler DMD. To make signal and idler photons distinguishable, \mathbf{a} and \mathbf{b} are not the same. The validity of Kronecker-type sensing matrices has been established and is of current interest in the CS community as attention shifts to higher-dimensional signals [33,34]. The measurement vector Y is obtained by counting coincident detections for the series of DMD configurations given by A .

A variety of reconstruction algorithms exists for Eq. (2), with their computational complexities dominated by repeatedly calculating AX and $A^T Y$ [35]. This calculation is especially unwieldy for correlation measurements, as the size of A is $M \times n^2$ for n -pixel DMDs. Using properties for Kronecker products [36], these can be more efficiently computed by

$$AX = \mathbf{diag}[\mathbf{b} \mathbf{sq}(X) \mathbf{a}^T], \quad (7)$$

$$A^T Y = \mathbf{vec}[\mathbf{b}^T \mathbf{od}(Y) \mathbf{a}], \quad (8)$$

where \mathbf{sq} and \mathbf{vec} reshape a vector to a square matrix and vice versa, \mathbf{diag} forms a vector from the diagonal elements

of a square matrix, and \mathbf{od} forms a square matrix with the operand vector on its diagonal.

C. Comparison to raster scanning

The compressive approach finds the joint probability distributions orders of magnitude faster than by raster scanning through two key improvements. The first improvement is simply the reduction in the number of measurements. To jointly raster scan an n -pixel space requires n^2 measurements. For a compressive measurement, sparsity is approximately n with dimensionality n^2 , so only $M \propto n \log(n)$ measurements are required. In practice, we find excellent results when M is only 3% of n^2 .

The second advantage of compressive measurements is that they more efficiently use available flux. For the raster scan, the total flux is distributed over, at best, n pairs of pixels in the case of perfect correlations. Conversely, the average flux per incoherent compressive measurement is independent of n , with each measurement receiving on average 1/4 the total flux. To maintain a constant SNR with increasing n (photons/measurement), the total measurement time scales as n^3 for raster scanning. Given a photon flux of Φ photons per second, the measurement time for a desired SNR is

$$t = n^2 t_{\text{meas}} = \frac{n^3 \text{SNR}^2}{\Phi}, \quad (9)$$

where t_{meas} is the time per measurement.

For incoherent compressive measurements, the acquisition time scales as $n \log(n)$. The compressive improvement therefore scales as $n^2 / \log(n)$. For $n = 32 \times 32 = 1024$, this improvement is of order 10^5 .

This scaling factor somewhat optimistically assumes that the reconstruction process yields an accurate result, despite a noisy signal. Unfortunately, the propagation of uncertainty through the reconstruction process remains a difficult problem, especially for nonideal real-world systems [37]. There has been much recent theoretical work on the topic for Gaussian [38–40] and Poissonian noise [41,42]. These results tend to require ideal sensing matrices or more complicated formulations to give provable performance bounds. As such, their findings are difficult to directly and quantitatively apply to an experiment. However, they do reveal pertinent features that indicate that CS can perform extremely well in the presence of noise.

A well-known characteristic of CS is a rapid phase change from poor- to good-quality reconstructions [43]. This phase change is often discussed as a function of increasing m , with the boundary $m \propto k \log(n/k)$. A similar phase transition occurs for decreasing measurement noise. Noise in our system is dominated by Poissonian shot noise, so this phase change occurs as the average number of detected photons increases. For some cases, these two phase transitions are linked [38]. A practical compressive measurement simply requires large enough m and photon flux Φ to be in the space of good reconstructions. Fortunately, simply obtaining a recognizable reconstruction generally indicates that the measurement conditions exceed this threshold.

Unlike a direct measurement, the information obtained by a series of y compressive measurements is contained in their deviation from the average value \bar{y} . In the presence of noise, these deviations must exceed the noise level. Assuming Poissonian shot noise, good reconstructions require $\text{std}(y) \geq \beta \sqrt{\bar{y}}$, where $\text{std}(y)$ is the standard deviation of the measurement vector, and β is a positive constant that is greater than 1.

The particular algorithm chosen to solve Eq. (2) also plays a role in the reconstruction’s accuracy. These algorithms often have provable performances on ideal signals but degrade when confronted with noisy or otherwise non-ideal conditions. In these circumstances, they have various strengths, including speed, accuracy, and sensitivity to user-selected parameters such as τ in Eq. (2). For more information on common reconstruction algorithms, see Refs. [24,35,44,45].

In practice, the best way to determine accuracy for a particular signal, sensing matrix, and reconstruction approach is simply repeated simulations or experiments. For our system, we reduce a $n = 32 \times 32$ measurement from a 310-d raster scan (SNR of 10) to an 8-h compressive acquisition, which is a 1000-fold improvement.

III. EXPERIMENT

The experimental apparatus is given in Fig. 1. Light from a 2.8-mW 325-nm HeCd laser is directed to a

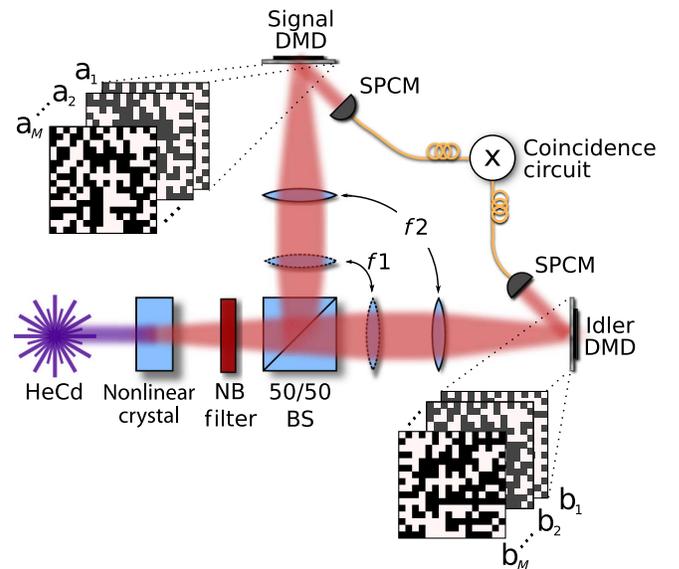


FIG. 1. Experimental setup. Photons generated via SPDC pass through a narrow-band (NB) filter and are split into signal and idler modes by a 50/50 beam splitter (BS). For position correlations, lenses $f1 = 125$ mm and $f2 = 500$ mm form a $4f$ imaging system with the crystal and DMDs placed in the object and image planes, respectively. For momentum correlations, $f1$ is removed and the DMD is placed in the focal plane of $f2 = 88.3$ mm. Photons striking DMD “on” pixels are directed to large-area SPCMs. Photon arrivals are then correlated by a coincidence circuit.

1-mm-long BiBO crystal that is oriented for type-I collinear SPDC. The generated daughter photons pass through a 650/13-nm narrow-band filter before separating into signal and idler modes at a 50/50 beam splitter. To measure position-position correlations, lenses $f_1 = 125$ mm and $f_2 = 500$ mm image the crystal onto signal- and idler-mode DMDs. For momentum-momentum correlations, f_1 is removed and the DMDs are placed in the focal plane of $f_2 = 88.3$ mm. DMD “on” pixels reflect light to large-area single-photon counting modules (SPCMs) that are connected to a correlating circuit.

To measure $p(u, v)$, a series of M random patterns are placed on the DMDs to form the sensing matrix A . For each set of patterns, joint detections are counted for acquisition times t_{aq} for a total measurement time $t = Mt_{\text{aq}}$ to make up the measurement vector y . The joint distribution $p(u, v)$ is reconstructed using a gradient projection solver for Eq. (2) with ℓ_1 regularization, which is commonly referred to as basis-pursuit denoising [44].

We measure at dimensions of $N = 256^2$, $N = 576^2$, and $N = 1024^2$ that correspond to DMD resolutions of 16×16 , 24×24 , and 32×32 pixels. The associated measurement numbers M are 2500, 10 000, and 30 000, so that M is only about $0.03N$. Acquisition times are 1 s for position

measurements and 1.5 s for momentum measurements, to average 1000 coincident detections per DMD configuration in all cases. Additionally, we perform representative simulations at 16×16 and 24×24 resolutions.

IV. RESULTS

A. Joint probability distribution

A simulation for measuring position-position correlations at 16×16 DMD resolution is given in Fig. 2. The object in Fig. 2(a) is the correlation function of Eq. (4). The simulation uses $m = 2500$ measurements and a photon flux of $\Phi = 5000$ photons/measurement that is multiplied by the ideal $p(u, v)$, which are conditions that are representative of the 1-s experimental acquisitions. Note that Φ is the total signal strength before interacting with the sensing matrix; the mean value of the measurement vector is $\Phi/4 = 1250$ detected photons. The values of the measurement vector are Poissonian distributed to simulate the effect of shot noise.

Figure 2(b) gives the reconstructed correlation function $p(u, v)$ between the signal and idler DMD pixels. The sharply defined diagonal line shows the expected positive correlations between the two DMDs. The DMD pixels are

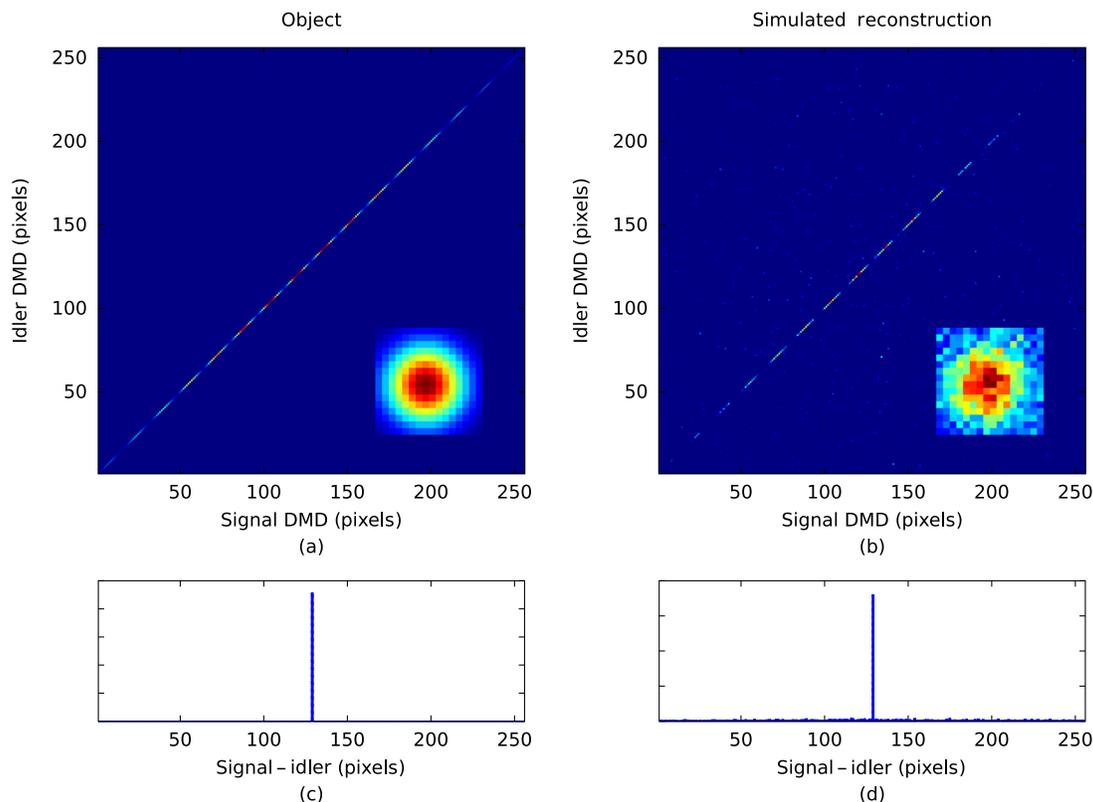


FIG. 2. 16×16 pixel simulation. The ideal object is given in (a). The object is incoherently sampled with $m = 2500$ random binary patterns. Poissonian noise corresponding to 5000 photons in the field (approximately 1250 detected) per measurement is added to the measurement vector. The reconstruction is shown in (b), with a MSE of 5×10^{-8} . The inset images in (a) and (b) show the signal photon’s 2D marginal distribution and give an image of the signal photon. (c) and (d) integrate along the antidiagonal to show that the reconstruction recovers the correlation width $\sigma_c < 1$ pixel with negligible error.

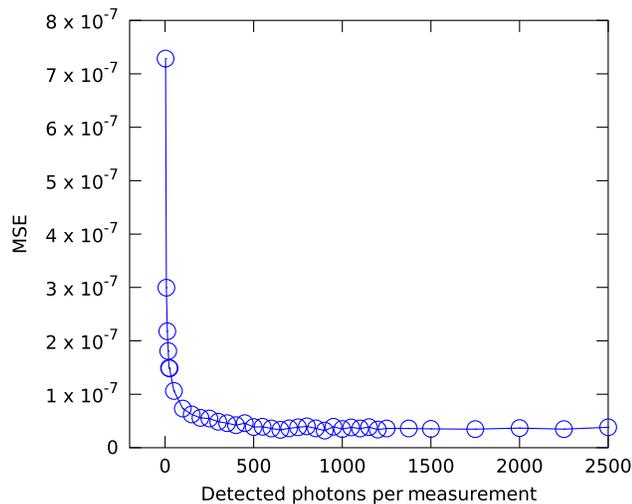


FIG. 3. Simulated MSE versus photon flux for $n = 256$ and $m = 2500$. The phase-change behavior versus photon number can be clearly seen. The experiment uses 5000 total (1250 detected) photons per measurement to comfortably exceed the phase change. The MSE approaches a value of 5×10^{-8} , which corresponds to a SNR of about 17.

listed in column-major order. The mean-squared error (MSE) for the reconstruction is 5×10^{-8} . The two-dimensional signal marginal distribution is inset, which provides an image of the signal beam. Figures 2(c) and 2(d) sum the result along the antidiagonal to show the correlation width σ_c . Qualitatively, the reconstruction closely resembles the original object, faltering only near the edges of the distribution where the signal falls beneath a noise floor. The reconstruction recovers a $\sigma_c < 1$ pixel with negligible error.

To demonstrate the reconstruction accuracy, simulations are performed for increasing photon flux Φ with DMD resolution 16×16 and $m = 2500$. The MSE versus Φ is given in Fig. 3. Reconstructions are normalized to the incident flux Φ for comparison to the ideal signal. The result shows the rapid phase change from poor to excellent reconstructions, with a MSE converging to 5×10^{-8} beyond the phase change.

The MSE can be used to roughly estimate the signal-to-noise ratio for a particular measurement of an average nonzero element. Assuming perfect pixel correlations and uniform marginal distributions, the energy in the signal is distributed over $1/n$ elements. The signal-to-noise ratio is then $1/n\sqrt{\text{MSE}}$. For $n = 256$ pixels and $\text{MSE} = 5 \times 10^{-8}$, the SNR is approximately 17. For comparison, using Eq. (9), a raster scan would require about 4 d to achieve a SNR of only 10. The simulated CS acquisition time is 42 min for 2500 1-s measurements.

Sample experimental reconstructions for position-position and momentum-momentum correlations at 16×16 pixel DMD resolution are given in Fig. 4. As in the simulations, the position-position result [Fig. 4(a)] shows a well-defined diagonal line that indicates positive

pixel correlations. Conversely, the momentum-momentum result [Fig. 4(b)] shows an antidiagonal line that shows the expected anticorrelations. Figures 4(c) and 4(d) sum the results along the antidiagonal (position-position) and the diagonal (momentum-momentum) to reveal an effective correlation width σ_{ce} of a single pixel. Our detection scheme is therefore as accurate as possible at this resolution, and our channel capacity remains detector limited.

B. Mutual information in the channel

Once $p(u, v)$ is recovered, the channel capacity is given by the classical mutual information that is shared between signal and idler DMD pixels:

$$I = -\sum_u p(u) \log p(u) - \sum_v p(v) \log p(v) + \sum_{u,v} p(u, v) \log p(u, v), \quad (10)$$

where, for example,

$$p(u) = \sum_v p(u, v) \quad (11)$$

is the signal particle's marginal probability distribution [18]. The entropic analysis is solely measurement based and does not require a wave function or density matrix reconstruction, a challenging task even for low-dimensional systems [46–48].

To estimate the uncertainty in the mutual information from shot noise and the reconstruction process, we perform 100 simulations at $n = 256$ pixel resolution and 31 simulations at $n = 576$ pixel resolution. Simulations are not performed at $n = 1024$ pixel resolution due to available computer time. In addition to the results from the raw reconstruction, thresholding is performed to provide noise reduction, where all values in the recovered $p(u, v)$ below a percentage of the maximum value are forced to zero. The simulated mutual information versus thresholding percentage is given in Fig. 5 for the $n = 256$ pixel simulations that are exemplified by Fig. 2. The error bars enclose 1 standard deviation from repeated simulations.

As the threshold increases from zero, the mutual information rises as a weak uncorrelated noise floor is removed. An optimal threshold is quickly reached, beyond which the threshold removes more signal than noise, reducing the mutual information. Note that the reconstructed mutual information is systematically lower than the actual mutual information in the ideal object. This discrepancy is due to remaining noise and difficulty in recovering parts of the signal toward the tail of the distribution.

The $n = 256$ far-field experimental result is included for comparison to the simulation. The experiment closely matches the simulation both for no thresholding and for beyond its optimal threshold but is smaller in the intermediate region. This deviation is likely due to experimental uncertainties that are not included in the simulation.

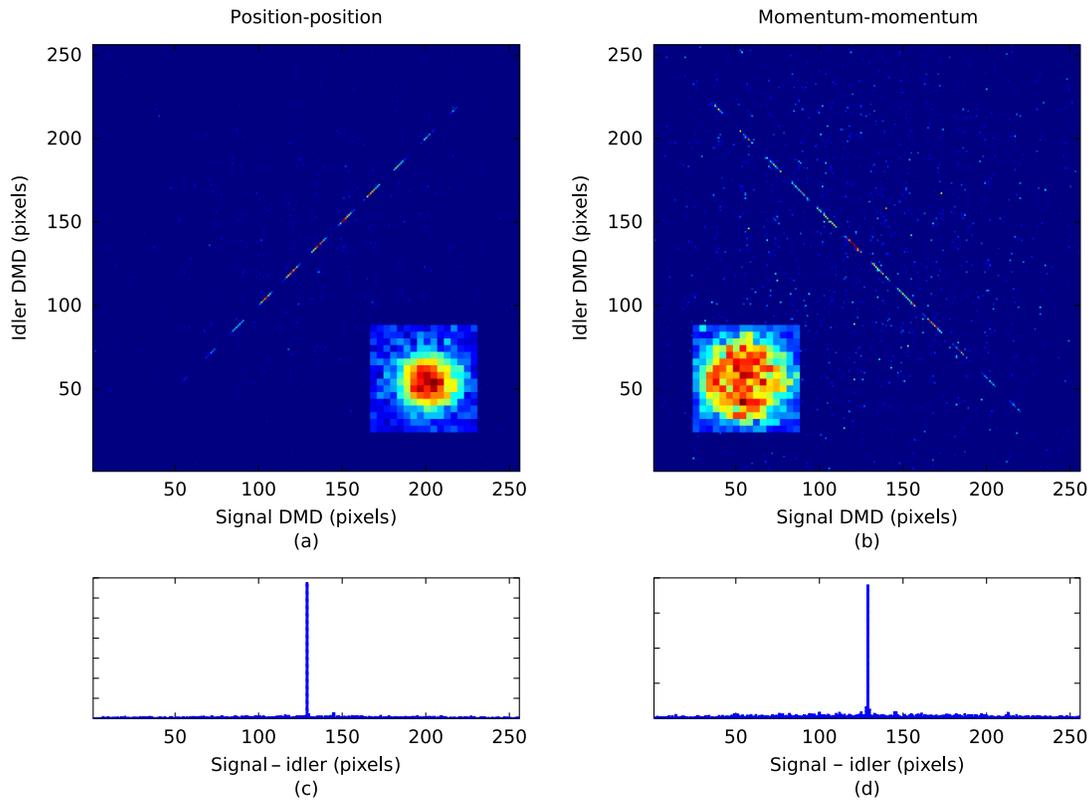


FIG. 4. Sample 16×16 experimental reconstructions. (a) and (b) give the joint probability distribution for position-position and momentum-momentum correlations where DMD pixels are listed in columnwise order. 2D marginal distributions for the signal photon are inset. (c) and (d) show correlation widths of only 1 pixel by summing over the (c) signal + idler and (d) signal - idler axes. Only 2500 (3% of raster scanning) measurements are needed, with a total acquisition time of about 40 min.

These errors include slight pixel misalignment between signal and idler DMDs, optical aberrations, detector dark noise, stray light, power fluctuations in the laser, and temperature stability of the nonlinear crystal. Figure 5 indicates that these experimental difficulties appear to increase the

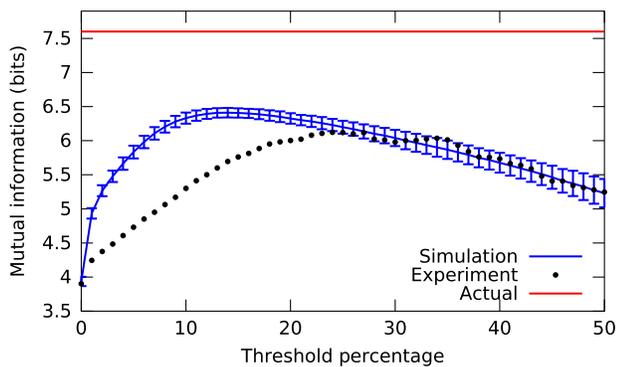


FIG. 5. Mutual information versus thresholding. The mutual information for reconstruction values above a thresholded percentage of the maximum is given for 100 $n = 256$ pixel simulations, with $m = 2500$ measurements and $\Phi = 5000$ photons per measurement. The solid red line gives the true mutual information for the simulated object. The black points give the $n = 256$ far-field experimental data for comparison.

uncorrelated noise floor rather than significantly affect the correlated part of the reconstruction.

Although thresholding is a simple postprocessing technique, it is applicable to how the entangled pixels might be used for communication. If a pair of entangled pixels has a correlated amplitude near or below the background noise, it will be difficult to use that particular mode for communication. Note that thresholding and similar noise reduction techniques cannot be used if a communication protocol encodes information on single instances of the state. However, any entanglement characterization will necessarily require many instances, so background noise can often be removed to obtain a more accurate measurement. This removal is similar to the technique in photonic quantum information of subtracting background noise from a measured signal. In CS, it is common to perform postprocessing or secondary optimization after maximizing sparsity, such as the debiasing routine in Ref. [44].

The experimental channel capacity versus DMD resolution for both position-position and momentum-momentum correlations is given in Fig. 6 for several levels of thresholding. The optimal threshold is that which maximizes the mutual information. At 256 and 576 pixel resolutions, optimal thresholds of 20% and 30% are used for position-position and momentum-momentum distributions,

respectively. At 1024 pixel resolution, noise is more significant, so the optimal thresholds increase to 30% and 40%. The error bars on $n = 256$ and $n = 576$ pixels measurements represent the simulated effect of shot noise and reconstruction uncertainty. The error bars conservatively include 2 standard deviations from the simulated result.

The joint probability distribution is also fit to the double-Gaussian wave function [Eq. (5)] to find effective widths σ_{ce} and σ_{pe} . When $\sigma_p \gg \sigma_c$, the mutual information between particles for Eq. (5) is the logarithm of the Federov ratio [49],

$$\log\left(\frac{\sigma_p^2}{\sigma_c^2}\right), \quad (12)$$

where the ratio is squared for two dimensions. While the Federov ratio technically applies to the continuous wave function, and the true σ_c is smaller than a DMD pixel, Eq. (12) still applies to the discretized measurement as long as the effective $\sigma_{ce} \gg \sigma_{pe}$.

Fitting yields the largest channel capacities, with a maximum of 8.4 bits for momentum-momentum correlations at 1024 pixel resolution, which is equivalent to 337 independent, identically distributed entangled modes.

Given that fitting more accurately characterizes the system and finds a larger mutual information, it is reasonable to question the usefulness of the direct computation of the mutual information. However, the two approaches suit different purposes. Fitting is useful if one is particularly

interested in the state itself. However, if one intends to use correlated pixels for some other purpose, such as communication, the direct calculation is more appropriate. In practice, the correlated pixels on the low-intensity tail of the distribution will be difficult to use, even if their amplitudes can be inferred by fitting. The reduced mutual information in the direct calculation reflects this difficulty.

The solid curve of Fig. 6 gives the maximum possible mutual information between two n -pixel detectors. Assuming perfect diagonal or antidiagonal correlations and uniform marginals, this maximum is simply $\log(n)$. Because we have Gaussian marginals, we do not expect to reach this bound, even with $\sigma_{ce} \approx 1$ pixel. By magnifying and using only the central part of the field, we could approach this upper limit.

C. Witnessing entanglement

Although we do not reconstruct a full density matrix, it is still possible to demonstrate nonclassical behavior by comparing position-position and momentum-momentum correlation measurements directly. This process has traditionally involved fitting the measurements to Eq. (5) and analyzing products or sums of conditional variances [50–52].

We recently presented a more inclusive entropic steering inequality for witnessing continuous variable entanglement with discrete measurements [20], where the sum of the

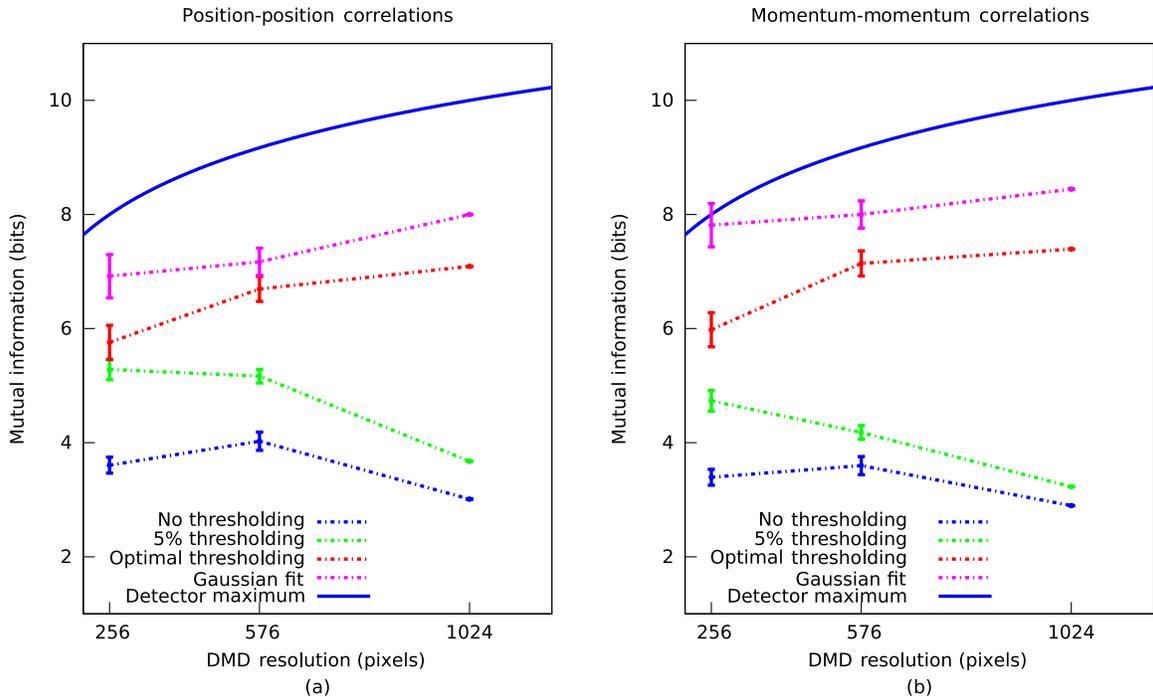


FIG. 6. Mutual information between signal and idler photons for (a) position-position and (b) momentum-momentum representations are presented as a function of detector resolution. Three levels of thresholding are shown, as well as a fit to Eq. (5). The dashed lines are guides for the eye. The error bars enclose 2 standard deviations from the expected uncertainty from simulations (not performed for $n = 1024$). The solid curve represents the maximum possible value for a particular detector resolution, given perfect correlations and uniform marginals.

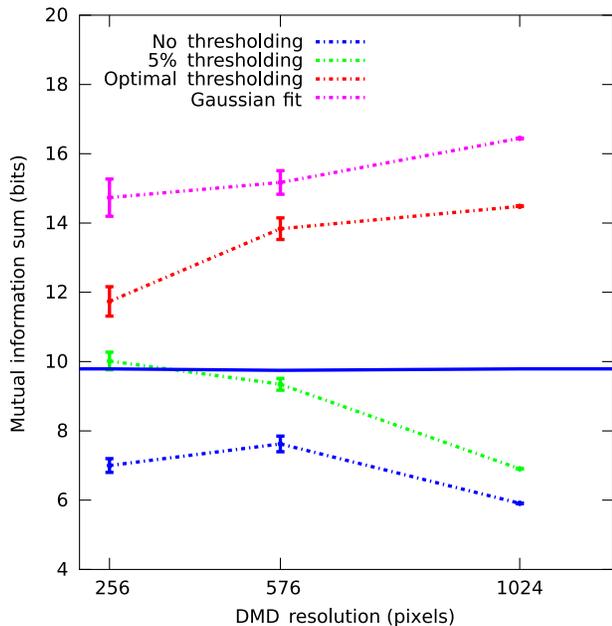


FIG. 7. The sum of position-position and momentum-momentum mutual information is presented as a function of detector resolution to demonstrate violation of an EPR steering inequality [Eq. (13)]. The solid line represents the threshold that must be *exceeded* to witness EPR steering. The error bars enclose 2 standard deviations from the expected uncertainty from simulations (not performed for $n = 1024$). Because simulations systematically underestimate the mutual information, the values above the bound are convincing.

classical mutual information between position-position and momentum-momentum correlations is classically bounded. For our system, all classically correlated measurements must satisfy

$$I_{xs,xi} + I_{ks,ki} \leq 2 \log\left(\frac{nd_x d_k}{\pi e}\right), \quad (13)$$

where d_x and d_k are the respective widths of the DMD pixels in the position and momentum bases. Note that $nd_x d_k$ is simply the bandwidth product for the DMD area and is independent of n if its total area does not change.

The sum of the classical mutual information in conjugate bases for each detector resolution is given in Fig. 7. The solid blue line provides the right-hand side of Eq. (13), which must be exceeded to witness EPR steering. The error bars for the $n = 256$ and $n = 576$ cases are derived from simulation and include 2 standard deviations. In all cases, we show EPR steering with both optimal thresholding and fitting to the double-Gaussian wave function [Eq. (5)]. Even at 5% thresholding, there is a violation for 16×16 dimensions. Recall that simulations (Fig. 3) systematically underrepresent the object's mutual information relative to measurement uncertainty, so measurement error is highly unlikely to overestimate this sum. For the fitted 32×32 dimensional result, we violate the classical bound by 6.6 bits.

V. CONCLUSION

In this article, we present a CS double-pixel camera for the characterization of the SPCD biphoton state with photon-counting detectors. This technique is very efficient and improves acquisition times over raster scanning by $n^2/\log(n)$ for n -pixel detectors. We image SPDC correlations at up to 1024 dimensions per detector and measure detector-limited mutual information of up to 8.4 bits. We also violate an entropic EPR steering bound, which indicates that these correlations are nonclassical. More broadly, our results suggest that compressive sensing can be extremely effective for analyzing correlations within large-dimensional signals (e.g., intensity-intensity correlations). Potential applications range from verifying security in spectral correlations for energy-time quantum key distribution [53] to imaging through scattering media [54].

ACKNOWLEDGMENTS

This work was supported by DARPA InPho Grant No. W911NF-10-1-0404.

- [1] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, *Phys. Rev.* **47**, 777 (1935).
- [2] J. C. Howell, R. S. Bennink, S. J. Bentley, and R. W. Boyd, *Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum- and Position-Entangled Photons from Spontaneous Parametric Down Conversion*, *Phys. Rev. Lett.* **92**, 210403 (2004).
- [3] J. B. Pors, Ph.D. thesis, Leiden University, 2011.
- [4] A. K. Ekert, *Quantum Cryptography Based on Bell's Theorem*, *Phys. Rev. Lett.* **67**, 661 (1991).
- [5] S. P. Walborn, D. S. Lemelle, M. P. Almeida, and P. H. Souto Ribeiro, *Quantum Key Distribution with Higher-Order Alphabets Using Spatially Encoded Qudits*, *Phys. Rev. Lett.* **96**, 090501 (2006).
- [6] S. P. Walborn, D. S. Lemelle, D. S. Tasca, and P. H. Souto Ribeiro, *Schemes for Quantum Key Distribution with Higher-Order Alphabets Using Single-Photon Fractional Fourier Optics*, *Phys. Rev. A* **77**, 062323 (2008).
- [7] C. H. Bennett and S. J. Wiesner, *Communication via One- and Two-Particle Operators on Einstein-Podolsky-Rosen States*, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [8] S. L. Braunstein and H. J. Kimble, *Dense Coding for Continuous Variables*, *Phys. Rev. A* **61**, 042302 (2000).
- [9] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, *Optical Imaging by Means of Two-Photon Quantum Entanglement*, *Phys. Rev. A* **52**, R3429 (1995).
- [10] A. F. Abouraddy, B. E. A. Saleh, A. V. Sergienko, and M. C. Teich, *Role of Entanglement in Two-Photon Imaging*, *Phys. Rev. Lett.* **87**, 123602 (2001).
- [11] D. S. Tasca, R. M. Gomes, F. Toscano, P. H. Souto Ribeiro, and S. P. Walborn, *Continuous-Variable Quantum Computation with Spatial Degrees of Freedom of Photons*, *Phys. Rev. A* **83**, 052325 (2011).

- [12] S. P. Walborn, D. S. Ether, R. L. de Matos Filho, and N. Zagury, *Quantum Teleportation of the Angular Spectrum of a Single-Photon Field*, *Phys. Rev. A* **76**, 033801 (2007).
- [13] H. Di Lorenzo Pires, C. H. Monken, and M. P. van Exter, *Direct Measurement of Transverse-Mode Entanglement in Two-Photon States*, *Phys. Rev. A* **80**, 022307 (2009).
- [14] V. C. Coffey, *Seeing in the Dark: Defense Applications of IR Imaging*, *Opt. Photonics News* **22**, 26 (2011).
- [15] M. A. Albota et al., *Three-Dimensional Imaging Laser Radars with Geiger-Mode Avalanche Photodiode Arrays*, *Lincoln Laboratory Journal* **13**, 351 (2002).
- [16] M. A. Itzler, M. Entwistle, M. Owens, K. Patel, X. Jiang, K. Slomkowski, S. Rangwala, P. F. Zalud, T. Senko, J. Tower, and J. Ferraro, in *Infrared Remote Sensing and Instrumentation XVIII*, SPIE Proceedings Vol. 7808 (SPIE-International Society for Optical Engineering, Bellingham, WA, 2010), p. 78080C.
- [17] M. P. Edgar, D. S. Tasca, F. Izdebski, R. E. Warburton, J. Leach, M. Agnew, G. S. Buller, R. W. Boyd, and M. J. Padgett, *Imaging High-Dimensional Spatial Entanglement with a Camera*, *Nat. Commun.* **3**, 984 (2012).
- [18] P. B. Dixon, G. A. Howland, J. Schneeloch, and J. C. Howell, *Quantum Mutual Information Capacity for High-Dimensional Entangled States*, *Phys. Rev. Lett.* **108**, 143603 (2012).
- [19] R. Bellman, *Dynamic Programming* (Dover, New York, 2003).
- [20] J. Schneeloch, P. B. Dixon, G. A. Howland, C. J. Broadbent, and J. C. Howell, *Witnessing Continuous Variable Entanglement with Discrete Measurements*, [arXiv:1210.4234](https://arxiv.org/abs/1210.4234).
- [21] D. L. Donoho, *Compressed Sensing*, *IEEE Trans. Inf. Theory* **52**, 1289 (2006).
- [22] E. J. Candès and J. Romberg, *Sparsity and Incoherence in Compressive Sampling*, *Inverse Probl.* **23**, 969 (2007).
- [23] R. G. Baraniuk, *Compressive Sensing [Lecture Notes]*, *IEEE Signal Process. Mag.* **24**, 118 (2007).
- [24] E. J. Candès and M. B. Wakin, *An Introduction to Compressive Sampling*, *IEEE Signal Process. Mag.* **25**, 21 (2008).
- [25] E. J. Candès and J. Romberg, *11-Magic: Recovery of Sparse Signals via Convex Programming* (unpublished).
- [26] E. J. Candès, J. Romberg, and T. Tao, *Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information*, *IEEE Trans. Inf. Theory* **52**, 489 (2006).
- [27] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Quantum State Tomography via Compressed Sensing*, *Phys. Rev. Lett.* **105**, 150401 (2010).
- [28] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, *Efficient Measurement of Quantum Dynamics via Compressive Sensing*, *Phys. Rev. Lett.* **106**, 100401 (2011).
- [29] O. Katz, Y. Bromberg, and Y. Silberberg, *Compressive Ghost Imaging*, *Appl. Phys. Lett.* **95**, 131110 (2009).
- [30] P. Zerom, K. W. C. Chan, J. C. Howell, and R. W. Boyd, *Entangled-Photon Compressive Ghost Imaging*, *Phys. Rev. A* **84**, 061804(R) (2011).
- [31] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, in *IEEE International Conference on Image Processing, 2006* (IEEE, New York, 2006), p. 1273–1276.
- [32] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, *Single-Pixel Imaging via Compressive Sampling*, *IEEE Signal Process. Mag.* **25**, 83 (2008).
- [33] M. F. Duarte and R. G. Baraniuk, in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2010* (IEEE, New York, 2010), p. 3650–3653.
- [34] M. F. Duarte and R. G. Baraniuk, *Kronecker Compressive Sensing*, *IEEE Trans. Image Process.* **21**, 494 (2012).
- [35] C. Li, W. Yin, and Y. Zhang, *TVAL3: TV Minimization by Augmented Lagrangian and ALternating Direction ALgorithms* (unpublished).
- [36] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, England, 1994).
- [37] Rebecca M. Willett, Roummel F. Marcia, and Jonathan M. Nichols, *Compressed Sensing for Practical Optical Imaging Systems: A Tutorial*, *Opt. Eng. (Bellingham, Wash.)* **50**, 072601 (2011).
- [38] D. L. Donoho, A. Maleki, and A. Montanari, *The Noise-Sensitivity Phase Transition in Compressed Sensing*, *IEEE Trans. Inf. Theory* **57**, 6920 (2011).
- [39] Yihong Wu and S. Verdu, *Optimal Phase Transitions in Compressed Sensing*, *IEEE Trans. Inf. Theory* **58**, 6241 (2012).
- [40] G. Reeves and M. Gastpar, in *46th Annual Conference on Information Sciences and Systems (CISS), 2012* (Curran, Red Hook, NY, 2012), p. 1–6.
- [41] R. M. Willett and M. Raginsky, in *IEEE International Symposium on Information Theory (ISIT), 2009* (IEEE, New York, 2009), p. 174–178.
- [42] Z. T. Harmany, R. F. Marcia, and R. M. Willett, in *IEEE/SP 15th Workshop on Statistical Signal Processing (SSP), 2009* (IEEE, New York, 2009), p. 634–637.
- [43] Surya Ganguli and Haim Sompolinsky, *Statistical Mechanics of Compressed Sensing*, *Phys. Rev. Lett.* **104**, 188701 (2010).
- [44] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, *Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems*, *IEEE J. Sel. Top. Signal Process.* **1**, 586 (2007).
- [45] P. R. Gill, A. Wang, and A. Molnar, *The In-Crowd Algorithm for Fast Basis Pursuit Denoising*, *IEEE Trans. Signal Process.* **59**, 4595 (2011).
- [46] S. M. Barnett and S. J. D. Phoenix, *Entropy as a Measure of Quantum Optical Correlation*, *Phys. Rev. A* **40**, 2404 (1989).
- [47] S. P. Walborn, B. G. Taketani, A. Salles, F. Toscano, and R. L. de Matos Filho, *Entropic Entanglement Criteria for Continuous Variables*, *Phys. Rev. Lett.* **103**, 160505 (2009).
- [48] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro, *Revealing Hidden Einstein-Podolsky-Rosen Nonlocality*, *Phys. Rev. Lett.* **106**, 130402 (2011).
- [49] M. V. Fedorov, Yu. M. Mikhailova, and P. A. Volkov, *Gaussian Modelling and Schmidt Modes of SPDC Biphoton States*, *J. Phys. B* **42**, 175503 (2009).
- [50] M. D. Reid, *Demonstration of the Einstein-Podolsky-Rosen Paradox Using Nondegenerate Parametric Amplification*, *Phys. Rev. A* **40**, 913 (1989).

-
- [51] Lu-Ming Duan, G. Giedke, J.I. Cirac, and P. Zoller, *Inseparability Criterion for Continuous Variable Systems*, *Phys. Rev. Lett.* **84**, 2722 (2000).
- [52] R. Simon, *Peres-Horodecki Separability Criterion for Continuous Variable Systems*, *Phys. Rev. Lett.* **84**, 2726 (2000).
- [53] I. Ali-Khan, C.J. Broadbent, and J.C. Howell, *Large-Alphabet Quantum Key Distribution Using Energy-Time Entangled Bipartite States*, *Phys. Rev. Lett.* **98**, 060503 (2007).
- [54] W. Gong and S. Han, *Correlated Imaging in Scattering Media*, *Opt. Lett.* **36**, 394 (2011).