Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases

Mhlambululi Mafu,[1] Angela Dudley,[2] Sandeep Goyal,[1] Daniel Giovannini,[3] Melanie McLaren,[2] Miles J. Padgett,[3]
Thomas Konrad,[1,4] Francesco Petruccione,[1] Norbert Lütkenhaus,[5] and Andrew Forbes[2,*]
[1]*School of Chemistry and Physics, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa*
[2]*CSIR National Laser Centre, P.O. Box 395, Pretoria 0001, South Africa*
[3]*School of Physics and Astronomy, SUPA, University of Glasgow, Glasgow, United Kingdom*
[4]*National Institute for Theoretical Physics (NITheP), University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa*
[5]*Institute for Quantum Computing & Department for Physics and Astronomy, University of Waterloo, 200 University Avenue West, N2L 3G1, Waterloo, Ontario, Canada*

(Received 26 June 2013; published 5 September 2013)

We present an experimental study of higher-dimensional quantum key distribution protocols based on mutually unbiased bases, implemented by means of photons carrying orbital angular momentum. We perform $(d+1)$ mutually unbiased measurements in a classically simulated prepare-and-measure scheme and on a pair of entangled photons for dimensions ranging from $d=2$ to 5. In our analysis, we pay attention to the detection efficiency and photon pair creation probability. As security measures, we determine from experimental data the average error rate, the mutual information shared between the sender and receiver, and the secret key generation rate per photon. We demonstrate that increasing the dimension leads to an increased information capacity as well as higher key generation rates per photon. However, we find that the benefit of increasing the dimension is limited by practical implementation considerations, which in our case results in deleterious effects observed beyond a dimension of $d=4$.

DOI: 10.1103/PhysRevA.88.032305 PACS number(s): 03.67.Dd, 03.67.Hk, 42.50.Ex

I. INTRODUCTION

Quantum key distribution (QKD) establishes a secure key between two parties, Alice and Bob, in which they can encode a secret message [1–3]. Protocols for QKD are classified as either prepare-and-measure (P&M) schemes or entanglement-based (EB) schemes. Examples of P&M schemes are BB84 [3], B92 [4], six-state [5], and SARG04 [6]. In general, P&M schemes can be translated into EB schemes, as discussed in [7] where the connection between the BB84 protocol and an EB protocol similar to E91 is made.

Mutually unbiased bases (MUBs) [8–10] have found many applications, for example in quantum state tomography [10–13] and quantum error correction codes [14,15], and are convenient in drawing up efficient QKD protocols. This is because projective measurements in one basis provides no knowledge of the state in any of the other bases [3,16,17]. Therefore if an eavesdropper measures in the incorrect basis, he or she will obtain no meaningful information but instead introduce a disturbance in the system, resulting in its detection. The simplest example of MUBs of dimension $d=2$ are the horizontal and vertical, diagonal and antidiagonal, and left- and right-handed polarization bases, as they are unbiased with respect to each other, forming a set of three MUBs. Although MUBs offer security against eavesdropping, encoding states in the polarization degree of freedom only allows a maximum of one bit of information transmitted per photon, which results in a limited key generation rate. Since systems with a higher-dimensional Hilbert space can store more information per carrier, the question arises as to whether QKD protocols using higher-dimensional MUBs also result in higher generation rates of secure key bits; indeed, such protocols can be expected to be more robust in terms of abstract noise measures [18,19]. Their actual performance in terms of secure key rate, however, depends on whether the amount of noise in higher-dimensional implementations grows faster with increasing dimension than their robustness against noise. The present article addresses this question for implementations using the orbital angular momentum (OAM) of photons. Beams that carry OAM have an azimuthal angular dependence of $\exp(i\ell\theta)$ [20], where ℓ is the azimuthal index and θ is the azimuthal angle. It has been shown theoretically that MUBs for higher-dimensional OAM states can be used to encode bits of information in alignment with the BB84 protocol [21–24]. A standard P&M implementation of a generalized BB84 protocol, relying on 11 OAM states and superpositions of these 11 OAM states, has previously been performed [25] using two of the 12 available MUBs. In this paper we experimentally investigate an entanglement-based scheme for QKD encoded in complete sets of higher-dimensional MUBs, which we first verify with a classically simulated P&M scheme. We implement our protocol with MUBs encoded in OAM states and present values for the corresponding average error rates, classical Shannon information, and secret key rates. As with all OAM protocols, our QKD protocol uses filter measurements that project onto one MUB element at a time; we provide the connection between these protocols to the established theory for protocols using full MUB measurements. To achieve this, we prove that the detection efficiency depends only on the basis choice and not on the elements within a basis; otherwise the security parameters of the protocol cannot be evaluated. This allows us to map our protocol to the key rates, thus arriving at the standard MUB protocol. By increasing the dimension d, we obtain an increase in the secret key rate which has been theoretically observed in recent papers [18,19], resulting in higher key generation rates for dimension $d=4$. Similarly, the Shannon mutual information increases, demonstrating an improvement in the information capacity.

[*] Corresponding author: aforbes1@csir.co.za

II. CHOICE OF MUTUALLY UNBIASED BASES

Two orthonormal bases, $\mathcal{M}_1 = \{|\phi_{(1,i)}\rangle, i = 0,1,\ldots,d-1\}$ and $\mathcal{M}_2 = \{|\phi_{(2,j)}\rangle, j = 0,1,\ldots,d-1\}$, of a d-dimensional Hilbert space $\mathcal{H}_d$ are said to be mutually unbiased if, and only if, all pairs of basis vectors $|\phi_{(1,i)}\rangle$ and $|\phi_{(2,j)}\rangle$ satisfy

$$|\langle\phi_{(1,i)}|\phi_{(2,j)}\rangle|^2 = \frac{1}{d}. \tag{1}$$

Physically, this means that for a system prepared in the basis $\mathcal{M}_1$ and measured with respect to basis $\mathcal{M}_2$, all outcomes are equally probable. This property of mutually unbiased bases makes them important for QKD protocols. Mutually unbiased bases were introduced by Schwinger [8] in 1960 as optimum incompatible measurement bases. In 1981, Ivonovic showed their application in quantum state discrimination [9]. Later Wootters and Fields [10] gave a constructive proof that there exist complete sets of MUBs for prime power dimensions and proved that for any dimension d there are not more than $d+1$ MUBs within any particular set of MUBs. The smallest prime dimension is 2, and for that an example of a complete set of MUBs consists of the eigenstates of the three Pauli spin operators $\sigma_z,\sigma_x,\sigma_y$, i.e.,

$$\{|0\rangle,|1\rangle\}; \tag{2}$$

$$\left\{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle),\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)\right\}; \tag{3}$$

$$\left\{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle),\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)\right\}. \tag{4}$$

Pauli operators can be generalized to higher dimension, known as the Weyl operators. These are unitary operators of the form X^kZ^l for $k,l \in \{0,1,\ldots,d-1\}$. The operator Z is diagonal in the standard basis $\{|0\rangle,|1\rangle,\ldots,|d-1\rangle\}$,

$$Z = \sum_{i=0}^{d-1}\omega^i|i\rangle\langle i|, \tag{5}$$

with $\omega = \exp(i2\pi/d)$, whereas the operator X reads

$$X = \sum_{i=0}^{d-1}|i+1\mathrm{mod}d\rangle\langle i|. \tag{6}$$

The eigenbases belonging to the different operators in the set $\{Z,XZ^l|l \in \{0,1,\ldots,d-1\}\}$ form a complete set of MUBs for any prime number d as the dimension of the underlying Hilbert space. For $d=2$, the operator X is identical with the Pauli operator σ_x, and the operator Z is given by the Pauli operator σ_z. In the present study of MUB-based QKD, a complete set of MUBs is implemented following the recipe above by means of photons carrying OAM. The MUBs are obtained by assuming that the standard basis (eigenbasis of the operator Z) is realized by single-photon states which correspond to an elementary excitation of Laguerre-Gaussian modes (LG$_\ell$) carrying an OAM value $l\hbar$. For $d=2$ we employ the LG$_\ell$ modes with $\ell = \pm 1$ to generate the standard basis. For $d=3$, our choice of the standard basis corresponds to LG$_\ell$ modes with OAM values $\ell = -1, 0, 1$:

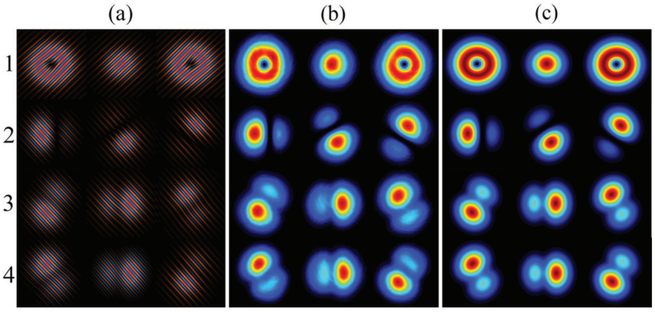

FIG. 1. (Color online) The states for each of the four MUBs for $d=3$. (a) Images representing the measurement filters (or holograms) for each of the 12 states. (b) Experimentally produced and (c) theoretically calculated intensity profiles of the LG$_\ell$ modes produced by each hologram. The first row (1) represents the well-known LG basis, sometimes called the OAM basis, as given by Eq. (7).

$$\left\{|-1\rangle \equiv \begin{pmatrix}1\\0\\0\end{pmatrix},|0\rangle \equiv \begin{pmatrix}0\\1\\0\end{pmatrix},|1\rangle \equiv \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}. \tag{7}$$

The remaining three bases are given in matrix notation with respect to the standard basis as

$$\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&\omega&\omega^2\\1&\omega^2&\omega\end{pmatrix}, \quad \frac{1}{\sqrt{3}}\begin{pmatrix}1&1&\omega\\1&\omega&1\\\omega&1&1\end{pmatrix},$$
$$\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&\omega^2\\1&\omega^2&1\\\omega^2&1&1\end{pmatrix}. \tag{8}$$

Here each matrix represents a complete orthonormal basis with its columns reflecting the basis vectors. In general, for prime dimension d, the standard basis consists of d LG$_\ell$ modes, while the remaining d bases pertain to superpositions of the LG$_\ell$ modes. Examples of the LG$_\ell$ modes and their superpositions are given in Fig. 1, which contains images of the measurement holograms and their corresponding intensity profiles. The first row represents the standard basis as given by Eq. (7), while rows 2–4 show the remaining three bases (for $d=3$), as given by Eq. (8).

III. FILTER-BASED MUB QKD PROTOCOL

We now describe our filter-based QKD protocol. In both scenarios, Alice (SLM A) prepares her mode in a state chosen randomly from one of the $(d+1)$ bases, while Bob (SLM B) performs a measurement on his mode by randomly selecting a state in one of the $(d+1)$ bases chosen out of $d(d+1)$ different basis settings but biased towards one basis. Each party then announces from which basis the filter measurement was chosen (sifting) and keeps measurements of whether they all arrived in the same basis. They later make announcements as to whether photon coincidences occurred (postselection). A coincidence event represents a conclusive result; otherwise it becomes inconclusive. This is followed by parameter estimation (error rate in the remaining data), error correction, and privacy amplification. The announcement step allows our filter-measurement-based QKD protocol to be mapped back to the original protocol, which uses full MUB measurements.

IV. AVERAGE ERROR RATE AND SECRET KEY RATE

In standard EB QKD protocols both parties perform measurements on the states that they receive, followed by a public announcement of their measurement basis. The two parties then compare a small portion of their measurements in order to obtain an estimate of the average error rate. This quantifies the error in the QKD protocol resulting from all sources of noise, such as noise in the transmission channel and errors in the measurements. Moreover, the noise could also be caused by an eavesdropper. The error rate refers to the probability that Alice sends the state $|\phi_{(\beta,k)}\rangle$, while Bob receives an orthogonal state $|\phi_{(\beta,k')}\rangle$. Given the MUB β, the corresponding average error rate in each basis Q^β is expressed as

$$Q^\beta = \sum_{\substack{k,k\\ k'\neq k}} \mathrm{tr}[|\phi^*_{(\beta,k)}\rangle\langle\phi^*_{(\beta,k)}| \otimes |\phi_{(\beta,k')}\rangle\langle\phi_{(\beta,k')}|\rho_{AB}], \tag{9}$$

where ρ_{AB} is the density matrix of the two-photon (joint) state. The total average error rate is the total error obtained as an average over the different MUBs, $\mathcal{L}$ [19], and is defined as

$$Q = \frac{1}{\mathcal{L}} \sum_{\beta\in\mathcal{L}} Q^\beta. \tag{10}$$

We use the full set of available MUBs; therefore $\mathcal{L} = d + 1$. Another important figure of merit for the performance of a QKD scheme is the secret key rate. It is given by the amount of information that one can send securely in a photonic QKD scheme. It equals the number of key bits per photon measured by both parties in the same basis that can be generated securely. The maximum secret key rate that one can achieve is $\log_2 d$ for a d-level system but is limited by an adversarial attack by Eve, which results in an observed error that requires Alice and Bob to perform error correction and privacy amplification. Both processes affect the secret key rate. The resulting key rate is given as [18,19]

$$\begin{aligned} r_{\min} = \log_2 d + \frac{d+1}{d} Q \log_2\left(\frac{Q}{d(d-1)}\right) \\ + \left(1 - \frac{d+1}{d}Q\right)\log_2\left(1 - \frac{d+1}{d}Q\right), \end{aligned} \tag{11}$$

where Q is the average error rate from Eq. (10). The secret key rate is given as the difference between the classical mutual information shared by Alice and Bob and the information shared by Alice and Eve as measured by the quantum mutual information. The quantum mutual information is also referred to as the Holevo quantity [2]. The Holevo quantity measures the information that one has on Alice's data as a result of Eve's interaction with the signals as they pass to Bob. The secret key rate can be written as

$$r = I(A{:}B) - \chi(X{:}E), \tag{12}$$

where $I(A{:}B)$ is the classical mutual information and $\chi(X{:}E) = H(X) - S(E) - S(X,E)$ is the quantum mutual information or Holevo quantity, where H and S denote the Shannon entropy and von Neumann entropy, respectively. The limit on the tolerable error rate that is safe for secret key generation can be improved by implementing a full set of $(d+1)$ MUBs [19,24]. Using a full set of MUBs results in an increase in the tolerable error rate in which we can still extract a reasonable secret key without compromising the security of the protocol. However, this happens at the cost of reducing the transmission rate, which is proportional to the probability $1/(d+1)$ that Alice and Bob choose the same basis. But in our protocol, this is not a problem since we make use of the asymmetric [26] basis choice, so one does not pay the high cost of sifting with MUBs. In order to calculate the maximum tolerable error rate $Q_{\max}$, the secret key rate $r_{\min}$ is set to zero.

V. EXPERIMENTAL SETUP

Our EB QKD protocol was implemented at the single-photon level on entangled photon pairs depicted in Fig. 2. The laser source was a mode-locked UV laser (Vanguard 355-2500), producing pulses of approximately 10 ps at a repetition rate of 80 MHz. The 355-nm wavelength beam was collimated and directed to pump a 3-mm-thick type-I beta barium borate (BBO) crystal, producing collinear frequency-degenerate entangled photon pairs at 710 nm. A beam splitter was used to separate the collinear signal and idler photons (depicted by arms A and B), which were directed and imaged (2×) from the plane of the crystal onto spatial light modulators (SLMs) by a 4-f telescope. The SLMs were used to execute the filter measurements and were encoded to manipulate both the phase and amplitude of the incident light [27–30], allowing only one particular superposition of the LG$_\ell$ modes to be detected by the detector, while all the others were blocked. False color images of the types of filters (or holograms) encoded on the SLMs are presented in Fig. 1. The projected mode obtained at the plane of the SLM, be it either Gaussian or non-Gaussian, depending on whether the filter either does or does not match the state of the incident photon, was imaged (0.004×) by a 4-f telescope onto a single-mode fiber. The fibers were connected to avalanche photodiodes (APDs) which

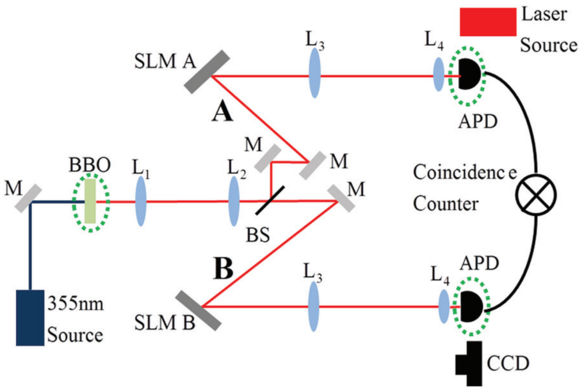

FIG. 2. (Color online) The experimental setup used to perform both the EB and P&M QKD protocols. The plane of the crystal was relayed imaged onto SLMs A and B with the use of lenses, L$_1$ and L$_2$ (f$_1$ = 200 mm and f$_2$ = 400 mm). Lenses L$_3$ and L$_4$ (f$_3$ = 500 mm and f$_4$ = 2 mm) were used to relay image the SLM planes to single-mode fibers.

detected the photon pairs via a coincidence counter with a gating time of 12 ns. The single count rates, S_A and S_B, and the coincidence count rates, C, were recorded simultaneously and accumulated over an integration time of 10 s.

An initial step in conducting our EB QKD protocol was to test it in a classically simulated P&M experiment. This served to confirm the correct filter measurements at more manageable photon count rates. Our experimental setup for the classically simulated P&M scheme can be illustrated with the use of Fig. 2, where the BBO crystal is considered to be reflective and the APD in arm A is replaced with a diode laser source at a wavelength equal to the downconverted photons (710 nm), and the APD in arm B with a CCD camera. This procedure is commonly referred to as backprojection or retrodiction [31] and has shown promise in classically simulating the downconverted photon experiment [32,33]. Conducting the experiment in this manner provided a quicker and simpler method for the verification of the experimental procedure.

VI. RESULTS AND DISCUSSION

By way of example, we consider $d = 3$ in our classically simulated P&M-based experiment. We scanned through all possible states, defined by Eqs. (7) and (8) and depicted in Fig. 1, on SLM A and SLM B. Figure 3(a) contains the cross-sectional intensity profiles recorded on the CCD (depicted in Fig. 2) when SLM A and SLM B scanned through the states pertaining to the first basis. It is evident that when SLM A and SLM B select the same (different) states, a Gaussian mode (singularity) appears on axis. The normalized on-axis intensities are depicted in Fig. 3(b) for the permutation of all the bases elements for $d = 3$. We note that the diagonal elements are approximately equal to $1/3$ $(1/d)$ and the elements corresponding to different bases are found to be approximately $1/9$ $(1/d^2)$. This validates the implementation of the filters (holograms) and their normalization. We note that there is some variance across the bases sets. This is likely due to the complex amplitude modulation implemented on a finite-resolution phase-only spatial light modulator. Our approach in obtaining the normalized joint probabilities is outlined in the Appendix together with a discussion on the detection efficiencies.

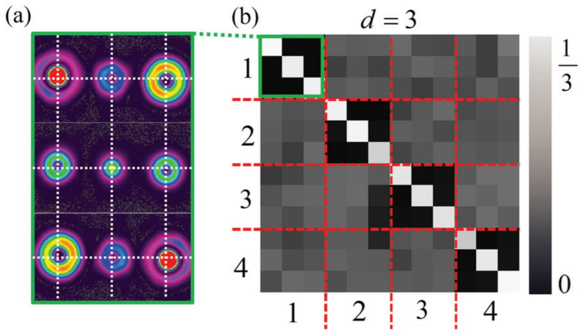

FIG. 3. (Color online) (a) Cross-sectional intensity profiles of the field recorded on the CCD for permutations of the first basis's states encoded on SLM A and SLM B. White crosshairs mark the axis of propagation. (b) The normalized intensity recorded at the CCD when SLM A (Alice) and SLM B (Bob) select one of the three states from one of the four bases.

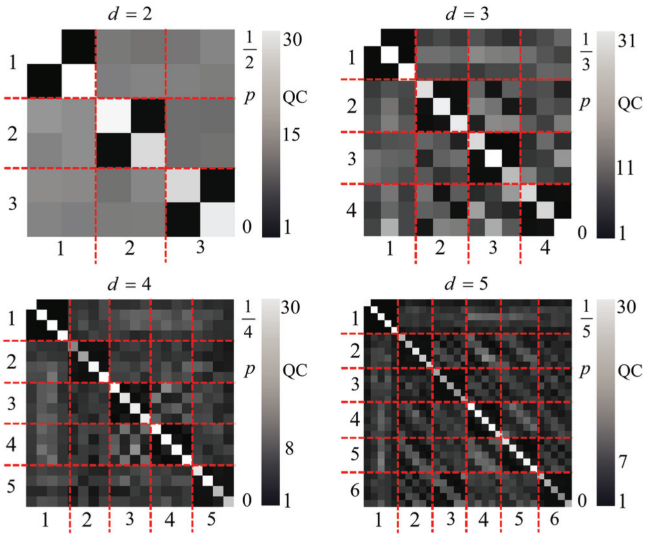

FIG. 4. (Color online) The normalized joint probabilities when SLM A (Alice) and SLM B (Bob) select one of the d states from one of the $d + 1$ bases for the EB scheme.

Following the successful implementation of the simulated P&M scheme, we proceeded to the EB scheme. For each permutation of the projective measurements by Alice and Bob in the EB scheme, the single count rates and coincidence count rates were recorded and the normalized joint probabilities calculated for $d = 2,3,4$, and 5 given in Fig. 4, together with the quantum contrast for each measurement. Coincidences peaked at approximately 300 counts per second for each dimension, roughly half the value measured when no hologram is programmed on the SLM ($\ell = 0$ mode on both SLMs.) The maximum efficiency of detection is limited by the APDs and SLMs at approximately 10%, while our measured detection efficiency with the intensity masking is only 1%. In studying the data in Fig. 4, it is evident that when the filter settings are the same, anticorrelations in all the bases are observed (denoted by the white diagonal elements). In performing the projective measurements, completely orthogonal filter settings result in no correlations (an inconclusive measurement), while

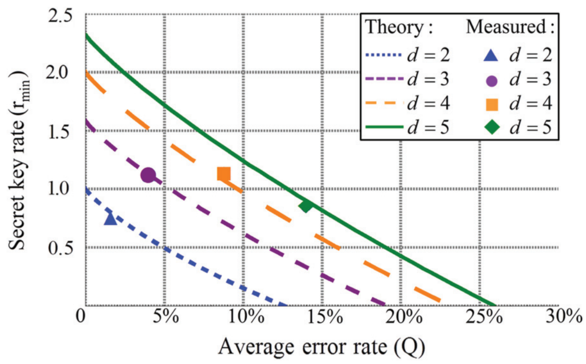

FIG. 5. (Color online) The secret key rate $r_{\min}$ as a function of the average error rate Q for different dimensions. The solid data points denote the measured values and the dashed curves the theoretical values calculated from Eq. (11).

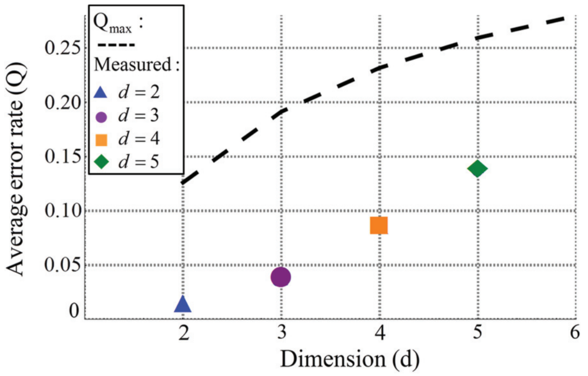

FIG. 6. (Color online) The measured average error rate (Q) and the maximum permissible error rate ($Q_{\max}$) evaluated when $r_{\min} = 0$.

the overlap between the remaining filter settings is given as the inverse of the dimension (i.e., $1/d$).

Based on the results from the normalized joint probabilities, we calculated the average error rate Q according to Eq. (10). We find that for $d = 2$, 3, 4, and 5, the average error rate is $Q = 0.016$, 0.040, 0.088, and 0.14, respectively. By using these values of Q together with Eq. (11) we calculate the secret key rate to be $r_{\min} = 0.7590$, 1.123, 1.139, and 0.8606 for $d = 2,3,4,$ and 5, respectively. Figure 5 contains the measured secret key rates plotted as a function of the measured average error rates for dimensions $d = 2$, 3, 4, and 5, denoted by the data points. The curves denote the theoretical secret key rate as a function of the average error rate, plotted with the use of Eq. (11). For each dimension d, the intersection between the dashed curves and the horizontal axis (i.e., where $r_{\min} = 0$) corresponds to the maximum permissible error rate ($Q_{\max}$) in order to enable the secure distribution of a secret key. Ideally, we want to minimize the error rate Q in order to maximize the secret key rate $r_{\min}$. These results are shown in a different format in Fig. 6, where it is now evident that all the measured error rates are well below the maximum permissible error rate.

The Shannon information for $d = 2,3,4,$ and 5 is calculated to be $I(A{:}B) = 0.9999$, 1.313, 1.478, and 1.487, respectively (depicted by the green data points in Fig. 7). While the Shannon mutual information increases monotonically, it seems to level off for $d = 4$ and 5. On the other hand, $r_{\min}$ first increases and then decreases for $d = 5$. This means that we have reached a finite limit on the dimension in which the protocol can encode, while still resulting in higher generation rates per photon. The difference between these two quantities [$I(A{:}B)$ and $r_{\min}$] is the mutual information between Alice and Eve, in other words the information that is shared between Alice and Eve (denoted by the red shaded region in Fig. 7). From our results it is evident that the noise (attributed to a disturbance by Eve) grows faster than the correlations between Alice and Bob that can be used to generate a key. As this is not expected theoretically, this may be due to the complexity associated with encoding higher-dimensional states holographically on pixelated, finite-resolution, spatial light modulators. Our detection efficiency is low because our filter measurements are based on intensity masking and serve as a proof-of-principle experiment.

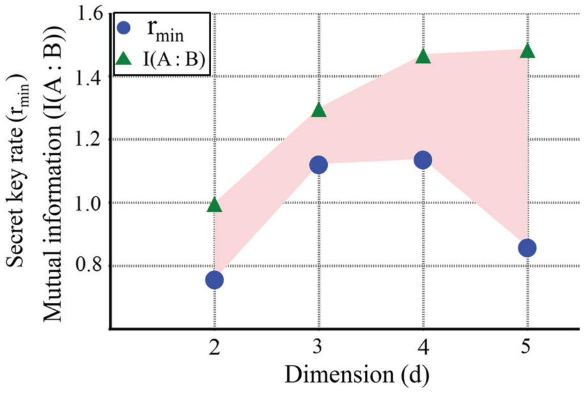

FIG. 7. (Color online) The Shannon mutual information $I(A{:}B)$ and the secret key rate $r_{\min}$ plotted as a function of the dimension. The shaded region denotes the mutual information between Alice and Eve.

VII. CONCLUSION

In this work, we have classically simulated a P&M and realized an EB QKD protocol for dimensions $d = 2$–5 using MUB encoded in the OAM degree of freedom. We show that our protocol, which is based on filter measurements, can be mapped back into the original MUB protocol, which uses full measurements. In particular, we verify our claim that detection efficiency depends on a basis choice and not on the element within a basis, an important consideration for the protocol to work. We show this explicitly for $d = 2$ and attest to the fact that this dependency holds for all dimensions. We infer from our measurements the average error rate, mutual information, and secret key generation rate per photon for each dimension. We observe that encoding in higher-dimensional MUBs leads to an increase in the encoding density per photon and increased key generation rates per photon. However, our implementation shows a decrease of generated secret key bits per carrier photon for dimension $d = 5$ compared to dimensions $d = 3$,4. Future studies are needed to determine whether this effect occurs independent of the particular implementation scheme and corresponds to a trend for higher dimensions.

ACKNOWLEDGMENTS

This work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation. T.K. acknowledges partial support from the National Research Foundation of South Africa [Grant No. 86325 (UID)].

APPENDIX

1. Shannon information

In order to analyze the security of our scheme, we employ the concept of mutual information given by Shannon. The Shannon entropy gives a measure of uncertainty for a random variable A with alphabet $\mathcal{A}$ and is defined as $H(A) = -\sum_{a\in\mathcal{A}} p(a)\log_2 p(a)$, where $p(a)$ is the probability of outcome a. The classical mutual information is defined as the amount by which the Shannon entropy on A decreases when one learns about B. The classical mutual information

$I(A{:}B)$ gives a degree of correlation between Alice (A), and Bob's (B) data and it is also an upper bound on the secret key rate. It is defined as $I(A{:}B) = H(A) + H(B) - H(A,B)$, where $H(A,B)$ is the joint entropy. The joint entropy is used to measure the total uncertainty about the pair (A,B). It is expressed as $H(A,B) = -\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} p(a,b)\log_2 p(a,b)$. After data processing, Alice and Bob apply a key map where their respective data is mapped to raw keys K and K'. In this step, the total probability distribution remains unchanged but the total classical mutual information changes to $I(A'{:}B)$, which is expressed as

$$I(A'{:}B) = H(A') + H(B) - H(A',B), \tag{A1}$$

where $H(A') = \sum_a \sum_{ij} p_{ij}^{aa} \log_2 p_{ij}^{aa}$. The joint entropy is defined in a similar manner as above.

2. Calculation of detection efficiencies

In this section, we show how to formalize and verify the claim that the detection efficiencies depend only on the bases but are the same for all elements within a basis. We demonstrate the calculation of detection efficiencies by comparing the expected and detected number of clicks for the case of qubit pairs ($d = 2$). For this purpose, we first calculate the expected number of detection events by following the light beam from the laser source to the detection device. Afterwards we relate them to the measured counts. By comparing the single count rates and the coincidence count rates we obtain an expression for the detection efficiency for each basis state.

a. Photon pair creation and action of the beam splitter. The state of the light exiting the laser source can be represented by a coherent state with complex parameter α which specifies the intensity and phase of the light:

$$|\alpha\rangle = \mathcal{D}(\alpha)|0\rangle = \exp(\alpha b_0^\dagger - \alpha^* b_0)|0\rangle, \tag{A2}$$

where $|0\rangle$ is the vacuum state, b_0 and $b_0^\dagger$ are annihilation and creation operators, respectively, with the index referring to OAM value $l = 0$. The operator $\mathcal{D}(\alpha)$ is called a displacement operator. The laser beam pumps a BBO crystal, creating pairs of photons with OAM values $\pm l$ by type-I parametric down conversion. This process can be modeled by the following transformation of creation operators:

$$b_0^\dagger \rightarrow \sum_\ell \sqrt{\chi_\ell}\, a_\ell^\dagger a_{-\ell}^\dagger, \tag{A3}$$

where χ_ℓ is the creation probability of a photon pair with OAM values $\pm\ell$ and $a_{\pm\ell}^\dagger$ are the corresponding creation operators. After passing through the BBO crystal, the light is sent to a 50:50 beam splitter, resulting in the following transformation:

$$a_\ell^\dagger \rightarrow \frac{1}{\sqrt{2}}(a_{\ell,A}^\dagger + a_{\ell,B}^\dagger), \tag{A4}$$

where A and B refer to the two beams exiting the beam splitter. Thus the combined action of the BBO crystal and the beam splitter reads

$$a_0^\dagger \rightarrow \sum_{\ell=0}^{\infty} \sqrt{\chi_\ell}\left(\frac{a_{\ell,A}^\dagger + a_{\ell,B}^\dagger}{\sqrt{2}}\right)\left(\frac{a_{-\ell,A}^\dagger + a_{-\ell,B}^\dagger}{\sqrt{2}}\right). \tag{A5}$$

It maps the displacement operator $\mathcal{D}(\alpha)$ to a squeeze operator $\mathcal{S}(\alpha\sqrt{\chi_\ell})$ given by

$$\begin{aligned}\mathcal{S}(\alpha\sqrt{\chi_\ell})&\\ = \exp\Bigg[&\alpha\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\frac{a_{\ell,A}^\dagger + a_{\ell,B}^\dagger}{\sqrt{2}}\right)\left(\frac{a_{-\ell,A}^\dagger + a_{-\ell,B}^\dagger}{\sqrt{2}}\right)\\ &-\alpha^*\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\frac{a_{\ell,A} + a_{\ell,B}}{\sqrt{2}}\right)\left(\frac{a_{-\ell,A} + a_{-\ell,B}}{\sqrt{2}}\right)\Bigg].\end{aligned} \tag{A6}$$

Thus the initial coherent state is transformed into a (two-mode) squeezed vacuum state: $|\tilde{\alpha}\rangle = \mathcal{S}(\alpha\sqrt{\chi_\ell})|0\rangle$. For a small value of $\alpha\sqrt{\chi_\ell}$, the state $|\tilde{\alpha}\rangle$ can be approximated to the first order in $\alpha\sqrt{\chi_\ell}$ as

$$\begin{aligned}|\tilde{\alpha}\rangle \approx \mathcal{N}\Bigg[&1 + \alpha\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\frac{a_{\ell,A}^\dagger + a_{\ell,B}^\dagger}{\sqrt{2}}\right)\\ &\times\left(\frac{a_{-\ell,A}^\dagger + a_{-\ell,B}^\dagger}{\sqrt{2}}\right)\Bigg]|0\rangle,\end{aligned} \tag{A7}$$

where $\mathcal{N}$ is the normalization constant. The vacuum does not play any role as far as photon detections are concerned; thus one can ignore the vacuum component. This results in the (un-normalized) state $|\psi\rangle$ which reads

$$|\psi\rangle = \alpha\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\frac{a_{\ell,A}^\dagger + a_{\ell,B}^\dagger}{\sqrt{2}}\right)\left(\frac{a_{-\ell,A}^\dagger + a_{-\ell,B}^\dagger}{\sqrt{2}}\right)|0\rangle, \tag{A8}$$

$$\begin{aligned}&= \frac{\alpha}{2}\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}(a_{\ell,A}^\dagger a_{-\ell,A}^\dagger + a_{\ell,A}^\dagger a_{-\ell,B}^\dagger\\ &\quad + a_{\ell,B}^\dagger a_{-\ell,A}^\dagger + a_{\ell,B}^\dagger a_{-\ell,B}^\dagger)|0\rangle.\end{aligned} \tag{A9}$$

b. Measurements. After the BBO crystal and the beam splitter filter measurements projecting onto individual basis modes were carried out independently in both beams A and B, the signal for each basis mode was detected by means of avalanche photodiodes. These detectors respond to incident photons but do not discriminate between a single photon and multiple photons. However, the probability for a click varies for different photon numbers. The probability to obtain a click in a filter measurement of mode s can be modeled by the expectation value of the effect P_s defined by

$$P_s = \sum_{n=1}^{\infty}\eta_s^{(n)}|n_s\rangle\langle n_s|, \tag{A10}$$

where $\eta_s^{(n)}$ represents the probability for n photons in mode s to trigger a detector click and reads [34]

$$\begin{aligned}\eta_s^{(n)} &= 1 - \left(1 - \eta_s^{(1)}\right)^n,\\ &\approx n\eta_s^{(1)} \text{ for small } \eta_s^{(1)}.\end{aligned} \tag{A11}$$

Because of photon loss on the path from source to detector and nonideal detection, only a fraction of the detection events expected under ideal conditions is measured in the experiment. We attribute any loss to nonideal detection. The probability of coincidence can be calculated as an expectation value of the operator $P_s \otimes P_{s'}$ with respect to the state $|\psi\rangle$ [Eq. (A9)] after the beam splitter.

From Eq. (A9) it is clear that only the single-photon components of state $|\psi\rangle$ can yield a click of detector A for OAM value ℓ, leading to a detection probability of

$$p_{\ell,A} = \langle\psi|P_\ell \otimes \mathbb{I}|\psi\rangle = \eta^{(1)}_{\ell,A}|\alpha|^2\chi_\ell/2. \qquad \text{(A12)}$$

Similarly, we can calculate the other probabilities as

$$p_{-\ell,A} = \langle\psi|P_{-\ell} \otimes \mathbb{I}|\psi\rangle = \eta^{(1)}_{-\ell,A}|\alpha|^2\chi_\ell/2, \qquad \text{(A13)}$$

$$p_{\ell,B} = \langle\psi|\mathbb{I} \otimes P_\ell|\psi\rangle = \eta^{(1)}_{\ell,B}|\alpha|^2\chi_\ell/2, \qquad \text{(A14)}$$

$$p_{-\ell,B} = \langle\psi|\mathbb{I} \otimes P_{-\ell}|\psi\rangle = \eta^{(1)}_{-\ell,B}|\alpha|^2\chi_\ell/2. \qquad \text{(A15)}$$

The probability of the coincidence count in detector A with OAM value ℓ and in detector B with OAM value $-\ell$ amounts to

$$p_{\ell,A,-\ell,B} = \langle\psi|P_\ell \otimes P_{-\ell}|\psi\rangle = \eta^{(1)}_{\ell,A}\eta^{(1)}_{-\ell,B}|\alpha|^2\chi_\ell/4. \qquad \text{(A16)}$$

For the measured count of clicks $C_{\ell,A}$ in detector A with OAM value ℓ and the measured count $C_{-\ell,B}$ in detector B with OAM value $-\ell$ we obtain the following expressions:

$$C_{\ell,A} = Np_{\ell,A}, \qquad \text{(A17)}$$

$$C_{-\ell,B} = Np_{-\ell,B}, \qquad \text{(A18)}$$

$$C_{\ell,A,-\ell,B} = Np_{\ell,A,-\ell,B}, \qquad \text{(A19)}$$

where N is the number of photon pairs created by consecutive pump pulses during the measurement period. For the coincidence counts $C_{\ell,A,-\ell,B}$ in the last equation it is assumed that photon loss in beams A and B are independent. Note that $p_{\ell,A,-\ell,B}/p_{\ell,A} = \eta^{(1)}_{-\ell,B}/2$ and hence one can calculate the efficiencies as follows:

$$\eta^{(1)}_{-\ell,B} = 2\frac{C_{\ell,A,-\ell,B}}{C_{\ell,A}}, \qquad \text{(A20)}$$

$$\eta^{(1)}_{\ell,A} = 2\frac{C_{\ell,A,-\ell,B}}{C_{-\ell,B}}. \qquad \text{(A21)}$$

For the SLM-filter setting $(|\ell\rangle \pm |-\ell\rangle)/\sqrt{2}$, which is a superposition of $\pm\ell$ OAM modes, the corresponding creation operators read $a^\dagger_\pm \equiv (a^\dagger_{\ell,A} \pm a^\dagger_{-\ell,A})/\sqrt{2}$. Thus we can represent $a^\dagger_{\ell,A}$ and $a^\dagger_{-\ell,A}$ in terms of $a^\dagger_\pm$ as

$$a^\dagger_{\pm\ell} = \frac{a^\dagger_{+,A} \pm a^\dagger_{-,A}}{\sqrt{2}}. \qquad \text{(A22)}$$

By substituting Eq. (A22) in Eq. (A9) we obtain

$$|\psi\rangle = \frac{\alpha}{4}\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\sqrt{2}\frac{(a^\dagger_{+,A})^2}{\sqrt{2}} - \sqrt{2}\frac{(a^\dagger_{-,A})^2}{\sqrt{2}} + 2a^\dagger_{+,A}a^\dagger_{+,B} - 2a^\dagger_{-,B}a^\dagger_{-,A} + \sqrt{2}\frac{(a^\dagger_{+,B})^2}{\sqrt{2}} - \sqrt{2}\frac{(a^\dagger_{-,B})^2}{\sqrt{2}}\right)|0\rangle. \qquad \text{(A23)}$$

Thus the probability of a click in detector A for the SLM setting $+$ amounts to $p_{+,A} = \eta^{(1)}_{+,A}|\alpha|^2\chi_\ell/2$, while the coincidence probability for the SLM setting $+$ in detector A and detector B reads $p_{+,A,+,B} = \eta^{(1)}_{+,A}\eta^{(1)}_{+,B}|\alpha|^2\chi_\ell/4$. The observed number of clicks is related to the expected detection counts as follows:

$$C_{+,A} = Np_{+,A}, \qquad \text{(A24)}$$

$$C_{+,B} = Np_{+,B}, \qquad \text{(A25)}$$

$$C_{+,A,+,B} = Np_{+,A,+,B}. \qquad \text{(A26)}$$

Since $p_{+,A,+,B}/p_{+,A} = \eta^{(1)}_{+,B}/2$, it follows for the efficiencies that

$$\eta^{(1)}_{+,B} = 2\frac{C_{+,A,+,B}}{C_{+,A}}, \qquad \text{(A27)}$$

$$\eta^{(1)}_{+,A} = 2\frac{C_{+,A,+,B}}{C_{+,B}}. \qquad \text{(A28)}$$

Similarly for SLM settings $(|\ell\rangle \pm i|-\ell\rangle)/\sqrt{2}$, the state $|\psi\rangle$ can be rewritten as

$$|\psi\rangle = \frac{\alpha}{4}\sum_{\ell=0}^{\infty}\sqrt{\chi_\ell}\left(\sqrt{2}\frac{(a^\dagger_{+y,A})^2}{\sqrt{2}} + \sqrt{2}\frac{(a^\dagger_{-y,A})^2}{\sqrt{2}} + 2a^\dagger_{+y,A}a^\dagger_{+y,B} + 2a^\dagger_{-y,B}a^\dagger_{-y,A} + \sqrt{2}\frac{(a^\dagger_{+y,B})^2}{\sqrt{2}} + \sqrt{2}\frac{(a^\dagger_{-y,B})^2}{\sqrt{2}}\right)|0\rangle, \qquad \text{(A29)}$$

where

$$a^\dagger_{\pm y,A} = \frac{a^\dagger_{\ell,A} \pm ia^\dagger_{-\ell,A}}{\sqrt{2}}. \qquad \text{(A30)}$$

Thus the relation for the efficiencies in this filter setting is obtained as

$$\eta^{(1)}_{+y,B} = 2\frac{C_{+y,A,+y,B}}{C_{+y,A}}, \qquad \text{(A31)}$$

$$\eta^{(1)}_{+y,A} = 2\frac{C_{+y,A,+y,B}}{C_{+y,B}}. \qquad \text{(A32)}$$

Using the expressions derived above, we calculated the detection efficiencies for the case of a two-level system for different SLM settings (Table I). We found that even though the detection efficiencies vary for different bases, the fluctuation in the values is very small for all the basis vectors within each basis, which proves the claim for qubits.

Furthermore, this method can be used to show that the detection efficiencies are independent of the basis vectors within each basis, regardless of the dimension. However, let us point out that the analysis of our measurement data indicated an anomaly for the detection efficiency for the OAM value $\ell = 0$, which is different from the other values of OAM. Although not so important in the present context, this case has to be investigated more carefully when it comes to actual key transmission and will be the subject of future work.

TABLE I. Detection efficiencies for different detectors projecting on different bases vectors. Here the first two vectors belong to the σ_z basis, the following two to the σ_x basis, and the last two to the σ_y basis.

Basis vectors	Detector A	Detector B
1	0.01504	0.02145
2	0.01517	0.02106
3	0.00536	0.00886
4	0.00503	0.00727
5	0.00508	0.00787
6	0.00556	0.00874

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[2] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[3] C. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, Vol. 175 (Bangalore, India, 1984).

[4] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[5] D. Bruß, Phys. Rev. Lett. **81**, 3018 (1998).

[6] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).

[7] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).

[8] J. Schwinger, Proc. Natl. Acad. Sci. USA **46**, 570 (1960).

[9] I. Ivonovic, J. Phys. A **14**, 3241 (1981).

[10] W. Wootters and B. Fields, Ann. Phys. **191**, 363 (1989).

[11] R. B. A. Adamson and A. M. Steinberg, Phys. Rev. Lett. **105**, 030406 (2010).

[12] A. Fernández-Pérez, A. B. Klimov, and C. Saavedra, Phys. Rev. A **83**, 052332 (2011).

[13] D. Giovannini, J. Romero, J. Leach, A. Dudley, A. Forbes, and M. J. Padgett, Phys. Rev. Lett. **110**, 143601 (2013).

[14] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. **78**, 405 (1997).

[15] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).

[16] S. M. Barnett, *Quantum Information* (Oxford University Press, Oxford, 2009).

[17] T. Durt, B.-G. Englert, I. Bengtsson, and K. Zyczkowski, Int. J. Quantum Inf. **08**, 535 (2010).

[18] L. Sheridan and V. Scarani, Phys. Rev. A **82**, 030301 (2010).

[19] A. Ferenczi and N. Lütkenhaus, Phys. Rev. A **85**, 052310 (2012).

[20] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, Phys. Rev. A **45**, 8185 (1992).

[21] H. Bechmann-Pasquinucci and A. Peres, Phys. Rev. Lett. **85**, 3313 (2000).

[22] S. Gröblacher, T. Jennewein, A. Vaziri, G. Weihs, and A. Zeilinger, New J. Phys. **8**, 75 (2006).

[23] I.-C. Yu, F.-L. Lin, and C.-Y. Huang, Phys. Rev. A **78**, 012344 (2008).

[24] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, Phys. Rev. Lett. **88**, 127902 (2002).

[25] B. Rodenburg, M. J. P. Lavery, M. Malik, M. N. O'Sullivan, M. Mirhosseini, D. J. Robertson, M. J. Padgett, and R. W. Boyd, Opt. Lett. **37**, 3735 (2012).

[26] H. Lo, H. Chau, and M. Ardehali, J. Cryptology **18**, 133 (2005).

[27] M. T. Gruneisen, W. A. Miller, R. C. Dymale, and A. M. Sweiti, Appl. Opt. **47**, A32 (2008).

[28] V. Arrizón, U. Ruiz, R. Carrada, and A. González, J. Opt. Soc. Am. A **24**, 3500 (2007).

[29] J. A. Davies, D. M. Cottrell, J. Campos, M. J. Yzuel, and I. Moreno, Appl. Opt. **38**, 5004 (1999).

[30] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Opt. Express **19**, 3542 (2011).

[31] D. Klyshko, Soviet Physics Uspekhi **31**, 74 (1988).

[32] M. G. McLaren, J. Romero, M. J. Padgett, F. S. Roux, and A. Forbes, arXiv:1306.2767.

[33] S. S. R. Oemrawsingh, J. A. de Jong, X. Ma, A. Aiello, E. R. Eliel, G. W. 't Hooft, and J. P. Woerdman, Phys. Rev. A **73**, 032339 (2006).

[34] T. Jennewein, M. Barbieri, and A. G. White, J. Mod. Phys. **58**, 276 (2011).