Eavesdropper's ability to attack a free-space quantum-key-distribution receiver in atmospheric turbulence

Poompong Chaiwongkhot,[1,2,*] Katanya B. Kuntz,[1,2,†] Yanbao Zhang,[1,2,3,4] Anqi Huang,[5,1,6] Jean-Philippe Bourgoin,[7,1,2] Shihan Sajeed,[8,1,6] Norbert Lütkenhaus,[1,2] Thomas Jennewein,[1,2,9] and Vadim Makarov[10,11,12,2]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

[2]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

[3]*NTT Basic Research Laboratories, NTT Corporation, Atsugi, Kanagawa 243-0198, Japan*

[4]*NTT Research Center for Theoretical Quantum Physics, NTT Corporation, Atsugi, Kanagawa 243-0198, Japan*

[5]*Institute for Quantum Information and State Key Laboratory of High Performance Computing, College of Computer, National University of Defense Technology, Changsha 410073, People's Republic of China*

[6]*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

[7]*Aegis Quantum, Waterloo, ON, Canada*

[8]*Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada M5S 3G4*

[9]*Quantum Information Science Program, Canadian Institute for Advanced Research, Toronto, ON, Canada M5G 1Z8*

[10]*Russian Quantum Center, Skolkovo, Moscow 143025, Russia*

[11]*Shanghai Branch, National Laboratory for Physical Sciences at Microscale and CAS Center for Excellence in Quantum Information, University of Science and Technology of China, Shanghai 201315, People's Republic of China*

[12]*NTI Center for Quantum Communications, National University of Science and Technology MISiS, Moscow 119049, Russia*

(Received 6 February 2019; published 14 June 2019)

The ability of an eavesdropper (Eve) to perform an intercept-resend attack on a free-space quantum-key-distribution (QKD) receiver by precisely controlling the incidence angle of an attack laser has been previously demonstrated. However, such an attack could be ineffective in the presence of atmospheric turbulence due to beam wander and spatial mode aberrations induced by the air's varying index of refraction. We experimentally investigate the impact turbulence has on Eve's attack on a free-space polarization-encoding QKD receiver by emulating atmospheric turbulence with a spatial light modulator. Our results identify how well Eve would need to compensate for turbulence to perform a successful attack by either reducing her distance to the receiver or using beam wavefront correction via adaptive optics. Furthermore, we use an entanglement-breaking scheme to find a theoretical limit on the turbulence strength that hinders Eve's attack.

DOI: 10.1103/PhysRevA.99.062315

I. INTRODUCTION

Quantum key distribution (QKD) allows two distant parties to exchange secret keys with—in theory—unconditional security [1,2]. However, in practice, a QKD system is often not perfect, and unconditional security cannot be guaranteed. Any imperfections in the physical implementation of a QKD scheme can lead to side channels that could be exploited by an eavesdropper (Eve) and compromise security [3–17]. Therefore, it is of utmost importance to perform security evaluations of practical systems, i.e., scrutinize vulnerabilities, determine useful testing methodologies, and assess the risk to formulate countermeasures for preventing successful attacks.

A widely studied implementation of QKD utilizes free-space communication between two parties (Alice and Bob) through the atmosphere [18–24], which allows for long-distance point-to-point links on the order of 100 km. This communication distance can be extended even further to the global scale by introducing satellite-based QKD systems [22–29]. However, free-space communication can be vulnerable to an eavesdropper attack, such as when Eve precisely controls the incidence angle of an attack laser directed at Bob's QKD receiver. Directing a laser in this way can induce a change in the measurement efficiencies of one (or more) detection channels, which enables Eve to do an intercept-resend (IR) attack that may compromise the system's security [13,30].

The success of this spatial mode attack depends on the eavesdropper's ability to precisely maintain specific beam angles to a free-space QKD receiver, which attacks different detection channels. Atmospheric turbulence could compromise or even prevent such an attack as turbulence causes a beam to randomly wander along its trajectory, as well as inducing various optical aberrations such as astigmatism, defocus, coma, etc. Stronger turbulence conditions result in a larger variance in the amount of beam wander [31]. Consideration of these physical limitations on Eve is not usually included in the theoretical security analysis of a system, but can be useful to verify whether an attack is feasible under more realistic conditions.

In this paper, we experimentally determine the minimum strength of atmospheric turbulence that could prevent a successful attack on our free-space polarization-based QKD

*poompong.ch@gmail.com
†katanyab@gmail.com

receiver by emulating atmospheric turbulence using a phase-only spatial light modulator (SLM). Since there are limitations on how well adaptive optics can correct for turbulence, our paper explores to what level Eve must correct her attack beam to still be successful [32,33]. We assume that the sender (Alice) and the receiver (Bob) only monitor the total count rates (as opposed to the rates of individual channels), and that they use a nondecoy state Bennett-Brassard 1984 (BB84) protocol [1]. We also assume that Eve has access to a weak coherent pulse source and state of the art photodetectors and does not have a quantum repeater. Furthermore, we assume that Eve cannot replace the quantum channel with a lossless channel. We find that an attack on our free-space receiver could still succeed if Eve can correct the tip-tilt mode for turbulence as strong as $r_0 = 1.53$ cm (assuming an initial beam diameter of 20 cm), where r_0 is the atmospheric coherence length. This result defines an "unsafe radius" of 543 m around Bob's receiver in typical sea level turbulence conditions where Eve's attack could be successful if done within this radius.

First we discuss our SLM setup used to emulate atmospheric turbulence, and how we verified its accuracy and reproducibility, in Sec. II. Then we describe the components and operation of our free-space polarization-based QKD receiver under test in Sec. III. In Secs. IV and V, we discuss the results from spatial mode attacks performed in various turbulence strengths, following a similar procedure to Sajeed *et al.* in Ref. [13]. Finally, in Sec. VI we discuss an entanglement breaking scheme proposed by Zhang and Lütkenhaus in Ref. [34], to theoretically verify if there exists an attack strategy for Eve, even if Alice and Bob know about their detection efficiency mismatch, and monitor the statistics of all possible detection outcomes. We conclude in Sec. VII.

II. TURBULENCE EMULATOR

We use a phase-only SLM to emulate a turbulent QKD channel in the laboratory. One advantage of using a SLM as opposed to performing the experiment outdoors is the ability to generate a range of turbulence strengths, from weak upper atmosphere to stronger sea level conditions. In addition, by performing our experiment in a laboratory, we are immune to the unpredictability of an outdoor environment, allowing us to repeat the same attack angles on our free-space QKD receiver under reproducible turbulence conditions.

Our model uses the "thin phase screen approximation," which emulates turbulence using a single random phase screen in the aperture of the receiver, as opposed to requiring two holograms to model multiple parameters that incorporate both phase and amplitude variations [35]. We assume that Eve's laser can mimic the intensity variations caused by turbulence (scintillation) [36]. Note that the absence of these fluctuations could arouse Alice and Bob's suspicion of an eavesdropper in the channel, although fluctuations on the time scale of scintillation at 1 s or less are rarely monitored in practice.

In order to reproduce the random statistics of turbulence, we load a series of 29 phase maps per turbulence strength on the SLM to distort the optical wavefront. The strength of the turbulence is completely characterized by the ratio of the initial beam diameter, D, to the atmospheric coherence length, r_0; turbulence dominates over diffractive effects when $D/r_0 \gg 1$.

We generate our phase holograms based on the well-known Kolmogorov model [37] that uses a weighted superposition of Zernike polynomials for the basis set [38]. There are several advantages to using Zernike polynomials to generate the holograms as their weights can be analytically calculated based on the turbulence strength [39]. Furthermore, Zernike polynomials directly relate to known optical aberrations, such as tip-tilt, defocus, astigmatism, coma, etc. Therefore, it is straightforward to characterize the SLM's ability to reliably and precisely emulate atmospheric turbulence by comparing calculated Zernike polynomial coefficients to those reconstructed by a measurement device, such as a wavefront sensor.

The radial phase function $\phi(\rho, \theta)$ that describes each hologram is given by a weighted sum of several Zernike polynomials as $\phi(\rho, \theta) = \sum_i c_i Z_i$, where Z_i and c_i are the Zernike polynomial and corresponding coefficient for the ith polynomial, respectively, following the Noll labeling convention and normalization constants [38]. We use 44 Zernike polynomials to ensure a complex spatial structure that can accurately emulate a range of atmospheric turbulence strengths.

Based on the Kolmogorov model [37,39], if we assume that the Zernike coefficients are normally distributed with mean zero, then c_i are random drawings from distributions with variance σ_{nm}^2 defined as

$$\begin{aligned} \sigma_{nm}^2 &= I_{nm}(D/r_0)^{5/3}, \\ r_0 &= 1.68\left(C_n^2 L k^2\right)^{-3/5}, \\ I_{nm} &= \frac{0.15337(-1)^{n-m}(n+1)\Gamma(14/3)\Gamma(n-5/6)}{\Gamma(17/6)^2\Gamma(n+23/6)}, \end{aligned} \tag{1}$$

where C_n^2 is the refractive-index structure constant of the atmosphere, L is the path length through the turbulent atmosphere that has a constant C_n^2, $k = 2\pi/\lambda$, λ is the laser wavelength, and Γ is the Gamma function. The indices n and m are related to the Zernike polynomial order following the Noll labeling convention, where $n \geqslant |m|$ and $n - m$ is even [38]. We note that the subscript "n" of C_n^2 is not related to the index "n" used in the Zernike polynomials, but instead to the refractive index of the atmosphere. A single value of C_n^2 is used when calculating σ_{nm}^2 over each of the n and m indices for each atmospheric strength modeled. A large C_n^2 (small r_0) value corresponds to stronger atmospheric turbulence. An example of stronger turbulent conditions that could be found at sea level corresponds to $r_0 = 1.00$ cm over $L = 1$ km for $D = 20$ cm at $\lambda = 532$ nm, whereas weaker conditions at high altitude correspond to $r_0 = 7.00$ cm [37].

Since Zernike polynomials directly relate to known optical aberrations, we can use simple equations and measurement devices (CCD camera and wavefront sensor), to independently verify and characterize our turbulence emulator. Figure 1 shows both the simulated and measured far-field intensity distributions of a beam after its wavefront has been distorted by the SLM hologram. Each hologram shown is one example from a set of 29 holograms per r_0 value used to emulate how different strengths of turbulence would affect a 20-cm beam at 532 nm. We experimentally image the far field by placing a camera in the focal plane of a lens that is located

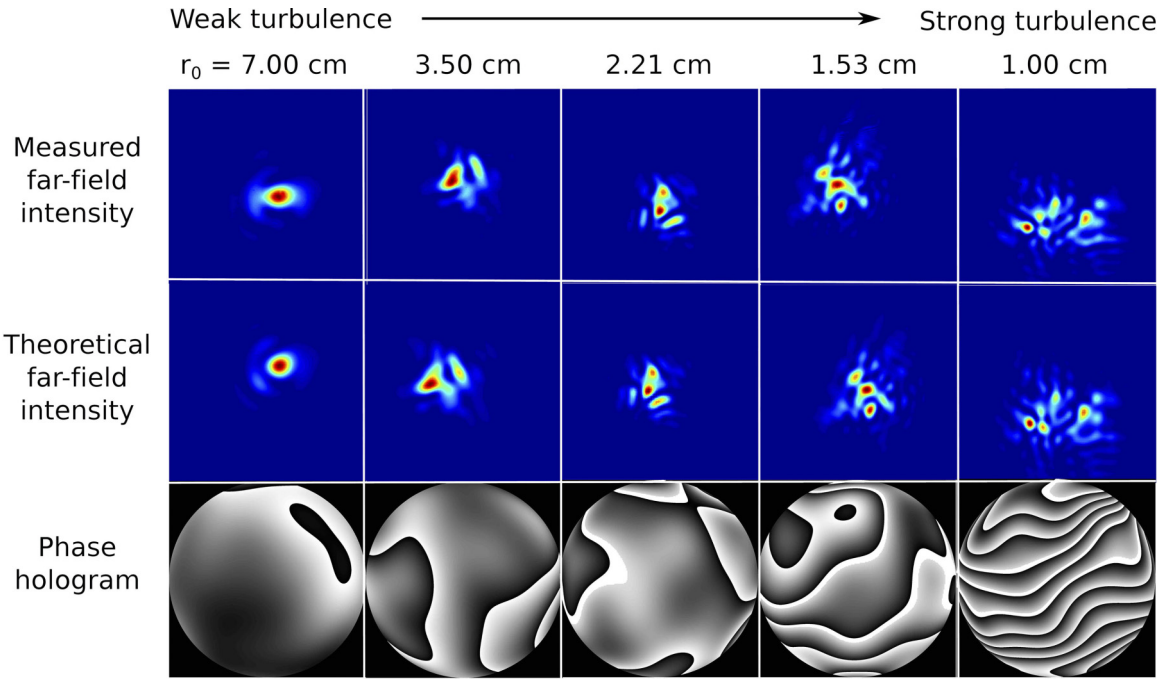

FIG. 1. Comparison between measured and theoretical far-field intensity distributions of a laser beam corresponding to one of 29 SLM phase holograms per turbulence strength (r_0) for a beam with $D = 20$ cm and $\lambda = 532$ nm. The grayscale in the holograms represents a 0 to 2π phase range. The results show our SLM setup accurately emulates a range of turbulence strengths.

one focal length from the SLM. This arrangement maps the phase wavefront imprinted on the beam by the hologram into an intensity distribution at the camera plane. Note that we include an additional x grating in the hologram (not shown for clarity) to spatially separate the first-order diffracted beam from the zeroth-order one, as only the first-order beam contains the pure phase wavefront. The zeroth-order (and higher-order) diffracted beams were carefully blocked shortly after the SLM.

We also verify our turbulence emulator by examining the centroid deviations caused by each hologram. This is an important characterization as beam displacements due to turbulence could dominate Eve's ability to repeatedly send a beam at precise angles to the receiver. Beam wander is the strongest effect on average as the tip-tilt coefficients ($n = 1$, $m = \pm 1$) have the largest weights overall [$I_{11} = 0.45$ from Eq. (1)], whereas defocus ($I_{20} = 0.02$) and astigmatism ($I_{22} = 0.02$) have a smaller contribution on average. Higher-order aberrations can also cause centroid displacement, especially in the case of stronger turbulence.

There is a direct relationship between the tilt angle variance of centroid displacement for two uncorrelated axes σ^2 and the turbulence strength r_0, which is given by [31]

$$\sigma^2 = 0.364\left(\frac{D}{r_0}\right)^{5/3}\left(\frac{\lambda}{r_0}\right)^{5/3}. \tag{2}$$

Since this equation is independent of the method used to emulate turbulence, we can verify whether the 29 chosen phase holograms accurately portray the statistics of atmospheric turbulence both theoretically via computer simulations of far-field intensity distributions and experimentally through our SLM setup. This independent verification ensures that the holograms are accurate, as well as that the SLM is correctly imprinting the phase mask onto the beam.

The centroid displacement data presented in Fig. 2 correspond to low-altitude sea level turbulence ($r_0 = 1.00$ cm for a 20-cm beam). The simulated centroid displacements from 500 holograms are shown in Fig. 2(a). Each data point corresponds to a unique hologram [Fig. 2(c)] and far-field intensity distribution [Fig. 2(d)]. The simulated centroids follow a Gaussian distribution with a standard deviation σ that is in agreement with Eq. (2). These results confirm that the phase holograms we calculated properly emulate the statistics of low-altitude sea level turbulence, irrespective of the SLM setup. Similar tests were performed to verify the sets of holograms for each r_0 value tested in this experiment.

We compare simulated and measured centroid displacements of 29 holograms per r_0 strength in Fig. 2(b). The number of holograms used in the hacking experiment was limited to reduce data acquisition time and stability issues while scanning. Therefore, we chose 29 holograms from a larger distribution of 500 to emulate each r_0 strength. The holograms were chosen based on their centroid displacements being approximately 0.5σ, σ, 2σ, and 3σ from the origin [along the dashed circles outlined in Figs. 2(a) and 2(b)], along with one histogram with no turbulence representing 0σ. The centroid results, along with the qualitative comparison between theoretical and measured far-field intensity distributions (Fig. 1), confirmed we had excellent agreement between theory and experiment for turbulence emulated by our SLM setup. The 29th hologram always emulates 0σ displacement with no turbulence. The contribution of each of the 29 holograms to the emulated turbulence in subsequent experiments

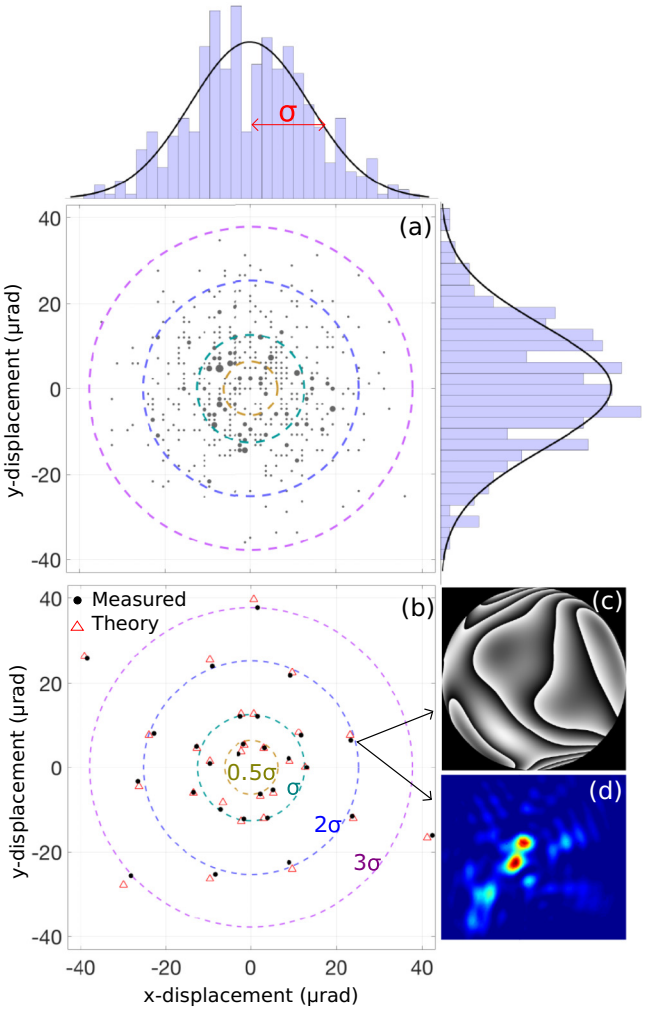

FIG. 2. Turbulence emulator characterization for $r_0 = 1.00$ cm, $D = 20$ cm, and $\lambda = 532$ nm. (a) Simulated centroid displacements corresponding to 500 phase holograms (σ is the two-axis standard deviation). The diameter of each data point is proportional to the count frequency. The centroid displacement distribution is normally distributed along both axes in agreement with Eq. (2). (b) Comparison between measured and simulated centroid displacements for a subset of 29 holograms. This subset was chosen to represent the normal statistical distribution of the 500-hologram set. The measured values are within error of most theoretical predictions (error bars for measured data are represented by diameter of data points). (c) Phase hologram and (d) far-field intensity distribution corresponding to one centroid data point.

is weighted by its probability of occurrence, which follows a Gaussian distribution. This probability of occurrence is a definite integral of normalized Gaussian distribution over the annulus formed by the adjacent radii shown in Fig. 2(b). We refer to each annulus by the name of its inner radius, near which its holograms are located. The 0σ annulus, extending from zero (where its hologram is located) to 0.5σ radius, has the weight of 0.1175. The 0.5σ annulus has the weight of 0.2760, 1σ has the weight of 0.4712, 2σ has the weight of 0.1242, and 3σ (extending to infinity) has the weight of 0.0111.

III. TEST SETUP FOR THE QKD SYSTEM

We use our turbulence emulator to study the effect of turbulence on free-space detection efficiency mismatch. Eve's experimental setup consists of two parts: the turbulence emulator (SLM) and the beam scanning unit, as shown in Fig. 3. Our source is a 532-nm continuous-wave laser that is first sent through a polarization beam splitter $\mathrm{PBS_E}$ (Thorlabs CCM1-PBS251) to transmit only horizontally polarized light to the SLM, which ensures phase-only modulation. The beam's wavefront after the SLM represents propagation through atmospheric turbulence of a particular strength. We use a quarter-wave plate $\mathrm{QWP_E}$ (Thorlabs AQWP10M-600) to rotate horizontal light to circularly polarized light to equalize the QKD receiver detector signals on the four polarization channels. Eve's scanning lens $\mathrm{L_E}$ is mounted on a two-axis motorized translation stage (Thorlabs MAX343/M), which scans the attack beam's angle. A half-wave plate $\mathrm{HWP_E}$ (Thorlabs AHWP10M-600) and neutral density filter $\mathrm{ND_E}$ (Thorlabs ND30A) are used to control Eve's intensity. Finally, the receiver is placed 13 m away from $\mathrm{L_E}$.

The QKD receiver under test is a prototype for a quantum communication satellite [26], which uses a passive basis choice to detect polarization-encoded light. Its telescope consists of a focusing lens L1 (diameter of 50 mm with a focal length $f = 250$ mm; Thorlabs AC508-250-A) and a collimating lens L2 (diameter of 5 mm with $f = 11$ mm; Thorlabs A397TM-A). The collimated beam of $\lesssim$2 mm diameter then passes through a 50:50 beam splitter BS (custom pentaprism [26]) and a pair of polarization beam splitters PBS1 and PBS2 (Thorlabs PBS121). The purpose of PBS2 is to increase the polarization extinction ratio in the reflected path from PBS1. The four lenses L3 (Thorlabs PAF-X-18-PC-A) focus the beams into four multimode fibers, each with a core diameter of 105 μm (Thorlabs M43L01), which are connected to single-photon detectors (Excelitas SPCM-AQRH-12-FC). We use one set of polarization optics and detectors to measure diagonal **D** and antidiagonal **A** polarizations by rotating them 45° relative to the horizontal **H** and vertical **V** polarization detectors. We note that this receiver under test does not contain any active pointing system or adaptive optics.

IV. ATTACK USING SPATIAL MODE DETECTION EFFICIENCY MISMATCH

This paper assumes that Alice and Bob generate a secret key using a nondecoy state BB84 protocol [1]. We also make the weaker assumption presented in Ref. [13] that they only monitor the total detection rate for evidence of Eve's attack rather than the counts of each channel. Additionally, we assume Alice and Bob also monitor only the average error rate over the four channels, and terminate the protocol if the average quantum bit error rate (QBER) over the four channels is higher than a 8% threshold [12].

The attack model we consider is an intercept-resend attack called the faked-state attack [4,40]. In this attack, Eve attempts to deterministically control Bob's basis choice and detection outcomes without terminating the protocol. To achieve this, Eve needs to maintain the expected detection rate between Alice and Bob, and keep the QBER below the termination

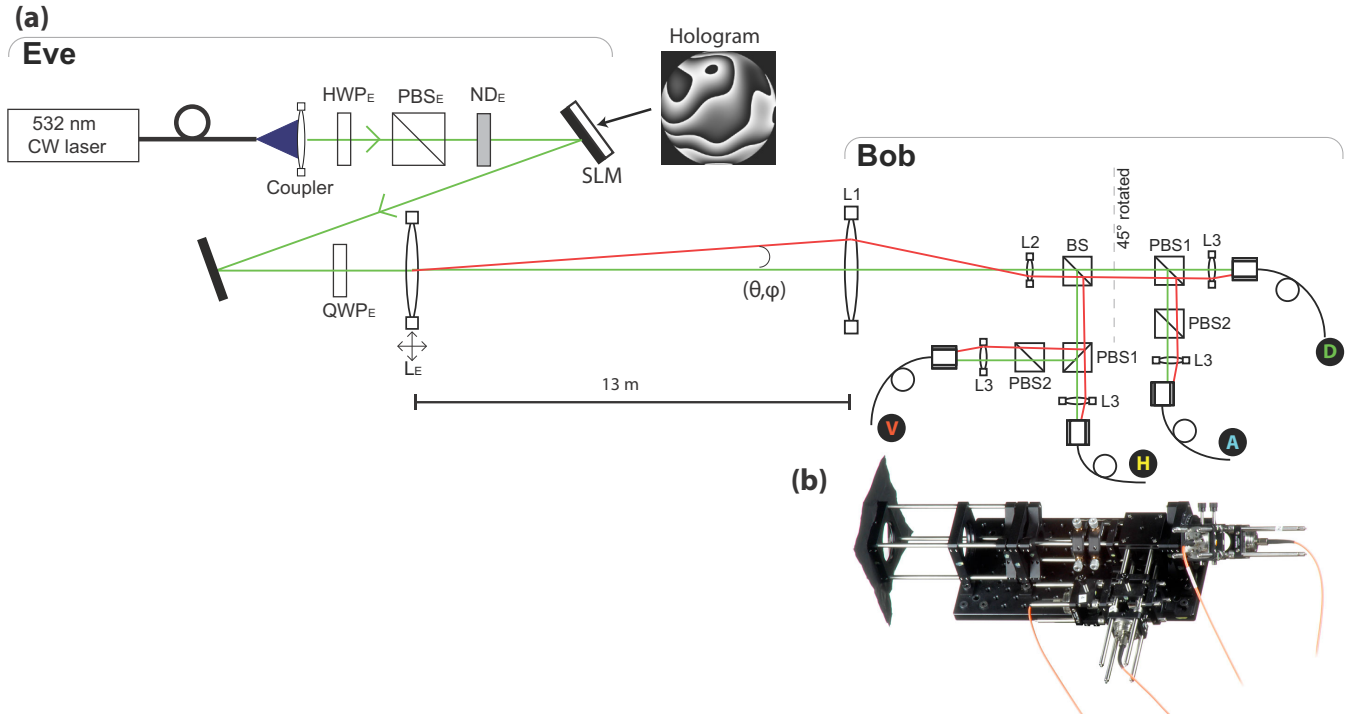

FIG. 3. Scanning setup. (a) Experimental setup of our spatial mode attack in a turbulent channel, top view (drawing not to scale). The green central ray that is parallel to the optical axis denotes normal alignment of Alice's beam into Bob's receiver. The red rays show the optical path of Eve's scanning beam when tilted at an angle (θ, ϕ) via lens L_E. CW, continuous wave; HWP, half-wave plate; QWP, quarter-wave plate; BS, beam splitter; PBS, polarization beam splitter; ND, neutral density filter; SLM, spatial light modulator; L, lens. (b) Photograph of the actual free-space QKD receiver for detecting polarization-encoded light.

threshold during her attack. In our practical attack model, we assume that Eve knows the attack angles for each polarization state, as well as the detection efficiency ratios between the detectors. Eve intercepts signals sent by Alice using an active basis choice receiver and superconducting nanowire detectors with an overall detection efficiency of 85%. This interception could be done right in front of Alice's setup, to negate the turbulence effect on Eve's measurement. She then generates a signal with the same polarization state as her measurement result, and sends it to Bob at the ideal attack angle. These fake signals may suffer from atmospheric turbulence in transmission to Bob.

We assume that Eve is restricted to today's technology and uses a weak coherent state for her resend signal. Thus, Eve can control the mean photon number μ of her pulses, as well as mimic scintillation caused by turbulence in the free-space channel to avoid arousing suspicion. Several free-space QKD systems employ pointing and tracking systems that use a bright beacon source and wavefront sensor [23,24,41] which could be adapted by Bob to monitor and correct beam wander. However, this pointing system uses a separate beacon laser at a different wavelength. This beacon laser does not need to be tampered with by Eve, and the pointing is unaffected by her attack. In the worst case, Eve could perform an intercept-and-resend attack on the beacon beam such that Bob's receiver is pointed according to her designated direction. Thus, this pointing and correction system cannot prevent the attack in our model.

To verify the possibility of a successful attack, we use an optimization program to find the mean photon number that Eve should use for each attack angle to match Bob's expected total detection probability while minimizing the QBER. Our detailed attack model and the optimization process are explained in Ref. [13].

We first characterize a spatial mode attack for a channel without turbulence ($r_0 = \infty$) before considering a turbulent channel. The optical alignment between the sender (Alice) and the receiver (Bob) is optimized by equalizing the detection count rates of the four polarization channels for a beam propagating through the center of the scanning lens L_E [i.e., along the green center ray shown in Fig. 3(a)]. This initial alignment represents normal operation which has a scanning angle $\phi = \theta = 0$. We then move the two-axis translation stage to adjust the position of lens L_E, and record the four detection efficiencies (**H**, **V**, **D**, and **A**) for different angles (θ, ϕ). In principle, the tip-tilt angles induced on the beam by the scanning lens are equivalent to including additional Zernike polynomial terms in the SLM hologram. Furthermore, the order in which the different Zernike polynomials are applied to the beam is interchangeable. As a result, our configuration of having the scanning lens follow the SLM is equivalent to Eve first steering the beam before it propagates through atmospheric turbulence. The scan is performed in 135-μrad steps, covering a range of ± 2.7 mrad, which corresponds to a lateral displacement of ± 35 mm along the front lens L1 of the QKD receiver.

In order for an angle to be a valid attack angle for channel k ($k = \mathbf{H}, \mathbf{V}, \mathbf{D}$, or $\mathbf{A}$), it must satisfy the condition that the probability of detection in channel k is δ_k times greater than the detection probabilities of the two channels in the other

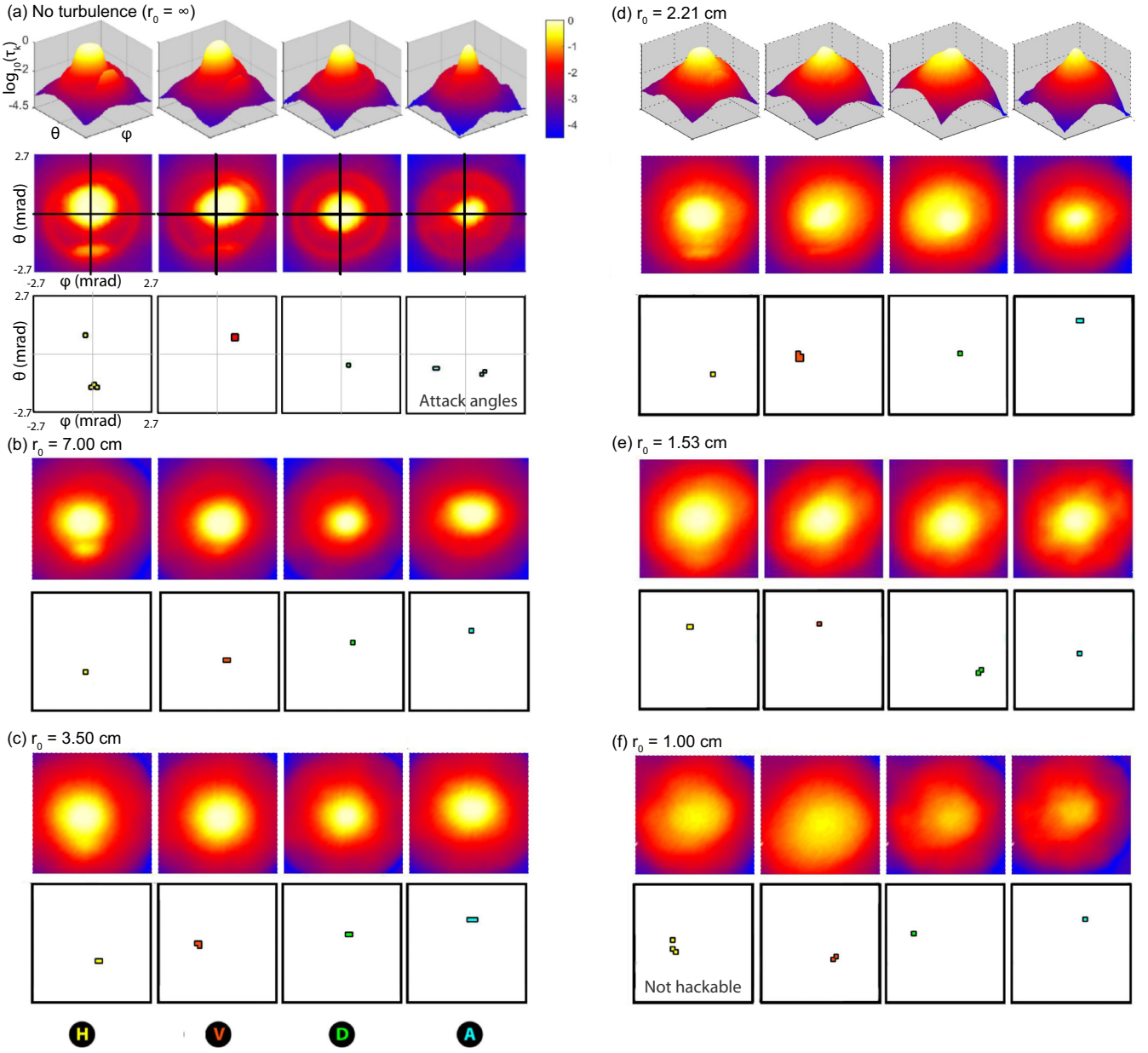

FIG. 4. Normalized count rates τ_k for each detector $k = \mathbf{H}, \mathbf{V}, \mathbf{D}$, or $\mathbf{A}$ at different incoming beam angles (θ, ϕ), and the corresponding attack angles for different turbulence strengths r_0. The attack angles for the four polarization detectors are shown left to right as horizontal **H** (yellow), vertical **V** (red), diagonal **D** (green), and antidiagonal **A** (light blue). The emulated turbulence corresponds to different r_0 values for an initial beam diameter $D = 20$ cm and $\lambda = 532$ nm. A smaller r_0 value corresponds to stronger atmospheric turbulence.

basis. For example, if $k = \mathbf{H}$, then $\min\{\tau_H/\tau_D, \tau_H/\tau_A\} > \delta_H$, where τ_k is the normalized detection probability defined as the ratio between the detection rate at the attack angle over the expected detection probability of Bob. We continuously increase the threshold δ_k until only a few attack angles satisfy these conditions. From the attacker's point of view, it is desirable to have δ_k as large as possible because a large value means an increased chance that detector k will click while minimizing the detection probabilities of the two other channels, which improves Eve's knowledge of Alice's state.

The scan results without turbulence ($r_0 = \infty$) for the four polarization channels are shown in Fig. 4(a), and the corresponding detection efficiency mismatch parameters are listed in Table I. There are noticeable features that cause efficiency mismatch, such as the side peak visible below the center peak in the **H** detector's map and the outer ring in all four detector maps. The valid attack angles for the **H** detector correspond to when the click probability is 22 times higher than **D** and **A** detectors (i.e., $\delta_H = 22$), and the normalized detection probability $\tau_H = 0.1$. Although the mismatch ratios on **D** ($\delta_D = 5$) and **A** ($\delta_A = 1.2$) channels are small, the mismatches in **H** and **V** ($\delta_V = 30$) channels are sufficient for a successful attack under our assumption that Alice and Bob only monitor the total count rate (not individual channels).

The optimized QBER as a function of transmission loss between Alice and Bob for a channel without turbulence

TABLE I. Detection efficiency mismatch parameters for attack data shown in Figs. 4 and 5. τ_k is the relative detection efficiency at an attack angle compared to the normal incidence case, and varies for different turbulence strengths due to changes in the scanning features that lead to valid attack angles. The value of the threshold of detection efficiency ratio δ_k decreases under stronger turbulence. If the δ_k are too low, it is impossible for Eve to find an optimal mean photon number for her resend signal that matches Bob's expected detection rate and does not induce error above the termination threshold. * denotes the turbulence strengths where an attack is not feasible.

	δ_k				τ_k			
r_0 (cm)	H	V	D	A	H	V	D	A
∞	22	30	5.0	1.2	0.1	0.03	0.3	0.001
7.00	20	5.0	1.03	3.5	0.3	0.4	0.8	0.7
3.50	8.0	2.5	1.08	2.3	0.5	0.15	0.85	0.5
2.21	4.5	1.8	1.15	2.21	0.4	0.2	0.85	0.2
1.53	3.0	2.0	1.7	1.25	0.45	0.3	0.85	0.02
1.00*	1.2	1.7	1.02	1.01	0.25	0.4	0.3	0.15

is shown in Fig. 5. In a practical scenario, Alice and Bob might experience transmission efficiency fluctuations in their quantum channel. As a result, they need to tolerate some deviation in their key rate from their estimated value. Shown in Fig. 5 is the QBER during Eve's attack as a function of the lowest transmission loss acceptable to Alice and Bob. In the next section, we examine the success of Eve's attack in the presence of turbulence.

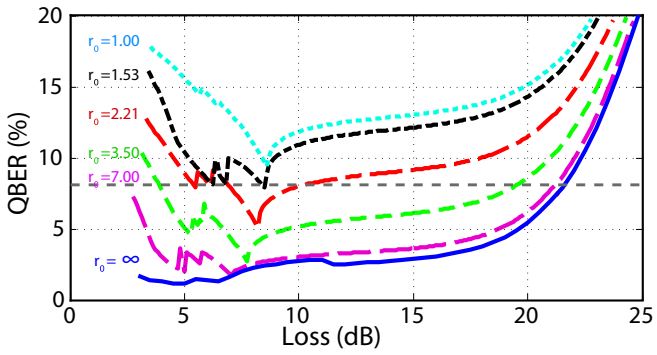

FIG. 5. Modeled attack performance. Quantum bit error rate (QBER) as a function of transmission loss for no turbulence (blue solid line) and different turbulence strengths corresponding to $r_0 =$ 7.00 cm (pink dashed line), 3.50 cm (green dotted line), 2.21 cm (red dot-dashed line), 1.53 cm (black dashed line), and 1.00 cm (cyan dashed line). The horizontal gray dashed line denotes the 8% threshold where Eve's attack is successful when QBER is below this value in our attack model. The maximum transmission loss where Eve's attack is successful decreases as turbulence strength increases. The mismatch ratios are too small in the case of 1.00 cm ($\delta_k \leqslant 2$ for all channels), and the optimization program could not find a solution with a QBER below 8% threshold given any transmission loss. The higher QBER at low loss (i.e., 3.5–7 dB) is because Eve has to send higher mean photon number states for channels with lower δ_k in order to match the expected detection rate of Bob.

V. PRACTICAL ATTACK UNDER TURBULENCE

To simulate our attack in the presence of atmospheric turbulence, we use a set of 29 holograms per turbulence strength, as described in Sec. II. We have performed scans of our QKD receiver for five different turbulence strengths: $r_0 = 7.00$, 3.50, 2.21, 1.53, and 1.00 cm. Our preliminary experiments that included tip-tilt wander caused by turbulence (i.e., the second and third terms of Zernike polynomials) showed that if Eve does not correct for beam wander caused by turbulence her attack is not feasible even under very weak turbulence ($r_0 = 7.00$ cm) corresponding to typical high-altitude atmospheric conditions. The beam wander from tip tilt alone was a strong enough disturbance to significantly hinder her attack. We then repeated the attack under the assumption that Eve can correct for tip-tilt beam wander using adaptive optics, such as with a deformable mirror or SLM. These corrections are implemented in our scans by setting the weight of the second and third terms of Zernike polynomials to zero.

In order to maintain accuracy and stability in our scans, we have chosen to cycle through all 29 holograms at one lens position before moving the translation stage to the next position. This method ensures each hologram is applied to the same scanning angle. We then repeat this scanning process for a total of 1681 angle positions, and record 29 separate detection rates per attack angle for each of the four polarization channels. To represent the Gaussian distribution of centroid displacements discussed in Sec. II, the final normalized detection efficiency of each detector τ_k is given by a weighted average of the detection rates from each hologram per scanning angle (θ, ϕ):

$$\tau_k(\theta, \phi) = \sum_{i=1}^{N} \Phi_i \tau_{k,i}(\theta, \phi), \tag{3}$$

where $\tau_{k,i}$ is the average detection efficiency of the k detector under the holograms selected from the ith radius. Φ_i is the probability of occurrence of the ith partition discussed in Sec. II. $N = 5$ is the number of partitions. We select one sample hologram for no turbulence; eight samples each for 0.5σ, 1σ, and 2σ partitions; and four samples from the 3σ partition. The samples are given the weight factor corresponding to the radius from the sample used to the next larger sample, thus representing the best case hologram from this range. This weight factor ensures that the samples form an optimistic (easier to hack) representation of the turbulence effect, and therefore ensure that any turbulence found to not be vulnerable to attacks is indeed safe under the parameter monitoring assumptions. The total detection rate τ_k is used to find valid attack angles under turbulent conditions using the same method as without turbulence. We then repeat this process for different turbulence strengths from very weak ($r_0 =$ 7.00 cm) to stronger turbulence emulating low-altitude sea level conditions ($r_0 = 1.00$ cm). A map of successful attack angles and the corresponding detection efficiency mismatch parameters are shown in Figs. 4(b)–4(f) and Table I.

Our scanning results in Table I show that, as the turbulence strength increases, the mismatch ratios δ_k are significantly reduced. We can see in Fig. 4 that the features that are responsible for efficiency mismatch become blurry and eventually

disappear as turbulence increases in strength, and it becomes harder for Eve to maintain a precise attack angle when $r_0 \leqslant 1.53$ cm. For stronger turbulence ($r_0 = 1.00$ cm), the only remaining hackable feature is the displacement of the center peaks due to a slight misalignment between the fiber couplers in each arm of the receiver. As a result, most of the attack angles at stronger turbulence are found closer to the center peak. However, they do not result in a successful attack for $r_0 < 1.53$ cm because the induced QBER is above the 8% termination threshold.

In order to perform a quantitative verification of an attack, we use an optimization program to find the minimal QBER as a function of transmission loss. The results in Fig. 5 show that the optimized QBER for an attack in stronger turbulence ($r_0 = 2.21$ cm) is higher than that of weaker turbulence ($r_0 = 7.00$ cm). If we assume that the QBER threshold for Alice and Bob to terminate the protocol is 8%, then the attack without turbulence is successful as long as the transmission loss between Alice and Bob is less than 21 dB, whereas in the presence of turbulence Eve can successfully attack this receiver for $r_0 \geqslant 2.21$ cm when the transmission loss is less than 10 dB but higher than 7 dB. Using Eq. (1), $r_0 = 2.21$ cm is equivalent to Eve having her resend setup approximately 0.5 km away from Bob's receiver in typical sea level turbulence conditions ($C_n^2 = 1.8 \times 10^{-14}$ m$^{-2/3}$). Eve is unable to match Bob's count rate for transmission loss below 3.5 dB even if she uses all four channels due to Eve's nonperfect detection efficiency. Therefore, the optimization program could not find a solution matching Bob's total detection rates for transmission losses below 3.5 dB.

The result for $r_0 = 1.53$ cm shows that there is only a small loss window (around 8.5 dB) where Eve can attack without inducing a QBER higher than the threshold. Using Eq. (1) and the value of C_n^2 given above, this r_0 corresponds to a distance of 1 km. At lower transmission loss (i.e., 3.5–7 dB), the expected detection rate at Bob is too high for Eve to match using a single channel, and therefore she must also use the other channels that have a lower δ_k. This causes the QBER to increase and results in the irregularities seen for loss below 7 dB when the number of channels being used is changed. The QBER curves become smoother at higher loss once Eve can fully replicate Bob's detection rates while only sending signals to a single polarization channel, which takes advantage of the greatest efficiency mismatch for an optimized attack. The mismatch ratios in the case of 1.00 cm ($\delta_k \leqslant 2$ for all channels) are too small for the optimization program to find a solution for a QBER below the threshold given any transmission loss.

Implementations of QKD can and should monitor counts at each detector to ensure they remain relatively balanced. The higher QBER obtained when Eve is forced to send states to channels with lower mismatch ratios illustrates how monitoring each channel would increase the difficulty of a successful attack. However, it is uncommon in practice to monitor individual count rates, and there are no current standards or established guidelines for allowable variation in detection rates. The added constraint to maintain precise detection rates would make hacking more difficult for an eavesdropper, but does not in itself prevent an attack. It also does not invalidate the current work of determining if bounds exist on the turbulence strength where QKD systems can be hacked.

VI. THEORETICAL LIMIT OF ATTACK UNDER TURBULENCE

The attack described in Sec. V is only one particular example of an intercept-resend attack. Other attacks in this class may exist, which shows that a QKD system with detection efficiency mismatch could be insecure if the security analysis does not take the mismatch into account. Whenever the observed and monitored data are compatible with an IR attack, no secret key can be obtained [42,43].

For this reason, it is useful to ask the question whether the data we observe are consistent with an IR attack or not. Along the way we can also answer the question whether a fine-grained analysis of the observations could exclude IR attacks, and thus potentially give a secure key where the coarse-grained analysis (which uses only average error rate and average detection rate) fails.

The handle to determine whether given data are compatible with an IR attack or not is the fact that IR attacks make the channel between Alice and Bob entanglement breaking. That is, this channel acting as one system of a bipartite entangled state will transform it into a separable bipartite state. So by verifying that the channel is not entanglement breaking, we can exclude the IR attacks. To do so, we do not require actual entanglement: we can probe the channel with nonorthogonal signal states, just as in any prepare-and-measure QKD setup, and use the formalism of the source-replacement scheme (see, for example, Ref. [44]) to formulate an equivalent thought setup that virtually uses an entangled state. The probabilities $p(ab|xy)$ between Alice's signal choice a and Bob's measurement result b for respective basis choices x and y can then be thought of as coming from measurements on this entangled state with both Alice and Bob performing measurements with POVM elements $M_A^{x,a}$ and $M_B^{y,b}$, respectively. If these observations serve as an entanglement witness, we have shown that the channel is not entanglement breaking.

We can formulate the entanglement verification problem as the optimization problem

$$\begin{aligned} &\text{find} && \rho_{AB} \\ &\text{subject to} && \rho_{AB} \geqslant 0 \text{ and } \rho_{AB}^{\Gamma_A} \geqslant 0 \\ &&& \mathrm{Tr}\left(\rho_{AB} M_A^{x,a} \otimes M_B^{y,b}\right) = p(ab|xy), \forall a, b, xy. \end{aligned} \tag{4}$$

Here Γ_A is the partial transpose operation on Alice's system. If the above optimization problem is not feasible, then the state ρ_{AB} is entangled [45]. In our previous work [34], we developed a method to solve the above optimization problem when detectors' efficiencies are mismatched and the dimension of the optical signal is unbounded.

In this paper, we did not measure the joint distribution $p(ab|xy)$ of Alice and Bob directly in the experiment. However, given the characterization of detection efficiency mismatch from our experiment, we can deduce the joint distribution of Alice and Bob from the case without efficiency mismatch according to our simulation model. Using the method developed in Ref. [34], we found that when there is no turbulence or very weak turbulence $r_0 = 7.00$ cm we

cannot verify entanglement. Thus, the channel is vulnerable. This result is in agreement with the results in Ref. [34].

However, when turbulence is stronger ($r_0 \leqslant 3.50$ cm), our calculation shows that entanglement can be verified. This means that there is no intercept-resend strategy for Eve that can match all of Alice and Bob's expected observations. This result is based on a strong condition where Eve needs to match all expected measurable parameters of Alice and Bob, whereas the results presented in Sec. V were under the practical assumptions that Alice and Bob monitor only coarse-grained information, namely, the total detection rate and error rate.

VII. CONCLUSION

We experimentally study how atmospheric turbulence in a free-space channel can affect an eavesdropper's ability to perform a spatial mode attack on a QKD receiver. We use a phase-only spatial light modulator to emulate atmospheric turbulence in the laboratory, the accuracy of which is verified by comparing measured far-field intensity distributions and centroid displacements to theoretical predictions. We then study a spatial mode detection efficiency mismatch attack under a range of atmospheric turbulence strengths to determine the maximum unsafe radius around the free-space QKD receiver. Our attack model is based on an intercept-resend attack under the practical assumptions that only the total detection rate and QBER are monitored by Alice and Bob. We find that for this particular receiver an eavesdropper could attack a nondecoy state BB84 system from up to about 1 km away in typical sea level turbulence conditions ($r_0 = 1.53$ cm for a 20-cm beam at 532 nm). This result is assuming Eve can correct for basic tip-tilt beam wander using conventional adaptive optics. Eve's chances of success will be further reduced if Alice and Bob choose to monitor individual detection channel statistics. In this case, we theoretically find that an IR attack is still possible for weaker turbulence ($r_0 \geqslant 7.0$ cm). The assumption that an eavesdropper has physical limitations is not usually included in the security analysis of a QKD system. If there is a chance that Eve is inside this secure zone around Bob's receiver, or has advanced adaptive optics capacities to correct for beam aberrations, then extra care regarding these types of attacks may be required.

ACKNOWLEDGMENTS

We thank Ben Davies and Brendon Higgins for assisting with our simulation code and data analysis. This work was funded by the US Office of Naval Research (ONR), Industry Canada, Canada Foundation for Innovation (CFI), The Natural Sciences and Engineering Research Council of Canada (NSERC) (Discovery program and CREATE project CryptoWorks21), Canadian Space Agency, Ontario MRIS, and the Ministry of Education and Science of Russia (program NTI center for quantum communications). P.C. was supported by Thai Development and promotion of science and technology teaching project (DPST) scholarship. K.B.K. was supported by Canada First Research Excellence Fund. A.H. was supported by China Scholarship Council.

P.C. and K.B.K. contributed equally to this work.

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (IEEE, New York, 1984), pp. 175–179.

[2] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[3] A. Vakhitov, V. Makarov, and D. R. Hjelme, J. Mod. Opt. **48**, 2023 (2001).

[4] V. Makarov, A. Anisimov, and J. Skaar, Phys. Rev. A **74**, 022313 (2006); **78**, 019905(E) (2008).

[5] B. Qi, C.-H. F. Fung, H.-K. Lo, and X. Ma, Quant. Inf. Comp. **7**, 73 (2007).

[6] A. Lamas-Linares and C. Kurtsiefer, Opt. Express **15**, 9388 (2007).

[7] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photonics **4**, 686 (2010).

[8] F. Xu, B. Qi, and H.-K. Lo, New J. Phys. **12**, 113026 (2010).

[9] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nat. Commun. **2**, 349 (2011).

[10] H. Weier, H. Krauss, M. Rau, M. Fürst, S. Nauerth, and H. Weinfurter, New J. Phys. **13**, 073024 (2011).

[11] P. Jouguet, S. Kunz-Jacques, and E. Diamanti, Phys. Rev. A **87**, 062313 (2013).

[12] S. Sajeed, I. Radchenko, S. Kaiser, J.-P. Bourgoin, A. Pappa, L. Monat, M. Legré, and V. Makarov, Phys. Rev. A **91**, 032326 (2015).

[13] S. Sajeed, P. Chaiwongkhot, J.-P. Bourgoin, T. Jennewein, N. Lütkenhaus, and V. Makarov, Phys. Rev. A **91**, 062301 (2015).

[14] V. Makarov, J.-P. Bourgoin, P. Chaiwongkhot, M. Gagné, T. Jennewein, S. Kaiser, R. Kashyap, M. Legré, C. Minshull, and S. Sajeed, Phys. Rev. A **94**, 030302(R) (2016).

[15] S. Sajeed, A. Huang, S. Sun, F. Xu, V. Makarov, and M. Curty, Phys. Rev. Lett. **117**, 250505 (2016).

[16] A. Meda, I. P. Degiovanni, A. Tosi, Z. Yuan, G. Brida, and M. Genovese, Light Sci. Appl. **6**, e16261 (2017).

[17] S. Sajeed, C. Minshull, N. Jain, and V. Makarov, Sci. Rep. **7**, 8403 (2017).

[18] R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, H. Weier, T. Scheidl, M. Lindenthal, B. Blauensteiner, T. Jennewein, J. Perdigues, P. Trojek, B. Ömer, M. Fürst, M. Meyenburg, J. Rarity, Z. Sodnik, C. Barbieri, H. Weinfurter, and A. Zeilinger, Nat. Phys. **3**, 481 (2007).

[19] C. Erven, C. Couteau, R. Laflamme, and G.Weihs, Opt. Express **16**, 16840 (2008).

[20] M. P. Peloso, I. Gerhardt, C. Ho, A. Lamas-Linares, and C. Kurtsiefer, New J. Phys. **11**, 045007 (2009).

[21] S. Nauerth, F. Moll, M. Rau, C. Fuchs, J. Horwath, S. Frick, and H. Weinfurter, Nat. Photonics **7**, 382 (2013).

[22] S.-K. Liao, H.-L. Yong, C. Liu, G.-L. Shentu, D.-D. Li, J. Lin, H. Dai, S.-Q. Zhao, B. Li, J.-Y. Guan, W. Chen, Y.-H. Gong, Y. Li, Z.-H. Lin, G.-S. Pan, J. S. Pelc, M. M. Fejer, W.-Z. Zhang, W.-Y. Liu, J. Yin, J.-G. Ren, X.-B. Wang, Q. Zhang, C.-Z. Peng, and J.-W. Pan, Nat. Photonics **11**, 509 (2017).

[23] C. J. Pugh, S. Kaiser, J.-P. Bourgoin, J. Jin, N. Sultana, S. Agne, E. Anisimova, V. Makarov, E. Choi, B. L. Higgins, and T. Jennewein, Quantum Sci. Technol. **2**, 024009 (2017).

[24] J. Yin, Y. Cao, Y.-H. Li, S.-K. Liao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, B. Li, H. Dai, G.-B. Li, Q.-M. Lu, Y.-H. Gong, Y. Xu, S.-L. Li, F.-Z. Li, Y.-Y. Yin, Z.-Q. Jiang, M. Li, J.-J. Jia, G. Ren, D. He, Y.-L. Zhou, X.-X. Zhang, N. Wang, X. Chang, Z.-C. Zhu, N.-L. Liu, Y.-A. Chen, C.-Y. Lu, R. Shu, C.-Z. Peng, J.-Y. Wang, and J.-W. Pan, Science **356**, 1140 (2017).

[25] R. Ursin, T. Jennewein, J. Kofler, J. Perdigues, L. Cacciapuoti, C. de Matos, M. Aspelmeyer, A. Valencia, T. Scheidl, A. Acin, C. Barbieri, G. Bianco, C. Brukner, J. Capmany, S. Cova, D. Giggenbach, W. Leeb, R. Hadfield, R. Laflamme, N. Lütkenhaus, G. Milburn, M. Peev, T. Ralph, J. Rarity, R. Renner, E. Samain, N. Solomos, W. Tittel, J. Torres, M. Toyoshima, A. Ortigosa-Blanch, V. Pruneri, P. Villoresi, I. Walmsley, G. Weihs, H. Weinfurter, M. Zukowski, and A. Zeilinger, Europhys. News **40**, 26 (2009).

[26] J.-P. Bourgoin, E. Meyer-Scott, B. L. Higgins, B. Helou, C. Erven, H. Hübel, B. Kumar, D. Hudson, I. D'Souza, R. Girard, R. Laflamme, and T. Jennewein, New J. Phys. **15**, 023006 (2013).

[27] J. Yin, Y. Cao, S.-B. Liu, G.-S. Pan, J.-H. Wang, T. Yang, Z.-P. Zhang, F.-M. Yang, Y.-A. Chen, C.-Z. Peng, and J.-W. Pan, Opt. Express **21**, 20032 (2013).

[28] G. Vallone, D. Bacco, D. Dequal, S. Gaiarin, V. Luceri, G. Bianco, and P. Villoresi, Phys. Rev. Lett. **115**, 040502 (2015).

[29] A. Carrasco-Casado, H. Takenaka, M. Fujiwara, M. Kitamura, M. Sasaki, and M. Toyoshima, in *SPIE 10660, Quantum Information Science, Sensing, and Computation X* (SPIE, Bellingham, Washington, 2018), p. 106600B.

[30] M. Rau, T. Vogl, G. Corrielli, G. Vest, L. Fuchs, S. Nauerth, and H. Weinfurter, IEEE J. Quantum Electron. **21**, 6600905 (2015).

[31] R. Tyson, *Principles of Adaptive Optics*, 3rd ed. (CRC, Boca Raton, FL, 2010).

[32] Y. Li, D. Chen, and X. Du, Proc. SPIE **5237**, 271 (2004).

[33] L. H. Lee, G. J. Baker, and R. S. Benson, J. Opt. Soc. Am. A **23**, 2602 (2006).

[34] Y. Zhang and N. Lütkenhaus, Phys. Rev. A **95**, 042319 (2017).

[35] B. Rodenburg, M. Mirhossein, M. Malik, O. S. Magana-Loaiza, M. Yanakas, L. Maher, N. K. Steinhoff, G. A. Tyler, and R. W. Boyd, New J. Phys. **16**, 033020 (2014).

[36] C. Erven, B. Heim, E. Meyer-Scott, J. P. Bourgoin, R. Laflamme, G. Weihs, and T. Jennewein, New J. Phys. **14**, 123018 (2012).

[37] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*, 2nd ed. (SPIE, Bellingham, Washington, 2005).

[38] R. J. Noll, J. Opt. Soc. Am. **66**, 207 (1976).

[39] L. Burger, I. A. Litvin, and A. Forbes, S. Afr. J. Sci. **104**, 129 (2008).

[40] V. Makarov and D. R. Hjelme, J. Mod. Opt. **52**, 691 (2005).

[41] J.-P. Bourgoin, N. Gigov, B. L. Higgins, Z. Yan, E. Meyer-Scott, A. K. Khandani, N. Lütkenhaus, and T. Jennewein, Phys. Rev. A **92**, 052339 (2015).

[42] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).

[43] M. Curty, O. Gühne, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. A **71**, 022306 (2005).

[44] A. Ferenczi, V. Narasimhachar, and N. Lütkenhaus, Phys. Rev. A **86**, 042327 (2012).

[45] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).